



Universidad de La Laguna

FACULTAD DE FÍSICA

# SESGO DE HALOS DE MATERIA OSCURA

*Memoria del Trabajo de Fin de Grado*

Autora:

Enia Lahcene Arzamendi

Tutor:

Dr. Andrés Balaguera-Antolínez

Marzo 2023

# Abstract

In the present work, the products of the Bias Assignment Method (BAM from now on) are going to be analyzed, in order to conclude to what extent, these products, conform to the theory.

BAM is a calibrated method representing one of the most recent bias mapping techniques aiming at generating simulated dark matter tracer catalogs quickly and accurately, by means of a trained learning approach with few reference  $N$ -body simulations. Indeed, after an evaluation of the accuracy of BAM [see e.g. 1] , it has been shown to generate assemblies with a percentage of accuracy at the two-point and three-point statistics level. The only three ingredients needed by this method are: i) an initial density field with the same initial conditions as the original reference simulation, but with a lower resolution than it, which will be the representation of the initial conditions of the BAM inputs, ii) a set of few dark matter halo catalogs whose seed corresponds to these initial conditions, iii) a gravity solver that evolves the low-resolution initial conditions to the redshift corresponding to the tracer catalog [2] .

The objective of this work will be to analyze the BAM products. The way to do this will be to study the quality of the power spectra in reproducing the two-point statistics of halos in Fourier space. To this end, the following calculations will be made for the power spectrums of one of the resulting catalogues. First, the mean value of this power spectrum at different redshift values, including their respective error bars. This average spectrum will be compared with the average value of the power spectrum, obtained from the nbodykit tool. Secondly, the variance of the spectrum will be calculated also for different redshifts, and it will be compared with the theoretical variance. The next step will be to calculate the correlation matrix of the spectrum for different values of the redshifts. Higher moments will be analyzed, in particular, the kurtosis and skewness of the spectrum, and both will be compared with the kurtosis and skewness of a Gaussian distribution with the same mean value and variance of the power spectrum that we are analyzing. In addition, a histogram of the power spectra will be made and it will be compared with the histogram of the previously mentioned Gaussian distribution, both for different values of the redshift and for different values of the  $k$  modes. The validity of these results will be determined with the Kolmogorov-Smirnov test. Once all these calculations have been generated, using the power spectra of the a dark matter density field (which has been obtained from the density field of the initial conditions, evolving it to the redshift of the halo catalog), the bias of the halos of said catalog will be analyzed. To do this, it will be calculated until what value of  $k$  the bias remains constant (since as we decrease the scale non-linearities arise) and if this maximum  $k$  value varied depending on the redshift. After this, the bias will be analyzed as a function of redshift and mode  $k$ , and this last one will be compared with the reference effective bias, also as a function of redshift and  $k$ . The last step will be to compare the power spectrum of the halos with the product of the effective reference bias and the spectrum of the underlying dark matter.

The results obtained show that, in general, the power spectrum of the halo catalog behaves correctly when we study the 2-point statistics.

Finally, it is concluded that BAM gives good results when it comes to generating precise and accurate covariance matrices in a fast way. Nevertheless, in the conclusion, we will also find some alternatives that could be carried out to obtain results with even better quality.

# Índice

<b>1</b>	<b>Introducción</b>	<b>5</b>
<b>2</b>	<b>Objetivos</b>	<b>6</b>
<b>3</b>	<b>Metodología</b>	<b>7</b>
3.1	Marco teórico: Cosmología y estructura a gran escala . . . . .	7
3.2	Evolución de la estructura a gran escala . . . . .	8
3.3	Observables cosmológicos . . . . .	10
3.4	Estadística de dos puntos . . . . .	12
3.5	Trazadores de la materia oscura . . . . .	14
3.5.1	Análisis estadístico del Universo a gran escala . . . . .	15
3.6	Espectro de potencias de halos de materia oscura . . . . .	16
3.7	Sesgo de los halos de materia oscura . . . . .	18
3.8	Comparación con modelo de sesgo de halos . . . . .	19
3.8.1	Función de masa de halos de materia oscura . . . . .	19
3.8.2	Sesgo efectivo . . . . .	19
<b>4</b>	<b>Análisis del espectro de halos de materia oscura</b>	<b>21</b>
4.1	Preámbulo . . . . .	21
4.2	BAM: Método de asignación de sesgo . . . . .	21
4.3	Análisis del espectro de potencia de halos en BAM . . . . .	23
4.3.1	Valores medios . . . . .	23
4.3.2	Varianza . . . . .	25
4.3.3	Matriz de correlación . . . . .	28
4.3.4	Momentos más altos . . . . .	30
4.3.5	Histograma de Espectros de Potencia $P(k)$ vs Histograma de Gaussianas . . . . .	32
4.4	Sesgo de los halos de materia oscura . . . . .	35
4.4.1	Sesgo en función de la escala: . . . . .	35
4.4.2	Sesgo en función del corrimiento al rojo y la escala . . . . .	36
4.4.3	Espectro de los halos de materia oscura . . . . .	38
<b>5</b>	<b>Conclusiones</b>	<b>40</b>

# Capítulo 1

## Introducción

Actualmente, nos encontramos con muchos proyectos de investigación (v.g. [3]) cuyo objetivo es la generación de una hoja de ruta hacia la construcción de catálogos simulados de galaxias que muestren un alto porcentaje de precisión en ciertas propiedades estadísticas dadas de la distribución espacial de las galaxias. Este tipo de experimentos de estructura a gran escala tales como [v.g. 4, 5, 6, 6] exigen estimaciones precisas de las matrices de covarianza y modelos exactos de agrupación de galaxias para su estudio [7, 8, 9, 10, 11, 12, 13]. Esta reconstrucción de grandes conjuntos de catálogos simulados, viene motivada por la falta de modelos analíticamente precisos que capturen la evolución gravitacional altamente no lineal que subyace a la distribución de las galaxias y los procesos físicos procedentes de estas (v.g. evolución no lineal de la materia oscura, el sesgo de galaxias, los efectos bariónicos [14, 15], distorsiones del espacio del corrimiento al rojo [16]). Sin embargo, aunque ya existen otros métodos con el mismo objetivo, es necesario un método que reduzca el tiempo/memoria computacional a la hora de generar miles de realizaciones, a la vez que genere matrices de covarianza exactas y precisas. Para esto, se han diseñado diversas alternativas, entre ellas el *Método de Asignación de Bias* (BAM), con el fin de construir un gran número de catálogos de galaxias simuladas de  $N$ -cuerpos con una mayor rapidez computacional. De hecho, se ha demostrado que BAM proporciona matrices de covarianza precisas para el espectro de potencia del halo, cuando se utiliza como entrenamiento referencia de gran volumen y se utilizan halos de materia oscura como trazadores principales de materia oscura. En el capítulo 4.2 se introducirá más detalladamente el funcionamiento de este método calibrado.

