

# The Cloud Factory II: gravoturbulent kinematics of resolved molecular clouds in a galactic potential

Andrés F. Izquierdo<sup>1,2,3★</sup>, Rowan J. Smith<sup>1</sup>, Simon C. O. Glover<sup>4</sup>, Ralf S. Klessen<sup>4,5</sup>,  
Robin G. Treß<sup>4</sup>, Mattia C. Sormani<sup>4</sup>, Paul C. Clark<sup>6</sup>, Ana Duarte-Cabral<sup>6</sup>  
and Catherine Zucker<sup>7</sup>

<sup>1</sup>Department of Physics and Astronomy, Jodrell Bank Centre for Astrophysics, University of Manchester, Oxford Road, Manchester M13 9PL, UK

<sup>2</sup>European Southern Observatory, Karl-Schwarzschild-Straße 2, D-85748 Garching bei München, Germany

<sup>3</sup>Leiden Observatory, Leiden University, NL-2300 RA Leiden, the Netherlands

<sup>4</sup>Zentrum für Astronomie, Institut für theoretische Astrophysik, Universität Heidelberg, Albert-Ueberle-Str 2, D-69120 Heidelberg, Germany

<sup>5</sup>Interdisziplinäres Zentrum für Wissenschaftliches Rechnen, Universität Heidelberg, INF 205, D-69120 Heidelberg, Germany

<sup>6</sup>School of Physics and Astronomy, Queens Buildings, The Parade, Cardiff University, Cardiff CF24 3AA, UK

<sup>7</sup>Harvard Astronomy, Harvard-Smithsonian Center for Astrophysics, 60 Garden St, Cambridge, MA 02138, USA

Accepted 2020 November 3. Received 2020 November 2; in original form 2020 May 12

## ABSTRACT

We present a statistical analysis of the gravoturbulent velocity fluctuations in molecular cloud complexes extracted from our ‘Cloud Factory’ Galactic-scale interstellar medium (ISM) simulation suite. For this purpose, we produce non-local thermodynamic equilibrium  $^{12}\text{CO } J = 1 - 0$  synthetic observations and apply the principal component analysis (PCA) reduction technique on a representative sample of cloud complexes. The velocity fluctuations are self-consistently generated by different physical mechanisms at play in our simulations, which include Galactic-scale forces, gas self-gravity, and supernova feedback. The statistical analysis suggests that, even though purely gravitational effects are necessary to reproduce standard observational laws, they are not sufficient in most cases. We show that the extra injection of energy from supernova explosions plays a key role in establishing the global turbulent field and the local dynamics and morphology of molecular clouds. Additionally, we characterize structure function scaling parameters as a result of cloud environmental conditions: some of the complexes are immersed in diffuse (interarm) or dense (spiral-arm) environments, and others are influenced by embedded or external supernovae. In quiescent regions, we obtain time-evolving trajectories of scaling parameters driven by gravitational collapse and supersonic turbulent flows. Our findings suggest that a PCA-based statistical study is a robust method to diagnose the physical mechanisms that drive the gravoturbulent properties of molecular clouds. Also, we present a new open source module, the PCAFACTORY, which smartly performs PCA to extract velocity structure functions from simulated or real data of the ISM in a user-friendly way.

**Key words:** gravitation – radiative transfer – turbulence – ISM: clouds – ISM: structure.

## 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 (Mac Low & Klessen 2004; 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; Klessen & Glover 2016). 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 (ISM;

Blitz 1993; Williams, Blitz & McKee 2000). These cloud complexes consist of interconnected molecular clouds (2–20 pc) that, at the same time, exhibit high degrees of sub-structuring over subsequent scales (Falgarone, Puget & Perault 1992) and filamentary signatures (André et al. 2010; Ragan et al. 2014; Smith, Glover & Klessen 2014b; Arzoumanian et al. 2019). Thus, studying the dynamics of molecular structures in different spatial regimes becomes highly relevant to uncover the nature and evolution of star formation properties.

A few decades ago, Larson (1979, 1981) discovered a systematic increase of the global velocity dispersion  $\Delta v$  (km s<sup>−1</sup>) with the projected size  $L$  (pc) of diverse molecular associations ( $\Delta v \propto L^{0.38}$ ) using mostly optically thin ( $^{13}\text{CO}$ ,  $\text{H}_2\text{CO}$ ,  $\text{NH}_3$ ) but also optically thick tracers ( $^{12}\text{CO}$ ). Larson interpreted this hierarchical behaviour to be a consequence of energy transport across successive spatial scales as it is reminiscent of the Kolmogorov structure law ( $\delta v \propto l^{1/3}$ , where lower case  $\delta v$  and  $l$  indicate internal velocity fluctuations and spatial scales), derived from the

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

statistical framework developed by Kolmogorov (1941) and Onsager (1949) for viscous incompressible (subsonic) turbulent fluids.<sup>1</sup> This linewidth–size relationship, often termed the (first) Larson law, would lay the groundwork for subsequent literature on the role of turbulence in setting dynamical signatures of the ISM.

Solomon et al. (1987) focused on a more homogeneous sample of clouds and reported a similar but slightly steeper linewidth–size relationship ( $\delta v \propto l^{1/2}$ ) using  $^{12}\text{CO}$  data. They interpreted this result as a consequence of virial equilibrium under the premise that the mean surface density of clouds is independent of size. However, this idea would be contradicted a few decades later by Heyer et al. (2009) who re-examined the same objects using a lower opacity tracer ( $^{13}\text{CO}$ ) and higher spectral and angular resolution. In any case, classical statistical-hydrodynamic theories (e.g. Kraichnan 1974; Fournier & Frisch 1983) derived the same velocity scaling index ( $\gamma_2 = 1/2$ ) for fluids in a compressible (supersonic) turbulent field, which thereby suggests that energy dissipation in molecular clouds not only occurs at small scales (where viscosity dominates) but can be driven by supersonic shocks at larger scales as well (McKee & Ostriker 2007).

Turbulence is essential not only for triggering primordial density enhancements and seeding star formation, but also for regulating 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 acts against gravity, which alongside support from magnetic forces, prevents runaway gravitational collapse (Falgarone et al. 1992; Federrath 2018). Further details about a gravoturbulent scenario for fragmentation in molecular clouds and its implications for star formation properties can be found in 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; Schneider et al. 2015). Conversely, low column densities dominated by turbulent supersonic motions yield lognormal PDFs (Vazquez-Semadeni 1994; Kainulainen et al. 2009) that can also exhibit non-Gaussian wings due to intermittency effects (Federrath et al. 2010a). 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; Kainulainen et al. 2009; Burkhardt 2018). Gravoturbulent mechanisms are hence crucial to establish stellar and core initial mass functions that may be closely related to the mass distribution of parental clouds (Padoan & Nordlund 2002; Hennebelle & Chabrier 2009).

A great deal of effort has also been expended on understanding the origin of non-thermal motions in the cold ISM. Heyer & 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  $\delta v$  ( $\text{km s}^{-1}$ ) associated with characteristic spatial scales  $l$  (pc) of the analysis region. They applied the algorithm to synthetic and real objects and found power-law dependencies analogous to the Larson linewidth–size relationship. Several studies were then carried out to connect the scaling parameters retrieved from this method to their intrinsic hydrodynamic structure function (Brunt & Heyer 2002; Brunt et al. 2003; Heyer & Brunt 2004; Roman-Duval et al. 2011; Brunt & Heyer 2013), responsible for describing the 3D velocity fluctuations field as a function of the spatial separation of particles in the fluid. Other works tested the sensitivity of the technique to different feedback conditions. Heyer, Williams & Brunt (2006) found distinct 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  $^{12}\text{CO}$  intensity, they infer a steeper relationship  $\delta v \propto l^{0.82 \pm 0.03}$  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 suggest that using  $^{13}\text{CO}$ , 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.

Magnetohydrodynamic simulations carried out by de Avillez & Breitschwerdt (2005) and Joungh, Mac Low & Bryan (2009) included a global Galactic context to consistently investigate the ISM evolution. They found that several observational properties of the ISM turbulence can be reproduced in supernova feedback-dominated scenarios. However, due to the achievable 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. (2010a) simulated synthetic turbulent fields made up of two different forcing components, solenoidal and compressive, within periodic uniform grids assuming isothermal gas. They suggest 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 discs. They used converging flows of accretion, incoming from the computational boundary with superimposed mean velocities and fluctuations, and obtained a linewidth–size relation  $\Delta v = 0.8(L/\text{pc})^{0.5} \text{ km s}^{-1}$  compatible with 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 supernovae 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) and H II regions (Peters et al. 2017; Haid et al. 2018), would demonstrate

<sup>1</sup>However, Larson (1981) also hinted at the possibility of supersonic turbulence in molecular clouds given the steeper scaling exponent compared to that of Kolmogorov’s law. Larson interpreted this as a lack of velocity fluctuations at small scales caused by energy dissipation at larger scales via supersonic shocks.

that the role of stellar feedback is also essential in configuring the turbulent field of molecular clouds (for a summary see also Klessen & Glover 2016). 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 considering neither the large-scale gravitational potential nor differential rotation.

Our Cloud Factory simulations seek to address this limitation by including both supernova feedback and the large-scale Galactic environment with high enough resolution to study the internal turbulence within clouds. We take into account the global Galactic context using a multicomponent gravitational potential and Galactic differential rotation while, at the same time, resolving selected molecular clouds with cell masses as small as  $0.25 M_{\odot}$ . We include stellar feedback in the 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-local thermodynamic equilibrium (LTE) radiative transfer calculations and the PCA technique on our cloud complexes to investigate the detailed signatures of non-thermal motions over 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 supernova 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, which can be used in future 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 radiative transfer set-up and statistical description of velocity fluctuations in fluids. In Section 5, we outline the general workflow and explain the three PCA extraction methods explored in the paper. We then present in Section 6 the results split by physical scenario (6.1), line-of-sight projection (6.2), time snapshot (6.3), and analysis scale (6.4). We provide a discussion on the resemblance of our self-consistently generated clouds to observational data and the role of supernovae feedback in Section 7, and wrap up with the conclusions of the work in Section 8. In Appendix A, we add supporting figures including edge-on cloud column densities, line emission, and optical depth profiles, and also show variations in PCA-derived parameters when assuming LTE and large velocity gradient (LVG) level populations for the radiative transfer.

## 2 THE CLOUD FACTORY SIMULATION SUITE

### 2.1 The hydrodynamic code and physical ingredients

The cloud complexes that we examine in this work are extracted from our Cloud Factory simulation suite (Smith et al. 2020, hereafter Paper I), which is built on a version of the AREPO code (Springel 2010; Pakmor et al. 2016) customized with a set of physical/chemical modules that account for various mechanisms taking place in the cold molecular ISM such as

- (i) the Galactic gravitational potential,
- (ii) time evolution of CO and hydrogen chemistry,

- (iii) ultraviolet (UV) extinction considering  $H_2$  and CO shielding properties, and dust absorption,
- (iv) star formation via sink particles, and
- (v) injection of (energy/momentum) feedback from supernova explosions.

We use an analytic description of the large-scale gravitational potential of the Galaxy to efficiently 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 by 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 potential are fully described in 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 UV photodissociation rate from de Jong (1977) and Falgarone & Puget (1985), which is a function of the gas number density and the strength of the UV portion of the interstellar radiation field (assumed in our simulations to be that of the solar neighbourhood derived by Draine 1978). We use the TREECOL algorithm (Clark, Glover & Klessen 2012) to compute the UV extinction of the medium by considering  $H_2$  and CO self-shielding, the shielding of CO by  $H_2$ , and the shielding of both by dust absorption.

Regarding the convergence of CO abundance in our simulations, at our highest resolution points for a number density of  $10^4 \text{ cm}^{-3}$  we have a resolution of 0.05 pc or smaller (see fig. 4 in Paper I), which compares well with the converging flow spatial resolution requirement from Joshi et al. (2019). It is unclear that we meet their convergence criterion everywhere as our resolution is spatially variable. However, when we plug a conservative estimate of the internal velocity dispersion in 0.1 pc scales in their criterion, we meet the resolution requirement at these densities.

The non-equilibrium hydrogen chemistry from Glover & Mac Low (2007a,b), implemented in our Cloud Factory, involves reactions between molecular ( $H_2$ ), atomic (H), and ionized ( $H^+$ ) hydrogen, electrons, cosmic rays, dust grains, and the UV radiation field. This encompasses  $H_2$  formation on grains, collisional and photodissociation of  $H_2$ , cosmic rays and collisional ionization of H, and  $H^+$  recombination in the gas phase or on dust grains. The net energy exchange due to radiative and chemical gas heating or cooling is computed using an atomic and molecular cooling function as outlined in Clark et al. (2019).

Our model of star formation uses a hybrid approach based on 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, Bonnell & Price (1995) and Federrath et al. (2010b), a cell and its neighbours have to be above a critical density  $\rho_c$  and satisfy energy checks to confirm that they are bound and the internal collapse is runaway: the cells must be located on a local minimum of the gravitational potential, outside the accretion radius of any other sink particle and have inwardly directed velocities and accelerations. 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 (Hubber, Walch & Whitworth 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 dominant driving mechanism of turbulence in star-forming galaxies (Mac Low & Klessen 2004; Padoan et al. 2016a). Either way, other mechanisms such as outflows/jets (Nakamura & Li 2007; Cunningham et al. 2011; Federrath et al. 2014), stellar winds (Dale & Bonnell 2008; Gatto et al. 2017; Peters 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 supernova explosions and (ii) a mixture of random supernovae and 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 yr 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 a realistic stellar initial mass function in order to compute the number 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). When the target mass is large ( $100 M_{\odot}$ ; see Section 2.2 below), the sink particles introduced in the simulation correspond to portions of clouds with size scales larger than individual star-forming cores, and hence a large fraction of gas in the sinks should not actually form stars. We account for this by adopting a low star formation efficiency for these sinks of 1–2 per cent based on the work of Krumholz & Tan (2007). When the simulation target mass is small ( $<10 M_{\odot}$ ), the sink particles more closely correspond to individual star-forming cores and so in this case we adopt a higher star formation efficiency of 33 per cent, based on Matzner & McKee (2000). In this approach, we also use random supernovae but at a lower rate of 1 event every 300 yr to account for Type Ia supernovae. The way in which energy from supernovae is released into the gas depends on whether the Sedov–Taylor phase of the supernovae expansion is resolved, 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 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 and in Tress et al. (2020).

## 2.2 Simulation set-up and refinement scheme

The initial gas distribution of our simulation is based on observational constraints and theoretical modelling of the Milky Way presented in McMillan (2017). They suggest an exponential profile for the H and  $H_2$  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 H, from which  $H_2$  will form self-consistently according to our chemical treatment as the Galaxy evolves.

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. In this stage, the cell target mass resolution is set to  $1000 M_{\odot}$  and the mesh refinement operates accordingly.

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 three 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 supernova 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, and the supernova feedback is mixed. By mixed feedback we mean that both the random and supernovae tied to sink particles are turned on (see our supernova implementation in Section 2.1). The target mass resolution of this phase is initially set to  $100 M_{\odot}$  but is further lowered down to  $10 M_{\odot}$  for the final 10 Myr of the middle phase.

In the final stage of our simulation, and in order to resolve sub-structures within the processed cold ISM, we further increase the resolution on individual cloud complexes of  $\sim 100$  pc radius within the high-resolution box by injecting Monte Carlo tracer particles (Genel et al. 2013) everywhere the gas density is above  $100 \text{ cm}^{-3}$ . 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. cell diameters of  $d_{\text{cell}} = 1 \text{ pc}$  at  $n = 5 \text{ cm}^{-3}$ , or  $d_{\text{cell}} = 0.03 \text{ pc}$  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.

Regardless of the stage of the simulation, we require that the Jeans length is resolved by at least four cells everywhere in the mesh to adequately check energy and bounding conditions and avoid artificial fragmentation (Truelove et al. 1997; Federrath et al. 2011). If sink creation densities are achieved but the gas fails to pass the energy checks (see Section 2.1), we continue to resolve the gas until it is unambiguously bound as long as it remains above the sink creation density.

## 2.3 The selected cloud complexes

In order to comprehensively investigate non-thermal motions in our synthetic clouds we use the same regions as in Paper I: A, B, C, and D, but include two more complexes,  $A_0$  and  $B_0$ , in which self-gravity is switched off.

These cloud complexes arise from contrasting environments in the Galaxy. Complexes  $A_0$  and A are (at the same location) in a dense spiral arm, whereas  $B_0$  and B are in an interarm, more diffuse region. Cloud complexes C and D, representative of the feedback-dominated scenario, are the densest regions in the high-resolution box and were born after a burst of clustered supernova feedback. However, complex D, the denser of the two, gets to form massive stars that undergo supernovae explosions over time, whereas complex C is only influenced by external feedback as it does not produce



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

Cloud complex	Galactic potential	Self-gravity	Supernova feedback		Description
			Random	On sinks	
A <sub>0</sub>	✓	✗	✓	✗	Inside arm
B <sub>0</sub>	✓	✗	✓	✗	Interarm region
A	✓	✓	✓	✗	Inside arm
B	✓	✓	✓	✗	Interarm region
C	✓	✓	✓	✓	No embedded SNe
D	✓	✓	✓	✓	Embedded SNe

embedded supernovae during the analysis time. This information is briefly summarized in Table 1. A figure showing the exact location of our cloud complexes can be found in [Paper I](#).

We consider two different time snapshots for complexes A<sub>0</sub>, B<sub>0</sub>, C, and D, and (for time evolution analysis) three snapshots for complexes A and B. In all cases, the first snapshot was extracted at a time when no massive sink particles had formed yet. Also, we explore three different cloud orientations, which we refer to as face-on, edge-on  $\phi=0^\circ$  and edge-on  $\phi=90^\circ$  views. In cylindrical coordinates, the face-on line-of-sight points towards the  $\{-\hat{e}_z\}$  direction, meaning that the cloud complex is viewed from above the Galaxy. The edge-on  $\phi=0^\circ$  and edge-on  $\phi=90^\circ$  lines-of-sight point, respectively, towards the  $\{+\hat{e}_\phi\}$  and  $\{-\hat{e}_R\}$  directions.

