Modeling Star Formation as a Markov Process in a Supersonic Gravoturbulent Medium

Evan Scannapieco © and Mohammadtaher Safarzadeh ©
Arizona State University School of Earth and Space Exploration, P.O. Box 871404, Tempe, AZ 85287, USA
Received 2018 July 15; revised 2018 September 3; accepted 2018 September 17; published 2018 September 26

Abstract

Molecular clouds exhibit log-normal probability density functions (PDF) of mass densities, which are thought to arise as a consequence of isothermal, supersonic turbulence. Star formation is then widely assumed to occur in perturbations in which gravitational collapse is faster than the rate of change due to turbulent motions. Here we use direct numerical simulations to measure this rate as a function of density for a range of turbulent Mach numbers, and show that it is faster at high densities than at low densities. Furthermore, we show that both the density PDF and rate of change arise naturally in a simple model of turbulence as a continuous Markov process. The one-dimensional Langevin equation that describes this evolution depends on only two parameters, yet it captures the full evolution seen in direct three-dimensional simulations. If it is modified to include gravity, the Langevin equation also reproduces the rate of material collapsing to high densities seen in turbulent simulations including self-gravity. When generalized to include both temperature and density, similar analyses are likely applicable throughout astrophysics.

Key words: methods: statistical – stars: formation – turbulence

1. Introduction

Many astrophysical systems exist in pseudo steady-states, in which the global properties are roughly constant over many dynamical times. The interstellar medium, for example, can be thought of in this manner, with material moving between different phases whose mass fractions and properties remain roughly constant (Cox 2005). Similarly, molecular clouds are well modeled as supersonic, isothermal systems that are in near-virial equilibrium and driven for sufficiently long times to exhibit a well-developed turbulent cascade (Federrath et al. 2010).

The densities and temperatures of any such system will be distributed according to a probability distribution function (PDF) that varies slowly with time, even though individual parcels of material are constantly changing. In the case of isothermal turbulence, the mass-weighted PDF is given as

$$P_M(s) \approx \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(s - s_0)^2}{2\sigma^2}\right],$$

where $s \equiv \ln(\rho/\rho_0)$, $\rho_0$ is the mean density, and the variance $\sigma^2 = 2s_0$ as required by mass conservation (Vazquez-Semadeni 1994; Padoan et al. 1997; Federrath et al. 2010). Direct numerical simulations find that $\sigma^2 = \ln(1 + b^2M^2)$, where $M = \sigma_z/c_s$ is the ratio of the mean velocity dispersion $\sigma_z$ to the sound speed $c_s$, and $b \approx 1$ and $1/3$ for compressive and solenoidal forcing, respectively (Padoan et al. 1997; Ostriker et al. 2001; Price et al. 2011).

While the structure and lifetime of dense regions have been studied (Falceta-Gonçalves & Lazarian 2011; Robertson & Goldreich 2018), the overall manner in which material moves within the turbulent PDF has never been directly measured. Nevertheless, it is this motion that controls which parcels of gas will form stars, as it determines which will collapse due to gravity before they are reshuffled to lower densities (Krumholz & McKee 2005; Hennebelle & Chabrier 2011; Padoan & Nordlund 2011; Federrath & Klessen 2012; Hopkins 2013; see, however, Elmegreen 2000; Murray 2011).

Here we measure this overall evolution for the first time, show that it can be described by two simple functions, and show how these results can be extended to reproduce the results of gravoturbulent simulations of star formation. The structure of this Letter is as follows. In Section 2 we describe our turbulent simulations and measurements of the evolution of the PDF. In Section 3 we apply these results to develop a Markov model of the evolution of the medium. In Section 4 we extend this model to study star formation in a steady-state gravoturbulent medium, and in Section 5 we show how similar models can be constructed to describe other scenarios of star formation as well as other astrophysical systems.

