A&A, 693, A38 (2025)
https://doi.org/10.1051/0004-6361/202451108
© The Authors 2024

Astronomy & Astrophysics

How do supernova remnants cool?

II. Machine learning analysis of supernova remnant simulations

P. Smirnova[1,⋆], E. I. Makarenko[1,⋆], S. D. Clarke[2], E. Glukhov[3], S. Walch[1,4], I. Vaezzadeh[1], and D. Seifried[1]

[1] I. Physikalisches Institut, Universität zu Köln, Zülpicher Str. 77, 50937 Köln, Germany
[2] Institute of Astronomy and Astrophysics, Academia Sinica, No. 1, Sec. 4, Roosevelt Rd., Taipei 10617, Taiwan
[3] Stony Brook University, 100 Nicolls Rd, Stony Brook, NY 11794, USA
[4] Center for Data & Simulation science, Universität zu Köln, Albertus-Magnus-Platz, 50923 Köln, Germany

Received 13 June 2024 / Accepted 16 November 2024

ABSTRACT

Context. About 15%-60% of all supernova remnants are estimated to interact with dense molecular clouds. In these high-density environments, radiative losses are significant. The cooling radiation can be observed in forbidden lines at optical wavelengths.

Aims. We aim to determine whether supernovae at different positions within a molecular cloud (with or without magnetic fields) can be distinguished based on their optical emission (e.g. Hα (λ 6563), Hβ (λ 4861), [O III] (λ 5007), [S II] (λ 6717, 6731), and [N II] (λ 6583)) using machine learning (e.g. principle component analysis and k-means clustering).

Methods. We have conducted a statistical analysis of the optical line emission of simulated supernovae interacting with molecular clouds that formed from the multi-phase interstellar medium modelled in the SILCC-Zoom simulations with and without magnetic fields. This work is based on the post-processing of simulations that have been carried out with the 3D (magneto)hydrodynamic code FLASH. Our dataset consists of 22 simulations. The supernovae were placed at a distance of either 25 pc or 50 pc from the molecular cloud's centre of mass. First, we calculated optical synthetic emission maps (taking into account dust attenuation within the simulation sub-cube) with a post-processing code based on MAPPINGS V cooling tables. Second, we analysed the dataset of synthetic observations using principle component analysis to identify clusters with the k-means algorithm. In addition, we made use of BPT diagrams as a diagnostic of shock-dominated regions.

Results. We find that the presence or absence of magnetic fields has no statistically significant effect on the optical line emission. However, the ambient density distribution at the site of the supernova changes the entire evolution and morphology of the supernova remnant. Due to the different ambient densities in the 25 pc and 50 pc simulations, we are able to distinguish them in a statistically significant manner. Although, optical line attenuation within the supernova remnant can mimic this result depending on the attenuation model that is used. That is why, multi-dimensional analysis of optical emission line ratios in this work does not give extra information about the environmental conditions (ambient density and ambient magnetic field) of supernova remnant.

Key words. magnetic fields – magnetohydrodynamics (MHD) – methods: statistical – ISM: supernova remnants

1. Introduction

Typically, massive stars of 8–40 $M_\odot$ explode as a type II supernova (SN) at the end of their lifetime. Around 15–60% of these SNe are estimated to interact with a nearby molecular cloud (MC) (Hewitt & Yusef-Zadeh 2009; Zhou et al. 2023). Interaction with the dense environment of a MC shapes the evolution of the SN remnant (SNR). First, the SN explodes, producing a medium mostly filled with hot gas (McKee & Ostriker 1997; Kavanagh et al. 2013; Alsabti & Murdin 2017). The ejecta undergoes free expansion, in which the shocks produced by the explosion significantly affect the gas, which becomes hot, ionised, and turbulent. In the next step, it evolves to the Sedov-Taylor (adiabatic blast wave) stage and the ejecta energy is transferred to the ambient gas, but the net energy is conserved (Sedov 1959; Truelove & McKee 1999; Haid et al. 2016). At this stage (typically around 10^4 years after the explosion, but the exact number depends on the ambient density distribution), the young SNR can be observed in the X-ray and even γ-ray regimes (Borkowski et al. 2001; Aharonian et al. 2004; Vink 2012; Sasaki et al. 2012; Slane et al. 2014). The third stage starts when radiative losses become significant, at which point a shell-like structure appears. Due to the cooling, the remnant starts to emit at UV and optical wavelengths, while further expanding into (snow-ploughing through) the ambient medium (Fesen 1985; Mavromatakis et al. 2002; Boumis et al. 2008; Fesen et al. 2024). About 20% of the SNRs in our Galaxy have such optical counterparts (Green 2019). In the final stage, the remnant dissolves in the interstellar medium (ISM) (Ostriker & McKee 1988).

The evolution of a SN in a homogeneous medium has already been well studied (McKee & Ostriker 1997; Cioffi 1988; Haid et al. 2016; Jiménez et al. 2019), but in reality, SNe are immersed in a complex ISM – a highly inhomogeneous environment with a wide range of densities and temperatures. Therefore, simulations of MCs are essential to our comprehension of the structure of the ISM, and as a result to modelling realistic SNRs. Recently, great progress has been made in simulations of the complex multi-phase ISM and the interaction of SNe with the turbulent ISM (e.g. de Avillez & Breitschwerdt 2005; Gatto et al. 2015; Walch et al. 2015; Walch & Naab 2015; Zhang & Chevalier 2019; Haid et al. 2019; Ganguly et al. 2023) or

[⋆] Corresponding authors; smirnova@ph1.uni-koeln.de, makarenko@ph1.uni-koeln.de with MCs in particular (Iffrig & Hennebelle 2015; Seifried et al. 2017; Seifried et al. 2018). These simulations help to answer many physical questions; for example the effect of SNe on the star formation rate (Padoan & Nordlund 2011; Gatto et al. 2015), and the energy and momentum contribution of SNe to the ISM (Walch & Naab 2015).

Makarenko et al. (2023) studied optical emission (Hα (λ6563), Hβ (λ4861), [O III] (λ5007), [S II] (λ6717, 6731), [N II] (λ6583)) from SNRs, employing the post-processing module CESS (Cooling Emission in the optical band from Supernovae in (M)HD Simulations)[1] used for the FLASH code. We employed the updated collision data output from MAPPINGS V (Sutherland & Dopita 2017). Makarenko et al. (2023) show that it is crucial to consider both the attenuation effect due to the dust within the SNR bubble using a simple radiative transfer, and a realistic density distribution from the surrounding ISM, as this significantly affects the line emission, and hence the diagnostics based on different lines. Correspondingly, the more complex simulations we have, the more difficult it is to disentangle and study the impact of each parameter (e.g., temperature, density, magnetic field) on the simulation outcome using non-statistical methods.

With the rapid development of supercomputers (and simulations), and an increasing number of telescopes (and accumulated observations), large datasets are now common in astrophysics (e.g. Perryman et al. 1997; Abdurro'uf et al. 2022; Smart et al. 2021). Because of this, the search for relations between physical parameters can be hampered by the large amount of data, and indeed correlations can be found between more than two parameters. In this case, tools such as unsupervised machine learning can be extremely useful: one such tool that is suitable for searching for correlations between a large number of parameters is principal component analysis (PCA). Principal component analysis transforms a large set of variables to a smaller one that still contains most of the information from the larger set, losing a little accuracy for simplicity. Einasto et al. (2011) presented a method that uses PCA to investigate the strength of correlations between the properties of superclusters of galaxies (data from SDSS DR7) and search for the presence of distance-dependent selection effects in the supercluster catalogue. Principal component analysis can also be applied for dimension reduction or to visualise data (Bressan et al. 2021). Several other algorithms can also be applied, such as t-distributed stochastic neighbour embedding (t-SNE), as is used by Anders et al. (2018), to better distinguish chemical sub-populations in the solar vicinity, rather than looking at 2D abundance maps. Once the dimensions are reduced, it is important to robustly find clusters with unsupervised methods.

The k-means algorithm (MacQueen 1967; Hartigan & Wong 1979) searches for proximity in multi-dimensional space (in our case in the reduced dimensional space). For example, the algorithm was used in Rubin & Gal-Yam (2016) to divide light curves into classes with a fixed number of clusters. The Sloan Digital Sky Survey (SDSS) used the k-means algorithm to identify clusters of different stellar spectral classes as well as rare objects and outliers (Sánchez Almeida & Allende Prieto 2013).

