Supplement for “Four Projects in Astrophysical Magnetohydrodynamics”

PI: David C. Collins (Florida State University)

We are requesting 151,000 SUs on *Stampede 2* and 76,800 Gb on Ranch to supplement our existing ACCESS allocation. Our allocation supports four projects in astrophysical plasmas. One of which is a study of the statistics of supersonic isothermal turbulence, wherein we have developed a number of interesting analytic models for the correlations between density and velocity in such flows. Our previous ACCESS request was conservative as we finished the models and performed preliminary runs. We have now finished the models, performed the preliminary simulations, and published the first paper on the topic. It is now time to complete the study.

We recognize that this request is somewhat larger than our initial request, and that *Stampede 2* is nearing the end of its life. We are confident that we can perform the simulations and analyze the results in a matter of weeks from the approval of this supplement. We are proposing 12 simulations that will each take about 9 days on 64 KNL nodes, and they will be run in three batches of four concurrent simulations. Thus the simulations will be completed by the end of March 2023.

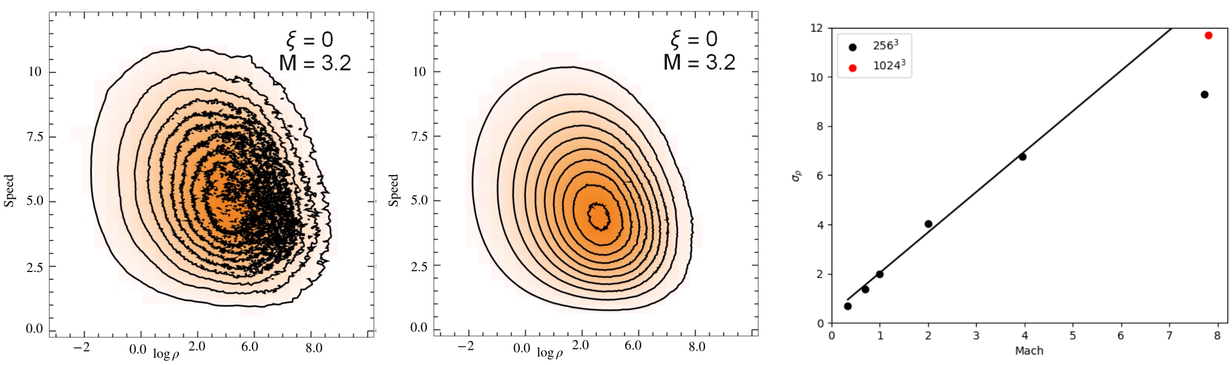

Figure 1: *(Left, Center)* The joint distribution between density (horizontal) and velocity (vertical) for low resolution preliminary runs. For uncorrelated variables, the joint distribution would be symmetric, while here we have a pronounced tilt. We have analytic models that reproduce these contours accurately. The left panel shows an averaging window of $30 t_{\rm dyn}$, while the right has averaged over 100 $t_{\rm dyn}$. Higher resolution can be run for only $10 t_{\rm dyn}$and achieve the same signal-to-noise. *(Right)* The relation between the Mach number and the variance in density, σ_ρ. Black points show low resolution runs at 256^3, while the red point was run with 1024^3 zones. The low resolution runs fail to reproduce the expected linear trend at large Mach numbers.

Motivation

Supersonic Isothermal Turbulence is ubiquitous in astrophysical flows. The low density of gas in the interstellar medium (ISM) allows for it to cool very effectively, so the temperature stays roughly constant. Explosions of supernovae drive flows with velocities more than 10 times the speed of sound in the ISM. An important aspect of isothermal flows is their high degree of compressibility, causing large swings in density. This is essential for processes like the formation of stars. We have developed analytic models of the relationship between density, ρ, and velocity, v, that will be useful in theoretical treatments of the interstellar medium. We now hope to verify these analytic models with high resolution simulations. Preliminary low resolutions have been successful, but the need for high resolution and statistically well-sampled turbulence is clear.

The simulations to be performed will have periodic boundary conditions and begin with a uniform density. Turbulence is generated by adding a fixed driving pattern to the velocity at large scale, and the chaos of the fluid dynamics distributes that energy to smaller and smaller scales until the energy is dissipated by molecular viscosity (or an appropriate approximation).