2. Simulations

To study the evolution of material in isothermal turbulence, we carried out direct numerical simulations over a range of Mach numbers using FLASH (Fryxell et al. 2000) version 4.2.1. We solved the hydrodynamics equations using an unsplit solver with third-order reconstruction (Lee 2013), and employed a hybrid Riemann solver, which uses both an extremely accurate but somewhat fragile Harten–Lax–van Leer–Contact (HLLC) solver (Toro et al. 1994) and a more robust, but more diffusive Harten–Lax and van Leer (HLL) solver (Einfeldt et al. 1991).

Each simulation was carried out in a $512^3$ periodic box of size $L_{box}$, within which turbulence was continuously driven (Eswaran & Pope 1988), by solenoidal modes (i.e., $\nabla \cdot F = 0$) in the range of wavenumbers $1 \leq L_{box}|k|/2\pi \leq 3$, such that the average forcing wavenumber was $k^{-1} \approx L_{box}/2/2\pi$. This driving choice was made because, at small scales, most of the turbulent kinetic energy is found in the solenoidal modes (Pan & Scannapieco 2010).

To ensure a nearly constant Mach number for each simulation, we made use of an adaptive scheme to update the driving conditions (Gray & Scannapieco 2016). At each time step, the rms velocity was measured and the strength of the kicks was enhanced by $\propto (\sigma_\perp/\sigma_z)^{-5}$, where $\sigma_\perp$ is the current global velocity dispersion, and $\sigma_z$ is the target value. In this
way, we carried out simulations with average Mach numbers of 2.03, 3.17, and 6.25 as shown in Figure 1.

Once the turbulence was fully developed, we assigned each cell a scalar that recorded the value of $s$ at a given time, $t_1$. These values were then advected in a manner similar to a mass fraction

$$\frac{\partial \rho s}{\partial t} + \nabla \cdot \rho s v = 0,$$

such that the current value of $s$ in a given cell is the mass-weighted averaged value of $s$ at the time when the passive scalar was painted on, with the weighting being computed over all of the Lagrangian fluid elements that are currently in the cell. This allows us to determine the difference between the value of a cell, $s_2$, at time $t_2$, and the value of the material from which it was comprised, $s_1$, at a previous time $t_1$. Four such scalars were advected, which were repainted onto the simulations at various times to study the change in $s$ over a fixed number of timesteps (250, 500, 1000, and 2000). Defining the eddy turnover time as $\tau_{\text{eddy}} = L_{\text{box}}/2\sigma_v$, these intervals corresponded to average time differences of $\Delta t/\tau_{\text{eddy}}$ of 0.097, 0.191, 0.383, and 0.767 in the Mach 2.03 simulation, 0.105, 0.210, 0.420, and 0.839 in the Mach 3.17 simulation, and 0.098, 0.197, 0.393, and 0.785 in the Mach 6.25 simulation. As a convergence test, the Mach 3 run was repeated at a resolution of 256$^3$ cells. For this run, the change in $s$ and the PDF matched our 512$^3$ results within 10% over the full range of $s$ values discussed below.

3. Markov Model

We then developed an analytic model of the evolution of $s$, approximating turbulence as a temporally homogeneous Markov process, a stochastic process for which the future state of any quantity of interest only depends on its current value.

While this is not strictly the case for a turbulent medium, in which velocity correlations are non-Gaussian and correlated in time (Hennebelle & Falgarone 2012; He et al. 2017), we found that this approximation nevertheless allowed us to accurately capture the evolution of $s$ in a simple way that provides insight into the timescales operating at low and high densities.

In this case, we can define the propagator density function $\Pi(ds|dt, s)$, which gives the change in $s$ over a time interval $dt$, given its initial value $s$. Note that $dt$ is an infinitesimal increment, but $ds$ need not be. However, we can greatly simplify $\Pi(ds|dt, s)$ if we further assume that the process is continuous, meaning that $\Pi(ds|dt, s)$ varies smoothly with $s$, and is zero outside an infinitesimally small neighborhood around $ds = 0$. In this case, the propagator is a Gaussian that is completely determined by only two characterizing functions (Gillespie 1996): the drift or advection function, $A(s)$, and the diffusion function, $D(s)$, as

$$\Pi(ds|dt, s) = \frac{1}{\sqrt{2\pi D(s)dt}} \exp\left\{-\frac{(ds - A(s)dt)^2}{2D(s)dt}\right\}. \tag{3}$$