Further, there have been many attempts recently to improve the classification of SNe in BPT (Baldwin, Phillips, & Terlevich) diagrams (Baldwin et al. 1981; Kauffmann et al. 2003; Kewley et al. 2019), since the classification of different astrophysical objects is not always unambiguous or does not reflect all physical features. The BPT diagrams allow for the main excitation source of an object (shocks or photo-ionisation) to be determined using optical line ratios. The idea of using multi-dimensional data classification for emission-line galaxies with the support vector machine algorithms has been tried both for observations (Stampoulis et al. 2019) and for theoretical models (Kopsacheili et al. 2020). There have been several attempts to improve the BPT diagrams using other optical line ratios (creating a multi-dimensional optical line ratio space), and modern machine learning techniques to better classify observed objects or to better constrain their physical conditions (e.g. Vogt et al. 2014; Ho 2019; Zhang et al. 2020; Ji & Yan 2020; Rhea et al. 2021). However, none of the new models constitute a universal tool that could be used in both theoretical work and observations. Here we test whether BPT diagrams are a sensitive diagnostic tool to uncover the environmental conditions of young SNRs.

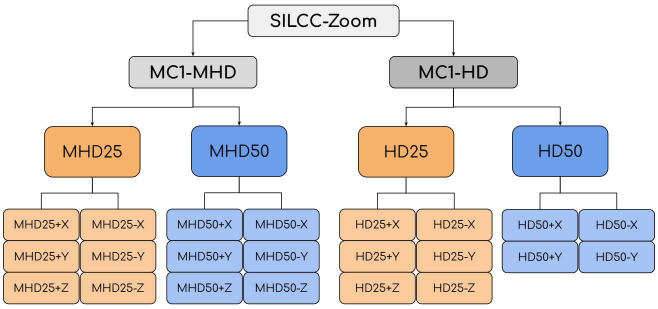

Fig. 1. Hierarchical structure of the dataset from the SILCC-Zoom project. We have in total 22 simulations: 'MC1-MHD' (with a magnetic field) and 'MC1-HD' (without a magnetic field). Each of the 'MD25', 'MHD50', 'HD25', 'HD50' datasets contains all possible positions of SN event (±X, ±Y, ±Z). Further details can be found in Seifried et al. (2018).

In this paper, we have performed a statistical study using unsupervised machine learning by post-processing (magneto-)hydrodynamic (MHD) simulations, comparing different initial conditions such as magnetic field strength and distances from the SN explosion to the centre of mass of the nearby MC that the remnant interacts with. We used multi-dimensional data consisting of multiple optical line ratios. The emission lines were calculated by post-processing the 3D simulations with the MAPPINGS V code. Our analysis reveals whether line ratios are sensitive to variations in the pre-shock density distribution and the presence of the magnetic field at the site of the SNR.

The paper is structured as follows. In Section 2, we describe the simulation setups, as well as the post-processing routine used to calculate optical emission lines. Section 3 describes the statistical methods (normalisation, pre-processing, and clustering) applied to our data and BPT diagrams. We present our results in Section 4 and discuss the importance of the ambient ISM properties and the attenuation effect for the optical emission of SNR. The conclusions are given in Section 5.

2. Dataset

We use SILCC-Zoom simulations of SNe interacting with MCs from Seifried et al. (2018) using the FLASH code (Fryxell et al. 2000) as an initial dataset. The dataset description can be found in Fig. 1 (for more details, see Walch et al. (2015); Girichidis et al. (2016); Seifried et al. (2017); Seifried et al. (2018)). In Section 2.1, we briefly summarise the simulation setup. In Section 2.2, we describe the post-processing tool from Makarenko et al. (2020, 2023) which was used to produce optical emission cubes and maps.

[1] https://github.com/kativmak/CESS

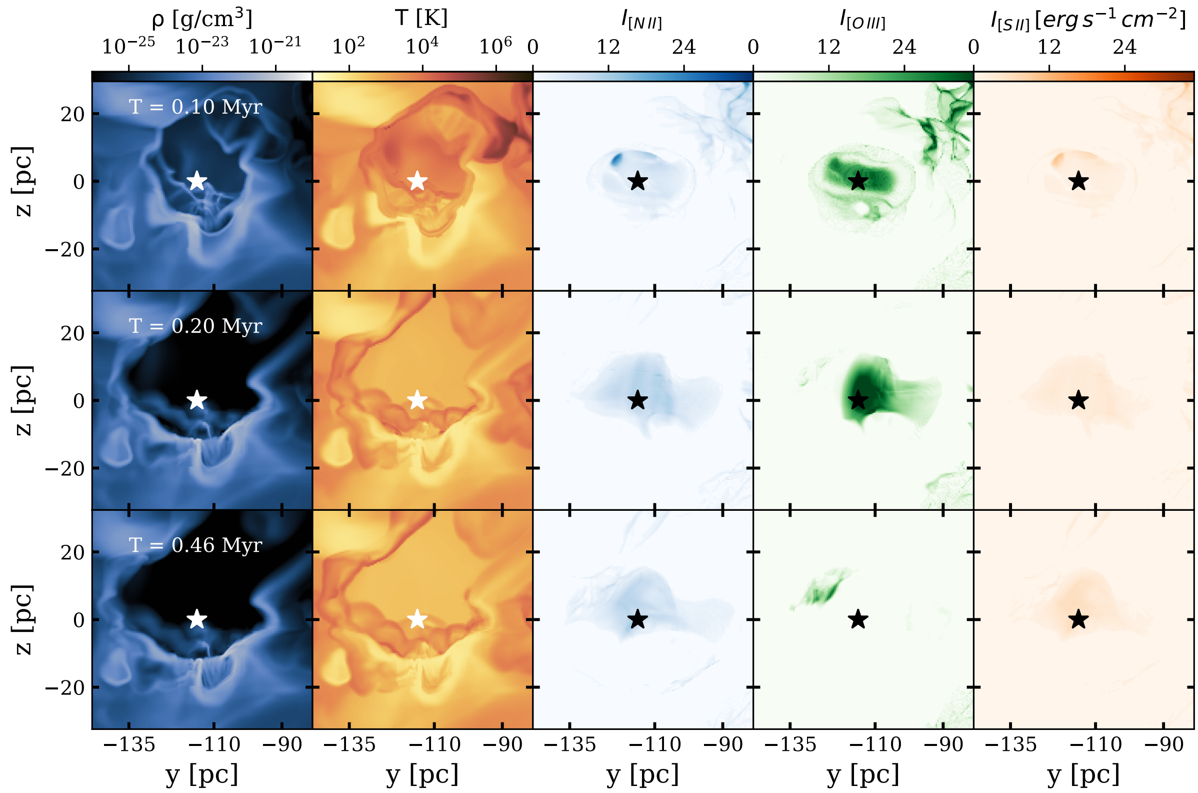

Fig. 2. Time evolution (from top to bottom) of (from left to right): density slice, temperature slice, [N II] (λ 6583 Å) intensity projection, [O III] (λ 5007 Å) intensity projection, and [S II] (λ 6731 Å) intensity projections for the 'MHD50+X' simulation. The star symbol in each panel represents the position of the SN explosion. The SN explosion disrupts part of the dense MC, as can be seen from the density slice (first column). This MC and SN interaction gives rise to the optical emission from the SNR. From the temperature evolution of the SNR bubble (second column), we can see the hot gas (10^{6-7} K) cooling down over 0.46 Myr, and forming the complex structures at the edge of the SNR and MC that are subsequently visible in the optical emission. Among three optical lines (third to fifth columns), the [O III] is the strongest as it is a 'volume-filling emission'. [N II] and [S II] originate from a thin cooling layer right behind the shock front.

