'print(‘Hello Future’): Developing Next Generation Astronomical Codes (HelloFuture)' Abstracts

Talks

Introducing L-Galaxies AD: Differentiable Semi-Analytic Galaxy Modelling for Scalable Inference

Authors

Andrew Green (1)

(1) University of Hertfordshire

Abstract

Computational advances are transforming our study of the Universe. Simulations now capture increasingly large and complex systems, while next-generation surveys produce data at unprecedented scale. These trends demand astronomy codes that can efficiently exploit parallel, data- and memory-intensive computing environments. At the same time, scalable inference in high-dimensional parameter spaces increasingly requires differentiable models that enable gradient-based methods.

Semi-analytic models (SAMs), such as L-Galaxies, provide a fast and flexible framework for modelling galaxy populations across cosmological volumes by evolving baryonic physics on pre-computed dark matter halo merger trees. Their efficiency makes them well-suited for inference tasks that require repeated model evaluations. However, as physical complexity and parameter dimensionality grow, traditional approaches such as Markov Chain Monte Carlo (MCMC) become increasingly costly and difficult to scale.

We present L-Galaxies AD, a new parallel C++ implementation of L-Galaxies (Guo et al. 2011; Henriques et al. 2020; Yates et al. 2024) designed for modern HPC systems and made fully differentiable. Unlike previous SAMs, L-Galaxies AD uniquely combines differentiable programming for both forward- and reverse-mode automatic differentiation (AD) with task- and process-level parallelism. Casting the SAM as a computational graph allows efficient gradient evaluation and scalable performance across distributed-memory architectures, providing a powerful foundation for modern inference workflows.

Our differentiable, parallel SAM will be used to generate extensive galaxy catalogues efficiently, and to directly constrain model parameters using Bayesian inference against super-high-resolution hydrodynamical simulations — demonstrating a scalable gradient-aware approach to simulation-based inference for galaxy evolution.

L-Galaxies AD demonstrates that embedding differentiability and parallelism within astrophysical models unlocks efficient, scalable inference for galaxy formation studies.

A Scalable Frozen-Core Approach to Topological Filament Finding in Large Cosmological Simulations

Authors

Ankit Singh (1), Frazer Pearce (1), Meghan Gray (1), Gustavo Yepes (2)

(1) University of Nottingham
(2) Departamento de Física Teórica, Universidad Autónoma de Madrid

Abstract

Cosmic filaments are a key component of the cosmic web, funnelling matter into clusters and shaping the environments in which galaxies form and evolve. Accurate reconstruction of this network over gigaparsec scales is increasingly important for cosmology, structure formation, and galaxy evolution studies. However, widely used topological filament-finding methods such as DisPerSE face a major computational limitation: they require the full Delaunay tessellation of the input point set to reside in memory, making them impractical for modern large-volume simulations. Naively subdividing the volume into independent regions is not a viable solution, as it produces inconsistent tessellations and discontinuous filament networks.

We present a frozen-core method that overcomes this bottleneck while preserving the global topology of the filamentary network. The simulation volume is decomposed into overlapping blocks, within which local tessellations are constructed and filtered using a circumsphere criterion that retains only tetrahedra consistent with the global tessellation. The resulting block-wise outputs are combined through a post-processing pipeline that performs core selection, duplicate removal, and boundary stitching, yielding a coherent, large-scale filament network.

We validate the method against a reference single-volume reconstruction on a 300 h^-1 Mpc subvolume of the MultiDark Planck 2 (MDPL2) simulation, achieving 99.6 per cent recovery of total filament length and 94.7 per cent filament matching. We then apply the method to the full (1 h^-1 Gpc)^3 MDPL2 simulation, containing 92 million haloes, and produce a filament catalogue on gigaparsec scales.

As a demonstration of its scientific potential, we measure the connectivity (\kappa), defined as the number of filaments attached to a halo, for 1200 haloes spanning masses M_{200c} = 10^{12}-10^{15.5} h^-1 Msun. We find a single power-law relation between halo mass and connectivity that extends smoothly from group to cluster scales, providing the first confirmation that the theoretically predicted scaling holds across three orders of magnitude in halo mass.

GPU-accelerating CMB cosmology with state of the art ODE solvers

Authors

Lawrence Berry (1), Will Handley (1), Oliver Hahn (2), Nils Schöneberg (3)

(1) University of Cambridge
(2) University of Vienna
(3) LMU Munich

Abstract

GPUs offer a way to massively parallelise certain types of physical models, allowing us to compare more models for the same compute budget (and hence do more science) whilst avoiding the use of emulators (which must be retrained every time you modify the underlying physics). More efficient forward models also allow us to tackle higher-dimensional problems (field-level inference) and more complex foreground systematics and nuisance parameters.

In particular, the so-called Einstein-Boltzmann equations that underly models of the CMB are linear ordinary differential equations that are highly amenable to parallelisation on a GPU, across both "k-modes" and parameters. Meanwhile, modern cosmological analyses are burdened with the technical debt of low-level codes written in FORTRAN and C, which are difficult to modify when trying to test new models of physics.

