



Universidad de Sonora
Departamento de Investigación en Física

Anteproyecto de tesis de Maestría en Ciencias (Física)

Encontrando amigos oscuros

Director:

Dr. Carlos Calcáneo Roldán

Estudiante:

L.F. Martín Alejandro Paredes Sosa

Hermosillo, Sonora a 5 de Junio 2020.

Resumen

Presentamos una propuesta para el desarrollo de una investigación asociada al estudio de halos de materia oscura usando técnicas de aprendizaje maquina (conocido técnicamente como *machine learning*). En esta propuesta se incorpora esta técnica a la búsqueda de sobre-densidades en simulaciones de materia oscura para lo cual tradicionalmente se usan otras técnicas.

Antecedentes

Una gran parte del trabajo que desarrollamos en cosmología en el grupo de Física Fundamental se sustenta en los siguientes antecedentes:

Existe un cúmulo de evidencia experimental y teórica que nos hace confiar en el modelo cosmológico actual, mismo que ha surgido de la confluencia entre la cosmología, la astrofísica y la física de partículas elementales. Las primeras describen el universo a gran escala, mientras la última se encarga de estudiar el mundo subatómico, su unión nos ha llevado a la realización de experimentos que están en la frontera de las posibilidades humanas, tanto materiales como intelectuales: LHC, WMAP, CDMS, LISA; por mencionar sólo algunos de los mas nombrados en la actualidad, incluso en la prensa popular.

El paradigma actual de la formación del Universo Observable se basa en una serie de suposiciones sencillas: Momentos después del gran estallido existían pequeñas imperfecciones en la distribución del plasma materia/energía que formaba al Universo. Al ir expandiéndose, el plasma se fue enfriando hasta dejar escapar fotones. Pequeñas imperfecciones en la distribución de materia fueron expandidas exponencialmente por un proceso de inflación de tal manera que se convirtieron en grumos de materia que se identifica comúnmente como las semillas de las galaxias. Los pequeños grumos van acumulándose cada vez mas grandes, a través de interacciones puramente gravitacionales, hasta formar halos dentro de los cuales, de nuevo a la atracción, se acumula materia bariónica hasta formar galaxias. Posteriormente las galaxias se agregan en cúmulos y así sucesivamente.

Para poder reproducir observaciones astrofísicas (como las curvas de velocidad circular de galaxias o la dispersión de velocidades dentro de un cúmulo de galaxias) ha sido necesario suponer que la mayoría de la materia contenida en los halos galácticos tiene características peculiares. Esta materia no puede corresponder a la materia bariónica que conocemos, su principal característica es que las partículas que forman esa materia solo deben interactuar con las partículas de la materia ordinaria a través de la fuerza gravitacional. Como dicha materia no interactúa termodinámicamente, no incrementa la temperatura de su distribución en el halo y por lo tanto no radia. Es por ello que se le ha denominado “Materia Oscura Fria” (CDM, Por sus siglas en inglés).

Hipótesis

En las simulaciones de materia oscura se determina la presencia de halos, o aglomeraciones de partículas, dentro de las mismas. Tradicionalmente se usa una técnica conocida como “amigos de amigos” (o *Friends of Friends*). Las técnicas contemporáneas de aprendizaje máquina (o *Machine learning*) pueden usarse en su lugar para lograr encontrar estas estructuras de una manera más rápida.

Justificación

Los argumentos principales para continuar con el estudio de halos de materia oscura también los hemos resaltado antes, ahora recordamos la esencia de ellos:

Uno de los modelos de formación de estructura estudiados más intensamente se basa en un universo dominado por CDM. Este modelo ha sido explorado en gran detalle por casi 30 años a través de simulaciones y modelado semi-analítico. Una propiedad importante de las simulaciones

es que los halos pequeños son los primeros en colapsarse y separarse de la expansión del Universo. Estos halos crecen lentamente al acretar grumos más pequeños o rápidamente al fusionarse con otro halo de tamaño comparable. En otras palabras, la formación de la estructura es jerárquica: los objetos pequeños se forman primero y los mas grandes después.