# Capítulo 2

## Objetivos

El objetivo de este trabajo consistirá en analizar estadísticamente un catálogo simulado de halos de materia oscura contruido a partir de BAM, observando la estadística de los espectros de potencia de un catálogo de 990 realizaciones, con el fin de determinar la calidad de los productos de este método. Esta valoración se va a realizar analizando hasta que punto es capaz BAM de modelar dicho espectro y reproducir la estadística de dos puntos de halos de dicho catálogo en el espacio de Fourier.

Para ello, se tendrá en cuenta que el espectro de potencias es una herramienta estadística fundamental para la descripción de campos de densidad, lo cual se introduce con más detalle en el capítulo 3.4 y este espectro de potencias se comparará con una distribución de densidad gaussiana para medir hasta que punto se ajusta a la teoría. Las mediciones estadísticas a realizar sobre los espectros de potencia, serán, entre otras:

- la media (primer momento)
- la varianza (segundo momento)
- la matriz de correlación
- curtosis y skewness (momentos mayores)
- test Kolmogórov-Smirnov

Por último, se hallará el promedio del sesgo del catálogo de halos de materia oscura en función de los modos  $k$  y en función del corrimiento al rojo  $z$  y se contrastarán dichos resultados.

# Capítulo 3

## Metodología

*This section begins with a brief introduction to modern cosmology, large-scale structure and its evolution. It continues with the definition of some cosmological observables that are relevant in this work, such as cosmological distances and redshift distortions, abundance and clustering, and a map of the distribution of 80,000 galaxies is presented. The next subsection introduces the two-point statistics, which is the key to analyze the BAM results. This subsection focuses on both the correlation function and its analog in Fourier space, the power spectrum. Both contain the cosmological information associated with the measurement of the clustering capacity of a given object, but in this paper we will focus on the power spectrum. Next, dark matter tracers and their different types of biases are described, and in addition, different methods for analyzing the statistics of the Universe on large scales, such as N-body simulations, approximate methods and statistical methods, are shown. Then, the following subsection focuses on the power spectrum of dark matter halos, how it evolves with the growth factor and how the halo bias can be calculated. Finally, the halo mass function and the effective bias are defined.*

### 3.1 Marco teórico: Cosmología y estructura a gran escala

La cosmología moderna se fundamenta en dos pilares, a saber, el Principio Cosmológico y la Teoría de la Relatividad General de Einstein [ver por ejemplo 17, 18]. El primero es la hipótesis de que el Universo es homogéneo e isótropo a grandes escalas ( $> 80$  Mpc). Esto nos permite asumir que las distancias se pueden medir con la métrica Friedmann-Robertson-Walker (FRW en adelante),

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu = dt^2 + a^2(t) \left[ \frac{dr^2}{1 - kr^2} + d\Omega^2 \right] \quad (3.1)$$

donde  $k$  es una constante que representa la curvatura del espacio, y  $a(t)$  corresponde al denominado *factor de escala cosmológico*, el cual relaciona distancias propias (físicas)  $r$  con distancias comóviles  $x$  de la forma  $x = ar$ , y  $d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$ . El segundo pilar se basa en la idea de que la estructura del espacio-tiempo está determinada por la distribución de la masa/energía en el Universo. Esto viene representado por las ecuaciones de campo de Einstein-Hilbert,

$$\mathcal{R}_{\mu\nu} - \frac{1}{2}g_{\mu\nu}\mathcal{R} = \frac{8\pi G_N}{c^2}\mathcal{T}_{\mu\nu}, \quad (3.2)$$

en donde  $\mathcal{R}_{\mu\nu}$  es el denominado *tensor de Ricci*,  $\mathcal{R} \equiv \mathcal{R}^{\mu\nu}g_{\mu\nu}$  es el *escalar de Ricci* y  $\mathcal{T}_{\mu\nu}$  es el denominado *tensor momento-energía*. Mientras que el lado izquierdo de esta expresión no indica las propiedades geométricas del espacio-tiempo, el lado derecho nos informa sobre el contenido de materia/energía en el universo responsable de dichas características geométricas.

En cosmología, la suposición más común es la de asumir que el contenido de materia energía a grandes escalas se puede representar a través de un fluido ideal, para el cual el tnesor momentum-energía se escribe de la forma

$$\mathcal{T}_{\mu\nu} = pg_{\mu\nu} + (p + \rho)u_\mu u_\nu, \quad (3.3)$$

en donde  $p$  y  $\rho$  denotan respectivamente la presión y la densidad de energía de la materia en el Univeso, y la cantidad  $u_\mu$  representa el cuadri-vector de velocidad de un elemento de fluido. Las distintas componentes de materia/energía en el Universo obedecen distitnas leyes de evolución temporal, así como distintas formas de ecuaciones de estado (es decir, la forma en que su presión y si densidad se relaciona): por ejemplo,  $p(\rho) = \omega\rho$ , en donde  $\omega = 0, 1/3, -1$  para la materia osruda, la radiación y la energía oscura. Si se asume que cada una de estas components evoluciona de forma independiente, la ley de conservación asociada a cada una nos permite escribir  $\rho(a) \sim a^{-n}$  donde de nuevo  $n = 3, 2, 0$  para la materia oscura, la radiación y la energía oscura,

La combinación de la métrica FRW y las ecuaciones de campo de Einstein-Hilbert generan las denominadas *ecuaciones de Friedman* [19]. Dichas ecuaciones representan la evolución del factor de escala como función del tiempo cosmológico, parametrizadas por las distintas componentes del universo (v.g., materia oscura, energía oscura, radiación).

$$\frac{\dot{a}}{a} = \left( \Omega_{\text{cdm}} a^{-3} + \Omega_{\text{rad}} a^{-4} + \Omega_{\text{vac}} a^{-3(1+\omega_x)} + a^{-2}(1 - \Omega_{\text{tot}}) \right)^{\frac{1}{2}} \equiv H(a) \quad (3.4)$$

en donde

$$\Omega_{\text{tot}}(t) \equiv \frac{8\pi G_{\text{N}} \rho_{\text{tot}}(t)}{3H^2(t)} \equiv \frac{\rho_{\text{tot}}(t)}{\rho_{\text{crit}}(t)}. \quad (3.5)$$

En esta expresiñon  $\rho_{\text{tot}}$  denota la desinsidad de energía total en el Universo, mientras que  $\rho_{\text{crit}}$  es la densidad crítica del Universo, cantidad con respecto a la cual se determina si el Universo es cerrado, abierto o plano. Viene definida como

$$\rho_{\text{crit}}(t) \equiv \frac{3H(t)^2}{8\pi G_{\text{N}}}, \quad (3.6)$$

Los distintos modelos cosmológicos difieren tanto en las contribuciones relativas a los distintos tipos de materia y energía, como en la naturaleza de estas. Las últimas mediciones provenientes de la radiación cósmica de fondo [20, 21] y de la aglomeración de galaxias [22, 4, 23] concluyen que la curvatura del Universo es estadísticamente compatible con  $k = 0$ .

