

The Cloud Factory II: Turbulent Line-widths of Resolved Molecular Clouds in a Galactic Potential

Andrés F. Izquierdo,^{1*} Rowan J. Smith,¹ Third Author^{2,3} and Fourth Author³

¹Jodrell Bank Centre for Astrophysics, School of Physics and Astronomy, University of Manchester, Oxford Road, Manchester M13 9PL, UK.

²Department, Institution, Street Address, City Postal Code, Country

³Another Department, Different Institution, Street Address, City Postal Code, Country

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

We present a statistical analysis of the turbulent line-widths of molecular cloud complexes extracted from high-resolution regions from our “Cloud Factory” galactic scale ISM simulation suite. The complexes are at similar evolutionary states but the turbulence is self-consistently generated by different physical mechanisms: a) The ISM dynamics are dominated purely by the large-scale potential of the galactic disc, differential rotation, and random supernovae feedback; b) The galactic potential and supernovae feedback are as above but gas self-gravity is turned on; c) The galactic potential and the self-gravity are turned on but now a burst of supernovae feedback tied to sites of star formation is triggered. To compare to observations, we perform radiative transfer simulations that predict the ^{12}CO $J=1-0$ line emission of the representative cloud complexes. Using the synthetic images we then apply the Principal Component Analysis (PCA) reduction technique and estimate a structure linewidth-size relation for each of the physical scenarios. The statistical analysis suggests that, even though purely gravitational effects are necessary to reproduce the standard observational laws and can recreate quiescent regions, they are not sufficient in most cases. We show that the extra injection of energy from clustered supernovae events plays a key role in establishing the global turbulent field and the local dynamics and morphology of molecular clouds in Milky Way-like galaxies. Once this is included, our Cloud Factory simulations generate molecular clouds matching observed scaling laws self-consistently, without the turbulence being put in by hand. We notice and characterise clustering of structure function parameters resulting from clouds in different physical ambiances, which include global diffuse (inner-arm) and dense (spiral-arm) environments, and embedded and outer-cloud supernovae feedback. For quiescent cloud complexes with low (isolated) stellar feedback, we present continuous time-evolving trajectories in the structure function parameter space, driven by gravitational collapse and supersonic turbulent flows at different scales. We observe several signs of intermittency that lead to variations in velocity structure functions depending on analysis scales, especially in our cloud complexes dominated by strong feedback from clustered supernova explosions. All these findings suggest that our PCA-based statistical study is a formidable method to diagnostic local and global surrounding conditions, evolutionary stages and physical mechanisms governing (real and synthetic) molecular clouds. We also present our new open source PCAFACORY package that collects all the tools developed to set up this work.

Key words: turbulence – molecular clouds – principal component analysis – radiative transfer

1 INTRODUCTION

The relative importance of the physical mechanisms involved in star formation has been subject to intense debate over the last decades. Undoubtedly, gravitational effects govern the concluding stages of individual star-forming systems (Krumholz & Tan 2007; Keto & Zhang 2010; Ballesteros-Paredes et al. 2011; Traficante

et al. 2018a,b), but additional factors may play a role on the larger scales where gas is assembled into molecular clouds and successive fragmentation takes place (Bergin & Tafalla 2007). Observational data from the latest generation of telescopes have confirmed that, far from being isolated systems, stars are formed within large-scale molecular cloud complexes (10–60 pc) that form in the cold interstellar medium (Mac Low & Klessen 2004). These cloud complexes consist of interconnected molecular clouds (2–20 pc) which, at the same time, exhibit high degrees of sub-structuring over subsequent

* E-mail: andres.izquierdo.c@gmail.com

scales (Falgarone et al. 1992) and filamentary signatures (Klessen et al. 2004; Smith et al. 2014b). Thus, studying the dynamics of molecular structures in different spatial regimes becomes highly relevant to uncover the nature and evolution of star formation properties.

Larson (1979, 1981) discovered a systematic increase of the global velocity dispersion Δv with the projected size L of diverse molecular associations ($\Delta v \propto L^{0.38}$) using mostly optically thin tracers (^{13}CO , H_2CO , NH_3). He interpreted this hierarchical behaviour to be a consequence of turbulent motions as it is reminiscent to the Kolmogorov's structure law ($\delta v \propto l^{1/3}$, where lower case δv and l indicate internal velocity and spatial scales), derived from the statistical framework developed by Kolmogorov (1941) and Onsager (1949) for viscous incompressible turbulent fluids. This linewidth-size relationship, the Larson's law, would then be the inspiration for a lot of succeeding literature on the role of turbulence in setting dynamical signatures of the ISM.

Solomon et al. (1987) focused on a more homogeneous sample of objects and claimed that turbulence in molecular clouds more closely resembles that of compressible fluids ($\delta v \propto l^{1/2}$), studied also by classical statistical-hydrodynamic theories (Kraichnan 1974; Frisch et al. 2001). This suggests that the energy diffusion can be driven by supersonic shocks at large scales rather than just by heat dissipation at small scales where viscosity dominates.

Turbulence is essential to trigger primordial density enhancements where star-forming regions emerge but also to regulate the onset of new stellar systems. Compressible (supersonic) turbulent velocity fields generate large-scale converging flows and strong density fluctuations, which, by the action of gravity, may end up collapsing and forming new stars in the most massive regions (Mac Low & Klessen 2004). At the same time, turbulence is a key mechanism for controlling star formation rates as it provides additional support that, along with magnetic forces, prevents runaway gravitational collapse (Falgarone et al. 1992; Federrath 2018). This gravoturbulent scenario of fragmentation in molecular clouds and its implications for star formation properties were detailed by Klessen et al. (2004).

Additionally, the interplay between supersonic turbulence and local gravitational forces produces particular gas density distributions. High column densities associated with massive regions, dominated by self-gravity, exhibit power-law probability density functions (PDFs) (Ballesteros-Paredes et al. 2011). Oppositely, low column densities dominated by turbulent supersonic motions yield log-normal PDFs (Vazquez-Semadeni 1994) that can deviate from a perfect profile due to intermittency effects (Federrath et al. 2010). The range of densities in cloud complexes is typically wide ($10^2 - 10^5 \text{ cm}^{-3}$, Mac Low & Klessen 2004), which implies that density distributions from realistic scenarios (see e.g. Schneider et al. 2002) are in general a combination of both profiles (Hennebelle & Chabrier 2008; Burkhardt 2018). Gravoturbulent mechanisms are hence crucial to establish stellar and core initial mass functions as they are closely related to the mass distribution of parental clouds (Padoan & Nordlund 2002).

Also, a great deal of effort has been guided to understand the origin of non-thermal motions in the cold ISM. Heyer & Peter Schloerb (1997) adapted the principal component analysis (PCA) reduction technique to investigate the turbulent behaviour of individual cloud complexes using spectroscopic data. The method consists in finding non-redundant representative components of (molecular) line emission data to extract velocity fluctuations δv associated to characteristic spatial scales l of the analysis region. They exploited the algorithm using synthetic and real objects and found power-law dependencies analogous to the Larson's linewidth-size

relationship. Several studies were then carried out to connect the scaling parameters retrieved from the method to intrinsic hydrodynamic structure functions (Brunt & Heyer 2002; Brunt et al. 2003; Heyer & Brunt 2004; Roman-Duval et al. 2011; Brunt & Heyer 2013). Other works tested the sensitivity of the technique to different energy contexts. Heyer et al. (2006) found different relationships for clouds inside ($\delta v = (1.00 \pm 0.04)l^{0.79 \pm 0.06}$) and outside ($\delta v = (0.70 \pm 0.03)l^{0.66 \pm 0.06}$) an ionization front driven by a cluster of massive stars in the Rosette cloud complex. Bertram et al. (2014) used numerical simulations of molecular clouds with imposed turbulent fields and noticed variations in PCA-derived exponents when changing mean densities and optical depths. Using ^{12}CO intensity, they found a steeper $\delta v \propto l^{0.82 \pm 0.03}$ for clouds with gas mean density $\bar{n} = 300 \text{ cm}^{-3}$, compared to the $\delta v \propto l^{0.59 \pm 0.02}$ for $\bar{n} = 100 \text{ cm}^{-3}$. Also, they found that using ^{13}CO , which is an optically thinner tracer, can lead to slightly different relationships ($\delta v \propto l^{0.74 \pm 0.02}$ for $\bar{n} = 300 \text{ cm}^{-3}$). These findings make the technique an interesting tool to investigate the nature of non-thermal motions in the ISM.

On the other hand, (M)HD simulations carried out by de Avillez & Breitschwerdt (2005) and Joung et al. (2009) included a global galactic context to consistently investigate the ISM evolution. They found that several observational properties of the ISM turbulence can be driven by supernovae feedback-dominated scenarios. However, due to the reachable spatial resolution ($\sim 1.5 \text{ pc}$) and the lack of local gravitational effects, they could not study the internal structure and dynamics of molecular clouds in detail.

Later, in order to uncover the structure of turbulent motions in molecular clouds, Federrath et al. (2010) simulated synthetic turbulent fields made up with two different forcing components, solenoidal and compressive, within periodic uniform grids assuming isothermal gas. They found that molecular clouds have generally different mixtures of forcing, in which the solenoidal component is associated with quiescent regions with low star formation activity, and the compressive component to regions dominated by sources of strong energy feedback. This is supported by observations of quiescent and active star-forming regions or a combination of both scenarios (Heyer et al. 2006; Hacar et al. 2016).

More sophisticated high-resolution simulations were then developed to try to explain the origin and nature of these turbulent motions. Klessen & Hennebelle (2010) provided analytic and numerical calculations including magnetic fields, self-gravity and a standard ISM cooling function to show that accretion processes can drive the observed turbulence on several scales, from galaxies to protostellar disks. They used converging flows of accretion, incoming from the computational boundary with superimposed mean velocities and fluctuations, and obtained a linewidth-size relation $\Delta v = 0.8L^{0.5}$ compatible with the Larson's law. This suggests that the turbulent scenario is similar to the classical energy cascade process from large to small scales, driven by outside-cloud phenomena. However, further work on energy injections from supernova explosions (Gatto et al. 2015; Walch et al. 2015; Girichidis et al. 2016; Pan et al. 2016), stellar outflows (Nakamura & Li 2007; Cunningham et al. 2011; Federrath et al. 2014), HII regions (Peters et al. 2017; Haid et al. 2018), would demonstrate that the role of stellar feedback is also essential in configuring the turbulent field of molecular clouds. In particular, ISM simulations presented by Padoan et al. (2016a,b, 2017) claimed that the structure and dynamics of molecular clouds are a natural consequence of a supernovae-driven scenario, and suggest that supernovae energy injection is necessary to set and maintain the turbulent cascade observed in molecular clouds. They generated random supernovae over a periodic cubic

box of 250 pc, with high (sub-parsec) spatial resolution, but at the cost of not considering the large scale gravitational potential nor differential rotation.

Our Cloud Factory simulations (Smith et al. 2019, submitted) seek to address this lack by including both supernovae feedback and large-scale galactic environment with enough resolution to study the internal turbulence within clouds. We take into account the global galactic context using a multi-component gravitational potential and galactic differential rotation, while, at the same time, we resolve selected molecular clouds in up to $0.25 M_{\odot}$. We include stellar feedback in form of supernovae, both randomly distributed across the galaxy and tied to sites of star formation, as well as local gravitational forces and molecular chemistry. In this work, we use full non-LTE radiative transfer calculations and the principal component analysis technique on our cloud complexes to investigate the detailed signatures of non-thermal motions in a wide range of spatial scales provided by our simulations. Full radiative transfer modelling is necessary to produce realistic synthetic observations that can be readily compared to observational data with analogous methods. Our analysis aims at investigating the role of clustered supernovae feedback and local and large-scale gravitational forces in configuring the velocity fluctuations field of the cold ISM. We conclude that our simulations are able to self-consistently generate cloud complexes, with realistic turbulent fields, that can be used in future for studies of clustered star formation in a galactic context.

We briefly present the main aspects of our Cloud Factory simulation suite and the selected cloud complexes in Section 2. Sections 3 and 4 are dedicated to the line emission theory and statistical description of turbulent fluids respectively. There, we also detail the radiative transfer and PCA setups used in the work. In Section 5 we outline the general workflow of the paper and explain the three PCA extraction methods implemented to analyse different spatial regimes. We then present in Section 6 our results based on the PCA study of ^{12}CO $J=1-0$ emission cubes extracted from our cloud complexes as a function of the physical scenario (6.1), line of sight projection (6.2), time evolution (6.3) and analysis scales (6.4). We wrap up with a short discussion on the resemblance of our self-consistently generated clouds to observational data in Section 7 and the conclusions Section 8. In Appendix A1 we show the results obtained if LTE approximations are assumed during the radiative transfer.

2 THE CLOUD FACTORY SIMULATION SUITE

2.1 The Hydrodynamic Code and Physical Ingredients

The cloud complexes that we examine across this work are extracted from our Cloud Factory simulation suite (Paper I, Smith et al. submitted), which is powered by a customised version of the AREPO code (Springel 2010; Pakmor et al. 2016). In order to account for a broad range of relevant physical processes taking place in the cold molecular ISM, this version uses the TREECOL algorithm (Clark et al. 2012) to compute ultraviolet extinction; the evolution of CO introduced by Nelson & Langer (1997) and the hydrogen chemistry of Glover & Mac Low (2007a,b); Glover et al. (2010); a sink particle-based hybrid model of star formation; a multipolar expansion to analytically model the galactic gravitational potential; and two different methods of injecting energy feedback from supernovae explosions. All these features are detailed in the following paragraphs.

AREPO solves the conservation laws of ideal hydrodynamics

on a moving or stationary unstructured Voronoi mesh by combining the strengths of Lagrangian methods such like the smoothed particle hydrodynamics (SPH) mesh-free technique (Gingold & Monaghan 1977; Lucy 1977) and Eulerian mesh-defined methods. Some of these strengths is that it i) allows to easily track thermodynamic quantities; ii) permits a continuous and automatic mesh (de-)refinement based on the grid cells distribution, meaning that a high range of spatial and mass resolutions can be achieved; (iii) accurately determines the exchange rates of physical conserved quantities between neighbouring cells via the finite-volume representation of the Euler equations; (iv) if required, consistently solves the Euler equations in presence of self-gravity and (v) avoid weaknesses like the low accuracy of SPH codes on boundary discontinuities and density jumps or the lack of Galilean invariance of Eulerian codes. For these reasons, AREPO is an excellent tool to deal with high dynamic range problems that in addition require several physical mechanisms taking place in a single domain.

We use an analytic description of the large-scale gravitational potential of the galaxy to cheaply determine and control its influence on the dynamics of mesh cells in each time step of the simulation. The potential is a combination of a dark-matter halo, a bulge, and a gas disc with thin and thick components. We use the best-fitting model of McMillan (2017), which is constrained to observations of the Milky Way. Additionally, we include a four-armed spiral component from Cox & Gómez (2002) and a consistent spiral perturbation to the potential, already implemented in Smith et al. (2014a). The density profiles spawning the large-scale gravitational potential are analytically described in our Paper I.

Our gas chemistry description adopts the approach of Nelson & Langer (1997), where the CO evolution is a simplified treatment that assumes a direct conversion between the C^+ and CO abundances (intermediate species are neglected). The CO formation is triggered by a radiative association between C^+ and H_2 to form hydrides that react afterwards with atomic oxygen. The CO destruction depends on the ultraviolet (UV) photo-dissociation rate from de Jong (1977) and Falgarone & Puget (1985), which is a function of the gas number density and the visual extinction of the medium scaled by (i) the strength of the UV portion of the interstellar radiation field (assumed in our simulations to be the same as that of the solar neighbourhood derived by Draine (1978)) and by (ii) a scattering factor that enhances the extinction. The UV extinction of the medium is calculated using the TREECOL algorithm (Clark et al. 2012), which considers the H_2 and CO self-shielding and the shielding of CO by H_2 and by dust absorption.

The non-equilibrium hydrogen chemistry implemented in our Cloud Factory involves reactions between molecular (H_2), atomic (HI) and ionised (H^+) hydrogen, electrons, cosmic rays, dust grains and the UV radiation field. This encompasses H_2 formation on grains, collisional and photo-dissociation of H_2 , cosmic rays and collisional ionisation of HI, and H^+ recombination in the gas phase or on dust grains. See the reaction formulas in Table 1 and their analytical treatment in Section 2.2 of Glover & Mac Low (2007a). The net energy exchange due to radiative and chemical gas heating or cooling is computed using the atomic and molecular cooling function as outlined in Clark et al. (2019).