While the propagator is not measurable directly from our simulations, we can constrain its properties by comparison with a simple statistical model. If the PDF of $s$ is exactly Gaussian, the only allowed continuous Markov process is an Ornstein–Uhlenbeck (OU) process, for which the characterizing functions take the form

$$A(s) = [s - s_0]/\tau_{\text{ev}} \quad \text{and} \quad D(s) = 2\sigma_v^2/\tau_{\text{ev}}, \tag{4}$$

where $\tau_{\text{ev}}$ is a characteristic timescale for the system to evolve, and $s_0 = s_0$, and $\sigma_v^2$ set the peak and the width of the PDF as in Equation (1) (Doob 1942; Gillespie 1996). The $\Delta s$ measured from our simulations is the change in $s$ over a fixed time interval given a final value of $s_2$, which in the OU case results in a linear relation

$$\Delta s(s_2, \Delta t) = [1 - e^{-\Delta t/\tau_{\text{ev}}}] [s_2 - s_0]. \tag{5}$$

Note that this relation approaches $s_2 - s_0$ at long times, which is the value expected if $s_2$ is uncorrelated with the value of log density at the beginning of the interval.

However, Equation (5) is not a good fit to the measured value of $\Delta s$ as a function $s_2$, which is shown in the top panel of Figure 2. Instead, the convergence toward $\Delta s_2 = s - s_0$ occurs much faster at higher values of $s_0$ than at lower values of $s_2$. In other words, material with large densities experiences changes much faster than material with lower densities, likely because the period between compressions by shocks is long compared to the timescale for the dispersal of the shocked material (Klessen et al. 2000; Vazquez-Semadeni et al. 2005; Glover & Mac Low 2007; Robertson & Goldreich 2018). This behavior can only be captured by moving beyond Equation (4).

A temporally homogeneous continuous Markov processes can also be expressed in terms of a Langevin equation as (Gillespie 1991):

$$s(t_0 + dt) = s(t_0) + A(s)dt + \mathcal{N}(0, 1) \times [D(s)dt]^{1/2}, \tag{6}$$

where $\mathcal{N}(0, 1)$ is a random number drawn from a normal distribution with mean of zero and variance of one. Note that this is a random walk as a function of time, rather than as a function of scale, as in Hopkins (2013). While recovering an exactly Gaussian PDF requires $\tau_{\text{ev}}$ to be constant, we are able...
to reproduce a nearly Gaussian PDF while also reproducing the observed evolution with density by replacing \( \tau_{ev} \) with a timescale that decreases with increasing \( s \), and shifting \( s_0 \) to preserve the position of the peak of the PDF. This allows us to systematically measure the timescale of turbulent density changes, which we find are well modeled by the empirical formula

\[
\tau_{ev}(s) = \frac{\tau_{edd}(s)}{3} \left[ 1 - \frac{1}{\pi} \arctan \left( \frac{s - s_0}{2} \right) \right].
\]

with \( s_0 = \frac{5}{4} s_0^* \). This varies slowly between \( \tau_{edd}/3 \) in underdense regions and 0 in the limit of very large \( s \), consistent with previous studies showing that the lifetimes of the densest regions are comparable to their (extremely short) local sound crossing times (e.g., Robertson & Goldreich 2018).

We evolved 200,000 particles based on Equations (6) and (7) and used these to predict the resulting \( \Delta s (s_2, \Delta t) \) and PDF. These results are presented in Figure 2, providing a good approximation to the evolution of \( \Delta s \) as a function of time and Mach number. Using the driving routines from Federrath et al. (2010), we ran an additional \( M = 3.3 \) hydro simulation with 50% of the energy in compressible modes, and again found a good fit to the PDF and \( \Delta s \) evolution with \( s_0 = 0.7 \) and Equation (7) unchanged. Thus \( \tau_{ev} \) is independent of \( M \) and the same for mixed and solenoidal driving, and for all cases it is significantly less than \( \tau_{edd} \). In fact, over the \( s \approx s_0 \pm 2 \) range in which we have made our measurements \( \tau_{ev}(s) \) varies from \( \approx \tau_{edd}/4 \) to \( \approx \tau_{edd}/12 \). Meanwhile, the effect of the Mach number is to increase the width of the PDF and shift its peak as required by mass conservation.