Los distintos modelos cosmológicos difieren en cuanto a las contribuciones relativas de la materia bariónica (protones, electrones y neutrones), la materia oscura y la energía, y en cuanto a la naturaleza de materia y energía oscuras. Los experimentos basados en la radiación cósmica de fondo y la aglomeración de galaxias han dmeostrado que el modelo  $\Lambda$ CDM es el que mejor ajusta las mediciones de las diferentes propiedades estadísticas (v.g. estadística de dos, tres puntos). Los resultados muestan a un Universo plano con  $\sim 4\%$  materia bariónica (es decir, de la que están hechas las estrellas y las galaxias), un  $\sim 30\%$  materia oscura fría y un  $\sim 66\%$  restante denominado energía oscura [24]. Sin embargo, la naturaleza exacta de la energía oscura y la materia oscura fría sigue siendo un misterio.

## 3.2 Evolución de la estructura a gran escala

La estructura a gran escala (EGE a partir de ahora) [19] se puede entender como el conjunto de características estadísticas de la distribución espacial de los distintos componentes del Universo. De acuerdo con el Principio Cosmológico, la distribución uniforme de la materia en

el universo no sería compatible con la existencia de galaxias o cúmulos de galaxias. Para llegar a este tipo de estructuras, a partir de un Universo casi homogéneo en sus orígenes tenemos que entender la evolución de las pequeñas perturbaciones que se originaron durante las etapas más tempranas después del Big Bang. En su etapa inicial, el Universo pasa a través de una fase de expansión acelerada denominada inflación [ver por ejemplo 25]. En esta etapa, regiones inicialmente en contacto causal se desacoplan (es decir, el radio de Hubble se mueve más rápido que la velocidad de la luz). El final de la inflación se denomina recalentamiento, en este proceso la inflación desintegra las partículas en un plasma térmico caliente denominado plasma primigenio. Regiones ya fuera de cualquier contacto causal comienzan su evolución a partir de pequeñas fluctuaciones de densidad. En gran parte de esta época, el Universo está dominado por la radiación, lo cual desacelera el proceso de ampliación de estructuras gravitacionalmente (el potencial gravitatorio decae) y el inicio de formación eficiente de estructuras se retrasa. Estos dos procesos contrapuestos, el efecto gravitatorio que tiende a aumentar la acumulación de perturbaciones de densidad y la radiación, que tiende a disiparla, producen oscilaciones en las perturbaciones de densidad, las cuales se propagan en el plasma como ondas acústicas. En otras palabras, en este momento nos encontramos en una lucha entre los componentes del plasma, partículas subatómicas y fotones, donde las primeras se encuentran constantemente absorbiendo y reemitiendo a los segundos, impidiéndoles recorrer largas distancias. El cambio comienza en el momento en el que hay una igualdad entre cantidad de materia y radiación, ese momento comienza cuando las perturbaciones de densidad del Fondo Cósmico de Microondas (FCM a partir de ahora) empiezan a tomar el control. A partir de ese momento, el potencial gravitatorio permanece estable, lo cual conlleva una acumulación más rápida de partículas de materia oscura en los pozos del potencial gravitatorio. De este modo, mientras la materia oscura comienza a colapsar bajo la influencia de la gravedad, la evolución del fluido bariónico-fotónico sigue estando gobernada por las ondas acústicas que más tarde quedarán impresas en el FCM. El crecimiento de las perturbaciones de densidad bariónicas no se acelera hasta 380.000 años después del big bang, tras el desacoplamiento de los fotones. Este crecimiento se debe a la caída de los bariones hacia las sobredensidades de materia oscura inducida por este desacoplamiento. Tras este desacople, la materia queda solo expuesta a la gravedad, y el patrón de las oscilaciones acústicas queda impreso en el FCM en el instante del desacople.

Tras el desacople, el FCM se desplaza hacia el rojo (su longitud de onda aumenta) debido a la expansión cósmica, de forma que abandona la zona visible del espectro electromagnético, de forma que el Universo ya no contiene ninguna fuente de luz visible, lo que da lugar al comienzo de la edad oscura. En esta época, la mayor parte de la materia bariónica reside en el hidrogeno neutro, que ocasionalmente absorbe fotones del FCM en transiciones de giro de 21 cm [26]. Al mismo tiempo las inhomogeneidades en las perturbaciones de densidad bariónica siguen aumentando, y cuando alcanzan un contraste de densidad de  $\delta\rho/\rho \sim 1$ , ocurre un desacople que se define como "turn-around", por el que la región con sobredensidad se desacopla de la expansión y empieza a colapsar. Este momento en el que el tamaño físico de la perturbación se encuentra en su máximo, marca la transición del régimen medianamente no lineal al régimen altamente no lineal en escalas pequeñas. Alrededor de cien millones de años después de la edad oscura, este proceso lleva a la formación de la primera generación de estrellas (estrellas de población III), seguida de la formación de las primeras galaxias y cuásares (galaxias con agujeros negros supermasivos en su centro) durante los siguientes cientos de millones de años.

El hidrogeno neutro absorbe parcialmente y reemite la radiación ultravioleta Lyman-alpha generada por las primeras estrellas antes de ser calentado por fuentes luminosas de rayos X (nucleos galácticos activos o el medio interestelar caliente). Esta época se denomina amanecer cósmico, la cual da lugar a un segundo rasgo característico de absorción de 21 cm en el FCM [27, 28].

Otro efecto importante de la primera luz estelar en el Universo es que ioniza gradualmente el gas de hidrógeno neutro. Este proceso recibe el nombre de reionización [29] y dura desde unos 250 millones de años hasta unos mil millones de años después del big bang. La reionización tiene lugar a causa de la energética radiación ultravioleta emitida por las primeras estrellas, galaxias y cuásares, los cuales separan los electrones de los protones revertiendo así los resultados de la recombinación.

Tras la época de la reionización, la formación de estructuras continua su curso, lo cual, gradualmente da lugar a lo que hoy en día se conoce como la red cósmica (distribución de materia que da forma a la apariencia del Universo hasta el día de hoy). La red cósmica está formada por halos (nudos densos) de materia oscura, conectados entre sí por estructuras filamentosas y rodeados por regiones desprovistas de materia (vacíos). En la figura 3.1 podemos ver esa compleja distribución, y como puede ser reemplazada por una distribución de halos de materia oscura. La materia bariónica traza la distribución de la materia oscura fría y esto da lugar a la formación de galaxias en regiones con sobredensidad. A medida que pasa el tiempo, estas galaxias forman cúmulos de galaxias y los supercúmulos más masivos se generan en el centro de halos densos de materia oscura.

