Multiple emission components in the Cygnus cocoon detected from Fermi-LAT observations

X. Astiasarain©, L. Tibaldo©, P. Martin©, J. Knödlseder1, and Q. Remy2©

1 IRAP, Université de Toulouse, CNRS, CNES, UPS, 9 avenue Colonel Roche, 31028 Toulouse, Cedex 4, France
   e-mail: xan.astiasarain@irap.omp.eu; luigi.tibaldo@irap.omp.eu; pierrick.martin@irap.omp.eu
2 Max Planck Institut für Kernphysik, Saupfercheckweg 1, 69117 Heidelberg, Germany

Received 29 November 2022 / Accepted 6 January 2023

ABSTRACT

Context. Star-forming regions may play an important role in the life cycle of Galactic cosmic rays (CRs), notably as home to specific acceleration mechanisms and transport conditions. Gamma-ray observations of Cygnus X have revealed the presence of an excess of hard-spectrum gamma-ray emission, possibly related to a cocoon of freshly accelerated particles.

Aims. We seek an improved description of the gamma-ray emission from the cocoon using ~13 yr of observations with the Fermi-Large Area Telescope (LAT) and use it to further constrain the processes and objects responsible for the young CR population.

Methods. We developed an emission model for a large region of interest, including a description of interstellar emission from the background population of CRs and recent models for other gamma-ray sources in the field. Thus, we performed an improved spectro-morphological characterisation of the residual emission including the cocoon.

Results. The best-fit model for the cocoon includes two main emission components: an extended component FCES G78.74+1.56, described by a 2D Gaussian of extension $r_{maj} = 4.4\pm0.1$ and a smooth broken power law spectrum with spectral indices $1.67\pm0.05$, $2.12\pm0.02$ below and above $3.0\pm0.02$ GeV, respectively; and a central component FCES G80.00+0.50, traced by the distribution of ionised gas within the borders of the photo-dissociation regions and with a power law spectrum of index $2.19\pm0.03$ that is significantly different from the spectrum of FCES G78.74+1.56. An additional extended emission component FCES G78.83+3.57, located on the edge of the central cavities in Cygnus X and with a spectrum compatible with that of FCES G80.00+0.50, is likely related to the cocoon. For the two brightest components FCES G80.00+0.50 and FCES G78.74+1.56, spectra and radial-azimuthal profiles of the emission can be accounted for in a diffusion-loss framework involving one single population of non-thermal particles with a flat injection spectrum. Particles span the full extent of FCES G78.74+1.56 as a result of diffusion from a central source, and give rise to source FCES G80.00+0.50 by interacting with ionised gas in the innermost region.

Conclusions. For this simple diffusion-loss model, viable setups can be very different in terms of energetics, transport conditions, and timescales involved, and both hadronic and leptonic scenarios are possible. The solutions range from long-lasting particle acceleration, possibly in prominent star clusters such as Cyg OB2 and NGC 6910, to a more recent and short-lived release of particles within the last 10–100 kyr, likely from a supernova remnant. The observables extracted from our analysis can be used to perform detailed comparisons with advanced models of particle acceleration and transport in star-forming regions.

Key words. acceleration of particles – cosmic rays – gamma rays: ISM – open clusters and associations: individual: Cygnus OB2 – open clusters and associations: individual: NGC 6910

1. Introduction

There is firm evidence that cosmic rays (CRs) at energies below 1 PeV originate from the Milky Way. Supernova remnants (SNRs) remain the leading candidate as sources of the majority of Galactic CRs, most likely through the process of diffusive shock acceleration, while alternative source classes including massive star-forming regions, the Galactic centre, pulsar wind nebulae (PWNe), and compact binary systems may bring complementary contributions over specific parts of the extended CR spectrum (see, for instance Gabici et al. 2019, and references therein).

Massive star-forming regions are of particular interest in this context (for instance Bykov et al. 2020). The clusters of OB stars at their centres are the progenitors of a variety of particle acceleration sites such as SNRs, pulsars, and PWNe, or compact binary systems. In addition, the collective action of powerful stellar winds and, after a few million to a few tens of million years, the explosion of massive stars into supernovae lead to the formation of super-bubbles (SBs), which are large cavities filled by a highly dynamical medium that, as a whole, may play a specific role in the life cycle of CRs. The isotopic abundances measured in CRs suggest that at least a fraction of the CR material is sourced from the winds of massive stars (Binns et al. 2008; Tatischeff et al. 2021).

Accelerated particles in distant locations can be revealed via the gamma-ray emission produced when they interact with interstellar gas, through inelastic collisions for nuclei or Bremsstrahlung for leptons, and radiation fields, through the inverse-Compton (IC) scattering by leptons. Therefore, star-forming regions are expected to be bright gamma-ray...
sources from the interactions of particles with the large masses of interstellar gas and the intense radiation fields available in these environments. Gamma-ray emission in the GeV and TeV energy ranges is detected towards a growing number of massive star-forming regions (for a review see for instance Tibaldo et al. 2021), and taken as evidence in favour of in situ CR acceleration. However, the clustering of energetic objects and interstellar clouds combined with the limited resolution of gamma-ray telescopes makes it difficult to firmly identify the acceleration sites and mechanisms, and to understand how particles propagate and interact through the region and eventually escape to merge into the large-scale CR population in the Galaxy.

Observational progress is matched by a flourishing development of models of particle acceleration and transport by stellar winds (Gupta et al. 2018; Bykov et al. 2020; Morlino et al. 2021) and SBs (Bykov 2001; Ferrand & Marcowith 2010; Tolksdorf et al. 2019; Vieu et al. 2022). The models show that these objects can be efficient particle accelerators and make a contribution to Galactic CRs. They predict a number of morphological and spectral signatures that can be looked for to test the physical processes at the origin of the observed gamma-ray signals.

Cygnus X is one of the best studied massive star-forming regions in the Milky Way. Cygnus X contains Cygnus OB2 that, with 78 confirmed O stars (Berlanas et al. 2020), is among the largest associations of massive stars in the Milky Way. It is composed of multiple substructures with a main group at ~1.76 kpc from the Earth and a foreground group at ~1.35 kpc (Berlanas et al. 2019) and at least two star-forming bursts ~3 and ~5 Myr ago (Berlanas et al. 2020). A second prominent massive stellar cluster in Cygnus X is NGC 6910 at a distance of ~1.73 kpc (Cantat-Gaudin et al. 2020), an age in the range from 5 to 10 Myr (Delgado & Alfaro 2000; Cantat-Gaudin et al. 2020), and a flat mass function pointing to a large number of massive stars (Kaur et al. 2020).

The Large Area Telescope (LAT) aboard the Fermi Gamma-ray Space Telescope (Atwood et al. 2009) unveiled a hard gamma-ray excess towards Cygnus X with an extension $r_{68} = 3.0^\circ \pm 0.3^\circ$ (Ackermann et al. 2011). The excess was interpreted as the signature of a cocoon of freshly accelerated particles. Gamma-ray emission in the energy range from hundreds of GeV to hundreds of TeV from the Cygnus cocoon was subsequently detected using ARGO-YBJ, HAWC, and LHAASO (Bartoli et al. 2020). The models show that these objects can be efficient particle accelerators and make a contribution to Galactic CRs. They predict a number of morphological and spectral signatures that can be looked for to test the physical processes at the origin of the observed gamma-ray signals.

2. Construction of interstellar gas maps

The distribution of interstellar gas towards the region of interest is a key ingredient of our analysis for two reasons: (1) it is necessary to model the strong foreground and background gamma-ray emission from the interactions of the large-scale Galactic CR population with interstellar gas in the direction of Cygnus, and thus be able to extract and characterise the emission of the cocoon; (2) it is used in the interpretation of the gamma-ray signal in terms of the underlying CR populations.

2.1. Atomic and molecular gas

We trace atomic gas using the 21 cm emission line from the hyperfine transition of atomic hydrogen H I. We use data from the Canadian Galactic Plane Survey (CGPS, Taylor et al. 2003) with an angular resolution of $1^\prime$ and a velocity resolution of $1.3$ km s$^{-1}$ in the region with Galactic longitude $75.5^\circ < l < 90^\circ$ and Galactic latitude $-3^\circ < b < 5^\circ$. Outside this region we use data from the all-sky HI4PI survey from Effelsberg and Parkes observations (HI4PI Collaboration 2016) with a lower angular resolution of $0.27^\circ$ and velocity resolution of $1.49$ km s$^{-1}$. We checked the consistency of the two surveys by comparing the data in the region covered by the CGPS.

We derived column densities $N$(H I) under the hypothesis of a uniform spin temperature. All results are shown for the reference spin temperature of 250 K suggested by emission-absorption spectrum pairs in the CGPS area (Dickey et al. 2009) and that was also found to best reproduce gamma-ray observations of the Cygnus region based on an earlier analysis (Ackermann et al. 2012a). This is a highly uncertain parameter that is not expected to be uniform along lines of sight and across the region. Therefore, the analysis was also performed for alternative uniform spin temperatures of 100 K (lower bound set by the brightness temperatures observed in the region), 400 K, and the optically thin case, which are used to set systematic uncertainties on relevant quantities.

Molecular hydrogen H$_2$ cannot be traced directly. We use the $^{12}$CO $J_{1-0}$ rotational line at 2.6 mm as surrogate tracer, under the usual hypothesis that $N$(H$_2$) column densities are directly proportional to the CO intensity (velocity-integrated brightness temperature) $W_{10}$ through a constant known as $X_{CO} = N$(H$_2$)/$W_{10}$. We use CO data from the composite survey by Dame et al. (2001) with an angular resolution of $0.125^\circ$ in the area considered in this paper and a velocity resolution of $1.3$ km s$^{-1}$. Data were noise-filtered using the moment-masking technique (Dame 2011).

The Doppler shift of the lines can be used to infer the gas velocity along the line of sight due to Galactic rotation, and therefore separate multiple structures. However, intrinsic velocity dispersion can cause biases in the estimates of gas column densities across adjacent structures. To address this problem, we used the line profile fitting technique described in Remy et al. (2017) to decompose emission from each line of sight into a combination of pseudo-Voigt functions. We built longitude-velocity and latitude-velocity diagrams based on the fit results for H I and CO, and defined in the longitude-latitude-velocity space boundaries that separate the gas into three structures along the line of sight, namely the local arm (including the Cygnus complex), the Perseus arm, and the outer arm and beyond. Figure 1 shows an example of longitude-velocity diagram in the longitude range $73^\circ < l < 87^\circ$ that is used in the following analysis.

The final H I and CO maps for the three regions along the line of sight are shown in Fig. 2. Since all surveys have different...

---

1 Throughout the paper we refer to a source extension as its 68% containment radius $r_{68}$. For a 2D Gaussian intensity distribution, $r_{68} = 1.51 r_{\sigma}$. 

---

A&A 671, A47 (2023)
Galactic longitude summed over −2° < b < 2°. Total column densities from each pseudo-Voigt profile were assigned to the velocity of the peak. The black lines show the boundaries that we defined to separate the three structures along the line of sight: local arm, Perseus arm, and outer Arm and beyond (from top to bottom).

2.2. Dark neutral medium (DNM)

A significant fraction of neutral interstellar gas cannot be traced by the H\textsubscript{i} 21 cm line nor by the \textsuperscript{13}CO \textsubscript{J=1−0} rotational line (Grenier et al. 2005) and is therefore missing in the maps described above. It can be referred to as the dark neutral medium (DNM) and it is thought to be a combination of opaque H\textsubscript{i} and diffuse H\textsubscript{2} at the atomic-molecular interface of clouds, or dense H\textsubscript{2} at the core of molecular clouds (Remy et al. 2017).

If dust and gas in the interstellar medium (ISM) were well mixed and the dust grains physical and chemical properties were the same everywhere, dust thermal emission would be proportional to total gas column densities along the line of sight. Therefore, we can derive a DNM map by subtracting from the dust thermal emission the components correlated with H\textsubscript{i} and CO. We use a map of the dust optical depth at 353 GHz obtained from component separation of Planck and IRAS data (Planck Collaboration XLVIII 2016) with an effective angular resolution of 5′ in high signal-to-noise regions.