Fig. 1 shows face-on projections of H, H<sub>2</sub>, and <sup>12</sup>CO column densities from the selected set of cloud complexes, 2 Myr after tracer refinement has commenced (edge-on  $\phi=90^\circ$  views can be found in Fig. A3). As a short comment, note that the CO density does not necessarily trace the density of hydrogen species, which is a consequence of collisional and photodissociation processes induced by supernova explosions and the interstellar radiation field in our simulations. This is particularly apparent in feedback-dominated complexes C and D, hinting at high amounts of CO-dark molecular gas (see e.g. Smith et al. 2014a), which is sensitive to variation of the feedback conditions in clouds.

### 3 RADIATIVE TRANSFER MODELLING

We perform radiative transfer simulations of our synthetic cloud complexes using the Line Modelling Engine code (LIME,<sup>2</sup> Brinch & Hogerheijde 2010) and the Polarized Radiation Simulator (POLARIS,<sup>3</sup> Reissl, Wolf & Brauer 2016). Both are flexible codes that predict molecular line and dust continuum emission from arbitrary 3D geometries in (sub-)millimetre and infrared wavelengths.

For a given grid point distribution, in our case with information on position, gas density, temperature, and velocity, the codes construct a Delaunay triangulation and its corresponding Voronoi mesh where they iteratively propagate photons and integrate the radiative transfer equation. In particular, LIME comprises two approaches to solve the line excitation problem, suitable for matter in different equilibrium states:

(i) An 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 that obey a Maxwell–Boltzmann distribution law. The code uses this to calculate

the level populations and the Kirchhoff’s law for thermal radiation to solve the transfer equation (see e.g. Rybicki & Lightman 1986).

(ii) 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. In each cell, the algorithm computes provisional level populations using the incoming local radiation and collisional rates and releases a number of photons proportional to the number of neighbouring cells. The calculation stops when the propagating photons escape the physical domain, 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 a free parameter.

On the other hand, besides LTE, POLARIS supports a LVG approximation that we also explore in this work. This approach assumes that velocity variations over a given size scale are larger than microturbulent and thermal velocities, which simplifies the computation of optical depths and level populations (Sobolev 1960).

Once the level populations are determined in either of the approaches, the codes integrate the radiative transfer equation along isotropic tracer rays that cross the Voronoi grid 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 3D  $n_{\text{xpix}} \times n_{\text{ypix}} \times n_{\text{chan}}$  position–position–velocity (PPV) cube containing intensity (in units of either Jy pix<sup>−1</sup>, Kelvin or L<sub>☉</sub> pix<sup>−1</sup>) or optical depth ( $\tau$ ) information as a function of the spectral channel (in m s<sup>−1</sup>).

Full radiative transfer modelling is necessary to produce synthetic emission maps that can be readily compared to observational data with analogous methods, especially at the present time 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 spatial and spectral finite resolution and natural constraints from projection and optical depth effects.

In the Appendix, we show variations in line emission profiles (Fig. A4) and optical depth maps (Fig. A5) after considering the radiative transfer approaches (LTE, LVG, and non-LTE) outlined in this Section, as well as their impact on PCA-derived structure function parameters (Fig. A6).

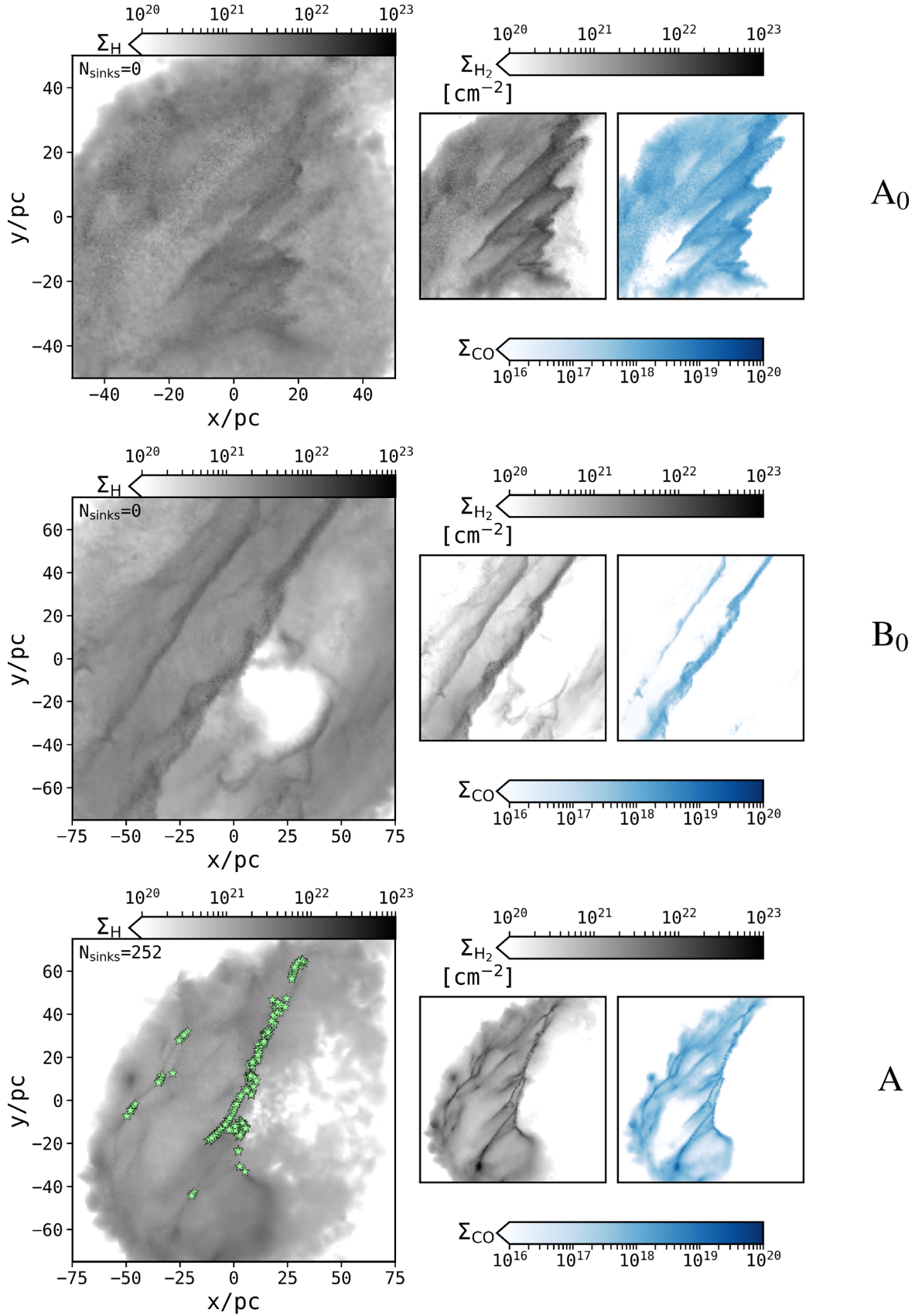
#### 3.1 Non-LTE <sup>12</sup>CO J = 1 – 0 line excitation

Our analysis focuses on the emission of carbon monoxide in its ground state rotational transition <sup>12</sup>CO J = 1 – 0, which, along with other isotopologues, have been used extensively to trace 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 the non-LTE implementation of LIME because, in our simulations, a considerable fraction of the H<sub>2</sub> density (the main collisional partner of <sup>12</sup>CO) is below the critical density ( $n_{\text{crit}} \sim 2 \times 10^3$  cm<sup>−3</sup>) to collisionally populate <sup>12</sup>CO at the upper level of its ground transition, which is valid within a wide range of temperatures (2–3000 K, Yang et al. 2010). This implies that the first rotational level of <sup>12</sup>CO is populated by different mechanisms in our cloud complexes; it is

<sup>2</sup><https://lime.readthedocs.io>

<sup>3</sup><http://www1.astrophysik.uni-kiel.de/~polaris>



**Figure 1.** Face-on projections of  $\text{H}$ ,  $\text{H}_2$ , and  $^{12}\text{CO}$  column densities ( $\Sigma$ ) from cloud complexes (labelled on the right) extracted 2 Myr after injecting tracer particles in the simulations. If any, sink particles are overlaid on  $\text{H}$  maps as the star markers.

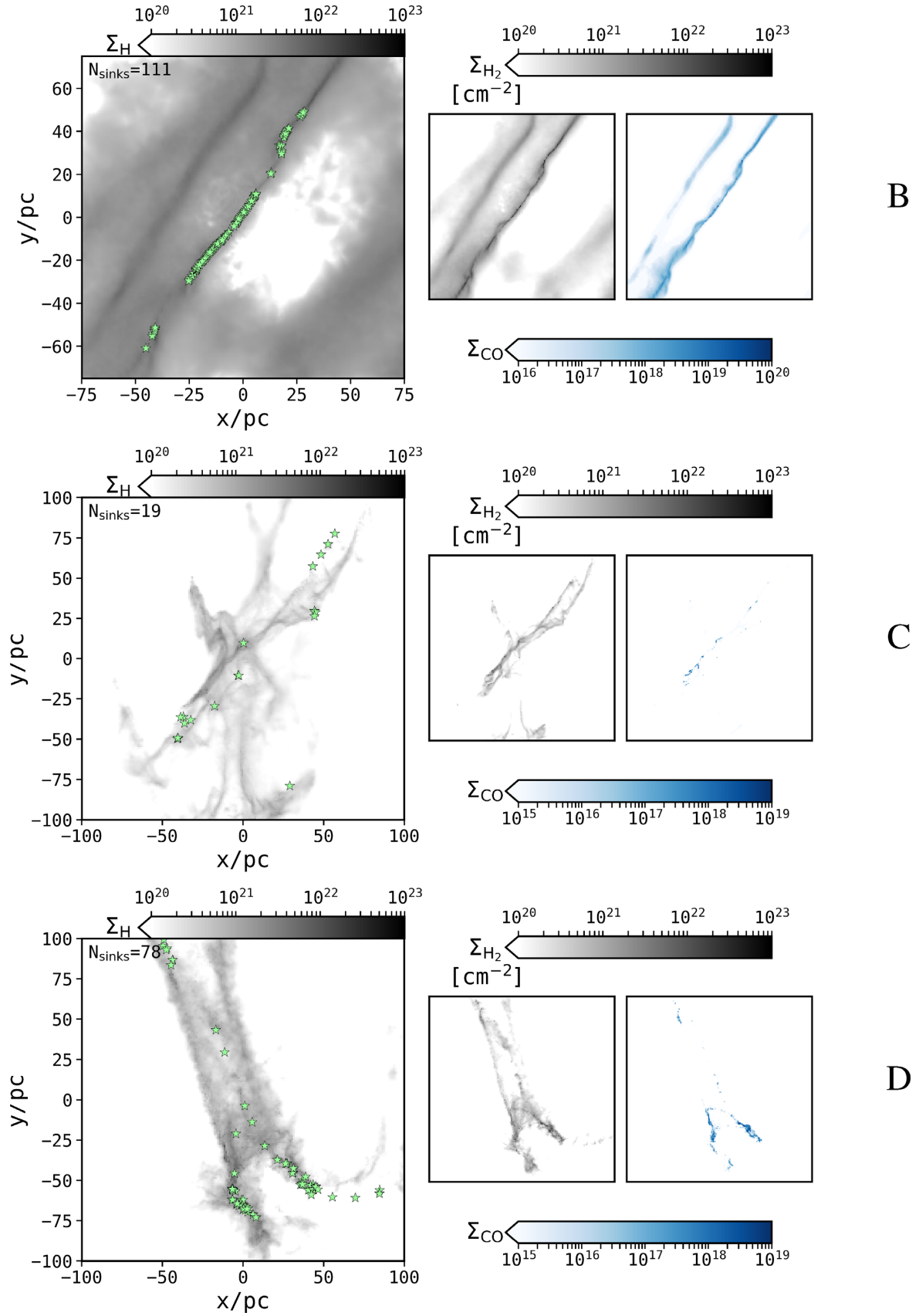


Figure 1. – continued.

dominated by radiation in diffuse regions and thermally controlled by collisions in dense regions (see section 12.4 of Wilson, Rohlfs & Hüttemeister 2013). We note that we do not consider collisions between CO and H in our modelling, which may be important for diffuse CO gas near the edges of clouds due to turbulent mixing.

### 3.2 New tools for handling AREPO-Like meshes in radiative transfer codes

#### 3.2.1 Input unstructured meshes and implementation of KDTREE algorithm in LIME

Our previous customization 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 on 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 on (highly) non-uniform Voronoi grids specially adapted to track the effects of physical mechanisms governing the gas dynamics at several time and length-scales. On top of that, our statistical analysis requires well-determined spatial scales to properly and unambiguously uncover the gravoturbulent nature of molecular clouds via velocity structure functions (see Section 4.1). For these reasons, we modified LIME to also handle unstructured meshes pre-processed with SF3DMODELS. To this end, we halt the default randomly weighted generation of grid points in LIME and force it to rather take the simulation points to reconstruct the Voronoi mesh where the radiative transfer is solved.

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 rather slow [of order  $O(N^2)$ ] with increasing 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,<sup>4</sup> which is pre-built in our customized 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 ‘sink-points’ (not to be confused with the sink particles that we use to represent star formation in the AREPO simulations), through which tracer rays emerge from the physical model to make up the synthetic image. Since the input unstructured grid can have any shape, large voids between sink-points and inner physical grid points are likely to exist, leading to artefacts 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 incorporated in SF3DMODELS. We distribute  $N/10$  dummy points randomly between a radius enclosing 90 per cent of the total gas mass and the maximum radial extent of the grid. However, this needs to be done with caution. For instance, including dummy points starting at small radii could induce multiple artificial holes in actual regions of the simulation and may also lead to an underestimation

of gas masses as the volume of the original grid cells would have to decrease to make room for the newly inserted dummies.

#### 3.2.2 Removal of twin AREPO cells after a refinement step

According to Springel (2010), a cell meeting user-defined refinement criteria is split along its centroid into two cells. Initially, the position locators of the newborn cells co-exist with the original cell centre. During subsequent time-steps, the new locators are separated via mesh-regularization techniques until they reach the actual centroids of the split cells. This bears the possibility of finding two or more cell locators at the same position in a single time snapshot of the simulation. Such subtlety can lead either to errors during the grid reconstruction by the triangulation algorithms of the radiative transfer codes, or it can slow down intermediate grid smoothing steps.

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 ids based on unique radially sorted cells. These unique cells are then compared only to their nearest neighbours to check if they share the exact same location in 3D. The algorithm recycles the mass of the twin cells into the surviving cell using a direct summation, whereas the other physical properties remain unchanged as they are approximately equal in all the twin cells. The algorithm returns a clean dictionary with unique cells and their new physical properties.

The SF3DMODELS package and the latest customized version of LIME are open source and documented online.<sup>5</sup>

### 3.3 Radiative transfer set-up

The front domain of the cloud complexes is deliberately set 2.4 kpc away from the observer, a typical distance to large nearby star-forming regions (e.g. W33, Immer et al. 2013). The pixel size of the PPV cubes is 26 arcsec, which translates into a projected physical resolution of 0.3 pc. This pixel size is a good compromise between resolving most of the cells from our AREPO meshes (which contain cells as small as  $\sim 0.03$  pc, see fig. 3 in Paper I) and the processing time of the radiative transfer. We assume an optimal scenario where no beam smearing or noise are considered. The spectral resolution, or channel width, for each cloud complex is determined using the relation  $\Delta_{\text{chan}} = (v_{\text{max}} - v_{\text{min}})/(n_{\text{chan}} - 1)$ , where  $v_{\text{max}}$  and  $v_{\text{min}}$  are the maximum and minimum projected velocities along each of the three lines-of-sight 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 PCA. Typical channel-widths range from  $\sim 0.2 \text{ km s}^{-1}$  in potential-dominated complexes to  $\sim 0.4 \text{ km s}^{-1}$  in feedback-dominated complexes.<sup>6</sup> For the calculation of level populations in the non-LTE set-up, we assume 50 iterations as the populations converge after approximately 30 cycles. We ran tests with 100 iterations for selected regions and found differences of  $< 5$  per cent in mean intensities, which is reasonable within the expected random uncertainty.

For consistency, we include micro-turbulence as an additional source of line broadening to account for non-thermal motions on 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

<sup>5</sup><https://star-forming-regions.readthedocs.io>

<sup>6</sup>We also analysed cubes with constant  $\Delta_{\text{chan}} = 0.1 \text{ km s}^{-1}$  for each physical scenario, at one time-step, and in all cases the PCA-derived parameters remained unchanged except for the associated errors that decreased systematically by a factor  $\leq 3$ .

<sup>4</sup><https://github.com/jtsiomb/kdtree>



in quadrature to the standard thermal broadening during the line transfer. A typical mass-weighted average temperature in our cloud complexes is 11 K, which yields a sound speed of  $0.25 \text{ km s}^{-1}$  for  $\text{H}_2$ . In cold dense cells, where the  $^{12}\text{CO}$  mass is non-negligible, micro-turbulence is generally much smaller than the bulk speed of the cells ( $\times 10$ – $100$  lower). Note that there may be additional unresolved small-scale turbulence or organized motions (e.g. driven by gravity) within the cell so our assumed micro-turbulence is a lower limit.

LIME takes the number density of hydrogen nucleons and CO abundance available from our simulations to compute the CO mass consistently. We assume that the CO collisional partners, the spin isomers of molecular hydrogen (para- $\text{H}_2$  and ortho- $\text{H}_2$ ), have a 1:1 ratio corresponding to the expected value for molecular gas with an age of a few Myr (see e.g. Flower, Pineau Des Forêts & Walmsley 2006). However, we note that the CO excitation rate is only weakly sensitive to the ortho-to-para ratio and hence our results should not be affected by this choice.

