structures, but also many odd features corresponding to dimples on the surfaces of larger ones.

So, in a simple picture of star formation, where one "clump" lasts a long time, and forms one star, it makes sense to think about breaking up the cloud into non-overlapping volumes and creating a "clump mass spectrum" from the result. But, in a turbulent picture, where hierarchical structure pervades the cloud and structures are transient, it seems foolheardy to break up the full volume of the cloud into non-overlapping clumps. The analysis is further complicated by the aforementioned 2 spatial+1 velocity dimension nature of the observational data. Segmentation based analysis algorithms typically ignore this subtlety and operate in three dimensions despite the fact that much of the clumpy structure identified is thought to result from chance superpositions along the line of sight (Ostriker et al) . We note that the segmentation of sparse two-dimensional maps via the CLUMPFIND is not as fraught since the standard application of the algorithm in this case is to relatively unblended data for source identification (similar to the standard Source EXtractor algorithm Bertin, E.; Arnouts, S, A&AS 1996).

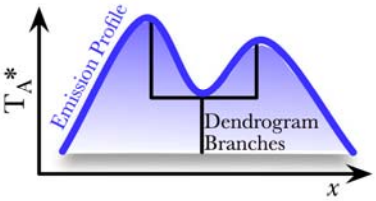
In this letter, we borrow heavily on techniques used in other fields to show that a novel application of commonly used structure trees (e.g. refs. from NSF proposal) to molecular line data provides a method to characterize the hierarchical sturcture use the results to identify physically relevant features in the data. While well-developed in other fields such as computer science and computational biology, the application of the tree methodology in astrophysics has been relatively lacking. In cosmological simulations, the merger history of galaxies is frequently parameterized as a function of redshift with a structure tree (e.g. Kauffmann & White, MNRAS 1993). In the field of star formation, Houlahan & Scalo (1990) proposed applying structure trees to extinction maps to characterize their hierarchical structure. Using this as an inspiration, we developed an algorithm that "abstracts" the hierarchical structure of a p-p-v data cube into an easily visualized respresentation; and then we use this abstraction to identify the structures relevant to star formation in the data.

DENDROGRAMS

A schematic illustration of the dendrogram process appears, showing the construction of a dendrogram from a hypothetical 1D emission profile (blue). The "dendrogram" (in black) graphically represents the peaks of the emission structure as separate *branches*. At the highest T_{A*} value that still spans both peaks, the two branches are connected into a single branch. By repeating this process for every peak of the emission, dendrograms abstract a complete topological description of the emission in a data cube[1]. We construct analogous dendrogram diagrams for 3D data cubes; however, to plot them most clearly in 2D, we flatten them, eliminating any meaningful spatial information on the *x*-axis.

Our new contribution to the dendrogram technique is the calculation *physical* properties for every object corresponding to a branch in the dendrogram [8]. To determine what features might be important, we calculate a virial parameter as the ratio of kinetic energy to gravitational binding energy (without external pressures or a magnetic field) for every branch in the dendrogram. We highlight in red every branch that corresponds, in our simplified model, to a self-gravitating object. The standard feature identification algorithms may find the objects at the top branches of the dendrogram trees, since these correspond to the peaks of "clumps," but would be unable to identify objects at the base of the dendrogram "tree."

In Figure 1 we show the dendrogram of the L1448 region in the star-forming molecular cloud Perseus. The original data (Figure 2) are ^{13}CO(1-0) emission taken from the COMPLETE survey of Star Forming Regions survey of the Perseus region (Ridge et al.). The main complex in L1448, represented by the dominant branch of the dendrogram, has gravitational and kinetic energies that are comparable and encompasses a large fraction of the region. The dendrogram indicates the importance of gravity over a large range of scales in the molecular cloud from the individual small clumps indicated with the "leaves" of the dendrogram down into the base of the structure tree. In contrast, the dendrogram indicates that the feature found at large velocities is dynamically as well as kinematically distinct from the majority of the region. In addition, the dendrogram identifies several "leaves" which have distinct regions where gravity dominates on the smallest scales. Such features are interesting since they represent where the molecular gas is closest to forming stars.



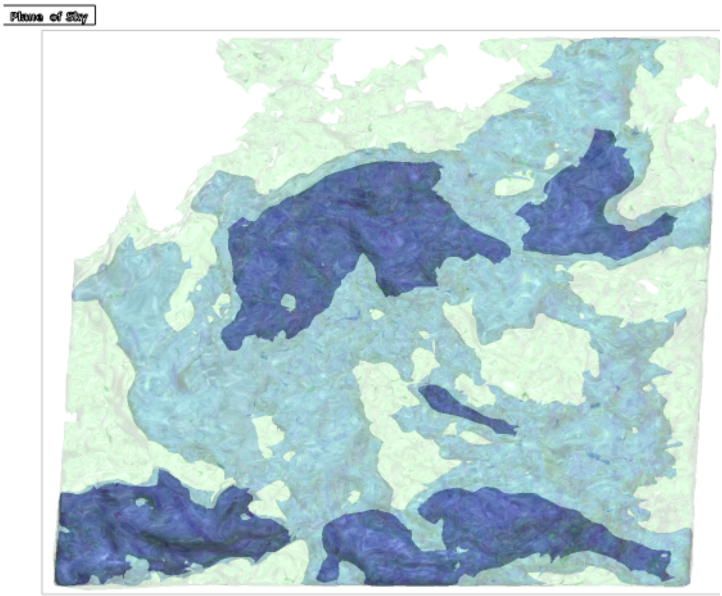


Figure 2 | Isointensity surface models of L1448 in ^{13}CO. The models are in *p-p-v* space, where the front of the cube is the plane of the sky. The intensity thresholds shown are chosen using the "dendrogram" procedure described.

---

[1] Dendrogram construction is completely determined by the data without relying on choices algorithm parameters, such as the step size in the CLUMPFIND algorithm or the relative weights of the c2 components in Gaussclumps.