To avoid biases from the missing DNM component in the determination of the components correlated with H\textsubscript{i} and CO, we used the iterative fitting procedure described in Tibaldo et al. (2015). Briefly, the procedure consists in an iterative fitting of the gas maps to the dust map where the positive part of the residuals is re-injected in the model at each iteration to compute unbiased values of the fit parameters and obtain an estimation of the missing DNM component. A DNM map was calculated for each uniform spin temperature considered. Figure 3 shows the DNM map obtained for the reference spin temperature value of 250 K.

2.3. Ionised gas

We derived an ionised gas column density map from the free-free emission measure \textit{EM}(l, b) extracted from component separation of Planck, WMAP, and 408 MHz data by Planck Collaboration X (2016). The free-free emission measure from Cygnus X is dominated by two strong peaks that, as indicated by 8 μm emission from dust, lie inside the cavities carved in the ISM by the intense star-forming activity in the region.

We calculated H\textsubscript{i} column densities under the assumption that ionised gas fills a sphere of radius 3.5° corresponding to ∼100 pc at a distance of 1.7 kpc and with an uniform density along each line of sight. The sphere is meant to model the ionised cavities at the hearth of Cygnus X. With \( r \) the radius of the sphere and \( d \) the distance to Cygnus X the electron volume density is:

\[
N_e(l, b) = \left( \frac{EM(l, b)}{2 \sqrt{r^2 - d^2 \sin^2(l - l_0) - d^2 \sin^2(b - b_0)}} \right)^{1/2},
\]

Therefore, for the column density we obtain:

\[
N_{\text{HII}}(l, b) = N_e(l, b) \times 2 \sqrt{r^2 - d^2 \sin^2(l - l_0) - d^2 \sin^2(b - b_0)},
\]

where \( l_0 \) and \( b_0 \) are the position of the sphere’s centre and \( EM(l, b) \) the emission measure in a given direction.

The final ionised gas column density map is displayed in Fig. 4. The angular resolution of the free-free emission measure map from Planck is 1°. The final ionised gas column density map was re-binned on the same grid as the MSX 8 μm map with a grid spacing of 1′′.

3. Gamma-ray analysis and results

3.1. Data selection

We analysed 13.25 yr of \textit{Fermi}-LAT data from the beginning of the mission on 4 August 2008 to 3 November 2021. We used the P8R3 data set (Atwood et al. 2013; Bruel et al. 2018) and selected events in the P8R3\_SOURCE class that is associated with instrument response functions P8R3\_SOURCE\_V3. This event selection has a level of background contamination sufficiently low to study the bright extended emission from Cygnus X. Furthermore, we restricted the analysis to time intervals in which the LAT configuration and data quality is appropriate for science analysis.

Events are separated in four independent data sets according to their PSF event type, that is the quality of the direction reconstruction. For each event type, we selected events above a minimum energy so that the point spread function (PSF) 68% containment radius is always better than 0.7′, which roughly corresponds to the characteristic size of the most prominent spatial structures in the gas maps. The minimum energy threshold used is 0.5 GeV. Lowering the minimum energy induced instabilities in the analysis due to bright emission from a few pulsars in the region. To reliably characterise extended emission at energies <0.5 GeV event selection based on the pulsars phases would be necessary, but this is beyond the scope of the current study. The maximum energy is 1 TeV for all event types.

To reduce contamination from the bright gamma-ray emission from the Earth atmosphere, we selected events within a cone from the local zenith with aperture \( \zeta_{\text{max}} \). The value of \( \zeta_{\text{max}} \) was chosen from a visual inspection of the distribution of counts as a function of zenith angle for each event type component in the given energy ranges. Energy and zenith angle selection for the four data sets are summarised in Table 1.
3.2. Region of interest and emission model

A major challenge in the characterisation of extended emission from Cygnus X is to model the bright interstellar emission from the large-scale population of CRs. Due to the large column densities of the ISM in this region, emission associated with gas is the dominant contribution at GeV energies (Ackermann et al. 2012a). Under the assumption that the large-scale CR densities are uniform on the spatial scales of interstellar complexes, we can model the foreground and background intensity associated with interstellar gas \( I_{\text{gas}} \) as a linear combination of the column density maps for the different gas phases and structures along the line of sight:

\[
I_{\text{gas}}(l, b, E) = q_{\text{LIS}}(E) \cdot \left[ \sum_{i=1}^{3} \left( A_i N_{\text{HI},i}(l, b) + B_i W_{\text{CO},i}(l, b) \right) + C \tau_{\text{DNM}}(l, b) \right],
\]

where \( q_{\text{LIS}}(E) \) is the local gas emissivity spectrum, that is the gamma-ray emission rate per hydrogen atom, from Casandjian (2015), derived from LAT data. The summation over \( i \) describes the combination of the three regions along the line of sight: local arm, Perseus arm, outer arm and beyond. The free parameters \( A_i, B_i, \) and \( C \) account at the same time for variations of the large-scale CR densities across the three regions, and for the \( X_{\text{CO}} \) ratios and the dust specific opacity \( \sigma_{353} = \tau_{353}/N(\text{H} I) \). The spectral shape of \( q_{\text{LIS}} \) is fixed throughout the paper. We know that spectral variations of the emissivity along the line of sight are small towards Cygnus (Ackermann et al. 2012a) and, in general, towards the outer Galaxy (Acero et al. 2016). Conversely, this implies that any spectral deviations from the local interstellar spectrum (LIS) in Cygnus X are not accounted for by the background model and characterised as part of the cocoon.

An additional diffuse component is given by IC emission from the large-scale population of CR leptons. We accounted for it using the GALPROP model \( ^5\text{Y}^2\delta^8\text{O}^{15}\text{O}^2 \) (Ackermann et al. 2012b). Finally, we need to account for the isotropic gamma-ray background, which is a combination of extra-galactic

A47, page 4 of 29
We note that the IC model is also subject to large uncertainties. However, morphological variations over our limited region of interest described below are expected to be small for conventional models. Moreover, uncertainties in the spectrum are mitigated by the fact that the isotropic background spectrum is derived from a fit to the LAT data.

The interstellar emission model, along with the LAT response, sets the choice of the region of interest (ROI) for the analysis. The longitude and latitude extents should be sufficiently large to separate the extended emission of the Cygnus cocoon from the large-scale background, and so that the different components of the background model can be reliably constrained by the data. We chose a ROI with Galactic longitude $73^\circ \leq l \leq 87^\circ$ and with Galactic latitude $|b| \leq 15^\circ$. The longitude interval leaves out complexes associated with Cygnus OB1 at $l < 73^\circ$ and with HB 21 at $l > 87^\circ$. The wider coverage in latitude makes it possible to better constrain emission from local H I, IC scattering, and the isotropic background.

We modeled individual sources within the region based on the most recent catalogue of gamma-ray sources detected by the LAT, 4FGL-DR3 (Abdollahi et al. 2022). All the sources within a square box of $40^\circ$ side centred at $l = 80^\circ$ and $b = 0^\circ$ were included to account for the spill-over due to the PSF.

For two extended sources with potential impact on the characterisation of the cocoon emission, we replaced the 4FGL-DR3 models with dedicated models provided by recent in-depth studies. The SNR Cygnus Loop is modelled following the analysis by Tutone et al. (2021) in the 0.1–100 GeV energy range. We used two templates based on X-ray (ROSAT 0.1–2.4 keV) and UV (GALEX 1771–2831 Å) data, each of them associated to a log-parabola spectrum. Emission from this source above 100 GeV is expected to be small, and we visually checked in the residuals that there were no excess or deficit of counts at the location of the Cygnus Loop.

The SNR Cygnus Loop is modelled following the analysis by Tutone et al. (2021) in the 0.1–100 GeV energy range. We used two templates based on X-ray (ROSAT 0.1–2.4 keV) and UV (GALEX 1771–2831 Å) data. However, due to the bright emission from PSR J2021+4026 (MAGIC Collaboration 2023), this possibility is not considered in our study.

The SNR Cygnus Loop is modelled following the analysis by Tutone et al. (2021) in the 0.1–100 GeV energy range. We used two templates based on X-ray (ROSAT 0.1–2.4 keV) and UV (GALEX 1771–2831 Å) data. Emission from this source above 100 GeV is expected to be small, and we visually checked in the residuals that there were no excess or deficit of counts at the location of the Cygnus Loop.

Figure 5 shows the gamma-ray count map in the region of interest and the excess counts associated with the cocoon obtained by subtracting from the data counts the best-fit model presented in Sect. 3.6. The cocoon, that is the excess shown in the right panel of Fig. 5, was initially modelled as in 4FGL-DR3: a 2D Gaussian with $r_{353} = 3'$ (Ackermann et al. 2011) and a spectrum described by a log-parabola function. The morphological and spectral models were refined later during the analysis.

3.3. Analysis framework

The analysis is performed using Fermi-Tools v2.0.8 and a modified version of Fermipy v1.0.1 that enables the use of catalogue 4FGL-DR3. Models are fit to the data via a binned maximum likelihood analysis with Poisson statistics. Emissions are binned on a grid with 10 bins per decade in energy and on maps with a pixel size of 0.1° in arrival direction.

Throughout the paper, we compare several models for the region and the source of interest. In the simpler cases we use the

---

2 We used files iso_P8R3_SOURCE_V3_PSFn_v1.txt with n the PSF event type, from https://fermi.gsfc.nasa.gov/ssc/data/access/lat/BackgroundModels.html
AIC is taken as the one best representing the data at the smaller maximum likelihood of the model. The model providing the smallest $k$ with $\text{AIC} = 2k - 2 \ln L$ is the maximum likelihood of a more parsimonious emission model with fewer free parameters (null hypothesis) and $L$ is the maximum likelihood of the more complex model that we want to test (test hypothesis). In the null hypothesis $T S$ is distributed as a $\chi^2$ with a number of degrees of freedom $n$ equal to the difference of degrees of freedom between the two models. This is only valid for nested models, that is if the model in the null hypothesis can be obtained from the model in the test hypothesis by fixing some of its parameters to values in the interior of the allowed range (for instance Protassov et al. 2002).

For non-nested models, we use the Akaike Information Criterion (AIC). The AIC of a model is defined as:

$$\text{AIC} = 2k - 2 \ln L,$$

with $k$ number of free parameters in the model and $L$ the maximum likelihood of the model. The model providing the smallest AIC is taken as the one best representing the data at the smaller cost in terms of free parameters according to information theory (for instance Burnham & Anderson 2002).

Throughout the paper, we use the method described in Bruel (2021) to assess the goodness of fit of the different models considered. Practically, we show the deviation of the data with respect to the model in units of significance based on the Poisson statistics using the so-called PS maps.

The analysis starts with a preliminary optimisation of the emission model via a procedure described in Appendix A.1. In subsequent steps, unless stated otherwise, we keep free the normalisations of the gas and IC components, as well as the normalisations and spectral parameters of the three pulsars PSR J2021+4026, PSR J2021+3651, and PSR J2032+4127, the two $\gamma$ Cygni extended components, and the two Cygnus Loop extended components. The normalisation of the subdominant isotropic background is fixed after the preliminary iterative optimisation of all the sources in the ROI due to the possible degeneracy with the IC component. The best-fit normalisation obtained is $0.91 \pm 0.02$ (with variations $\leq 0.01$ for the different spin temperature values).

### 3.4. Morphological analysis

In this section, we aim at characterising the morphology of the cocoon. As a first step, we optimised the position and extension of the 2D Gaussian model used in Ackermann et al. (2011) and the LAT catalogues to describe the extended cocoon emission, following the methodology described in Appendix A.2.

The corresponding PS map is displayed in Fig. 6 (panel a, top row). An excess appears in the central region of the cocoon, in part reminiscent of the two main peaks of ionised gas column density within the Cygnus X cavities (Fig. 4). Therefore, we tested the addition of a central component in the cocoon region using two alternative models: either the ionised gas template clipped at the boundaries of the cavities (defined as contours above $1.85 \times 10^{-6} \text{ W m}^{-2} \text{ sr}^{-1}$ emission at 8 $\mu$m), or two Gaus-
sians with free extensions and positions, initialised at the main peaks in the ionised gas map. All newly added sources on top of the Gaussian model for the extended cocoon component here and elsewhere in this section are modelled using a power-law spectrum.