Una consecuencia importante de este paradigma cosmológico es que las galaxias se forman al centro de los halos de materia oscura y que, como cerca del 80 % de la materia del halo es materia oscura, ésta domina las interacciones entre las galaxias. Así, estudiando estas interacciones idealizadas esperamos entender cuáles son los factores que influyen en la evolución de las galaxias y en general la evolución de la estructura del Universo.

La razón principal para estudiar las interacciones entre halos de materia oscura es que su dinámica nos permite entender la evolución de las galaxias dentro de cúmulos de galaxias o bien de satélites dentro de la Vía Láctea. Las galaxias se observan en diferentes ambientes, desde completamente aisladas en el campo hasta agrupadas densamente en cúmulo de galaxias. Una parte importante de estudiar la evolución de galaxias es entender las transformaciones que sufren mientras orbitan dentro de este ambiente denso.

Objetivos

Tradicionalmente en la búsqueda de estructura en los halos de materia oscura se utiliza un método conocido como “*Friends of Friends*”, o amigos de amigos, en los que se recogen sobre-densidades usando distancias características entre partículas de una simulación.

En este trabajo nos proponemos aprender técnicas similares a las propuestas por Yongseok y Ji-hoon [1] o Liao, Guo et al. [3] y usar algoritmos novedosos de aprendizaje de máquina, mejor conocido como “*machine learning*” (ML de aquí en adelante) para encontrar subestructuras en halos de materia oscura.

Metas

- Completar un algoritmo de ML para la determinación de subestructuras de halos de materia oscura en una simulación cosmológica.
- Comparar el método de ML con los métodos tradicionales para encontrar estructuras.

Metodología

REVISAR

- Revisar la bibliografía relevante para el estudio de las simulaciones cosmológicas de N-cuerpos.
- Aprender la metodología de Machine Learning a través de cursos y estudio individual.
- Conseguir simulaciones cosmológicas y aprender a procesarlas. (Apoyo del profesor Calcáneo y el profesor Santana).
- Estudiar los conceptos relevantes de las simulaciones de N-cuerpos que permitan explorar el problema de encontrar halos.
- Estudiar conceptos importantes de Machine Learning.
- Estudiar conceptos importantes de Machine Learning, aplicados al problema de Halos.
- Trabajar en la programación de los algoritmos de Friends of Friends.
- Trabajar en la programación de los algoritmos de Machine Learning.
- Aplicar algoritmos a una simulación cosmológica.

- Comparar los resultados obtenidos entre ambos algoritmos.
- Presentar y discutir los temas revisados y los resultados obtenidos periódicamente en coordinación con otras personas asociadas al grupo de trabajo.

Resultados esperados

A lo largo del proyecto y como parte de su conclusión final se espera:

- Simulación cosmológica propia, o bien descarga de datos de alguna fuente confiable.
- Algoritmo de ML para detección de halos de materia oscura
- Catálogo de halos de subestructuras de materia oscura.
- Resumen del trabajo realizado en una disertación para obtener el grado de Maestría en Ciencias.

Infraestructura disponible

La investigación que nos proponemos realizar se apoya fuertemente en el cómputo mediante cálculos numéricos y el desarrollo y análisis de simulaciones. Gran parte del tiempo se invertirá en mejorar y diseñar nuevos algoritmos por lo que es de importancia fundamental la disponibilidad tanto de un buen equipo de cómputo como de software y compiladores actualizados que permitan la implementación y el desarrollo del trabajo.

Actualmente realizamos el trabajo numérico en un servidor Intel con procesador Core 2 Quad Q9550, 8Gb de memoria RAM DDRII y capacidad de disco duro combinada de 1Tb y contamos con una impresora Láser Dell Multifuncional.

En lo que concierne al material bibliográfico, el Departamento de Física cuenta con un acervo considerable, mismo que en su oportunidad habrá de actualizarse. Entre otras revistas de investigación se tiene acceso a las siguientes: The Physical Review D, Physical Review Letters, The Astrophysical Journal, Journal of Mathematical Physics, Revista Mexicana de Física, Revista Mexicana de Astronomía y Astrofísica. Además se tiene acceso en línea a la colección de Revistas del IOP, AIP y de la APS.