Our model of star formation uses a hybrid approach of sink particles that can represent either individual stellar systems or clusters of stars depending on the target mass resolution of the region where they form. In order to become a sink particle, following Bate et al. (1995), a cell must be above a critical density ρ_c and satisfy energy checks to confirm that it is bound and the internal collapse is runaway: it must be located at a gravitational minimum, outside the

accretion radius of any other sink particle and have inward velocities and accelerations. The sink particles behave like non-gaseous bodies that interact gravitationally with the surrounding medium and can accrete material from neighbouring bound cells that are within a given sink accretion radius. Sinks are especially helpful to set up a natural halt threshold in the code and prevent excessive mesh refinements (Huber et al. 2013) but at the same time to keep track of the sites where stellar feedback will be injected in form of thermal energy and/or momentum from supernovae explosions (Gatto et al. 2015; Walch et al. 2015; Girichidis et al. 2016; Padoan et al. 2016a,b; Pan et al. 2016). In this paper, we only consider supernovae as they are the most energetic source of stellar feedback and generally accepted to be the dominating driving mechanism of turbulence in star-forming galaxies (Mac Low & Klessen 2004; Padoan et al. 2016a). Either way, other mechanisms such like outflows/jets (Nakamura & Li 2007; Cunningham et al. 2011; Federrath et al. 2014), stellar winds (Dale & Bonnell 2008; Peters et al. 2017; Gatto et al. 2017) and local photoionizing radiation (Peters et al. 2017; Haid et al. 2018) are also present in real scenarios.

We consider two ways of injecting stellar feedback: (i) purely random supernovae explosions and (ii) supernovae tied to star formation sites. For the first approach, we randomly distribute the supernovae according to the gas density profile of the galactic disc. We adopt a rate of 1 event every 50 years as estimated from Milky way observations of Gamma-ray emission in massive stars (Diehl et al. 2006). Our second approach produces bursts of strong feedback from the spiral arms. It assumes a star formation efficiency and stellar Initial Mass Function in order to compute the amount of massive stars ($> 8 M_{\odot}$) that will undergo supernovae explosions at the end of their lifetime. We use the stellar mass function from Kroupa (2002) and calculate the number of massive stars out of the stellar content of sink particles following Sormani et al. (2017). Since our low resolution target mass ($\sim 100 M_{\odot}$) correspond to molecular cloud scales, we estimate a rough star formation efficiency in sinks of 1–2% based on the work of Krumholz & Tan (2007). In contrast, the high resolution target masses ($10–0.25 M_{\odot}$) achieve significantly higher formation efficiencies as they already conform protostellar cores and protostars, so we rather use 33% as in Matzner & McKee (2000). Moreover, the way the energy is released into the gas depends on whether the Sedov-Taylor phase of the supernovae expansion is resolved by a sufficient number of cells, similar to the approach of Hopkins et al. (2014) and Gatto et al. (2015) and introduced analytically by Blondin et al. (1998). In our case, if the expansion phase is resolved by 32 cells we can inject thermal energy directly into the surrounding gas, otherwise we inject terminal momentum pointing radially outwards. Further details of our supernovae model can be found in Paper I.

2.2 Simulation Setup and Refinement Scheme

In Paper I we generated resolved molecular clouds in a galactic potential using the following procedure.

The initial gas distribution of our simulation was based on observational constraints and theoretical modelling of the Milky Way presented in McMillan (2017). They suggest an exponential profile for the HI and H₂ densities as a function of the radius of the galaxy disc. We rather take the mass contribution from both profiles and start with a single gas distribution consisting only of HI, from which H₂ will form self-consistently according to our chemical treatment as the galaxy evolves. The domain of our simulation comprises radii between 4 kpc and 12 kpc as we are interested in analysing properties of clouds outside the Milky Way’s central bar.

In the first stage of the simulation we let our galaxy evolve for 150 Myr under the effects of the large-scale potential and the energy/momentum feedback from random supernovae to naturally form spiral arms and reach a steady state. For mesh refinement purposes we set the cell target mass to be $1000 M_{\odot}$ during this period.

Next, we start the middle phase of the simulation by turning on a co-rotating 3 kpc high resolution box centred at a galactic radius of 8 kpc. This phase lasts for around 70 Myr, or two spiral arm passages. In this phase we launch 3 runs undergoing different physical mechanisms as follows: (a) a potential dominated scenario in which the ISM dynamics respond only to the large-scale gravitational potential and the random supernovae feedback as it was set up during the first stage of the simulation, (b) same as the previous case but this time gas self-gravity between cells is included, and (c) a feedback dominated scenario in which both the large-scale potential and self-gravity effects operate; there is sink particle formation and the supernovae feedback is mixed. By mixed feedback we mean that both the random and clustered supernovae tied to sink particles are turned on (see our supernova implementation in Sec. 2.1). For the random supernovae we assume a rate of one every 350 years in the mixed case. The target mass of this phase is initially set to $100 M_{\odot}$ but is further lowered down to $10 M_{\odot}$ for the final 10 Myr.

Regardless of the stage of the simulation, we require at all points in the AREPO mesh that the Jeans length is resolved by at least 4 cells to adequately check energy and bounding conditions of the region and avoid artificial fragmentation (Truelove et al. 1997). If sink creation densities are achieved but the gas fails to pass the energy checks (see Sec. 2.1), we continue to resolve the gas until it is unambiguously bound as long as the creation density persists satisfied.

In the final stage of our simulation, and in order to resolve substructures within the processed cold ISM, we further increase the resolution on individual cloud complexes of ~ 100 pc radius within the high resolution box by injecting Monte Carlo tracer particles (Genel et al. 2013) everywhere the gas density is above 100 cm^{-3} . The target mass is lowered down to $0.25 M_{\odot}$ where tracer particles are present, which allows us to achieve high spatial resolutions (e.g. $d_{\text{cell}} = 1 \text{ pc}$ at $n = 5 \text{ cm}^{-3}$, $d_{\text{cell}} = 0.03 \text{ pc}$ at $n = 10^5 \text{ cm}^{-3}$). For this target mass, we set a sink creation density of $\rho_c = 10^4 \text{ cm}^{-3}$, which according to Mac Low & Klessen (2004) (and based on the size scales reachable in this phase) corresponds either to individual star-forming clumps or protostellar cores. Unlike Paper I, in this phase we split the potential-dominated scenario in two cases, with and without self-gravity, in order to explore the effects of local gravitational forces on the dynamical signatures retrieved from our cloud complexes.

2.3 The Simulated Cloud Complexes

In order to investigate the turbulence in our self-consistent clouds we extract the same clouds as in Paper I, but with one small difference. In addition to complexes A and B we also include complexes A₀ and B₀ in which self-gravity is not turned on.

Our selected cloud complexes arise in very different galactic environments. Cloud complexes A and A₀ are at the same location in a dense spiral arm, whereas B and B₀ are in an inter-arm, more diffuse region of the Galaxy. Cloud complexes C and D, representing the feedback-dominated case, are the densest regions in the high resolution box where there are no massive sink particles yet. Complex D (the most massive of the two), however, gets to form new massive stars during subsequent time steps and, as a conse-

Cloud Complex	Galactic potential	Self-gravity	Supernovae feedback		Description
			Random	On sinks	
A ₀	✓	✓			Inside arm
B ₀	✓	✓			Inter-arm region
A	✓	✓	✓		Inside arm
B	✓	✓	✓		Inter-arm region
C	✓	✓	✓	✓	No embedded SNe
D	✓	✓	✓	✓	Embedded SNe

Table 1. Cloud complexes analyzed in this work and physical mechanisms operating in each, with a short description about their surrounding environment.

quence, gains further energy feedback from supernovae explosions taking place within itself. Complex C, conversely, has no embedded supernovae and hence the energy feedback is external only. This information is briefly summarized in Table 1. Also, a figure showing the exact location of these complexes can be found in Paper I.

We will refer to the cloud orientations explored in this work as the face-on, edge-on $\phi=0^\circ$ and edge-on $\phi=90^\circ$ views. The face-on line of sight points towards the $-\hat{z}$ direction, meaning that the cloud complex is viewed from above the galaxy, whereas the edge-on $\phi=0^\circ$ and edge-on $\phi=90^\circ$ line of sights point along the $+\hat{y}$ and $-\hat{x}$ directions respectively.

Figure 1 shows HI, H₂ and CO column density distributions of the cloud complexes, 2 Myr after tracer refinement has commenced, from face-on and edge-on $\phi=90^\circ$ orientations. Note from the green contours on the face-on view that the CO does not uniformly trace the H₂-dense regions of the complexes, which is a natural consequence of collisional and photo-dissociation processes induced in our simulations by nearby supernovae feedback and the global Interstellar Radiation Field (e.g. Smith et al. 2014a).

3 RADIATIVE TRANSFER MODELLING

We perform radiative transfer simulations of our synthetic cloud complexes using the Line Modelling Engine code (LIME, Brinch & Hogerheijde 2010). LIME is a flexible code for predicting molecular line and dust continuum emission from arbitrary 3D geometries in (sub-)millimetre and infrared wavelengths.

By default, LIME generates grid points according to the input density distribution, but can alternatively take any other user-defined physical distribution. Then, the code constructs a Delaunay triangulation and its corresponding Voronoi grid to do photon propagation and integrate the transfer equations. To solve the line excitation problem, LIME comprises two approaches suitable for matter in different equilibrium states.

(i) A Local Thermodynamic Equilibrium (LTE) approximation in which the radiative features of the gas are fully and uniquely determined by the local kinetic temperature and its internal properties, namely, the atomic/molecular level populations are dominated by particle collisions and obey a Maxwell-Boltzmann distribution law. The code uses this distribution to calculate the level populations and the Kirchoff's law for thermal radiation to solve the transfer equation (see e.g. Rybicki & Lightman 1986).

(ii) Additionally, LIME has a non-LTE mode for media in which the level populations are not only ruled by collisions but also have

a non-negligible contribution from the local radiation field. This problem needs to be addressed iteratively over the physical domain, taking into account the outgoing radiation from all the grid cells with each other. LIME solves this by propagating randomly oriented packages of photons from every cell of the grid, along lines of the Delaunay triangulation. The algorithm computes provisional level populations in each cell using the incoming local radiation and collisional rates. Each cell releases a number of photons proportional to the number of neighbouring cells. When the propagating photons escape the physical domain, the calculation stops and the whole process is repeated, ideally, until convergence (i.e. populations in equilibrium) is reached. As convergence depends on the input physical distributions, the number of iterations is not fixed by LIME but left as free parameter for the user to decide when to stop.

Once the level populations are determined in either of the modes, LIME integrates the transfer equations along isotropic tracer rays that cross the Voronoi grid (associated to the Delaunay triangulation), which stores the input and derived physical properties of each cell, until they hit the border of the physical domain. The resulting specific intensity is then used to compute the observed intensity at the distance, spatial and spectral resolutions established by the user.

The output FITS file is a 3-dimensional $N_{\text{pix}} \times N_{\text{pix}} \times N_{\text{chan}}$ position-position-velocity (PPV) cube containing intensity (in units of Jy pix^{-1} , Brightness temperature or $L_\odot \text{ pix}^{-1}$) or optical depth (τ) information as a function of the spectral channel (in m s^{-1}).

Full radiative transfer modelling is necessary to produce realistic synthetic observations that can be readily compared to observational data with analogous methods, especially in the present times with the advent of new telescopes and techniques that allow resolving non-ideal and highly coupled regimes. Evidently, the resulting cubes also inherit observational limitations such as uniform spatial and spectral resolutions and natural constraints from projection and optical depth effects.

The original version of LIME is publicly available¹, however, in Section 3.2 we present new features of an already customised version of LIME that allow the user to run radiative transfer simulations on AREPO-like unstructured meshes as those used in this work.

3.1 non-LTE $^{12}\text{CO J}=1-0$ Line Excitation

Our analysis is focused on the carbon monoxide emission at its ground state rotational transition $^{12}\text{CO J}=1-0$, which has been widely employed as a tracer of molecular, relatively dense, cold gas (van Dishoeck & Blake 1998; van Dishoeck 2004) as well as in previous statistical studies (Larson 1981; Heyer & Brunt 2004; Roman-Duval et al. 2011; Bertram et al. 2014) of both simulated and observed molecular associations.

We use non-LTE calculations because, in our simulations, a considerable fraction of the gas H₂ density (the main collisional partner of ^{12}CO) is below the critical density ($n_{\text{crit}} \sim 2 \times 10^3 \text{ cm}^{-3}$) to collisionally populate ^{12}CO at the upper level of its ground transition, which is valid within a wide range of temperatures for this transition (2 – 3000 K, Yang et al. 2010). This implies that the first rotational level of ^{12}CO is populated by different mechanisms in our cloud complexes; it is dominated by radiation in diffuse regions and thermally controlled by collisions in dense regions (see Sec. 12.4 of Wilson et al. 2013). The critical density is defined as $n_{\text{crit}} = A_{ij}/C_{ij}$, which involves the Einstein emission coefficient

¹ <https://lime.readthedocs.io>

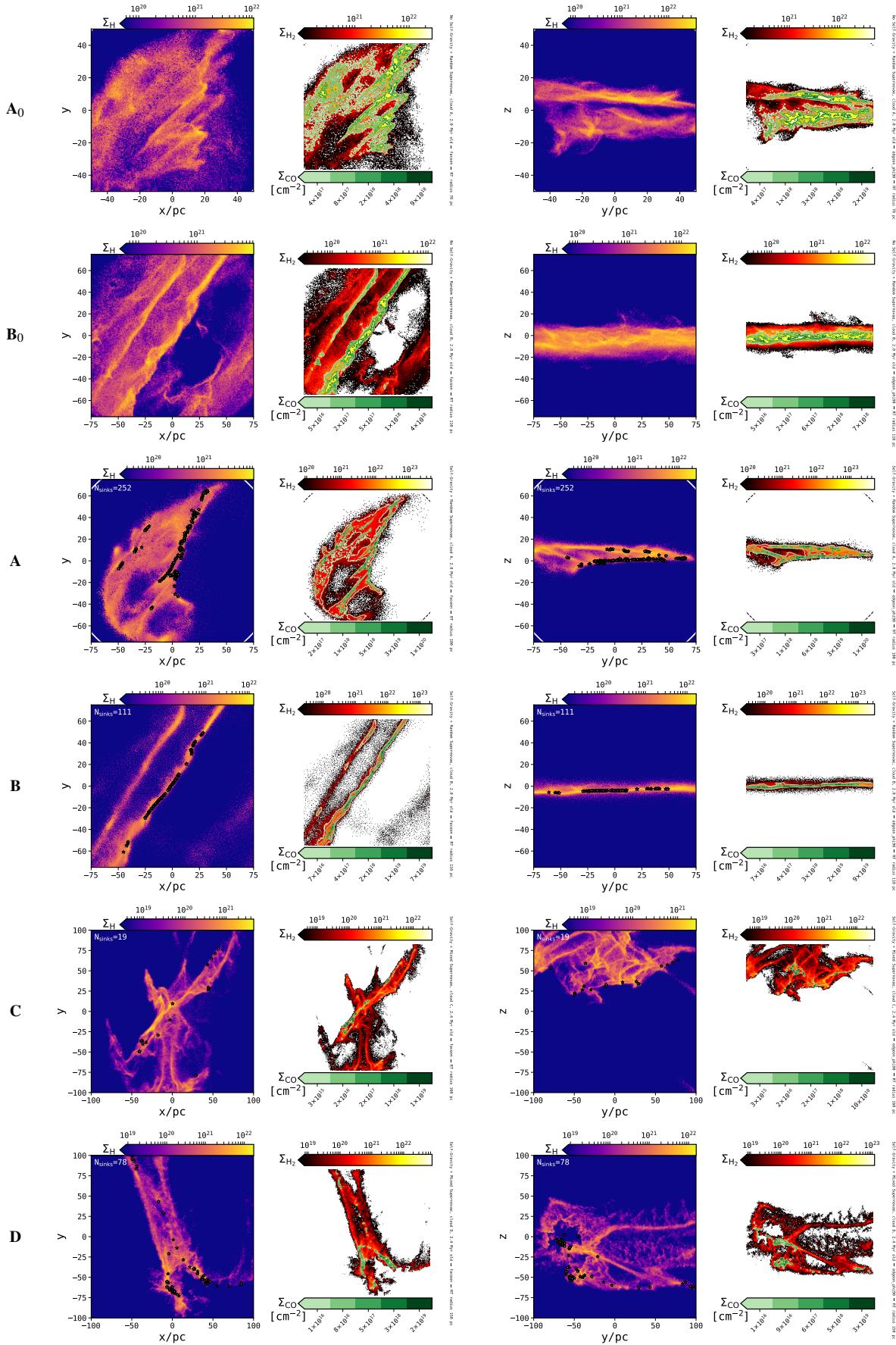


Figure 1. Face-on (left-column panels) and edge-on ($\phi=90^\circ$ right-column panels) projections of HI, H₂ and ¹²CO (green contours) column densities of cloud complexes extracted 2 Myr after injecting tracer particles. When small enough, the radiative transfer domain is shown as a circle enclosing the region. If any, sink particles are overlaid on HI maps with star markers.