We have therefore written a new CMB code in Python and JAX (DISCO-EB) which can take advantage of the massive computational throughput of modern GPUs, but is easy to modify with new physics. As part of this work, we have developed new ODE solvers that are better optimised for solving many small ODEs at once and are 100x faster than the current Python/JAX state-of-the-art (diffrax). The result is precision similar to that of existing codes but at vastly reduced $ inference cost. The ODE solvers themselves are packaged separately and could have broad application in the scientific community.

This deep dive into the world of GPU programming has taught us that simply "Jaxifying" one's code is not always the best solution, and that it pays to understand how a GPU works under the hood. Moreover, as GPU manufacturers begin to optimise their chips for a deep learning community that prefers lower precision matrix multiplication operations, the scientific community must adapt its codes to make use of GPUs that have fewer 64 bit precision cores but more "tensor cores". We present a number of examples of future-proofing to this effect.

Differentiable and Emulated Radiative Transfer for On-the-Fly Astrophysical Simulations

Authors

Lorenzo Branca (1)

(1) IWR-Heidelberg University

Abstract

Radiative transfer remains one of the main computational bottlenecks in astrophysical simulations because it is nonlocal, time dependent, and tightly coupled to gas dynamics. In this contribution, we present a two-level strategy for accelerating on-the-fly radiative transfer in turbulent astrophysical environments. First, we introduce Ray-trax, a GPU-oriented, fully differentiable 3D ray-tracing framework written in JAX, designed for the time-dependent emission–absorption problem on turbulent hydrodynamic fields. Ray-trax is vectorized across rays and sources, supports arbitrary numbers of frequency bins without changes to the core architecture, and exposes end-to-end gradients, enabling inverse problems and future coupling to differentiable hydrodynamics. The method is validated against analytical solutions and demonstrated in turbulent media.

Second, building on this radiative-transfer setting, we present a neural surrogate approach for emulating 3D monochromatic radiative transfer in both static and time-dependent regimes. The emulator uses a Fourier Neural Operator combined with U-Net components and achieves more than two orders of magnitude speedup while maintaining average relative errors below 3%; in the repository results, the steady-state model reaches about 6,750× speedup with 2.6% error and the recurrent time-dependent model about 600× with 2.9% error.

Together, these two approaches define a practical pathway toward fast, differentiable, and machine-learning-accelerated radiative transfer for next-generation hydrodynamical simulations: Ray-trax provides a physically grounded and differentiable solver, while the neural emulator offers a route to further reduce computational cost in production settings.

Attention-Based Preprocessing Framework for Improving Rare Transient Classification

Authors

Xinyue Sheng (1), Tuan Dung Pham (2), Zichi Zhang (2)

(1) Astrophysics Research Centre, Queen's University Belfast
(2) School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast

Abstract

Data preprocessing plays a critical role in determining the final performance of machine-learning models, particularly when training data are extremely imbalanced. With the rapidly growing number of transients discovered by current and upcoming wide-field imaging surveys, machine learning has become an essential tool for prioritizing events for follow-up using light-curve and host-galaxy information. However, the identification of rare transient classes remains challenging due to severe class imbalance, and feature extraction from host images is often compromised by bright foreground sources, especially when the true host galaxy is faint or distant.

I will present a data-augmentation pipeline for both images and light curves designed to address these challenges, and demonstrate its application to improving the classification of Superluminous Supernovae Type I (SLSNe-I) and Tidal Disruption Events (TDEs) using our existing NEEDLE framework. For imaging data, the method employs a Similarity Index to reject artefacts and a masking procedure that removes unrelated sources while preserving the transient and its host galaxy. This focuses the classifier on physically relevant pixels and enables arbitrary rotations for effective class upsampling.

For time-domain data, observed multi-band light curves are modeled using a two-dimensional Gaussian Process. Synthetic, data-driven light curves are then generated via resampling and redshifting, and cross-matched with host-galaxy images from the same class to create unique yet realistic training examples. A customized focal loss function is further introduced to account for the remaining class imbalance.

Models trained on the augmented dataset show a substantial improvement in performance. For classifications with confidence ≥ 0.8, we achieve purities of 75% (43%) and completenesses of 75% (66%) for SLSNe-I (TDEs), demonstrating the effectiveness of the proposed approach for rare-transient identification.

AstroPhot: fitting everything everywhere all at once in astronomical images (Online)

Authors

Connor Stone (1), Renée Hložek (1)

(1) University of Toronto

Abstract

Astrophot is a mature tool for parametric image modelling of astronomical images. One has the flexibility to construct arbitrarily complex models of arbitrarily many objects in a scene over arbitrarily many images at different epochs/wavelengths. This is done through an intuitive object oriented interface, and assisted by a large suite of detailed Jupyter notebook tutorials and documentation describing every model. It operates in PyTorch or JAX allowing for accelerated CPU or GPU processing; further, with a caskade backend it is able to trivially interface with other codes such as caustics to add gravitational lens modelling capabilities. In this talk I will detail some of the applications for which AstroPhot is already in use, as well as ongoing transient scene modelling work for the LSST.