2.1. Simulations

In the SILCC project (Walch et al. 2015; Girichidis et al. 2016), we modelled the evolution of a section of a galactic disc (500 pc × 500 pc × ± 5 kpc), with an initial gas surface density of $\Sigma_{\rm gas}$ = 10 M$_\odot$ pc^{-2} with initial conditions similar to the ones present in the solar neighbourhood. The simulations took into account SN feedback, magnetic fields (in MHD runs, $B_{0,x} \sim 4\,\mu$G is the initial field strength in the disc mid-plane) and self-gravity with a non-equilibrium treatment of the H_2 and CO chemistry of the gas. The SNe rate is in agreement with the Kennicutt–Schmidt relation for our surface density (Robert & Kennicutt 1998) (15 Myr^{-1}). The SNe explosions heat and stir the gas and a complex multi-phase ISM appears. A zoom-in strategy was used to resolve the formation of MCs to scales of ~0.12 pc in the SILCC-Zoom simulations (Seifried et al. 2017; Seifried et al. 2018). We selected two basic simulations: one hydrodynamical (HD) without magnetic fields and one MHD with magnetic fields: MC1-HD and MC1-MHD. The general properties of the MCs that we shall study further are given in Table 1.

After the formation of MCs (we denote this time as t_0), and some further evolution, a new SN was exploded at time $t_{\rm SN} = t_0$ + 1.53 Myr. In Seifried et al. (2018), a new simulation was run for each SN exploding at different distances (25 pc or 50 pc simulations are used here), and at positions with respect to the centre of mass of the cloud along the x, y and z directions in a zoom-in region (see Fig. 1). Each SN explosion was modelled by adding 10^{51} erg of thermal energy into the radius, $R_{\rm inj}$, around the explosion centre. We made sure to resolve the Sedov-Taylor radius with at least 4 simulation cells, corresponding to 0.48 pc (Gatto et al. 2015). We followed each SNR evolution for 0.46 Myr. It is interesting to consider different SN positions because of the variation in density distribution at the site of the SN explosion ($10^{-27.5}$–$10^{-21.5}$ g cm^{-3} at 25 pc, and 10^{-28}–10^{-21} g cm^{-3} at 50 pc). As a result, the optical emission, which arises from interactions with a dense medium, will also look different. The aim of this paper is to explore how this difference is manifested.

Examples of the time evolution of temperature and density are shown in Fig. 2, from top to bottom in columns 1 and 2, respectively. As is seen in the upper right corner of the temperature evolution, the SN blows out the gas from the cavity towards

Table 1. Main parameters of MCs.

Simulation	Total mass (M$_\odot$)	Volume (pc^3)	H_2 fraction (M$_\odot$)
MC1-HD	7.3×10^4	88 × 78 × 71	2.1×10^4
MC1-MHD	7.8×10^4	88 × 78 × 71	1.3×10^4

Notes. Name of the simulation dataset (first column), total mass (first column), considered sub-volume volume (second column), and molecular hydrogen fraction within the selected sub-volume (third column).

Table 2. Median percentage of attenuation for each line with the error in each dataset.

Dataset	Median % of the line attenuation				
	[S II]	[N II]	[O III]	Hα	Hβ
MHD25	18 ± 10	25 ± 9	39 ± 10	32 ± 10	19 ± 11
HD25	34 ± 5	39 ± 5	40 ± 7	40 ± 5	34 ± 5
MHD50	34 ± 14	38 ± 12	43 ± 12	39 ± 10	35 ± 14
HD50	37 ± 9	36 ± 10	34 ± 14	33 ± 10	28 ± 13

Notes. Mean percentage of attenuation (Cols. 2–6), the name of the dataset (Col. 1). The attenuation of each optical line is taken into account, as is described in Section 2.2.

the region of lower density. During this process, the gas cools down and starts to emit in the optical ([N II], [O III], and [S II] lines; third, fourth, and fifth columns of Fig. 2 correspondingly). At the end of the evolution (t = 0.46 Myr), the optical emission fades away, as the majority of the SNR bubble is cooled below typical optical emission temperatures (less than a few 10^4 K). Because the different parts of the bubble evolve on different timescales (due to the complex shock-cloud interaction), we can still see optical emission at t = 0.46 Myr in the central region and in the upper right part of the SNR bubble.

2.2. Optical emission post-processing

To prepare our simulations and reduce the computational cost of post-processing, we cut out sub-cubes using the biggest SNR bubble radius as a border (which can be measured at t = 0.46 Myr). Each sub-cube is 64.8 pc on a side. In this case, we took into account the emission mainly from the SNR and avoided significant contamination from the background (previous SN events) in our analysis.

We followed the method of Makarenko et al. (2023), which introduces the post-processing module CESS for the FLASH code, which uses the collision data from MAPPINGS V (Sutherland & Dopita 2017; Sutherland et al. 2018) to reproduce optical emission maps of simulated SNRs.

In summary, the post-processing procedure is as follows. First, from the .hdf5 simulation output, we cut out a sub-cube of the region containing the SNR (64.8 pc on a side), and re-gridded the AMR grid to a uniform grid (every cell is 0.12 pc^3). Then, we calculated the emitted luminosity for every cell in the 3D computational domain using the temperature-cooling rate dependency, known from MAPPINGS V. After that, we integrated along a given line of sight, taking into account the attenuation (projecting the 3D cube to a 2D map):

$$F_{\rm tot} = \int F_i e^{-\tau_{\rm i}}\, ds \tag{1}$$

where $F_{\rm i}$ is the flux of the cell i, $\tau_{\rm i}$ is the optical depth, and ds is the area of the cube. The optical depth, $\tau_{\rm i}$, was calculated in the following manner:

$$\tau_i = \kappa_{\rm abs}\rho_i V_i^{1/3} f_{\rm d}, \tag{2}$$

where $\kappa_{\rm abs}$ is the dust absorption cross section per mass of dust (cm^{2}g^{-1}), $\rho_{\rm i}$ is the density of cell i, V_i is the cell volume, and $f_{\rm d}$ is the dust-to-gas ratio ($f_{\rm d}$ = 0.01 is fixed in our simulations). The dust absorption cross-section was taken from Weingartner & Draine (2001) (Milky Way dust with R$_{\rm V}$ = 4.0). The attenuation percentage for each line in our dataset can be found in Table 2. The attenuation effect can be high for the optical emission lines due to gas and dust attenuation. Here we took into account only attenuation within the SNR simulation domain. This attenuation effect leads to fewer SNRs detected in the optical regime in our Galaxy compared to radio and X-ray (Green 2019). Thus, we reproduced optical emission line maps (Hα (λ 6563 Å), Hβ (λ 4861 Å), [O III] (λ 5007 Å), [S II] (λ 6717 Å, 6731 Å), and [N II] (λ 6583 Å) that show the same features as in real observations of SNRs interacting with the dense medium. We investigated the following optical line ratios: [O III] (λ 5007)/Hα, [S II] (λ 6731)/Hα, [N II] (λ 6583)/Hα, and [O III] (λ 5007)/Hβ. These line ratios are typically used in BPT diagrams as shown in Section 3.4. As a result, we have a 4D dataset. We calculated line ratios at each timestep (0.02 Myr) during the SNR evolution time (0.46 Myr) along the x axis from one side of the cube in 22 simulations (see Fig. 1).

3. Statistical analysis

3.1. Data normalisation

In order to make a comparison between data points, the dataset must be normalised according to the chosen transformation function. We would like the normalised variables to be in units relative to the standard deviation of the sample. To get the normalised variables, we performed the following normalisation for each data point:

$$z = \frac{d-\mu}{\sigma},$$

where d is initial (raw) data, μ is the mean of the sample, and σ is the standard deviation. This is standard practice for PCA to remove the effects of different means and scales between the features. As an input for this procedure, we took optical line ratios in a linear space.

3.2. Dimension reduction: Principal component analysis

We have four optical line ratios (see Section 2.2) and their time evolution over 0.46 Myr. This yields a 4D dataset. We first reduced our dataset to 2D, to aid in data interpretation, while retaining as much information as possible (for 3D, see Appendix C, Fig. C.1; the third dimension contributes less than 10% significance, and therefore using 2D does not lead to lower accuracy). Principal components (PCs) are directions in feature space along which the original data exhibits the greatest variance. By retaining the two PCs with the largest variances, we minimise the loss of information during dimensionality reduction. The larger the variance of the PC axis, the larger the dispersion of the data along it (the more information it carries), and therefore the less information is lost. This enables us to visualise the data in two dimensions.

