MEMORIAS DE LA XXVII ESCUELA DE VERANO EN FÍSICA
JUNIO 10 - 21, 2019

Eds: José Récamier Rocío Jáuregui

Instituto de Física, UNAM, 10 - 14 de junio 2019
Instituto de Ciencias Físicas, UNAM, 17 - 21 de junio de 2019

Contenido

Contribuciones
Introducción
Agradecimientos
Profesores participantes
Alumnos participantes

- **Contribuciones**
  Antonio M. Juárez Reyes 1
  Aplicaciones recientes de espectroscopia en problemas relevantes a la sociedad
  Sebastien Fromenteau 12
  Numerical and Statistics methods for Cosmology
  Guillermo Hinojosa Aguirre 24
  La física fundamental y aplicaciones de los iones negativos
  François Leyvraz 34
  Entrelazamiento, no-localidad y otras particularidades de la mecánica cuántica.
  Frédéric S. Masset 45
  Pplanetary migration in gaseous protoplanetary disks
  I Ramos-Prieto, A Paredes, H. Moya-Cessa y J. Récamier 64
  Evolución aproximada para un sistema compuesto por dos Hamiltonianos de Jaynes-Cummings acoplados
  Thomas Stegmann 77
  A brief introduction to the NEGF method for electron transport at the nanoscale
  J Alberto Vázquez, R. Medel Esquivel, I Gómez Vargas 89
  Cosmología observacional con Redes Neuronales Artificiales
  María Barrera, Carlos Villarreal 95
  Redes reguladoras complejas en la Medicina y la Biología

- **Introducción**

  La XXVII Escuela de Verano en Física fue organizada con apoyo del Posgrado en Ciencias Físicas por el Instituto de Física y el Instituto de Ciencias Físicas de la Universidad Nacional Autónoma de México. Se llevó a cabo en las instalaciones del Instituto de Física en Ciudad Universitaria, del 10 al 14 de junio de 2019 y en las instalaciones del Instituto de Ciencias Físicas, en Cuernavaca, Morelos, del 17 al 21 de junio de 2019.

  En esta Escuela se impartieron 18 cursos y 11 conferencias. Los cursos y conferencias cubrieron un amplio espectro con temas como óptica cuántica, sistemas complejos, colisiones atómicas y moleculares, mecánica cuántica, química cuántica, movimientos colectivos, cosmología y biofísica, entre otros.

Rocío Jáuregui, IF UNAM
José Récamier, ICF UNAM
Febrero, 2021

Cosmología observacional con Redes Neuronales Artificiales

J. Alberto Vázquez,[1, a] Ricardo Medel Esquivel,[2] and Isidro Gómez-Vargas[2]

[1] *Instituto de Ciencias Físicas UNAM, Av. Universidad s/n, Col. Chamilpa, Cuernavaca, Morelos, 62210, México.*

[2] *CICATA-Legaria, Instituto Politécnico Nacional, Ciudad de México, CP 11500, México.*

En estas memorias se describe el empleo de la inferencia Bayesiana en la cosmología observacional y cómo esta técnica se puede combinar con Redes Neuronales Artificiales para optimizar el tiempo de cómputo. En el transcurso del escrito, se pretende dar al lector un panorama muy general de la cosmología observacional, el torrente de datos que genera y la necesidad de incorporar algoritmos avanzados, tanto estadísticos como de inteligencia artificial, en este campo de investigación.

Palabras clave: cosmología observacional, estimación de parámetros, redes neuronales artificiales

I. INTRODUCCIÓN

Durante la última década, el rápido avance en el desarrollo de instrumentos observacionales altamente sofisticados, así como el incremento en el poder de cómputo han guiado al surgimiento de la *Cosmología de alta precisión*. En particular, experimentos desarrollados para medir las anisotropías observadas en la radiación cósmica de fondo, las distancias luminosas a partir de explosiones de Supernovas tipo Ia y la formación de estructura a gran escala. Por ejemplo, la Figura 1 muestra la distribución de varios millones de galaxias en las escalas más grandes del cosmos, cada punto representa una galaxia. Son estas escalas del Universo las que estudia la cosmología y, en particular, la cosmología observacional se enfoca en hacer uso de este tipo de información para validar los modelos teóricos que intentan explicar matemáticamente la formación, estructura y evolución del Universo.

