



Inference de paramètres cosmologiques à partir de prédictions statistiques théoriques d'ordre élevé

Spécialité Astrophysique

Niveau d'étude Bac+5

Formation Master 2

Unité d'accueil [DAP/LCS](#)

Candidature avant le 16/06/2022

Durée 6 mois

Poursuite possible en thèse oui

Contact [Codis Sandrine](#)

+33 1 69 08 57 64

jean-luc.starck@cea.fr

Autre lien <http://jstarck.cosmostat.org>

Résumé

Sujet détaillé

Le satellite Euclid, qui sera lancé en 2023, observera le ciel dans les domaines optique et infrarouge, et mesurera les distorsions gravitationnelles jusqu'à des redshifts très élevés. L'effet de lentille gravitationnelle faible est considéré comme l'un des outils les plus prometteurs de la cosmologie pour contraindre les modèles. Les lentilles faibles sondent l'évolution des structures de la matière noire et peuvent aider à distinguer l'énergie noire des modèles de gravité modifiée. Grâce aux mesures de cisaillement, nous pourrions reconstruire une carte de masse de matière noire de 15 000 degrés carrés. La cartographie de masse implique la construction de cartes bidimensionnelles utilisant des mesures de forme de galaxie, représentant la densité de matière totale intégrée le long de la ligne de visée. Les cartes de masse sur des petits champs ont souvent été utilisées pour étudier la structure et la distribution en masse des amas de galaxies, alors que les cartes à grand champ ne sont possibles que depuis peu, en raison des stratégies d'observation de relevés de galaxies tels que CFHTLenS, HSC, DES et KiDS. Les cartes de masse contiennent des informations cosmologiques non gaussiennes significatives et peuvent être utilisées pour identifier des amas massifs ainsi que pour effectuer une corrélation croisée entre le signal de lentille et les structures d'avant plan.

Les paramètres cosmologiques sont traditionnellement estimés à l'aide d'une vraisemblance gaussienne basée sur des prédictions théoriques de statistiques de second ordre telles que le spectre de puissance ou les fonctions de corrélation à deux points. Cela nécessite de construire des matrices de covariance, et donc de nombreuses simulations à n corps très lourdes. Cette approche présente également plusieurs inconvénients supplémentaires : premièrement, les statistiques de second ordre capturent toutes les informations disponibles dans les données uniquement dans le cas des champs aléatoires gaussiens, tandis que la distribution de la matière est hautement non gaussienne et présente de nombreuses caractéristiques telles que des filaments, des

feuilletés ou des amas. Deuxièmement, la matrice de covariance est dépendante de la cosmologie et le bruit n'est généralement pas gaussien, ces deux aspects étant généralement mal pris en compte. Enfin, tous les effets systématiques tels que les masques, l'alignement intrinsèque, les effets baryoniques sont très difficiles à prendre en compte. Pour toutes ces raisons, une nouvelle approche a récemment émergé, appelée inférence de paramètres cosmologiques sans vraisemblance, basée sur une modélisation "forward". Il a le grand avantage de ne plus avoir besoin de matrices de covariance, évitant le stockage d'un énorme ensemble de données simulées (nous avons généralement besoin de 10 000 réalisations à n corps pour chaque ensemble de paramètres cosmologiques). De plus, cela nous ouvre la porte à l'utilisation d'informations statistiques d'ordre élevé et il est relativement simple d'inclure tous les effets systématiques. Il présente cependant deux inconvénients sérieux, le premier est le besoin d'énormes ressources GPU pour traiter des relevés tels qu'Euclid et le second est que la solution repose sur la précision des simulations, ce qui pourrait conduire à des discussions infinies au cas où les résultats seraient différents de ce qui est attendu. Grâce à une percée récente (Codis, 2021), nous disposons désormais d'outils théoriques pour prédire, pour un ensemble donné de paramètres cosmologiques, la fonction de probabilité de densité multi-échelle (pdf) de cartes de convergence comme celle qui sera observée avec Euclid.

L'objectif de ce travail est de développer une approche hybride, consistant en une inférence de paramètres cosmologiques sans vraisemblance qui serait basée sur la prédiction théorique statique d'ordre élevé plutôt que sur des simulations à n corps. Il aurait donc l'avantage des deux approches décrites précédemment, car il n'aura pas besoin de stocker un énorme ensemble de données pour calculer une matrice de covariance et il ne nécessitera pas d'énormes ressources CPU/GPU comme méthode de modélisation avancée. Cette frugalité intense rendra cette approche hautement compétitive pour contraindre le modèle cosmologique en utilisant des statistiques d'ordre élevé dans les futurs relevés.

Pour atteindre cet objectif, la première étape sera de construire un émulateur de carte, similaire à ce qui a été fait pour les statistiques à 2 points (c'est-à-dire la méthode flask), mais qui respecte avec précision les prédictions d'ordre élevé. En utilisant cet émulateur, il sera alors possible de l'utiliser comme contournement dans un code d'inférence récemment développé. Cela permettra l'utilisation de statistiques d'ordre élevé telles que la norme l_1 de la transformée en ondelettes de la carte convergence pour contraindre les paramètres cosmologiques, la norme l_1 étant ici une statistique extrêmement puissante (Ajani et al, 2021). La méthode sera d'abord appliquée sur le relevé CFIS, puis sur Euclid.

Mots clés

Simulation, TensorFlow, inference

Compétences

TensorFlow, deep learning

Logiciels

TensorFlow, python

Likelihood-free cosmological parameter inference using theoretical high- order statistics predictions

Summary

Full description

The Euclid satellite, to be launched in 2023, 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. Cosmological parameters are traditionally estimated using a Gaussian likelihood based on theoretical predictions of second order statistic such as the power spectrum. This requires to build a covariance matrices, and therefore need a lot of very heavy n-body simulations. To overcome these difficulties, a new approach has recently emerged, called likelihood- free inference which is based on a forward modelling. It has the great advantage to not need covariance matrices anymore, avoiding the storage of huge simulated data set. Furthermore, it paves the way to use high order statistical information and is relatively straightforward to include all systematics effect. It has however two serious drawbacks: the need of huge GPU ressources and very accurate simulations. Thanks to a recent breakthrough (Barthelemy+20), we have now theoretical tools to predict, for a given set of cosmological parameters, the multi-scale density probability function of convergence maps such as the one that will be observed with Euclid.

The goal of the stage work is to develop an hybrid approach, consisting in a likelihood- free cosmological parameter inference which would be based on the high order statics theoretical prediction rather than n-body simulations. It would therefore have the advantage of both previously described approaches, as it will not need to store huge data set to compute a covariance matrix and it will not require huge CPU/ GPU ressources as the forward modelling method. This intense frugality will make this approach highly competitive to constraint the cosmological model using high order statistics in future surveys. To achieve this goal, the first step for the M2 internship will be to build a map emulator, similar to what has been done for 2 point statistics, but which respects accurately the high order predictions. Using this emulator, it will then be possible to use it as a bypass in a recently developed Likelihood Free Inference code. The method will then be used first on the CFIS survey, and then on Euclid.

Keywords

Likelihood-free cosmological parameter inference, Simulation, TensorFlow

Skills

TensorFlow, deep learning

Softwares

TensorFlow, python