The loading represents the weight (or contribution) of a specific original variable to the PC. The loadings for each line ratio are shown in Table 3 for each PC of the dataset. Despite PC1 and PC2 being represented by a mixture of features, we focussed on the ones with the highest variance. It is easy to see that for PC1 the line ratio [S II]/Hα is the most significant. S$^+$ is a known tracer of SNRs, as it is a collisionally excited ion that can form in a large recombination zone behind the SNR shock. For PC2, it is [O III]/Hα and [O III]/Hβ that are the most significant. O^{++} is forming within a significant region of the SNR bubble, as it has a higher excitation temperature than S$^+$ or N$^+$. Due to this, the observed area f the SNR bubble, where we can observe O^{++},

Table 3. Loading for principal axes in feature space.

Dataset	PC	Coefficients of the line ratio			
		[S II]/Hα	[N II]/Hα	[O III]/Hα	[O III]/Hβ
MC1$_{\rm att}$	PC1	–0.62	0.26	0.36	0.39
	PC2	0.12	–0.52	0.70	0.70

Notes. Loadings for PC1 or PC2 (Col. 2)represent the directions of maximum variance in the dataset for each line ratio (Cols. 3–6). The variance captured by PC1 and PC2 is spread across multiple features, indicating that those features contribute together to the variance in a similar manner.

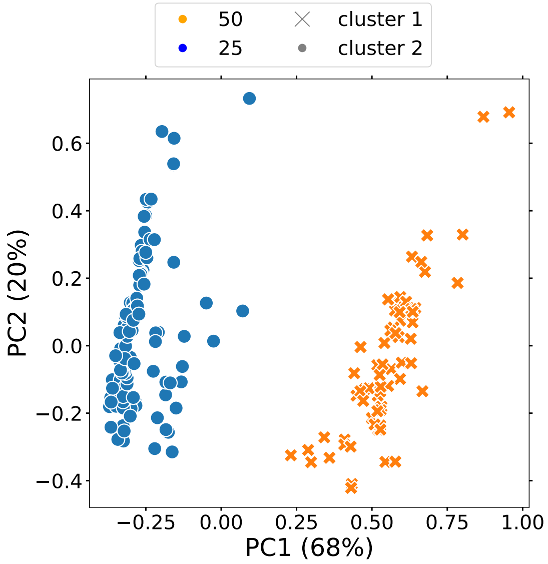

Fig. 3. 'MHD25', 'MHD50', 'HD25', 'HD50' datasets (e.g. 'MC1$_{\rm att}$') after using the PCA algorithm. 'MHD25'+'HD25' is shown in orange, and 'MHD50'+'HD50' is in blue. The predicted clusters are marked with circles and crosses. The higher the percentage for each PC, the higher the relative variance in the dataset that is observed in the direction of the corresponding eigenvector (for the absolute values, see Table 3). We used $n_{\rm clust}$ = 2 in the k-means algorithm. The data is clearly separated into two distinct clusters.

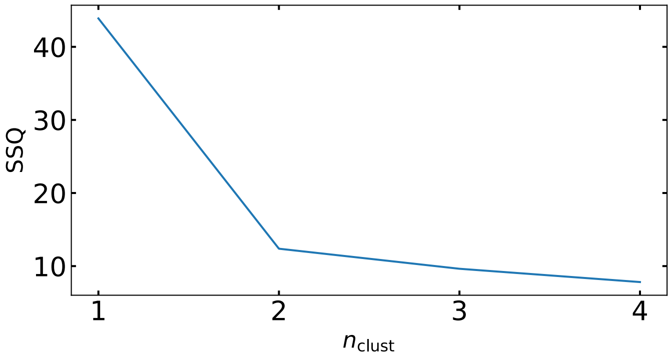

Fig. 4. Elbow plot for all dataset ('MC1$_{\rm att}$'). On the x-axis the number of clusters is presented, while on the y axis the total within the sum of squared distances (SSQ) is plotted. We see a kink in the line at $n_{\rm clust}$ = 2, which indicates that the optimal number of clusters to be used by the k-means algorithm is 2.

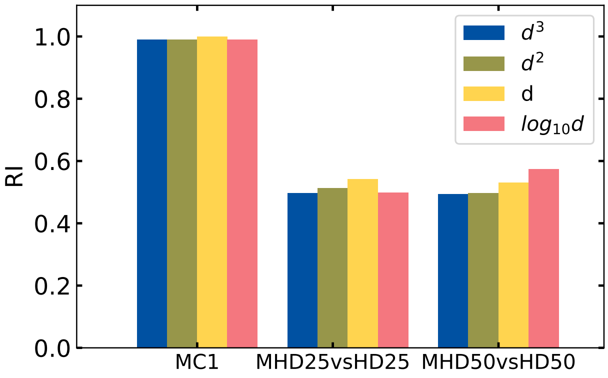

Fig. 5. Rand index value on the y-axis for different datasets on the x-axis: all data try to define two clusters ('MC1', first column), 'MHD25' vs. 'HD25' (second column), and 'MHD50' vs. 'HD50' (third column). The colours show the different data transformations: polynomial with blue and green, standard with yellow, and logarithm with red (see the legend for details). The higher the Rand index, the more effective the clustering. The best results were obtained with standard and polynomial data for the first column (all data, distance difference).

differs in various simulations, and can be used as a parameter to distinguish them. The result of the PCA algorithm is shown in Fig. 3 and Appendix A, Fig. A.1. We note that the data in the figures is already clustered (circles and crosses). We describe the clustering process in Section 3.3.

Apart from the PCA, we also tried to reduce the dimensionality of the dataset using the t-SNE algorithm (see Appendix B, Fig. B.1). This is a non-linear algorithm, in contrast to PCA. We found no qualitative difference between its results and those from PCA, so in the following analysis, we have only used the PCA algorithm.

3.3. Clustering: k-means and Rand index

After reducing the dimensionality of the dataset, we grouped the points in 2D space into clusters using the k-means algorithm. This algorithm seeks to minimise the sum of the squared Euclidean distances between each point and the centroid of the cluster to which it has been assigned. For the k-means algorithm, we need to provide the number of expected clusters as an input parameter. The elbow method (Thorndike 1953) is a technique to determine the number of clusters ($n_{\rm clust}$) to use in the k-means clustering algorithm. First, the elbow method calculates the sum of the squares of the distances between all points and then calculates the mean. When we choose $n_{\rm clust}$ = 1, the sum of the squared distances within the cluster is the largest. As the value of $n_{\rm clust}$ increases, the sum of squared distances (SSQ) within the cluster decreases. Finally, to determine the best choice of $n_{\rm clust}$, we plot $n_{\rm clust}$ versus the SSQ, as is shown in Fig. 4. At the point $n_{\rm clust}$ = 2, the SSQ decreases dramatically, or forms an 'elbow'. This point is considered the optimal value of the number of clusters. In Fig. 3 the colours represent the real dataset, and the symbols represent the identified clusters after performing the k-means algorithm (and vice versa in Fig. A.1).

To verify the clusters we identify with the k-means algorithm, we can use the Rand index. The Rand index indicates the similarity between the given (original) clusters and the predicted clusters. It is defined as RI = $\frac{a}{b}$, where a is the number of agreeing pairs (original clusters vs predicted clusters), and b is the total number of pairs. The closer the Rand index is to 1.0, the better the clustering. Fig. 5 shows the Rand index for different samples. For example, there were 198 data points in the 'MC1$_{\rm att}$' dataset, and for 'MHD25'+'HD25', 98 data points