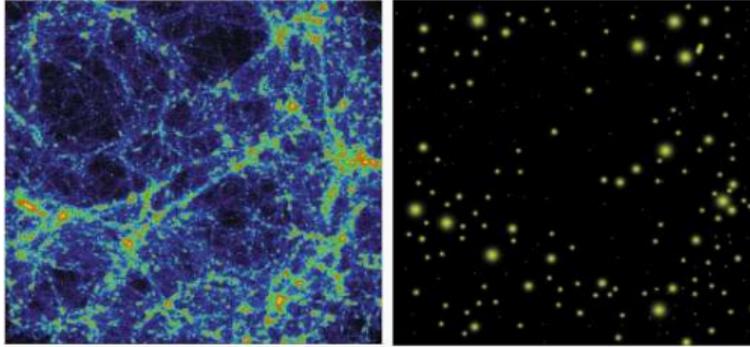


Figure 3.1: A la izquierda se encuentra la distribución de la materia oscura de una simulación, mientras que la imagen a la derecha es una representación de la misma reemplazada por halos de materia oscura [Adaptado de 30]

La dinámica y la evolución de la red cósmica es un proceso altamente no lineal, por lo que para su estudio es necesario la generación de simulaciones de  $N$ -cuerpos que se describen más a fondo en el capítulo 4.

### 3.3 Observables cosmológicos

La distribución de estructuras en el Universo nos brinda información sobre las fluctuaciones en el universo primitivo, así como sobre la distribución de la materia oscura, la formación de galaxias y una medida de la velocidad de expansión del universo.

- Distancias cosmológicas y distorsiones por desplazamiento al rojo

$$r(z) = \int_0^z \frac{dz}{H(z)} \quad (3.7)$$

en donde  $z$  es el desplazamiento al rojo cosmológico,

$$z = \frac{1}{a} - 1 \quad (3.8)$$

- Abundancia: el número de objetos (galaxias, halos de materia oscura) por unidad de volumen es una cantidad que depende del modelo cosmológico y de sus parámetros. Las cantidades que usualmente se estudian en este aspecto son la función masa de halos de materia oscura o la función de luminosidad de galaxias.
- Agrupación (estadística de  $n$ -puntos): Se trata de los objetos galácticos con mayor densidad de la EGE los cuales se enuecentran ligados gravitacionalmente. El grado de agrupación se mide a través de la función de correlación, que cuantifica el exceso de probabilidad de encontrar  $n$  objetos a una dada distancia respecto de una distribución de Poisson. También se puede utilizar el espectro de potencias, que es equivalente a la función de correlación, pero en el espacio de Fourier.

Anteriormente he comentado que la mayor parte de la masa del Universo es oscura y que en sus orígenes esta masa estaba distribuída de forma bastante suave. Sin embargo, si nosotros quisiéramos medir la distribución de la materia oscura no podríamos hacerlo directamente ya que esta no emite ningún tipo de radiación electromagnética ni tampoco interactúa en ninguna forma con ella. Por lo tanto, hay que recurrir a otro método para su medición, como es el considerar que las galaxias luminosas que vemos hoy en día, pueden ser trazadores (aunque más adelante veremos que son sesgados) de la distribución de la materia oscura.

En la figura 3.2 podemos observar la distribución espacial de 80.000 galaxias en el 2dFGRS (es un sondeo de redshift de galaxias elaborado por el Observatorio Anglo-Australiano (AAO) [31]), en el cual se estima la distancia de las galaxias a partir del redshift medido. Si observamos la figura es evidente que la distribución de las galaxias no es aleatoria, si no que estas galaxias están distribuídas en grupos y cúmulos de mayor o menor densidad de galaxias, o en estructuras filamentosas que rodean regiones desprovistas de galaxias. Estas regiones vacías tienen diámetros de hasta 100 Mpc y contienen ninguna o muy pocas galaxias. [19]

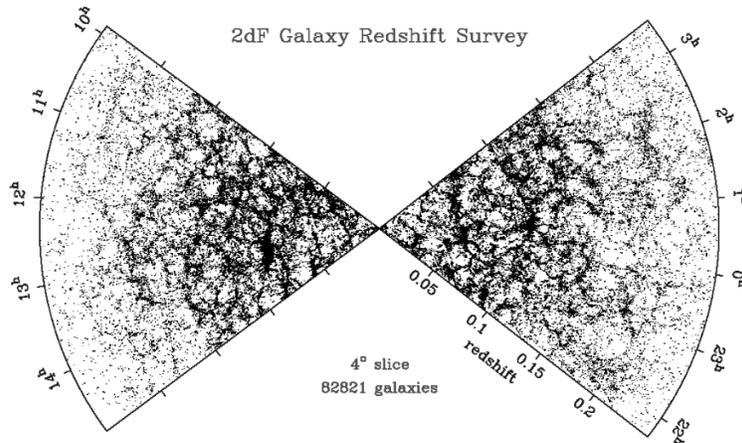


Figure 3.2: Distribución de 80.000 galaxias en el 2dFGRS en un corte de  $4^\circ$  proyectado sobre el plano de desplazamiento al rojo/ascensión derecha. [Adaptado de 31]

Para poder hacer mediciones en la distribución de galaxias vamos a considerar a estas distribuciones como un conjunto de puntos en el espacio y analizaremos las correlaciones espaciales entre estos puntos estadísticamente. Para ello vamos a introducir la estadística de dos-puntos y las herramientas que utilizaremos, las cuales son: la función de correlación (en el dominio real) y el Espectro de Potencias (en el dominio de Fourier).

### 3.4 Estadística de dos puntos

La estadística de dos puntos es la herramienta estadística más importante para caracterizar la estructura a gran escala del Universo, y por ende para extraer información cosmológica a través del análisis del agrupamiento de galaxias [ver v.g. 32]. La estadística de dos puntos se puede definir en el espacio de configuración, es decir, como función de la separación entre objetos; a esta cantidad se le denomina *función de correlación*, y es un concepto ampliamente utilizado otros campos de la física. Igualmente, la estadística de dos puntos puede ser definida el espacio de Fourier, y se le denomina *espectro de potencias*. Dentro de cada una de estas dos categorías, existen versiones que dependen de la descripción espacial que se use, por ejemplo, la función de correlación de dos puntos en 2 o 3 dimensiones. [33, 18, 34]

#### Función de correlación

Consideremos una distribución homogénea de objetos con densidad  $\rho$ . Asumimos que la distribución de objetos obedece a un proceso de Poisson, con lo cual probabilidad de encontrar un objeto en un elemento de volumen infinitesimal  $dV$  es  $dP = \rho dV$ . La probabilidad conjunta de encontrar un par de objetos a una distancia  $r$ , en volúmenes  $dV_1$  y  $dV_2$  es igual al producto de las probabilidades independientes:  $dP_{12} = \rho^2 dV_1 dV_2$ . No obstante, la distribución espacial de objetos cuyo origen está asociado a la evolución de fluctuaciones en densidad, y por ende, al efecto de la gravedad, no puede ser descrita con un modelo homogéneo de Poisson. Se introduce así el proceso *inhomogéneo* de Poisson, caracterizado por:

$$dP_{12} = \rho^2(1 + \xi(r_{12}))dV_1 dV_2, \quad (3.9)$$

en donde  $\xi(r)$  es la denominada *función de correlación de dos puntos* (FC en adelante). A partir de esta expresión, se entiende la FC como una medida del exceso de probabilidad de encontrar dos objetos separados por una distancia determinada, con respecto a la misma probabilidad en una distribución homogénea. En términos de promedios sobre ensambles estadísticos, la FC se define como

$$\xi(\vec{x}) = \langle \delta(\vec{r} + \vec{x})\delta(\vec{r}) \rangle, \quad (3.10)$$

donde la fluctuación en densidad  $\delta$  se define como

$$\delta(\vec{r}) = \frac{\rho(\vec{r}) - \langle \rho \rangle}{\langle \rho \rangle} \quad (3.11)$$

con  $\langle \rho \rangle$  como el valor medio de la densidad. Este valor medio debe, en principio y siguiendo la definición de la FC, ser calculado como un valor esperado dentro de un ensamble estadístico <sup>1</sup> de universos. Sin embargo, dado que en cosmología sólo tenemos acceso a un Universo, determinamos este valor medio como una integral sobre el volumen de la cantidad en cuestión. Esta aproximación es válida siempre y cuando el volumen bajo inspección pueda ser considerado como una muestra representativa del Universo.