## 4 STATISTICAL DESCRIPTION OF VELOCITY FLUCTUATIONS IN MOLECULAR ASSOCIATIONS

### 4.1 Velocity structure function of a fluid

The concept of a generalized function to describe non-thermal velocity fluctuations in a 3D fluid originated with the work of Kolmogorov (1941). The author 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 statistically random variables. Using similarity hypotheses on time and length-scale energy dissipation rates, the author found that the (averaged) velocity dispersion is a function of the spatial separation between test points of the fluid. This work 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 the effect of viscosity exceeds the magnitude of velocity fluctuations on 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

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

where  $A$  is a dimensionless constant,  $k = 2\pi/l$  is the wavenumber associated with 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 kind considered in Kolmogorov's scenario are rare in molecular clouds, where non-thermal motions are not negligible but rather transonic or supersonic (Mac Low & Klessen 2004; McKee & Ostriker 2007). In particular, supersonic motions lead to shock-dominated turbulence (Burgers-like turbulence, Burgers 1939; Kraichnan 1974; Fournier & Frisch 1983; Passot, Pouquet & Woodward 1988; Frisch, Bec & Villone 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 decay faster at high-order wavenumbers:

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

Moreover, it is well known from early studies that the gradient of energy dissipation in non-ideal turbulent fluids depends on the properties of the medium (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 (see e.g. Federrath et al. 2010a),

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

The exponent  $\beta$  is known as the spectral index of the energy spectrum in a 3D turbulent fluid and is an intrinsic property of the velocity distribution in the medium. In addition, this power-law dependence is supported in Onsager (1949) by the fact that the total vorticity of a fluid is in general a linear combination of the wavenumber scale vector  $\vec{k}$ .

Using equation (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-sized) 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 is straightforward to derive the root-mean-square (rms) velocity,

$$\langle |\delta v(l)|^2 \rangle^{1/2} = v_0 l^{\gamma_2}; \quad \text{with } \gamma_2 = (\beta - 1)/2, \quad (5)$$

where  $v_0$  and  $\gamma_2$  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) and Frisch (1995) in order to (statistically) explain multicomponent turbulent motions, which the scaling parameters of the rms-velocity field are unable to model comprehensively (see Brunt et al. 2003). In molecular clouds, multicomponent turbulence arise from several dissipation mechanisms such as shocks, magnetic fields, radiative cooling, and heat diffusion (Boldyrev, Nordlund & Padoan 2002), but also from energy-injection mechanisms like stellar feedback, both affecting different time and length-scales of the region. The velocity structure function is written as

$$S_p(l) = \langle |\delta v(l)|^p \rangle; \quad \text{with } \delta v = v(r) - v(r + l), \quad (6)$$

where the exponent  $p$  accounts for the order of the velocity fluctuations  $\delta v$  and provides information about the degree of coherence of the velocity field when subject to spatial variations  $l$ . This relationship can be modelled as a power law of the size scale  $S_p(l) \propto l^{\zeta_p}$  analogous to the approach followed in equations (4) and (5) for the rms velocity, then

$$\langle |\delta v(l)|^p \rangle^{1/p} = v_0 l^{\gamma_p}; \quad \text{with } \gamma_p = \zeta_p/p, \quad (7)$$

being  $\gamma_p$  (or the equivalent  $\zeta_p$ ) the intrinsic scaling exponent of order  $p$  and  $v_0$  the magnitude of velocity fluctuations known as scaling coefficient or normalization of the velocity structure function.<sup>7</sup> In Section 4.2.1, we expand the discussion on structure function variations with changing  $p$ -orders.

As a final comment, note from equation (5) that Kolmogorov's law ( $\beta = 5/3$ ) and the Burgers-like turbulence ( $\beta = 2$ ) yield rms scaling exponents,  $\gamma_2 = 1/3$  and  $\gamma_2 = 1/2$ , respectively.

<sup>7</sup>Note that this expression satisfies the upper limit of the Hölder condition for the  $(l, \delta v)$  metrics, which implies that the velocity distribution is uniformly continuous within the spatial domain of the fluid.

## 4.2 Principal component analysis

In this sub-section, we briefly summarize the main aspects of the PCA technique applied to the study of the structure and dynamics of molecular associations in the ISM.

PCA is a statistical multivariate method that transforms an input data set 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 data set, 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 for the study of ISM dynamics in the work of Heyer & Schloerb (1997), who described the formalism of the technique when considering 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.

Using the PPV intensity cube  $T_{ij} = T(\mathbf{r}_i, v_j)$ , composed of  $n = n_x \times n_y$  pixels, and  $n_c$  velocity channels, the method first finds the associated covariance matrix defined as

$$S_{jk} = \frac{1}{n} \sum_{i=1}^n T_{ij} T_{ik}. \quad (8)$$

The method then computes the corresponding eigenvalues  $\lambda_m$ , and eigenvectors  $\mathbf{u}_m$ , of the covariance matrix, by solving the eigenequation  $S\mathbf{u} = \lambda\mathbf{u}$ . The subscript  $m$  stands for the  $m$ th principal component of the covariance matrix. Next, the intensity cube is projected on to the eigenvectors to construct the associated eigen images:

$$I_m(\mathbf{r}_i) = \frac{1}{n} \sum_{j=1}^{n_c} T_{ij} \mathbf{u}_{mj}. \quad (9)$$

The characteristic width of the 3D autocorrelation functions of the eigenvectors  $\mathbf{u}_m$  and eigenimages  $I_m$ , i.e. the velocity and spatial lags for which the autocorrelation functions have decreased by a factor of  $1/e$  from their peak values, determine, respectively, the  $m$ th order velocity fluctuations  $\delta v_m$  and the spatial scales  $\delta l_m$ . Further details can be found in BH13 and references therein.

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

$$\delta v = v_0 l^\alpha, \quad (10)$$

linking velocity fluctuations  $\delta v$  to characteristic size scales  $l$ , where the normalization  $v_0$  is the magnitude of velocity fluctuations in a given cloud complex. Note that we use  $\alpha$  to refer to the PCA-derived scaling exponent, but in general it differs from the intrinsic exponent  $\gamma_p$  of the velocity structure function as we briefly discuss in Section 4.2.1. In any case, this pseudo-structure function seems analogous to the empirical linewidth–size relationship,

$$\Delta v = CL^\Gamma, \quad (11)$$

found by Larson (1979, 1981) after computing global velocity dispersions  $\Delta v$  from an ensemble of varied molecular associations as a function of their projected sizes  $L$ . We use capitals  $C$  and  $\Gamma$  to indicate that the scaling coefficient and exponent, respectively,

were derived from a global linewidth–size analysis that combined data from independent clouds rather than from a local-scale study of velocity fluctuations within individual clouds.

Both relationships were systematically investigated in subsequent works, some of which are listed in Table 2 for further comparison with our results. Also, both of them resemble a structure function-like dependence (see equation 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 as summarized in Sections 4.2.1 and 4.2.2.

We use the PCA module included in the TURBUSTAT<sup>8</sup> package (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 per cent of the total variance), but not too much (i.e.  $\gg 99$  per cent) 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 ( $v_0$ ,  $\alpha$ ) are computed separately by our PCAFACTORY package according to the three extraction methods summarized in Section 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  $\alpha$  with the spectral index  $\beta$ , which is intrinsic to 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 sporadic 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  $\beta$  between 1.2 and 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 calibration because, as BH13 suggest, their analytic result should be seen as a supportive element rather than a replacement of the previous empirical estimations given the approximations used in their derivation. The RD11 calibration can be written as

$$\beta = (0.20 \pm 0.05) + (2.99 \pm 0.09)\alpha, \quad (12)$$

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

The retrieved spectral indices  $\beta$  are related to the second-order structure function scaling exponent  $\gamma_2$  via equation (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 multicomponent nature triggered by several energy dissipation and injection mechanisms.

Brunt et al. (2003) demonstrated that equation (12) is a good link between  $\alpha$  and  $\gamma_2$ , which holds even for orders  $p \neq 2$  as long as the velocity field of the region is non-intermittent and therefore reproducible by a single  $\gamma$ . For completeness of details,

<sup>8</sup><https://turbustat.readthedocs.io>

**Table 2.** Scaling parameters from Larson-like and PCA-derived structure functions reported in selected literature. For comparison, calibrations indicated in Section 4.2.1 are used to translate parameters to rms structure function parameters when not provided. From top to bottom, the horizontal rules separate classical, observational and theoretical parameters.

Reference	Marker	Larson-like exponent $\Gamma$	PCA-derived exponent $\alpha$	Structure function $v_0$ (km s $^{-1}$ )	$\gamma$	Short description
Kolmogorov (1941)	---	—	—	—	1/3	• Kolmogorov’s law for ideal incompressible turbulence.
Kraichnan (1974)	---	—	—	—	1/2	• Compressible supersonic shock-dominated turbulence (Burger’s turbulence).
Larson (1979, 1981)	Ⓐ	0.38	—	1.1	0.38	• Molecular clouds, clumps, cores, H II regions, mapped by different tracers.
Solomon et al. (1987)	Ⓒ	$0.50 \pm 0.05$	—	$1.0 \pm 0.1$	$0.50 \pm 0.05$	• $^{12}\text{CO}$ emission from homogeneous sample of 273 clouds.
Heyer & Schloerb (1997)	Ⓐa	—	$0.43 \pm 0.04$	$1.23 \pm 0.08$	$0.24 \pm 0.10$	• $^{12}\text{CO}$ emission of cloud complex Sh 155.
Heyer & Schloerb (1997)	Ⓐb	—	$0.55 \pm 0.03$	$0.78 \pm 0.05$	$0.42 \pm 0.09$	• $^{12}\text{CO}$ emission of cloud complex Sh 235.
Heyer & Brunt (2004)	ⒶH	—	$0.65 \pm 0.01$	$0.87 \pm 0.02$	$0.57 \pm 0.07$	• $^{12}\text{CO}$ emission from 27 molecular clouds.
Heyer et al. (2006)	Ⓐa	—	$0.74 \pm 0.04$	$0.73 \pm 0.03$	$0.71 \pm 0.12$	• Rosette cloud complex as a whole.
Heyer et al. (2006)	Ⓐb	—	$0.79 \pm 0.06$	$1.00 \pm 0.04$	$0.79 \pm 0.04$	• Rosette cloud complex, zone I: inside ionization front, feedback from H II region, using $^{12}\text{CO}$ (similar for $^{13}\text{CO}$ ).
Heyer et al. (2006)	Ⓐc	—	$0.66 \pm 0.06$	$0.70 \pm 0.03$	$0.59 \pm 0.14$	• Rosette cloud complex, zone II: outside ionization front.
Bolatto et al. (2008)	Ⓐ4	$0.60 \pm 0.10$	—	$0.76 \pm 0.27^a$	$0.60 \pm 0.10$	• Molecular clouds in extragalactic systems, using $^{12}\text{CO}$ .
Roman-Duval et al. (2011)	---	—	$0.62 \pm 0.20$	—	$0.53 \pm 0.35$	• Average from 368 molecular clouds, with a resolution limit of 1 FWHM (48 arcsec), using $^{13}\text{CO}$ emission.
Hacar et al. (2016)	ⒶM	—	—	$0.66^a$	0.58	• Musca cloud as a whole, isolated from stellar sources of feedback, using $^{13}\text{CO}$ and $\text{C}^{18}\text{O}(2-1)$ emission.
Rice et al. (2016)	---	$0.49 \pm 0.04$	—	$0.66 \pm 0.09^a$	$0.49 \pm 0.04$	• Dendrogram-based catalogue of 611 outer Galaxy clouds.
Rice et al. (2016)	---	$0.52 \pm 0.03$	—	$0.87 \pm 0.09^a$	$0.52 \pm 0.03$	• Dendrogram-based catalogue of 428 inner Galaxy clouds.
Miville-Deschênes, Murray & Lee (2017)	---	$0.63 \pm 0.30$	—	$0.83^a$	$0.63 \pm 0.30$	• $^{12}\text{CO}$ catalogue of 8107 clouds in the Galactic plane.
Traficante et al. (2018b)	---	$0.09 \pm 0.04$	—	—	$0.09 \pm 0.04$	• Gravity-dominated regions, core and clump scales.
Federrath et al. (2010a)	---	—	$0.66 \pm 0.05$	—	$0.59 \pm 0.13$	• Simulations with purely solenoidal forcing ( $\nabla \cdot f = 0$ ).
Federrath et al. (2010a)	---	—	$0.76 \pm 0.09$	—	$0.74 \pm 0.19$	• Simulations with purely compressive forcing ( $\nabla \times f = 0$ ).
Klessen & Hennebelle (2010)	Ⓐ5	—	—	0.8	0.5	• Simulations with continuous accretion of diffuse material.
Bertram et al. (2014)	---	—	$0.82 \pm 0.03$	—	$0.83 \pm 0.11$	• Non-isothermal simulations of chemically evolving clouds. Values for $^{12}\text{CO}$ emission and mean density $n = 300 \text{ cm}^{-3}$ .
Padoan et al. (2017)	Ⓐ6	—	—	$0.82^{a,b}$	$0.5 \pm 0.1$	• Simulated clouds with supernovae-driven turbulence.

Notes. <sup>a</sup>Converted from 1D to 3D velocity dispersion assuming isotropy. <sup>b</sup>Extracted manually from their fig. 4. We provide a graphical version of this summary in Fig. A1.

it is straightforward to show from equation (7) that  $\gamma_p$  is constant ( $\gamma_p = \gamma_2 \equiv \gamma$ ) if and only if  $\zeta_p$  depends linearly on  $p$  ( $\zeta_p = p\gamma$ ). This is not the case for intermittent velocity fluctuations, in which the exponent  $\gamma_p$  is no longer constant but depends on the order  $p$  of the function. Furthermore, they concluded on intermittent fields that the PCA-derived exponent  $\alpha$ , after calibration, is better correlated with structure function exponents of orders  $p = 1/2, 1$ , which means that the translation given by the combination of equations (5) and (12) is preferentially a measure of the intrinsic scaling exponent  $\gamma_p$  of low order rather than rms velocity fluctuations. Qualitatively, fields that are intermittent exhibit extreme differences in the magnitude of velocity fluctuations and/or density distribution over a given spatial scale (Brunt et al. 2003), which is why multiple structure functions would be needed to fully describe such a scenario.

#### 4.2.2 Universality of turbulence in the molecular ISM

Based on earlier works, Larson (1981) collated 3D rms velocity dispersions along with projected sizes  $L$  (in pc) of varied 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  $\sim 1000$  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  $\Delta v = (1.0 \pm 0.1)L^{0.5 \pm 0.05}$  (see

also Falgarone, Pety & Hily-Blant 2009, for a compilation of data covering five orders of magnitude in spatial scale). For the velocity dispersion calculation, they extracted centroid and line-width-based velocity differences and added them in quadrature.

These relationships describe multiple evaluations of the structure function in particular cases  $S_2(l = L)$ , which, when combined, turn out to follow the same functional form as if they were part of a single fluid, despite the fact that most of the objects analysed were born far from each other and do not interact.

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 and self-similarity over different scales of the molecular ISM in 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:

$$\begin{aligned} C &\approx v_0 \\ \Gamma &\approx \gamma_2. \end{aligned} \quad (13)$$

For standardization purposes, we use this turbulence universality condition and the second-order calibrations (equations 5 and 12) to translate, respectively, Larson-like exponents  $\Gamma$  and PCA-derived exponents  $\alpha$  to rms scaling exponents  $\gamma_2$ , which we will simply call  $\gamma$  from now on.



#### 4.2.3 Optical depth effects on PCA pseudo-structure functions

Larson (1981) considered optically thin regions, mostly traced by  $^{13}\text{CO}$ ,  $\text{H}_2\text{CO}$ , and  $\text{NH}_3$ , and only two  $^{12}\text{CO}$  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 from line saturation. However, numerical and observational studies have concluded that PCA-derived scales are nearly insensitive to the opacity regime of the emission (Heyer & Schloerb 1997; Brunt 2003; Brunt & Mac Low 2004; Heyer et al. 2006; Brunt, Heyer & Mac Low 2009; Roman-Duval et al. 2011). This is a strong point of the method as it allows us to study a broad 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 (see also the discussion in Bertram et al. 2014).

In Section 6.2, we discuss optical depth effects by comparing the retrieved scaling parameters from cloud complexes at different orientations.

### 5 GENERAL METHODOLOGY

We implement three PCA extraction methods to investigate the response of the retrieved pseudo-structure functions to different analysis scales:

(i) The Mixed method: This method consists of deriving PCA scales ( $l$ ,  $\delta v$ ) from small (30 pc) cloud portions,<sup>9</sup> which are semi-automatically extracted from the simulated cloud complexes ( $\sim 200$  pc). The resulting scales are then combined to construct a single fit equation (often called ‘Mixed fit’) that represents the pseudo-structure function of the complex. This is the method that we use the most throughout the work.

(ii) The Complex method: In this method, we analyse the cloud complexes as they are, without any sub-portioning.

(iii) The Individual Cloud method: For this method, we do not combine the PCA-derived scales from the individual (30 pc) portions but rather use them to compute their pseudo-structure functions separately, as individual objects. This will allow us to do further statistics, especially to analyse cloud environmental effects on the retrieved scaling parameters. Henceforth, we refer to individual portions interchangeably as ‘individual clouds’ or just ‘clouds’.