A_{ij} for the transition $i \rightarrow j$ and the collision rate coefficient C_{ij} associated to a given collisional partner at a kinetic temperature T_k . The ^{12}CO main collisional partners are the spin isomers of molecular hydrogen, known as parahydrogen and orthohydrogen, where the parahydrogen is a lower-energy state with less magnetic moment than orthohydrogen.

In Appendix A1, we investigate how our line profiles would vary if LTE approximations were used instead. There, we also present the impact of considering LTE simulations on the retrieved structure function parameters (see Sec. 4.1) of our cloud complexes.

3.2 New Tools for handling AREPO–like meshes in LIME

3.2.1 Input Unstructured Meshes and Implementation of *kdtree* Algorithm in LIME

Our previous customisation of LIME allowed the user to compute the radiative transfer solutions on analytic 3D models of star-forming regions generated with the SF3DMODELS package (Izquierdo et al. 2018). These models were based in a uniform-Cartesian grid and then mapped by LIME via random grid points weighted by the density distribution of the model. Cartesian grids are computationally efficient and a good approach to problems with low-dynamic range. Our current simulations, however, are based in (highly) non-uniform Voronoi grids especially adapted to track the effects of physical mechanisms governing the gas dynamics at several time and length scales. On top of that, the statistical analysis of this work requires well-determined spatial scales in order to properly and unambiguously uncover the turbulent nature of our cloud complexes via the velocity structure function (see Sec. 4.1).

For these reasons, we adapted LIME so that it could also handle input unstructured meshes pre-processed with SF3DMODELS. This new mode turns off the default random-generated grid in LIME and forces it to take the exact AREPO cell locators to later reconstruct its own Voronoi grid for the radiative transfer.

Another addition to the code is motivated by the fact that, unlike Cartesian grids, finding neighbouring cells in unstructured grids is non-trivial. This is necessary during the ray-tracing algorithm in which the radiative transfer equations are computed iteratively over subsequent cells. Clearly, a ‘brute force’ search that minimizes the distance from the test cell to the whole set of cells is too slow (of order $O(N^2)$) with increasing the number of cells N . To solve this, we implement a k-d tree algorithm that splits the search domain in representative areas to discard unnecessary distance checks to remote cells. We use the third-party, open source, C library *kdtree*², which is pre-built in our customised version of LIME and does not need any particular installation.

Additionally, the LIME domain is spherical and surrounded on its surface by randomly distributed points called ‘sinkpoints’ (not to be confused with the sink particles that we use to represent star formation in the AREPO simulations), throughout which tracer rays pop out from the physical model to make up the synthetic image. Since the input unstructured grid can have any shape, big voids between sinkpoints and inner physical grid points are likely to exist, leading to artifacts at the borders of the image as the reconstructed Voronoi cells are larger there than in the original mesh. We soften this effect by including empty ‘dummy’ grid points close to the border of the radiative transfer domain using our new GRID.FILLGRID module of SF3DMODELS. We generate $N/10$ dummy

points and distribute them randomly between a radius enclosing the 90% of the total gas mass and the maximum radial position of the grid. This needs to be done conscientiously; for instance, including dummy cells from small radii would cause too many artificial holes in the region and hence an underestimation of actual gas masses as the volume of grid cells is obviously reduced to open space for the new dummy cells.

3.2.2 Removal of Twin AREPO Cells after a Refinement Step

According to Springel (2010), a cell fulfilling user-defined refinement criteria is split along its centroid into two cells. Right away, the position locators of the newborn cells coexist with the original cell centre. During few subsequent time steps, the new locators are separated away via mesh-regularization techniques until they reach the actual centroids of the split cells. This opens the possibility of finding two or more cell locators at the same position in a single time snapshot of the simulation, which might lead to errors during the grid reconstruction by the triangulation algorithms of the radiative transfer codes or slow down intermediate smoothing steps if any.

We have written an efficient iterative algorithm included in the AREPO.UNIQUECELLS module of SF3DMODELS to get rid of these ‘twin’ cells. Broadly speaking, the algorithm constructs an array of cell id’s based on radially-sorted non-repeated cell positions and inspects whether the cells are twins or not only within ‘holes’ of the id’s array. The algorithm recycles the mass of the twin cells into the survivor cell using a direct summation, whereas the other physical properties remain unchanged as they are approximately equal in all the twin cells. It returns a clean dictionary with unique cells and their new physical properties.

The SF3DMODELS package and the latest customised version of LIME are open source and documented³. New contributions are welcome.

3.3 Radiative Transfer Setup

The front of the cloud complexes domains are deliberately set at 2.4 kpc from the observer, a typical distance to nearby molecular cloud complexes (e.g. W3, W33, M17). The pixel size of the PPV cubes is 26'', which translates into a projected physical resolution of 0.3 pc. This pixel size is a good compromise between resolving most of the cells in our AREPO meshes (which contain cells as small as ~ 0.03 pc) and the processing time of our images. We determine the spectral resolution, or channel width, for each cloud complex and orientation considering the relation $\Delta_{\text{chan}} = v_{\max}/(n_{\text{chan}} - 1)$, where v_{\max} is the maximum projected velocity along each of the 3 line-of-sights explored in this work (face-on, edge-on $_{\phi=0^\circ}$ and edge-on $_{\phi=90^\circ}$). The number of channels is constant ($n_{\text{chan}} = 101$) to ensure the same dimensionality of the principal component analysis for all the cloud complexes. For the non-LTE runs we assume 50 iterations as the populations level out around 30 cycles. We tested 100 iterations for some regions and found differences of $< 5\%$ on the computed mean intensities, which is well within the random nature of the algorithm.

For consistency, we include micro-turbulence as an additional source of line broadening to account for non-thermal motions in scales smaller than the cell size. This contribution is assumed to be equal to the sound speed of each cell of the mesh and added in

² <https://github.com/jtsiomb/kdtree>

³ <https://star-forming-regions.readthedocs.io>

quadrature to the standard thermal broadening during the line transfer. For cold dense cells, where the ^{12}CO mass is non-negligible, micro-turbulence is generally much smaller than their bulk speed ($\times 10 - 100$ lower). Note there may be additional unresolved small scale turbulence within the cell so our assumed micro-turbulence is a lower limit.

The CO mass is consistently calculated by providing LIME with all the hydrogen species available from our HD simulations and specifying that the CO abundance refers to hydrogen nucleons. We assume that the CO collisional partners, the spin isomers of molecular hydrogen (para-H₂ and ortho-H₂), split 50/50 the H₂ distribution as they can reach equilibrium at timescales similar to the evolution time of our simulation snapshots ($t \sim 1$ Myr) according to Flower et al. (2006). They computed the evolution of the para-ortho ratio using simulations of cold ($T = 10 - 30$ K), dense ($n_{\text{H}} = 10^4$ cm⁻³) gas, starting from a steady ratio 1:3 as orthohydrogen has 3 possible equivalent energy states (see also Vaytet et al. 2014).

4 STATISTICAL DESCRIPTION OF TURBULENT MOTIONS IN MOLECULAR ASSOCIATIONS

4.1 Velocity Structure Function of a Turbulent Medium

The concept of a generalized function to describe non-thermal velocity fluctuations in a 3D fluid came up with the work of Kolmogorov (1941). He considered an incompressible viscous fluid with very large Reynolds numbers ($\rightarrow \infty$), namely, locally dominated by isotropic turbulent motions, and assumed all the components of the turbulent velocity to be homogeneous and statistical random variables. Using similarity hypotheses on time and length-scale energy dissipation rates, he found that the (averaged) velocity dispersion is a function of the spatial separation between test points of the fluid. He envisioned the idea of rapid successive transport of turbulent kinetic energy, from large (low order) to small (high order) scales of the fluid, as a cascading process. Higher order scales end up dispersing energy in the form of heat as viscosity exceeds the magnitude of velocity fluctuations in small scales. Later, Onsager (1949) used an analogous theoretical basis to demonstrate that 3D vorticities can accelerate the turbulent cascade and explain the rapid viscous dissipation of energy with increasing wavenumber. Based on this, he found the characteristic energy spectrum of Kolmogorov-like fluids

$$E(k)dk = CQ^{2/3}k^{-5/3}dk, \quad (1)$$

where C is a dimensionless constant, $k = 2\pi/l$ is the wave-number associated to a given size scale l in a Fourier expansion of the velocity field, and $E(k)dk$ stands for the kinetic energy distribution within an interval dk , which is being dissipated as heat at a rate Q .

However, incompressible flows of the Kolmogorov's scenario are rare in molecular clouds, where non-thermal motions are not negligible but rather transonic or supersonic (McKee & Ostriker 2007). In particular, supersonic motions lead to shock-dominated turbulence (Burgers-like turbulence, Kraichnan 1974; Frisch et al. 2001), which serves as a mechanism of energy diffusion at large scales apart from just heat dissipation at small scales. This makes the energy spectrum to decay faster at high order wave-numbers,

$$E(k) \propto k^{-2} \quad (2)$$

Moreover, it is well known from early studies that the gradient of energy dissipation in non-ideal turbulent fluids depends

on the medium properties (Kraichnan 1974) and the turbulent cascade must reflect this on different size scales. Hence, a power law generalization of the energy spectrum is a reasonable approach to account for intrinsic velocity fluctuations as a function of the input environment

$$E(k) \propto k^{-\beta}. \quad (3)$$

The exponent β is known as the spectral index of the energy spectrum in a three-dimensional turbulent fluid and is an intrinsic property of the velocity distribution in the medium. In addition, this powerlaw dependence is supported in Onsager (1949) by the fact that the total vorticity of a fluid is in general a linear combination of the wave-number scale vector \vec{k} .

Using eq. 3, it is possible to compute the mean square velocity fluctuations at a given size scale l by summing up the energy contributions from higher order (smaller) scales as follows

$$\langle |\delta v(l)|^2 \rangle \propto \int_{2\pi/l}^{\infty} E(k)dk \propto \int_{2\pi/l}^{\infty} k^{-\beta} dk \propto l^{(\beta-1)}; \quad (\beta > 1), \quad (4)$$

from which the root-mean-square (rms) velocity is straightforward,

$$\langle |\delta v_l|^2 \rangle^{1/2} = v_0 l^{\gamma_{\text{rms}}}, \text{ with} \\ \gamma_{\text{rms}} = (\beta - 1)/2, \quad (5)$$

where v_0 and γ_{rms} are the intrinsic scaling parameters of the rms velocity fluctuations field.

A generalized description of velocity fluctuations, the so-called velocity structure function, was introduced by Kolmogorov (1941) and further developed by Anselmet et al. (1984); Frisch (1995) in order to (statistically) explain multi-component turbulent motions that the scaling parameters of the rms-velocity field are unable to model comprehensively (see Brunt et al. 2003). In molecular clouds, multi-component turbulence arise from several dissipation mechanisms such as shocks, magnetic fields, radiative cooling, and heat diffusion (Boldyrev et al. 2002), but also from energy-injection mechanisms like stellar feedback, both affecting different time and length scales of the region. The structure function is usually presented as

$$S_p(l) = \langle |\delta v(l)|^p \rangle \propto l^{\zeta_p} \iff \langle |\delta v(l)|^p \rangle^{1/p} \propto l^{\zeta_p/p} \\ \therefore \langle |\delta v(l)|^p \rangle^{1/p} = v_0 l^{\gamma_p} \quad (6)$$

with $\gamma_p = \zeta_p/p$

where p stands for the order of the velocity fluctuations δv and provides information about the degree of coherence of the velocity field when subject to spatial variations. The intrinsic scaling exponent γ_p is the original exponent ζ_p divided by the order p , and v_0 is the magnitude of velocity fluctuations known as scaling coefficient or normalisation of the structure function. This expression satisfies the upper limit of the Hölder Condition for the metrics δv and l , which implies that the distribution of velocities is uniformly continuous within the spatial domain of the fluid. In Section 4.2.1, we expand the discussion on structure function variations with changing p -orders.

As final comment, note from eqs. 1, 2 and 5 that the Kolmogorov's law and the Burgers-like turbulence yield rms structure function scaling exponents $\gamma_{\text{rms}} = 1/3$ and $\gamma_{\text{rms}} = 1/2$ respectively.

4.2 Principal Component Analysis

In this subsection we briefly summarise the main aspects of the principal component analysis (PCA) technique as an application for revealing the structure and dynamics of molecular associations in the ISM.

PCA is a statistical multivariate method that transforms an input dataset of n , possibly correlated, variables into a new object spanned by a set of m orthogonal uncorrelated components (called principal components) in such a way that the variance of the initial dataset, which can be seen as the amount of information as a function of the original variables, is (non-redundantly) maximized along subsequent principal components. This property allows the dimensionality of the analysis to be reduced to only the components that hold most of the variance of the data, i.e $m \leq n$.

The method's theoretical framework was originally presented by Pearson (1901) and Hotelling (1936) but adapted for the first time to the study of ISM dynamics in the work of Heyer & Peter Schloerb (1997), who described the formalism of the technique when considering position-position-velocity (PPV) data cubes and demonstrated its ability to retrieve velocity fluctuations within characteristic spatial scales from synthetic models with well-known line profiles and noise level. However, it was not until the study of (Brunt & Heyer 2013, hereafter BH13) that the method, applied to spectroscopic images, acquired a formal theoretical foundation.

Heyer & Peter Schloerb (1997) applied the PCA technique to a sample of 4 cloud complexes and found, for each, a correlation of the form

$$\delta v = v_0 l^\alpha, \quad (7)$$

linking velocity fluctuations (δv) to characteristic size scales (l) in each complex. Note that we use α (instead of the intrinsic γ) to refer to the PCA derived scaling exponent. This seems analogous to the empirical linewidth-size relationship

$$\Delta v = C L^\Gamma \quad (8)$$

found by Larson (1979, 1981) after mixing global velocity dispersions (Δv) from an ensemble of several molecular associations, as a function of their projected sizes (L). We use capitals C and Γ to indicate that the scaling coefficient and exponent, respectively, were derived from global-cloud scales rather than from inner-cloud scale analysis.

Both relationships were systematically reported in subsequent works, some of which are listed in Table 2 along with a short description. Also, both of them resemble a structure function-like dependence (see eq. 6), but actually do not represent the natural behaviour of non-thermal motions mainly due to projection and radiative transfer effects. It is possible, however, to connect these pseudo-structure functions to intrinsic structure functions via calibration relations and universality principles of turbulence demonstrated in later works and shortly summarized in Sections 4.2.1 and 4.2.2.

We use the PCA module included in the TURBUSTAT Python package⁴ (Koch et al. 2019) to retrieve characteristic size and spectral scales ($l, \delta v$) from our cloud complexes. We constrain the PCA algorithm to keep most of the data variance in principal components (95–99% of the total variance), but not too much (i.e. $\gg 99\%$) in

order to avoid artificial clustering of points at the minimum recoverable scales of intensity cubes. The PCA pseudo-structure functions and their corresponding scaling parameters (v_0, α) are computed separately by our new open source PCAFACtory package, according to three PCA extraction methods focused on different analysis windows (see Sec. 5).

4.2.1 Calibration from PCA to Structure Function Scaling Parameters

Brunt & Heyer (2002) tested the sensitivity of the technique to different energy spectra $E(k) \propto k^{-\beta}$ from simulated data and found an empirical calibration that relates the PCA derived scaling exponent α with the spectral index β , which is intrinsic of the 3D velocity distribution. Equivalently, Roman-Duval et al. (2011, hereafter RD11) obtained almost the same calibration using a wider range of synthetic scenarios with different intermittency (both spatial and temporal fluctuations in the turbulent flow) and power spectra for density and velocity. It is valid below a relatively high level of density variability and for spectral indices β between 1.2 – 2.6. Later, BH13 derived a slightly steeper but very similar calibration from an analytic point of view. However, in this work we use the RD11's calibration because, as BH13 suggest, their analytic result should be seen as a supportive element rather than a replacement to the previous empirical estimations given the approximations used in their derivation, then

$$\beta = (0.20 \pm 0.05) + (2.99 \pm 0.09)\alpha. \quad (9)$$

This calibration is useful to connect PCA scaling parameters derived from genuine observables to the intrinsic energy distribution within a given 3D fluid volume. Observables in molecular clouds results from line-of-sight projected averages of velocity fluctuations in the best of the cases, but there might also be optical depth effects involved.

The retrieved spectral indices β are related to the 2nd structure function scaling exponent γ_{rms} via eq. 5, however, as mentioned in Section 4.1, the rms velocity is generally not a full representation of the fluctuations field in a fluid due to its multi-component nature triggered by several energy dissipation and injection mechanisms.

