

The Dust Between Us and the Big Bang

PI: David C. Collins (Florida State University)

We are requesting 2.4×10^5 SUs on Frontera and 1.8×10^5 Gb on the Ranch archival system in order to perform several studies of the interstellar medium (ISM). This is part of an effort to observe gravitational waves from the big bang, imprinted in the cosmic microwave background (CMB). These observations are hampered by the dust and plasma in our own galaxy, which outshines the CMB. We will simulate the plasma in the ISM in two ways to characterize its observational properties. One study will be high resolution simulations of plasma turbulence, and the other suite will be simulations of isolated galaxies. The turbulence simulations give us a detailed study of a typical patch of the ISM, while the galaxy simulations will study the large scale structure. The disk requests represents the newly created simulations as well as our archive of turbulence and star formation simulations.

1 Background

Gravitational waves generated at the beginning of the Universe leave an imprint in the polarization of the cosmic microwave background (CMB). The CMB is the oldest light observable in the Universe, and gives us a snapshot of its very early stages. Observing the polarization of the CMB will give us a measurement of gravitational waves that were produced at the formation of the Universe. Unfortunately, the dusty plasma in our own Milky Way also produces polarized light that is brighter than the primordial signal. Our simulations will study the polarization properties of the light from the plasma in Milky Way type galaxies. Understanding this polarization is essential for its removal from future measurements of the CMB.

The Planck satellite (Planck Collaboration et al. 2020) measured the polarization of the entire sky. Polarization, having direction, needs two quantities for its description. The most useful option are the parity-even E -mode and parity-odd B -mode. The only source of B -mode polarization in the CMB are gravitational waves, so searching for B -modes in the CMB will prove quite profitable. However, the dusty, magnetized plasma in our own galaxy also produces B -mode polarization. So this must be understood so it can be removed.

The Planck measurements found that both E - and B -mode are relatively uniform power-laws in wavenumber, with spectra given by

$$C_\ell^{EE} = A^{EE} k^{\alpha^{EE}}, \quad (1)$$

for E -mode and a similar expression for B -mode. They found that the slopes are roughly equal, $\alpha^{EE} \sim \alpha^{BB} \sim -2.5$, and the B -mode have roughly half the power of the E -mode, so $A^{EE}/A^{BB} \sim 2$. This signal has three components: the CMB signal that we are interested in; thermal dust emission near the Sun; and high altitude synchrotron electrons. Both the dust and the synchrotron electrons are aligned by the magnetic field that threads the galaxy, and the light they emit is thus polarized.

A curious finding in the Planck data is a correlation between the total signal, T -mode, which is parity-even, and the parity-odd B -mode. In principle, these should be uncorrelated, but the Planck satellite measured a non-zero correlation at the 5% level. This TB may come from some unseen property of the plasma, or it could come from large scale structure in the local magnetic field.

With a suite of MHD turbulence simulations run on *Stamped 2*, we have shown (Stalpes et al 2023, in prep) that the slopes of the polarization are directly related to the state of the turbulent gas, specifically the r.m.s. velocity and magnetic field strength. Our suite of 21 simulations varied the Mach number, $\mathcal{M} = v/c_s$, the ratio of r.m.s. velocity to sound speed, and Alfvén mach number, $\mathcal{M}_A = v/v_A$, where v_A is the speed of magnetic disturbances, in a suite of driven turbulence simulations. Increasing \mathcal{M} makes for more filamentary structure in the cloud, and slopes consistent with what Planck observed, and increasing magnetic field suppresses small scale structure. Our suite employed $\mathcal{M}=(\frac{1}{2}, 1, 2, 3, 4, 5, 6)$ and $\mathcal{M}_A=(\frac{1}{2}, 1, 2)$. We perform synthetic observations of polarized emission of the gas, and find that both E and B are distributed like a power law, and their slopes are roughly linearly related to