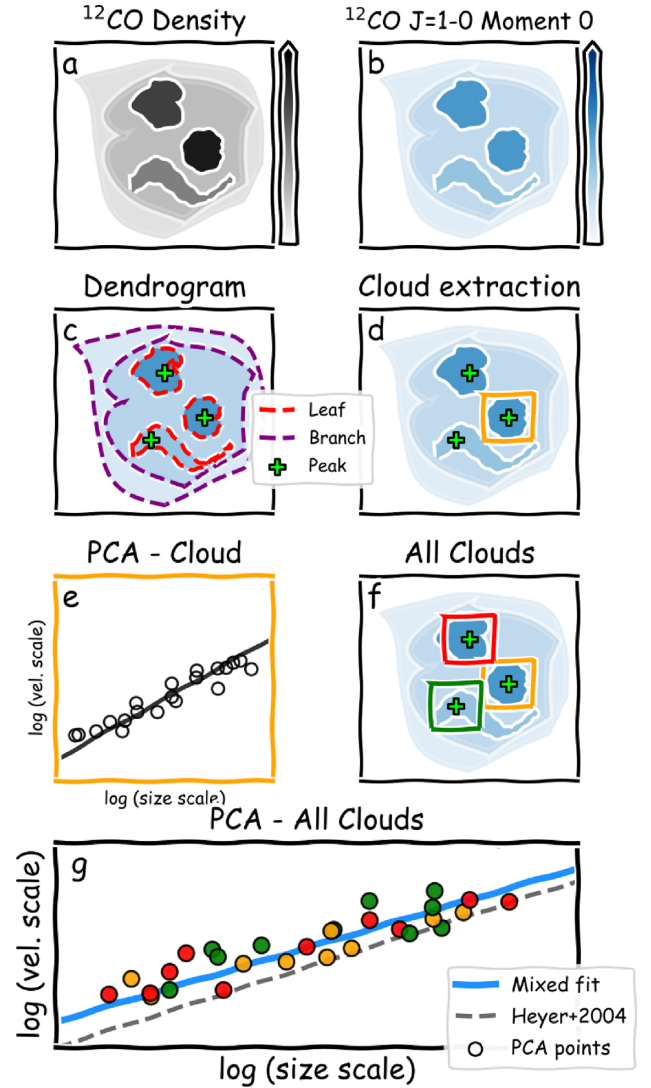
The Complex method is straightforward to implement. We take the intensity cubes obtained from the radiative transfer calculations and pass them directly through the TURBUSTAT.PCA module to compute spatial ( $l$ ) and velocity scales ( $\delta v$ ) from our cloud complexes, as single objects. The Mixed and Individual Cloud methods require, however, further steps that we explain below in detail.

First, we determine zeroth moment maps using the cloud complex intensity cubes. The zeroth moment is defined as the integrated intensity along the velocity (or frequency) axis  $\psi$  as follows:

$$M_0(x, y) = \int_{\psi} T(x, y; \nu) d\nu, \quad (14)$$

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

Then, we use the ASTRODENDRO<sup>10</sup> package with each zeroth moment map as input to compute a hierarchical tree, called a



**Figure 2.** Basic workflow to retrieve PCA pseudo-structure functions of cloud complexes formed in our Cloud Factory simulation suite. This figure illustrates the Mixed PCA extraction method.

dendrogram, which divides the map into closed sub-regions denoted by equal-valued pixels.

From the computed dendrograms, we focus on extracting the innermost, irreducible, sub-structures of the region, called ‘leaves’, which are usually within larger scale ‘branches’ that may also be part of more extended structures (see panel c in Fig. 2). To construct the dendrograms, we set the three main parameters,  $\text{min\_value} = 0.1 \bar{M}_0$ ,  $\text{min\_delta} = \bar{M}_0$  and  $\text{min\_npix}$  between 100 and 400, where  $\bar{M}_0$  is the mean value of the zeroth moment. For reference, typical values for complex D are  $\text{min\_value} = 9.3$ ,  $\text{min\_delta} = 92.9$ , and  $\text{min\_npix} = 300$ . Next, we find peak pixels in ‘leaves’ and centre  $(30 \text{ pc})^2$  portions on them that we use afterwards to slice the original  $^{12}\text{CO}$  cubes. Ideally, these portions (or individual clouds) should cover the cloud complex moment map as much as possible so that the least intensity is left out of the analysis, regardless of the velocity channel. In order to achieve this, we set the dendrogram parameters in such a way that the ‘leaves’ are enough in number and sufficiently separated from each other. Finally, PCA is computed on individual clouds to construct the Mixed and Individual Cloud fits. Fig. 2 is a schematic representation of this procedure.

<sup>9</sup>This size is typical of giant molecular clouds, used as individual objects for PCA in, e.g. Heyer & Brunt (2004).

<sup>10</sup><http://www.dendrograms.org>



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 consists of the mean scaling parameters ( $v_0$ ,  $\alpha$ ) obtained from the  $N$  random fits, and the errors correspond to the standard deviations ( $\sigma_{v_0}$ ,  $\sigma_\alpha$ ) from the mean values. We also report the associated coefficient of determination,  $R^2$ , and the reduced chi-square,  $\chi^2$ , to evaluate the goodness of the fits. This is all recorded in the supplementary data base provided with the manuscript.

Despite the fact that we assign a single power-law structure function to each cloud complex, all of them can be seen as a composition of different power laws, each of which is associated with a different portion of the complex (one per colour in Fig. 4). The structure functions extracted from individual clouds (e.g. Fig. 7, bottom row) are all collected in the Supplementary file. The dispersion of individual cloud scaling parameters is related to the level of intermittency of the hosting complex (see Section 4.2.1).

This pipeline and analysis tools are all built-in and executed by our PCAFACTORY package.

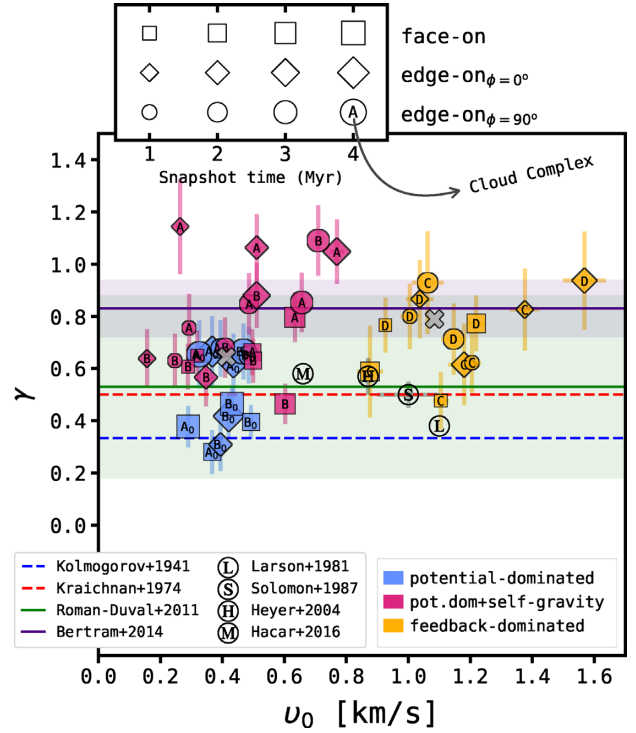
## 6 RESULTS

In this section, we present the PCA of a set of cloud complexes extracted from our Cloud Factory simulation suite using  $^{12}\text{CO}$  J = 1 – 0 intensity PPV cubes. We report the impact on velocity structure functions after adding physical mechanisms such as gas self-gravity or supernova feedback on star-forming sites (Section 6.1). Likewise, we present variations driven by different line-of-sight projections (Section 6.2) and time snapshots (Section 6.3). Lastly, we investigate cloud environmental effects and explore variations in velocity fluctuations when applying PCA locally on individual molecular clouds or globally on full cloud complexes (Section 6.4). We study six cloud complexes (two per physical scenario) as summarized in Table 1, but since we consider (three) different lines-of-sight and (two/three) time snapshots, there are in practice 42 different objects: 12 for the potential-dominated scenario with no self-gravity; 18 for the potential-dominated scenario with self-gravity, for which we extracted an extra time snapshot to study the time evolution of clouds under local gravitational effects (see Section 6.3); and 12 for the feedback-dominated scenario.

For reasons of space, the corresponding PCA figures and line emission profiles from all the cloud complex configurations can be found in the supplementary file provided along with this manuscript.<sup>11</sup> Additionally, the whole catalogue of ( $l$ ,  $\delta v$ ) and ( $v_0$ ,  $\alpha$ ) pairs extracted from individual clouds and cloud complexes is also available online. The software developed to carry out this work, the PCAFACTORY, is open source and available on GitHub.<sup>12</sup>

### 6.1 Structure function dependence on the physical scenario

In Fig. 3, we present the structure function scaling parameters ( $v_0$ ,  $\gamma$ ) derived for all our cloud complexes. The scaling parameters were retrieved using the Mixed PCA extraction method (see Section 5), which combines PCA-derived scales ( $l$ ,  $\delta v$ ) of smaller cloud portions to find the associated pseudo-structure function of each cloud



**Figure 3.** Structure function parameters ( $v_0$ ,  $\gamma$ ) of cloud complexes extracted from different physical scenarios (marker colour), orientations (marker style), and time snapshot (marker size) after tracer refinement has commenced. The parameters are derived from PCA pseudo-structure functions using the Mixed method (see Section 5 and Fig. 2). The grey crosses are the best cluster centres obtained from a K-means clustering analysis. The empty circles and the horizontal lines, with errors as the shades, are selected parameters from previous literature (see Table 2).

complex. The translation from PCA-derived scaling exponents ( $\alpha$ ) into the corresponding structure function exponents ( $\gamma$ ) assumes the second-order calibrations given by equations (5) and (12). The ( $v_0$ ,  $\gamma$ ) pairs are computed for different line-of-sight projections and time snapshots for each cloud complex. For comparison, we include representative observational and synthetic values reported in previous literature as 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. To assess this, we perform a K-means clustering analysis along with a Silhouette Coefficient score test that suggests that the parameter distribution is better represented by two clusters (score = 0.56), followed by three clusters (score = 0.51). The cluster centres are also shown in Fig. 3. Overall, the structure function scaling exponents ( $\gamma$ ) 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 alone is not sufficient to fully determine the physical nature of cloud complexes. Scaling coefficients ( $v_0$ ) must also be taken into account to effectively distinguish between clouds.

We summarize in Table 3 the mean and standard deviations of scaling parameters derived from all the cloud complex configurations, using three PCA extraction methods detailed in Section 5. For the potential-dominated case with no self-gravity, the mean scaling exponent  $\bar{\gamma} = 0.51 \pm 0.15$  corresponds to that of the classical Burger-like turbulence ( $\gamma = 1/2$ ). Turning on self-gravity

<sup>11</sup> <https://github.com/andizq/andizq.github.io/tree/master/pcafactory-data>

<sup>12</sup> <https://github.com/andizq/pcafactory>

**Table 3.** Mean and standard deviation of structure function parameters derived from different PCA extraction methods. The mean is taken over all time snapshots and lines of sight. Conventions: a. Potential dominated; b. Potential dominated with self-gravity; c. Feedback dominated.

PCA extraction method	Physical scenario	Cloud complex	$\overline{v}_0$ (km s <sup>-1</sup> )	$\overline{\gamma}$
Mixed	a	A <sub>0</sub> & B <sub>0</sub>	0.41 ± 0.06	0.51 ± 0.15
	b	A & B	0.46 ± 0.17	0.78 ± 0.19
	c	C & D	1.14 ± 0.18	0.74 ± 0.14
Complex	a	A <sub>0</sub> & B <sub>0</sub>	0.31 ± 0.05	0.48 ± 0.15
	b	A & B	0.36 ± 0.14	0.74 ± 0.19
	c	C & D	1.13 ± 0.20	1.02 ± 0.18
Individual clouds	a	A <sub>0</sub>	0.36 ± 0.09	0.64 ± 0.26
	a	B <sub>0</sub>	0.48 ± 0.11	0.64 ± 0.31
	b	A	0.49 ± 0.20	1.01 ± 0.26
	b	B	0.45 ± 0.22	0.82 ± 0.35
	c	C	1.37 ± 0.55	0.84 ± 0.48
	c	D	1.20 ± 0.46	0.85 ± 0.29

increases scaling exponents to  $\overline{\gamma} = 0.78 \pm 0.19$  while maintaining low scaling coefficients  $\overline{v}_0 = 0.46 \pm 0.17$  km s<sup>-1</sup>. On the other hand, feedback-dominated complexes show similar exponents but suffer a substantial increase in scaling coefficients to  $\overline{v}_0 = 1.14 \pm 0.18$  km s<sup>-1</sup>.<sup>13</sup> Additionally, we show in Table 4 a comparison between PCA parameters derived from <sup>12</sup>CO and <sup>13</sup>CO J = 1 – 0 cubes for some cloud complexes and orientations. We assume that the <sup>13</sup>CO abundance is a factor of 69 lower than that of <sup>12</sup>CO (see Visser, van Dishoeck & Black 2009). Despite the contrast in opacity between the two tracers, no significant differences are found for scaling exponents. This indicates that both tracers are valid descriptors of cloud kinematics as long as velocity fluctuations are supersonic (generally satisfied), which softens opacity effects. We find that even in edge-on orientations, line emission profiles from small cloud portions can span over a wide range of velocities and exhibit substructure suggesting that the emission can still arise from different depths in the cloud (see e.g. Fig. A4). Scaling coefficients, however, decrease systematically in the case of <sup>13</sup>CO, which we attribute to the fact that <sup>13</sup>CO lines are less extended in velocity channels and hence the retrieved fluctuations decrease for all size scales.

For illustration, in Fig. 4 we show the PCA relationships derived for three cloud complexes, B<sub>0</sub>, B, and D, each from a different physical scenario. All of them are seen face-on and extracted at similar time-steps after starting tracer refinement so that the effects of varying physical scenario prevail. The PCA already hints at differences in pseudo-structure functions ( $\delta v$ ) from case to case, increasing both the scaling exponent ( $\gamma$ ) and coefficient ( $v_0$ ) as we add more physical processes in the simulation. However, if we look at the full set of results for  $\gamma$  shown in Fig. 3, which includes values for the other cloud complexes and other projections, we see that while there is still a clear sign that including self-gravity increases  $\gamma$ , the difference between the feedback- and (both) potential-dominated cases is less clear.

The coloured circles in the left-hand panel of Fig. 4 are the PCA-derived scales ( $l, \delta v$ ) from selected cloud portions correspondingly illustrated as coloured squares in the right-hand panel, overlaid on the <sup>12</sup>CO J = 1 – 0 zeroth moment map of the cloud complex. The crosses in the left-hand panel are the PCA-derived scales from the

cloud complex as a whole. The pie chart shows the number of PCA scales extracted from each cloud portion with the total net scales in the middle. Yet, a cloud portion is subject to rejection so long as (a) the centre of its associated window is enclosed by another window or (b) it has no scales retrievable by the PCA. This is shown by the coloured round boxes at the bottom of the left-hand panel. The blue line is the associated fit to the combined points from cloud portions (Mixed method) and the pink line is the fit to the points derived from the cloud complex as a whole (Complex method). The blue- and pink-shaded regions are the uncertainties of these fits (see details in Section 5). For reference, the dashed grey line is the best fit found by Heyer & Brunt (2004) using PCA-derived scales from 27 molecular clouds, which, in this case, most closely resembles the feedback-dominated cloud complex D.

Also, note that the number of PCA points is more sensitive to higher zeroth moment values than to larger <sup>12</sup>CO flux extents, as one might think at first. This is because the zeroth moment is related to intensity variance along the velocity axis. If density fluctuations and velocity intermittence in the region are low, a high-valued pixel in zeroth moment spans into a wide range of values in the covariance matrix of the line cube. Thus, it is equivalent to the amount of information that principal components can keep and hence the PCA algorithm will raise more ( $l, \delta v$ ) pairs. Furthermore, the number of pairs usually depends on the characteristic scaling exponent of the region. There are fewer pairs (per number of cloud portions) for cloud complexes with lower exponents (see also Brunt & Heyer 2002), which is in turn related to the diagonality of the covariance matrix, driven by a typically higher degree of line-centroid variation in higher exponent velocity fields (see e.g. fig. 1 of Brunt & Heyer 2013).

On the other hand, the fact that the velocity scaling coefficients are generally low for complexes A and B, at all times, reflects a high level of velocity coherence due to weak stellar feedback in 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 weak 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 Paper I.

Conversely, clustered supernova feedback plays a significant role in taking cloud complexes out of that coherent state. As previously found in Paper I, supernovae tied to star formation sites make

<sup>13</sup>We warn the reader that the reported  $\gamma$  should be used with caution as it more reliably traces low-order fluctuations (rather than rms velocities) when the region is intermittent (see Section 4.2.1).

**Table 4.** Comparison between PCA-derived scaling parameters from LVG  $^{12}\text{CO}$  and  $^{13}\text{CO}$   $J = 1 - 0$  cubes for different cloud complexes and orientations, using the Mixed extraction method. We also show the distribution of parameters for both isotopologues in the LVG panel of Fig. A6.

Complex	Time-step (Myr)	Orientation	$^{12}\text{CO}$		$^{13}\text{CO}$	
			$v_0$ (km s $^{-1}$ )	$\alpha_{\text{PCA}}$	$v_0$ (km s $^{-1}$ )	$\alpha_{\text{PCA}}$
A <sub>0</sub>	2	face-on	$0.30 \pm 0.01$	$0.66 \pm 0.03$	$0.16 \pm 0.01$	$0.57 \pm 0.05$
A <sub>0</sub>	2	edge-on- $\phi=90^\circ$	$0.49 \pm 0.01$	$0.69 \pm 0.03$	$0.27 \pm 0.01$	$0.79 \pm 0.05$
B	3	face-on	$0.77 \pm 0.01$	$0.61 \pm 0.02$	$0.60 \pm 0.02$	$0.73 \pm 0.03$
B	3	edge-on- $\phi=90^\circ$	$0.88 \pm 0.03$	$0.87 \pm 0.05$	$0.49 \pm 0.02$	$0.93 \pm 0.05$
C	1	face-on	$0.83 \pm 0.03$	$0.75 \pm 0.07$	$0.59 \pm 0.03$	$0.74 \pm 0.09$
C	1	edge-on- $\phi=90^\circ$	$0.91 \pm 0.04$	$0.94 \pm 0.09$	$0.79 \pm 0.04$	$0.88 \pm 0.10$
D	1	face-on	$0.79 \pm 0.02$	$0.88 \pm 0.04$	$0.79 \pm 0.03$	$0.99 \pm 0.05$
D	1	edge-on- $\phi=90^\circ$	$1.04 \pm 0.04$	$0.92 \pm 0.06$	$0.87 \pm 0.04$	$0.90 \pm 0.07$

complexes C and D have shorter and less massive filaments. This type of feedback also induces stronger shear that reduces filament lifetimes and consequently also their star formation efficiencies. In this work, this manifests as high scaling coefficients, which indicates a larger degree of velocity fluctuations in complexes C and D. Again, this is 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 valuable information about the cold dense molecular ISM.