#### Espectro de potencias

El espectro de potencias  $P(k)$  es la transformada de Fourier de la función de correlación. Para ver su forma, tomemos la transformada de Fourier de un campo de sobredensidad  $\delta(\vec{r})$ :

$$\delta(\vec{k}) = \int_{-\infty}^{\infty} \delta(\vec{r}) e^{-i\vec{k}\vec{r}} d^3r \quad (3.12)$$

<sup>1</sup>Se trata de una colección hipotética de sistemas (por ejemplo, termodinámicos) que siguen las mismas leyes físicas representada a través de propiedades estadísticas. En el contexto cosmológico, un ensamble se representa como una serie de universos independientes y diferentes punto a punto, que comparten las mismas propiedades estadísticas [35] (v.g. valores medios, dispersiones, covarianzas).

El campo de fluctuaciones en densidad  $\delta(\vec{r})$  es una función real, por lo que, su transformada compleja tiene que satisfacer  $\delta(\vec{k}) = \delta(-\vec{k})^*$ , en donde "\*" denota conjugación compleja. Por lo tanto:

$$\delta(\vec{k}')^* = \int_{-\infty}^{\infty} \delta(\vec{r}') e^{i\vec{k}'\vec{r}'} d^3r' \quad (3.13)$$

Si se sustituyen estos valores para obtener  $\langle \delta(\vec{k})\delta(\vec{k}')^* \rangle$ , se obtiene:

$$\langle \delta(\vec{k})\delta(\vec{k}')^* \rangle = \left\langle \int_{-\infty}^{\infty} d^3r d^3r' \delta(\vec{r})\delta(\vec{r}') e^{-i\vec{k}\vec{r}} e^{i\vec{k}'\vec{r}'} \right\rangle = \int_{-\infty}^{\infty} d^3r d^3r' \langle \delta(\vec{r})\delta(\vec{r}') \rangle e^{-i(\vec{k}\vec{r} - \vec{k}'\vec{r}')} \quad (3.14)$$

siendo  $\vec{r}' = \vec{r} + \vec{x}$  y , por lo tanto,  $\langle \delta(\vec{r})\delta(\vec{r}') \rangle = \langle \delta(\vec{r})\delta(\vec{r} + \vec{x}) \rangle$ , que como vimos en la ecuación 3.10, es la definición de la función de correlación  $\xi(\vec{x})$ . Por otra parte:  $d^3r' = d^3x$ . Por lo tanto la expresión 3.14, puede escribirse como:

$$\langle \delta(\vec{k})\delta(\vec{k}')^* \rangle = \int_{-\infty}^{\infty} d^3r d^3x \xi(\vec{x}) e^{-i\vec{k}\vec{r} + i\vec{k}'\vec{r} + i\vec{k}'\vec{x}} = \int_{-\infty}^{\infty} d^3r d^3x \xi(\vec{x}) e^{-i\vec{r}(\vec{k} - \vec{k}')} e^{i\vec{k}'\vec{x}} \quad (3.15)$$

Utilizando la transformada de Fourier de la delta de Dirac:

$$\langle \delta(\vec{k})\delta(\vec{k}')^* \rangle = \int_{-\infty}^{\infty} d^3x \xi(\vec{x}) e^{i\vec{k}'\vec{x}} \int_{-\infty}^{\infty} d^3r e^{-i\vec{r}(\vec{k} - \vec{k}')} = \int_{-\infty}^{\infty} d^3x \xi(\vec{x}) e^{i\vec{k}'\vec{x}} \int_{-\infty}^{\infty} (2\pi)^3 \delta^D(\vec{k} - \vec{k}') \quad (3.16)$$

Considerando entonces que el espectro de potencias es la tranformada de Fourier de la función de correlación:

$$\langle \delta(\vec{k})\delta(\vec{k}')^* \rangle = (2\pi)^3 \delta^D(\vec{k} - \vec{k}') P(\vec{k}) \quad (3.17)$$

Siendo  $P(\vec{k}) = P(|\vec{k}|) = P(k)$ , y por lo tanto:

$$\langle \delta(\vec{k})\delta(\vec{k}')^* \rangle = (2\pi)^3 \delta^D(\vec{k} - \vec{k}') P(k) \quad (3.18)$$

La función de correlación como el espectro de potencias contienen la misma información cosmológica y dicha información está asociada a una medida de la capacidad de aglutinamiento de cualquier objeto (v.g. materia oscura, trazadores de la materia oscura). Sin embargo, cada una de estas versiones de la estadística de dos puntos esta sometida a diversas fuentes de efectos sistemáticos y a la hora de ser medidas pueden tener diversos niveles de complejidad numérica.

Con el fin de medir el espectro de potencia, se han desarrollado diferentes métodos para obtener un estimador de esta medida. En este capítulo se va a desarrollar uno de estos métodos [36].

En primer lugar, se supone que las galaxias forman una muestra de Poisson [] del campo de densidad  $1 + f(r)$ :

$$P(\text{el elemento } \delta V \text{ contiene una galaxia}) = \delta V \bar{n}(r) [1 + f(r)] \quad (3.19)$$

donde  $\bar{n}$  es la densidad espacial media esperada de las galaxias dado el criterio de selección angular y de luminosidad. Como se ha introducido antes, el espectro de potencias es la transformada de Fourier de la FC:

$$P(k) = P(\mathbf{k}) \equiv \int d^3r \xi(r) e^{i\mathbf{k}\cdot\mathbf{r}} \quad (3.20)$$

donde  $\xi(\mathbf{r}) = \xi(r) = \langle f(r')f(r'+r) \rangle$ .