Brunt et al. (2003) demonstrated that eq. 9 is a good bridge between α and $\gamma_{rms} \equiv \gamma_2$, which even also holds for orders $p \neq 2$ as long as the velocity field of the region is non-intermittent and therefore reproducible by a single γ . For completeness of details, it is straightforward showing from eq. 6 that γ_p is constant ($\gamma_p = \gamma_2 \equiv \gamma$) if and only if ζ_p depends linearly on p ($\zeta_p = p\gamma$). This is not the case for intermittent velocity fluctuations, in which the exponent γ_p is no longer constant but depends on the order p of the function. Furthermore, they concluded on intermittent fields that the PCA derived exponent α , after calibration, is better correlated with structure function exponents of orders $p = 0.5, 1$, which means that the translation given by the combination of eqs. 5 and 9 is preferentially a measure of the intrinsic scaling exponent γ_p of low-order rather than rms velocity fluctuations. In Section 6.4, we semi-quantitatively study intermittency effects on the retrieved scaling parameters by considering different analysis windows in our cloud complexes (smaller windows reduce intermittent velocity fluctuations).

⁴ <https://turbustat.readthedocs.io>

4.2.2 Universality of Turbulence in the Molecular ISM

Based on earlier works, Larson (1981) collated of three-dimensional rms velocity dispersions along with the (maximum) projected sizes L (in pc) of diverse molecular associations and derived the relation $\Delta v = 1.10L^{0.38}$ for $0.1 \lesssim L \lesssim 100$ pc. This can even be extended up to larger-scale ~ 1000 pc interstellar motions (Larson 1979). He found complementary relationships that connect mass and density to the dynamics of molecular clouds. Similarly, but from a more homogeneous sample, Solomon et al. (1987) found the relation $\Delta v = (1.0 \pm 0.1)L^{0.5 \pm 0.05}$. For the velocity dispersion calculation, they extracted centroid and line-width based velocity differences and added the contributions in quadrature.

None of these relationships is a direct representation of a true velocity structure function in the sense that the studied objects do not correspond to a unique self-interacting environment. It is rather an evaluation of the structure function in a particular case ($l = L, p = 2$) for each object, which, when combined, turn out to have the same functional form as if they were part of a single fluid.

Heyer & Brunt (2004) applied PCA on individual molecular clouds and performed Monte Carlo simulations to prove that this resemblance is a consequence of the turbulence universality within the molecular ISM across the galaxy, which hints at common formation mechanisms for molecular clouds. For this reason, let us adopt the following 1-to-1 translations from Larson-like to structure function parameters,

$$\begin{aligned} C &\approx v_0 \\ \Gamma &\approx \gamma_{\text{rms}}. \end{aligned} \quad (10)$$

For standardization purposes, we use the 2nd order calibrations and the universality condition of turbulence to translate, respectively, PCA derived exponents α and Larson-like exponents Γ to rms structure function scaling exponents γ_{rms} , to which we will refer from now on as just γ .

4.2.3 Optical Depth Effects on PCA Pseudo-Structure Functions

Larson (1981) considered optically thin regions, mostly traced by ¹³CO, H₂CO and NH₃, and only two ¹²CO optically thick regions where the large-scale velocity variations were dominant (compared to the smaller-scale fluctuations derived from linewidths) in order to avoid any effects of line saturation. However, modelling and observational studies have concluded that PCA derived scales are nearly insensitive to the opacity regime of the emission (Heyer & Peter Schloerb 1997; Brunt 2003b; Brunt & Mac Low 2004; Heyer et al. 2006; Brunt et al. 2009; Roman-Duval et al. 2011). This is a strong point of the method as it allows studying a broader range of objects/scales traced by opaque emission. Brunt & Heyer (2013) suggest that this is a consequence of the centroid velocity not being affected by saturation in the optically thick regime as long as it is symmetric to the line central frequency.

In Section 6.2, we discuss optical depth effects by comparing the retrieved scaling parameters from cloud complexes at different orientations. Also, we perform PCA runs straight on optical depth maps (tracing cold, dense gas) from selected clouds and present a short discussion on discrepancies to ¹²CO J=1–0 intensity derived ($l, \delta v$) scales in Appendix ??.

5 GENERAL METHODOLOGY

We implement three PCA extraction methods to investigate the response of retrieved pseudo-structure functions to different analysis windows in our cloud complexes.

Our cloud complexes are much bigger than the individual clouds considered in nearby regions as used by Larson (1981); Brunt (2003b); Heyer & Brunt (2004). Thus, the main (and default) method of our work consists in using PCA derived scales ($l, \delta v$) from smaller cloud portions (30 pc) distributed on each cloud complex (~ 200 pc), and combine them all to construct a single fit equation representing the pseudo-structure function of the complex. We call this method the “Mixed method”.

In the second method, we analyse the cloud complexes as they are, without any kind of sub-portions. This method is named the “Complex method”. Intermittency of density and velocity fields is more likely to happen in this case as all the structures of the cloud complex are taken into account in a single analysis window. See some of the expected implications of having intermittent fields in 4.2.1.

For the third method, we do not combine the PCA derived scales from individual portions but rather use them to compute their pseudo-structure functions separately, as individual objects. Given that these square portions are 30 pc wide, a typical size of Giant Molecular Clouds, we call this method the “Individual Cloud method”. In the succeeding sections, we refer to individual portions interchangeably as “individual clouds” or just “clouds” to keep this convention. The Complex method extraction procedure is straightforward. We take the intensity cubes and pass them directly through the TURBUSTAT.PCA module to compute spatial and velocity scales of each cloud complex as a whole. The Mixed and Individual Cloud methods require, however, a couple of further steps explained below in detail.

First, we use the intensity position-position-velocity (PPV) cubes obtained from the radiative transfer solutions of our cloud complexes to determine their associated moment 0 map. The moment 0 is defined as the integrated intensity along the velocity (or frequency) axis as follows.

$$M_0(x, y) = \int_l I(x, y; v) dv, \quad (11)$$

where x and y refer to the pixel location on the synthetic cube.

Then, we use the ASTRODENDRO [http] package with each moment 0 map to work out a hierarchical tree, called dendrogram, which divides the input 2D-field into closed subregions according to equal-valued pixels. We consider moment 0 because we want to track and study most of the ¹²CO J=1–0 emission regions as possible, regardless of the analysis frequency.

From the computed dendograms, we focus on extracting the outermost, irreducible, substructures of the tree, called “leaves”, which are usually the smallest objects. Next, we find the peak pixels of leaves and centre 30 pc square portions around them in order to cover the whole emission map. We slice the ¹²CO cubes using the portions coordinates and assign a colour code to each region that will be consistent on every future figure. Finally, PCA is computed on the cube slices to construct the Mixed and Individual Cloud fits. Figure 2 is a cartoon representation of the procedure described up to this point.

For the fitting process we randomise the PCA derived scales ($l, \delta v$) N times using their uncertainties ($\sigma_l, \sigma_{\delta v}$) as Gaussian standard deviations. We run 1000 random realisations for the Mixed and Complex fits and 200 for each Individual Cloud fit. The reported fit

Reference	Marker	Larson-like exponent	PCA derived exponent	Structure function		Short description
		Γ	α	v_0	γ	
Kolmogorov (1941)	---	–	–	–	1/3	<ul style="list-style-type: none"> • Kolmogorov’s law for ideal incompressible turbulence.
Kraichnan (1974)	---	–	–	–	1/2	<ul style="list-style-type: none"> • Compressible supersonic shock-dominated turbulence (Burger’s turbulence).
Larson (1979, 1981)	(L)	0.38	–	1.1	0.38	<ul style="list-style-type: none"> • Molecular clouds, clumps, cores, H II regions, mapped by different tracers.
Solomon et al. (1987)	(S)	0.5 ± 0.05	–	1.0 ± 0.1	0.5 ± 0.05	<ul style="list-style-type: none"> • Homogeneous sample of 273 molecular clouds, from ^{12}CO emission.
Heyer & Peter Schloerb (1997)	(1a)	–	0.43 ± 0.04	1.23 ± 0.08	0.24 ± 0.10	<ul style="list-style-type: none"> • ^{12}CO emission of cloud complex Sh 155.
Heyer & Peter Schloerb (1997)	(1b)	–	0.55 ± 0.03	0.78 ± 0.05	0.42 ± 0.09	<ul style="list-style-type: none"> • ^{12}CO emission of cloud complex Sh 235.
Heyer & Brunt (2004)	(H)	–	0.65 ± 0.01	0.87 ± 0.02	0.57 ± 0.07	<ul style="list-style-type: none"> • ^{12}CO emission from 27 molecular clouds.
Heyer et al. (2006)	(3a)	–	0.74 ± 0.04	0.73 ± 0.03	0.71 ± 0.12	<ul style="list-style-type: none"> • Rosette cloud complex as a whole.
Heyer et al. (2006)	(3b)	–	0.79 ± 0.06	1.00 ± 0.04	0.79 ± 0.04	<ul style="list-style-type: none"> • Rosette cloud complex, zone I: inside ionization front, feedback from H II region, using ^{12}CO (similar for ^{13}CO).
Heyer et al. (2006)	(3c)	–	0.66 ± 0.06	0.70 ± 0.03	0.59 ± 0.14	<ul style="list-style-type: none"> • Rosette cloud complex, zone II: outside ionization front.
Bolatto et al. (2008)	(4)	0.60 ± 0.10	–	0.76 ± 0.27*	0.60 ± 0.10	<ul style="list-style-type: none"> • Molecular clouds in extragalactic systems, using ^{12}CO.
Federrath et al. (2010)	—	–	0.66 ± 0.05	–	0.59 ± 0.13	<ul style="list-style-type: none"> • Simulations with purely solenoidal forcing ($\nabla \cdot f = 0$).
Federrath et al. (2010)	—	–	0.76 ± 0.09	–	0.74 ± 0.19	<ul style="list-style-type: none"> • Simulations with purely compressive forcing ($\nabla \times f = 0$).
Klessen & Hennebelle (2010)	(5)	–	–	0.8	0.5	<ul style="list-style-type: none"> • Simulations with continuous accretion of diffuse material.
Roman-Duval et al. (2011)	—	–	0.62 ± 0.20	–	0.53 ± 0.35	<ul style="list-style-type: none"> • Average from 368 molecular clouds, with a resolution limit of 1 FWHM (48''), using ^{13}CO emission.
Bertram et al. (2014)	—	–	0.89 ± 0.05	–	0.93 ± 0.14	<ul style="list-style-type: none"> • Simulated clouds under non-isothermal regime, chemically evolving. ^{12}CO emission, mean density $n = 300 \text{ cm}^{-3}$.
Hacar et al. (2016)	(M)	–	–	0.66*	0.58	<ul style="list-style-type: none"> • Musca cloud as a whole, isolated from stellar sources of feedback, using ^{13}CO and $\text{C}^{18}\text{O}(2-1)$ emission.
Padoan et al. (2017)	(6)	–	–	0.82*,†	0.5 ± 0.1	<ul style="list-style-type: none"> • Simulated clouds with supernovae-driven turbulence.
Traficante et al. (2018b)	—	0.09 ± 0.04	–	–	0.09 ± 0.04	<ul style="list-style-type: none"> • Gravity-dominated regions, core and clump scales.

Table 2. Scaling parameters from Larson-like and PCA derived linewidth-size relationships in previous literature. *Converted into isotropic 3D velocity dispersion. †Extracted manually from their Figure 4 (missing: estimate the uncertainty of this extraction).

consists of the mean scaling parameters (v_0 , α) obtained from the N random fits, and the errors correspond to the standard deviations (σ_{v_0} , σ_α) from the mean values.

These steps are semi-automatically followed by our PCAFACTORY package. The dendrogram and PCA algorithms still need user-defined parameters, based on the input region(s), in order to perform optimally.

6 RESULTS

In this section we present Principal Component Analysis (PCA) of a set of cloud complexes extracted from our Cloud Factory simulation suite using ^{12}CO J=1–0 intensity PPV cubes. We report the effects of changing physical scenario (Section 6.1), line of sight projection (Section 6.2) and evolutionary time (Section 6.3). We also explore how the retrieved kinematic information changes when applying PCA locally on individual molecular clouds or globally on complexes as a whole (Section 6.4). We study 6 cloud complexes as summarised in Table 1, but, as we consider different line of sights

and snapshot times, there are 42 independent configurations in total: 12 for the potential-dominated scenario with no self-gravity and 18 for the self-gravitating case (we extracted 1 snapshot time extra in this case to study the evolution of local gravitational effects in Sec. 6.3), and 12 for the feedback-dominated scenario.

For reasons of space, the corresponding PCA figures and tables from all the cloud complex configurations can be found in the supplementary file provided along with this manuscript⁵. Additionally, the whole catalog of (l , δv) and (v_0 , α) pairs extracted from individual clouds and cloud complexes is also available online⁵. The software developed to carry out this work, called PCAFACTORY, is open source and can be accessed on GitHub⁶. Please refer to this paper if you find these complementary files and/or this software useful for your research results.

⁵ <https://github.com/andizq/andizq.github.io/tree/master/pcafactory-data>

⁶ <https://github.com/andizq/pcafactory>

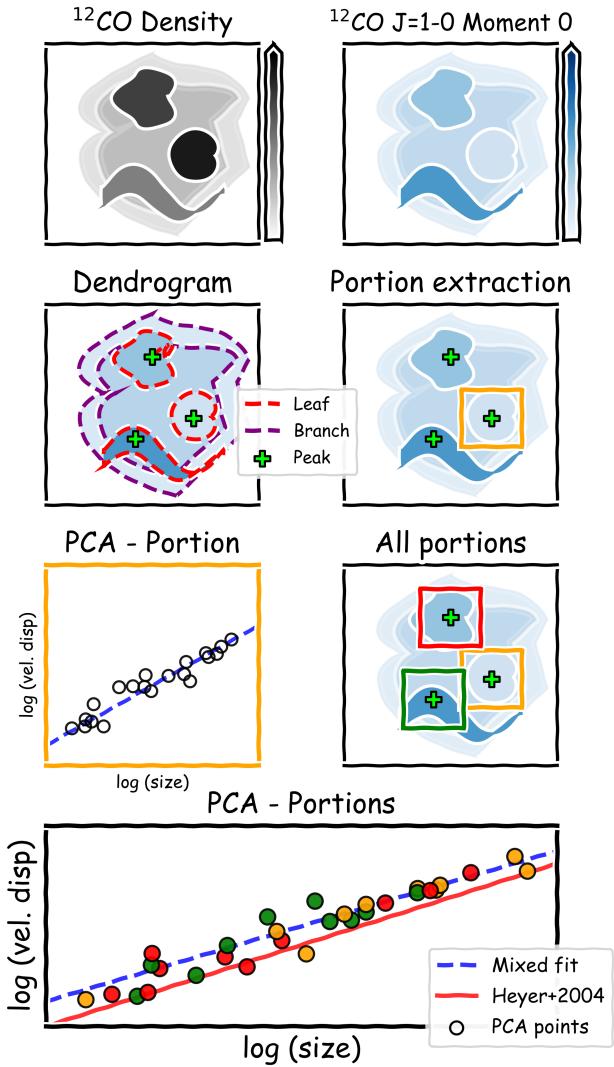


Figure 2. General workflow to retrieve PCA pseudo-structure functions and linewidth-size scale relationships of cloud complexes made up in our Cloud Factory simulation suite. This figure illustrates the Mixed PCA extraction method.