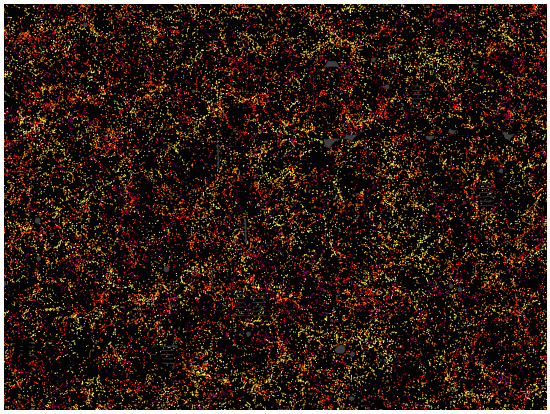


Figura 1. Distribución de varios millones de galaxias en las escalas más grandes del cosmos. Fuente: SDSS [1]

Esta imagen, la Figura 1, conjunta el trabajo proveniente del Sloan Digital Sky Survey (SDSS), un proyecto fundado en el año 2000, cuyo observatorio se encuentra ubicado en Apache Point, Nuevo México, y que tiene como objetivo capturar imágenes y señales de las profundidades del Universo.

Así como este gran proyecto se enfoca en recabar información sobre la distribución de galaxias y oscilaciones acústicas de bariones desde la Tierra, existen también satélites, como COBE[1], WMAP[2] y Planck[3], que proveen un laboratorio celeste para la cosmología y permiten estudiar otros fenómenos cósmicos desde el espacio.

Esta nueva etapa en el estudio del Universo permite combinar una serie de observaciones a diferentes escalas y, por tanto, probar nuestras teorías con mayor exactitud. Por si fuera poco, dentro de un par de años, el telescopio LSST[4] fotografiará toda la bóveda celeste con la cámara más grande jamás construida (3.2 Gigapixeles), y se estima que recopilará alrededor de 15 Terabytes por noche. Se vislumbra que la gestión y minería de datos del telescopio sean la parte técnica más difícil del proyecto. Por otro lado, el próximo radiotelescopio más grande del mundo, SKA[5], colectará aproximadamente 14 exabytes de información por día, suficientes datos como para llenar 15 millones de iPods de 64 GB a diario; además, su computadora central tendrá el poder de, más o menos, 100 millones de unidades de procesamiento. La cosmología, así como otras áreas de la ciencia, ha entrado a una nueva etapa donde la cantidad de datos juega un papel imprescindible para el entendimiento del Universo.

Jim Gray, premio Turing en 1998 y desarrollador de la base de datos del SDSS, pronosticó el advenimiento de un cuarto paradigma en la ciencia. J. Gray mencionaba que el primer paradigma ocurrió hace mil años, cuando la ciencia se limitaba a ser descriptiva y empírica; el segundo tuvo su inicio con las investigaciones de Isaac Newton, al adquirir un carácter teórico; el tercero surgió en el siglo XX, con la incorporación de las simulaciones computacionales. El cuarto paradigma sería la ciencia de

[a] javazquez@icf.unam.mx

[1] `science.nasa.gov/missions/cobe`
[2] `map.gsfc.nasa.gov`
[3] `https://bit.ly/2lQJbfK`
[4] `www.lsst.org`
[5] `https://bit.ly/2kexoY5`

datos o *e-ciencia*, una combinación de ciencia teórica con análisis computacional de una cantidad exorbitante de datos [2].

La cosmología contemporánea, con su gran caudal de datos y la necesidad de analizarlos de manera exhaustiva, cristaliza en una realidad la visión de Jim Gray del cuarto paradigma científico.

En general, el análisis de los datos cosmológicos requiere de gran variedad de técnicas estadísticas y computacionales. Por ello, las redes neuronales artificiales y otros algoritmos de inteligencia artificial, en años recientes, son útiles dentro de la cosmología. En particular, en estas memorias, mencionaremos cómo se pueden usar las redes neuronales artificiales dentro de la estimación de parámetros cosmológicos.