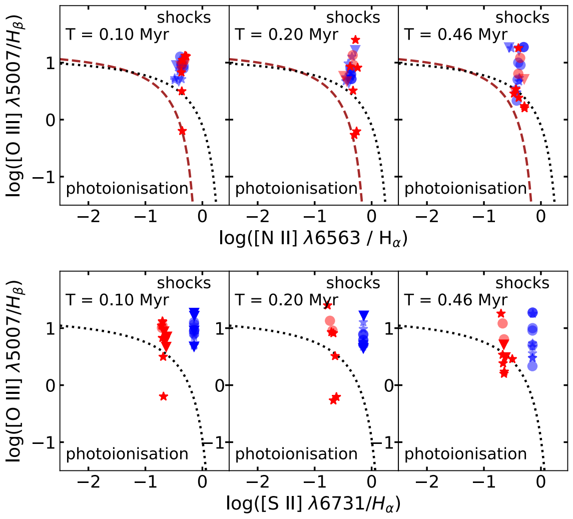

Fig. 6. Time evolution (from left to right) of the classical BPT diagram (upper row) and sulphur BPT diagram (lower row) for ‘MHD25’, ‘HD25’ (blue) and ‘MHD50’, ‘HD50’ (red). Star symbols show MHD simulations and circle symbols show HD simulations. The dotted reference line is Kewley et al. (2001), and the dashed line is from Kauffmann et al. (2003) for the upper row. The dotted reference line for the lower row is from Kewley et al. (2001). For both figures, reference lines separate the photoionisation region (star-forming, lower left corner) and the shock-dominated region (upper right corner). The SNRs in the upper figure are typically located in the “mixed region” (between the reference lines) or in the region of shocks as collisions are the main excitation mechanism for strong optical lines. For the lower figure, SNRs are typically located in the shocks region. The sulphur BPT diagram can classify SNRs according to the different ambient densities (‘MHD25’+‘HD25’ vs. ‘MHD50’+‘HD50’), as was shown with k-means.

were correctly linked to the clusters (initial and predicted point coincided), which leads us to RI = 98/198 ~ 0.5. Various transformations of the initial data (polynomial – d^2 or d^3, unscaled – d, and logarithmic – $\log_{10} d$) were also considered before proceeding with the PCA so as not to bias the result. Despite this, the best Rand index is obtained for the unscaled data, for the dataset ‘MC1$_{\rm att}$’. In this case, we can almost perfectly identify the ‘MHD25’+‘HD25’ and ‘MHD50’+‘HD50’ groups again after the statistical analysis if we compare the resulting clusters from the k-means algorithm with the initial dataset. We note that we do not use the information about labels for the initial dataset in our analysis. The Rand index for ‘MHD50+HD50’ and ‘MHD25+HD25’ is ~0.5, which means that we cannot distinguish different clusters for simulations with and without a magnetic field.

3.4. BPT diagram

The BPT diagram is a diagram based on strong optical line ratios (typically, [O III]/Hβ and [N II]/Hα). It helps to classify the dominant mechanism of ionisation in the observed object (e.g. an individual object like a SNR or in the whole galaxy). Two reference lines (Kewley et al. 2001; Kauffmann et al. 2003) (defined by theoretical modelling and SDSS catalogue analysis York et al. 2000) separate objects that are ionised by hot stars (star-forming emission or H II regions, lower left part) or by hard radiation of shocks (upper right part). Mostly, SNRs are located between the theoretical lines (in the so-called ‘mixed region’) or in the shock-dominated region. The BPT diagram for our dataset (‘MC1$_{\rm att}$’) is shown in the upper row in Fig. 6. Colours indicate different datasets: ‘MHD25’, ‘HD25’ – blue; ‘MHD50’, ‘HD50’ – red. The shape of the markers shows the presence or absence of a magnetic field: HD is represented by circles, and MHD be stars. The mean values for different simulations of SNRs start at $t = 0.1$ Myr at the border of the “mixed region”, or at the lower limit of the shocks region, then at $t = 0.2$ Myr and move up from the classification line as the gas cools down and shocks start to be observable in the optical band. Finally, at $t = 0.4$ Myr, shocks are mostly dissipated, optical line ratios have become weaker, and SNRs move back to the mixed region. This is a typical SNR evolution on the BPT diagram. We can conclude that this BPT diagram does not reveal different evolutionary paths for our dataset.

As was concluded in the PCA analysis above, the [S II](λ 6731)/Hα line ratio is the most important to detect SNRs and trace the various initial conditions (e.g. density distribution) at the SNR site. Therefore, we also plot a sulphur BPT diagram in the bottom row of Fig. 6. The colours and symbols are the same as in the classical BPT diagram and the separation line is taken from Kewley et al. (2001). During the evolution of the SNR, our mean value is clearly in the shock-dominated region, apart from a few points in the photoionisation region. The sulphur BPT diagram is sensitive enough to distinguish between the different ambient density distributions at the site of the SNRs. That is why at any moment of the evolution, our line ratios are divided into two groups: ‘MHD25’+‘HD25’ and ‘MHD50’+‘HD50’. The [S II](λ 6731)/Hα line ratio is also very similar for these two groups of simulations and almost does not change. Due to the known dependence of the [S II] doublet on the electron ambient density (Smith et al. 1993; Draine 2011), the similarity of the line ratios could be a consequence of the ambient media distribution at the SNR site. This is discussed further in Section 4.

4. Results

4.1. Different distances of supernovae to the molecular cloud

To test the influence of the initial density distribution on the optical emission of a SNR, two positions of the SN relative to the centre of the MC were considered. The first is located at 25 pc, and the second one at 50 pc. The PCA and t-SNE algorithms both show the statistical difference between these two cases, as we can see from the following k-means clustering and the Rand index in Fig. 5. The reason for this is the presence or absence of a denser medium near the site of the SN explosion. For strong optical emission in forbidden lines, a fairly dense environment that will be compressed by the shock wave is required, which will then radiate at low energies of the optical range in the late stages of the evolution of the SNR. This could cause the rise of optical emission in 25 pc simulations at a different time as each evolutionary stage of the SN strongly depends on the surrounding medium.

To explain why we see a statistical difference between the 25 pc and 50 pc simulations, we performed an analysis of the surrounding (ambient) ISM, as is detailed below. First, we defined the shock front (forward and reverse, but only forward shock detection was used) of the SNR using a shock finding routine based on Lehmann et al. (2016). This allows us to determine the shock cells based on the conditions of velocity divergence and the density gradient. Second, we cast six rays starting from the SN explosion position to locate the position of the primary shock with respect to the explosion centre. It was then possible to mask the SNR bubble, and calculate the median density and interquartile range only of the ambient (unshocked) medium. Taking into

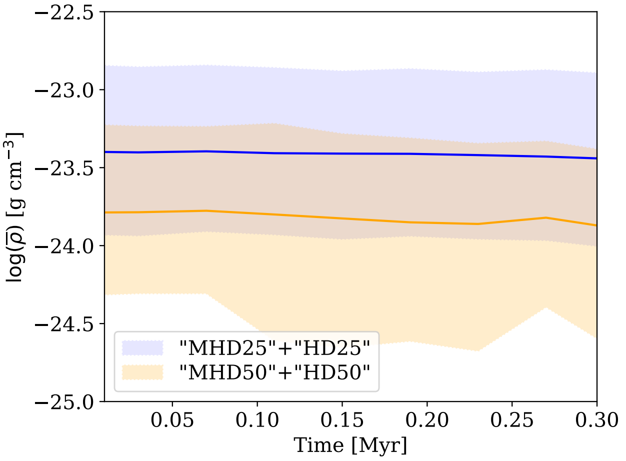

Fig. 7. Evolution of the mean density of the ambient medium for the 'MHD25'+'HD25' (blue) and 'MHD50'+'HD50' (orange) datasets (the mean of means for both datasets). We note that the mean was originally calculated in linear space over each set and presented in log space in the figure. The shaded region shows the interquartile range (25th and 75th percentiles of the data). The 'MHD25'+'HD25' dataset has higher values overall (by around 0.4 orders of magnitude) than the 'MHD50'+'HD50' dataset.

account the time evolution of the media surrounding our SNR (0.4 Myr), we can check how different the distribution of the medium is at 25 pc compared to 50 pc. The result of this procedure for the whole dataset is shown in Fig. 7. The mean ambient density for 'MHD25'+'HD25' (blue) is, on average, higher than for 'MHD50'+'HD50' (orange), as can be inferred from the evolution of the mean and interquartile range around 0.4 orders of magnitude. Although this difference does not appear to be significant, a difference is seen for optical radiation. Optical emission is brighter in older SNRs (where the shock wave velocities are usually less than 200 km s^{-1}). This brightness is determined by the density of the ambient medium encountered by the shock front.

If our SNe were placed at different positions in the MC they would encounter different ambient densities that would not necessarily depend on their radial distance from the MC centre. We would be able to differentiate them only in the case of different mean densities: for example, as in the cases of 25 pc and 50 pc in our simulations.