4. Gravity and Star Formation

Calculating the rate at which stars form requires extending the model to include gravitational collapse. Current analytic models do this by assigning a critical overdensity, \( s_{crit} \), above which gravity overcomes turbulent motions. In this case, the normalized star formation rate per free fall time becomes

\[
\text{SFR}_{ff} = \frac{s_{ff}}{\phi_{ff}} \int_{s_{ff}}^{s_{crit}} ds \frac{P_M(s)}{t_{ff}(s_{ff})},
\]

where \( \phi_{ff} \) is a “replenishment factor” that accounts for the timescale over which the log-normal is assumed to be replenished, and the efficiency factor \( \epsilon_{collapse} \) accounts for the fraction of the mass with density larger than \( s_{crit} \) that actually collapses. Various models in the literature make different assumptions about \( s_{crit} \) and \( s_{ff} \) (Krumholz & McKee 2005), \( s_{ff} = s_{crit} \) Padoan & Nordlund (2011) or \( s_{ff} = s \) (Hennebelle & Chabrier 2011; Federrath & Klessen 2012), calibrating \( \epsilon_{collapse}/\phi_{ff} \) to direct numerical simulations. Recently, Burkhart (2018) developed a model in which the log-normal PDF was combined with a high-density power-law tail in a continuous and differentiable way, which has the effect of leading to higher star formation rates for lower power-law slopes, as seen in observations and simulations.

This behavior also arises naturally through a simple extension of our Markov model. If we approximate each
perturbation as a constant density sphere collapsing from rest from a large initial radius, conservation of energy gives \( ds/dt = (24\pi G\rho)^{1/2} \). Adding this term to Equation (6), we arrive at a Langevin equation that describes supersonic, isothermal turbulence including self-gravity:

\[
s(\tilde{t}_0 + \tilde{d}) = s(\tilde{t}_0) + \tilde{A}(s)\tilde{d} + N(0, 1) \times [\tilde{D}(s)\tilde{d}]^{1/2} + (90/\alpha)^{1/2} L_t e^{1/2} \tilde{d},
\]

where \( \tilde{A} \equiv \tau_{edd} A, \tilde{D} \equiv \tau_{edd} D \) and \( \tilde{t} \equiv t/\tau_{edd} \) are dimensionless such that \( \tilde{A}(s) \) and \( \tilde{D}(s) \) are purely a functions of \( s, \alpha \equiv 5\sigma_t^2 R/GM \) is the virial parameter with \( R \) and \( M \) the cloud radius and mass, and \( L_t \) is the turbulent driving scale in units of the cloud radius. In the left panel of Figure 3, we compare the results of this modified Langevin equation with full three-dimensional gravoturbulent simulations (Kritsuk et al. 2010). These simulations drive turbulence to a steady state with \( M = 6 \) and a corresponding log-normal distribution with \( \sigma_s = 1.73 \), and then “turn-on” the force of gravity abruptly with \( t_{ff,0} = 0.7 \tau_{edd} \).

To model this evolution, we start with material distributed according to a log-normal PDF and allow it to evolve according to Equation (9), multiplying the gravity term by \( \theta(t - t_{ff,0} e^{-s/2}) \) to approximate the abrupt turn-on of gravity. Note that our goal here is to produce a model that simply reproduces the physics included in such simulations, not to argue that such simulations capture the full physics of star formation. The resulting evolution is shown in the right panel of Figure 3. Here we see a power law developing toward the high-density part of the PDF whose slope matches the numerical results.

Typically, star formation simulations convert gas into sink particles representing collapsing cloud cores once they cross a threshold value. The normalized star formation rate per free fall time (SFRff) is then computed as the rate at which gas is converted into sink particles divided by the mass in the simulation. In the Langevin approach, this corresponds to placing an absorbing barrier at a threshold density and measuring the rate at which particles cross it. This method cannot pick out a physical scale, but it can give an overall rate.