En cuanto a recursos de software contamos con los paquetes Mathematica (versión 11.3), Super-Mongo (versión 2.4.36), sistema operativo Linux con compiladores y graficadores de fuente abierta instalados. Realizamos el trabajo de edición en \LaTeX

Acciones de movilidad

- Participación en escuelas de entrenamiento y presentación de póster.

Recursos económicos disponibles

- Beca de CONACYT (\$11,885.17).

Habilidades que desarrollará el estudiante para la terminación del proyecto.

REVISAR

- Capacidad para actualización en un tema de investigación de frontera,
- Capacidad para identificar un problema relevante y para hacer un planteamiento adecuado para su resolución,
- Capacidad para resolver un problema de manera numérica, implementando aproximaciones relevantes, fundamentadas y adecuadas al contexto del problema,
- Capacidad para analizar los resultados obtenidos,
- Capacidad para reportar sus resultados en formato científico,

Calenderización

En la siguiente tabla indicamos la calendarización general que seguiremos para realizar el proyecto.

Etapas	Fechas	Descripción de la etapa
I	Jun - Ago 2020	Preparación de anteproyecto inicial. Preparación bibliográfica básica. y revisión preliminar.
II	Ago - Dic 2020	Presentación de anteproyecto con el comité de seguimiento. Estudio de modelaje de Machine learning. Optativa de Métodos numéricos. Estudio de simulaciones de N-Cuerpos.
III	Ene - Jun 2021	Programación de algoritmos. Optativa de Cosmología Aplicación de algoritmos
IV	Jun - Ago 2021	Edición de tesis. Presentación de examen de grado.

Referencias

- [1] Machine-assisted semi-simulation model (MSSM): estimating galactic baryonic properties from their dark matter using a machine trained on hydrodynamic simulations. J. Yongseok, K. Ji-hoon. *MNRAS*, **489**:3565,2019.
- [2] The Quijote simulations. Villaescusa-Navarro, F., Hahn, C., Massara, E., et al. 2019, arXiv e-prints, arXiv:1909.05273
- [3] HIKER: a halo-finding method based on kernel-shift algorithm. Sun, S., Liao, S., Guo, Q., et al. 2019, arXiv e-prints, arXiv:1909.13301
- [4] Gravitational N-Body Simulations – Tools and Algorithms. S. J. Aarseth. *Cambridge University Press*, Cambridge, UK, Primera edición, 2003. ISBN:0521432723.
- [5] Galactic Dynamics. J. Binney, S. Tremaine. *Princeton University Press*, Princeton, NJ, segunda edición, 2008. ISBN:978-0691130279.
- [6] Resolving Cosmic Structure Formation with the Millennium-II Simulation. M. Boylan-Kolchin, V. Springel, S. D. M. White, A. Jenkins, G. Lemson. *MNRAS*, **398**:1150,2009.
- [7] N-Body realizations of compound galaxies. L. Hernquist. *Astrophys. J. Suppl.*, **86**:389, 1993.
- [8] Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Data Processing, Sky Maps, and Basic Results. G. Hinshaw, J. L. Weiland, R. S. Hill, N. Odegard, D. Larson, C. L. Bennett, J. Dunkley, B. Gold, M. R. Greason, N. Jarosik, E. Komatsu, M. R. Nolta, L. Page, D. N. Spergel, E. Wollack, M. Halpern, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, E. L. Wright *Astrophys. J. Suppl.*, **180**:225, 2009.
- [9] The structure of cold dark matter halos. J. F. Navarro, C. S. Frenk, S. D. M. White. *ApJ*, **462**:563, 1996.
- [10] The cosmological simulation code GADGET-2. V. Springel. *MNRAS*, **364**:1105, 2005.
- [11] PKDGRAV, comunicación privada con J. Stadel.
- [12] Improved numerical modeling of clusters of galaxies. E. van Kampen. *MNRAS*, **273**:295, 1995.
- [13] The Graininess of Dark Matter Haloes. M. Zemp, J. Diemand, M. Kuhlen, P. Madau, B. Moore, D. Potter, J. Stadel, L. Widrow. *MNRAS*, **394**:641, 2009.