Advantages of GPU-accelerated approach for solving the Parker equation in the heliosphere

S. Della Torre[b,d], G. Cavallotto[a,b], D. Besozzi[c], M. Gervasi[a,b], G. La Vacca[a,b], M. Nobile[d] and P.G. Rancoita[b]

[a]*Dipartimento di Fisica, Università di Milano-Bicocca,* [b]*INFN Sezione di Milano Bicocca,*
[c]*Dipartimento di Informatica, Sistemistica e Comunicazione, Università di Milano-Bicocca,*
[d]*Dipartimento di Scienze Ambientali, Informatica e Statistica, Università di Ca' Foscari Venezia*
[d]*Italian Research Center on High Performance Computing, Big Data and Quantum Computing, spoke 3 (Italy)*

Solar Modulation of Galactic Cosmic Rays

Solar modulation is a reduction of galactic cosmic rays (GCR) fluxes compared to flux intensities outside the heliosphere (i.e. the Local Interstellar Spectrum, LIS). The particle propagation in the interplanetary medium can be described as a process that transforms the LIS (at the heliosphere boundary) to the modulated spectra observed in the heliosphere. All relevant physical processes involved (i.e. diffusion, convection, magnetic drift and energy loss) are embedded into the Parker Transport Equation, named after Eugene Parker who first proposed it in the 1960s.

$$\frac{\partial U}{\partial t} = -\nabla \cdot (U\vec{V}) + \nabla \cdot \left[\tilde{K} \cdot \nabla U\right] + \frac{(\nabla \cdot \vec{V})}{3} \frac{\partial}{\partial T} \left(\alpha_{\mathrm{rel}} T U\right)$$

U = Cosmic Rays number density per unit interval of kinetic energy

$$\frac{\partial f}{\partial t} = -\nabla \cdot (f\vec{V}) + \nabla \cdot \left[\tilde{K} \cdot \nabla f\right] + \frac{(\nabla \cdot \vec{V})}{3p^2} \frac{\partial}{\partial p} \left(p^3 f\right)$$

f = Omnidirectional distribution function per unit interval of particle momentum

Diffusion | Convection | Drift | Energy Loss

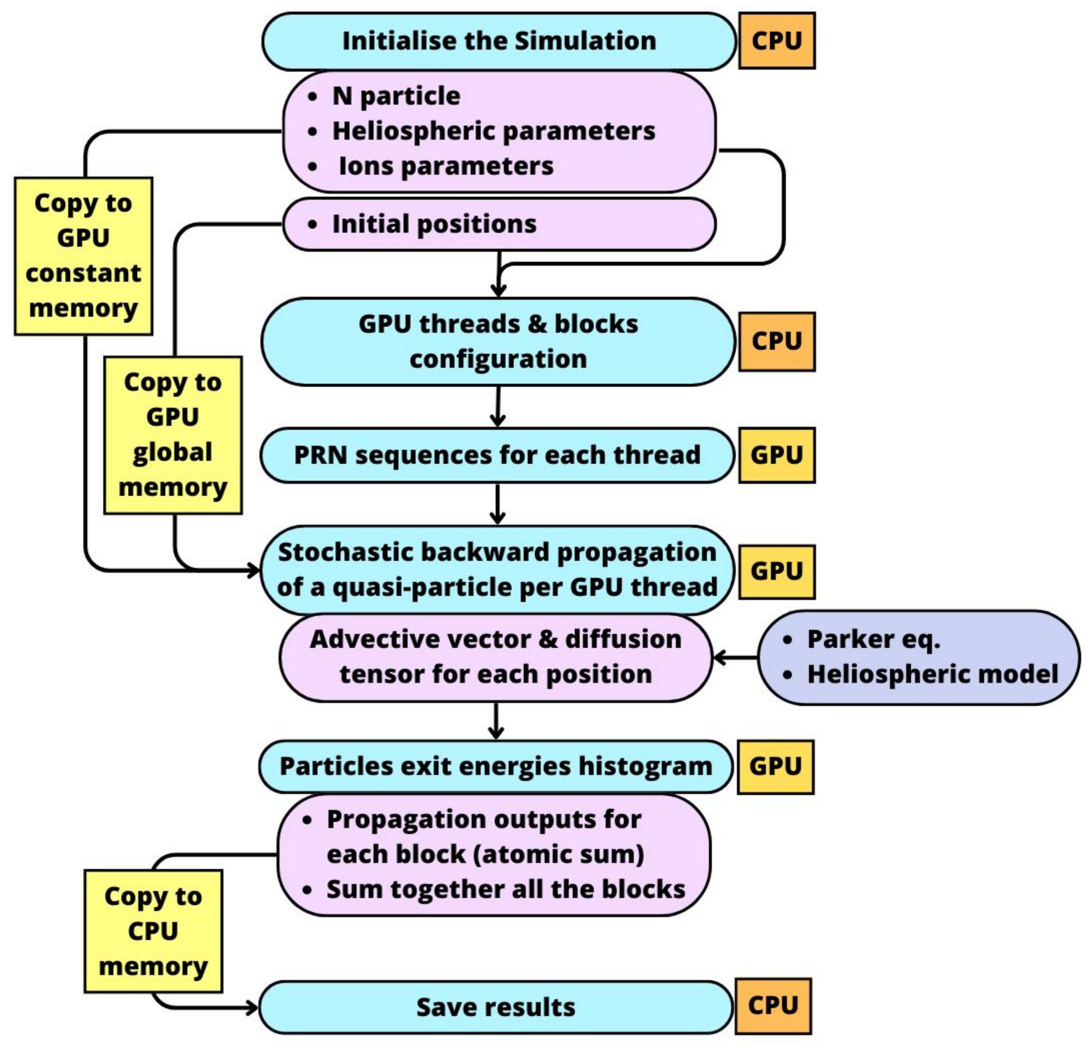
Solar modulation varies according to the intensity level of solar activity, the intensity and polarity of the solar magnetic field, and are energy- and charge-sign-dependent.