II. LA INFERENCIA BAYESIANA

El modelo ΛCDM es considerado el "modelo estándar" en cosmología, puesto que es la parametrización más sencilla que toma en cuenta todos los aspectos y componentes fundamentales del Universo. Este modelo incluye una componente de propiedades exóticas para dar explicación a la actual expansión acelerada del Universo, comúnmente denominada energía oscura ($\Omega_{DE} = 1 - \Omega_m$), y cuya propuesta más simple es descrita a través de una constante cosmológica (Λ). Recientes observaciones indican que la energía oscura contribuye al 70 % del contenido total del universo (Ω_{DE}), mientras que el 30 % restante corresponde a la suma de la materia oscura fría (CDM) y materia ordinaria o barionica Ω_m. La inferencia Bayesiana, en conjunto con las observaciones cosmológicas, nos permiten deducir los componentes del Universo [3]. Por ejemplo, la Figura 2 muestra dos tipos de gráficas que son muy importantes en el análisis de datos cosmológicos. La primera es conocida como distribución *posterior* 1-dimensional, donde se visualiza la región más probable de un parámetro, en este caso, de la densidad total de materia Ω_m. Por otro lado, la imagen de la derecha representa la distribución *posterior* 2-Dimensional, con las regiones de probabilidad de dos parámetros, en este caso, la densidad de materia Ω_m junto con la densidad bariónica Ω_b; la parte interna de estos elipses representa la región de confianza 1σ (68.3 %), mientras que la externa corresponde a 2σ (95.4 %). Este tipo de análisis se obtiene después de realizar un proceso de inferencia bayesiana, que se basa en datos observacionales así como en un modelo teórico dado.

El teorema de Bayes es la base de la inferencia bayesiana y, en la prueba de modelos teóricos, tiene el siguiente aspecto:

$$P(\theta|D,H) = \frac{P(D|\theta,H)P(\theta|H)}{P(D|H)},$$

donde D representa el conjunto de datos observacionales, H es la hipótesis (el modelo) y θ el conjunto de

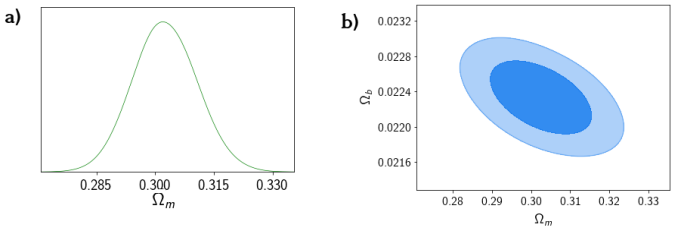


Figura 2. Distribución *posterior* sobre el contenido del Universo: a)1-Dimensional, b)2-Dimensional.

parámetros (un modelo que describe una linea, por ejemplo, tendría como parámetros libres la pendiente de una recta y su ordenada al origen). Para el caso en cuestión, un ejemplo de análisis en cosmología, estas cantidades describen los siguientes aspectos:

- D - Datos/Observaciones cosmológicas, i.e. aquellos provenientes del SDSS.
- H - Modelo cosmológico, i.e. modelo ΛCDM.
- θ - Conjunto de parámetros que mejor describen las observaciones, i.e. contenido de energía oscura Ω_{DE}.

Además, la probabilidad *previa* $P(\theta|H)$[6] representa nuestro conocimiento de los parámetros θ antes de considerar los datos observables. Esta probabilidad se modifica a través de la *verosimilitud* $P(D|\theta,H)$ al incluir datos experimentales D. El objetivo final es obtener la probabilidad *posterior* $P(\theta|D,H)$, que representa el estado de nuestro conocimiento de los parámetros del modelo al considerar la información de los datos. La constante de normalización $P(D|H)$, o *evidencia bayesiana*, es el promedio de la *verosimilitud* sobre la probabilidad *previa*:

$$P(D|H) = \int d^N\theta P(D|\theta,H)P(\theta|H),$$

donde N es la dimensión del espacio de parámetros. Esta cantidad, fundamental en la comparación de modelos, es una constante que puede omitirse en la estimación de parámetros.

La *verosimilitud* se determina al suponer una distribución estadística para los datos D, la probabilidad *previa* se determina a partir de información conocida con antelación. Cuando se calcula la distribución *posterior*, se conocen los valores más probables de los parámetros basados en el modelo teórico y en los datos observacionales.

En el cálculo de las integrales de alto grado de complejidad, que surgen de manera natural en la inferencia bayesiana, se utilizan métodos númericos especiales, conocidos como Métodos de Monte Carlo vía Cadenas de Markov (MCMC, por sus siglas en inglés).

[6] Las probabilidades involucradas son condicionales.