6.1 Structure Function Dependence on the Physical Scenario

In Figure 3 we summarize the structure function scaling parameters (ν_0, γ) derived for all our cloud complexes. The scaling parameters were retrieved using the Mixed PCA extraction method (see Sec. 5), which combines PCA derived scales ($\delta\nu, l$) of inner-complex smaller cloud portions to find the associated pseudo-structure function of each cloud complex. Individual cloud portions are 30 pc wide boxes tracked down on the ^{12}CO $J=1-0$ moment 0 map of the complex (see Sec. 5). This size is typical of Giant Molecular Clouds (GMCs), so we will refer to individual portions as cloud portions, or just clouds, from here on. We find N molecular clouds within our cloud complexes. The translation from PCA derived scaling exponents (α) into the corresponding structure function exponents (γ) assumes the 2nd order calibrations given by eqs. 5 and 9. The (ν_0, γ) pairs are computed on different orientations and time snapshots for each cloud complex. For comparison, we include representative observational and synthetic values reported in previous literature as

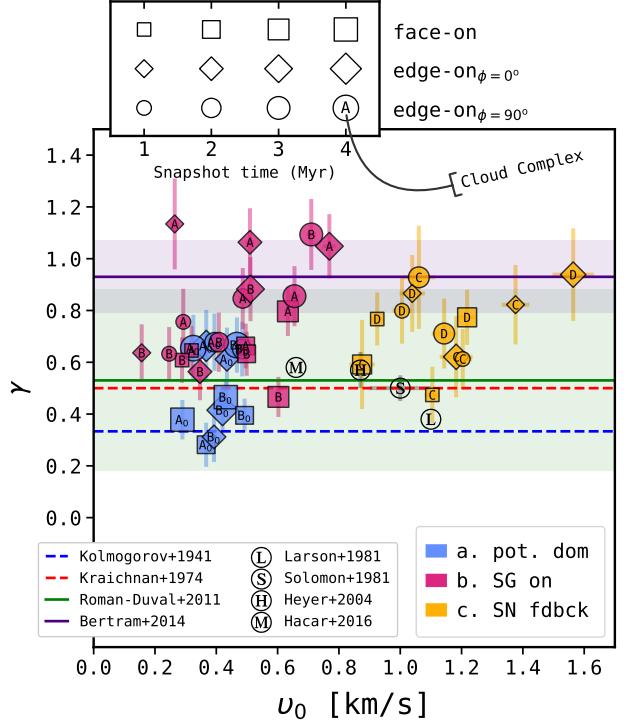


Figure 3. Structure function parameters (ν_0, γ) of cloud complexes extracted from different physical scenarios (marker colour), orientations (marker style) and snapshot times after tracer refinement has commenced (marker size). The parameters are derived from PCA pseudo-structure functions using the Mixed method with individual clouds of cloud complexes (see Sec. 5 and Fig. 2). For comparison, we include previous values reported in literature as transparent markers or horizontal lines with their associated errors as shaded regions. Find the legend and marker code of literature values in Table 2.

well as theoretical regimes for subsonic and supersonic turbulence (see Table 2).

The most obvious aspect is that the scaling parameters agglomerate in different zones according to the physical mechanisms governing the cloud complexes. Note however that overall, the structure function scaling exponents (γ) are similar regardless of the physical scenario and lie mostly around the range of values obtained by RD11 and Bertram et al. (2014), suggesting that the structure exponent γ alone is not sufficient to fully determine the nature of cloud complexes.

Figure 4 displays the PCA applied on cloud complexes B₀, B and D. The coloured circles in the left panel are the PCA derived scales ($\delta\nu, l$) from selected cloud portions correspondingly illustrated as coloured squares in the right panel. The crosses are the PCA derived scales from the cloud complex as a whole. The pie chart in the left panel shows the number of PCA scales extracted from each cloud portion and the total net scales in the middle. Some cloud portions are subject to rejection if their centres overlap with other cloud centres or if there are no scales retrieved by the PCA algorithm on them. This is shown by the coloured round boxes at the bottom of the left panel, which also specify, if any, the number of PCA scales found in the rejected regions. The dashed blue line is the associated fit to the combined cloud portion points (Mixed method) and the dashed green line is the fit to the points derived from the cloud complex as a whole (Complex method). The blue and green shaded regions are the uncertainties of these fits (see

details in Sec. 5). We also find the associated fit to individual cloud portions (Cloud method) but only present the corresponding shaded regions in grey. For comparison we show the best fit (red solid line) found by Heyer & Brunt (2004) using PCA derived scales from 27 molecular clouds. The right panel background is the ^{12}CO J=1–0 moment 0 map associated to each cloud complex.

All the cloud complexes are viewed from a face-on projection and have similar evolutionary times so that nearly only the effects of varying physical mechanisms are manifest. The first two rows of the figure are associated to the same complex but one is dominated by the large-scale galactic potential on its own (complex B_0) and the other has self-gravity between gas particles (complex B), both extracted 2 Myr after turning on tracer-particle refinement. As expected, the number of cloud portions detected by our algorithm is considerably larger in complex B_0 due to the absence of local gravitational forces. This fact allows the gas to spread out on the field of view. Consequently, both the ^{12}CO density and line intensity with respect to complex B are reduced as can be noticed in the moment 0 maps on the right-hand side of the figure. For complex D, in the feedback dominated case, the number of cloud portions is even lower. This is consistent because in this scenario, the ^{12}CO density and its spatial extension are drastically lowered down by molecular dissociation processes undergone during energy/momentum injections from nearby supernovae tied to star formation sites.

On the other hand, the number of PCA points is more sensitive to higher moment 0 values than to larger ^{12}CO flux extents as one might think in first instance. This is because moment 0 is related to intensity variance along the velocity axis in the sense that a high-valued pixel in moment 0 is likely to span across quite a wide range of values in the covariance matrix representation of the line cube, especially when density fluctuations and velocity intermittence in the region are low. Thus, it is also equivalent to the amount of information that principal components can keep and hence the PCA algorithm will raise more ($l, \delta v$) pairs. (Check this statement and the mineigval parameters used for the clouds (percentage of variance kept) as the number of PCA points is highly dependent on it. If they are too different I should remove this paragraph).

PCA spotlights clear variations on the retrieved structure functions from different physical mechanisms. From cloud complexes B_0 , B and D analysed in Figure 4, the scaling exponent (α , and so it does γ) steadily increases as a function of the amount of active physics in the cloud complex. For the potential-dominated scenario, the case with no self-gravity (B_0) yields a structure function scaling exponent $\gamma = 0.39 \pm 0.06$ ($\alpha = 0.53 \pm 0.01$). Turning on gas self-gravity (B) leads to a higher $\gamma = 0.63 \pm 0.09$ ($\alpha = 0.69 \pm 0.02$), whereas the mixed supernovae feedback case (D) results in the steepest $\gamma = 0.77 \pm 0.10$ ($\alpha = 0.78 \pm 0.03$). Nevertheless, based on Figure 3, this trend is only clear when comparing cloud complexes in the potential-dominated cases (with and without self-gravity). In the feedback-dominated case, the scaling exponents lie in the middle of the other two scenarios in overall.

PCA derived velocity-scaling coefficients (v_0) respond a bit differently. This time, there is a gap between values noticeable only for feedback-dominated cloud complexes. In the potential-dominated cases, self-gravity shifts complexes to slightly higher v_0 with respect to non self-gravitating complexes when studied altogether, but this is a small effect. From Figure 4 we do not find any sensible distinction in the particular case of complexes B_0 and B at $t = 2$ Myr. At this time snapshot, we get the structure function scaling coefficients $v_0 = 0.49 \pm 0.00$, $v_0 = 0.50 \pm 0.01$ and $v_0 = 0.92 \pm 0.02 \text{ km s}^{-1}$ for B_0 , B and D. Interestingly though, in Section 6.3 we track the evolution of structure function parameters

and find that the scaling coefficients increase steadily in time due to local and sustained gravitational forces. This does not occur in cloud complexes without self-gravity.

We summarise in Table 3 the mean and standard deviations of scaling parameters derived from all our cloud complex configurations, using our 3 PCA extraction methods detailed in Section 5.

The fact that velocity scaling coefficients are generally low for complexes A and B, at all times, reflects a high level of velocity coherence due to lower stellar feedback affecting these regions. This favours the development of long filamentary structures present in both complexes (stretched out by differential rotation of the galaxy), their preservation over time, and a sustained emergence of stellar systems represented by sink particles. This agrees with the less effective feedback provided by randomly distributed supernovae, and also with the low velocity gradients in complex B and in long filaments of complex A reported in our Paper I of the Cloud Factory series.

Conversely, the strong mixed supernova feedback plays a significant role in taking cloud complexes out from that coherent state and producing more realistic objects. As previously found in our Paper I, clustered feedback tied to star formation sites makes complexes C and D to have shorter and less massive filaments. This type of feedback also induces stronger shears that reduce filaments lifetime scale and in consequence their formation efficiency of stellar systems. In this work this is manifest in the high scaling coefficients found for the feedback-dominated scenario, which indicates a larger degree of velocity fluctuations in complexes C and D. This is again in good agreement with the physical mechanisms we know are governing the regions and suggests that the velocity-scaling coefficient of the structure function can provide an excellent representation of the nature of the cold dense molecular ISM.

6.2 Structure Function: Cloud Orientation Effects

We vary the line of sight direction of cloud complexes during the radiative transfer calculations to explore any influence on the retrieved structure functions. As shown by the different symbols in Figure 3, there are no systematic variations on scaling parameters when changing the orientation of cloud complexes; they rather depend on the particular geometry of each complex and hence on their time evolution itself.

Edge-on projections of cloud complexes with low scale-heights exhibit flat and continuous CO distributions but actually consist of both nearby and distant structures that conform the cloud complex as a whole (see Fig. 1). Thus, background emission is usually shielded by foreground gas in these orientations, especially if optically thick tracers are used as in our case with ^{12}CO . This makes the retrieved structure function representative only for a smaller fraction of the gas.

Projection effects for A_0 and B_0 , which are our least realistic cases, seem to be driven by column density variations. From Figure 3, complex A_0 yields comparable scaling exponents both for the edge-on ($\phi = 0^\circ$) and edge-on ($\phi = 90^\circ$) orientations ($\gamma \approx 0.6$ at any time) but lower values for the face-on views ($\gamma \approx 0.3$). The column densities are very high when this complex is viewed edge-on (either $\phi = 0^\circ$ or $\phi = 180^\circ$) compared to the face-on view ($\approx 3\times$ higher, see Fig. 1). This intensifies for larger scales as velocity fluctuations get contribution from further gas mass as a function of the enclosed area, although this effect should be attenuated on high optical depth regions ($\tau >> 1$).

Cloud complex B_0 does not follow the same pattern. In this

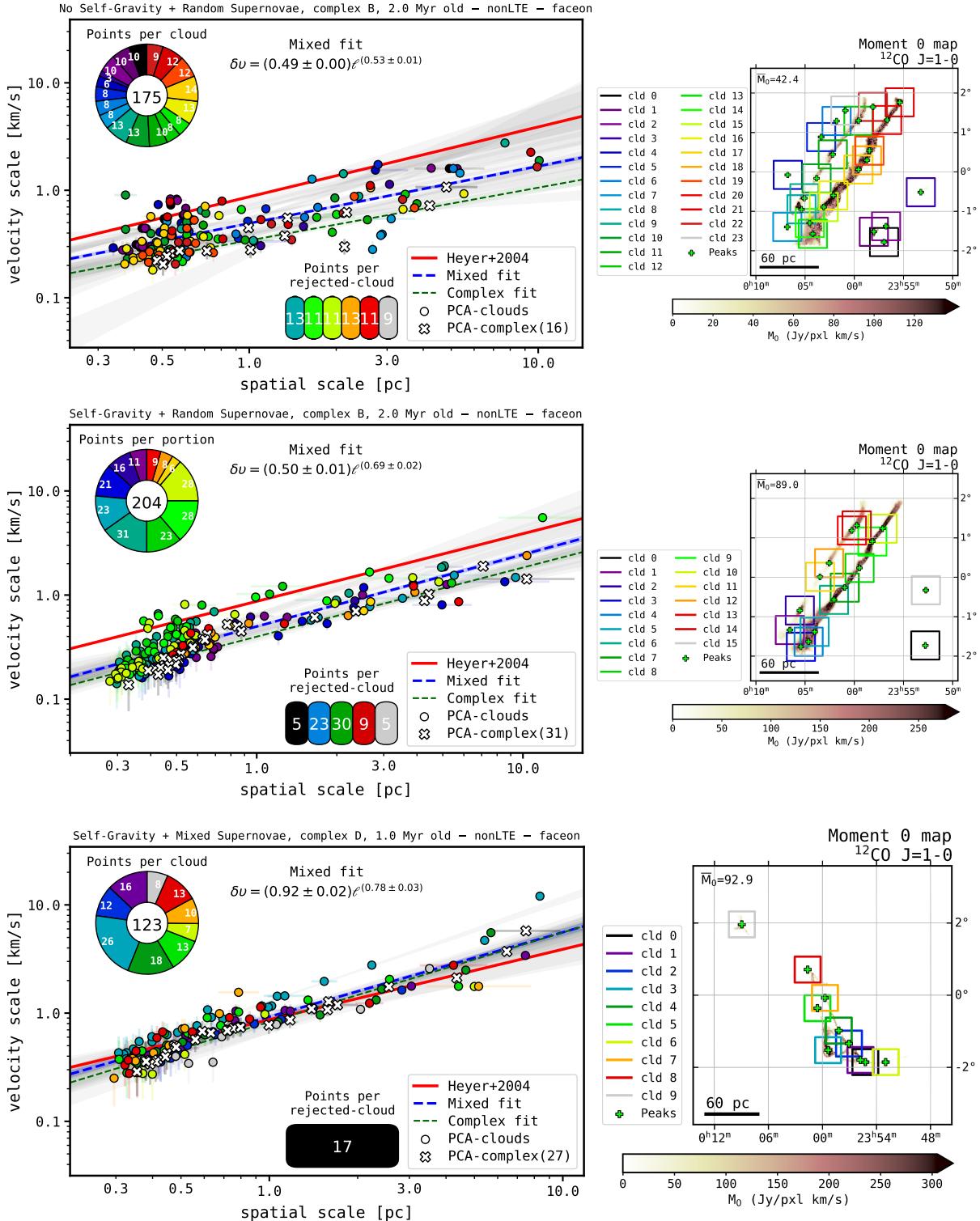


Figure 4. *Left column:* Principal Component Analysis from the face-on view of cloud complexes B₀ (top), B (middle) and D (bottom), each representing different physical scenarios after the tracer refinement has commenced. The coloured circles are the PCA derived scales ($\ell, \delta v$) from the corresponding coloured portions on the right panel. The white crosses are the PCA points from the cloud complex as a whole. The pie chart indicates the amount of points extracted from each portion and the total number of points on its centre. The number of cloud points is shown in the lower right corner of the plot. For reference, the red line is the PCA best fit found by Heyer & Brunt (2004) from an ensemble of 27 molecular clouds. The blue and green lines are, respectively, the fit from considering all the portion-derived points simultaneously (Mixed fit) and the fit from the Complex points (Complex fit). A shaded region around a fit represents the standard deviation from the mean fit computed for N random runs. For the single portions we only show the fit shaded regions (in grey) to avoid a confusing net of lines. *Right column:* velocity-integrated intensity or moment 0 (M_0) map of the ^{12}CO J=1-0 synthetic cube of the cloud complex computed by LIME in LTE mode. The upper limit of the colorbar is $\bar{M}_0 + 2\sigma$. The plus markers are the M_0 peak values in dendrogram leaves of the map and the 30×30 pc² coloured regions are the individual cloud portions selected for the PCA study.

PCA Mode	Physical Scenario	Complex	\bar{v}_0	$\bar{\gamma}$
Mixed clouds	potential-dominated	$A_0 \cup B_0$	0.41 ± 0.06	0.51 ± 0.15
	pot-dom+self-gravity	$A \cup B$	0.46 ± 0.17	0.78 ± 0.19
	feedback-dominated	$C \cup D$	1.14 ± 0.18	0.74 ± 0.14
Entire Complex	potential-dominated	$A_0 \cup B_0$	0.31 ± 0.05	0.48 ± 0.15
	pot-dom+self-gravity	$A \cup B$	0.36 ± 0.14	0.74 ± 0.19
	feedback-dominated	$C \cup D$	1.13 ± 0.20	1.02 ± 0.18
Individual clouds	potential-dominated	A_0	0.36 ± 0.09	0.64 ± 0.26
	potential-dominated	B_0	0.48 ± 0.11	0.64 ± 0.31
	pot-dom+self-gravity	A	0.49 ± 0.20	1.01 ± 0.26
	pot-dom+self-gravity	B	0.45 ± 0.22	0.82 ± 0.35
	feedback-dominated	C	1.37 ± 0.55	0.84 ± 0.48
	feedback-dominated	D	1.20 ± 0.46	0.85 ± 0.29

Table 3. Mean and standard deviations of structure function parameters derived from different PCA extraction methods. (Include column of number of objects N)

case, variations induced by line of sight projections are smaller. This is in well agreement with our previous interpretation because gas column densities are this time nearly isotropic, with values $< 2\times$ between face-on and edge-on column densities.

When compared to PCA applied over cloud complexes as a whole, the variability of complex B_0 decreases even more (see Figure 5). For B_0 , all the scaling parameters lie around $v_0 = 0.35$ km s $^{-1}$ and $\gamma = 0.25$ regardless of the orientation. This suggests that variations of scaling parameters in this complex are due to individual cloud fluctuations driven either by localized collisions or random supernovae feedback rather than to any generalized feature of the complex itself. Oppositely, cloud complex A_0 does maintain the same behaviour when studied as a whole, namely, the face-on orientations still yield the lowest scaling exponents at all times and the gap between values is roughly the same. Both results strengthen our argument that larger column densities yield steeper scaling exponents as long as they do not get optically thick too fast, and equivalently, isotropic column densities result in comparable scaling parameters. We further explore differences from changing PCA extraction methods in Section 6.4.