In summary,

- (i) cloud complexes governed by different physical mechanisms produce separate clusters of points in the  $(v_0, \gamma)$  parameter space (see e.g. Fig. 3).
- (ii) the scaling coefficient  $v_0$  is an excellent reference to distinguish between potential-dominated and feedback-dominated cloud complexes, whereas the role of  $\gamma$  is only discernible when self-gravity is turned off (see Fig. 3 and Table 3).
- (iii) Low velocity scaling coefficients are associated with (quiescent) regions with coherent velocity fields, which favours the development and preservation of long filamentary structures. Conversely, strong stellar feedback takes cloud complexes out of their coherent state, producing smaller and short-lived structures.

## 6.2 Structure function: cloud orientation effects

We also vary the line-of-sight projection of cloud complexes during the radiative transfer calculations to explore the influence of density/velocity anisotropies and optical depth effects on the retrieved structure functions. As shown by the different markers in Fig. 5, there are no systematic variations in scaling parameters when changing the orientation of cloud complexes; they rather depend on the particular geometry of each complex and hence on their evolution over time.

Edge-on projections of cloud complexes with low scale heights (A<sub>0</sub>, B<sub>0</sub>, A, and B) exhibit flat and continuous CO distributions but actually consist of both nearby and distant structures that comprise each cloud complex as a whole (see Figs 1 and A3). Background emission is hence susceptible to being blocked by foreground gas with similar line-of-sight velocities if optically thick tracers are used as in our case with  $^{12}\text{CO}$ . However, opacity effects are softened by the fact that velocity fluctuations are generally supersonic and, in consequence, part of the background emission can still reach the observer.

Projection effects for A<sub>0</sub> and B<sub>0</sub>, which are our least realistic cases, seem to be driven by column density variations. As shown in Fig. 5, complex A<sub>0</sub> 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).

Cloud complex B<sub>0</sub> does not follow the same pattern. In this case, variations induced by line-of-sight projections are smaller. This is in good agreement with our previous interpretation because gas column densities are this time nearly the same for all projections, with differences smaller than 30 per cent between face-on and edge-on column densities.

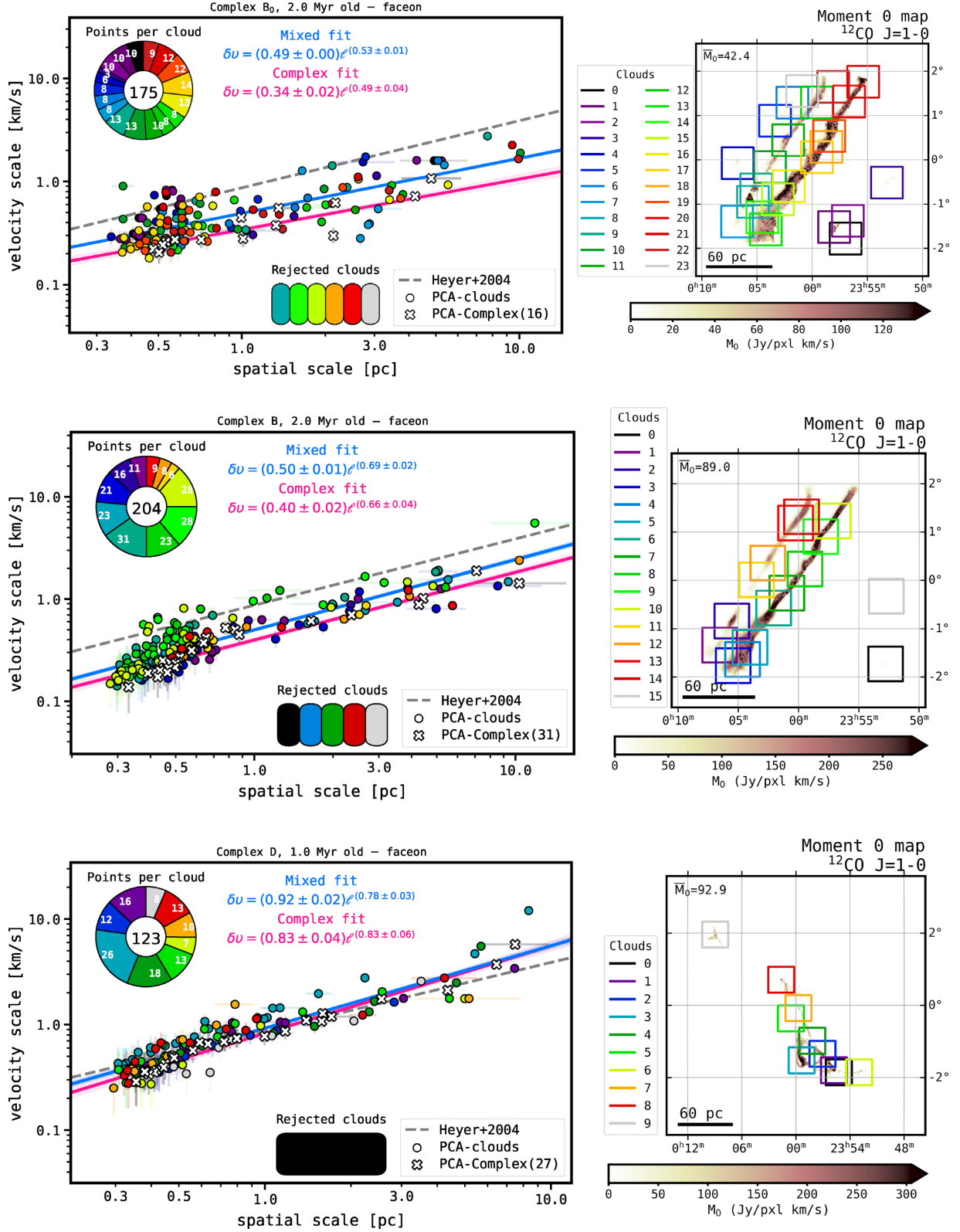
When compared to PCA applied over cloud complexes as a whole, the variability of complex B<sub>0</sub> decreases even more (see Fig. A2). For B<sub>0</sub>, all the scaling parameters lie around  $v_0 = 0.35$  km s $^{-1}$  and  $\gamma = 0.4$  regardless of orientation. This suggests that variations in scaling parameters in this complex are due to individual cloud velocity fluctuations driven either by local collisions or random supernova feedback. On the other hand, cloud complex A<sub>0</sub> 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 to edge-on values is roughly the same. Both results strengthen our argument that larger column densities yield steeper scaling exponents, and equivalently, isotropic column densities result in similar scaling parameters. We emphasize, however, that these are small variations that in no way resemble the variations induced by changing feedback conditions.

Complexes A and B show similar signatures but their parameters are strongly affected by local gravitational forces, displaying rather systematic variations over time. Hence, we prefer to leave this for Section 6.3 where cloud complexes are analysed from their evolutionary context.

For feedback-dominated complexes, C and D, the edge-on  $\phi=0^\circ$  orientation (parallel to the  $\phi$ -axis) produces high scaling exponents and the highest scaling coefficients in most of the cases. This is because for both complexes the axes of filamentary structures are preferentially stretched out along the  $\phi$ -direction due to differential rotation of the Galaxy. This makes the edge-on  $\phi=0^\circ$  projection contain more gas mass within smaller projected scales than the other orientations. 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 as optical depth effects are softened (see Fig. A5).

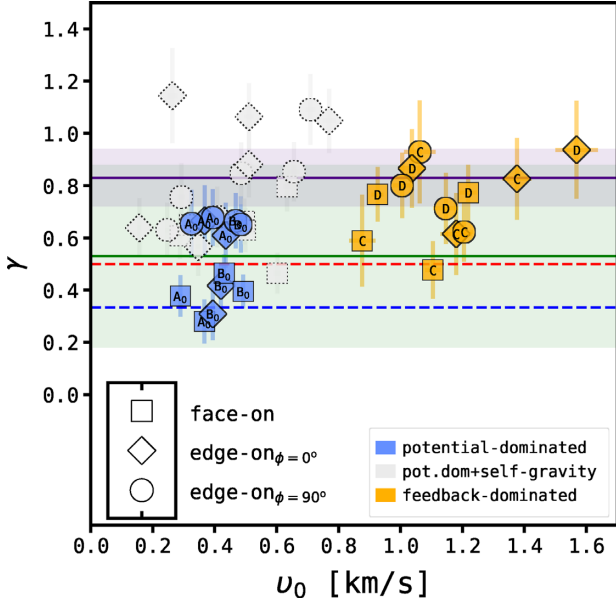
In summary,

- (i) in potential-dominated complexes without self-gravity, higher column densities lead to steeper scaling exponents. For comparable column densities, variations in scaling parameters are smaller and



**Figure 4.** *Left column:* Velocity structure function from the face-on view of cloud complexes B<sub>0</sub> (top), B (middle), and D (bottom) at selected time snapshots. The coloured circles are the PCA-derived scales ( $\ell, \delta v$ ) from the same-coloured (30 pc)<sup>2</sup> clouds on the right-hand panel. The white crosses are the PCA-derived scales from the cloud complex as a whole. The pie chart indicates the number of scales extracted from each cloud and the total number of scales in the middle. Overlapping cloud portions are rejected and shown in colours at the bottom of the left-hand panels. For reference, the dashed line is the PCA-derived best fit found by Heyer & Brunt (2004) from an ensemble of 27 molecular clouds. The blue and fuchsia lines are, respectively, the fit from individual clouds simultaneously (Mixed method) and the fit to the points from the cloud complex as a whole (Complex method). *Right column:* Cloud portions extracted for PCA, overlaid on the synthetic <sup>12</sup>CO J = 1 – 0 zeroth moment maps of the cloud complexes.





**Figure 5.** Structure function parameters ( $v_0, \gamma$ ) of cloud complexes extracted from different physical scenarios (marker colour) and orientations (marker style) using the Mixed method. The solid lines (and their errors as shades) are reference literature values whose legends can be found in Table 2. The markers in grey are potential-dominated self-gravitating complexes studied separately in Section 6.3 as a function of time.

driven by localized cloud-scale disturbances (e.g. cloud collisions, isolated supernovae).

(ii) in feedback-dominated complexes, where optical-depth effects are less prominent, projections with larger column densities yield higher scaling parameters. Especially for edge-on  $\phi=0^\circ$  lines of sight, which are parallel to the long axis of filamentary structures.

### 6.3 Structure function: time evolution

Fig. 6 follows the evolution of structure function parameters of cloud complexes for each physical scenario and orientation. Potential-dominated complexes with no self-gravity yield stable scaling exponents ( $\gamma$ ) over time but present a systematic reduction of velocity scaling coefficients ( $v_0$ ). This suggests a sustained lack of turbulence fueling sources to compensate the natural accelerated energy decay over length-scales characteristic of Kolmogorov-like fluids (see e.g. Onsager 1949), which, on small scales, rapidly lose energy via viscous dissipation. Also, this indicates that in these complexes, driving of large-scale velocity fluctuations by random supernovae is unable to maintain the level of turbulence within the clouds (see also Ibáñez-Mejía et al. 2017 and Seifried et al. 2018 who find similar results).

Turning on self-gravity dramatically changes this behaviour. In this case, the scaling coefficient (or magnitude of velocity fluctuations,  $v_0$ ) increases with time for complexes A and B. The gradient of this increment appears linear and is similar for both complexes, which allows us to capture the time evolution of velocity scaling coefficients with the following relationship:

$$v_0(t) = 0.07 + 0.19 \left( \frac{t}{\text{Myr}} \right) [\text{km s}^{-1}]. \quad (15)$$

As expected, we also find a (nearly linear) sustained increase in mean surface densities over time. For instance, cloud complex B yields mean surface number densities  $\Sigma_H(t = 1 \text{ Myr}) = 7.65 \times 10^{19} \text{ cm}^{-2}$ ,  $\Sigma_H(t = 2 \text{ Myr}) = 1.07 \times 10^{20} \text{ cm}^{-2}$ , and  $\Sigma_H(t = 3 \text{ Myr}) = 1.44 \times$

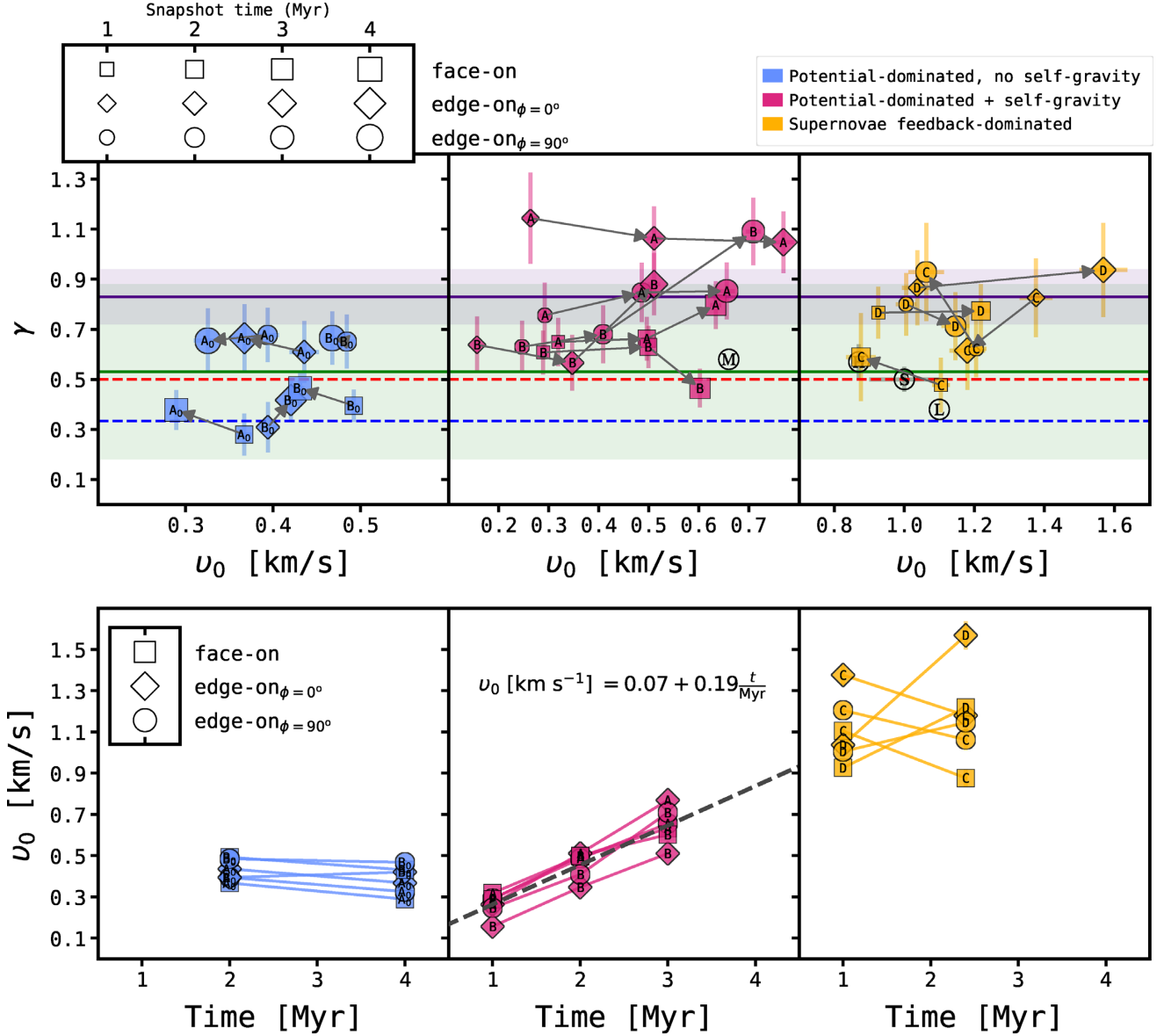
$10^{20} \text{ cm}^{-2}$ . This is in excellent agreement with the dependence of  $v_0$  on the cloud surface density found by Heyer et al. (2009), and suggests that the physics behind such a dependence is dominated by local gravitational interactions.

Conversely, the evolution of scaling exponents  $\gamma$  is less predictable, which might indicate that local gravitational forces trigger velocity fluctuations across a wide range of spatial scales. To assess this, we zoomed into cloud complex B to track the evolution of individual filament scaling parameters as a response to local gravitational effects (see Fig. 7). At  $t = 1 \text{ Myr}$ , when there are still no sink particles, individual cloud parameters cluster around a common zone in the ( $v_0, \gamma$ ) space. Overall, scaling exponents are steeper ( $\bar{\gamma} = 0.75$ ) and velocity fluctuations weaker ( $\bar{v} = 0.28 \text{ km s}^{-1}$ ) compared to those of the potential-dominated case with no self-gravity ( $\bar{\gamma} = 0.51$ ,  $\bar{v} = 0.42 \text{ km s}^{-1}$ ). We propose that this is a consequence of global (large-scale) collapse processes that reduce the degree of turbulence in the region by increasing coherence.