Explicado de manera simplificada, el método consiste en realizar un muestreo de las distribuciones de probabilidad bajo la integral, es decir, en tomar aleatoriamente una gran cantidad M de valores posibles en el espacio N-dimensional de parámetros acordes a la densidad de probabilidad expresada por la *verosimilitud*, de tal modo que se satisfagan las condiciones del Teorema Ergódico, que es una generalización de la ley fuerte de los grandes números válida para cadenas de Markov. Este muestreo, entonces, permite aproximar la integral como una sumatoria:

$$\int d^N\theta P(D|\theta,H)P(\theta|H) \approx \frac{1}{M}\sum_{i=1}^{M}\mathcal{L}(\theta_i),$$

donde $\mathcal{L}(\theta)$ es una función construída a partir de la *verosimilitud* y depende de la cantidad que se desea inferir.

Las mayores dificultades técnicas de los métodos MCMC radican en el proceso de muestreo, pues es deseable explorar todo el espacio de búsqueda, es decir, ser capaces de elegir al azar cualquier valor posible en el espacio de parámetros. Sin embargo, las cadenas de Markov son secuencias de valores aleatorios con la característica distintiva de que cada valor depende exclusivamente del valor anterior; característica que finalmente induce la convergencia de la cadena de Markov a una distribución de probabilidad específica pero también reduce la región de exploración aleatoria.

Tales dificultades técnicas han estimulado la mejora de los métodos MCMC y la búsqueda de nuevos esquemas de exploración estócastica, como el muestreo anidado, siendo en la actualidad un campo de investigación muy activo.

En el caso particular de la cosmología, dado que los modelos cosmológicos suelen tener al menos seis parámetros, diversos grupos de investigación han desarrollado códigos especializados que agrupan varios algoritmos para hacer la estimación de parámetros, por ejemplo: CosmoMC [4], CosmoSIS [5] y SimpleMC [6].

Al lector que desee ahondar en la inferencia bayesiana se le recomienda la Ref. [7] como texto de divulgación introductorio, la Ref. [8] para una introducción con mayores detalles técnicos, la Ref. [9] si se desea un tratamiento más formal, [3] es un buen tratamiento en el contexto cosmológico y en [10] se puede encontrar un texto de divulgación, similar al presente, con mayor énfasis en determinado tipo de algoritmos que permiten la inferencia bayesiana y comparación de modelos.

III. REDES NEURONALES ARTIFICIALES

En la sección anterior, se mencionó que los algoritmos involucrados en la inferencia bayesiana de un modelo cosmológico tienen que explorar un espacio de parámetros con muchas dimensiones. Las integrales involucradas para el cálculo de la función de densidad de probabilidad de *verosimilitud* y de la *evidencia bayesiana* se tornan muy complejas. Una alternativa para optimizar estos cálculos es mediante una red neuronal artificial.

La arquitectura más simple de las redes neuronales artificiales, pertenece a la clase de algoritmos del *aprendizaje automático* denominados supervisados.

En la Figura 3 se esquematiza, a grandes rasgos, las partes que componen a una red neuronal artificial simple. Cada neurona recibe datos de entrada de todas las neuronas que le anteceden. La suma ponderada es un valor que contiene la información de los valores de entrada multiplicada por los pesos w que se les asignan. Después, esta suma es el parámetro de una función no lineal de activación. La salida final de todas las neuronas interconectadas, es la que se compara con el valor esperado y con la cual se calcula el error que permite entrenar a la red neuronal.

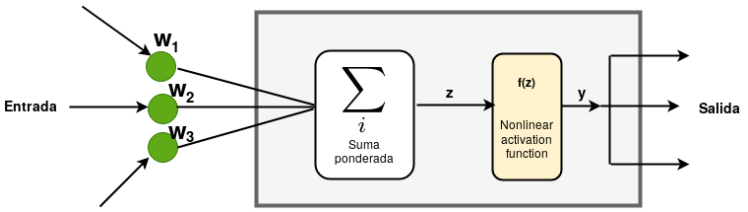


Figura 3. Esquema simplificado de una red neuronal artificial

El entrenamiento consiste el minimizar una función objetivo, llamada función de costo, que mide la diferencia entre las predicciones (aleatorias al inicio del proceso) y los valores reales conocidos de antemano. Esta minimización del error (medido por la función de costo) se efectúa después de varias iteraciones donde se aplican los algoritmos del *descenso del gradiente* y *retropropagación* (ver Figura 4).

