

Sujet de stage M2

Proposants, laboratoire

Cuby Jean-Gabriel, jean-gabriel.cuby@lam.fr

Le stage sera co-encadré par une personne du Centre de données Astrophysique de Marseille - CeSAM Laboratoire d'Astrophysique de Marseille. CNRS, CNES & Université Aix-Marseille (AMU)

Descriptif du sujet

Sujet. Modélisation de galaxies et de quasars très distants et outils de sélection par analyse Bayésienne et apprentissage machine pour la mission Euclid de l'agence spatiale européenne.

Contexte. Les galaxies et quasars à très grand décalage spectral (*redshift*, $z > 7$) nous éclairent sur les processus de formation et d'évolution des galaxies dans l'Univers primordial. L'observation de l'Univers à très grand décalage spectral est au centre de beaucoup des grands projets de l'astrophysique des décennies à venir (Euclid, JWST, WFIRST, SKA, ATHENA, LISA).

Méthodologie et Programme de travail. La sélection des objets à très grand décalage spectral se fait sur la base de leurs couleurs très (infra-)rouges. Une des difficultés provient de la similitude de ces couleurs avec celles d'étoiles très froides de l'environnement local. Des outils statistiques doivent alors être développés afin d'affiner la sélection grâce à une modélisation précise des populations d'objets très lointains et des contaminants (voir par exemple Pipien et al. (2018) pour une description d'un modèle Bayésien). La précision et la limite de validité de ces modèles risquent toutefois d'être challengées par la grande sensibilité d'Euclid. Les méthodes de *machine learning* et/ou de *deep learning* voient leur apparition en astrophysique (voir par exemple arXiv:1806.06607) et permettent d'accéder à des domaines de l'espace des paramètres où les modèles sont peu contraints.

Le travail consistera à poser les bases d'un développement d'outils de sélection des objets à grand décalage spectral pour la mission Euclid :

- i. Développement d'outils de modélisation d'objets à grand décalage spectral et de contaminants. Il s'agira ici de développer quelques outils simples permettant de générer des populations d'objets (catalogues) à partir de leurs spectres (connus) et de leurs fonctions de distribution spatiale, morphologique, et en luminosité.
- ii. Développement d'un modèle Bayésien adapté aux données d'Euclid. Il s'agira ici d'adapter un modèle Bayésien existant et de le tester à partir des populations d'objets qui auront été générées auparavant (point i.).
- iii. Test de méthodes d'apprentissage machine: à partir des catalogues simulés au point i. et de la modélisation effectuée au point ii., il s'agira ici d'effectuer des tests préliminaires avec des outils d'apprentissage machine. Idéalement, ces tests permettront une première comparaison avec les résultats de l'analyse Bayésienne.

Ce travail sera effectué sous une double responsabilité 'astrophysique' et 'logicielle' au sein du Laboratoire d'Astrophysique de Marseille. Ce travail pourra se poursuivre en thèse (un co-financement du CNES est d'ores et déjà acquis).

Bibliographie

- [An 800-million-solar-mass black hole in a significantly neutral Universe at a redshift of 7.5](#), Bañados et al., 2018, *Nature*, Volume 553, Issue 7689
- [High-redshift quasar selection from the CFHQSIR survey](#), Pipien et al., 2018, *A&A*, 617, id.A127
- [On the nature of the luminous Ly \$\alpha\$ emitter CR7 and its restframe UV components](#), Sobral et al., arXiv:1710.08422
- [A Remarkably Luminous Galaxy at \$z=11.1\$ Measured with Hubble Space Telescope Grism Spectroscopy](#), Oesch et al., *The Astrophysical Journal*, Volume 819, Issue 2, article id. 129, 11 pp. (2016)
- [Probabilistic selection of high-redshift quasars](#), Mortlock et al., *Monthly Notices of the Royal Astronomical Society*, Volume 419, Issue 1, pp. 390-410 (2012)
- [A luminous quasar at a redshift of \$z = 7.085\$](#) , Mortlock et al., *Nature*, Volume 474, Issue 7353, pp. 616-619 (2011)
- [Euclid Definition Study Report \(Red book\)](#). arXiv:1110.3193 (pages 29-30)