The right panel of Figure 3 shows our predicted star formation rate compared to full simulations, with SFRff computed when 20% of the gas is converted into stars. We note that different groups adopt different threshold densities, and we have adjusted our comparisons appropriately. Furthermore, for each simulation set, we have only picked the highest-resolution run. For example, we compare our results to the simulations presented in PHN12 that have highest Alfvén Mach numbers at \( M = 10 \), as we focus on pure hydro simulations. From the simulations presented in F15, we choose the pure hydro simulation without the implementation of jets. We varied the absorbing barrier in our Markov estimate of \( t_{ff,0} \) and converted into sink particles divided by the mass in the simulation. The maximum resolution of each simulation is indicated in the legend. The points are color coded based on their Mach number, where \( M = 3, 5, 9, 10, 11, \) and 50 is shown in red, orange, green, blue, navy, and purple, respectively. To match the values used in the simulations, we have imposed absorbing barriers at log densities of \( \rho = 9 \) (Padoan & Nordlund 2011), 11.5 (Padoan et al. 2012), 5.8 (M = 3), 6 (M = 9, 10, 11), 4.1 (M = 50) (Federrath & Klessen 2012), and 8.75 (Federrath 2015). The dashed and dotted-dashed lines envelope factors of 2 and 4 from the one-to-one relationship.
5. Discussion

The analysis here has been limited to models of star formation due to gravity acting within molecular clouds that are purely hydrodynamic and constantly driven in near-virial equilibrium. Yet, the overall method is quite general and can be applied to alternative pictures. For example, a similar analysis could be used to quantify the evolution of the density PDF in clouds undergoing global hierarchical collapse (e.g., Vázquez-Semadeni et al. 2017), impacted by protostellar jets (e.g., Cunningham et al. 2018), or supported by dynamically important magnetic fields (e.g., Padoan et al. 2012; Federrath 2015).

The method presented here also provides a language for expressing the motion of material within pseudo steady state systems that display a range of both densities and temperatures. In this case, the evolution is governed by two vector functions, $A(s)$ and $D(s)$, which determine the rate at which material advects and diffuses within the two-dimensional phase space of temperature and density as

$$\Pi(ds|dt, s) = \frac{e^{-\frac{1}{2\Delta t_s D_1(s)}(ds - A(s))^2}}{\sqrt{2\pi D_1(s)}} \frac{e^{-\frac{1}{2\Delta t_s D_2(s)}(ds - A(s))^2}}{\sqrt{2\pi D_2(s)}}, \quad (10)$$

where $s \equiv (\ln \rho, \ln T)$. These functions can be measured from simulations in a similar manner to the one applied in Section 3 by computing $\Delta s(s)$, the change in $s$ given its final value at the end of a time interval, $\Delta t$. From these constraints, two-dimensional Langevin equations can be constructed and applied to address issues such as molecule formation (e.g., Cunningham et al. 2018), impacted by protostellar jets (e.g., Cunningham et al. 2018), or supported by dynamically important magnetic fields (e.g., Padoan et al. 2012; Federrath 2015).

Our work is related to generative models of log-normal distributions, in which a multiplicative process, $\mathcal{X} = F \mathcal{X}_{f-1}$, leads to a log-normal distribution for a random variable $\mathcal{X}$ if $F$ is log-normally distributed. Furthermore, trivial variations of a multiplicative process can generate power-law and log-normal distributions, as well as log-normal distributions with power-law tails (Mitzenmacher 2004). The possibility of re-formulating our work in this way (e.g., Basu & Jones 2004) is left for future investigation.

We thank Tom Abel, Marcus Brüggen, Paolo Padoan, and Enrique Vázquez-Semadeni for useful discussions, and the referee, Mark Krumholz, for his detailed comments. This work was supported by NSF grant AST14-07835 and NASA theory grant NNX15AK82G. We thank the Texas Advanced Computing Center (TACC) and the Extreme Science and Engineering Discovery Environment (XSEDE) for providing HPC resources via grant TG-AST130021.