Complexes A & B show similar signs but they are convolved with time evolution effects of local gravitational forces. For this reason we prefer to analyze them in terms of the evolution over time in Section 6.3 where the outcome is clearly systematic.

For feedback-dominated complexes, C and D, the edge-on $_{\phi=0^\circ}$ orientation (parallel to the y-axis) produces the highest and steepest scaling parameters in most of the snapshot times. This is because for both, the axes of long filamentary structures are preferentially stretched out along the y-direction due to differential rotation of the galaxy. This makes the edge-on $_{\phi=0^\circ}$ projection contain more gas mass in few analysis regions than the other projections. Additionally, the scale-height of these regions is naturally high due to nearby (in C) and internal (in D) energy feedback from supernova explosions, which enables more gas to contribute to the retrieved velocity fluctuations field as optical depth effects are softened.

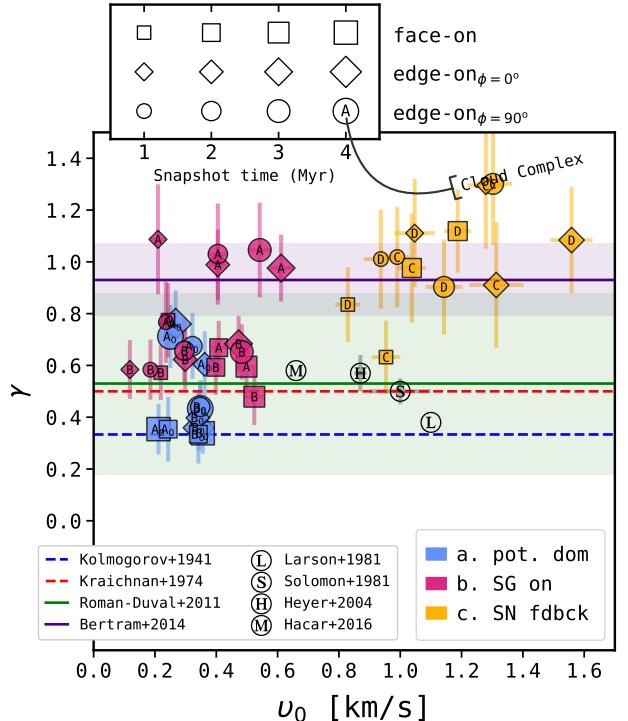


Figure 5. Same as Figure 3 but this time, the structure function parameters (v_0, γ) are derived from PCA applied to cloud complexes as a whole (i.e. using the Complex method, see Sec. 5.)

6.3 Structure Function: Time Evolution

Figure 6 connects structure function parameters of subsequent cloud complexes in time from each physical scenario and orientation. We compare the Mixed and Complex PCA extraction methods but do not find any major difference between them.

Potential-dominated complexes with no self-gravity have a sys-

tematic reduction of velocity scaling coefficients (v_0), while the scaling exponents (γ) raise over time. This suggests a sustained lack of turbulence fueling sources (lowering scaling coefficients) and exhibits the natural accelerated energy decay over length scales of Kolmogorov-like fluids (see e.g. Onsager 1949), which, in small scales, rapidly lose energy via viscous dissipation (scaling exponents increase). This agrees with the retrieved scaling exponents for this scenario; most of them lie on, or close to the Kolmogorov's $\gamma = 1/3$. Evidently, random supernovae remain insufficient to fulfill the role of turbulence fueling source.

Turning on self-gravity completely breaks that tendency. In complexes A and B, the scaling coefficient (or magnitude of velocity fluctuations, v_0) increases with time. The gradient of this increment appears linear and is similar for both complexes. This allows us to capture the time evolution of mean scaling coefficients with the following relations:

$$\bar{v}_0(t) = 0.10 + 0.18 \left(\frac{t}{\text{Myr}} \right) [\text{km s}^{-1}] \quad (12)$$

$$\bar{v}_0(t) = 0.05 + 0.16 \left(\frac{t}{\text{Myr}} \right) [\text{km s}^{-1}], \quad (13)$$

where eqs. 11 and 5 hold for the Mixed and Complex PCA extraction methods respectively (see Sec. 5).

The evolution of the scaling exponents is chaotic in this scenario, suggesting that local gravitational forces can provide velocity fluctuations across a wide range of spatial scales depending on the evolutionary stage of the clouds and their response to this field. This randomness also hints at differences in the evolution phase of (smaller) internal structures in the complexes. We zoom into cloud complex B to investigate time evolution of internal structures from a gravitational point of view.

In Figure 7 we track the evolution of individual filaments in complex B. At $t = 1$ Myr, when there are no sink particles yet, individual cloud parameters are all around a common zone in the (v_0, α) space. In overall, scaling exponents are steeper and velocity fluctuations quieter (provide numbers) compared to those of the potential-dominated case with no self-gravity (provide numbers). We propose that this is a consequence of global (large-scale) collapse processes that reduce the degree of turbulence in the region.

Later, at $t = 2$ Myr, there is a burst of sink particles in the long dense filament ($N_{\text{sinks}} > 100$) but none in the diffuse filament yet. This produces a rupture of individual cloud parameters from both filaments in two clear subgroups as shown in the middle panel of Fig. 7. The PCA scaling exponents of the shorter filament are this time shallower because local gravitational fragmentation has commenced. This favours smaller scales given by the Jeans length of the medium and is likely to be a feature of pre-core stages in molecular clouds. The long dense filament jumped up this phase as it formed several individual clumps/cores much faster. Such a rapid emergence of stellar systems drastically lowers down velocity fluctuations in small scales and consequently raises up scaling exponents (γ). This is a consequence of our sink implementation that reduces the surrounding (bound) gas mass of the core to point-mass gravitating particles with no information of gas velocities. We emphasize that in real observations, especially where cores are not resolved, high scaling exponents should still hold in this stage but the magnitude of turbulence (v_0) must be even higher because gas from cores/clumps do actually contribute to the magnitude of velocity fluctuations in the region. Hence, this 'burst' of new cores establishes multiple point-like centres of collapse affecting inter-core gas in medium and large scales predominantly. The zone in the

long filament where sink particles are more numerous has a mean sink separation of ≈ 0.5 pc. Note that both fragmentation processes in pre-core stages, as well as the subsequent formation of cores, continuously increment the magnitude of velocity fluctuations (v_0) across the cloud complex.

Finally, at $t = 3$ Myr, the short filament commence the formation of cores at its upper tip. For that particular cloud (cld 13 at $t = 2$, cld 12 at $t = 3$ Myr), both scaling parameters increase and tend to the zone occupied formerly by clouds from the long filament but not as sharply as them because of the much lower core formation rates. Cloud portions with no cores remain with the same γ but keep raising v_0 up as expected because fragmentation is still carrying on. The long filament proceeds increasing both parameters as the number of cores grows over time. Interestingly, new molecular associations appear on the right side of the complex and yield scaling parameters close around the same zone where clouds from both of the filaments started at $t = 1$ Myr, which strengthen the idea that quiescent clouds have gravity-driven time-dependant trajectories in the (v_0, α) space. This behaviour is also hold for cloud complex A, where similarly there are bursts of cores and quiescent zones throughout the region.

Note that scaling parameters retrieved from the Mixed and Complex PCA extraction methods follow the averaged values between both filaments at $t = 2$ Myr, but describe preferentially the left filament at $t = 3$ Myr. This is a consequence of intermittent velocity fields between both filaments and the fact that the left filament might be better represented by first-order velocity structure functions, dominating the PCA derived properties of the complex (Brunt 2003a).

Feedback-dominated cloud complexes evolve accordingly to their particular environment. As seen in Figure 6, the common pattern is that complex C migrates toward lower scaling coefficients v_0 whereas complex D moves toward higher v_0 . In complex C, external feedback from supernovae explosions set large turbulence driving scales that decay over time through smaller scales as there are no internal fueling sources of turbulence. Local gravitational effects do not look to play the same role as in complexes A & B because complex C is much more diffuse and so mass instabilities are more difficult to reach. In complex D, however, energy re-injection from internal supernovae seems to sustain large turbulence driving scales by competing against local collapse through expanding supernova bubbles. In addition, internal clustered feedback triggers up the magnitude of velocity fluctuations in time, which results in the region being destroyed considerably faster than complex C.

6.4 Structure Function: Inner Cloud Complex Variability

We explore three PCA extraction methods to address any variations on scaling parameters that might arise when different analysis windows are used to retrieve pseudo-structure functions. As a reminder, the three approaches are the Mixed, the Complex and the Individual Cloud methods. The Mixed method consists in combining PCA derived scales from cloud portions to construct the structure function of the original complex; the Complex method applies the PCA algorithm on whole complexes, without sub-portions; and the Cloud method extracts structure functions from cloud portions as if they were individual objects. See further details of the methods in Section 5.

Figure 8 shows the distribution of structure function scaling parameters derived from each of the PCA extraction methods and physical scenarios, and Table 3 summarises the mean values and standard deviations. The Mixed and the Complex methods yield very

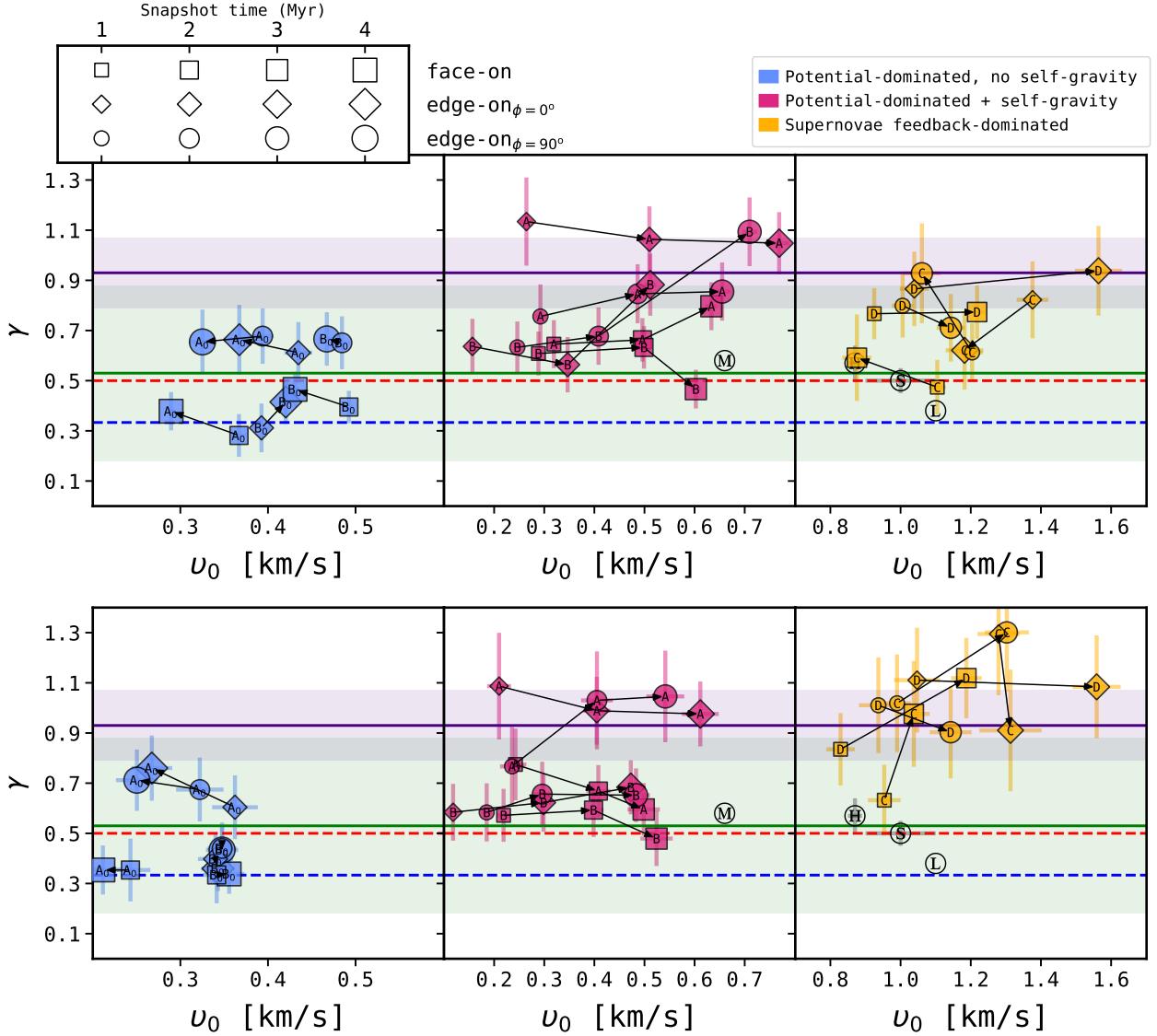


Figure 6. Figures 3 (top row) and 5 (bottom row) split by physical scenario to illustrate variations of structure function parameters over time. Arrows join the sequence in time of cloud complexes from equal orientations. The plots share the same axis scale for γ but not for v_0 , whose axes were stretched or shortened to facilitate the analysis. Same literature values are kept here for reference, see marker code and description in Table 2.

similar parameter scattering (represented by the size of concentric ellipses) for all the physical scenarios. However, studying cloud complexes as a whole with the Complex method shifts the potential-dominated cases (A_0 , B_0 , A & B) to lower scaling coefficients v_0 ($\Delta \sim 0.1 \text{ km s}^{-1}$) and the feedback-dominated cases (C, D) to much higher scaling exponents γ ($\Delta \sim 0.22$) compared to the Mixed method. We attribute these variations to intermittency of density and velocity fields that make the PCA scaling exponents not follow 2nd order velocity fluctuations but lower orders only (see Brunt et al. 2003; Roman-Duval et al. 2011). Similarly, the difference in scaling coefficients is due to the PCA derived scales prioritizing regions of low-order velocity fluctuations within the Complex (see e.g. Fig. 4). This is a more prominent effect in the Complex method where the PCA is computed from entire cloud complexes and hence intermittent fields are more likely to appear. Especially in the feedback-dominated cases where fluctuations in the turbulent flow are common due to localized and nearby supernovae.

Figure 8 also presents the scaling parameter distribution for the Individual Cloud method, which computes the structure function of individual molecular clouds and treat them as individual objects. Due to the much higher number of objects, the parameter scattering is naturally larger for all the scenarios compared to the Mixed and Complex methods. However, the feedback-dominated scenario produces the widest range of scaling parameters as a consequence of the variation in localized internal and external supernovae feedback plus the local gravitational influence. Figure 9 splits individual cloud distributions depending on the origin cloud complex to show variations driven by the surrounding environment. Parameter changes due to different density and energy contexts are evident. The potential-dominated cases inside spiral-arms, A₀ & A, generally yield steeper scaling exponents than the more diffuse inner-arm regions B₀ & B. For the feedback-dominated scenario both cloud complexes produce similar scaling parameters, however, complex D spans a smaller range of parameters from its individual clouds. This

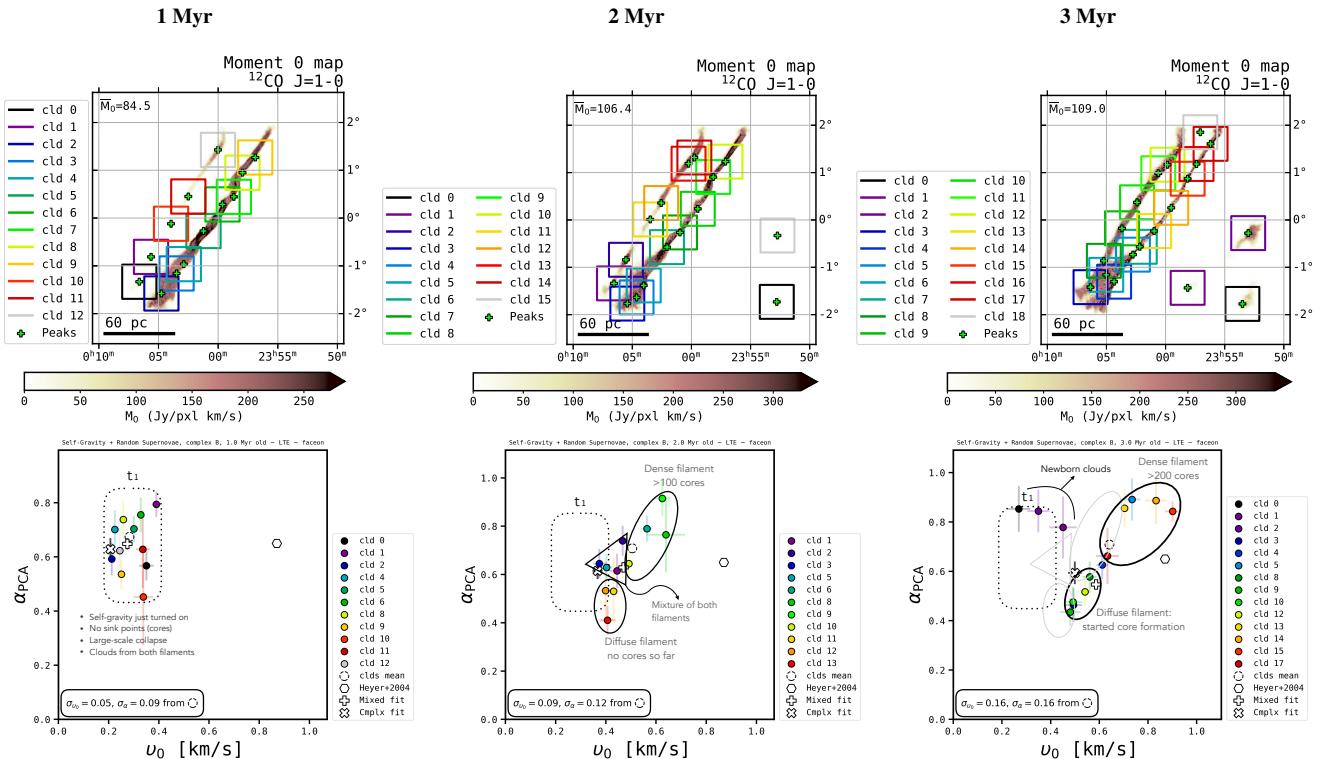


Figure 7. Time evolution of cloud complex B due to local gravitational forces. The top row shows moment 0 maps at different times since the beginning of tracer refinement. The bottom row panels display the PCA derived pseudo-structure function parameters from individual clouds of the corresponding complex on the top. The standard deviation of cloud parameters is shown in the lower left corner of panels. The colour code of individual clouds in top panels matches the marker colours in bottom panels. We illustrate how individual clouds migrate over time in the (v_0 , α) space as a correlation with the evolution of governing gravitational processes across different structures of the region.

is due to the new embedded supernova explosions that destroy cloud structures faster in complex D compared to the quieter complex C.