A continuación, en la figura 3.3 se representa el espectro de potencias hoy para dos modelos cosmológicos diferentes ( $\Lambda$ CDM <sup>2</sup> y sCDM <sup>3</sup>). En los siguientes capítulos (mandar a dicha sección) se describirá más a fondo el espectro de potencias de la materia oscura. Se puede observar que a escalas pequeñas el espectro se invierte y cae cuanto mayor es el modo  $k$ . Esta inversión del espectro de potencias se llama turnover y se da a mayor o menos escala según el modelo cosmológico elegido, ya que este turnover corresponde a la entrada del modo en el horizonte en igualdad de materia/radiación. En la figura 3.3, el espectro de  $\Lambda$ CDM, se invierte en escalas más grandes ya que tiene un  $a_{eq}$  mayor. Esto se debe a la cantidad de materia de cada modelo; un universo con más materia tiene un  $a_{eq}$  más temprano, esto significa que solo las escalas muy pequeñas ingresan en el horizonte durante la época dominada por la radiación y en consecuencia el turnover ocurre en escalas más pequeñas. Otro aspecto a tener en cuenta es que por encima de  $k_{nl} = 0.2h\text{Mpc}^{-1}$  no se pueden ignorar las no linealidades y por lo tanto no podemos comparar el espectro de la figura a partir de dicho  $k$  a la distribución de materia hoy. [34]

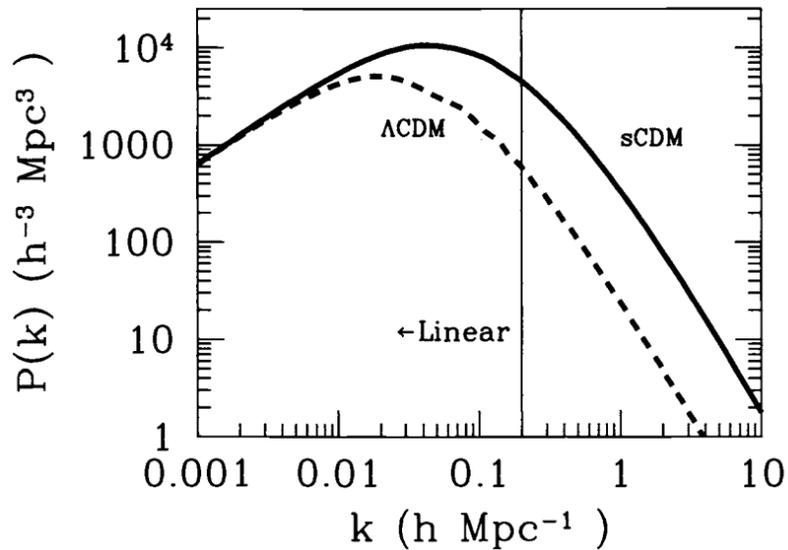


Figure 3.3: Espectro de potencias para dos modelos cosmológicos de Materia Oscura Fría ( $\Lambda$ CDM y sCDM). El espectro se ha normalizado para coincidir en escalas grandes. [Tomado de 34]

### 3.5 Trazadores de la materia oscura

La información que podemos recopilar para inferir las propiedades del Universo a gran escala sólo la podemos entender a través de la distribución espacial de la materia que emite radiación electromagnética, es decir, galaxias y cúmulos de galaxias. Sin embargo, existe un sesgo entre las propiedades estadísticas de la distribución espacial de dichos trazadores y las de la materia oscura subyacente. Entender este sesgo es fundamental para lograr inferir las

<sup>2</sup>Modelo cosmológico de materia oscura fría con una constante cosmológica  $\Lambda$  asociada a la energía oscura

<sup>3</sup>Modelo cosmológico de materia oscura fría estándar

características más importantes del Universo. Sin embargo, hay un paso intermedio para llegar a un entendimiento claro de las propiedades de la materia oscura. El modelo del halo [30], ampliamente aceptado en el contexto del modelo estándar cosmológico y soportado por resultados observacionales, establece que los halos de materia oscura son la el vínculo entre materia oscura y las galaxias. Por un lado, estos objetos se forman a partir del colapso gravitacional de regiones del Universo que se desacoplan del flujo de Hubble (expansión). Por otro lado, dichos objetos, al ser los más masivos generan pozos de potenciales dentro de los cuales los bariones caen para formar las galaxias.

Si  $\delta_{\text{tr}}$  denota la sobredensidad de trazadores, y  $\delta_{\text{dm}}$  la de la materia, podemos identificar distintos tipos de sesgos como:

- Local y determinista: Radica en asumir la sobredensidad de halos proporcional a la sobredensidad de materia oscura,  $\delta_{\text{tr}} = b\delta_{\text{dm}}$
- Local y estocástico: Este caso es más general, y asume que la sobredensidad de trazadores obedece una relación estadística con respecto a la materia oscura, esto es, existe una distribución de probabilidad  $P(\delta_h|\delta_m)$  tal que el valor esperado de sobredensidad es:

$$\langle \delta_h | \delta_m \rangle = \int d\delta_h \delta_m P(\delta_h | \delta_m) \quad (3.21)$$

- No-local: Se asume que el sesgo depende de la posición en la que se está mediando y las posiciones en un radio determinado alrededor de dicha celda.

En general, el sesgo de los trazadores (halos de materia oscura) posee todas estas propiedades. Otro tipo de sesgo del que se habla es aquél que está vinculado con las propiedades de los halos de materia oscura, en particular, cómo halos de materia oscura con distintas propiedades (tales como su masa) tiene una distribución espacial distinta con respecto a la materia oscura.

### 3.5.1 Análisis estadístico del Universo a gran escala

El análisis estadístico de los distintos experimentos de estructura a gran escala busca determinar, dadas las características del experimento y dado un modelo subyacente, los valores e incertezas en los parámetros cosmológicos que caracterizan dicho modelo. Incluso, dichos análisis se pueden extender a determinar si un modelo es el adecuado para describir las observaciones.

La maquinaria para determinar parámetros cosmológicos se basa en una parte en el nivel de predicción de un modelo (por ejemplo,  $\Lambda$ CDM), en las características del experimento (por ejemplo, sensores de galaxias, o de radiación cósmica de fondo) y por último, en la capacidad de generar estimaciones robustas de las matrices de covarianza de los distintos observables cosmológicos. Este trabajo se enfoca en este último aspecto.

Las matrices de covarianza para distintos observables cosmológicos contienen la información sobre la incerteza que se obtiene al hacer una medición y el modo en que la misma medición hecha a distintos puntos (modos) se correlacionan. Estas cantidades se definen como promedios sobre ensambles estadísticos, lo cual impone la necesidad de generar herramientas que nos permitan definir, a partir de un solo Universo observable, un ensamble de universos. Es de uso común utilizar del orden de  $> 1000$  réplicas de Universos para definir un ensamble a partir del cual poder determinar de forma precisa las matrices de covarianza. Para alcanzar dicho objetivo, se emplean distintos métodos, a saber:

- Simulaciones de  $N$ -cuerpos (N-BODY a partir ahora)[ver v.g. 37, 38], a partir de las cuales se pueden construir los denominados catálogos sintéticos (mock catalogs, en inglés). Estas

simulaciones representan soluciones exactas a las ecuaciones de la dinámica de partículas de materia oscura en el espacio de fase desde altos valores de  $z$  (régimen lineal) hasta bajos desplazamientos, en donde la evolución es altamente no lineal. Una de las grandes desventajas de este método es su alto costo computacional.