4.2. Magnetohydrodynamic versus hydrodynamic runs

The presence of a magnetic field at the explosion site does not directly affect the optical emission from the SNR. However, the magnetic field influences the morphological evolution of the SNR (as well as the shock waves). Moreover, due to compression, the magnetic field strength at the rim of the SNR bubble grows from 4 to 100 µG. Singly ionised particles (S$^+$, N$^+$) are usually formed in approximately the same temperature zone behind the shock wave. If there are any thermal instabilities combined with a magnetic field or other shock wave parameters favouring regions where hydrogen is ionised (even when N$^+$ and S$^+$ are able to recombine) the optical line ratios may change. Therefore, the difference in the optical line ratio depends on the initial conditions and magnetic field presence. This can be noted from the grid of shock models calculated using MAPPINGS V (Allen et al. 2008).

We have investigated if the magnetic field (initial condition B-field of 4 µG) influences the optical line ratios. For all simulations, we have two simulation runs – one with the magnetic field (MHD) and one without it (HD). Studying the optical lines we could not confirm any statistically significant deviation depending on the magnetic field. Thus, we show that the optical line ratios do not depend on the magnetic field during the SNR stage to a large extent, or the strength of the magnetic field should be significantly higher. The magnetic field is also expected to be much more important for high-energy bands (e.g. UV, X-ray, γ-ray).

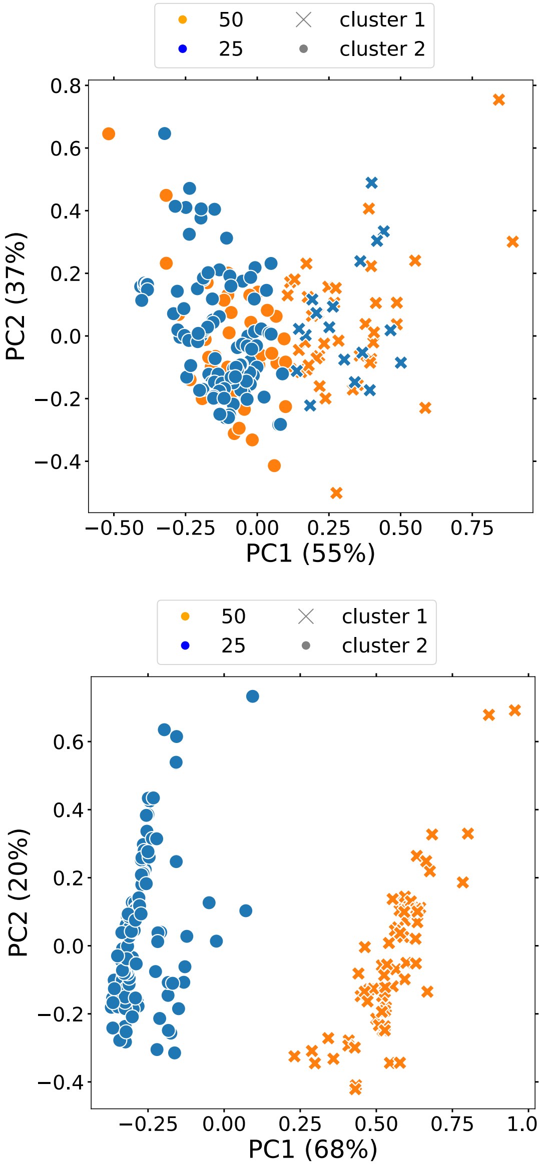

Fig. 8. Clustering of all data ('MC1') with k-means after PCA algorithm for line ratios without attenuation ('MC1$_{\rm unatt}$', top panel) and with attenuation ('MC1$_{\rm att}$', bottom panel). Colours represent real datasets; symbols represent predicted datasets. The best result shows the line ratios with attenuation (bottom panel). There, we can separate the different distances from the SN explosion to the centre of mass of the MC.

4.3. Attenuation effect

Makarenko et al. (2023) demonstrate that attenuation within the SNR cube should be taken into account, as it changes the classification of an object as an SNR based on the optical line ratios. In this work, we see a similar pattern. In Figs. 8 and 9, attenuation

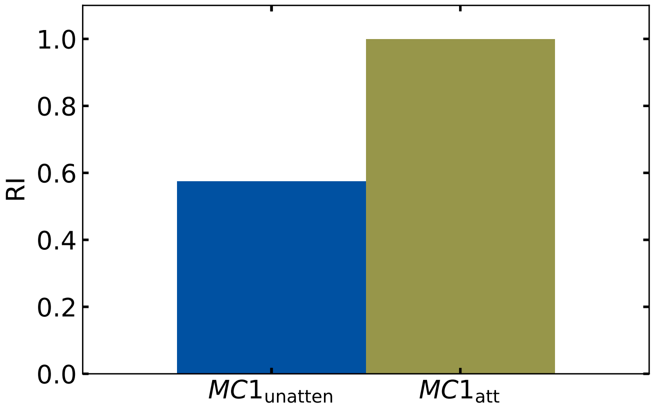

Fig. 9. Rand index value (y-axis) for unattenuated ('MC$1_{\rm unaten}$') and attenuated ('MC$1_{\rm att}$') line ratios from Fig. 8. The line ratio with attenuation (green) shows the best Rand index.

plays a crucial role in the search for the effect of a magnetic field or the distribution of matter at the site of an SNR on optical emission. To check that the [S II](λ6731)/Hα line ratio is a reliable tracer of the local properties at the SNR site, we need to be sure that attenuation is not mimicking some form of noise for the unattenuated optical lines brightnesses. We tested this by applying random noise to the unattenuated brightness values, at a level corresponding to the range of attenuation between its maximum and minimum values (approximately 80% and 20%, respectively; see Paper I, Makarenko et al. (2023), Figure 7), to examine the edge case. Therefore, we did not use the average values provided in Table 2. We repeated this analysis 1000 times. As a result, we can obtain point separations, for example, on a sulphur BPT diagram in Appendix D, Fig. D.1 using noise. This means that the [S II](λ6731)/Hα line ratio is not a universal determinant of the presence of shock waves from an SNR in the environment, and can be confused for other effects. Therefore, we advise caution when using this type of BPT diagram to come to conclusions about correlations between the line ratios and initial conditions at the site of the SNR.

5. Conclusions

In this paper, we have analysed optical emission from the SILCC-Zoom simulation dataset of SNRs interacting with MCs with and without magnetic fields. We have post-processed a dataset of 22 simulations using the CESS package.

To perform a statistical study, we further used the PCA algorithm to pre-process the data, and clustered it using the k-means algorithm. The Rand index allows us to compare how well the initial labels of the dataset fit with the clusters obtained without any prior information on the data. The Rand index has a value of 1 for the predicted different positions of the SNe with respect to MCs (dataset 'MC$1_{\rm att}$'), which means that our initial classification matches the clustered data. This means that in our simulations we can distinguish different distances to the centre of the MC (25 pc and 50 pc) due to the difference in the mean ambient medium density at the site of the SN explosion. As there is no universal trend of density with radius in MCs, it is not possible to link optical emission with distance from the MC centre. Therefore, the mean ambient density at the site of the SN is a relevant quantity in determining subsequent optical emission. The Rand index shows no statistically significant differences between the simulations with or without a magnetic field, so we cannot distinguish them. Analysing the nitrogen BPT diagram of the optical emission, all our SNRs are mainly located in the mixed region, and the variation between various simulations is minimal. For the sulphur BPT diagram, we see the same division of the dataset as in our clustering analysis, depending on the ambient density distribution at the SN explosion site. Due to that, we performed an independent analysis of unattenuated optical line luminosity adding random noise. We show that we can mimic the results of the sulphur BPT diagram with attenuation using unattenuated optical lines and some random noise. Multi-dimensional analysis of optical emission line ratios does not give extra information about the environmental conditions of the SNR. Therefore, we propose to not blindly trust the optical line diagnostic as a probe for the environment near the SNR and as a classification tool for the SNRs.

Nevertheless, we can conclude that realistic modelling of the ISM is an essential component of SNR modelling. The density distribution at the explosion site will affect not only the attenuation of optical emission, but also the rate of evolution of the SNR. This leads to different amounts of optical emission throughout the simulations, as was shown for a single SNR in Paper I (Makarenko et al. 2023). Finally, the use of statistical analysis of a large dataset is necessary at present. This allows for a less biased assessment of the importance of different parameters (density distributions and magnetic field) on the optical emission of SNRs.