As expected, due to the strong supernovae field set in our feedback-dominated scenario, the mean scaling parameters lie close to values resulting from hypersonic turbulence samples such as the ionised zone of the Rosetta cloud reported by Heyer et al. (2006) (marker ⑩ in Figs. 8 and 9) and the Federrath et al. (2010) simulations using purely compressive forces (solid blue line —).

7 DISCUSSION

7.1 Comparison to observational and synthetic structure functions

From Figure 8, the potential-dominated complexes without self-gravity are mostly clustered between the (incompressible) subsonic (Kolmogorov 1941; Onsager 1949) and supersonic shock-dominated turbulence (Kraichnan 1974; Frisch et al. 2001) regimes, or just above the latter. This suggests that the large-scale gravitational forces along with random (isolated) supernovae feedback are able to reproduce classical scaling exponents (γ) from theoretical fluids. They also agree with observational exponents found by Larson (1981) and Solomon et al. (1987), and with simulations of pure solenoidal turbulence forcing performed by (Federrath et al. 2010). The solenoidal turbulence is associated to regions of low star formation activity like the quiescent zone of the Rosette cloud complex (Heyer et al. 2006) or the Musca cloud (Hacar et al. 2016). However,

the magnitude of velocity fluctuations (v_0) remains too low in this scenario to fully reproduce those observations.

Turning on self-gravity leads in general to steeper structure function scaling exponents (γ). Depending on the evolutionary stage of cloud complexes, gravity-driven (non-thermal) motions such like collapse and/or accretion flows can increase the velocity dispersion. This results in a sustained increment of the magnitude of velocity fluctuations (v_0), triggered by the emergence of local potential wells in the region. In pre-core stages this occurs on small scales given by the core accretion zone, and in core stages on medium/large scales given by the separation between cores (Smith et al. 2016). Complex B is an excellent sample to study this as it is made of two quiescent filamentary structures that evolve differently. At early stages, one filament is dense and prominent and the other is diffuse and discontinuous. Figure 7 shows that the PCA derived parameters from individual clouds of this complex are clearly different depending on the analysed filament. Both v_0 and α are in general higher on the long-dense filament, which agrees very well with the idea that massive regions yield higher velocity dispersion for larger column densities due to bound clumps and cores within them undergoing gravitational collapse (Ballesteros-Paredes et al. 2011).

It is particularly interesting to compare the results from this physical scenario with real Giant Molecular Clouds that have little stellar feedback. This is the case for the Musca cloud in the Musca-Chamaeleonis molecular complex reported by Hacar et al. (2016) as the first observational evidence of a filament that is largely coherent with negligible internal turbulence. They found two 1-D observational relations of the form $\delta v_{1D} = v_0 l^\gamma$ that depend on the scale

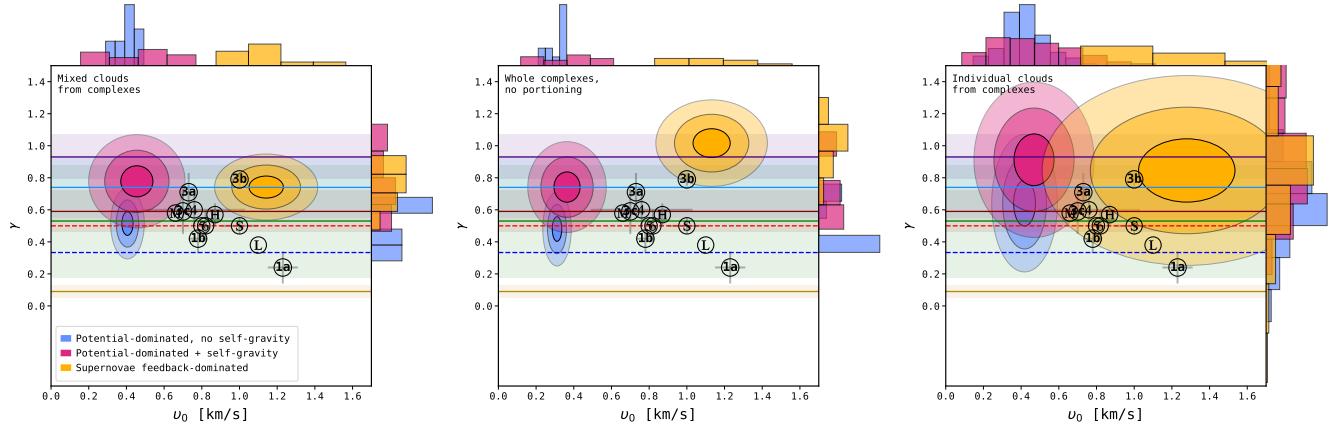


Figure 8. Distribution of structure function parameters depending on the PCA extraction method. The size of concentric ellipses correspond to 1,2 and 3 standard deviations from the mean parameter values for each physical scenario. The first two panels represent the Mixed and Complex methods (shown as markers in Figs. 3 and 5), and the third panel corresponds to the scaling parameters of individual clouds from cloud complexes. We also plot histograms on the top and right axes of the panels to show localized fluctuations. The histograms are normalized to enclose a unit area. Individual clouds are more numerous and exhibit Gaussian-like distributions along both directions, so we fit and overlay Gaussian PDFs in that case. For comparison, we show the full set of literature values listed in Table 2. Mean parameters and standard deviations obtained here are summarized in Table 3.

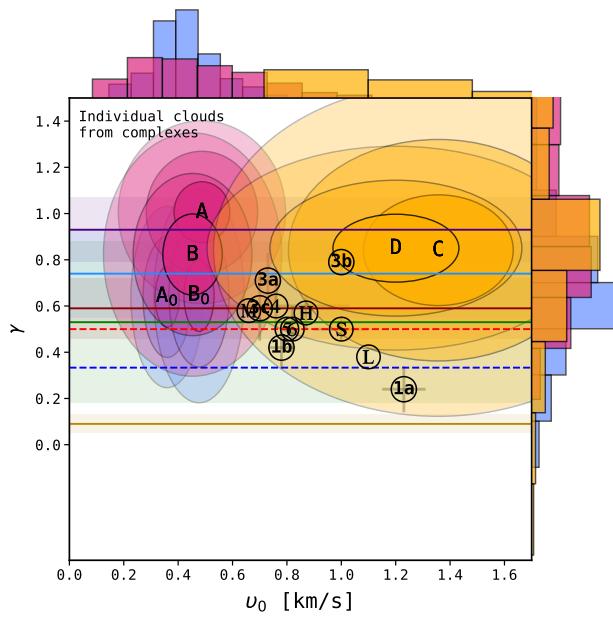


Figure 9. Same as the third panel of Figure 8 but discriminating individual clouds according to their original Complex.

size range of the cloud: (i) a transonic $\delta v = 0.55l^{0.25}$ for scales < 1.0 pc and (ii) a supersonic $\delta v = 0.66l^{0.58}$ in scales between 1.0 and 3 pc. To ease the comparison to our 3-dimensional parameters, we intentionally removed the 1D subscript as we are assuming isotropic 3-dimensional velocity dispersion and hence multiplying v_0 by a $\sqrt{3}$ factor.

Our self-gravitating cloud complexes without strong clustered feedback are consistent with the scaling parameters derived from this quiescent region in the supersonic regime, especially the most evolved complexes (see Fig. 6). Additionally, smaller individual cloud portions where there are few cores manage

to reproduce the Musca's parameters in the transonic regime as is the case with the diffuse filament in Figure 7. Thus, we attribute the power-law break noticed in the Musca cloud to localised evolution of parameters driven by gravitational effects (see the clusters of parameters on different (v_0 , α) enclosed by solid line ellipses in Fig. 7).

Furthermore, Hacar et al. (2016) also found that the separate analysis of individual portions in the Musca cloud leads to a wide range of scaling parameters, suggesting the presence of local fluctuations that can substantially differ from the structure function of the cloud as a whole. This resembles our result that individual cloud parameters exhibit a high degree of scattering when compared to the parent larger-scale cloud complex parameters. We found that this level of scattering is related to several gravitational stages governing different scales of the cloud/complex, which at the same time, is related to the magnitude of turbulence and density distribution of the region.

Structure function scaling exponents (γ) from our feedback-dominated cloud complexes are consistent with Federrath et al. (2010) simulations with purely compressive forces or with the value reported by Bertram et al. (2014) for ^{12}CO emission of molecular clouds with artificial turbulent fields. When considering scaling coefficients (v_0), some of our cloud complexes can reproduce observations from Larson (1981); Solomon et al. (1987); Heyer & Brunt (2004) (see Fig. 3) but in general most of them are around the point 3c (in Figs. 8 and 9) from the zone II of the Rosette cloud complex, which is dominated by strong stellar feedback from nearby massive stars (Heyer et al. 2006). This is compatible with the burst of supernova explosions that provide strong energy feedback to our simulated regions. Also, this indicates that structure functions from other typical clouds (Heyer & Peter Schloerb 1997; Bolatto et al. 2008) and simulations (Klessen & Hennebelle 2010; Padoan et al. 2017) intermediate between our scenarios in Figs. 8 and 9 require feedback but not strong bursts of it.

8 CONCLUSIONS

We have performed principal component analysis (PCA) of full radiative transfer simulations of synthetic molecular cloud complexes from galactic initial conditions. The cloud complexes are self-consistently generated using our Cloud Factory galactic-scale ISM simulation suite. We explore PCA-derived velocity structure functions from 3 different physical scenarios set up in our Cloud Factory: a) one where the ISM dynamics is dominated by the large-scale galactic potential, with (isolated) supernovae explosions randomly distributed across the galaxy, b) same as the previous case but self-gravity is turned on, and c) a feedback-dominated scenario where supernova explosions are random but also tied to star formation sites, which results in strong clustered stellar feedback. Large-scale potential and local gravitational effects are both active in this case.

Regardless of the physical scenario, we find that 40 out of the 42 cloud complexes analysed from our Cloud Factory zooms agree with turbulent scaling exponents reported in the literature (approx. 0.2-1.0). Clearly, large scale gravitational forces alone when combined with turbulent decay are enough to explain this behaviour. Under weak influence from (isolated) supernova explosions, local gravitational forces can make structure functions of molecular clouds evolve over time and reproduce observations of quiescent molecular associations.

We report time-dependant trajectories in the structure function parameter space driven by local gravitational effects and supersonic turbulent flows. The magnitude of velocity fluctuations (v_0) increases steadily for self-gravitating regions with low stellar feedback. Typically, just assembled clouds evidence low magnitudes and then migrate over the parameter space following the mentioned trajectories. The structure function scaling exponents (γ) are globally more chaotic because depend on the stage of gravitational collapse, which varies locally as a function of the gas conditions. This can explain the power-law break found by Hacar et al. (2016) in the structure function of a quiescent molecular cloud for different scale-size regimes.

However, gravitational forces alone (when combined with random supernovae feedback) are not enough to reproduce both the normalisation and the scaling exponent of molecular clouds with active star formation. We find that clustered supernovae feedback tied to sites of star formation is key to self-consistently make clouds that reproduce both the scaling exponent and coefficient of most of the linewidth-size relationships extracted from observations with similar scale sizes and resolutions to those used in our simulations. These simulations produce self-consistently generated clouds that can be used in future for studies of clustered star formation in a galactic context.

On the other hand, intermittency effects are manifest within our cloud complexes. This is particularly clear for regions dominated by strong feedback from clustered supernovae, where structure functions can change considerably depending on the analysis scales.

Also, there are clear variations in the PCA-derived structure functions when different environmental conditions are tested. Our cloud complexes can be either immersed in dense (spiral-arms; complexes A₀ and A) or diffuse (inner-arm; complexes B₀ and B) regions of the Galaxy, and get feedback influence from inner+outer (complex D) or from only outer supernova explosions (complex C). We find that self-gravitating regions inside a denser context yield higher scaling exponents overall, and that the influence from strong embedded supernova explosions reduces the parameters dispersion (due to faster disruption of clouds) compared to cloud complexes with no embedded supernovae.

Our results suggest that our PCA-based statistical study is a formidable method to diagnostic local and global surrounding conditions, evolutionary stages and physical mechanisms governing (real and synthetic) molecular clouds. The analysis tools developed in this work are all condensed in our new open source PCAFACTORY package.

ACKNOWLEDGEMENTS

AFI acknowledges the studentship funded by the UK's Science and Technology Facilities Council (STFC) through the Radio Astronomy for Development in the Americas (RADA) project, grant number ST/R001944/1. RJS gratefully acknowledges support from an STFC Ernest Rutherford Fellowship (grant ST/N00485X/1), without which this work would not have been possible.