- Métodos aproximados [ver v.g. 39, 40], una ruta alternativa para evitar el alto costo computacional de las simulaciones de  $N$ -cuerpos es ejecutar algoritmos rápidos, los cuales reproduzcan en una medida variable, las estadísticas a gran escala de estas simulaciones. Sin embargo, la gran desventaja es que se pierde precisión en la física a pequeña escala, ya que los métodos aproximados no son capaces de resolver las subestructuras del halo.
- Métodos estadísticos. Cuando generar ensambles es prohibitivo, una opción ampliamente utilizada consiste en dividir la muestra en subvolumenes, con métodos denominados *Jack-knife* o *bootstrap*. El tamaño de estos subvolumenes tiene que ser tal que sea una muestra representativa (estadística, ente hablando). (Referencias)

### 3.6 Espectro de potencias de halos de materia oscura

La evolución de las perturbaciones cosmológicas se divide en tres etapas [34]:

1. Todos los modos se encuentran fuera del horizonte ( $k\eta^4 \ll 1$ ) y el potencial se mantiene constante.
2. Los modos entran en el horizonte y el universo pasa de la época de dominación de radiación ( $a \ll a_{eq}$ <sup>5</sup>) a la época de dominación de materia ( $a \gg a_{eq}$ ). Los modos de pequeña escala ( $k$  grande) entran en el horizonte antes de ( $a_{eq}$ ), evolucionan de forma muy diferente a los modos a gran escala ( $k$  pequeña) que entran en el horizonte mucho después de  $a_{eq}$ .
3. Todos los modos evolucionan idénticamente, permaneciendo constantes para  $\Omega_m = 1$ . Esta tercera etapa se denomina tiempos tardíos.

Basándonos en el capítulo 3.4 el espectro de la materia oscura viene dado por:

$$\langle \delta_{dm}(\bar{k}') \delta_{dm}(\bar{k}) \rangle = (2\pi)^3 \delta_p^3(k - k') P_{dm}(k) \quad (3.22)$$

Esta forma asume isotropía y homogeneidad.

En las siguientes líneas se desarrollará una vía para obtener el espectro de potencias de la materia oscura rastreándolo hacia atrás en el tiempo.

En tiempos tardíos, el potencial se puede expresar en relación a un espectro inicial ( $P_{dm}(k, z = z_{ini})$ ) como:

$$P_{dm}(k, z) = D_+^2(z) T^2(k) P_{dm}(k, z = z_{ini}), \quad (3.23)$$

donde  $D_+^2$  es el factor de crecimiento y  $T(k)$  es la función de transferencia que describe la evolución de las perturbaciones a través de las épocas de cruce de horizontes y transición de dominación de radiación/materia, la cual para tiempos tardíos a gran escala, se aproxima a la unidad, por lo que podemos escribir el espectro como:

$$P_{dm}(k, z) = D_+^2(z) P_{dm}(k, z = z_{ini}). \quad (3.24)$$

<sup>4</sup>Horizonte comóvil ( $\eta$ ): Distancia comóvil máxima recorrida por la luz desde el comienzo del universo

<sup>5</sup>Época de igualdad entre radiación y materia

Para obtener una expresión para el factor de crecimiento, se parte de las siguientes tres ecuaciones, las dos primeras correspondientes al conjunto de ecuaciones que rige la evolución de la materia oscura y la tercera la ecuación algebraica para el potencial gravitatorio [v.g. 34]:

$$\dot{\delta} + ik\nu = -3\dot{\Phi} \quad (3.25)$$

$$\dot{\nu} + \frac{\dot{a}}{a}\nu = ik\dot{\Phi} \quad (3.26)$$

$$k^2\Phi = 4\pi G a^2 \left[ \rho_{dm}\delta + 4\rho_r\Theta_{r,0} + \frac{3aH}{k}(\rho_{dm}\nu + 4\rho_r\Theta_{r,1}) \right]. \quad (3.27)$$

Se procede escribiendo las ecuaciones anteriores en términos de  $y \equiv \frac{a}{a_{eq}} = \frac{\rho_{dm}}{\rho_r}$  (siendo nuestro caso el límite  $y \gg 1$ ) y se reemplazan las derivadas respecto del tiempo por derivadas respecto a  $y$ , siendo  $\dot{y} = aHy$ . Además, en esta época del universo, no hay dependencia de las perturbaciones de radiación ya que estas son subdominantes, y tampoco hay dependencia de  $\frac{aH\nu}{k}$  porque los modos están ya dentro del horizonte, por lo que  $\frac{aH}{k} \ll 1$ . Además, como se puede desprestigiar la radiación  $4\pi G\rho_{dm} = (3/2)H_0^2\Omega_m a^{-3}$ , por lo que las ecuaciones anteriores 3.25, 3.26, 3.27 se pueden escribir respectivamente como:

$$\delta' + \frac{ik\nu}{aHy} = -3\Phi' \quad (3.28)$$

$$\nu' + \frac{\nu}{y} = \frac{ik\Phi}{aHy} \quad (3.29)$$

$$k^2\Phi = \frac{3}{2}H_0^2\frac{\Omega_m}{a}\delta. \quad (3.30)$$

Ahora el objetivo es pasar de las ecuaciones 3.28 y 3.29 a una ecuación de segundo orden para  $\delta$ , diferenciando 3.28:

$$\delta'' + ik\nu \left( \frac{d(aHy)^{-1}}{dy} - \frac{1}{aHy^2} \right) = \frac{3\Omega_m H_0^2}{2y^3 a^2 H^2 a_{eq}} \delta. \quad (3.31)$$

Utilizando la ecuación de continuidad y diviendo por  $a_{eq}^2$ :

$$\frac{d^2\delta}{da^2} + \left( \frac{d\ln(H)}{da} + \frac{3}{a} \right) \frac{d\delta}{da} - \frac{3\Omega_m H_0^2}{2a^5 H^2} \delta = 0 \quad (3.32)$$

A partir de ahora se utilizará  $a$  como variable. Para el límite  $y \gg 1$  todos los factores  $a_{eq}$  desaparecen.

Hay dos soluciones para esta ecuación, la primera es  $\delta \propto H$  y la segunda es la solución creciente. Esta última es la solución de interés, ya que implica que a medida que el tiempo avanza,  $\delta$  crece y, por ende, la sobredensidad aumenta, aumentando así la cantidad de masa para la formación de estructuras.