For the models described above, extended deviations are still apparent (see Fig. 6 top row). The largest excess appears in the western part of the cocoon at $l \sim 78.8^\circ$ and $b \sim 3.7^\circ$. It does not overlap with any known sources or structures in the gas maps. A second extended region of positive deviations appears at the edge of our ROI, at $l \sim 84.6^\circ$ and $b \sim -5.6^\circ$ towards the southern arc of the Cygnus SB as imaged in soft X-rays (Cash et al. 1980). We added two additional Gaussian components to model those excesses, hereafter referred to as western and off-field excesses. We initialise the Gaussian centres on the excess peaks, and fit their positions and extensions. This results in a significant likelihood improvement ($\Delta \text{ln } L \sim 200$) for all models of the central cocoon component.

The different models for the cocoon central component are compared in Table 2. The addition of the central cocoon component on top of the 2D Gaussian for the extended one provides a significant improvement in likelihood ($\Delta \text{ln } L > 240$). Conversely, a model including the ionised gas map without the extended cocoon Gaussian component resulted in a marked degradation of the likelihood ($\Delta \text{ln } L = -600$).

The model including the ionised gas template for the central cocoon component provides the largest likelihood and the smallest AIC, and therefore it is the one favoured by our analysis. It is strongly preferred over two additional Gaussian components at the peaks in the ionised gas distribution ($\Delta \text{AIC} < -124$), which strengthens the evidence for a correlation between part of the

---

#### Fig. 5. Map of data counts in the full 0.5 GeV–1 TeV energy range (left) and map of excess counts obtained by subtracting from the data counts the best-fit model presented in Sect. 3.6 except for the components associated to the cocoon, namely FCES G78.74+1.56, FCES G80.00+0.50, and FCES G78.83+3.57 (right). The dashed contours correspond to the 8 $\mu$m emission from MSX data at $1.85 \times 10^{-6} \text{ W m}^{-2} \text{ sr}^{-1}$. The contours correspond to the column density of the ionised gas template at $3.5 \times 10^{20} \text{ H cm}^{-2}$. The circle and the dashed circle correspond to the position and $r_{\text{eff}}$ of FCES G78.83+3.57 and FCES G78.74+1.56 respectively.
gamma-ray signal and the ionised matter distribution. This conclusion is supported by visual inspection of the deviations in Fig. 6 (bottom row, panels b and c).

We also tested the full ionised gas map, that is not clipped at the boundaries of the cavities, but this yielded a smaller likelihood. The interpretation of this result is not straightforward. The reason may be physical, for example related to confinement of the particles in the cavities. It could also be related to limitations in the analysis, such as systematic biases in the emission measure map extracted from Planck data, approximations in the derivation of the ionised gas column density, or degeneracies with other gas templates outside the two main peaks in the map.

Spatial parameters for the multiple overlapping extended sources may be degenerate to some level. To robustly determine their values, we performed an iterative fit of the positions and extensions of all 2D Gaussian sources discussed in this section for the best model. The iterations proceed until the log-likelihood improvement between two iterations is smaller than 1. The iterative fit converged after 6 iterations, with a total improvement in \( \ln L \) of 30.

---

Fig. 6. PS maps for different models considered in the morphological analysis. On the top panels we can see the PS maps for different morphological models: (a) one extended Gaussian, (b) one extended Gaussian plus two smaller Gaussians at the peaks of the ionised gas column density distribution, (c) one extended Gaussian plus the ionised gas template. On the bottom panels, we can see the PS maps for the same models with the addition of two Gaussians for the western and off-field excesses in the model (ultimately labelled FCES G78.83+3.57 and FCES G85.00−1.78). The bin size is 0.1° and the maps were smoothed for display with a kernel of size 0.13°.
3.5. Spectral analysis

In this section, we aim at characterising the spectral properties of the FCES sources. Based on the best morphological model derived in the previous section, and for each emission component listed in Table 3, we tested three spectral models: a simple power law (PL), a log-parabola (LP), and a smooth broken power law (SBPL). The expressions for these models are given in Appendix A.3.

<table>
<thead>
<tr>
<th>Model</th>
<th>$\Delta \ln \mathcal{L}$</th>
<th>$\Delta \text{AIC}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extended Gaussian + Western + Off-field</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Extended Gaussian + 2 Gaussians (IG peaks) + Western + Off-field</td>
<td>316 [246, 316]</td>
<td>$-622 [-622, -492]$</td>
</tr>
<tr>
<td>Extended Gaussian + IG template + Western + Off-field</td>
<td>435 [379, 435]</td>
<td>$-868 [-868, -746]$</td>
</tr>
</tbody>
</table>

Notes. In $\mathcal{L}$ and AIC values are provided as differences with respect to the simplest model with only one 2D Gaussian for the extended emission of the cocoon. The intervals correspond to the minimum and maximum spin temperatures considered. The Western and Off-field components are named later FCES G78.83+3.57 and FCES G85.00–1.78.

Table 3 provides the best-fit morphological parameters and $TS$ for the extended emission components considered in the morphological analysis.

<table>
<thead>
<tr>
<th>Name</th>
<th>$I(\circ)$</th>
<th>$b(\circ)$</th>
<th>$r_{68}(\circ)$</th>
<th>$TS$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCES G78.74+1.56 (CoExt)</td>
<td>$78.7 \pm 0.1^{+0.0}_{-0.2}$</td>
<td>$1.56 \pm 0.06^{+0.07}_{-0.02}$</td>
<td>$4.4 \pm 0.1^{+0.1}_{-0.1}$</td>
<td>2751 [2436, 2824]</td>
</tr>
<tr>
<td>FCES G80.00+0.50 (CoCent)</td>
<td>$...$</td>
<td>$...$</td>
<td>$...$</td>
<td>1267 [1267, 1301]</td>
</tr>
<tr>
<td>FCES G78.83+3.57 (CoWest)</td>
<td>$78.8 \pm 0.1^{+0.0}_{-0.1}$</td>
<td>$3.6 \pm 0.1^{+0.0}_{-0.0}$</td>
<td>$0.9 \pm 0.1^{+0.0}_{-0.0}$</td>
<td>93 [93, 106]</td>
</tr>
<tr>
<td>FCES G85.00–1.78 (OffExc)</td>
<td>$85.0 \pm 0.4^{+0.4}_{-0.2}$</td>
<td>$-1.8 \pm 0.25^{+0.3}_{-0.2}$</td>
<td>$6.4 \pm 0.2^{+0.1}_{-0.1}$</td>
<td>684 [680, 723]</td>
</tr>
</tbody>
</table>

Notes. The first uncertainties are statistical. The second uncertainties and TS variations are systematic from varying the spin temperature.

The models are compared in Table 4. For the components CoCent and CoWest the best-fit model is the simple PL, with the LP providing a negligible improvement in log-likelihood. The fit of the SBPL for these two components did not converge, presumably due to lack of curvature in the spectrum. Conversely, for CoExt and OffExc the models with curvature (LP or SBPL) provide a large improvement in $\mathcal{L}$. From the Akaike criterion, we can conclude that the SBPL is favoured. Spectral parameters for the best-fit models are presented in the next subsection.

The spectral indices of the components CoCent and CoWest are compatible with each other. If we fix the spectral index of CoWest to the best-fit value for CoCent, we obtain a decrease in log-likelihood of 1.2 [1,2,1.8] (1.1 [1,1,1.3])$\sigma$, where here and in the following variations or uncertainties refer to the different spin temperatures). This demonstrates that the two components have compatible spectral shapes. However, if we model CoExt or CoCent with the same spectral shape as CoExt and a free normalisation, we observe a decrease in log-likelihood of 1.74 [1.9,2.36] and 70.6 [66.6,70.6] ($\Delta \text{AIC} = 145.2 [137.2,145.2]$), respectively. So both CoCent and CoWest have spectra incompatible with the spectrum of CoExt, this time suggesting a different origin of the gamma-ray emission.

We then computed the spectral energy distribution (SED) of the four sources. To this end we performed independent analyses over 4 energy bins per decade between 500 MeV and 1 TeV. For this part of the analysis all spectral-shape parameters are fixed and only normalisations are allowed to vary. However, the normalisations of the gas maps are fixed to the best-fit values obtained for the entire energy range to preserve the local emissivity shape. The IC normalisation is also fixed to the best-fit value obtained for the entire energy range due to the large degeneracy with the extended components. Finally, the normalisations of all pulsars are fixed above 5 GeV because their emission fades off rapidly. The results are shown in Fig. 7. For flux densities (and
obtained for a spin temperature of 100 K, but with a difference in spin temperature values considered. The largest likelihood is the Geminga-like pulsar PSR J1957+... spectrum for CoExt is similar at energies (Abeysekara et al. 2021), confirming earlier indications of a and connects to the cocoon spectrum measured by HAWC CoCent. The spectrum of CoExt extends to higher energies quadrature with those from the spin temperature choice. all derived quantities later) we include in the systematic uncer-
}
A47, page 9 of 29

Table 4. Statistical comparison of different spectral models for the extended sources in Cygnus X.

<table>
<thead>
<tr>
<th>Component</th>
<th>$\Delta \ln L_{\text{LP-PL}}$</th>
<th>$\Delta \ln L_{\text{SBPL-PL}}$</th>
<th>$\Delta \text{AIC}_{\text{SBPL-PL}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCES G78.74+1.56 (CoExt)</td>
<td>33 [27, 35]</td>
<td>42 [36, 43]</td>
<td>−14 [−14, −14]</td>
</tr>
<tr>
<td>FCES G80.00+0.50 (CoCent)</td>
<td>1 [1, 1]</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>FCES G78.83+3.57 (CoWest)</td>
<td>1 [1, 2]</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>FCES G85.00–1.78 (OffEx)</td>
<td>26 [24, 27]</td>
<td>38 [35, 38]</td>
<td>−22 [−22, −20]</td>
</tr>
</tbody>
</table>

Notes. The columns show the log-likelihood differences between a log-parabola and a power-law model, $\Delta \ln L_{\text{LP-PL}},$ or between a smooth broken-power-law and a power-law model, $\Delta \ln L_{\text{SBPL-PL}},$ and the Akaike information criterion difference between a smooth broken-power-law and a log-parabola model, $\Delta \text{AIC}_{\text{SBPL-PL}}.$ The intervals correspond to the minimum and maximum values obtained from variation of the spin temperature.

Table 5. Best-fit values after the final optimisation of the model.

<table>
<thead>
<tr>
<th>Component</th>
<th>Model</th>
<th>$N_0$ (cm$^{-2}$ s$^{-1}$ MeV$^{-1}$)</th>
<th>Spectral parameters</th>
<th>$E_b$ (GeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCES G78.74+1.56 (CoExt)</td>
<td>SBPL</td>
<td>$8.6 \pm 0.6^{+0.2}_{-0.9} \times 10^{-11}$</td>
<td>$1.67 \pm 0.05^{+0.02}_{-0.01}$</td>
<td>$2.12 \pm 0.02^{+0.00}_{-0.01}$</td>
</tr>
<tr>
<td>FCES G80.00+0.50 (CoCent)</td>
<td>PL</td>
<td>$4.7 \pm 0.2^{+0.0}_{-0.0} \times 10^{-11}$</td>
<td>$2.19 \pm 0.03^{+0.00}_{-0.01}$</td>
<td>...</td>
</tr>
<tr>
<td>FCES G78.83+3.57 (CoWest)</td>
<td>PL</td>
<td>$4.8 \pm 0.6^{+0.0}_{-0.0} \times 10^{-12}$</td>
<td>$2.3 \pm 0.1^{+0.0}_{-0.0}$</td>
<td>...</td>
</tr>
<tr>
<td>FCES G85.00–1.78 (OffEx)</td>
<td>SBPL</td>
<td>$1.7 \pm 0.5^{+0.7}_{-0.0} \times 10^{-11}$</td>
<td>$0.7 \pm 0.3^{+0.2}_{-0.0}$</td>
<td>$2.12 \pm 0.04^{+0.00}_{-0.00}$</td>
</tr>
</tbody>
</table>

Notes. The first uncertainties are statistical and the second uncertainties are systematic and result from varying the spin temperature. $N_0$ is given at the reference energy of 1 GeV.

3.6. Final global fit

After the selection of the best spectral models we performed a final optimisation of the ROI, including free normalisation and spectral-shape parameters for the FCES sources. The final spectral parameters of the FCES sources are displayed in Table 5.

Figure 8 illustrates the quality of the ROI model after all the optimisation steps. The PS map show that we obtained a model of the ROI with no deviation above 3$\sigma$ and the fractional deviation in the bottom left panel show no deviation above 10% in the central part of the ROI. This model serves as a reference for the following spectro-morphological analysis. The largest deviation with a PS value close to 3$\sigma$ lies at $l \sim 84^\circ$, $b \sim 12^\circ$ (at 1$\sigma$ from the Geminga-like pulsar PSR J1957+5033; see Saz Parkinson et al. 2010). The study of this excess is left for another work.