We have developed analytic models for several quantities in such flows. The probability distribution function for both density and velocity are traditionally modeled as a lognormal and Maxwellian, respectively, but our first paper has shown that neither is accurate, and we have new models for these PDFs. The next thing we are verifying is the joint PDF of density and velocity, $f_{\rho,v}(\rho, v)$. This can be seen in Figure 1. If density and velocity are uncorrelated, which is often assumed, the joint PDF can be written as the product of marginalized PDFs, $f_{\rho,v}(\rho, v) = f_\rho(\rho) f_v(v)$. Our preliminary low resolution simulations show that this is not the case, and we have developed a correction function, $g(\rho, v)$ such that

$$f_{\rho,v}(\rho, v) = f_\rho(\rho) f_v(v) + g(\rho, v).$$

This will allow us to make accurate predictions of the density and its moments based only on moments of the velocity.

Supersonic isothermal turbulence can be parametrized by two free parameters: the Mach number $\mathcal{M}_S$, which is the ratio of velocity to the local sound speed; and the forcing parameter, ξ, which determines the amount of compressive motions in the driving velocity pattern. ξ ranges between 0 and 1, with 0 representing only rotational motions, and 1 denoting only compressive motions. Typical mach numbers in the interstellar medium range from 1 to about 10, though much faster flows do occur. More compressibility in the driving pattern alters the density statistics and thus the correlations, so it is essential for our models to account for both parameters.

Our preliminary runs consisted of twelve low-resolution simulations at 256^3 and five high-resolution runs at 1024^3. The low-res runs used $\mathcal{M}_S$=(1,2,4,8) and $\xi = (0, 0.5, 1)$. The high-res run used a fixed value of $\xi = 0.5$. Fixing ξ allowed us to verify a subset of parameter space and explore resolution effects. Our low resolution study shows that our model is accurate for all values of ($\mathcal{M}_S, \xi$) that were appropriate for that resolution, but the high-res runs show that these are not numerically converged. Now it is time to run $\xi = 0, 1$ to verify that the models are fully general.

We will verify our models using a suite of twelve 1024^3 simulations. We will use $\mathcal{M}_S$ = (1,2,4,10), with forcing parameters $\xi = 0, 0.5, 1$. We will only use three values of ξ for the full suite, which brackets the range of possible parameters. This range of Mach numbers is important as the flow properties change from weakly supersonic $\mathcal{M}_S$=1 to highly supersonic $\mathcal{M}_S$=10. Higher values of $\mathcal{M}_S$ are of interest, but less common so will be omitted at this time.

A resolution of 1024^3 is necessary to properly reproduce the density statistics. Low resolution simulations have large zones, thus it is difficult to get large densities, which in turn impacts the mean and variance measured for a given Mach number. This can be seen in Figure **??**, which shows the standard deviation in density, σ_ρ, vs. the Mach number for a suite of low resolution runs (black points) and our recently run 1024^3 (red point). It is theoretically expected that the two are linearly related, $\sigma_\rho = b\mathcal{M}_S$, but as can be seen from the Figure, the low resolution runs fall short. The high resolution run is more in line with the linear relation. Even higher resolution than 1024^3 would be useful and interesting, but as the cost scales like the fourth power of size, we cannot justify such an expense at this time.

Simulations will be run for $10t_{\rm dyn}$, where the dynamical time $t_{\rm dyn}$=$0.5/\mathcal{M}_S$. This is the time for the driving flow to cross the box, which has a pattern size of one half of the box. This integration time is important to reduce the noise in the joint distributions. This can be seen in Figure 1, which shows the joint $\rho - v$ distribution, $f_{\rho,v}(\rho, v)$. First observe the center panel. If the two were uncorrelated, the distribution would be symmetric, instead of skewed. Our analytic formula reproduces these contours with a few-percent accuracy. The effect of time integration can be seen by comparing the left and center panels. The left panel was averaged over 50 $t_{\rm dyn}$, while the right was averaged over 100 $t_{\rm dyn}$, substantially increasing the signal-to-noise. The high resolution runs, having more data points, achieve convergence about an order of magnitude faster, so we will only run the high resolution runs for $10t_{\rm dyn}$. Since the noise decreases roughly like shot noise, $1/\sqrt{t}$, further integration improves the noise slowly.

