IRFU : Institut de recherche sur les lois fondamentales de l'univers

Saclay

DPhN/LSN

Analyse des mesures gravitationnelles de l'interféromètre LISA: Machine learning pour l'estimation rapide de paramètres des binaires galactiques

Spécialité Mathématiques appliquées

Niveau d'étude Bac+5

Formation Master 2

Unité d'accueil DPhN/LSN

Candidature avant le 28/02/2021

Durée 6 mois

Poursuite possible en thèse oui

Contact MOUTARDE Herve
+33 1 69 08 73 88
herve.moutarde@cea.fr

Autre lien http://jerome-bobin.fr

Résumé

Le travail de cette proposition de stage porte sur le développement d'une méthode rapide pour l'identification d'un grand nombre de systèmes binaires galactiques par l'utilisation de méthodes d'apprentissage automatique.

Sujet détaillé

A la suite des premières détections directes d'ondes gravitationnelles en 2015 par les interféromètres au sol LIGO et VIRGO, une nouvelle fenêtre d'observation de notre Univers s'est ouverte, ouvrant la voie à la mission spatiale LISA (Laser Interferometer Space Antenna). Cet observatoire spatial, dont le lancement est prévu en 2034, permettra la détection directe d'ondes gravitationnelles indétectables par les interféromètres terrestres.

A la différence des observatoires au sol, sensibles à de rares signaux d'ondes gravitationnelles et soumis à un bruit de mesure dominant, l'interféromètre spatial LISA devrait pouvoir observer un grand nombre de signaux distincts. Les estimations actuelles des quantités et types de source prévoient entre autres 60 millions de systèmes binaires galactiques émettant de manière continue, 10 à 100 signaux annuels provenant de trous noirs supermassifs, et de 10 à 1000 signaux annuels issus de systèmes binaires avec des rapports de masses très élevés.

Un des objectifs scientifiques de LISA est l'étude de la formation et de l'évolution des systèmes binaires galactiques : des naines blanches, mais aussi des étoiles à neutrons ou des trous noirs d'origine stellaires. LISA devrait permettre la caractérisation d'environ 25000 systèmes binaires galactiques.

L'un des défis posés par l'analyse des données de la mission LISA sera l'identification rapide, robuste et fiable des quelques 25000 systèmes binaires attendus, pouvant être détectés de manière individuelle. Plus précisément, l'identification de chaque système binaire consiste en l'estimation des paramètres des modèles physiques décrivant la forme des signatures gravitationnelles de tels systèmes. Ce problème devient d'autant plus difficile que les données

réelles seront soumises à un certain nombre de bruits et d'artefacts à prendre en compte (par exemple des données incomplètes liées à des interruptions de mesures), qui affectent significativement notre capacité à détecter et identifier les signaux issus de systèmes binaires galactiques.

A l'heure actuelle, l'approche la plus classique est fondée sur l'utilisation de méthodes Monte-Carlo Markov Chains (MCMC - voir [2]) pour l'estimation de ces paramètres par maximum de vraisemblance. Cependant, le traitement conjoint de milliers de binaires par de telles méthodes requiert l'estimation d'un grand nombre de paramètres, conduisant à un coût calculatoire très élevé et une difficulté réelle à faire fonctionner des méthodes MCMC de manière efficace.

Le travail de cette proposition de stage porte sur le développement d'une méthode rapide pour l'identification d'un grand nombre de systèmes binaires galactiques par l'utilisation de méthodes d'apprentissage automatique. A cette fin, les problèmes auxquels nous nous attaquerons seront:
* la mise en oeuvre d'une méthode d'apprentissage statistique pour l'estimation de parmètres de binaires galactiques,
* son extension à l'estimation conjointe des incertitudes d'estimation, élément essentiel permettant de caractériser les erreurs d'estimation et sa fiabilité,
* l'évaluation des résultats sur des simulations réalistes du futur observatoire LISA.
Ce travail reposera sur l'application de méthodes statistiques pour l'analyse des données, en particulier par le déploiement de méthodes Deep Learning pour l'estimation paramétrique et la quantification d'incertitudes.

Ce sujet comporte une dominante de traitement du signal et de programmation soignée; le candidat recherché doit avoir de solides connaissances en statistiques et en traitement du signal. Des connaissances préalables en apprentissage statistique sont bienvenues.