Figure 9 shows the final likelihood values for the different spin temperature values considered. The largest likelihood is obtained for a spin temperature of 100 K, but with a difference in log-likelihood <1 with respect to the reference value of 250 K. The largest difference $\Delta \ln L \sim 6$ is found for the optically thin case.

After the final global fit the normalisation of the $H_1$ emissivity in the local arm is 0.97 ± 0.01–0.17. The normalisation is in reasonable agreement with the average value for the local neighbourhood from Casandjian (2015). Under the hypothesis that the same CR population interacts with atomic, molecular and dark gas in the local arm, we can use the normalisation of the gas maps to infer the DNM in Fig. 3 closely resembles the structures in the local arm and the Cygnus complex, which is less obvious due to the lack of knowledge on the distribu-

diagram {A47, page 9 of 29}
increases by $\sim 16\%$ for a spin temperature of 100 K and decreases by $\sim 5\%$ for the optically thin case.

### 3.7 Spectro-morphological analysis

#### 3.7.1 Extension and position versus energy

We first tested if the best-fit spatial model of the CoExt and CoWest components changes as a function of energy. The component OffExc is left aside in the spectro-morphological analysis because it is displaced regarding the Cygnus X region which we aim to study in this paper, and it lies on the border of the ROI, and therefore its characterisation may not be optimal.

We fitted the extension and position of CoExt and CoWest in five energy bands: 0.5–1.6 GeV, 1.6–5 GeV, 5–16 GeV, 16–50 GeV, and 50–1000 GeV. As shown in Sect. 3.5, CoWest has a softer spectrum: above $\sim 5$ GeV the source flux becomes very low and its TS is below 25. Therefore, fitting the position and extension of the source above $\sim 5$ GeV is not possible. In this section, the diffuse components, that is the gas maps and the IC component, are fixed, while the two components of Cygnus Loop are fixed above 5 GeV because of their very steep spectrum. OffExc is fixed due to its off-centre position.

The results are shown in Fig. 11. The top panel shows the extension as a function of energy for both sources. There is no indication of an evolution of the extension as a function of energy, and the values in different energy bands are compatible with that obtained in the broadband analysis. The lower panels show the best-fit centroid positions for the two components. For CoExt, all positions are compatible with each other and the broadband fit within 1σ. For CoWest, we can see a hint of evolution of the position in the first two bins, but the two values are compatible within 2σ with the value from the broadband fit over the full energy range.

#### 3.7.2 Spectral variations across the extended sources

In this section, we search for spectral variations across the emitting regions for CoExt and CoCent. We started by examining CoExt. To this end, we replaced the Gaussian model with a combination of rings and segments. We tested several combinations but here we describe the profile obtained with a combination of: a central disk of radius 0.7°; five rings of external radius 1.7°, 2.7°, 3.7°, 4.7° and 6°, that can be decomposed into four segments spanning 90° in azimuth; and two large rings of external radius 7.4° and 8.9°. This somewhat arbitrary setup was chosen to ensure a minimal TS (at least 25) in every segment and ring (see Fig. B.1). Eventually, however, the wider ring has a low TS ($\sim 10$) therefore its parameters have to be interpreted with some caution.

All the components are modelled using a LP spectrum with parameters initiated at the best-fit values found in Sect. 3.5. The LP model was chosen instead of the SBPL model for this part of the analysis because it yields more stable results when fitting
several free components at once. For this section the two components of Cygnus Loop are fixed due to their off-centre position. OffExc is also fixed due to its off-centre position and proximity with the border of the ROI.

We also tested a combined description of CoExt and CoCent via rings and segments by removing the ionised gas template from the emission model, but the highly structured central part was poorly described by the latter model. Thus, we decided to proceed with the ionised gas map in the model, and we used the combination of segments and rings to only represent the description of the emitting region, that is there are no significant variations of the spectrum as a function of distance from the centre. Conversely, we find a model improvement in step C, which demonstrates that the emission is not azimuthally symmetric in intensity. The likelihood improvement is equally shared representing a 2D Gaussian with concentric rings and a central disk. However, this is not a cause for concern as the decrease in log-likelihood is small. Step B does not provide an improvement in likelihood. The values of $\Delta \ln L$ and $\Delta$AIC for the four steps are provided in Table 6.

For step C, a few different orientations for the segments were tested, and we present the results for the one yielding the best likelihood. The values of $\Delta \ln L$ and $\Delta$AIC for the four steps are provided in Table 6.

The degradation in step A is due to the approximation of representing a 2D Gaussian with concentric rings and a central disk. However, this is not a cause for concern as the decrease in log-likelihood is small. Step B does not provide an improvement in the description of the emitting region, that is there are no significant variations of the spectrum as a function of distance from the centre. Conversely, we find a model improvement in step C, which demonstrates that the emission is not azimuthally symmetric in intensity. The likelihood improvement is equally shared by all the rings concerned, and it is not surprising given the diversity of regions inside and outside the plane spanned by each broad ring. A further improvement in the likelihood is obtained.

Table 6. Variations of $\ln L$ and AIC for the decomposition of FCES G78.74+1.56 in rings and segments.

<table>
<thead>
<tr>
<th>Step</th>
<th>$\Delta \ln L$</th>
<th>$\Delta$AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>$-9 [-13, -9]$</td>
<td>$32 [32, 40]$</td>
</tr>
<tr>
<td>B</td>
<td>$7 [7, 14]$</td>
<td>$-18 [4, 18]$</td>
</tr>
<tr>
<td>C</td>
<td>$58 [47, 58]$</td>
<td>$-86 [-86, -64]$</td>
</tr>
<tr>
<td>D</td>
<td>$35 [35, 42]$</td>
<td>$-10 [-24, -10]$</td>
</tr>
</tbody>
</table>

Notes. Values are provided as difference with respect to the previous step, and, for step A, with respect to the global analysis presented in Sect. 3.5. See the text for details.

Beta, see Appendix A.3 for a definition) were fixed to the initial values from the analysis in Sect. 3.5.

B The parameters $\alpha$ and $\beta$ for the disk and rings were free.

C The innermost five rings (beyond the central disk) were decomposed into four azimuthal segments. Only the normalisations were free and the spectral parameters were fixed to the values obtained in the step B.

D All spectral parameters were free.

We replaced the Gaussian by the aforementioned combination of seven rings and a central disk. Only the normalisation of each template was free, while the spectral parameters ($\alpha$ and

Fig. 8. Final deviation maps for the best-fit model over the entire energy range, on the left as fractional deviation, and on the right using the PS map. The bin size is 0.1° and the size of the smoothing kernel 0.13°.

Fig. 9. Difference in log-likelihood in fits of the final model for different spin temperatures.

Fig. 10. Neutral gas column density in the local arm and Cygnus region, obtained by summing $N$(H), $N$(H2) and an estimate of the column density for the DNM, with conversion factors calibrated on the gamma-ray analysis.
in step D, showing also the presence of azimuthal spectral variations, mostly driven by the two innermost rings and the central disk with an improvement in log-likelihood of 24 (ΔAIC = −15) when only those components have the spectral shape free.

However, the azimuthal variations in the first two rings could be explained by spectral variations across CoCent. To check this hypothesis, we sliced the ionised gas template vertically at $l = 79.8^\circ$ to separate the two main lobes of ionised gas and repeated the step D. This results in an improvement of the log-likelihood smaller than one, meaning that no spectral variation is detectable between the two sides of the source CoCent.

We conclude that the best model is the one combining the ionised gas template to describe CoCent and with CoExt decomposed and fitted as in step D. This is used as a basis for the interpretation of the results in the next section, where the emission profiles extracted from the data is also shown. Some additional plots illustrating the results are provided in Appendix B.

4. Discussion

4.1. The cocoon and its landscape

Our analysis shows that the Cygnus cocoon in the LAT energy band is best described by at least two spatial components with different spectra: a central component, CoCent, with a power law spectrum of index $2.19 \pm 0.03^{+0.00}_{-0.01}$, and an extended component, CoExt, with a smooth broken power law spectrum with indices $1.67 \pm 0.05^{+0.02}_{-0.01}$ below $3.0 \pm 0.6^{+0.12}_{-0.02}$ GeV and $2.12 \pm 0.02^{+0.00}_{-0.01}$ above. A third newly discovered extended emission component,
The striking spatial coincidence of the brightest part of the gamma-ray signal and the contours of the cavity, and to a lesser extent the bright arc of 8 µm emission, but does not coincide with any over-densities in neutral or ionised gas densities (see Figs. 2–4). Its centroid lies at approximately 1° from the γ Cygni SNR, that, if we assume a distance to the Earth of 1.7 kpc, corresponds to a physical distance of ~30 pc.

Under the hypothesis that the observed gamma-ray emission is of hadronic origin we can convert the excess map into an emissivity map. To this aim we divided the excess cube in the analysis energy bins by the exposure cube and the total, neutral plus ionised, gas column density map. The latter quantity is an upper limit to the relevant gas column densities because gas could be distributed over a larger distance along the line of sight compared to the volume probed by the particles in the cocoon. However, we expect most of the gas in this region to be concentrated around the star-forming complex in Cygnus X, and we do not have an alternative simple prescription to estimate the foreground and background column densities to be subtracted. The results are displayed in Fig. 13.

On one hand, we can see an emissivity peak in the cocoon central area coincident with the peaks in the ionised gas distribution (modelled by CoCent in our analysis) with broad wings extending to several degrees from the centre (CoExt). On the other hand, we see a marked peak at the position of CoWest and around the γ Cygni SNR and NGC 6910 stellar cluster. Although position and spectral similarity to CoCent suggest that this source is related to the cocoon, the interpretation is not obvious. CoWest may be related to gas missing in our model, or else to a nearby source or some peculiar transport configuration that results in an accumulation of particles in this region. In the following for simplicity we concentrate on the interpretation of the two brightest sources in the cocoon area, namely CoCent and CoExt.
extent the resemblance with the extended X-ray emission structure, have suggested that both phenomena may have a common origin: the abundant massive-star population of the region. The most prominent stellar clusters in the regions, the Cyg OB2 association and NGC 6910 cluster, are natural candidates, powerful enough to accelerate particles able to produce non-thermal emission at the observed level.

We evaluated the properties of these two objects, following what was done in Ackermann et al. (2011). For Cygnus OB2, we considered 78 O stars (Berlanas et al. 2020) and a power law mass function of index 1.09 (Wright et al. 2010). For NGC 6910 we assumed a power law mass function of index 0.74 (Kaur et al. 2020) normalised according to Fig. 9 of their paper. We evaluated mass loss rates, cluster wind terminal velocities, and mechanical power of the winds by separating stars in four groups, namely O5 to O3, O9 to O5, B5 to B0, and B8 to B5. The sample is limited to stars heavier than B8 due to the validity range for the reference mass-loss rate model adopted. We assumed standard properties of O stars from Martins et al. (2005) and for B stars from Cox (2000). We used the parametric wind model by Vink et al. (2000). This yields a mass loss rate of $5.1 \times 10^{-4} M_\odot$ yr$^{-1}$ for Cyg OB2 and of $2.7 \times 10^{-4} M_\odot$ yr$^{-1}$ for NGC 6910. The mechanical power of the winds is evaluated to $8 \times 10^{38}$ erg s$^{-1}$ for Cyg OB2 and $4 \times 10^{38}$ erg s$^{-1}$ for NGC 6910. The collective wind terminal velocity therefore is $\sim 2200$ km s$^{-1}$ for both clusters. We show in the next section that such powers are sufficient to account for the observed signal in some scenarios.

We can estimate the physical and angular sizes of the cluster wind termination shock and shocked wind bubble using the formulae in Morlino et al. (2021), which follow the simple models in Weaver et al. (1977); Gupta et al. (2018). If we assume ages of 5 Myr (Berlanas et al. 2020; Kaur et al. 2020) for both clusters and interstellar gas densities of $5 \text{ H cm}^{-3}$ we obtain a size of the wind termination shock of 40 pc for Cyg OB2, and of 33 pc for NGC 6910. The total size of the wind bubble is 200 pc for Cyg OB2, and 180 pc for NGC 6910. Figure 12 shows that the sizes of the termination shocks are comparable to that of the central emission component, with CoWest being located at the edge of the termination shock from NGC 6910, while the sizes of the wind bubbles compare well to that of the cocoon extended emission component.