\mathcal{M} and \mathcal{M}_A . After inverting our linear relationship, we predict that the observed values of α^{EE} and α^{BB} are best matched by plasma with $\mathcal{M}=4.7$ and $\mathcal{M}_A=1.5$. While the ISM has many patches of gas with a variety of parameters, our preliminary this represents a “typical” patch of sky. Our new simulations will verify this prediction, and provide high resolution mock datasets with similar spectral properties to the observed microwave sky. The study presented in Stalpes et al were runs were at a modest resolution of 512^3 with limited inertial range. Our simulations will improve the resolution to 2048^3 which will improve the fidelity of the results, and test our prediction of the typical \mathcal{M} and \mathcal{M}_A of the local galaxy.

The results from Stalpes et al (2023) show that turbulence alone can account for a TB correlation only at the 2% level at best. These were relatively low resolution studies, it is possible that the turbulent correlation is lower. This would imply the TB correlation seen in the sky comes from large scale features in the dust and magnetic field morphology. To properly test this hypothesis, we will perform simulations of Milky Way sized disk galaxies that have high resolution where the dust is, and also capture the circumgalactic medium (CGM) around the galaxy. Including both the disk and the CGM is essential as the two form a coupled system, and at present it is not clear how far above the galaxy the synchrotron electrons reside. The dust and star formation primarily resides in a thin disk, with a scale height of 100 pc. The interstellar medium is a dynamical environment, driven by supernova explosions and star formation. These explosions expel material from the disk into the CGM, which then cools and falls back to the disk, providing fuel for more star formation. Thus to properly simulate the morphology of the magnetic field in and around a galaxy, one must also properly model the surrounding CGM. All of these pieces are important for a complete picture of the polarized sky.

We propose two suites of simulations. The first is a continuation of the moderate resolution turbulent boxes presented in Stalpes et al (2023). In that study, we predict that a Mach number of 4.7 and an Alfvén Mach number of 1.5 reproduce the sky. We will perform two simulations at 2048^3 . These simulations will test our prediction, and provide high quality datasets that reproduce the salient features of the sky that can be used to test foreground removal algorithms and ISM models. While these simulations are quite large, the excellent scaling of our code on Frontera and our years of experience guarantee success. The second suite of simulations is a set of three isolated galaxies that will simultaneously provide high resolution on the mid plane, and properly resolve the environment and boundary conditions of the galaxy. These will be used initially to test CMB foreground models, and will also have many applications beyond the CMB. These galaxies will be a stack of nine fixed resolution levels at roughly 512^3 per level, with the finest level concentrated in a thin disk. Employing fixed resolution and consistent grid size allows us to optimize the performance of the simulations.

We will introduce the code in Section 2, outline and motivate the simulation design in Section ??, and provide a detailed accounting of the performance of the simulations in Section 4.

2 Methods

The code we will use is Enzo (Bryan et al. 2014; Collins et al. 2010), an open source code that has been used for a number of astrophysics applications (e.g. Abel et al. 2002; Correa Magnus et al. 2023). Enzo is an adaptive mesh refinement (AMR) code that can dynamically add resolution elements as the system requires it, using the strategy of Berger & Colella (1989) and Balsara (2001). We will use the constrained transport (CT) module (Collins et al. 2010; Gardiner & Stone 2005) that conserves the divergence of the field to machine precision. It uses FFT-based gravity for the root grid and multigrid relaxation for gravity on fine grids. The base MHD solver is a higher order Godunov method. We will use Grackle (Smith et al. 2017) to handle the chemistry and thermodynamics.

3 Simulations

We will perform two suites of simulations, a suite of driven *turbulence* simulations and a suite of isolated *galaxy* simulations.