Table 1: The accounting for the request. The total cost for each simulation, $SU = \frac{N_Z N_U N_N}{N_C \zeta} \frac{1}{3600}$. For the proposed 12 simulations, $N_Z = 1024^3$, $N_N = 64$, $N_C = 4096$, $\zeta = 10^5$ =(core updates)/(processor second). $N_U = T/\Delta t$ is the number of steps, and depends on $\mathcal{M}_S$ as described in the text. $T_{wall} = SU/N_N$ is presented in hours.

M_S	ξ	T	ΔT	N_U	T_{wall}	SU
1.0	0.0	5	1.4×10^{-5}	3.7×10^{5}	269.7	1.7×10^{4}
2.0	0.0	2.5	9.0×10^{-6}	2.8×10^{5}	202.3	1.3×10^{4}
4.0	0.0	1.25	5.4×10^{-6}	2.3×10^{5}	168.6	1.1×10^{4}
10.0	0.0	0.5	2.5×10^{-6}	2.0×10^{5}	148.3	9.5×10^{3}
1.0	0.5	5	1.4×10^{-5}	3.7×10^{5}	269.7	1.7×10^{4}
2.0	0.5	2.5	9.0×10^{-6}	2.8×10^{5}	202.3	1.3×10^{4}
4.0	0.5	1.25	5.4×10^{-6}	2.3×10^{5}	168.6	1.1×10^{4}
10.0	0.5	0.5	2.5×10^{-6}	2.0×10^{5}	148.3	9.5×10^{3}
1.0	1.0	5	1.4×10^{-5}	3.7×10^{5}	269.7	1.7×10^{4}
2.0	1.0	2.5	9.0×10^{-6}	2.8×10^{5}	202.3	1.3×10^{4}
4.0	1.0	1.25	5.4×10^{-6}	2.3×10^{5}	168.6	1.1×10^{4}
10.0	1.0	0.5	2.5×10^{-6}	2.0×10^{5}	148.3	9.5×10^{3}
					Total SU	1.5×10^{5}
					Disk (Gb)	7.7×10^{4}

Timing

The total cost for each simulation can be found as $SU = T_{wall} N_N$, where

$$T_{wall} = \frac{N_Z N_U}{N_C \zeta} \frac{1}{3600}. \tag{1}$$

SU is the cost for the simulation, T_{wall} is its duration, N_N is the number of nodes, N_Z is the number of zones, N_U is the number of updates, N_C the number of cores, and ζ = (zone-updates)/(core-second) is the performance of the code. From the scaling study, we find that $\zeta = 10^5$. Due to the excellent scaling of fixed-resolution Enzo, we will use $N_C = 4096$. We will use 64 cores per node and 4096 cores, so $N_N = 64$. N_Z is set by our target resolution of 1024^3. N_U is the number of updates, which depends on M_S in the following way.

The number of updates is found by $N_U = T_{sim}/\Delta t$, the total time over the size of a step. $T_{sim} = 10 t_{\rm dyn}$ per our noise requirement. The time step size, Δt, is determined by a typical Courant condition that the signal cannot propagate more than half a zone in a timestep,

$$\Delta t = \eta \frac{\Delta x}{v_{max} + c_s} \propto \frac{1}{1 + M_S} \tag{2}$$

where Δx is the zone size, and $v_{max} + c_s$ is the maximum signal speed over the whole domain. It was verified with our suite of Mach 8 1024^3 runs that Δt does not depend on the forcing parameter, ξ. The peak velocity, v_{max}, is not predictable due to the chaos of the turbulence, but is thankfully found to scale with the Mach number, and we calibrate to our recent high res simulations. Table 1 shows a breakdown of each simulation, the total time, time step size, wall time, and SUs.

We are also requesting 76,800 Gb of long term storage on Ranch to store these simulations. For each simulation we will store 10 frames per $t_{\rm dyn}$. Each frame contains 1024^3 double precision values for each of 8 fields (density, 3 components of velocity, and 3 components of the driving field, and the internal energy) for a total of 64Gb per output. This gives a total of 76800 Gb for the whole suite. With our existing 96 Tb archive, this gives a total request of 1.8×10^5 Gb.