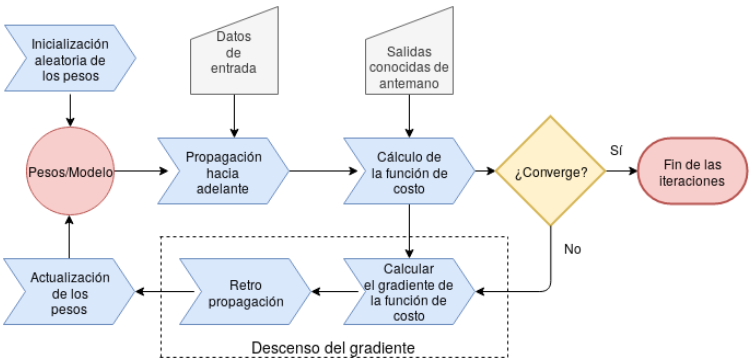


Figura 4. Entrenamiento de una red neuronal

En la Fig. 5 se puede observar un perceptrón multicapa, donde las neuronas se ordenan en paredes, y todas las neuronas de una capa reciben información de las que les preceden.

Debido a que la inferencia bayesiana es un método muy bien estructurado, las redes neuronales se pueden entrenar mediante el muestreo del espacio de parámetros y con los cálculos de las integrales que el algoritmo en

función esté realizando. Cabe mencionar, que el ajuste de parámetros de un modelo cosmológico puede requerir más de 10^6 iteraciones, donde se puede invertir el 80 % de las iteraciones en entrenar una red neuronal como la de la Fig. 5 y en las restantes utilizar la red neuronal entrenada para hacer el cálculo de las complejas integrales, y por tanto disminuir considerablemente el tiempo de computo.

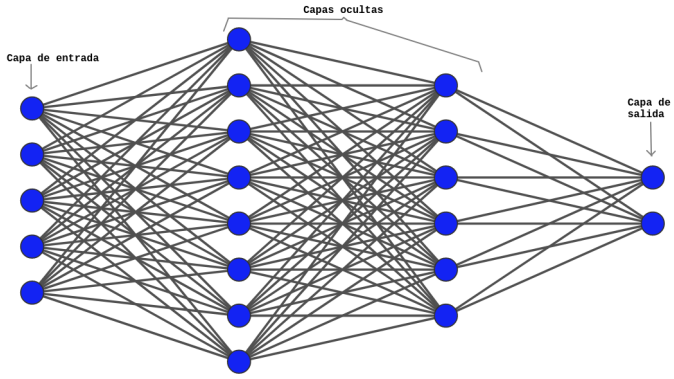


Figura 5. Red Neuronal Artificial con dos capas ocultas

IV. EJEMPLO

A modo de ejemplo, estimamos los parámetros del modelo cosmológico LCDM con el algoritmo conocido como muestreo anidado elipsoidal [11] y, también, con ese mismo algoritmo complementado mediante redes neuronales artificiales. La implementación que hemos realizado fue incorporada dentro del código SimpleMC.

Los datos observacionales utilizados fueron SN-IA, cronómetros cósmicos, oscilaciones acústicas de bariones y datos de la radiación cósmica de fondo obtenidos por el satélite Planck. En la Figura 6, se pueden apreciar los resultados obtenidos mediante ambos métodos.

Además, en la Tabla I, se comparan los tiempos que consumieron ambos métodos, así como sus cálculos de la evidencia bayesiana. Esta diferencia de tiempo, a pesar de parecer pequeña ($\sim$ 13 %), corresponderá a una mejora considerable cuando se consideren modelos cosmológicos más y una cantidad elevada de datos, como la que esperamos en los siguientes años.

	sin RNA	con RNA
tiempo (minutos)	3.194	2.819
$log(Evidencia)$	$-35{,}611 \pm 0{,}233$	$-35{,}67 \pm 0{,}231$