REFERENCES

- Anselmet F., Gagne Y., Hopfinger E. J., Antonia R. A., 1984, *Journal of Fluid Mechanics*, **140**, 63
- Ballesteros-Paredes J., Hartmann L. W., Vázquez-Semadeni E., Heitsch F., Zamora-Avilés M. A., 2011, *MNRAS*, **411**, 65
- Bate M. R., Bonnell I. A., Price N. M., 1995, *MNRAS*, **277**, 362
- Bergin E. A., Tafalla M., 2007, *ARA&A*, **45**, 339
- Bertram E., Shetty R., Glover S. C. O., Klessen R. S., Roman-Duval J., Federrath C., 2014, *MNRAS*, **440**, 465
- Blondin J. M., Wright E. B., Borkowski K. J., Reynolds S. P., 1998, *ApJ*, **500**, 342
- Bolatto A. D., Leroy A. K., Rosolowsky E., Walter F., Blitz L., 2008, *ApJ*, **686**, 948
- Boldyrev S., Nordlund Å., Padoan P., 2002, *ApJ*, **573**, 678
- Brinch C., Hogerheijde M. R., 2010, *A&A*, **523**, A25
- Brunt C. M., 2003a, *ApJ*, **583**, 280
- Brunt C. M., 2003b, *ApJ*, **584**, 293
- Brunt C. M., Heyer M. H., 2002, *ApJ*, **566**, 276
- Brunt C. M., Heyer M. H., 2013, *MNRAS*, **433**, 117
- Brunt C. M., Mac Low M.-M., 2004, *ApJ*, **604**, 196
- Brunt C. M., Heyer M. H., Vázquez-Semadeni E., Pichardo B., 2003, *ApJ*, **595**, 824
- Brunt C. M., Heyer M. H., Mac Low M. M., 2009, *A&A*, **504**, 883
- Burkhart B., 2018, *ApJ*, **863**, 118
- Clark P. C., Glover S. C. O., Klessen R. S., 2012, *MNRAS*, **420**, 745
- Clark P. C., Glover S. C. O., Ragan S. E., Duarte-Cabral A., 2019, *MNRAS*, **486**, 4622
- Cox D. P., Gómez G. C., 2002, *ApJS*, **142**, 261
- Cunningham A. J., Klein R. I., Krumholz M. R., McKee C. F., 2011, *ApJ*, **740**, 107
- Dale J. E., Bonnell I. A., 2008, *MNRAS*, **391**, 2
- Diehl R., et al., 2006, *Nature*, **439**, 45
- Draine B. T., 1978, *ApJS*, **36**, 595
- Falgarone E., Puget J. L., 1985, *A&A*, **142**, 157
- Falgarone E., Puget J. L., Perault M., 1992, *A&A*, **257**, 715
- Federrath C., 2018, *Physics Today*, **71**, 38
- Federrath C., Roman-Duval J., Klessen R. S., Schmidt W., Mac Low M. M., 2010, *A&A*, **512**, A81
- Federrath C., Schröder M., Banerjee R., Klessen R. S., 2014, *ApJ*, **790**, 128
- Flower D. R., Pineau Des Forets G., Walmsley C. M., 2006, *A&A*, **449**, 621
- Frisch U., 1995, Turbulence. The legacy of A.N. Kolmogorov
- Frisch U., Bec J., Villone B., 2001, *Physica D Nonlinear Phenomena*, **152**, 620
- Gatto A., et al., 2015, *MNRAS*, **449**, 1057
- Gatto A., et al., 2017, *MNRAS*, **466**, 1903
- Genel S., Vogelsberger M., Nelson D., Sijacki D., Springel V., Hernquist L., 2013, *MNRAS*, **435**, 1426
- Gingold R. A., Monaghan J. J., 1977, *MNRAS*, **181**, 375

- Girichidis P., et al., 2016, *MNRAS*, **456**, 3432
- Glover S. C. O., Mac Low M.-M., 2007a, *ApJS*, **169**, 239
- Glover S. C. O., Mac Low M.-M., 2007b, *ApJ*, **659**, 1317
- Glover S. C. O., Federrath C., Mac Low M.-M., Klessen R. S., 2010, *MNRAS*, **404**, 2
- Hacar A., Kainulainen J., Tafalla M., Beuther H., Alves J., 2016, *A&A*, **587**, A97
- Haid S., Walch S., Seifried D., Wünsch R., Dinnbier F., Naab T., 2018, *MNRAS*, **478**, 4799
- Hennebelle P., Chabrier G., 2008, *ApJ*, **684**, 395
- Heyer M. H., Brunt C. M., 2004, *ApJ*, **615**, L45
- Heyer M. H., Peter Schloerb F., 1997, *ApJ*, **475**, 173
- Heyer M. H., Williams J. P., Brunt C. M., 2006, *ApJ*, **643**, 956
- Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, *MNRAS*, **445**, 581
- Hotelling H., 1936, *Biometrika*, **28**, 321
- Hubber D. A., Walch S., Whitworth A. P., 2013, *MNRAS*, **430**, 3261
- Izquierdo A. F., Galván-Madrid R., Maud L. T., Hoare M. G., Johnston K. G., Keto E. R., Zhang Q., de Wit W.-J., 2018, *MNRAS*, **478**, 2505
- Joung M. R., Mac Low M.-M., Bryan G. L., 2009, *ApJ*, **704**, 137
- Keto E., Zhang Q., 2010, *MNRAS*, **406**, 102
- Klessen R. S., Hennebelle P., 2010, *A&A*, **520**, A17
- Klessen R. S., Ballesteros-Paredes J., Li Y., Mac Low M. M., 2004, in Lamers H. J. G. L. M., Smith L. J., Nota A., eds, Astronomical Society of the Pacific Conference Series Vol. 322, The Formation and Evolution of Massive Young Star Clusters. pp 299–308 ([arXiv:astro-ph/0403469](https://arxiv.org/abs/astro-ph/0403469))
- Koch E. W., Rosolowsky E. W., Boyden R. D., Burkhardt B., Ginsburg A., Loepky J. L., Offner S. S. R., 2019, *AJ*, **158**, 1
- Kolmogorov A., 1941, Akademija Nauk SSSR Doklady, **30**, 301
- Kraichnan R. H., 1974, *Journal of Fluid Mechanics*, **62**, 305
- Kroupa P., 2002, *Science*, **295**, 82
- Krumholz M. R., Tan J. C., 2007, *ApJ*, **654**, 304
- Larson R. B., 1979, *MNRAS*, **186**, 479
- Larson R. B., 1981, *MNRAS*, **194**, 809
- Lucy L. B., 1977, *AJ*, **82**, 1013
- Mac Low M.-M., Klessen R. S., 2004, *Reviews of Modern Physics*, **76**, 125
- Matzner C. D., McKee C. F., 2000, *ApJ*, **545**, 364
- McKee C. F., Ostriker E. C., 2007, *ARA&A*, **45**, 565
- McMillan P. J., 2017, *MNRAS*, **465**, 76
- Nakamura F., Li Z.-Y., 2007, *ApJ*, **662**, 395
- Nelson R. P., Langer W. D., 1997, *ApJ*, **482**, 796
- Onsager L., 1949, *Il Nuovo Cimento*, **6**, 279
- Padoan P., Nordlund Å., 2002, *ApJ*, **576**, 870
- Padoan P., Pan L., Haugbølle T., Nordlund Å., 2016a, *ApJ*, **822**, 11
- Padoan P., Juvela M., Pan L., Haugbølle T., Nordlund Å., 2016b, *ApJ*, **826**, 140
- Padoan P., Haugbølle T., Nordlund Å., Frimann S., 2017, *ApJ*, **840**, 48
- Pakmor R., Springel V., Bauer A., Mocz P., Munoz D. J., Ohlmann S. T., Schaal K., Zhu C., 2016, *MNRAS*, **455**, 1134
- Pan L., Padoan P., Haugbølle T., Nordlund Å., 2016, *ApJ*, **825**, 30
- Pearson K., 1901, *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, **2**, 559
- Peters T., et al., 2017, *MNRAS*, **466**, 3293
- Roman-Duval J., Federrath C., Brunt C., Heyer M., Jackson J., Klessen R. S., 2011, *ApJ*, **740**, 120
- Rybicki G. B., Lightman A. P., 1986, Radiative Processes in Astrophysics
- Schneider N., Simon R., Kramer C., Stutzki J., Bontemps S., 2002, *A&A*, **384**, 225
- Smith R. J., Glover S. C. O., Clark P. C., Klessen R. S., Springel V., 2014a, *MNRAS*, **441**, 1628
- Smith R. J., Glover S. C. O., Klessen R. S., 2014b, *MNRAS*, **445**, 2900
- Smith R. J., Glover S. C. O., Klessen R. S., Fuller G. A., 2016, *MNRAS*, **455**, 3640
- Solomon P. M., Rivolo A. R., Barrett J., Yahil A., 1987, *ApJ*, **319**, 730
- Sormani M. C., Treß R. G., Klessen R. S., Glover S. C. O., 2017, *MNRAS*, **466**, 407
- Springel V., 2010, *MNRAS*, **401**, 791
- Traficante A., Fuller G. A., Smith R. J., Billot N., Duarte-Cabral A., Peretto N., Molinari S., Pineda J. E., 2018a, *MNRAS*, **473**, 4975
- Traficante A., et al., 2018b, *MNRAS*, **477**, 2220
- Truelove J. K., Klein R. I., McKee C. F., Holliman II J. H., Howell L. H., Greenough J. A., 1997, *ApJ*, **489**, L179
- Vaytet N., Tomida K., Chabrier G., 2014, *A&A*, **563**, A85
- Vazquez-Semadeni E., 1994, *ApJ*, **423**, 681
- Walch S., et al., 2015, *MNRAS*, **454**, 238
- Wilson T. L., Rohlfs K., Hüttemeister S., 2013, Tools of Radio Astronomy, doi:[10.1007/978-3-642-39950-3](https://doi.org/10.1007/978-3-642-39950-3).
- Yang B., Stancil P. C., Balakrishnan N., Forrey R. C., 2010, *ApJ*, **718**, 1062
- de Avillez M. A., Breitschwerdt D., 2005, *A&A*, **436**, 585
- de Jong T., 1977, *A&A*, **55**, 137
- van Dishoeck E. F., 2004, *ARA&A*, **42**, 119
- van Dishoeck E. F., Blake G. A., 1998, *ARA&A*, **36**, 317

APPENDIX A: RADIATIVE TRANSFER EFFECTS: LTE VS NON-LTE

A1 Structure Function: Variations from LTE considerations

Assuming LTE conditions during the radiative transfer calculations generally causes populations of ^{12}CO levels to be overestimated, which results in optically thicker lines (see Sec. 3). Figure A1 compares LTE and non-LTE profiles from edge-on $_{\phi=90^\circ}$ views of the cloud complexes explored in Fig. 4. A lot of line sub-structures are smoothed under LTE approximations. This reduces variations in velocity fluctuations over different spatial scales and leads to a systematic flattening of PCA derived scaling exponents from LTE ^{12}CO $J=1-0$ cubes for all our cloud complexes and PCA extraction methods. This can be seen by comparison between LTE Figure A2 and Table A1 against their corresponding non-LTE Figure 8 and Table 3.

This paper has been typeset from a TeX/LaTeX file prepared by the author.

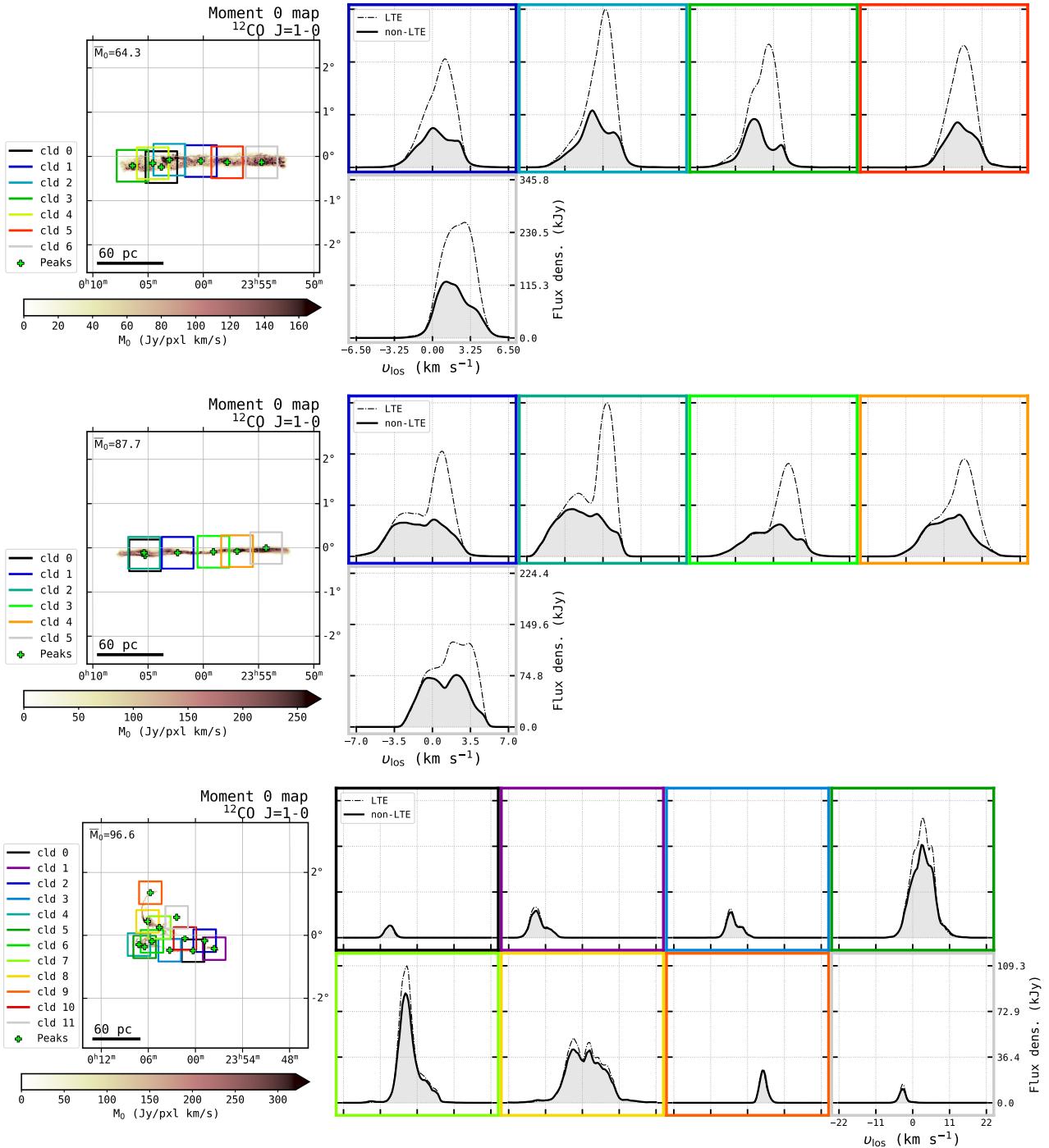


Figure A1. LTE (dotted lines) and non-LTE (solid lines and shaded regions) ^{12}CO $J=1-0$ emission profiles of individual clouds in turbulent cloud complexes generated with our Cloud Factory simulation suite. The colours of the axes spines in the right panels correspond to the colour code in the left panels (moment 0 maps) and represent the cloud portions where line profiles are extracted from. Missing cloud profiles correspond to overlapping regions in the left panels that are not taken into account during the PCA runs. In the right panels, tick spacing and labels are the same for all the plots so they are only shown in the last panel. These cloud complexes correspond to the edge-on $\phi=90^\circ$ views of the complexes analysed in Fig. 4.

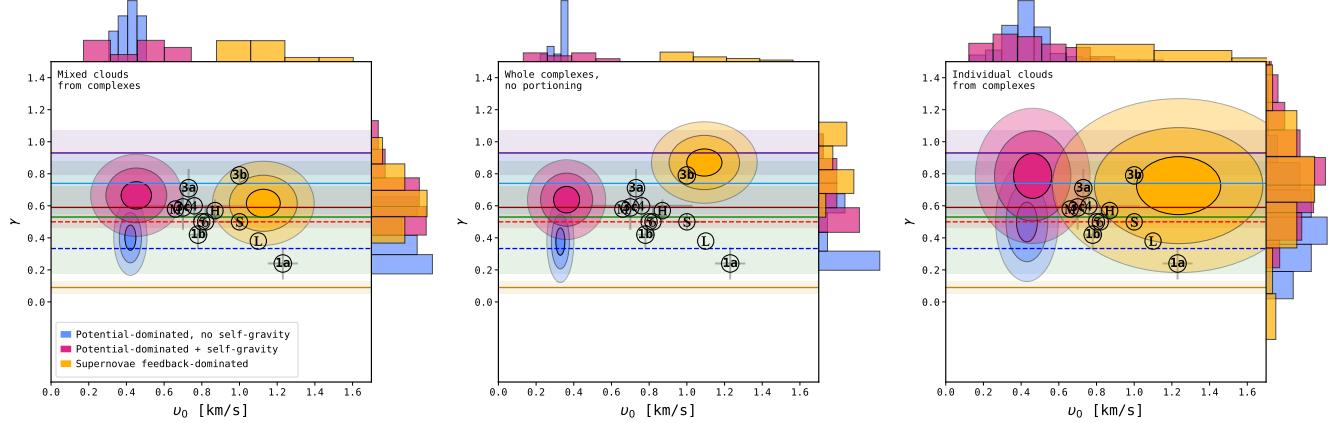


Figure A2. Same as Fig. 8 but under LTE assumptions for the radiative transfer (see Sec. 3)

PCA Mode	Physical Scenario	Complex	\bar{v}_0	$\bar{\gamma}$
Mixed clouds	potential-dominated	$A_0 \cup B_0$	0.42 ± 0.06	0.40 ± 0.16
	pot-dom+self-gravity	$A \cup B$	0.45 ± 0.16	0.67 ± 0.17
	feedback-dominated	$C \cup D$	1.14 ± 0.18	0.59 ± 0.21
Entire Complex	potential-dominated	$A_0 \cup B_0$	0.33 ± 0.05	0.38 ± 0.17
	pot-dom+self-gravity	$A \cup B$	0.36 ± 0.14	0.64 ± 0.17
	feedback-dominated	$C \cup D$	1.08 ± 0.20	0.81 ± 0.24
Individual clouds	potential-dominated	A_0	0.37 ± 0.09	0.58 ± 0.22
	potential-dominated	B_0	0.49 ± 0.09	0.43 ± 0.25
	pot-dom+self-gravity	A	0.48 ± 0.20	0.88 ± 0.26
	pot-dom+self-gravity	B	0.44 ± 0.21	0.69 ± 0.27
	feedback-dominated	C	1.26 ± 0.43	0.61 ± 0.56
	feedback-dominated	D	1.26 ± 0.52	0.80 ± 0.35

Table A1. LTE mean and standard deviations of structure function parameters derived from different PCA extraction methods.