While these results tend to lend support to the idea that Cyg OB2 and NGC 6910 may be the sources ultimately responsible for the observed gamma-ray emission, we emphasise that the modelling of the winds and bubbles is without any doubts oversimplified. Several effects can be expected to affect the results and weaken the similarity of the gamma-ray emission and expected SB signatures. Generally, the classical theory from Weaver et al. (1977) is known to underestimate the radiative losses, hence overestimate the size of the bubble. Hydrodynamical simulations reveal enhanced radiative losses due to instabilities at the interfaces. This result in a bubble being $\sim 40\%$ smaller than predicted by the classical analytical solution for well-developed bubbles, in agreement with observations (Krause & Diehl 2014). At earlier stages, before SB breakout from the parent molecular cloud, the dense and fractal medium surrounding the cluster drives turbulent mixing and efficient cooling, resulting in a reduction of $\sim 30\%$ in bubble size for parameters relevant to Cyg OB2 and NGC 6910 (Lancaster et al. 2021). More fundamentally, the simple bubble model from Weaver et al. (1977); Gupta et al. (2018); Morlino et al. (2021) may not be straightforwardly applied to Cyg OB2, which is not a compact cluster but instead presents multiple substructures with a 50% containment radius of stellar members spanning 0.2°, that is 5 pc at a distance of 1.7 kpc (Berlanas et al. 2019). Furthermore, as illustrated in Fig. 12, the bubbles from the two stellar clusters may have interacted and it is not clear how this would have impacted the development of the whole region and whether this should have left specific signs that we should now see.

Therefore, the connection of the observed gamma-ray signal with gas structures imprinted by the development of a SB is far from obvious. Actually, the separation of the cocoon emission into CoExt and CoCent, as well as the correlation of the innermost bright signal with ionised gas, can be interpreted in a way that weakens the link between the gamma-ray emission and the cavity delineated by photo-dissociation regions. Indeed, the emission could arise from the interaction of freshly accelerated CRs with the ionised gas inside the cavity, the latter playing no role. These CRs extend much beyond the limits of the cavity, as was already clearly observed in Ackermann et al. (2011), and their interactions with the ambient medium give rise to the source CoExt.

The origin of the CRs powering the cocoon emission can therefore be unrelated to the Cyg OB2 association and NGC 6910 cluster (in the sense that they play no role as a whole, but they can harbour or have harboured the actual source). The central region actually contains a handful of extremely energetic objects and potential particle sources.

We can list: the $\gamma$ Cygni SNR (G78.2+0.1) with a probable distance of 1.7 to 2.6 kpc from association with the $\gamma$ Cygni nebula (Leahy et al. 2013) and dynamic properties estimated in Leahy et al. (2020) as ejecta mass of $5 M_\odot$, age of $9.4^{+2.3}_{-1.6}$ kyr, and supernova (SN) energy of $6.3^{+2.8}_{-2.0} \times 10^{50}$ erg; the $\gamma$ Cygni pulsar, PSR J2021+4026, associated with the $\gamma$ Cygni SNR and with a spin-down power of $1.2 \times 10^{35}$ erg s$^{-1}$ (Ray et al. 2011); PSR J2032+4127, a pulsar in a highly eccentric binary system with a Be-type star (Lyne et al. 2015), probably part of Cyg OB2, with an orbital period of 45-50 yr, a spin-down power of $1.5 \times 10^{35}$ erg s$^{-1}$, and a characteristic age of $\sim 200$ kyr (Ho et al. 2017).
Furthermore, the centroid of the emission from CoExt and the peak in the emissivity map do not coincide with any of the potential particle accelerators, stellar clusters or others (Figs. 12 and 13). Therefore, we tried to account for our observations in a generic way, with a simple diffusion model based on an unspecified source and not exclusively relevant to massive star clusters and their associated SBs. We introduce in the following sections the model framework used and the parameter setups yielding satisfactory fits to the data.

4.2. A simple diffusion-loss framework for the cocoon

Given the layout of the emission exposed in the previous subsection, with significantly extended radiation from a region reaching well beyond the vicinity of potential sources, it seems reasonable to consider that gamma rays are produced by particles that were released by one or several sources some time ago and were transported in the surrounding medium since then. In this section, we aim to provide a quantitative assessment of this idea.

We interpret the observations in the framework of a one-zone diffusion-loss transport model where particles are continuously injected at a point in space for some duration and then experience diffusive transport in a uniform and isotropic medium. This is very likely an overly simplistic description of the processes at stake because there may be multiple sources, not all of them can be assumed to be of negligible size, the medium is probably not uniform over the few hundreds of parsecs probed by the emission, and there may be other transport processes than diffusion. Yet, our goal is to draw a few key inferences from the observables and we defer more advanced modelling efforts to subsequent publications. Moreover, we show later that such a modelling with a very limited number of free parameters can yield a fairly good representation of the observables.

The full formalism of the model framework is provided in Appendix C. Ultimately, the main parameters of the model are: injection luminosity $Q_\text{inj}$, power law injection spectrum slope $\alpha$, characteristic injection duration $t_\text{inj}$, diffusion duration $t_\text{diff}$, and diffusion coefficient normalisation $D_0$. We explored a large parameter space for these four parameters and fitted the predictions to the results of the gamma-ray analysis.

The diffuse emission from the Cygnus cocoon is very extended, with an angular size of $4.4\pm0.1^\circ$ for CoExt that translates into a $\sim 130$ pc length at a distance of 1.7 kpc. A more compact and central emission component CoCent is correlated with the distribution of ionised gas within a radius of about 50 pc; the spectrum of CoExt is flat and that of CoCent, although significantly softer, is also pretty hard compared to interstellar emission on larger scales.

Given these observables, we proceeded to educated guesses for the main parameters of the model, considering first the case of a hadronic scenario. The typical extent of the emission provides a constraint on the diffusion length, that is on the product of diffusion coefficient and diffusion time:

$$r_d = \sqrt{4D t_\text{diff}} \gtrsim 100 \text{ pc},$$

$$D = D_\text{inj}(10 \text{ GeV}) = 10^{29} \text{ cm}^2 \text{ s}^{-1} \Rightarrow t_\text{diff} \approx 7.5 \text{ kyr},$$

$$D = D_\text{inj}(10 \text{ GeV}) = 10^{27} \text{ cm}^2 \text{ s}^{-1} \Rightarrow t_\text{diff} \approx 0.75 \text{ Myr}.$$  

If diffusion has the average properties inferred for transport over large scales in the Galaxy (Trotta et al. 2011), defined by the coefficient $D_\text{inj}$, particles need less than 10 kyr to fill a volume that would account for the extent of the observed emission at a distance of 1.7 kpc. Conversely, if diffusion is for some reason strongly suppressed by one to two orders of magnitudes as inferred for a variety of sources including star-forming regions (Aharonian et al. 2019; Abyeysekara et al. 2017; Abramowski et al. 2015), and is characterised by coefficient $D_\text{supp}$, then about 1 Myr is needed.

The emissivity enhancement inferred for the cocoon is comparable to the local emissivity within a factor of two to three depending on the energy range, such that the CR energy density $u_\text{CR}$ in the region is similar to the one in the solar neighbourhood, $u_\text{CR, local}$. This makes it possible to constrain the properties of particle injection, namely its power $L_\text{inj}$ and typical duration $t_\text{inj}$.

$$4\pi r_d^3 \times u_\text{CR} \approx \frac{1}{2} L_{\text{inj}}t_{\text{inj}},$$

$$u_\text{CR} \approx u_{\text{CR, local}} \approx 1 \text{ eV cm}^{-2},$$

$$L_{\text{inj}}t_{\text{inj}} \approx 4 \times 10^{50} \text{ erg},$$

$$L_{\text{inj}} = 10^{38} \text{ erg s}^{-1} \Rightarrow t_{\text{inj}} \approx 0.1 \text{ Myr}.$$  

As computed in the previous subsection, the mechanical luminosity of the most prominent star clusters in Cygnus is in the range $4\times10^{38}$ erg s$^{-1}$. Such a power source can deliver particle injection at a level of $10^{38}$ erg s$^{-1}$, pending efficient particle acceleration with a yield of $\sim 10^{-30\%}$ (by some unspecified mechanism at this stage). In that case, the inferred CR density enhancement can be attained if injection lasts over about 100 kyr (and particles accumulate in the volume, see the discussion in the next paragraph). If particle acceleration is less efficient or the source is less powerful, by about an order or magnitude, then injection has to proceed on Myr time scales. Alternatively, a supernova producing $10^{50}$ erg of accelerated particles and releasing the majority of them over $\sim 3\times10$ kyr would provide an injection power of $3\times10\times10^{38}$ erg s$^{-1}$ and thus allow short-lived injection.

Scenarios with $t_{\text{inj}}$ much smaller than $t_{\text{diff}}$ are not viable because particles spread out and leave the volume too rapidly, which results in too flat intensity profiles and too steep spectra (because energy-dependent diffusion depletes the particle population at the high end of the spectrum). So $t_{\text{inj}}$ has to be comparable to or greater than $t_{\text{diff}}$, with the additional constraint that sufficient energy should be released within a time $t_{\text{diff}}$ to match the observed level of emission. In practice, this means that: for average interstellar diffusion, the region is filled over a $\sim 10$ kyr timescale, thus requiring a strong enough source with injection power $\sim 10^9$ erg s$^{-1}$ typical of a SN; alternatively, weaker sources such as the star clusters in Cygnus with injection power $\sim 10^{37-38}$ erg s$^{-1}$ require moderate to strong diffusion suppression and transport occurring over hundreds to thousands of kyr.

These considerations remain mostly valid in the case of a leptonic scenario. The main difference with a hadronic scenario is the importance of energy losses, mostly from synchrotron radiation and inverse-Compton scattering. Yet, for an interstellar magnetic field strength $B = 3 \mu$G and optical and infrared interstellar radiation fields with total energy density of about 1 eV cm$^{-3}$, such as those predicted in the large-scale model of Popescu et al. (2017) at the Galactic position of the cocoon, particles with energy below $300\text{–}400 \text{ GeV}$ have a cooling time of 1 Myr or more. The transport of electrons over the distances and time scales considered above may therefore be little affected by energy losses in many scenarios. The actual situation is however far more complex because radiation densities in the innermost region of the cocoon are much stronger than the large-scale...
4.3. Possible diffusion scenarios for the cocoon

We tested the hypothesis that components CoExt and CoCent are produced by a single population of non-thermal particles. In that context, CoCent is gas-related emission (pion decay in hadronic scenarios, and Bremsstrahlung in leptonic scenarios) that not including a contribution from background CRs provides much better fits to the data. Atomic and molecular gas in the cocoon region enter twice in the analysis: in the fit to the Fermi-LAT data over a large ROI, where they trace emission from the background population of CRs, and in the interpretation of the extended emission from CoExt and CoCent, where they are associated to an additional population of particles. Conversely, the ionised gas template enters just once, and the associated emission may therefore comprise contributions from background CRs and from an additional population of particles. In practice, what is fitted to the spectrum of CoCent is the spectrum of the extended region can be considered close to the local one; see Sect. 3.6). We tested both options and, as illustrated below, it turns out that not including a contribution from background CRs provides much better fits to the data.

4.3.1. Hadronic scenarios

To begin with, we present the result of complete calculations for hadronic scenarios. Following the above discussion of the most likely diffusion-loss model setups given the observables at hand, we present the results of four scenarios dubbed H1, H2, H3 and H4, the parameter sets of which are presented in Table 7. The H1 and H2 scenarios feature constant injection of a hard spectrum of protons and mainly differ by the diffusion time (300 kyr instead of 100 with respect to the interstellar average). The H3 and H4 scenarios correspond to shorter-lived injection over 3 or 30 kyr and transport over 10 or 100 kyr in a medium with no or moderate diffusion suppression (by a factor 10 or 100 with respect to the interstellar average). Scenario H4 actually corresponds to a scaled version of scenario H3 (multiplying injection and diffusion times and dividing diffusion normalisation and injection power by the same amount, that is ten), such that both setups are completely identical in terms of predictions.