Le stage se fera au sein des l'Institut de Recherche sur le lois Fondamentales de l'Univers (IRFU), sur une thématique à l'interface entre l'analyse statistique de données et la physique des ondes gravitationnelles.
Contacts : Jérôme BOBIN jerome.bobin@cea.fr ou Hervé MOUTARDE herve.moutarde@cea.fr

Références
[1] A.Blelly, J.Bobin, H.Moutarde, "Sparsity Based Recovery of Galactic Binaries Gravitational Waves", Phys. Rev. D, en cours de publication, 2020.
[2] T.B.Littenberg, "A detection pipeline for galactic binaries in LISA data", Phys.Rev.D, 84, 063009, (2011)

Mots clés

Machine learning, statistiques, estimation paramétrique, binaires galactiques, ondes gravitationnelles

Compétences

Logiciels

Python

Analysis of the gravitational measurements of the LISA interferometer: Machine learning for the rapid estimation of galactic binary parameters

Summary

The work of this internship proposal focuses on the development of a rapid method for the identification of a large number of galactic binary systems using machine learning methods.

Full description

Following the first direct detections of gravitational waves in 2015 by the LIGO and VIRGO ground interferometers, a new window of observation of our Universe has opened, paving the way for the LISA (Laser Interferometer Space Antenna) space mission. This space observatory, scheduled for launch in 2034, will allow the direct detection of gravitational waves undetectable by ground-based interferometers.

Unlike ground-based observatories, which are sensitive to rare gravitational wave signals and subject to dominant measurement noise, the LISA space interferometer should be able to observe a large number of distinct signals. Current estimates of source quantities and types include 60 million continuously emitting galactic binary systems, 10 to 100 annual signals from supermassive black holes, and 10 to 1000 annual signals from binary systems with very high mass ratios.

One of the scientific objectives of LISA is the study of the formation and evolution of galactic binary systems: white dwarfs, but also neutron stars or black holes of stellar origin. LISA should allow the characterization of about 25000 galactic binary systems.

One of the challenges posed by the analysis of the LISA mission data will be the rapid, robust and reliable identification of the expected 25,000 or so binary systems that can be detected individually. Specifically, the identification of each binary system consists of estimating the parameters of physical models describing the shape of the gravitational signatures of such systems. This problem becomes all the more difficult as the actual data will be subject to a number of noise and artifacts to be taken into account (e.g. incomplete data due to measurement interruptions), which significantly affect our ability to detect and identify signals from galactic binary systems.

At present, the most classical approach is based on the use of Monte-Carlo Markov Chains methods (MCMC - see [2]) for the estimation of these parameters by maximum likelihood. However, the joint processing of thousands of binaries by such methods requires the estimation of a large number of parameters, leading to a very high computational cost and a real difficulty to make MCMC methods work efficiently.

The work of this internship proposal focuses on the development of a rapid method for the identification of a large number of galactic binary systems using machine learning methods. To this end, the problems to be addressed will be:
* the implementation of a statistical learning method for the estimation of galactic binary parmeters,
* its extension to the joint estimation of estimation uncertainties, an essential element for characterizing estimation errors and its reliability,
* Evaluation of the results on realistic simulations of the future LISA observatory.
This work will be based on the application of statistical methods for data analysis, in particular through the deployment of Deep Learning methods for parametric estimation and quantification of uncertainties.

This subject has a strong emphasis on signal processing and careful programming; the successful candidate must have a strong knowledge of statistics and signal processing. Previous knowledge of statistical learning is welcome.

The internship will take place at the Institut de Recherche sur le Lois Fondamentales de l'Univers (IRFU), on a theme at the interface between statistical data analysis and gravitational wave physics.
Contacts: Jérôme BOBIN jerome.bobin@cea.fr or Hervé MOUTARDE herve.moutarde@cea.fr

References:

[1] A.Blelly, J.Bobin, H.Moutarde, "Sparsity Based Recovery of Galactic Binaries Gravitational Waves", Phys. Rev. D, in publication, 2020.
[2] T.B.Littenberg, "A detection pipeline for galactic binaries in LISA data", Phys.Rev.D, 84, 063009, (2011)

Keywords

Machine learning, statistics, parametric estimation, galactic binaries, gravitational waves

Skills

Softwares

Python