Post-doctoral project

Bayesian inference for cosmology: Inferring the initial fields of our cosmic neighborhood

Sep. 2026 – Dec. 2027

Jenny SORCE, Pierre-Antoine THOUVENIN,

Centre de Recherche en Informatique, Signal et Automatique de Lille

CRIStAL UMR 9189 - Lille, France

Dates: 18 months; starting date around September 2026, which can be adjusted depending on the availability of the successful candidate.

Contacts: Jenny SORCE (jenny.sorce@univ-lille.fr),
https://sorcej.github.io/Jenny.G.Sorce/
Pierre-Antoine THOUVENIN (pierre-antoine.thouvenin@centralelille.fr),
https://pthouvenin.github.io/

Laboratory: Centre de Recherche en Informatique, Signal et Automatique de Lille (UMR 9189 CRIStAL), Villeneuve d'Ascq, France.

Keywords: Inverse problem, cosmological simulation model, physics-informed training, Bayesian inference, MCMC algorithms.

---

1 Project overview

According to the standard cosmological model, about 95% of the Universe is dark. Recent large survey analyses reveal tensions with this model. For instance, the local measurement of the expansion rate and the estimate of the Universe homogeneity differ by more than three standard deviations from those inferred with the first light of the Universe. These discrepancies are at the heart of a heated debate in cosmology to determine whether these tensions require new physical models to be acccounted for, or are mere consequences of systematic biases in the observation processing pipeline. Part of this pipeline relies on cosmological simulations to act as the missing ground truth. However, the simulations only reproduce the statistics of the local cosmic web. A new type of simulations, qualified as constrained, is emerging. Initial velocity and density fields of such simulations stem from observational constraints.

This post-doctoral project is aimed at inferring the initial velocity and density fields of the local cosmic web from today's luminosity distances and observational redshifts measurements. This high-dimensional astrophysical inverse problem is challenging. In particular, it will leverage a large number of measurements (Prideaux-Ghee et al., 2023; Bayer et al., 2023). The absence of ground truth data calls for reliable estimators with associated uncertainty quantification. This motivates the use of Markov chain Monte-Carlo (MCMC) algorithms to access posterior distributions. The hierarchical model relies on a costly cosmological simulator to describe the evolution of cosmological objects from the initial conditions.

A first step has already been carried out by replacing the costly black-box simulator in the inference algorithm by a tractable surrogate model trained on a grid of simulations (Prost et al., 2025), in the spirit of Raissi et al. (2019); Jindal et al. (2023); Dai et al. (2023). Building on Prost et al. (2025), this postdoctoral project will focus on the design of a high-dimensional MCMC algorithm to infer the parameters of interest inspired by Durmus et al. (2018); Vono et al. (2021); Coeurdoux et al. (2024a,b).

Keywords. Inverse problem, cosmological simulation model, Bayesian inference, MCMC algorithms.

2 Scientific context

The project is part of the *Chaire WILL UNIVERSITWINS* (*UNIVERSe dIgital TWINS*) led by Jenny Sorce (funded by the Université de Lille under the initiative of excellence). The successful candidate will be jointly supervised by Jenny Sorce (CNRS Researcher in cosmology) and Pierre-Antoine Thouvenin (Assoc. Prof., Centrale Lille), and hosted in the CRIStAL lab (UMR 9189), Lille, France. The work will be conducted in collaboration with Jean Prost (Assoc. Prof., ENSEEIHT) in the IRIT lab.

More than 2000 GPU.hours have already been secured for the project at TGCC on the Irene/Rome partition. They will be used to finetune, validate and deploy the surrogate model to perform Bayesian inference. Access to the medium scale computing center from the University of Lille is also ensured.

3 Profile and requirements

PhD in signal/image processing, computer science or applied mathematics. The project requires a strong background in data science and/or machine learning (statistics, optimization), signal & image processing. Very good Python coding skills are expected. A B2 English level is mandatory.

Knowledge in C++ programming, as well as experience or interest in parallel/distributed code development (MPI, OpenMP, CUDA, ...) will be appreciated.

4 Application procedure

Applicants are invited to email the following documents as a single `.pdf` file to the co-advisors:

- a detailed curriculum, including a list of publications;
- link to the PhD manuscript (or PhD project if upcoming defense);
- reports from PhD reviewers if available;
- a cover letter;
- references: recommendation letters or names of 2 researchers recommending your application.

For further information, please contact the co-advisors of the project:

- Jenny Sorce, jenny.sorce@univ-lille.fr
- Pierre-Antoine Thouvenin, pierre-antoine.thouvenin@centralelille.fr.

References

James Prideaux-Ghee, Florent Leclercq, Guilhem Lavaux, Alan Heavens, and Jens Jasche. Field-based physical inference from peculiar velocity tracers. *Mon. Not. R. Astron. Soc.*, 518(3):4191–4213, Jan. 2023. doi: 10.1093/mnras/stac3346.

Adrian E. Bayer, Chirag Modi, and Simone Ferraro. Joint velocity and density reconstruction of the universe with nonlinear differentiable forward modeling. *Journal of Cosmology and Astroparticle Physics*, 2023 (06):046, jun 2023. doi: 10.1088/1475-7516/2023/06/046. URL https://dx.doi.org/10.1088/1475-7516/2023/06/046.

Jean Prost, Pierre-Antoine Thouvenin, Jenny Sorce, and Pierre Chanais. Un réseau de neurones augmenté pour simuler l'évolution de la distribution de matière noire. *Gretsi 2025*, 2025.

M. Raissi, P. Perdikaris, and G. E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, Feb. 2019. ISSN 0021-9991. doi: 10.1016/j.jcp.2018.10.045.

Vaibhav Jindal, Drew Jamieson, Albert Liang, Aarti Singh, and Shirley Ho. Predicting the initial conditions of the universe using deep learning, 2023. URL http://arxiv.org/abs/2303.13056.

Zhenyu Dai, Ben Moews, Ricardo Vilalta, and Romeel Dave. Physics-informed neural networks in the recreation of hydrodynamic simulations from dark matter, 2023. URL http://arxiv.org/abs/2303.14090.

Alain Durmus, Éric Moulines, and Marcelo Pereyra. Efficient Bayesian Computation by Proximal Markov Chain Monte Carlo: When Langevin Meets Moreau. *SIAM J. Imaging Sci.*, 11(1):473–506, 2018. ISSN 1936-4954. doi: 10.1137/16M1108340. URL https://epubs.siam.org/doi/10.1137/16M1108340.

Maxime Vono, Nicolas Dobigeon, and Pierre Chainais. Asymptotically exact data augmentation: Models, properties, and algorithms. *J. Comput. and Graph. Stat.*, 30(2):335–348, 2021. ISSN 1061-8600. doi: 10.1080/10618600.2020.1826954.

Florentin Coeurdoux, Nicolas Dobigeon, and Pierre Chainais. Plug-and-Play split Gibbs sampler: Embedding deep generative priors in Bayesian inference. *IEEE Trans. Image Process.*, 33:3496–3507, 2024a. ISSN 1941-0042. doi: 10.1109/TIP.2024.3404338. URL https://ieeexplore.ieee.org/document/10541919?source=authoralert.

Florentin Coeurdoux, Nicolas Dobigeon, and Pierre Chainais. Normalizing flow sampling with Langevin dynamics in the latent space. *Mach. Learn.*, 2024b. ISSN 1573-0565. doi: 10.1007/s10994-024-06623-x. URL https://doi.org/10.1007/s10994-024-06623-x.