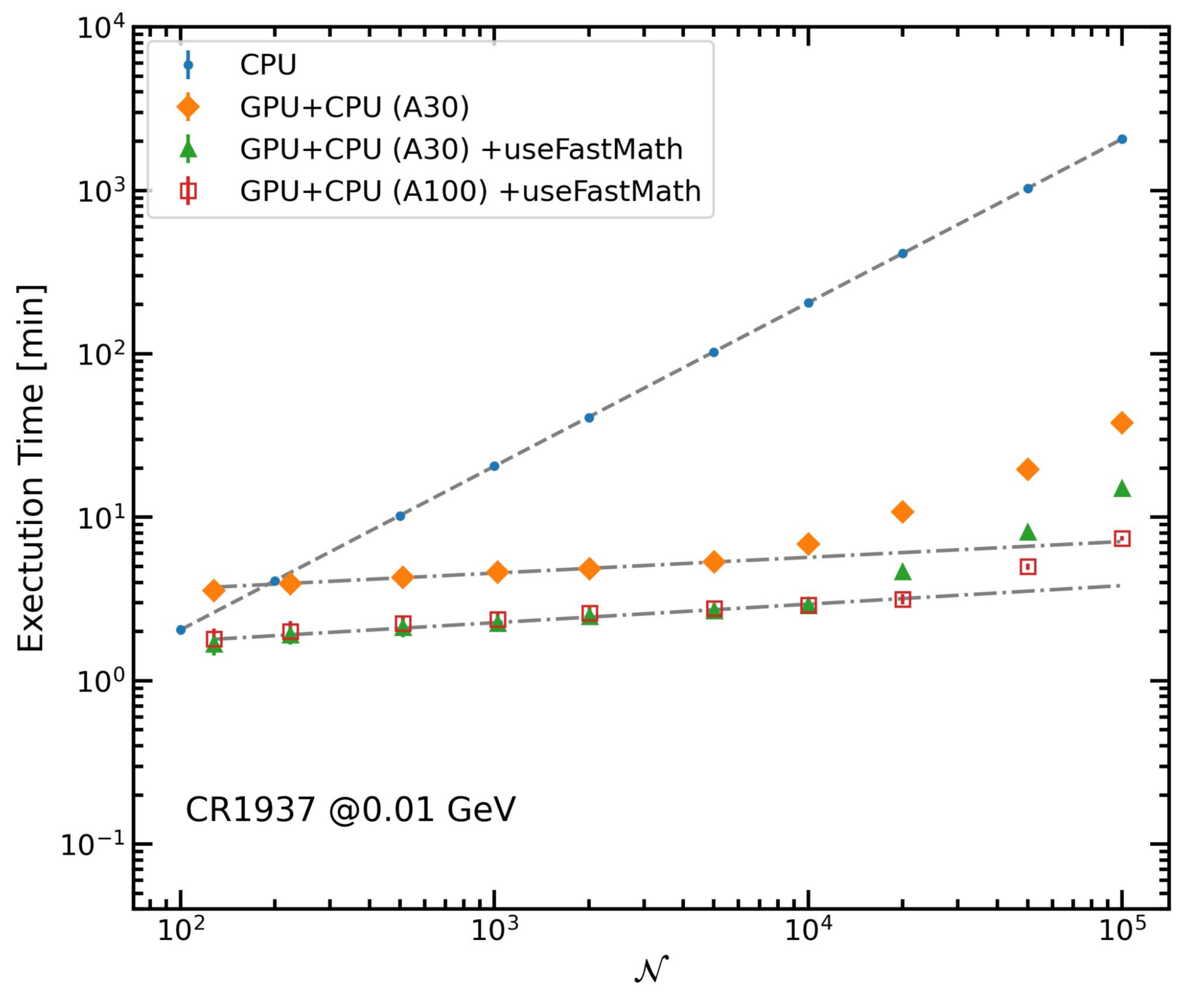
The numerical integration

Parker Equation can be solved numerically by means of Monte Carlo integration of the equivalent system of stochastic differential equations.

The numerical integration evolves a number *N* of *quasi-particle objects* from the observation zone to the heliosphere boundary. These objects represent the statistical properties of an ensemble of particles, thus, a) they have meaning only when treated statistically and b) they should be independent from the others.

We ported the original algorithm for CPU-only to CUDA-C language to optimize the architecture of NVIDIA-Ampere hardware. In this approach, the evolution of quasi-particle objects is assigned to different GPU threads.
In the case of multiple GPUs, the algorithm assigns to each GPU a subset of energies to be simulated: in this case, it is not necessary to share memory across devices, each CPU-thread can proceed independently from the others, and only the final histograms are transferred to the host without the need for further merging.





The performance

To test the computational performance, the CPU-only code was executed on a server with CPU Intel(R) Xeon(R) 2.10GHz. The CPU+GPU code was executed on servers with NVIDIA A30 and NVIDIA A100 GPUs boards.
We also tested the use of the compilation flag `--use_fast_math` which resulted, for the studied cases, in a negligible loss of numerical accuracy but led to a huge improvement in lowering the runtime.