3.1 Turbulence

In a turbulent system, such as the ISM or stirring cream in one’s coffee, energy is injected at a scale, L , and the fluid dynamics shreds that energy into a cascade of ever decreasing size scale until the dissipation length is reached. The dissipation length scale, $\ell_D \propto \mathcal{M}^{-3/4}$, decreases as the mean kinetic energy increases. There are three primary regimes in a turbulent system; the large scale driving, the intermediate scale *inertial range*, and the small scale dissipation. The driving scale is a boundary condition, and the dissipation is done by the numerics. The inertial range is where simulations can teach us about nature, as it does not depend on the user-supplied boundary conditions. In simulations, the inertial range is proportional the linear resolution. We have performed a suite of simulations of the ISM at a resolution of 512^3 , and now wish to build on that success with two simulations at 2048^3 .

By using these linear relations found in our preliminary work, we predict that the ISM has a “typical” value of $\mathcal{M} = 4.7$ and $\mathcal{M}_A = 1.5$. This prediction was done at low resolution with limited inertial range. We will further test this prediction against high resolution simulations, which further separate the boundary condition and numerical shortcomings from the signal of interest.

The actual ISM is filled with many clouds with a range of Mach numbers, so we will explore this predicted “typical” cloud as well as one additional cloud at higher Mach number. These two simulations aim to maximize the useful portion of the inertial range. The first simulation will be at the typical $\mathcal{M} = 4.7$ and $\mathcal{M}_A = 1.5$, and the second will be at $\mathcal{M} = 8$ and $\mathcal{M}_A = 1.5$. This higher Mach number is challenging to do at moderate resolution, as the length of the shocks and the dissipation scale gets close to the grid resolution. The temptation to repeat the whole suite of 21 simulations is large, especially considering the poor inertial range of the lower Mach number runs; however this would be quite expensive in terms of disk and SU usage.

The simulations will be run for long enough to gather meaningful statistics. The relevant timescale is the dynamical time, $t_{dyn} = L/\mathcal{M}$, where L is the size of the velocity pattern, $L = 1/2$. This gives the time it takes for a typical eddy to pass over itself. We will run for $5t_{dyn}$. It takes about $2t_{dyn}$ to establish a steady-state as the dissipation structures are created by the cascade. We will then run for 3 additional dynamical times to obtain a statistically well sampled set. While more time would be desirable, this should be sufficient to obtain meaningful results.

The size of these simulations will make them somewhat challenging to perform and analyze. Our recent successes outlined in the Progress document demonstrates that we are quite ready for this scale, and the Scaling document shows that Enzo will perform admirably.

3.2 Galaxy

The next generation of CMB telescopes will reach an angular resolution of 1° in the near future, and $1''$ by the end of the mission. This translates to linear size of (1.7pc, 0.03pc) at 100pc from the observer. Our near term stretch goal is to create a galaxy with 1.7pc resolution for the dust at the mid-plane of the galaxy, while also including the CGM at 1Mpc. This is somewhat beyond our reach at the moment, as an next step we will aim for 6.25pc resolution at the disk and 0.8Mpc for the outer box. We will do this with a stack of 8 levels of static AMR, with approximately 512^3 zones on each level. We will quantize our resolution into grids of 32^3 zones, rather than the dynamic allocation of grids that is Enzo’s speciality, in order to optimize memory usage and performance.

A spiral galaxy is essentially a spinning pancake of gas and stars in a ball of dark matter and hot gas. Thus our finest level will be wide and flat ($100\text{pc} \times 20,000\text{pc} \times 20,000\text{pc}$) at 6.25pc resolution. The next two level increase the aspect ratio of the refined region, increasing a factor of 3 in height and about 20% in width. The remaining levels will be cubic regions of 512^3 . We will quantize our refinement regions into blocks of 32^3 in order to optimize the performance of the simulation.

We will simulate for a total of 1Gyr. A galaxy of this size has an orbital period of about 250Myr, this gives 8 rotations of the galaxy. The initial conditions take about 500Myr to reach a roughly steady state; stellar explosions keep the galaxy from collapsing to the mid-plane, but take a few Myr to happen. We will then give the galaxy several orbital times to establish the magnetic field morphology and component distribution.