Figure 14 shows the total fitted spectra for CoExt and CoCent for model setups H1 and H2. The predicted shape for the CoExt spectrum is too steep. A steeper predicted spectrum for the CoCent spectrum is in good agreement with the data, while that for the CoCent spectrum is too steep. A steeper predicted spectrum for CoCent is obtained because higher-energy particles leave the innermost regions more rapidly than lower-energy particles, and also because it contains a contribution from the background CR population that has a steeper spectrum.
Spectral energy distribution (GeV/cm$^2$/s)

$\chi^2$ to the emission from a background population of CRs (see text). The full
is done first, with the injection luminosity as fitting parameter, followed
by a second fit to the spectrum of FCES G80.00$^{+0.50}_{-0.50}$. A fit of the model to the spectrum of FCES G78.74$^{+1.56}_{-1.56}$ + H$_2$ (bottom). A fit of the model to the spectrum of FCES G78.74 + 1.56 is done first, with the injection luminosity as fitting parameter, followed
by a second fit to the spectrum of FCES G80.00$^{+0.50}_{-0.50}$, with ionised gas
column density as fitting parameter. This last fit includes a contribution to
the emission from a background population of CRs (see text). The full
extent of the error bars corresponds to the quadratic sum of the statis-
tical and systematic uncertainties, while the caps mark the contribution
of the statistical uncertainty only. The first $\chi^2$ corresponds to the fit to
FCES G78.74$^{+1.56}_{-1.56}$ and the second one to FCES G80.00$^{+0.50}_{-0.50}$.

Interestingly, a much better fit to the spectrum of CoCen is obtained when not adding a local emissivity contribution to the
model spectrum for the central region, at the expense of higher
fitted column density for the ionised gas. This is illustrated in the
top two panels of Fig. 15. Emission from the ionised gas in the innermost
regions of the cocoon would then arise only from CRs produced in Cygnus. Background pre-existing CRs
may have been evacuated in the past during the SB growth,
for instance by advection in the stellar winds. Alternatively,
the spectral signature of these pre-existing CRs may have been
absorbed by another component in the fit to the LAT observa-
tions (for instance by the molecular gas or DNM templates that
have a high degree of correlation with ionised gas in the central
region). Since the spectral fit is so much better when not includ-
ing the contribution from background CRs for CoCen (this is
true also in leptonic scenarios), we present in the following only
results produced with this approach. We note, however, that this
has almost no influence on the intensity and emissivity profiles
presented thereafter.

For model setup H1 (resp. H2), the fit implies a proton
injection luminosity of $3.6 \times 10^{36}$ erg s$^{-1}$ (resp. $3.2 \times
10^{37}$ erg s$^{-1}$), which would correspond to proton injection
efficiencies $<0.5\%$–$1\%$ (resp. $<5\%$–10\%) for the Cyg OB2 or
NGC 6910 clusters. Such low efficiencies are consistent with
the assumed flat injection spectra with $\alpha = 2.0$, at least in
the framework of diffusive shock acceleration. For model setup H3
(resp. H4), the fits implies a much higher proton injection luminos-
ities of $2.8 \times 10^{37}$ erg s$^{-1}$ (resp. $2.8 \times 10^{38}$ erg s$^{-1}$). This
would correspond either to a high particle acceleration efficiency
of $\sim 20\%$ in an SN with a $10^{51}$ erg explosion kinetic energy,
with subsequent release of accelerated particles over a timescale

![Fig. 14. Fitted model spectra for the FCES G78.74$^{+1.56}_{-1.56}$ and FCES G80.00$^{+0.50}_{-0.50}$ emission components, for model setup H1 (top) and H2 (bottom). A fit of the model to the spectrum of FCES G78.74$^{+1.56}_{-1.56}$ is done first, with the injection luminosity as fitting parameter, followed by a second fit to the spectrum of FCES G80.00$^{+0.50}_{-0.50}$, with ionised gas column density as fitting parameter. This last fit includes a contribution to the emission from a background population of CRs (see text). The full extent of the error bars corresponds to the quadratic sum of the statistical and systematic uncertainties, while the caps mark the contribution of the statistical uncertainty only. The first $\chi^2$ corresponds to the fit to FCES G78.74$^{+1.56}_{-1.56}$ and the second one to FCES G80.00$^{+0.50}_{-0.50}$.

![Fig. 15. The top two panels are the same as in Fig. 14, without a contribution to the emission from a background population of CRs in the fit to the spectrum of FCES G80.00$^{+0.50}_{-0.50}$. The bottom panel is the corresponding figure for scenario H3. The first $\chi^2$ corresponds to the fit to FCES G78.74$^{+1.56}_{-1.56}$ and the second one to FCES G80.00$^{+0.50}_{-0.50}$.](image-url)
Fig. 16. Intensity and emissivity radial profiles in three different gamma-ray energy bands for the FCES G78.74+1.56 and FCES G80.00+0.50 emission components, compared to predictions for model setup H1. In the intensity plots, the intensity distribution corresponding to the best-fit two-dimensional Gaussian model is displayed for comparison as a dotted line. In the emissivity plots, the local emissivity and its uncertainty in each energy range are displayed for comparison as a dotted line and a shaded band. The data points correspond to a decomposition of the emission into the ionised gas template, a central disk, two outer rings, and five intermediate rings split azimuthally into four segments. For the latter, we displayed the corresponding angular range only for one segment in each ring and introduced a small horizontal shift of the others, for readability. The full extent of the error bars corresponds to the quadratic sum of the statistical and systematic uncertainties, while the caps mark the contribution of the statistical uncertainty only.

Figure 16 displays the predicted intensity and emissivity profiles for both central and extended model components in scenario H1, compared to the values inferred from the spectromorphological analysis in segments, in three different energy bands: 0.5-2.2 GeV, 2.2-10 GeV, and 10 GeV-1 TeV. The intensity and emissivity are shown as radial profiles, with error bars indicating statistical and systematic uncertainties. The data points are compared to model predictions, with the best-fit two-dimensional Gaussian model shown as dotted lines and shaded bands. The caps on the error bars indicate the contribution of the statistical uncertainty.

The analysis of 3 kyr (resp. 30 kyr), or to a lower acceleration efficiency in an SN more energetic than in the canonical picture. The first option, with relatively high efficiency, may be conflicting with our assumption of a flat injection spectrum.
bands: 0.5–2 GeV, 2–10 GeV, and 10 GeV–1 TeV. The comparison for other scenarios is shown in Appendix D. To compare the different model setups, we provide in each panel the $\chi^2$, computed from all significant intensity data points using statistical uncertainties only. There is, however, no formal fit of the model to the measured intensity profiles, and this $\chi^2$ is just a figure of merit to characterise each model setup. The agreement is overall quite good from the centre up to beyond 8°, especially considering the simplicity of the model and the limited number of parameters. Most measurements are within a factor two of the predictions. When subtracting the contribution from the central model component, relatively flat emissivity profiles are obtained for the extended model component, in agreement with the trend inferred from the data analysis. The results lend support to the idea that CoExt and CoCent are produced by the same population of particles. All four model setups are overall equally good at accounting for the data, despite widely different parameter sets.

4.3.2. Leptonic scenarios

We now present the result of complete calculations for leptonic scenarios, in which the emission can be produced by non-thermal Bremsstrahlung and inverse-Compton scattering. We considered two scenarios dubbed L1 and L2, the parameter sets of which are presented in Table 7. Both scenarios feature constant injection of a hard spectrum of electrons and mainly differ by the diffusion time (1 or 3 Myr) and level of diffusion suppression (by a factor 10 or 100 with respect to the interstellar average). The magnetic field is assumed to have a strength $B = 3 \, \mu G$, and the interstellar radiation field model is taken from the large-scale model of Popescu et al. (2017) at the Galactic position of the cocoon (at a 1.7 kpc distance from us). We tested the effect of a stronger interstellar radiation field model, such as the one used in the original Cygnus cocoon paper (Ackermann et al. 2011), and obtained a much poorer fit to our measurements. This stems mostly from the shorter propagation range and different relative contributions of Bremsstrahlung and inverse-Compton to the emission. This scenario is however extreme since it enforces very strong inverse-Compton losses over an extended volume, whereas in reality enhanced radiation fields are expected only in the innermost regions of the cocoon. This is a caveat of the model framework used in this work, which cannot handle inhomogeneous environments. We also tested a leptonic version of scenario H3, but that yields a poor fit to the data.

The fits of the model to the spectra of CoExt and CoCent are displayed in Fig. 17. They are overall pretty satisfactory and yield $\chi^2$ similar to those obtained with the hadronic models, although slightly higher. The diffusion time range in leptonic scenarios seems rather constrained: small ages ≤1 Myr tend to produce too hard spectra for the extended component, while ages ≥3 Myr result in too steep spectra. The injection luminosities resulting from these fits are $6.8 \times 10^{35}$ and $2.4 \times 10^{36}$ erg s$^{-1}$ for the L1 and L2 scenarios, respectively. This translates into electron injection efficiencies at the sub-percent level at most if the mechanical luminosity from Cyg OB2 and NGC 6910 is the power source for particle acceleration. We checked that the corresponding synchrotron emission does not exceed the radio luminosity from Cyg OB2 and NGC 6910.

The corresponding predicted intensity profiles are compared to the measurements in Fig. 18 for model L1. As for the spectra, the fit is slightly degraded compared to that obtained with hadronic models. The best scenario is L1, with diffusion suppressed by two orders of magnitude with respect to the interstellar average and a diffusion time of 3 Myr. A lower level of diffusion suppression results in too flat intensity profiles, underestimating the data in the inner regions and exceeding them at large distances from the injection point. As mentioned above, a smaller diffusion time does not help because it yields a too hard spectrum for CoExt.

We also tested the hypothesis that the observed emission actually is a pulsar halo, following the discovery of very extended gamma-ray emission around some middle-aged pulsars (Abeysekara et al. 2017). In such a scenario, PSR J2032+4127 appears as an interesting candidate because of its location close to the peak of the emission and characteristic age of ~200 kyr. We used the phenomenological two-zone diffusion-loss halo model implementation presented in Martin et al. (2022), using as baseline key parameters: the spin-down power, estimated distance, and characteristic age of PSR J2032+4127 from the ATNF data base4; a broken power law injection spectrum with indices

---

4 https://www.atnf.csiro.au/research/pulsar/psrcat/
The fit of the predicted emission properties to the observed spectra and intensity profiles is relatively good (see Appendix D), although not at the level of those obtained in scenarios H1-H4 and L1-L2. Yet, the implied present-day injection luminosity is of the order of $10^{36}$ erg s$^{-1}$, an order of magnitude larger than the spin-down power of PSR J2032+4127, which dismisses this pulsar as the possible source of the halo. This result seems to be robust against variations of the main model parameters. One cannot exclude, however, that another currently unknown pulsar with the right properties exists in this active star-forming region.

4.4. The origin of the cocoon

To summarise, our extended set of observables for CoCent and CoExt, including a radial profile for the extended component over nearly $10^7$, together with intensity measurements and emissivity estimates in three energy bands from 0.5 GeV to 1 TeV, can be accounted for reasonably well from a simple diffusion-loss model with a small number of free parameters. Several pretty different model setups seem to provide viable explanations of the observations, which suggests that more developed modelling frameworks and, more likely, additional observational data need to be considered in future studies.

An important result is that the data can be explained from one single population of injected particles. This population spans the full region of extended component CoExt, and gives rise to the central component CoCent by interacting with ionised gas in the innermost regions. Both hadronic and leptonic scenarios are viable, although it should be confirmed that leptonic scenarios are still valid in a more realistic modelling framework including non-uniform inverse-Compton losses in the strongly varying radiation fields of the region. All solutions have in common to require a flat particle spectrum at the source, with a power law index 2.0, which points to very recent acceleration.

The solutions are, however, very different in terms of energetics and time scales involved. Setups H1 and L1 feature continuous injection, strong diffusion suppression in the region (by a factor 100 with respect to the large-scale interstellar average), over a spatial extent of more than 200 pc), transport proceeding over several Myr (in agreement with age estimates for Cyg OB2 and NGC 6910), and low acceleration efficiencies (at the sub-percent level in the hadronic scenario if Cyg OB2 and NGC 6910 are the mechanical power source). Setups H2 and L2 are very similar with continuous injection, moderate diffusion suppression (by a factor 10 with respect to the large-scale interstellar average), a more recent injection and transport process over the last 0.3–1 Myr, and five-to-ten times higher acceleration efficiencies with respect to H1 and L1.