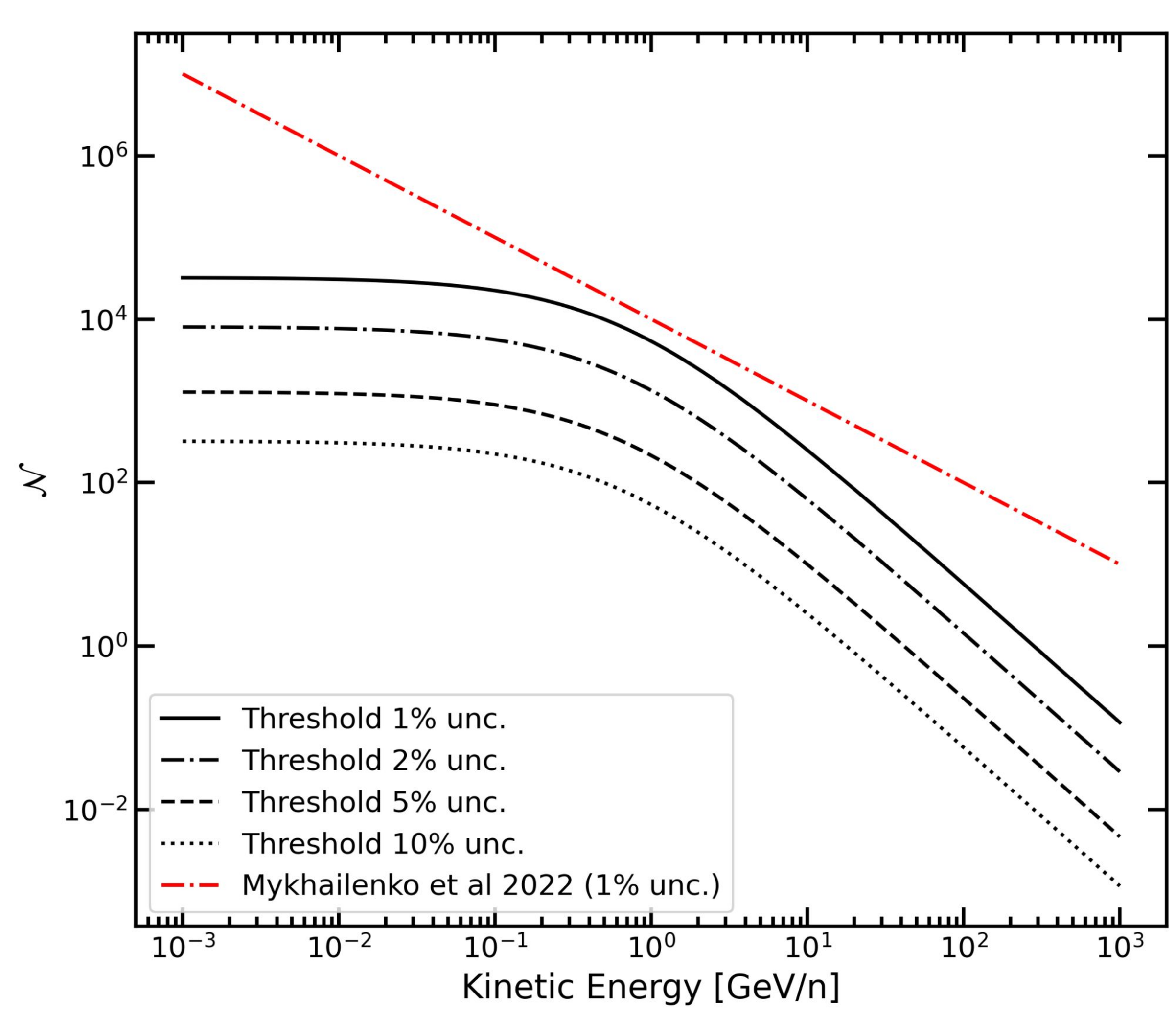
Due to the linearity of the algorithm, the execution time of CPU-only code scales as a power law of N with spectral index 1, i.e. a linear function in logarithmic scale. On the other hand, the execution time of CPU+GPU-code shows two different regimes: up to N ~ 10^4 it scales as a power law with spectral index ~ 0.1, then the spectral index becomes steeper.
The threshold of changing regime is different for A30 and A100. With A30 the spectral index changes for N> 10^4, while using A100 boards it appens when N> 2 x 10^4.

The numerical uncertainties

We studied the Monte Carlo uncertainties for the 2D-SDE propagation model due to the number of quasi-particle objects simulated (N).
We considered different Carrington Rotations where, for each case, we computed the numerical results up to 200 times. For each energy bin, we computed the mean simulated fluxes and the corresponding standard deviation, which represents the numerical uncertainties.
We found a relationship between the generated statistics and the numerical uncertainties that varies in the offset only among different Carrington Rotations. This led to the definition of an empirical *Safety Threshold* which allows one to estimate a *safe* number of quasi-particle objects (*N*) to be simulated for kinetic energy *T* in order to have a numerical uncertainty lower than a certain level:

$$\mathcal{N}_{\mathrm{SafetyThreshold}} = 2 \cdot \sigma^2_{r,sim} \cdot \left(\frac{T^c + d}{a \cdot T^b}\right)^2$$

where a = 1.29 ± 0.07, b = −0.85 ± 0.01, c = −0.85 ± 0.01, and d = 1.5 ± 0.1.



Conclusions

We evaluated the improved speed-up by using GPUs with respect to CPUs for several injected statistics. From those tests, it was evident that for a large number of injected particles (i.e. less numerical uncertainties) CPU+GPU code is significantly faster than using CPU-only. This performance, together with the availability of high-performance GPUs at affordable cost and the possibility to install several GPUs on a relatively small cluster, allows the application of parameters scanning algorithms to improve the knowledge of model parameters space.

Contacts: stefano.dellatorre@mib.infn.it

Acknowledgements:
Activity supported within the ASIF implementation agreement 2017-15-HD.0 ASI-INFN and 2021-36-HH.0 involving ASI and Milano-Bicocca University. Software development is supported by FondazioneICSC , Spoke 3 Astrophysics and Cosmos Observations. National Recovery and Resilience Plan (Piano Nazionale di Ripresa e Resilienza, PNRR) Project ID CN_00000013