Tabla I. Tiempos de ejecución y evidencias bayesianas

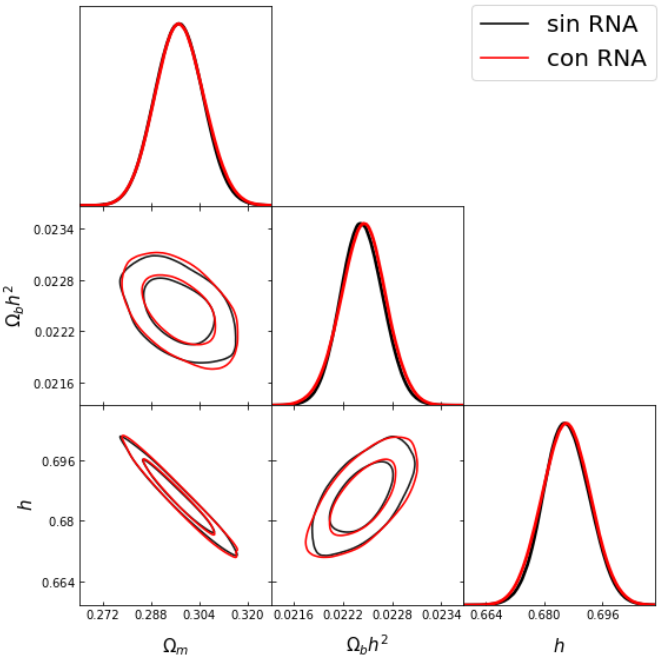


Figura 6. Comparación de resultados

V. CONCLUSIONES

La incorporación de redes neuronales dentro de la estimación de parámetros cosmológicos, permite disminuir el tiempo computacional al aprender a calcular las integrales de las funciones de *verosimilitud*. Además, se conserva la precisión tanto en la estimación de los valores más probables de los parámetros como en el cálculo de la evidencia bayesiana.

AGRADECIMIENTOS

J.A.V. agradece el apoyo proporcionado a los proyectos FOSEC SEP-CONACYT Investigación Básica A1-S-21925 y UNAM-PAPIIT IA102219. I. G. V y R. M. E. agradecen el apoyo brindado por el CONACYT a través de sendas Becas de Posgrado.

[1] Astronomers map a record-breaking 1.2 million galaxies to study the properties of dark energy. `https://bit.ly/2lMDQ92`. Accessed: 2019-09-12.

[2] T Hey, S Tansley, and K Tolle. Jim gray sobre la eciencia: un método científico transformado. In *El cuarto paradigma. Descubrimiento científico intensivo en datos*, chapter 1, pages 17–23. UAM, México, 2014.

[3] Luis E Padilla, Luis O Tellez, Luis A Escamilla, and J Alberto Vazquez. Cosmological parameter inference with bayesian statistics. *arXiv preprint arXiv:1903.11127*, 2019.

[4] Antony Lewis. Efficient sampling of fast and slow cosmological parameters. *Physical Review D*, 87(10):103529, 2013.

[5] Joe Zuntz, Marc Paterno, Elise Jennings, Douglas Rudd, Alessandro Manzotti, Scott Dodelson, Sarah Bridle, Saba Sehrish, and James Kowalkowski. Cosmosis: modular cosmological parameter estimation. *Astronomy and Computing*, 12:45–59, 2015.

[6] Éric Aubourg, Stephen Bailey, Julian E Bautista, Florian Beutler, Vaishali Bhardwaj, Dmitry Bizyaev, Michael Blanton, Michael Blomqvist, Adam S Bolton, Jo Bovy, et al. Cosmological implications of baryon acoustic oscillation measurements. *Physical Review D*, 92(12):123516, 2015.

[7] P Castro Ortega. *El Teorema de Bayes. Aproximándonos a la verdad*. RBA coleccionables, España, 2017.

[8] R Medel Esquivel, I Gómez-Vargas, JA Vázquez, and R. García Salcedo. An introduction to markov chain monte carlo. *Boletın de Estadıstica e Investigación Operativa*, 2019. Accepted but not published yet.

[9] Devinderjit Sivia and John Skilling. *Data analysis: a Bayesian tutorial*. OUP Oxford, 2006.

[10] I Gómez-Vargas, R. Medel Esquivel, R. García Salcedo, and JA Vázquez. Una aplicación de las redes neuronales en la cosmología. *Komputer Sapiens*, 11-2, 2019. DOI: 10.13140/RG.2.2.31865.52321. Accepted but not published yet.

[11] Pia Mukherjee, David Parkinson, and Andrew R Liddle. A nested sampling algorithm for cosmological model selection. *The Astrophysical Journal Letters*, 638(2):L51, 2006.