



Proposition de sujet de thèse

Titre de la thèse: Euclid weak lensing: PSF and galaxy shape measurements and the impact of errors on cosmological parameters

Laboratoire : IRFU/SAP

Nom du responsable du thèse: Jean-Luc Starck (CEA France)

Nom du co-responsable du thèse: Martin Kilbinger (CEA France)

Coordonnées : jstarck@cea.fr

☎ : 01 69 08 57 64

Thesis Abstract :

The Euclid satellite, to be launched in 2020, will observe the sky in the optical and infrared, and will be able to map large scale structures and weak lensing distortions out to high redshifts. Weak gravitational lensing is thought to be one of the most promising tools of cosmology to constrain models. Weak lensing probes the evolution of dark-matter structures and can help distinguish between dark energy and models of modified gravity. Thanks to the shear measurements, we will be able to reconstruct a dark matter mass map of 15000 square degrees. These shear measurements are derived from the galaxy shapes, which are blurred by the PSF (point-spread function) of the optical imaging system. One of the main problems to achieve the scientific goals is therefore the need to model the point spread function (PSF) of the instrument with a very high accuracy. The PSF field can be estimated from the stars contained in the acquired images. It has to take into account the spatial and spectral variation of the PSF. An additional problem to take care of is the subsampling of the images. Once the PSF is correctly modelled, we need to derive the shear from galaxy shapes.

The goal of the PhD consists first in finding a robust method to build the 3D PSF field (two spatial and one spectral dimensions) using interpolations tools on the adequate manifold. As this manifold is unknown, it has to be learned using recent learning techniques based on sparse data representation. The PSF field can then be used as ingredient for galaxy shape measurements. Both PSF estimation and shape measurement have some limitations, and are affected by statistical and systematic errors. An important step will be the study of the propagation of these errors from measurements to cosmological parameters.

Résumé de la thèse:

Le projet spatial Euclid, dont le lancement est prévu en 2020, observera le ciel en optique et en infrarouge et permettra de construire des cartes de très grandes échelles afin mesurer les distorsions gravitationnelles jusqu'à des redshifts très élevés. Grâce à ces mesures de cisaillement gravitationnel, nous pourrons reconstruire des cartes de matières noires de 15000 degrés carrés, soit presque la moitié du ciel. Ces mesures de shear sont dérivées de l'analyse des formes de galaxies, qui sont floutées dues l'optique du télescope.

L'un des problèmes majeurs pour atteindre les objectifs scientifiques est donc la nécessité de modéliser la fonction d'étalement du point (Point Spread Function (PSF) en anglais) du satellite, et de mesurer la forme des galaxies avec une très grande précision et corrigée de la PSF.



Le champ de PSF peut être calculé à partir des étoiles contenues dans les images observées. Il doit prendre en compte la variation spatiale et spectrale de la réponse de l'instrument. Une difficulté supplémentaire vient du problème de sous échantillonnage des images.

Le but de cette thèse est de trouver une méthode robuste permettant de construire le champ de PSF tridimensionnel (deux dimensions spatiales et une dimension spectrale) en utilisant des outils d'interpolation sur la variété associée aux PSFs. Cette variété n'étant pas connue, il faudra l'apprendre via des outils modernes d'apprentissage basés sur la parcimonie de l'information.

Ce champ de PSF pourra ensuite être utilisé dans la mesure des formes des galaxies. L'impact des erreurs de mesures (PSF et mesure de forme) sur l'estimation des paramètres cosmologiques sera étudié.

Résumé étendu de la thèse:

Le contexte

Le projet spatial Euclid, qui observera le ciel en optique et en infrarouge, permettra de construire des cartes de très grandes échelles, et de mesurer les distorsions gravitationnelles jusqu'à des redshifts très élevés. Grâce à ces mesures de cisaillement gravitationnel, nous pourrons reconstruire des cartes de matières noires de 15000 degrés carrés, soit la moitié du ciel.

Le CEA a une très forte implication dans le projet EUCLID. La France a obtenu la responsabilité de *Organization Unit Level 3*, et le CEA a un rôle important dans les groupes de travail correspondants. Nous sommes donc fortement impliquées dans le développement d'algorithmes.

L'un des problèmes majeurs pour atteindre les objectifs scientifiques est la nécessité de modéliser la fonction d'étalement du point (Point Spread Function ou PSF en anglais) du satellite (PSF) avec une très grande précision. Ce champ de PSF peut être calculé à partir des étoiles contenues dans les images observées. Il doit prendre en compte la variation spatiale et spectrale de la réponse de l'instrument. Une difficulté supplémentaire vient du problème de sous échantillonnage des images.

Le but de cette thèse est de trouver une méthode robuste permettant de construire le champ de PSF tridimensionnel (deux dimensions spatiales et une dimension spectrale) en utilisant des outils d'interpolation sur la variété associée aux PSFs. Cette variété n'étant pas connue, il faudra l'apprendre via des outils modernes d'apprentissage basés sur la parcimonie de l'information.

Ce champ de PSF pourra ensuite être utilisé dans la mesure des formes des galaxies. L'impact des erreurs de mesures (PSF et mesure de forme) sur l'estimation des paramètres cosmologiques sera étudié.