Setups H3 and H4 describe an even more recent event, with injection lasting 3 or 30 kyr, transport proceeding over 10 or 100 kyr in a medium where diffusion is not or only moderately suppressed, from a much more powerful source with properties that eventually seem relevant to an SN. Support for these scenarios would imply finding evidence of a middle-aged remnant in the region. The γ Cygni SNR (G78.2+2.1), with its estimated age ~10 kyr, would be an interesting candidate source for scenario H3. Another option could be the remnant that resulted from the explosion giving birth to PSR J2032+4127, ~100–200 kyr ago. Its age is comparable to the diffusion time involved in scenario H4 and is high enough that the remnant has most likely gone undetectable by now. Interestingly, the typical injection

1.8 and 2.2 below and above a break energy of 500 GeV respectively; an injection starting time of 40 kyr; diffusion suppression by a factor of 50 within 50 pc of the pulsar, with a power law dependence in rigidity with index 1/3; a surrounding magnetic field with strength $B = 3 \mu G$ and the interstellar radiation field model from the original Cygnus cocoon paper (Ackermann et al. 2011). We neglected the effect of proper motion on the emission morphology.

---

Fig. 18. Intensity radial profiles in three different gamma-ray energy bands for the FCES G78.74+1.56 and FCES G80.00+0.50 emission components, compared to predictions for model setup L1. The intensity distribution corresponding to the best-fit two-dimensional Gaussian model is displayed for comparison as a dotted line. The $\chi^2$ value correspond to the deviation from the 2D Gaussian fit.
time of 30 kyr and diffusion suppression by a factor 10 in sce-
nario H4 are reminiscent of the results obtained in Nava et al. 
(2019) for the non-linear diffusion of 0.1−1 TeV CRs escaping 
from a supernova remnant in a hot ionised medium.

For comparable model setups, the difference in injection effi-
ciency between leptonic and hadronic scenarios is of an order of 
magnitude at most. This means that mixed lepto-hadronic sce-
narios either imply a relatively high electron-to-proton ratio at 
injection, or a predominance of the hadronic contribution to the 
emission if the electron-to-proton ratio at injection is expected to 
have more classical values of $10^{-2}$ at most.

All the potential sources discussed above, Cyg OB2, 
NGC 6910, the γ Cygni SNR, and the unobserved SNR asso-
ciated with PSR J2032+4127 are displaced with respect to the 
centroid of the emission (Fig. 12). This could suggest an inho-
mogenous transport scenario if one of these sources is indeed 
responsible for the origin of the particle population.

Possible improvements to the modelling presented here 
include the possibility of multiple and extended sources, for 
instance Cyg OB2 and NGC 6910 releasing accelerated particles 
at their respective super-wind termination shocks, and a more 
complete transport scheme, including inhomogeneous diffusion 
and (or) energy losses over such a large volume, or the effect 
of advection. The latter point is particularly relevant in the case 
of high levels of diffusion suppression, as it may dominate the 
transport of the lowest-energy particles and alter their diffusion. 
The contribution from pre-existing CRs, for instance their lep-
tonic emission from inverse-Compton scattering in the dense 
photon fields of the main clusters (Orlando & Strong 2007) or 
their hadronic emission after reacceleration in the turbulent int-
erior of the region (Tolksdorf et al. 2019), certainly deserves a 
more sophisticated treatment than done here. Meanwhile, we can 
qualitatively compare our results to more sophisticated models 
of particle acceleration and transport at cluster wind termina-
tion shocks and in SBs from the literature.

Morlino et al. (2021) presents a model of particle accelera-
tion at the winds of star clusters that predicts a spectrum similar 
to the cocoon central component CoCent, which spans a region 
with size comparable to that of the wind termination shocks from 
Cyg OB2 and NGC 6910 (see Fig. 12). The spatial distribution 
of low-energy particles in their model can be rather flat up to 
a few times the termination shock radius, which is comparable to 
the extended component CoExt. However, their model predicts 
that higher energy particles are more tightly confined around the 
shock, which is in contrast with our results of a harder spectrum 
for CoExt than CoCent. In addition, as already discussed above, 
there is no clear identification of a super-wind termination shock 
in the region, nor is it clear that there is a super-wind emanating 
from Cyg OB2.

Alternatively, Vieu et al. (2022) have shown that their model 
for particle acceleration and transport in SBs can reproduce the 
overall spectrum of the Cygnus cocoon measured by the LAT 
and HAWC in the case of efficient confinement in the bubble 
shell. In these conditions, they show that particle densities are 
rather uniform inside the SB, which might be in good agreement 
with the flat radial emissivity profile of the CoExt component. 
However, it is not obvious that their model can reproduce the 
morphological properties of the observed emission, especially 
its centrally peaked nature, if particles are efficiently trapped in 
an outer shell (the location of which remains unclear in Cygnus).

Recently, Fornieri & Zhang (2022) presented a model of 
gamma-ray emission from Cygnus featuring two CR sources 
(Cyg OB2 and the γ Cygni SNR) and a description of particle 
transport that takes into account the detection of multiple plasma 
modes in the region (Zhang et al. 2020). As a consequence CR 
diffusion is predicted to be inhomogenous, which results in con-
finement for a long time in the central cavities where plasma 
modes are predominantly magnetosonic, and a more rapid dif-
dusion in the nearby Alfvénic-dominated regions. However, 
their model does not take into account ionised gas in the central 
cavities, which seems required to explain the emissions observed 
from CoCent. Furthermore, they calculate a diffusion coefficient 
for physical parameters, such as gas density and temperature, that 
are not necessarily relevant for the entire gamma-ray emitting 
region. Overall, it is not clear if their model can explain the large 
extension of CoExt.

4.5. FCES G85.00−1.78

The introduction of FCES G85.00−1.78 (OffExc), significantly 
Improves the likelihood of the model (Sect. 3.4). However, the 
best-fit Gaussian model is only partially contained in our anal-
ysis region, and therefore the results may be subject to large 
uncertainties. A proper characterisation of this excess is left for 
future studies, and in this section we only provide general con-
siderations on possibilities concerning its origin. There is no 
obvious correlation between OffExc and structures in the gas 
maps (Figs. 2 and 3).

No SNRs are found overlapping with OffExc in SNRCat5 
(Ferrand & Safi-Harb 2012). On the other hand, the ATNF 
1.67 pulsar catalogue6 (Manchester et al. 2005) lists three pul-
sars within the $r_{98}$ area of OffExc with a spin-down power 
$>10^{34}$ erg s$^{-1}$. Among those PSR J2111+4606, the closest to 
the source centroid at an offset of 3.3$, has a spin-down power of 
$1.4 \times 10^{36}$ erg s$^{-1}$, a characteristic age of 17.5 kyr, and an uncer-
tain distance. It may be reminiscent of some middle-aged pulsars 
powering large gamma-ray sources such as HESS J1825−137 
(H.E.S.S. Collaboration 2019; Principe et al. 2020) or the pulsar 
halos around PSR J0633+1746 or PSR B0656+14 (Abeysekara 
et al. 2017). In the latter case, the offset of the pulsar from the 
centre of the emission can be explained by a combination of 
proper motion and time-dependent injection (Zhang et al. 2021).

OffExc is bordered on the east by X-ray emission in the 
Cygynus SB (Cash et al. 1980). The corresponding X-ray struc-
ture, dubbed as S-ARC 3 in Uyaniker et al. (2001) has been 
associated to the stellar association Cyg OB4. However, the very 
existence of Cyg OB4 is questioned based on parallax distances 
de Zeeuw et al. 1999). Cantat-Gaudin et al. (2020) reports ten 
stellar clusters with more than 100 members at a probability 
>70% within the $r_{98}$ area of OffExc. Among those, seven are 
located at distances from the Earth <1.8 kpc, which would corre-
respond to a physical size <200 pc. Most of them are quite far 
away from the gamma-ray emission centroid, with the closest at 
2.8$ being NGC 7082 at 1.339 kpc from Earth and with an 
estimated age of 61 Myr, larger than typical stellar clusters with 
established detections in gamma rays (for instance Tibaldo et al. 2021).

The spectrum of OffExc has a particularly strong break with 
a steep slope at low energies and a hard spectrum above a few 
GeV that is strikingly similar to the one of CoExt. Establishing a 
physical connection between the two emission components is 
not obvious. OffExc may be produced by particles escaping the 
volume encompassed by CoExt, after an energy-dependent trans-
port process that depleted the particle spectrum at the low-energy 
end. This is reminiscent of the illumination of neighbouring

5 http://snrcat.physics.umanitoba.ca/ 
6 http://www.atnf.csiro.au/research/pulsar/psrcat/
gas clouds by CRs escaping from a nearby source (for instance Tang 2019). Alternatively, steep slopes at low energies has been suggested as a signature of particles reacceleration in SBs (for instance Tolksdorf et al. 2019), and there may be radial gradients in such a mechanism.

5. Summary and conclusions

We presented an analysis of the gamma-ray emission from the Cygnus region based on ~13 yr of Fermi-LAT data. The extraction of the emission from the so-called Cygnus cocoon was performed from a dedicated modelling of interstellar emission from the region. Compared to the analysis presented in Ackermann et al. (2011), we used almost seven times more data, produced with an improved reconstruction scheme corresponding to enhanced instrument performance. The data analysis is based on a much larger catalogue of gamma-ray sources and dedicated results for major sources in the fields. We also used improved gas tracer data and an iterative procedure to derive the dark neutral gas map.

As a result, the emission from the cocoon is now separated into two main components: first, a central component, FCES G80.00+0.50 (CoCent), traced by a model for the distribution of ionised gas within the borders of photo-dissociation regions, and having a power law spectrum with index 2.19 ± 0.03+0.02−0.01; second, an extended component, FCES G78.74+1.56 (CoExt), that can be modelled with a 2D Gaussian intensity distribution of extension $r_{68} = 4.4^\circ \pm 0.1^\circ +0.1^\circ$ and a smooth broken power law spectrum with spectral indices $1.67 \pm 0.05+0.02−0.01$ and $2.12 \pm 0.02+0.01−0.01$ below and above $3.0 \pm 0.3_{-0.2}^{+0.2}$ GeV, respectively. Emission from this component is significantly detected out to nearly 10$^\circ$ from the approximate centre of the star-forming region. Its total spectrum is significantly different from that of CoCent, and it exhibits significant spectral variations in azimuth in the innermost $\leq 3^\circ$.

Two additional extended emission components were significantly detected during the analysis. Source FCES G78.83+3.57 (CoWest) overlaps with a bright arc of 8 μm emission on the border of the central cavities in Cygnus X, and has a spectrum statistically compatible with CoCent. Although the spectral similarity and spatial proximity suggests a common origin, CoWest does not show any obvious correlation with known gas structures. Another source, FCES G85.00−1.78 (OffExc), is offset by several degrees with respect to Cygnus X and its spectrum is significantly different from all the other extended components studied, so a common origin seems unlikely. The centroid of OffExc lies on the edge of our analysis region, and therefore its current characterisation may be inaccurate. A proper study of this component is left for follow-up work.

The extended set of observables resulting from our analysis for the two brightest sources making up the cocoon, CoCent and CoExt, can be accounted for reasonably well from a simple diffusion-loss framework with a small number of free parameters, under several model setups. In all viable scenarios, one single population of non-thermal particles with a flat injection spectrum at the central source is sufficient and both hadronic and leptonic options are viable. Particles span the full extent of source CoExt as a result of diffusion, and give rise to source CoCent by interacting with ionised gas in the innermost regions. Possible solutions are very different in terms of energetics, transport conditions, and time scales involved. Some scenarios involve continuous injection during ~0.3–3 Myr, transport in a medium with moderately to strongly suppressed diffusion with respect to the large-scale interstellar average, and injection luminosities in the $10^{36}$–$10^{37}$ erg s$^{-1}$ range. They could describe a process by which the observed gamma-ray emission is powered by particle acceleration in the prominent star clusters Cyg OB2 and NGC 6910. Alternatively, a hadronic solution exists involving a more recent event, with injection lasting 3–30 kyr and transport proceeding over 10 kyr in a medium where diffusion is not or only moderately suppressed, and a much more powerful source with injection luminosity ~$10^{38}$ erg s$^{-1}$. Such a scenario seems more relevant to a single supernova explosion.

Possible improvements beyond this simple interpretation framework include accounting for multiple and extended sources, a more advanced description of particle transport in an inhomogeneous medium and including advection, and, last but not least, the description of physically motivated acceleration mechanisms. The observables extracted from our analysis are made available in machine-readable format and can be used in the future to perform detailed comparisons with more sophisticated models.