Acknowledgements. We thank the anonymous referee for constructive comments which helped us to improve the manuscript. E.I.M. S.W. and D.S. acknowledge the support of the Deutsche Forschungsgemeinschaft (DFG) via the Collaborative Research Center SFB 1601 'Habitats of massive stars across cosmic time' (subprojects B1, B4 and B5). The following PYTHON packages were used: NUMPY (Harris et al. 2020), SCIPY (Virtanen et al. 2020), MATPLOTLIB (Hunter 2007), YT (Turk et al. 2011), SCIKIT-LEARN (Pedregosa et al. 2011), PANDAS (McKinney et al. 2010).

References

Abdurro'uf, Accetta, K., Aerts, C., et al. 2022, ApJS, 259, 35
Aharonian, F. A., Akhperjanian, A. G., Aye, K. M., et al. 2004, Nature, 432, 75
Allen, M. G., Groves, B. A., Dopita, M. A., Sutherland, R. S., & Kewley, L. J. 2008, ApJS, 178, 20
Alsabti, A. W., Murdin, P. 2017, Handbook of Supernovae (Springer)
Anders, F., Chiappini, C., Santiago, B. X., et al. 2018, A&A, 619, A125
Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, PASP, 93, 5
Borkowski, K. J., Lyerly, W. J., & Reynolds, S. P. 2001, ApJ, 548, 820
Boumis, P., Alikakos, J., Christopoulou, P. E., et al. 2008, A&A, 481, 705
Bressan, G., Barnaba, C., Peresan, A., & Rossi, G. 2021, Phys. Earth Planet. Interiors, 320, 106787
Cioffi D. F., M. C. F. B. E. 1988, ApJ, 334, 252
de Avillez, M. A., & Breitschwerdt, D. 2005, A&A, 436, 585
Draine, B. T. 2011, Physics of the Interstellar and Intergalactic Medium (Princeton Series in Astrophysics)
Einasto, M., Liivamägi, L. J., Saar, E., et al. 2011, A&A, 535, A36
Fesen R. A., Blair W. P., . K. R. P. 1985, ApJ, 292, 29
Fesen, R. A., Drechsler, M., Strottner, X., et al. 2024, arXiv e-prints [arXiv:2403.00317]
Fryxell, B., Olson, K., Ricker, P., et al. 2000, ApJS, 131, 273
Ganguly, S., Walch, S., Seifried, D., Clarke, S. D., & Weis, M. 2023, MNRAS, 525, 721
Gatto, A., Walch, S., Low, M. M. M., et al. 2015, MNRAS, 449, 1057
Girichidis, P., Walch, S., Naab, T., et al. 2016, MNRAS, 456, 3432
Green, D. A. 2019, J. Astrophys. Astron., 40, 36
Haid, S., Walch, S., Naab, T., et al. 2016, MNRAS, 460, 2962
Haid, S., Walch, S., Seifried, D., et al. 2019, MNRAS, 482, 4062
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Nature, 585, 357
Hartigan, J. A., & Wong, M. A. 1979, JSTOR: Appl. Statist., 28, 100
Hewitt, J. W., & Yusef-Zadeh, F. 2009, ApJ, 694, L16
Ho, I. T. 2019, MNRAS, 485, 3569
Hunter, J. D. 2007, Comput. Sci. Eng., 9, 90
Iffrig, O., & Hennebelle, P. 2015, A&A, 576, A95

Ji, X., & Yan, R. 2020, MNRAS, 499, 5749

Jiménez, S., Tenorio-Tagle, G., & Silich, S. 2019, MNRAS, 488, 978

Kauffmann, G., Heckman, T. M., Tremonti, C., et al. 2003, MNRAS, 346, 1055

Kavanagh, P. J., Sasaki, M., Points, S. D., et al. 2013, A&A, 549, A99

Kewley, L. J., Dopita, M. A., Sutherland, R. S., Heisler, C. A., & Trevena, J. 2001, ApJ, 556, 121

Kewley, L. J., Nicholls, D. C., Sutherland, R., et al. 2019, ApJ, 880, 16

Kopsacheili, M., Zezas, A., & Leonidaki, I. 2020, MNRAS, 491, 889

Lehmann, A., Federrath, C., & Wardle, M. 2016, MNRAS, 463, 1026

MacQueen, J. B. 1967, in Proceedings of the fifth Berkeley Symposium on Mathematical Statistics and Probability, 1, eds. L. M. L. Cam, & J. Neyman (University of California Press), 281

Makarenko, E. I., Walch, S., Clarke, S. D., & Seifried, D. 2020, in J. Phys. Conf. Ser., 1640, 012009

Makarenko, E. I., Walch, S., Clarke, S. D., et al. 2023, MNRAS, 523, 1421

Mavromatakis, F., Boumis, P., Papamastorakis, J., & Ventura, J. 2002, A&A, 388, 355

McKee, C. F., & Ostriker, J. P. 1997, ApJ, 218, 148

McKinney, W., et al. 2010, in Proceedings of the 9th Python in Science Conference, 445, Austin, TX, 51

Ostriker, J. P., & McKee, C. F. 1988, Rev. Mod. Phys., 60, 1

Padoan, P., & Nordlund, Å. 2011, ApJ, 730, 40

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, J. Mach. Learn. Res., 12, 2825

Perryman, M. A. C., Lindegren, L., Kovalevsky, J., et al. 1997, A&A, 323, L49

Rhea, C., Rousseau-Nepton, L., Prunet, S., et al. 2021, ApJ, 923, 169

Robert, C., & Kennicutt, J. 1998, ApJ, 498, 541

Rubin, A., & Gal-Yam, A. 2016, ApJ, 828, 111

Sánchez Almeida, J., & Allende Prieto, C. 2013, ApJ, 763, 50

Sasaki, M., Pietsch, W., Haberl, F., et al. 2012, A&A, 544, A144

Sedov, L. I. 1959, Similarity and Dimensional Methods in Mechanics (Academic Press)

Seifried, D., Walch, S., Girichidis, P., et al. 2017, MNRAS, 472, 4797

Seifried, D., Walch, S., Haid, S., Girichidis, P., & Naab, T. 2018, ApJ, 855, 81

Slane, P., Bykov, A., Ellison, D. C., Dubner, G., & Castro, D. 2014, Space Sci. Rev., 188, 187

Smart, R. L., Sarro, L. M., Rybizki, J., et al. 2021, A&A, 649, A6

Smith, R. C., Kirshner, R. P., Blair, W. P., Long, K. S., & Winkler, P. F. 1993, ApJ, 407, 564

Stampoulis, V., van Dyk, D. A., Kashyap, V. L., & Zezas, A. 2019, MNRAS, 485, 1085

Sutherland, R. S., & Dopita, M. A. 2017, ApJS, 229, 34

Sutherland, R., Dopita, M., Binette, L., & Groves, B. 2018, MAPPINGS V: Astrophysical plasma modeling code, Astrophysics Source Code Library [record ascl:1807.005]