We will run three simulations. The first will be our unmagnetized fiducial run; the second will have a weak uniform magnetic field, and the third will have supernova deliver magnetic energy. From these we will examine the development of the large scale magnetic field, its impact on the star formation behavior, and synthetic observations from within for use in CMB foreground removal.

This full 8 level AMR structure has been tested on Frontera, and found to perform as expected from the scaling study. The scaling study was restricted to one level of refinement, as this is sufficient to engage all of the AMR overhead.

These simulations will be extremely useful for the study of the CMB foregrounds and star formation, and will also serve as a proving ground for our target of 1pc resolution. The next step will be to repeat this work with 1024^3 zones per level, which will give 3pc resolution. This will require some optimization of the code and layout to be performant enough, but is well within our reach. With 2048^3 zones per level, we will achieve 1pc resolution, though the expected cost of 100 million SU may prove prohibitive.

4 Request Details

The cost for each simulation is found as

$$SU = t_{wall} N_N \quad (2)$$

$$t_{wall} = \frac{\sum_{\ell} N_Z N_U}{\zeta} \frac{1}{N_C}, \quad (3)$$

where SU is the total cost; t_{wall} is the run time; N_N is the number of nodes; N_Z is the number of zones per level ℓ ; N_U is the number of updates per level; ζ is the performance measured in (zone-updates)/(processor-second); and N_C is the total number of cores. The cost, $zeta$, is determined in the scaling document. Total disk usage is found as

$$Disk = 4N_Z N_F N_D, \quad (4)$$

where each of N_Z zones contains N_F fields and we store N_D dumps for each run. The values of N_N , N_C , N_F , N_D and ζ can be found in Table 1. The values of N_Z and N_U are outlined in Table 2. We will outline the process for estimating each.

Table 1: The compute layout for each simulation suite. Shown is the number of nodes, N_N ; the number of cores, N_C ; the number of fields per zone, N_F ; the number of datasets stored per simulation, N_D ; and the performance in (zone-update)/(core-second), ζ .

Suite	N_N	N_C	N_F	N_D	ζ
<i>turbulence</i>	586	32,768	17	10	4×10^5
<i>galaxy</i>	32	512	27	20	2×10^5

The number of zones, N_Z , is computed from the problem geometry. All of these simulations will use fixed or static resolution, so N_Z is known. For the *turbulence* simulations, $N_Z = 2048^3$. For the galaxy simulations, N_Z will be computed level-by-level from the grid layout. We construct each level in grid patches of 32^3 zones each. Consistent patch size allows us to optimize the memory usage and performance of the simulation. The finest level is a thin pancake of $100 \times 100 \times 1$ grids. This covers the midplane with a resolution of 6.25 pc for the entire disk, to the dust scale height of ± 100 pc and a side length of 20kpc. Resolution increases by a factor of 2 each level. The next two levels increase the aspect ratio of the refined region ($54 \times 54 \times 3$ and $30 \times 30 \times 11$ grids, respectively), and the remaining levels are $16 \times 16 \times 16$ grids. This grid structure gives an outer size of 8.2×10^5 pc on a side with 6.25pc resolution over the whole disk at the midplane.

Table 2: The compute layout for each simulation suite. Shown is the number of nodes, N_N ; the number of cores, N_C ; the number of fields per zone, N_F ; the number of datasets stored per simulation, N_D ; and the performance in (zone-update)/(core-second), ζ .