L'Apprentissage

Nous n'avons de l'information sur les PSFs qu'à la position des étoiles. La construction du champ de PSFs nécessite donc une interpolation de cette information, spatialement **et** spectralement. En général cette interpolation est faite à partir d'une simple moyenne pondérée des étoiles suivant leur distance Euclidienne au point d'interpolation, soit directement dans l'espace des pixels (comme dans le package PSFext) soit dans l'espace PCA (*principal component analysis*). Aucune de ces deux solutions n'est adaptée, car aucun de ces deux espaces ne sont adéquat pour représenter les PSFs. L'invariance par rotation ou par dilation de la représentation est en pratique nécessaire, et d'une manière plus générale les PSFs vivent sur une variété particulière, mais qui n'est pas connue. De nouvelles approches doivent être développées, qui ne conduisent pas forcément à des décomposition linéaires, comme par exemple le *Multiscale Dictionary Learning* [1] ou encore le dictionary Learning avec invariance [2].

La superresolution

Les données sont sous échantillonnées et un problème de super résolution se rajoute au problème d'interpolation. Des travaux récents [3] ont montré comment une contrainte de parcimonie peut améliorer de manière significative la reconstruction d'une PSF Euclid super résolue. Cette technique de super résolution devra donc s'intégrer dans le mécanisme d'interpolation.



La synergie instrument-algorithme

Le CEA (équipe de Jerome Amiaux) a développé un banc de test optique permettant d'acquérir des PSFs Euclid réalistes. Des données existent déjà, et d'autres pourront être obtenues en fonction des besoins. Les résultats devront être également validés sur ces simulations en collaboration avec l'équipe instrumentation.

La Propagation des erreurs

La propagation des erreurs systématiques et statistiques sur les paramètres cosmologiques est un problème loin d'être résolu. Des méthodes, développées dans le cadre de la collaboration Euclid, intègre des estimations d'erreur. Par exemple, les méthodes Bayésiennes explore la distribution a posteriori des paramètres de forme et des méthodes de maximum de vraisemblance donnent des estimations d'erreur correspondant au modèle. Les futurs relevés grand-champs comme Euclid seront dominés par les erreurs systématiques, et non statistiques, il est donc important de pouvoir quantifier et exploiter cette information lors de l'estimation des paramètres cosmologiques.

References:

- 1) M. Maggioni, S. Minsker and N. Strawn, Multiscale Dictionary Learning: Non-Asymptotic Bounds and Robustness. <http://arxiv.org/abs/1401.5833>, 2014.
- 2) S. Beckouche, J.-L. Starck, G. Peyre and J. Fadili, "Astronomical Image Denoising using Dictionary Learning", **Astronomy and Astrophysics**, 556, id.A132, pp 14, 2013. [arXiv:1401.5833](http://arxiv.org/abs/1401.5833).
- 3) F.M. Ngolè Mboula *, J.-L. Starck, S. Ronayette, K. Okumura, J. Amiaux, Super-resolution method using sparse regularization for point spread function recovery, **A&A**, 575, id.A86, 2015.
- 4) M. Kilbinger 2014, **RoPP**, 78, 8, id. 086901, 2015. <http://www.arxiv.org/abs/1411.0115>.
- 5) R. Mandelbaum et al. 2014, **ApJS**, 368, 715, <http://www.arxiv.org/abs/1308.4982>

L'environnement scientifique:

La thèse se déroulera au sein du groupe de recherche pluridisciplinaire, CosmoStat, au service d'Astrophysique du CEA sous la direction de Jean-Luc Starck et de Martin Kilbinger. L'équipe est très impliquée dans le projet Euclid et le weak lensing, avec de nombreuses responsabilités (Jean-Luc Starck est leader de l'unité OULE3, en charge de définir les algorithmes utilisés pour dériver les produits finaux et également du Work Package OULE3-Weak Lensing. Martin Kilbinger est responsable OU-Shear France, l'unité en charge des mesures de formes et co-responsable du Weak Lensing Scientific Working Group, et Sandrine Pires est en charge de l'intégration des codes de weak lensing dans le Data Processing Center). Jean-Luc Starck a une grande visibilité internationale dans le domaine AstroStatistics, il est en particulier vice-président de l'IAA (International Astrostatistics Association) en charge de la section Astronomy/Astrophysics. Il a été lauréat ERC Advanced Grant, et a écrit trois livres. L'équipe a une très forte expertise dans différents domaines (mesure de PSFs, mesure de formes, problèmes inverses, weak lensing, etc.), et l'étudiant sera à l'interface entre les scientifiques et les ingénieurs qui travaillent sur l'optique de l'instrument. L'étudiant thésard disposera donc d'un environnement idéal pour mener à bien ses travaux. A l'issue de sa thèse, l'étudiant maîtrisera des outils mathématiques sophistiqués, le savoir faire pour estimer les paramètres cosmologiques à partir de données weak lensing, et aura une très bonne connaissance du projet Euclid. Il pourra alors soit continuer dans le domaine académique, soit exploiter ses compétences dans le domaine industrie.