Thorndike, R. 1953, Psychometrika, 18, 267

Truelove, J. K., & McKee, C. F. 1999, ApJS, 120, 299

Turk, M. J., Smith, B. D., Oishi, J. S., et al. 2011, ApJS, 192, 9

Vink, J. 2012, A&A Rev., 20, 49

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261

Vogt, F. P. A., Dopita, M. A., Kewley, L. J., et al. 2014, ApJ, 793, 127

Walch, S., & Naab, T. 2015, MNRAS, 451, 2757

Walch, S., Girichidis, P., Naab, T., et al. 2015, MNRAS, 454, 246

Weingartner, J. C., & Draine, B. T. 2001, ApJ, 548, 296

York, D. G., Adelman, J., Anderson, John E., J., et al. 2000, AJ, 120, 1579

Zhang, D., & Chevalier, R. A. 2019, MNRAS, 482, 1602

Zhang, X., Feng, Y., Chen, H., & Yuan, Q. 2020, ApJ, 905, 97

Zhou, X., Su, Y., Yang, J., et al. 2023, ApJS, 268, 61

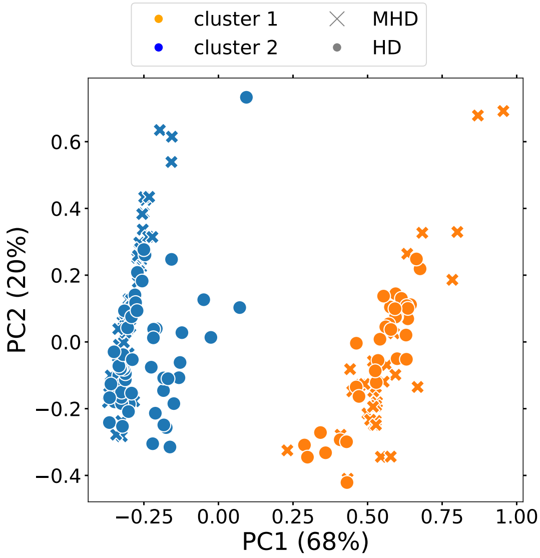

Fig. A.1: 'MHD25', 'MHD50', 'HD25', 'HD50' datasets (e.g. '$\mathrm{MC1_{att}}$') after using the PCA algorithm. 'MHD' simulations are marked with crosses, and 'HD' simulations with circles. The predicted clusters are marked with blue and orange. The higher the percentage for each PC the higher the relative variance in the dataset that is observed in the direction of the corresponding eigenvector (for the absolute values see Table 3). We use $n_{\mathrm{clust}} = 2$ in the k-means algorithm. The data is clearly separated into two distinct clusters by distance: 25 pc and 50 pc (see Fig. 3), while MHD and HD labels are mixed.

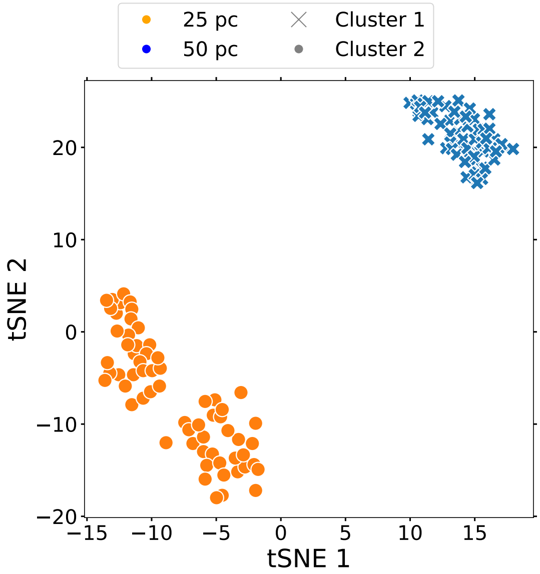

Fig. B.1: 'MHD25', 'MHD50', 'HD25', 'HD50' datasets (or '$\mathrm{MC1_{att}}$') after using the t-SNE algorithm. We use $n_{\mathrm{clust}} = 2$ in the k-means algorithm. Even though three clusters can be visually distinguished, this separation is not related to the presence or absence of a magnetic field. It is also not related to other physical parameters that we might associate with each cluster. That is why, the data is separated into two distinct clusters as for the PCA algorithm for this dataset.

Appendix A: PCA analysis for MHD vs HD

In Fig. A.1 we present the results of the PCA algorithm for the whole dataset ('$\mathrm{MC1_{att}}$') divided by k-means in two clusters ($\mathrm{n_{clust}} = 2$). We add labels to identify simulations with (MHD) and without magnetic field (HD). We note that the labels are not used in the PCA or k-means, we use them only to check the classification of the dataset. The data is clearly separated into two distinct clusters, but not by magnetic field presence (the colours do not match the shapes, as described in the figure caption). Therefore we can conclude that we can not distinguish between simulations with and without magnetic field. For the Rand index (see Fig. 5) it reaches a value of 0.5 which also means that the labels do not match the resulting clusters (for non-normalised Rand index).

Appendix B: t-SNE statistical analysis

Non-linear algorithms such as t-SNE for the dimension reduction of the dataset perform better than the linear ones (as PCA), especially for the preservation of local structures (e.g. clusters) of data. However, for our dataset, there was no difference between these algorithms. An example of t-SNE and k-means algorithm for the '$\mathrm{MC1_{att}}$' dataset is shown in Fig. B.1. We can define the same clusters as with the PCA algorithm in Fig. 3. Therefore, we use in this work PCA as it is computationally less expensive.

Appendix C: 3D representation of the statistical analysis (PCA)

In this work, we used a 2-D representation of the data after the PCA analysis (i.e. we took only the two most important components). Here we show the 3D distribution of the data (three PCs) in Fig. C.1. The third dimension contributes little to cluster classifications (only 10–13%). In addition, the third dimension is not helpful for better-separating points in space, therefore a 2D representation is optimal for the k-means clustering.

Appendix D: Sulphur BPT diagram with noise

To assess whether [S II](λ 6731)/Hα is reliable to identify different ambient density distributions at the SNR site, we plot the sulphur BPT diagram for the unattenuated optical emission lines with added random noise. Typically, [S II](λ 6731)/Hα is widely used to classify an object as an SNR as it is relatively strong and easy to observe. However, we found that attenuation plays an important role in this line ratio and can be confused with random noise as in the upper panel of Fig. D.1.

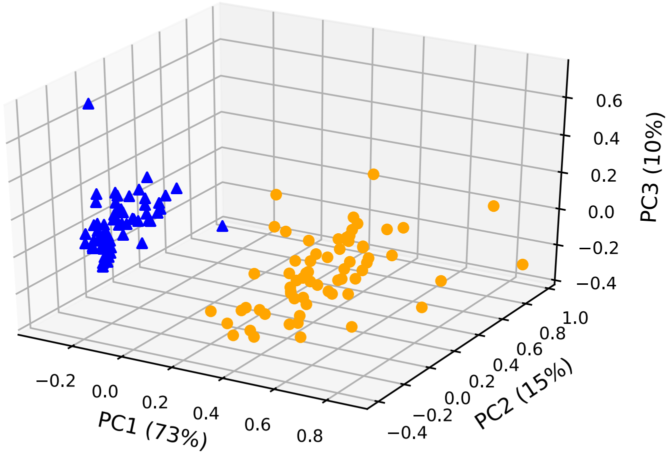

Fig. C.1: 3D representation of the data 'MHD25'+'HD25' (blue) and 'MHD50'+'HD50' (orange) after the PCA step. The predicted clusters are marked with circles and triangles. The contribution of each PC is as follows: 73% for PC1, 15% for PC2, and 10% for PC3. As in Fig. 3, we can still clearly see two clusters. However, visually, PC3 does not allow more efficient separation of clusters. Thus, in this work, 2D visualisation was mainly used (only the two PCs).

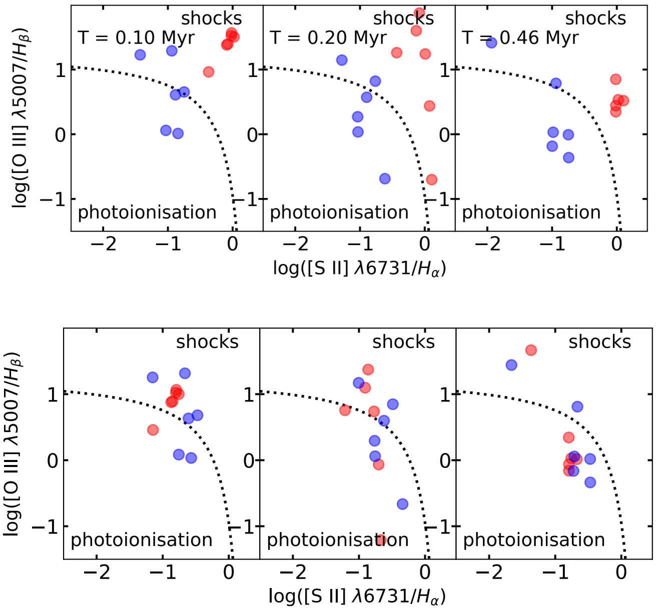

Fig. D.1: Sulphur BPT diagram for unattenuated line emission with added random noise. Colours are the same as in Fig. 6. On the lower panel, noise is higher than the real attenuation percentage (around 80% ± 10%, see Table 2). On the upper panel, noise is lower than the calculated attenuation (around 20% ± 10%). The upper panel's circles can be well separated during the whole time evolution (from left to right).