suite	$(\mathcal{M}, \mathcal{M}_\text{A})$	T	Δt	N_Z	SU	
<i>turbulence</i>	(4.7, 1.5)	0.64	1.1×10^{-6}	2048^3	6.4×10^4	
<i>turbulence</i>	(8, 1.5)	0.31	6.6×10^{-7}	2048^3	5.0×10^4	
Turbulence: Disk				$1.3 \times 10^4 \text{Gb}$	SU: 1.1×10^5	
suite	Level	grid layout	$T[\text{Gyr}]$	$\Delta t[\text{Gyr}]$	N_Z	SU
<i>galaxy</i>	8	$100 \times 100 \times 1$	2	2.3×10^{-3}	$(689)^3$	2.5×10^4
<i>galaxy</i>	7	$54 \times 54 \times 3$	2	4.6×10^{-3}	$(659)^3$	1.1×10^4
<i>galaxy</i>	6	$30 \times 30 \times 11$	2	9.1×10^{-3}	$(687)^3$	6.2×10^3
<i>galaxy</i>	5	$16 \times 16 \times 16$	2	1.8×10^{-2}	$(512)^3$	1.3×10^3
<i>galaxy</i>	4	$16 \times 16 \times 16$	2	3.7×10^{-2}	$(512)^3$	6.4×10^2
<i>galaxy</i>	3	$16 \times 16 \times 16$	2	7.3×10^{-2}	$(512)^3$	3.2×10^2
<i>galaxy</i>	2	$16 \times 16 \times 16$	2	1.5×10^{-1}	$(512)^3$	1.6×10^2
<i>galaxy</i>	1	$16 \times 16 \times 16$	2	2.9×10^{-1}	$(512)^3$	8.0×10^1
<i>galaxy</i>	0	$16 \times 16 \times 16$	2	5.8×10^{-1}	$(512)^3$	4.0×10^1
One Galaxy: Disk				$7.0 \times 10^3 \text{Gb}$	SU: 4.4×10^4	
Three Galaxies: Disk				$2.1 \times 10^4 \text{Gb}$	SU: 1.3×10^5	
Archive: Disk				$1.5 \times 10^5 \text{Gb}$		
Total: Disk				$1.8 \times 10^5 \text{Gb}$	SU: 2.5×10^5	

The number of updates, $N_U = T/\Delta t$ is estimated by setting T by the problem goals, and estimating Δt from the solver behavior and prior results. As described above, the *turbulence* simulations will run for $T = 5t_{\text{dyn}} = 5/(2\mathcal{M})$ (in dimensionless code units), to statistically resolve the turbulence. The *galaxy* simulations will run for $T = 2 \text{ Gyr}$, which represents roughly 8 orbits of the galaxy. The timestep size, Δt , is found by a standard Courant condition,

$$\Delta t = \eta \frac{\Delta x}{(v + c_f)_{\text{max}}}, \quad (5)$$

where $(v + c_f)_{\text{max}}$ is the maximum signal velocity on each resolution level. Here, v is the velocity and c_f is the fast MHD speed. It is not possible to predict the exact value of this signal speed, so we calibrate to lower resolution simulations and use Equation 5 to rescale. For the *turbulence* simulations, we calibrate Δt to the fiducial simulations in Stalpes et al (2023). For the *galaxy* simulations, we calibrate to a low-resolution preliminary galaxy.

Table 2 shows a summary of the request. The first two rows show each of the *turbulence* simulations. The remaining rows outline the *galaxy* simulations, estimating N_Z and $N_U = T/\Delta t$ and the total cost, SU . Also shown in the table is the projected disk usage for the new simulations, as well as our archive already on Ranch of recently run hydrodynamical turbulence simulations and star formation simulations.

References

- Abel, T., Bryan, G. L., & Norman, M. L. 2002, *Science*, 295, 93
- Balsara, D. S. 2001, *J. Comput. Phys.*, 174, 614
- Berger, M. J. & Colella, P. 1989, *J. Comput. Phys.*, 82, 64