Later, at  $t = 2 \text{ Myr}$ , there is a burst of star formation in the long dense filament ( $N_{\text{sinks}} > 100$ ) but none in the diffuse filament yet. This produces a clear separation of individual cloud parameters coming from each filament as shown in the middle panel of Fig. 7. The PCA scaling exponents of the short diffuse filament are this time shallower because gravitational fragmentation has commenced. This favours small-scale interactions in the region and is likely to be a signature of pre-core stages in molecular clouds. The long dense filament skipped this phase as it formed several individual clumps/cores much faster. Such a rapid emergence of stellar systems drastically lowers velocity fluctuations on small scales and, hence, increases the scaling exponent ( $\gamma$ ). This ‘burst’ of new cores also establishes multiple point-like centres of collapse affecting inter-core gas predominantly on intermediate and large scales. For reference, the region in the long filament where sink particles are more numerous has a mean sink separation of  $\approx 0.5 \text{ pc}$  (see Fig. 1). As a side note, a consequence of our sink particle implementation is that there is some missing information from gas velocities on small scales because it reduces the bound gas mass around the cores to point-mass gravitating particles. Hence, in real observations, shallower (but still high) scaling exponents should be found in core stages because gas from cores/clumps do still contribute to small-scale velocity fluctuations in the region via line broadening.

Finally, at  $t = 3 \text{ Myr}$ , the short filament commences the formation of cores at its upper tip. For that particular cloud (cld 13 at  $t = 2 \text{ Myr}$ , cld 12 at  $t = 3 \text{ Myr}$ ), both scaling parameters increase and favour the zone occupied formerly by clouds from the long filament, but not as high as them, because of the much lower core formation rates. As expected, cloud portions with no cores retain the same  $\gamma$ , but with higher  $v_0$ , because fragmentation is still carrying on. Both parameters continue to increase for the long filament 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 to the same zone where clouds from both of the filaments started at  $t = 1 \text{ Myr}$ , which strengthen the idea that quiescent clouds have gravity-driven time-dependent trajectories in the ( $v_0, \alpha$ ) space. This behaviour also holds for cloud complex A, namely individual cloud scaling parameters also respond differently depending on local core formation stages. Note that both fragmentation processes in pre-core stages, as well as in subsequent formation of cores, continuously lead to an increase in the magnitude of velocity fluctuations ( $v_0$ ) across the cloud complex.

Feedback-dominated cloud complexes evolve according to their particular environment. As seen in Fig. 6, the common pattern is that complex C migrates towards lower scaling coefficients  $v_0$ , whereas



**Figure 6.** Top row: Structure function parameters split by physical scenario to illustrate variations over time. The arrows show the evolution of cloud complex parameters for a given line of sight. Literature values are shown for reference; marker codes and descriptions are summarized in Table 2. Bottom row: time evolution of scaling coefficients  $v_0$  for each cloud complex and orientation. The dashed line in the middle panel is the linear best fit obtained for complexes in the potential-dominated scenario with self-gravity.

complex D moves towards higher  $v_0$ . Local gravitational effects do not play the same role in these complexes as in quiescent regions because mass instabilities are this time more difficult to reach. In complex C, external feedback from supernova explosions set large turbulence driving scales that decay over time through smaller scales as there are no internal fueling sources of turbulence. Other mechanisms of internal feedback (beyond the scope of this work) such as winds or photoionizing radiation could potentially help inject fresh energy into the cloud (see e.g. Peters et al. 2017). In complex D, however, energy re-injection from internal supernovae seems to sustain and increase the level of turbulence through expanding supernova bubbles.

In summary,

(i) random supernovae alone are unable to sustain turbulence in molecular clouds at a level consistent with observations.

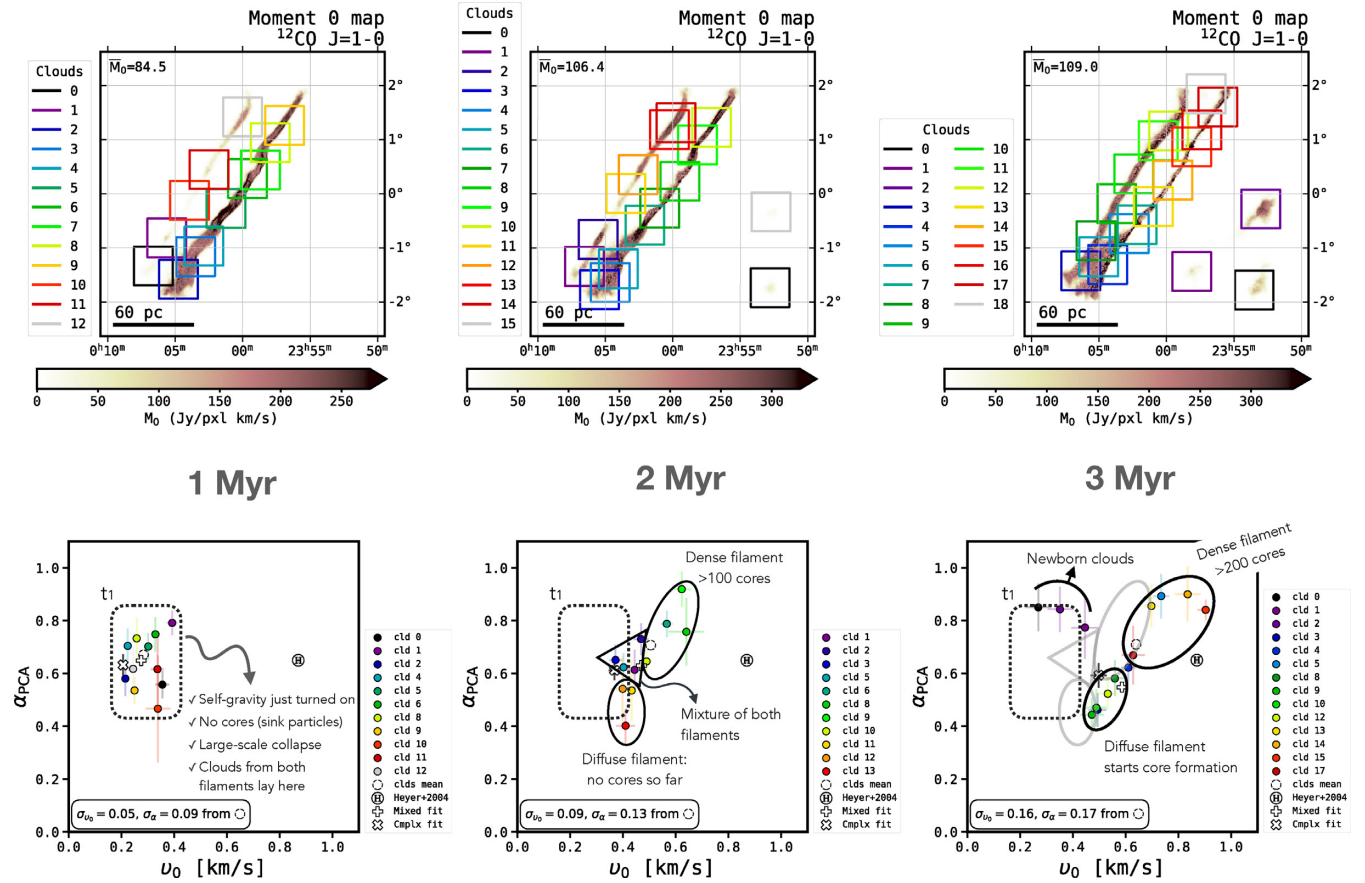
(ii) quiescent self-gravitating molecular clouds have time-dependent trajectories in the  $(v_0, \gamma)$  parameter space.

(iii) there is a common zone in the parameter space where molecular clouds are born.

(iv) the scaling coefficient  $v_0$  increases steadily over time. This is associated with gravitational fragmentation processes (both in pre-core and core formation stages) that increase the magnitude of velocity fluctuations over the cloud lifetime.

(v) the evolution of scaling exponents  $\gamma$  is less predictable. It is associated with variations in the characteristic driving size-scale of gravitational evolutionary stages in cloud portions of a complex. Thereby, large-scale collapse, pre-core, and core formation stages, which may all take place at the same time in a cloud complex, respond differently in the  $(v_0, \gamma)$  space.

(vi) mean surface densities in quiescent self-gravitating cloud complexes also increase steadily over time, suggesting a connection



**Figure 7.** Time evolution of scaling parameters in cloud complex B. The top row shows zeroth moment maps at different time-steps after commencing tracer refinement. The bottom row shows the PCA-derived pseudo-structure function parameters ( $v_0$ ,  $\alpha$ ) from individual clouds in complex B. The standard deviation of cloud parameters is shown in the lower left corner of the panels. The colour code of individual portions in top panels matches the marker colours in bottom panels. Individual clouds migrate over time in the ( $v_0$ ,  $\alpha$ ) space in response to different gravitational processes taking place in the complex.

between the evolutionary stage of local gravitational processes and the dependence of  $v_0$  on surface density reported in previous literature.

(vii) pre-core stages favour small-scale fluctuations ruled by turbulence and local fragmentation, yielding shallower exponents.

(viii) core stages increase scaling exponents as cores act as multiple point-like centres of collapse affecting inter-core gas at medium and large scales predominantly.

(ix) Local gravitational effects do not seem to play the same role in feedback-dominated cases because mass instability is more difficult to achieve and supernovae effects prevail.

(x) In feedback-dominated complexes, driving of large-scale velocity fluctuations by clustered supernovae can sustain the level of turbulence within the clouds.

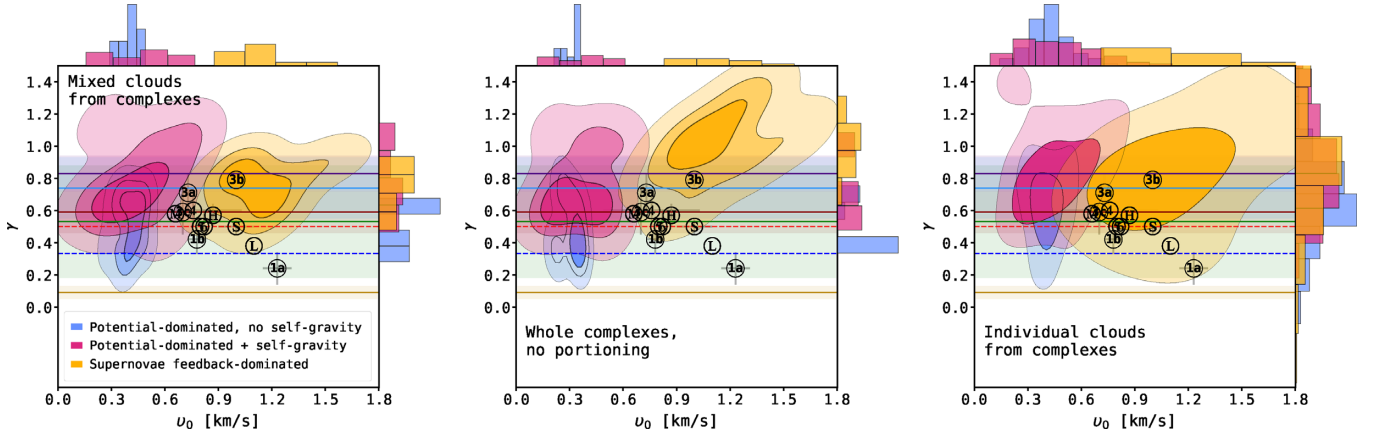
(xi) Unlike external feedback, internal clustered supernovae substantially increase the magnitude of velocity fluctuations  $v_0$  over time, which in turn destroys the cloud faster.

#### 6.4 Structure function: variability within cloud complexes and environmental effects

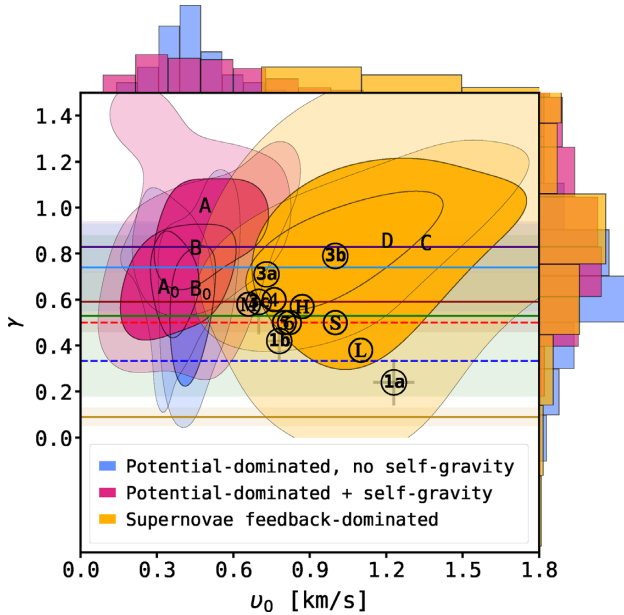
We explore three PCA extraction methods to address any variations in scaling parameters that might arise when different analysis scales are used to retrieve structure functions. As a reminder, the three approaches are the Mixed, the Complex, and the Individual

Cloud methods. The Mixed method consists of combining PCA-derived scales from cloud portions to construct the structure function representative of their hosting complex; the Complex method applies the PCA algorithm on whole complexes, without sub-portioning; and the Individual Cloud method extracts PCA-derived structure functions from cloud portions as if they were individual objects, without mixing them. See further details in Section 5.

Fig. 8 shows the distribution of structure function scaling parameters derived from each of the PCA extraction methods and physical scenarios, and Table 3 summarizes the mean values and standard deviations. The Mixed and the Complex methods yield very similar parameter dispersion for all the physical scenarios. However, studying cloud complexes as a whole with the Complex method shifts the potential-dominated cases (A<sub>0</sub>, B<sub>0</sub>, A, and B) to lower scaling coefficients  $v_0$  ( $\Delta v_0 \sim -0.1$  km s<sup>-1</sup>) and the feedback-dominated cases (C, D) to higher scaling exponents  $\gamma$  ( $\Delta \gamma \sim 0.22$ ) compared to the Mixed method. We attribute these variations to intermittency of density and velocity fields that make the PCA scaling exponents not to follow second-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 describing low-order velocity fluctuations within the Complex (see e.g. Fig. 4). This effect is more prominent in the Complex analysis method, in which PCA is computed on entire cloud complexes and hence intermittent fields are more likely to appear; especially



**Figure 8.** Distribution of structure function parameters for different PCA extraction methods: from left to right, Mixed, Complex, and Individual Clouds (see Section 5). The filled contours correspond to 33, 68, and 95 percent levels. Mean parameters and standard deviations for each method and scenario are summarized in Table 3. The horizontal lines and the circles are the literature parameters listed in Table 2.



**Figure 9.** Same as the third panel of Fig. 8 but discriminating individual clouds according to their hosting cloud complex. The lines and the circles refer to the literature values indicated in Table 2.

in the feedback-dominated cases where extreme fluctuations in the turbulent flow are expected.

In Fig. 8, we also present the scaling parameter distribution for the Individual Cloud method, which computes the structure function of individual molecular clouds and treats 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 supernova feedback plus the local gravitational influence.

Fig. 9 splits individual cloud distributions depending on the origin cloud complex to illustrate the influence of the surrounding environment. Parameter variations due to different density contexts are not significant for potential-dominated complexes though. On the other hand, both cloud complexes in the feedback-dominated

scenario produce similar scaling parameters, however, complex D spans a smaller range of parameters as noticed from its individual cloud values. This is due to the supernova explosions embedded in complex D that destroy cloud structures faster than in the quieter complex C.

As expected, due to the prescribed burst of supernovae in our feedback-dominated scenario, the mean scaling parameters lie close to values resulting from hypersonic-turbulence samples such as the ionized zone of the Rosette cloud reported by Heyer et al. (2006; marker  $\textcircled{3b}$  in Figs 8 and 9) and the Federrath et al. (2010a) simulations using purely compressive forces (the solid blue line  $\text{—}$ ).

In summary,

- (i) using different analysis scales can reveal density and velocity intermittent fields.
- (ii) feedback-dominated cases are more prone to intermittency. This is evidenced as higher dispersion of individual cloud scaling parameters than potential-dominated complexes.
- (iii) environmental conditions of clouds may also split scaling parameters ( $\nu_0, \gamma$ ) into separate clusters of points for each physical scenario.
- (iv) the feedback-dominated scenario produces the largest range of scaling parameters.
- (v) the parameter distribution from clouds with embedded supernovae is more confined in  $\gamma$  than those with external feedback only. We attribute this to the faster disruption of coherent structures in the former case.

## 7 DISCUSSION

### 7.1 Comparison to observational and synthetic structure functions

From Fig. 8, the potential-dominated complexes without self-gravity are mostly clustered between (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 large-scale gravitational forces along with random (isolated) supernova feedback are able to reproduce classical scaling exponents ( $\gamma$ ) 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. (2010a). The solenoidal turbulence is associated with regions of low star formation activity such as 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 the observations.

Turning on self-gravity generally produces steeper structure function scaling exponents ( $\gamma$ ). Depending on the evolutionary stage of cloud complexes, gravity-driven (non-thermal) motions such as collapse and/or accretion flows can increase the velocity dispersion. This results in a sustained increase 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 characterized 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 local gravitational effects as it is composed of two quiescent filamentary structures that evolve differently. At early stages, one filament is dense and prominent and the other is diffuse and discontinuous. Fig. 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  $\alpha$  are in general higher for the longer, denser 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 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 observational relations of the form  $\delta v = v_0 l^\gamma$  that depend on the scale size range of the cloud: (i) a transonic  $\delta v = 0.55 l^{0.25}$  for scales  $< 1.0$  pc and (ii) a supersonic  $\delta v = 0.66 l^{0.58}$  on scales between 1.0 and 3.0 pc. We multiplied their reported  $\delta v$  by a factor of  $\sqrt{3}$  (assuming isotropic 3D velocity fluctuations) to facilitate comparison with our 3D parameters. Our self-gravitating cloud complexes without clustered feedback are consistent with the scaling parameters derived from this quiescent region in the supersonic regime, especially the most evolved complexes (see middle panel of Fig. 6). Additionally, some cloud portions in complex B also manage to reproduce Musca's parameters in the transonic regime. This is the case for clouds in the diffuse filament of complex B where core formation has just commenced at  $t = 3$  Myr (see Fig. 7).