### ORCID iDs

Evan Scannapieco @ https://orcid.org/0000-0002-3193-1196
Mohammadthar Safarzadeh @ https://orcid.org/0000-0002-1827-7011

### References

Basu, S., & Jones, C. E. 2004, MNRAS, 347, L47
Burkhardt, B. 2018, arXiv:1801.05428
Cox, D. P. 2005, ARA&A, 43, 337
Cunningham, A. J., Krumholz, M. R., McKee, C. F., & Klein, R. I. 2018, MNRAS, 476, 771
Doob, J. L. 1942, AnMat, 43, 351
Einfeldt, B., Roe, P. L., Munz, C. D., & Sjogreen, B. 1991, JCoPh, 92, 273
Elmegreen, B. G. 2000, ApJ, 530, 277
Esvaran, V., & Pope, S. B. 1988, CF, 16, 257
Falceta-Gonçalves, D., & Lazarian, A. 2011, ApJ, 735, 99
Federrath, C. 2015, MNRAS, 450, 4035
Federrath, C., & Klessen, R. S. 2012, ApJ, 761, 156
Federrath, C., Roman-Duval, J., Klessen, R. S., Schmidt, W., & Mac Low, M. M. 2010, A&A, 512, A81
Fryxell, B., Olson, K., Ricker, P., et al. 2000, ApJS, 131, 273
Gillespie, D. T. 1991, Markov Processes: An Introduction for Physical Scientists (Amsterdam: Elsevier)
Gillespie, D. T. 1996, PhilRvE, 54, 2084
Glover, S. C. O., & Mac Low, M.-M. 2007, ApJ, 659, 1317
Gray, W. J., & Scannapieco, E. 2016, ApJ, 818, 198
Gray, W. J., & Scannapieco, E. 2017, ApJ, 849, 132
He, G., Jin, G., & Yang, Y. 2017, AnRFM, 49, 51
Hennebelle, P., & Chabrier, G. 2011, ApJL, 743, L29
Hennebelle, P., & Falgarone, E. 2012, A&A, 540, 1653
Hopkins, P. F. 2013, MNRAS, 430, 1653
Klessen, R. S., Heitsch, F., & Low, M.-M. M. 2000, ApJ, 535, 887
Kritsuk, A. G., Norman, M. L., & Wagner, R. 2010, ApJL, 727, L20
Krumholz, M. R., & McKee, C. F. 2005, ApJ, 630, 250
Lee, D. 2013, JCoPh, 243, 269
Mitzenmacher, M. 2004, Internet Mathematics, 1, 193
Murray, N. 2011, ApJ, 729, 133
Oppenheimer, B. D., Segers, M., Schaye, J., Richings, A. J., & Crain, R. A. 2018, MNRAS, 474, 4740
Ostriker, E. C., Stone, J. M., & Gammie, C. F. 2001, ApJ, 546, 980
Padoan, P., & Nordlund, Å. 2011, ApJ, 730, 40
Padoan, P., Haugbølle, T., & Nordlund, Å. 2012, ApJL, 759, L27
Padoan, P., Nordlund, Å., & Jones, B. J. T. 1997, MNRAS, 288, 145
Pan, L., & Scannapieco, E. 2010, ApJ, 721, 1765
Price, D. J., Federrath, C., & Brun, C. M. 2011, ApJL, 727, L21
Richings, A. J., Schaye, J., & Oppenheimer, B. D. 2014, MNRAS, 440, 3349
Robertson, B., & Goldreich, P. 2018, ApJ, 854, 88
Safarzadeh, M., & Scannapieco, E. 2016, ApJL, 832, L9
Toro, E. F., Spruce, M., & Speares, W. 1994, ShiWav, 4, 25
Vázquez-Semadeni, E., Gonzalez-Samaniego, A., & Colín, P. 2017, MNRAS, 467, 1313
Vázquez-Semadeni, E., Kim, J., Shadmehri, M., & Ballesteros-Paredes, J. 2005, ApJ, 618, 344
Walch, S., Girichidis, P., Naab, T., et al. 2015, MNRAS, 454, 238