- Bryan, G. L., Norman, M. L., O'Shea, B. W., Abel, T., Wise, J. H., Turk, M. J., Reynolds, D. R., Collins, D. C., Wang, P., Skillman, S. W., Smith, B., Harkness, R. P., Bordner, J., Kim, J.-h., Kuhlen, M., Xu, H., Goldbaum, N., Hummels, C., Kritsuk, A. G., Tasker, E., Skory, S., Simpson, C. M., Hahn, O., Oishi, J. S., So, G. C., Zhao, F., Cen, R., Li, Y., & Enzo Collaboration. 2014, *ApJS*, 211, 19
- Collins, D. C., Xu, H., Norman, M. L., Li, H., & Li, S. 2010, *ApJS*, 186, 308
- Correa Magnus, L., Smith, B. D., Khochfar, S., O'Shea, B. W., Wise, J. H., Norman, M. L., & Turk, M. J. 2023, arXiv e-prints, arXiv:2307.03521
- Gardiner, T. A. & Stone, J. M. 2005, *J. Comput. Phys*, 205, 509
- Planck Collaboration, Akrami, Y., Ashdown, M., Aumont, J., Baccigalupi, C., Ballardini, M., Banday, A. J., Barreiro, R. B., Bartolo, N., Basak, S., Benabed, K., Bernard, J. P., Bersanelli, M., Bielewicz, P., Bond, J. R., Borrill, J., Bouchet, F. R., Boulanger, F., Bracco, A., Bucher, M., Burigana, C., Calabrese, E., Cardoso, J. F., Carron, J., Chiang, H. C., Combet, C., Crill, B. P., de Bernardis, P., de Zotti, G., Delabrouille, J., Delouis, J. M., Di Valentino, E., Dickinson, C., Diego, J. M., Ducout, A., Dupac, X., Efstathiou, G., Elsner, F., Enßlin, T. A., Falgarone, E., Fantaye, Y., Ferrière, K., Finelli, F., Forastieri, F., Frailis, M., Fraisse, A. A., Franceschi, E., Frolov, A., Galeotta, S., Galli, S., Ganga, K., Génova-Santos, R. T., Ghosh, T., González-Nuevo, J., Górski, K. M., Gruppuso, A., Gudmundsson, J. E., Guillet, V., Handley, W., Hansen, F. K., Herranz, D., Huang, Z., Jaffe, A. H., Jones, W. C., Keihänen, E., Keskitalo, R., Kiiveri, K., Kim, J., Krachmalnicoff, N., Kunz, M., Kurki-Suonio, H., Lamarre, J. M., Lasenby, A., Le Jeune, M., Levrier, F., Liguori, M., Lilje, P. B., Lindholm, V., López-Caniego, M., Lubin, P. M., Ma, Y. Z., Macías-Pérez, J. F., Maggio, G., Maino, D., Mandolesi, N., Mangilli, A., Martin, P. G., Martínez-González, E., Matarrese, S., McEwen, J. D., Meinhold, P. R., Melchiorri, A., Migliaccio, M., Miville-Deschênes, M. A., Molinari, D., Moneti, A., Montier, L., Morgante, G., Natoli, P., Pagano, L., Paoletti, D., Pettorino, V., Piacentini, F., Polenta, G., Puget, J. L., Rachen, J. P., Reinecke, M., Remazeilles, M., Renzi, A., Rocha, G., Rosset, C., Roudier, G., Rubiño-Martín, J. A., Ruiz-Granados, B., Salvati, L., Sandri, M., Savelainen, M., Scott, D., Soler, J. D., Spencer, L. D., Tauber, J. A., Tavagnacco, D., Toffolatti, L., Tomasi, M., Trombetti, T., Valiviita, J., Vansyngel, F., Van Tent, B., Vielva, P., Villa, F., Vittorio, N., Wehus, I. K., Zacchei, A., & Zonca, A. 2020, *A&A*, 641, A11
- Smith, B. D., Bryan, G. L., Glover, S. C. O., Goldbaum, N. J., Turk, M. J., Regan, J., Wise, J. H., Schive, H.-Y., Abel, T., Emerick, A., O'Shea, B. W., Anninos, P., Hummels, C. B., & Khochfar, S. 2017, *MNRAS*, 466, 2217