Furthermore, Hacar et al. (2016) also found that an independent 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 scatter 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 (see Fig. 7), which at the same time, is closely related to the density distribution of the region. Thus, we attribute the broken power law reported for the Musca cloud to asynchronous evolution of parameters driven by local gravitational effects.

Scaling exponents ( $\gamma$ ) derived from our feedback-dominated cloud complexes are consistent with Federrath et al. (2010a) simulations with purely compressive forces or with the value reported by Bertram et al. (2014) for  $^{12}\text{CO}$  emission of molecular clouds with artificial

turbulent fields. When considering both scaling parameters ( $\gamma$ ,  $v_0$ ), our cloud complexes and individual clouds can reproduce a range of observations including those of Larson (1981), Solomon et al. (1987), and Heyer & Brunt (2004; see Figs 8 and 9). The distribution of parameters is centred around the point ③ corresponding to 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 inject strong energy feedback in our simulations.

## 7.2 Are supernovae important for driving turbulence in molecular clouds?

It is useful to briefly discuss the role that supernovae play in driving cloud-scale turbulence in these simulations, given their overall importance for the energy balance of the ISM (see e.g. Mac Low & Klessen 2004). We find, in common with several previous studies (Ibáñez-Mejía et al. 2017; Seifried et al. 2018), that randomly distributed supernovae that explode with a rate similar to the supernova rate in the Milky Way are unable to drive turbulence in molecular clouds at a level consistent with observations of Galactic GMCs. As Seifried et al. (2018) explore in some detail, the reason for this is that in the random supernova scenario, the chances of supernovae exploding close to the clouds on a regular basis are small and so any turbulence injected into the clouds by a nearby supernova tends to decay away long before the next nearby supernova occurs. We therefore conclude that supernovae are not important for driving the turbulence observed in quiescent (i.e. non-star forming) molecular clouds and rather gravity plays a more dominant role (see e.g. Klessen & Hennebelle 2010). On the other hand, in cloud complexes actively forming massive stars, supernova explosions do not occur with a spatially random distribution, but instead are highly correlated with the gas distribution. In this case, the rate of nearby supernova explosions is much higher than with the random distribution, allowing the supernovae to play a dominant role in driving the turbulence in these clouds. Future work will include the analysis of other important driving mechanisms such as stellar winds, jets, and photoionizing radiation, as well as the influence of magnetic fields.

## 8 CONCLUSIONS

We have performed PCA on full non-LTE radiative transfer simulations of molecular cloud complexes, self-consistently generated using our Cloud Factory Galactic-scale ISM simulation suite. We explore PCA-derived velocity structure functions from three 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 feedback. Large-scale potential and local gravitational effects are both active in this case.

Regardless of the physical scenario, we find that all the cloud complexes analysed from our Cloud Factory zooms agree with distinct types of turbulence reported in the literature. Clearly, large-scale gravitational forces alone when combined with turbulent decay are enough to reproduce Kolmogorov's and Burgers-like turbulence scaling exponents, but scaling coefficients remain too low compared with observations. None the less, under weak influence of isolated

supernova explosions, local gravitational forces can make structure functions evolve over time and reproduce observations of quiescent molecular clouds.

We report time-dependent 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 display low magnitudes and then migrate through the  $(v_0, \gamma)$  parameter space as star-forming cores emerge within. The scaling exponents ( $\gamma$ ) are generally less predictable because they depend upon the stage of gravitational collapse, which varies locally as a function of the boundness conditions of sub-structures in the cloud complex. This could explain power-law breaks and variations in structure function parameters observed for different size-scales in quiescent molecular clouds.

However, gravitational forces alone (when combined with random supernovae feedback) are not enough to reproduce both the scaling coefficient and exponent of molecular clouds with active star formation. We find that clustered feedback from supernovae tied to sites of star formation is key to self-consistently generate clouds that reproduce the scaling parameters reported by observations with similar size-scales and resolutions to those used in our simulations.

Our results suggest that a PCA-based statistical study is a robust method to diagnose the physical mechanisms driving gravoturbulent fluctuations in molecular clouds by providing a quantitative description of the velocity field. The analysis tools developed in this work are all condensed in our new open source PCAFACTORY package.

## ACKNOWLEDGEMENTS

The authors would like to thank C. Brunt, C. Dickinson, and the anonymous referee for their inspiring discussions and comments that improved the results, conclusions, and key theoretical aspects of this work. 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, as well as support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Ref no. FOR 2634/1 TE 1024/1-1. AFI also acknowledges support by the Agencia de Educación Superior de Medellín (Sapiencia), attached to the Alcaldía de Medellín, and the Centro de Ciencia y Tecnología de Antioquia (CTA), through the grant Premio Medellín Investiga 2019, which encourages high-impact research and innovation in the city of Medellín, Colombia. RJS gratefully acknowledges support from an STFC Ernest Rutherford Fellowship (grant ST/N00485X/1), without which this work would not have been possible. SCOG and RSK acknowledge support from the DFG via the Collaborative Research Center (SFB 881, Project-ID 138713538) ‘The Milky Way System’ (sub-projects A1, B1, B2 and B8) and from the Heidelberg cluster of excellence (EXC 2181 – 390900948) ‘STRUCTURES: A unifying approach to emergent phenomena in the physical world, mathematics, and complex data’, funded by the German Excellence Strategy. The authors gratefully acknowledge support by the state of Baden-Württemberg through bwHPC and the German Research Foundation (DFG) through grant INST 35/1134-1 FUGG. They furthermore acknowledge the data storage service SDS@hd supported by the Ministry of Science, Research and the Arts Baden-Württemberg (MWK) and the German Research Foundation (DFG) through grant INST 35/1314-1 FUGG. This work used the COSMA Data Centric system at Durham University, operated by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility ([www.dirac.ac.uk](http://www.dirac.ac.uk)). This equipment was funded by a

BIS National E-infrastructure capital grant ST/K00042X/1, DiRAC Operations grant ST/K003267/1 and Durham University.

## DATA AVAILABILITY

The data underlying this article are publicly available on [yt.Hub](https://yt.hub), which is a data management platform ‘supported in part by the Gordon and Betty Moore Foundation’s Data-Driven Discovery Initiative through Grant GBMF4561 to Matthew Turk and the National Science Foundation under Grant No. ACI-1535651’. The data can be accessed by following instructions at <https://github.com/andizq/andizq.github.io/tree/master/pcafactory-data>.

## REFERENCES

- André P. et al., 2010, *A&A*, 518, L102  
 Anselmet F., Gagne Y., Hopfinger E. J., Antonia R. A., 1984, *J. Fluid Mech.*, 140, 63  
 Arzoumanian D. et al., 2019, *A&A*, 621, A42  
 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  
 Blitz L., 1993, in Levy E. H., Lunine J. I., eds, *Protostars and Planets III*. p. 125, Univ. Arizona Press, Tucson  
 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., 2003, *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  
 Burgers J. M., 1939, *Mathematical Examples Illustrating Relations Occurring in the Theory of Turbulent Fluid Motion*. Vol. 17. Kon. Ned. Akad. Wet. Verh., Amsterdam  
 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  
 de Avillez M. A., Breitschwerdt D., 2005, *A&A*, 436, 585  
 de Jong T., 1977, *A&A*, 55, 137  
 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  
 Falgarone E., Pety J., Hily-Blant P., 2009, *A&A*, 507, 355  
 Federrath C., 2018, *Phys. Today*, 71, 38  
 Federrath C., Roman-Duval J., Klessen R. S., Schmidt W., Mac Low M. M., 2010a, *A&A*, 512, A81  
 Federrath C., Banerjee R., Clark P. C., Klessen R. S., 2010b, *ApJ*, 713, 269  
 Federrath C., Sur S., Schleicher D. R. G., Banerjee R., Klessen R. S., 2011, *ApJ*, 731, 62  
 Federrath C., Schrön M., Banerjee R., Klessen R. S., 2014, *ApJ*, 790, 128  
 Flower D. R., Pineau Des Forêts G., Walmsley C. M., 2006, *A&A*, 449, 621

- Fournier J. D., Frisch U., 1983, *J. Mec. Theor. Appl.*, 2, 699
- Frisch U., 1995, *Turbulence: The Legacy of A. N. Kolmogorov*. Cambridge Univ. Press, Cambridge
- Frisch U., Bec J., Villone B., 2001, *Phys. D Nonlinear Phenom.*, 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
- 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
- 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
- Hennebelle P., Chabrier G., 2009, *ApJ*, 702, 1428
- Heyer M. H., Brunt C. M., 2004, *ApJ*, 615, L45
- Heyer M. H., Schloerb F. P., 1997, *ApJ*, 475, 173
- Heyer M. H., Williams J. P., Brunt C. M., 2006, *ApJ*, 643, 956
- Heyer M., Krawczyk C., Duval J., Jackson J. M., 2009, *ApJ*, 699, 1092
- 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
- Ibáñez-Mejía J. C., Mac Low M.-M., Klessen R. S., Baczynski C., 2017, *ApJ*, 850, 62
- Immer K., Reid M. J., Menten K. M., Brunthaler A., Dame T. M., 2013, *A&A*, 553, A117
- 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
- Joshi P. R., Walch S., Seifried D., Glover S. C. O., Clarke S. D., Weis M., 2019, *MNRAS*, 484, 1735
- Joung M. R., Mac Low M.-M., Bryan G. L., 2009, *ApJ*, 704, 137
- Kainulainen J., Beuther H., Henning T., Plume R., 2009, *A&A*, 508, L35
- Keto E., Zhang Q., 2010, *MNRAS*, 406, 102
- Klessen R. S., Glover S. C. O., 2016, *Saas-Fee Advanced Course*, 43, 85
- 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, *ASP Conf. Ser. Vol. 322, The Formation and Evolution of Massive Young Star Clusters*. Astron. Soc. Pac., San Francisco, p. 299
- Koch E. W., Rosolowsky E. W., Boyden R. D., Burkhart B., Ginsburg A., Loeppky J. L., Offner S. S. R., 2019, *AJ*, 158, 1
- Kolmogorov A., 1941, *Akademiia Nauk SSSR Dokl.*, 30, 301
- Kraichnan R. H., 1974, *J. Fluid Mech.*, 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
- Mac Low M.-M., Klessen R. S., 2004, *Rev. Mod. Phys.*, 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
- Miville-Deschênes M.-A., Murray N., Lee E. J., 2017, *ApJ*, 834, 57
- 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., Schaaf K., Zhu C., 2016, *MNRAS*, 455, 1134
- Pan L., Padoan P., Haugbølle T., Nordlund Å., 2016, *ApJ*, 825, 30
- Passot T., Pouquet A., Woodward P., 1988, *A&A*, 197, 228
- Pearson K., 1901, *London Edinburgh Dublin Phil. Mag. J. Sci.*, 2, 559
- Peters T. et al., 2017, *MNRAS*, 466, 3293
- Ragan S. E., Henning T., Tackenberg J., Beuther H., Johnston K. G., Kainulainen J., Linz H., 2014, *A&A*, 568, A73
- Reissl S., Wolf S., Brauer R., 2016, *A&A*, 593, A87
- Rice T. S., Goodman A. A., Bergin E. A., Beaumont C., Dame T. M., 2016, *ApJ*, 822, 52
- 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*. Wiley-VCH, Germany
- Schneider N., Simon R., Kramer C., Stutzki J., Bontemps S., 2002, *A&A*, 384, 225
- Schneider N. et al., 2015, *A&A*, 578, A29
- Seifried D., Walch S., Haid S., Girichidis P., Naab T., 2018, *ApJ*, 855, 81
- 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
- Smith R. J. et al., 2020, *MNRAS*, 492, 1594
- Sobolev V. V., 1960, *Moving Envelopes of Stars*. Harvard Univ. Press, Cambridge
- 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
- Tress R. G., Smith R. J., Sormani M. C., Glover S. C. O., Klessen R. S., Mac Low M.-M., Clark P. C., 2020, *MNRAS*, 492, 2973
- Truelove J. K., Klein R. I., McKee C. F., Holliman J. H. II, Howell L. H., Greenough J. A., 1997, *ApJ*, 489, L179
- van Dishoeck E. F., 2004, *ARA&A*, 42, 119
- van Dishoeck E. F., Blake G. A., 1998, *ARA&A*, 36, 317
- Vazquez-Semadeni E., 1994, *ApJ*, 423, 681
- Visser R., van Dishoeck E. F., Black J. H., 2009, *A&A*, 503, 323
- Walch S. et al., 2015, *MNRAS*, 454, 238
- Williams J. P., Blitz L., McKee C. F., 2000, in Mannings V., Boss A. P., Russell S. S., eds, *Protostars and Planets IV*. Univ. Arizona Press, Tucson, p. 97
- Wilson T. L., Rohlfs K., Hüttemeister S., 2013, *Tools of Radio Astronomy*. Springer Nature, Switzerland AG
- Yang B., Stancil P. C., Balakrishnan N., Forrey R. C., 2010, *ApJ*, 718, 1062

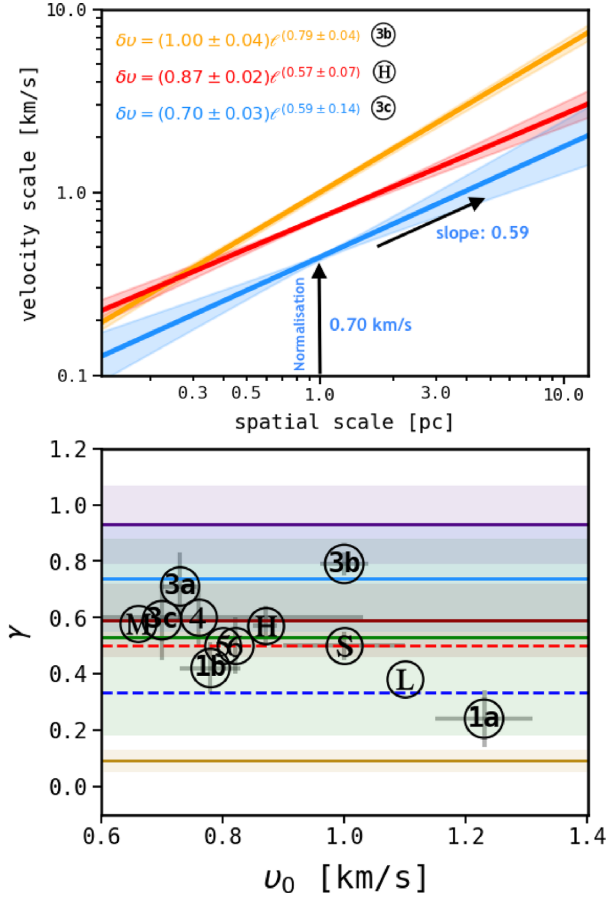
## SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

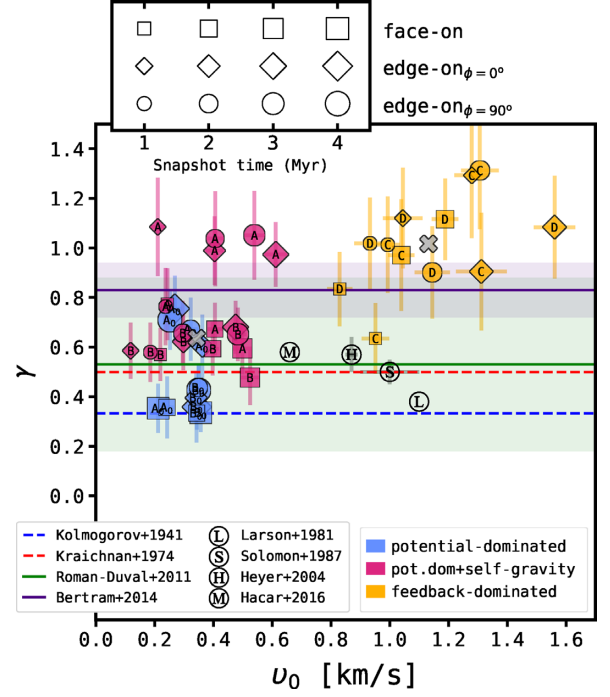
**Izquierdo\_Gravoturbulent\_statistics\_supfile\_v3.zip**

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

## APPENDIX A: SUPPORTING FIGURES

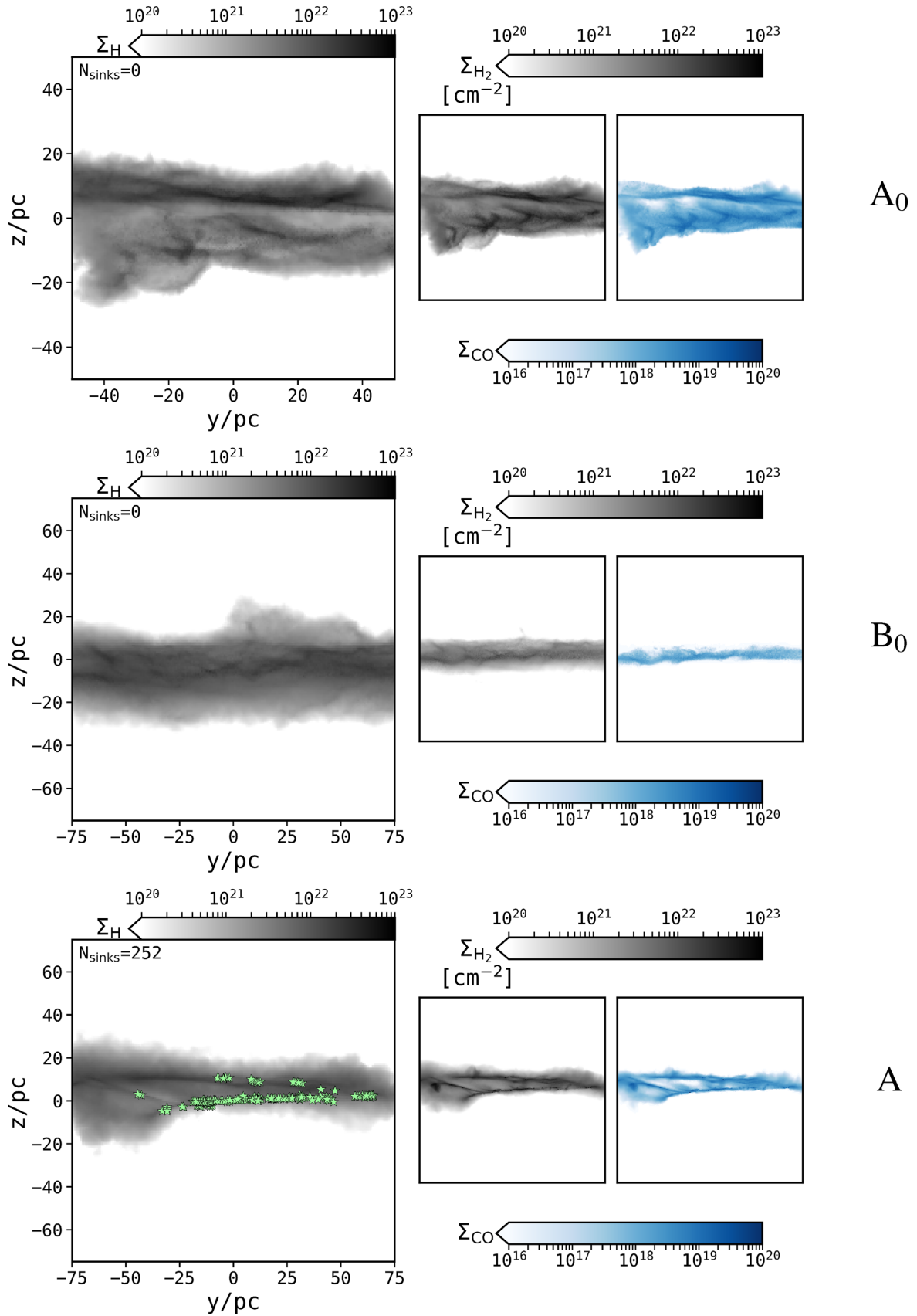


**Figure A1.** Comparison of selected structure function scaling parameters reported in previous literature. Line and marker codes are listed in Table 2.