From the observational perspective, significant advances in gamma rays can be expected from instruments with improved sensitivity and angular resolution. The upcoming Cherenkov Telescope Array (Cherenkov Telescope Array Consortium 2019) above a few tens of GeV will provide a sensitivity an order of magnitude better than previous ground-based instruments and an angular resolution reaching a few arcmin. Proposed space missions dedicated to the MeV to GeV domain (de Angelis et al. 2018; McEnery & Agamo Team 2020) may also improve a few times the angular resolution and by one or two orders of magnitude the sensitivity compared to Fermi, and provide observations in an energy range, the sub-MeV and MeV domain, poorly observed until now. Complementary advances in the characterisation of interstellar gas (for instance Emig et al. 2022), and of multi-wavelength and multi-messenger emission from the cocoon (for instance Mizuno et al. 2015; Yoast-Hull et al. 2017) are also key to improving our understanding of particle acceleration and transport in this region.

Acknowledgements. The Fermi-LAT Collaboration acknowledges generous ongoing support from a number of agencies and institutes that have supported both the development and the operation of the LAT as well as scientific data analysis. These include the National Aeronautics and Space Administration and the Department of Energy in the United States, the Commissariat à l’Énergie Atomique and the Centre National de la Recherche Scientifique/Institut National de Physique Nucléaire et de Physique des Particules in France, the Agenzia Spaziale Italiana and the Istituto Nazionale di Fisica Nucleare in Italy, the Ministry of Education, Culture, Sports, Science and Technology (MEXT), High Energy Accelerator Research Organization (KEK) and Japan Aerospace Exploration Agency (JAXA) in Japan, and the K. A. Wallenberg Foundation, the Swedish Research Council and the Swedish National Space Board in Sweden. Additional support for science analysis during the operations phase is gratefully acknowledged from the Istituto Nazionale di Astrofisica in Italy and the Centre National d’Études Spatiales in France. This work performed in part under DOE Contract DE-AC02-76SF00515. This work was supported by the “Agence Nationale de la Recherche” through grant ANR-19-CE31-0014 (GAMALO project, PI: P. Martin). This work makes use of NuPy (Harris et al. 2020), AstroPy (Astropy Collaboration 2022), Matplotlib (Hunter 2007), SciPy (Virtanen et al. 2020), and the colourmaps in the CMasher package (van der Velden 2020). The authors would like to thank E. Orlando and M. Pesce-Rollins for their helpful comments on the manuscript, as well as I. A. Grenier for insightful conversations about the project.

Note added in proof. Our calculation of stellar cluster properties in Sect. 4.1 considered stars reaching 60 $M_\odot$, which corresponds to type O3. It was pointed out to us by G. Morlino that the most massive star observed in NGC 6910 is of type O9.6V (Kaur et al. 2020), thus with a mass <20 $M_\odot$, and that, based on the relationship between cluster mass and stellar mass upper limit from
Weidner & Kroupa (2004), the most massive stars in that cluster are expected to reach $\sim 25 M_\odot$. Such an upper limit on the stellar mass would yield a reduction of more than an order of magnitude in the cluster mechanical power. In this case, NGC 6910 could not be the only source of accelerated particles in scenario H2, as it would require $\gtrsim 100\%$ efficiency in the conversion of mechanical energy into particle energy, but remains a viable candidate source in scenario H1, with a $\sim 10\%$ conversion efficiency. For the more massive OB2 association the impact of the stellar mass upper limit from Weidner & Kroupa (2004) on the estimated cluster luminosity is only of a factor of a few and does not alter significantly our conclusions.

References

Aharonian, F., Yang, R., & de Oña Wilhelmi, E. 2019, Nat. Astron., 3, 561
Berland, K., & Anderson, D. 2002, Model Selection and Multimodel Inference: a Practical Information-theoretic Approach (Berlin: Springer Verlag)
Cohen, O., & Tatischeff, V. 2014, J. Open Source Softw., 5, 2004
Appendix A: Data analysis

We provide in this appendix a series of technical details about the data analysis.

Appendix A.1: Preliminary model optimisation

The preliminary optimisation of the emission model went through the following steps:

1. a simultaneous fit of the normalisation of bright sources with \( TS > 10^4 \) and predicted photons counts > 500;
2. an iterative fit of the normalisation and spectral shape of all the sources in the ROI by order of intensity and significance (method optimize of Fermipy);
3. a further simultaneous fit of the normalisation of sources with \( TS > 10^4 \) and predicted photons counts > 500, and also of the spectral shape of sources with the same \( TS \) condition and predicted photons counts > 1000.

The thresholds on predicted photon counts and \( TS \) were chosen to optimise all parameters for the brightest sources in the ROI (three pulsars, Cygnus Loop, \( \gamma \) Cygnii). We are more restrictive for the spectral shape to reduce the number of free parameters in the analysis, which otherwise can become unstable.

Appendix A.2: Morphological fits

In the morphological characterisation stages, the optimisation of the related components is performed in a two-step process.

1. Extension optimisation: we perform a first scan over a coarse grid of extension values, followed by a second finer scan around the first optimum and then the fit of a parabola to determine the final extension value and its uncertainty.
2. Position optimisation: we compute a map of log-likelihood values over a position grid of \( 2^\circ \times 2^\circ \) with a 0.2\(^\circ\) binning and determine the best-fit position and its uncertainty from the fit of an ellipse.

Appendix A.3: Spectral models

In the spectral characterisation of the various components of the cocoon, we consider the following models: a simple power law (PL) of expression

\[
\frac{dN}{dE} = N_0 \left( \frac{E}{E_0} \right)^{-\gamma},
\]

with \( N_0 \) flux at the reference energy \( E_0 \) and \( \gamma \) spectral index; a log-parabola (LP) of expression

\[
\frac{dN}{dE} = N_0 \left( \frac{E}{E_0} \right)^{-\alpha + \beta \ln \left( \frac{E}{E_0} \right)},
\]

with \( N_0 \) flux at the reference energy \( E_0 \), \( \alpha \) slope parameter, and \( \beta \) curvature parameter; and a smooth broken power law (SBPL) of expression

\[
\frac{dN}{dE} = N_0 \left( \frac{E}{E_0} \right)^{-\gamma_1} \left[ 1 + \left( \frac{E}{E_b} \right)^{\frac{\gamma_2 - \gamma_1}{\kappa}} \right]^{-\kappa},
\]

with \( N_0 \) flux at the reference energy \( E_0 \), \( E_b \) break energy, \( \gamma_1 \) spectral index at energies \( \ll E_b \), \( \gamma_2 \) spectral index at energies \( \gg E_b \), and \( \kappa \) the smoothing parameter fixed to 0.2.

Appendix B: Additional results on the spectro-morphological analysis

Figure B.1 shows the \( TS \) values for the best decomposition of CoExt, as in step D described in Section 3.7.2. The figure illustrates how the decomposition was designed to conserve a minimum \( TS \) of 25 except for the largest ring where the requirement could not be met.

We used the intensities derived in Section 3.7.2 to derive emissivities. To do so, we divided the flux associated with each segment, ring, or disk by the total neutral gas column density in the local arm (atomic, molecular and DNM) integrated over solid angle for each segment, ring, or disk area (shown in Figure B.2). See Section 4.1 for a discussion on uncertainties in the gas column densities relevant for the emissivity computation.

Figure B.3 shows maps of intensities and emissivities for CoExt as decomposed in step D in section and in three different energy bands 0.5 – 2.236 GeV, 2.236 – 10 GeV and 10 – 1000 GeV.
Appendix C: Diffusion-loss model framework

In Section 4.2, we introduce a simple diffusion-loss framework to account for the observed properties of the Cygnus cocoon. We provide here the full formalism of this framework.

The transport equation governing the evolution of the particle distribution $f$ in momentum $p$, position $r$, and time $t$ is:

$$\frac{\partial f}{\partial t} = \nabla \cdot (D \cdot \nabla f) - \frac{\partial}{\partial p} [\dot{\rho} \cdot f + Q]. \quad (C.1)$$

where $D$ is a spatial diffusion coefficient, $\dot{\rho}$ a momentum loss term, and $Q$ a source term. The momentum loss term $\dot{\rho}$ includes losses that are uniform in space and arise from radiative processes, hadronic interactions for accelerated protons, and Bremsstrahlung, inverse-Compton scattering, and synchrotron radiation for accelerated electrons (Schlickeiser 2002). The source term $Q(p, t)$ is assumed to be point-like in space and to have a constant power-law with exponential cut-off spectral shape:

$$Q(p, t) = Q_0 \left(\frac{p}{10 \text{ GeV}/c}\right)^{-\alpha} \cdot e^{-p/p_{\text{cut}}} \cdot e^{-t/t_{\text{cool}}}. \quad (C.2)$$

Particles are injected with a power law spectrum in momentum with index $\alpha$. The cut-off momentum is set to a high value of 1 PeV/c that our gamma-ray data are not sensitive to. To investigate the possibility that injection is not constant in time, and may have occurred over a finite duration some time ago followed by a longer diffusion time, we implemented (somewhat arbitrarily) an exponential decay of the source term. Integrating the source term $Q(p, t)$ over particle energies yields the time-dependent injection luminosity $L_{\text{inj}}(t)$. The diffusion coefficient $D(p)$ assumed to be constant in space and time is defined as:

$$D(p) = \beta \cdot D_0 \cdot \left(\frac{p}{10 \text{ GeV}/c}\right)^{\delta} \quad \text{with} \quad \delta = 1/3, \quad (C.3)$$

where $\beta = v/c$ with $v$ the velocity of a particle and $c$ the velocity of light. The power-law dependence in momentum, with an index corresponding to a Kolmogorov scaling for the magnetic turbulence spectrum, is adapted from models of large-scale CR propagation in our Galaxy (Trotta et al. 2011; Orlando 2018). We note that alternative expressions were considered recently in the light of more accurate direct CR measurements at Earth (Evoli et al. 2019; Génolini et al. 2019), and that the considered deviations from a pure power law may have consequences in the energy range we are interested in here.

The solution to the transport equation is obtained as (Atoyan et al. 1995):

$$f(p, r, t) = \int_{\max[0, t_{\text{cool}}(p, t) - t]}^{\infty} \frac{\dot{\rho}(p_0) \cdot Q(p_0, t_0)}{\pi^{3/2} r_d^3} \cdot e^{-r_d^2/2} \cdot dt_0, \quad (C.4)$$

with $r_d$ diffusion distance. For a present-day momentum $p$, the integration runs either over the full injection and transport history, or over the recent period spanning a cooling time $t_{\text{cool}}$ from the cut-off momentum down to $p$, with cooling time $t_{\text{cool}}(p_0, p) = \int_p^{p_0} \frac{dp}{\dot{\rho}}. \quad (C.5)$

This three-dimensional spherically symmetric distribution of particles is integrated along the line of sight for any angular offset from the centre $\theta$ and given the distance $d$ to the source:

$$\Phi(p, \theta, t) = 2 \cdot d^2 \cdot \int_0^{\infty} f(p, \sqrt{\theta^2 d^2 + t^2}, t) \, dt. \quad (C.6)$$

The resulting angular distribution of particles is then used to compute non-thermal emissions at any point in the region of interest. To this aim, we used the naima package (Zabalza 2015) in the approximation of isotropic radiation fields in the case of inverse-Compton scattering and using a nuclear enhancement factor of 1.845 in the case of pion decay (Mori 2009).

Appendix D: Possible diffusion-loss scenarios

We display here the results obtained for some of the diffusion scenarios considered in the interpretation of the results (see Section 4.3). The intensity and emissivity profiles for hadronic scenarios H2 and H3 or H4 are presented in Figs. D.1 and D.2, and the intensity profiles for model leptonic scenario L2 are presented in Figure D.3. The full set of results for the leptonic pulsar halo scenario is presented in Figs. D.4 and D.5.
Fig. D.1: Same as Figure 16, for model setup H2.
Fig. D.2: Same as Figure 16, for model setup H3 or H4.
Fig. D.3: Same as Figure 18 for model setup L2.

Fig. D.4: Same as Figure 17 for the pulsar halo scenario.
X. Astiasarain et al.: Multiple emission components in the Cygnus cocoon detected from Fermi-LAT observations

Fig. D.5: Same as Figure 18 for the pulsar halo scenario.