**Figure A2.** Same as Fig. 3 but using the complex method (see Section 5).





**Figure A3.** Edge-on  $\phi=90^\circ$  projections of  $\text{H}$ ,  $\text{H}_2$ , and  $^{12}\text{CO}$  column densities ( $\Sigma$ ) from cloud complexes (labelled on the right) extracted 2 Myr after injecting tracer particles in the simulations. If any, sink particles are overlaid on  $\text{H}$  maps as star markers.

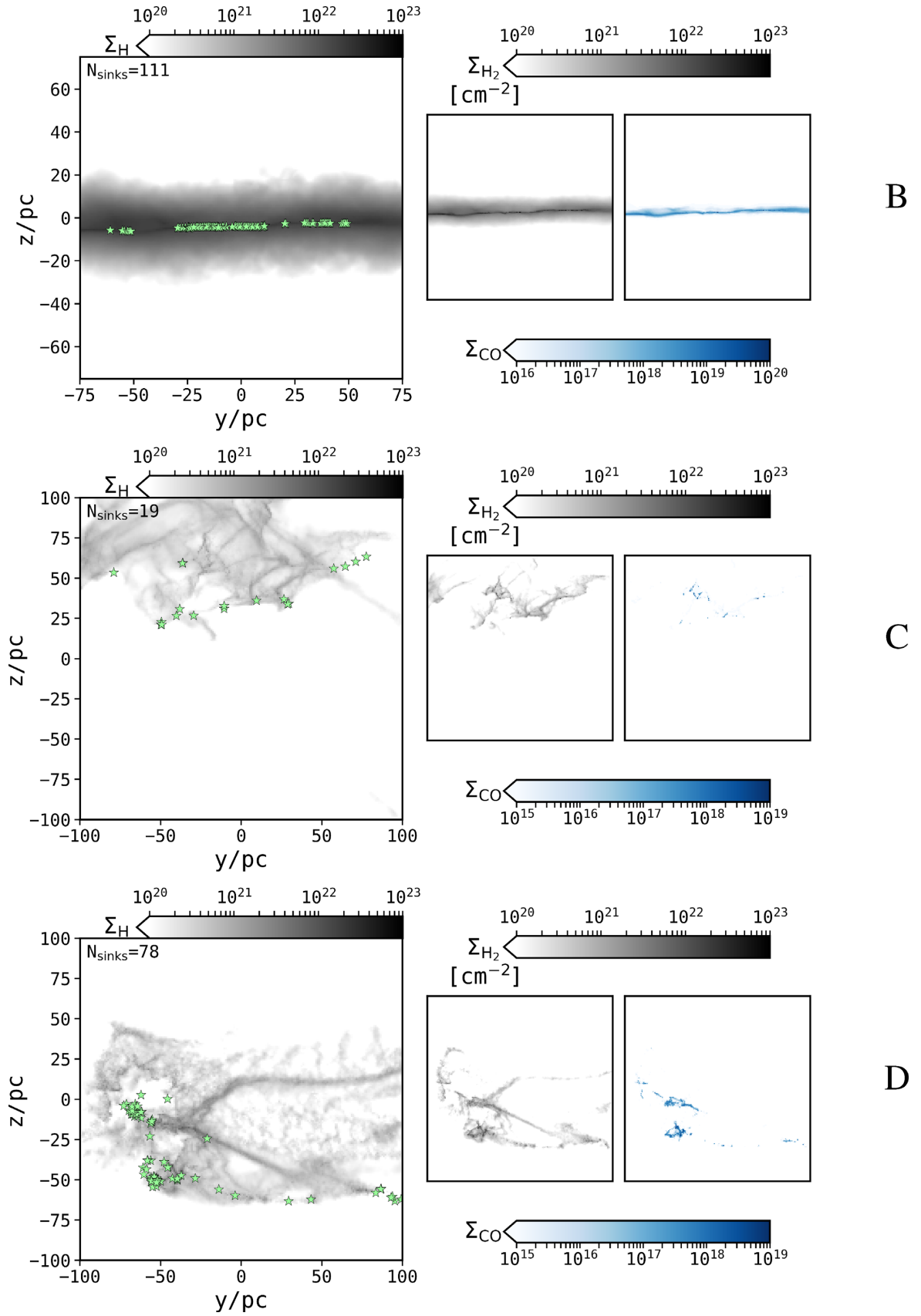
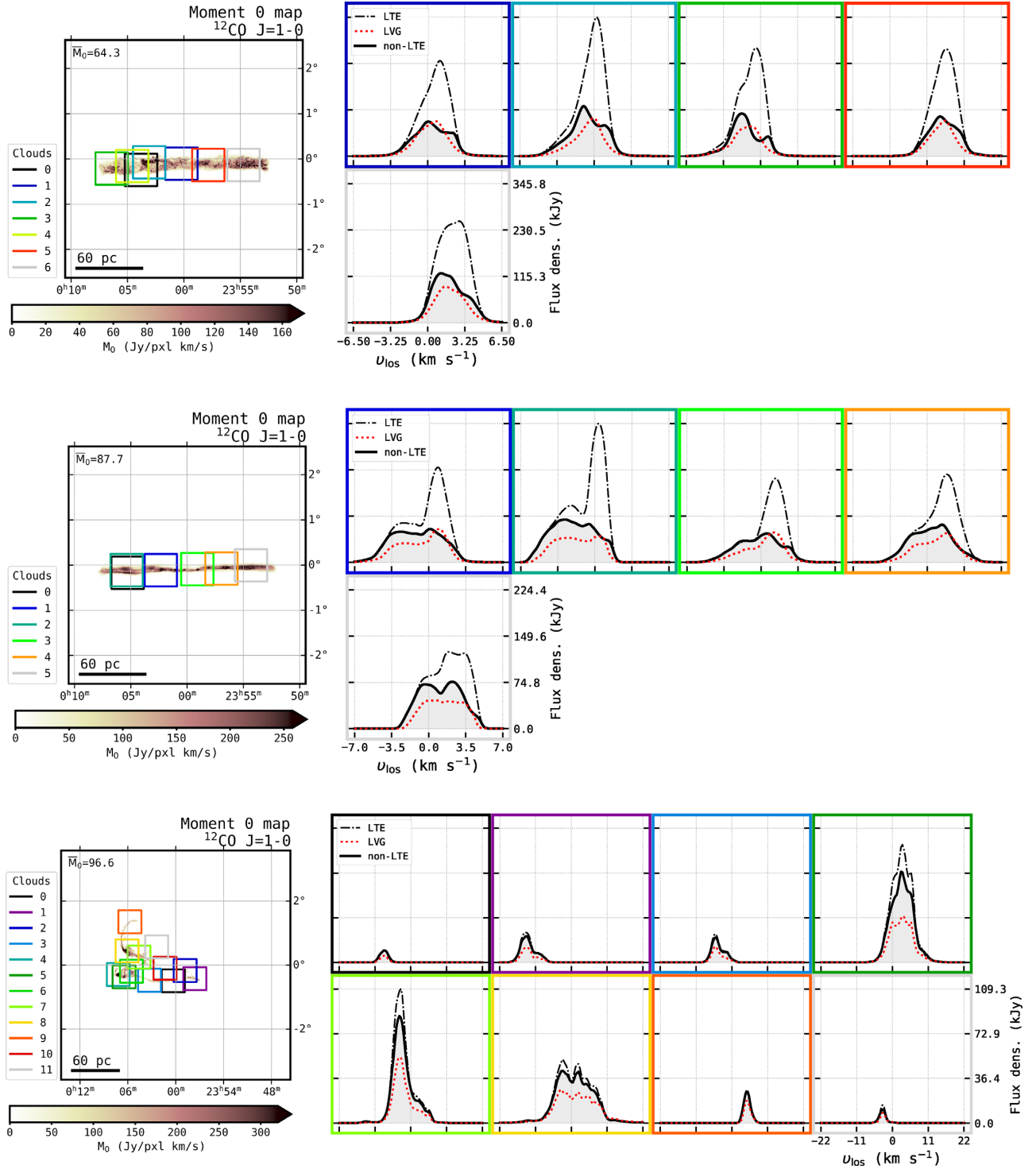


Figure A3. – continued.



**Figure A4.** LTE (the dashed lines) LVG (the red-dotted lines) and non-LTE (the solid lines and the shades)  $^{12}\text{CO}$   $J=1-0$  emission profiles of individual clouds in cloud complexes B<sub>0</sub> (top), B (middle), and D (bottom). These are edge-on  $\phi=90^\circ$  views of the complexes analysed in Fig. 4. The colours of the axes spines in the right-hand panels correspond to the colour code in the left-hand panels and indicate from which cloud the line profile is extracted.

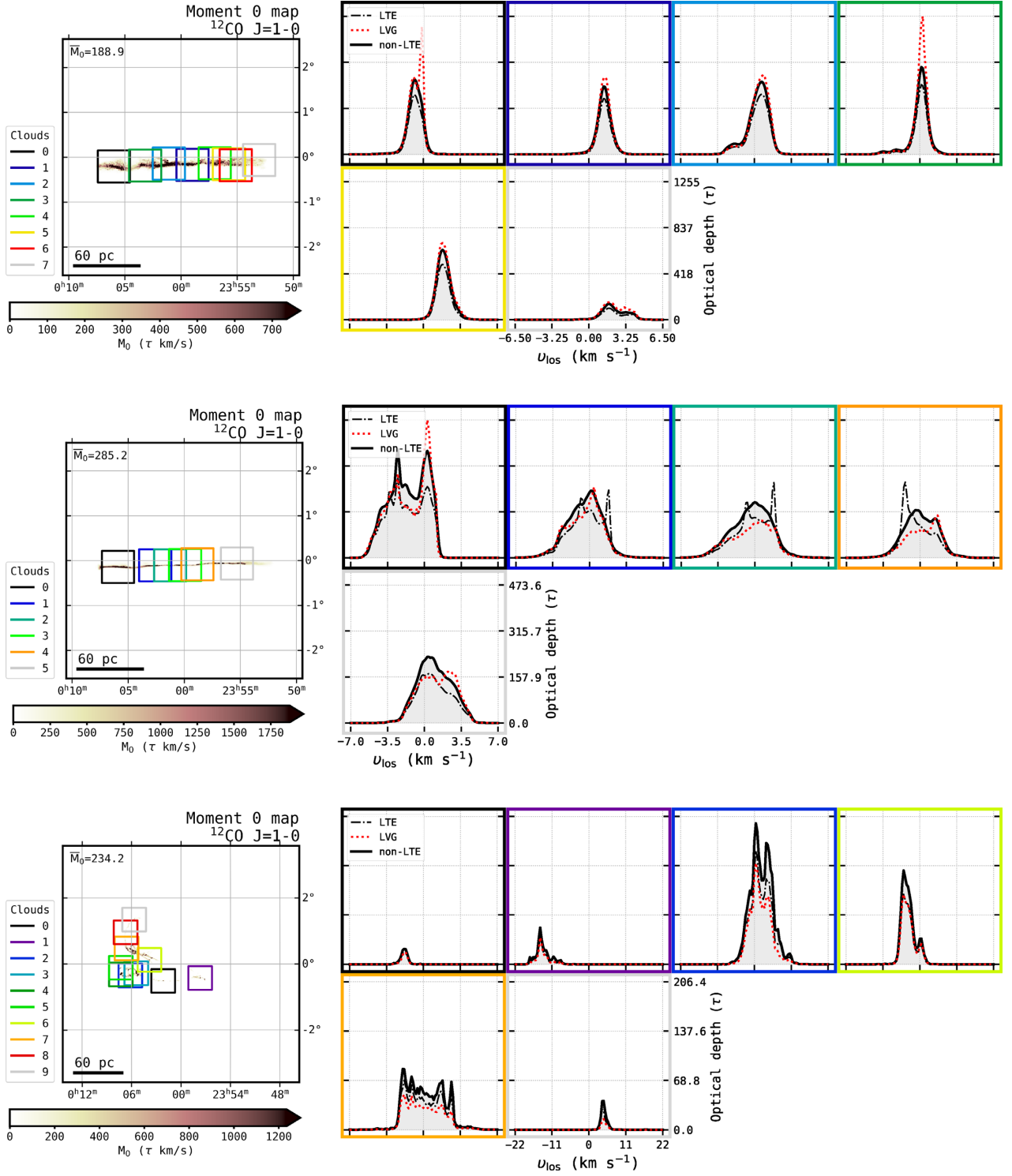
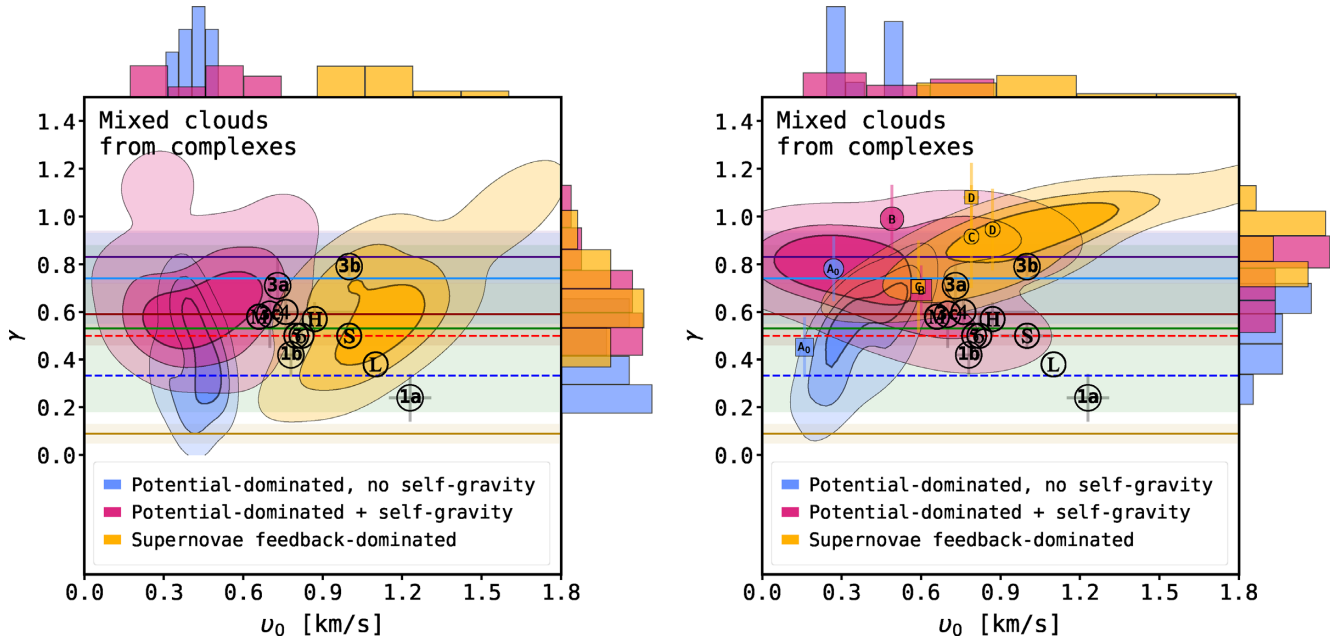


Figure A5. Same as Fig. A4 but using optical depth maps.





**Figure A6.** Same as left-hand panel of Fig. 8 but using LTE (left) and LVG (right) level populations for the radiative transfer of  $^{12}\text{CO}$  from our cloud complexes. The coloured markers in the right-hand panel are the  $^{13}\text{CO}$  parameters presented in Table 4.

This paper has been typeset from a  $\text{\LaTeX}$  file prepared by the author.