Para obtener la solución creciente, intentamos una solución de la forma  $u = \frac{\delta}{H}$ . La ecuación de evolución para  $u$  viene dada por:

$$\frac{d^2u}{da^2} + 3 \left[ \frac{d\ln(H)}{da} + \frac{1}{a} \right] \frac{du}{da} = 0. \quad (3.33)$$

Integrando la ecuación de primer orden para  $u'$ :

$$\frac{du}{da} \propto (aH)^{-3}. \quad (3.34)$$

Integrando de nuevo y teniendo en cuenta que  $u = \frac{\delta}{H} \dot{\delta} = uH$ :

$$D_1(a) \propto H(a) \int^a \frac{da'}{(a'H(a'))^3}. \quad (3.35)$$

Teniendo en cuenta que la relación entre el potencial y su valor justo después del régimen de la función de transferencia ( $a > a_{late}$ ) se define como:

$$\frac{\Phi(a)}{\Phi(a_{late})} \equiv \frac{D_1(a)}{a}, \quad (3.36)$$

para tiempos de dominación de materia,  $D_1$  debe ser igual a  $a$  y  $H = H_0 \Omega_m^{1/2} a^{-3/2}$ , por lo que el factor de crecimiento debe ser:

$$D_1(a) = \frac{5\Omega_m}{2} \frac{H(a)}{H_0} \int_0^a \frac{da'}{(a'H(a')/H_0)^3}, \quad (3.37)$$

donde:

$$H(a)^2 = H_0^2 \left[ \Omega_m a^{-3} + \Omega_r a^{-4} + \Omega_k a^{-2} + \Omega_v a^{-3(1+w)} \right], \quad (3.38)$$

siendo  $H_0$  es el parámetro de Hubble (función de Hubble en  $a = 1$ ) y  $\Omega_m$ ,  $\Omega_r$ ,  $\Omega_k$  y  $\Omega_v$  los parámetros cosmológicos, que para este trabajo, se utilizarán los valores determinados por la misión Planck [41]:  $\Omega_m = 0.3089$ ,  $\Omega_k = 0$ ,  $\Omega_v = 0.692$  y  $\Omega_r = 1 - \Omega_m - \Omega_v - \Omega_k$ .

### 3.7 Sesgo de los halos de materia oscura

Una forma de obtener información sobre el sesgo de los halos es a través de los espectros de materia de los halos y de la materia oscura subyacente [v.g. 42]:

$$\langle b(k; z)^2 \rangle = \left\langle \frac{P(k; z)_h(k)}{P(k; z)_{dm}(k)} \right\rangle, \quad (3.39)$$

donde  $P_{dm}(k; z)$  representa al espectro de potencia de la materia oscura y  $P(k; z)_h(k)$  es el espectro de potencias de los halos.

Sin embargo, es importante tener en cuenta que al calcular el sesgo, a partir de cierta escala van a aparecer irregularidades, ya que a medida que disminuimos la escala surgen no linealidades y demás efectos que hemos comentado anteriormente. Por lo tanto, para obtener un valor correcto del sesgo es necesario preguntarse hasta que valor del modo  $k$  se mantiene el sesgo constante. Para ello, se estudiará la tendencia del sesgo, con el fin de obtener el valor máximo de  $k$  – *bin* al que el sesgo es constante y, además, se analizará este sesgo para distintos valores del corrimiento al rojo con el fin de ver si este valor máximo depende de  $z$  o no.

Por lo tanto, para nuestro caso, el promedio del sesgo vendrá definido como:

$$\langle b(z_\alpha) \rangle = \frac{1}{N_r} \frac{1}{N_k} \sum_{i=0}^{N_r-1} \sum_{j=0}^J \left[ \frac{P_h^i(k_j; z_\alpha)}{P_{DM}^i(k_j; z_\alpha)} \right]^{\frac{1}{2}}, \quad (3.40)$$

donde  $N_r$  es el número de referencias que se utilizarán,  $N_k$  es el número de modos máximo hasta el que el bias se mantiene constante, es decir, el número de modos hasta  $k$ , siendo  $k < k_j \equiv k_{max}$ . El primer promedio se realiza sobre el número de realizaciones total y el segundo es la suma sobre modos  $k < k_j \equiv k_{max}$ .

Para hallar el sesgo en función de los modos, para un valor del corrimiento al rojo dado ( $z_l$ ), se calculará:

$$\langle b(k, z_l) \rangle = \frac{1}{N_k} \sum_{j=0}^J \left[ \frac{P_h^i(k_j; z_l)}{P_{DM}^i(k_j; z_l)} \right]^{\frac{1}{2}} \quad (3.41)$$

## 3.8 Comparación con modelo de sesgo de halos

### 3.8.1 Función de masa de halos de materia oscura

La función de masa de halos de materia oscura (fmh, a partir de ahora) mide la abundancia de objetos de halos de materia oscura en un volumen. Los cúmulos y las galaxias se encuentran dentro de estos halos de materia oscura y representan picos en el campo de densidad de la materia. De esta forma, muchas cantidades físicas observables están relacionadas con esta función, como por ejemplo, la función de masa de las galaxias. Por esta razón, la fmh es interesante ya que es el ingrediente fundamental para una modelización precisa de los observables.

Los propiedades de los picos en el campo de densidad pueden analizarse de forma rigurosa a través del formalismo de los picos, pero este formalismo tiene importantes limitaciones a la hora de hallar la fmh. Este formalismo de picos supone que la masa que colapsará para formar objetos no lineales, en nuestro caso, halos de materia oscura, puede identificarse en el campo de densidad inicial, aplicando, en primer lugar, un filtro de la escala adecuada que lo suavice y en segundo lugar, localizando después todos los picos por encima de cierto umbral. De esta forma, este formalismo sirve para calcular la abundancia y las propiedades de agrupación de los picos de densidad del campo de sobredensidad suavizado. Sin embargo, cuando se aplica a los halos de materia oscura, plantea una dificultad, la cual se denomina el problema de 'Cloud-In-Cloud' <sup>6</sup>, e ilustra que derivar una función de masa para objetos colapsados utilizando el formalismo de picos requiere tratar la estadística de los picos a medida que se varía  $R$ . [19]

Por lo tanto, lo que es necesario para predecir la función de masa de los objetos colapsados es:

1. un método para dividir el campo de densidad lineal  $\delta(x)$  a tiempo temprano en un conjunto de regiones disjuntas cada una de las cuales formará en algún momento posterior un único objeto colapsado
2. calcular las propiedades estadísticas de esta división

Dada las dificultades anteriormente descritas, es habitual utilizar un formalismo más sencillo aunque menos riguroso para hallar la función de masa de objetos colapsados, este es el formalismo de Press-Schechter, a través del cual el número de halos por unidad de masa por unidad de volumen se define como [43, 44]:

$$\frac{dn}{dM} = f(\sigma) \frac{\rho_m}{M} \frac{d \ln \sigma^{-1}}{dM} \quad (3.42)$$

donde  $\sigma$  es la raíz cuadrada de la varianza de un filtro top hat esférico que contiene una masa  $M$ ,  $\rho_m = \rho_{crit} \Omega_m$  es la densidad de masa media y  $f(\sigma)$  es la función de multiplicidad del halo.

### 3.8.2 Sesgo efectivo

El sesgo efectivo viene definido como [v.g. 45, 46, 47]:

$$b_{\text{eff}}(M_{\text{min}}, z) = \frac{\int_{M_{\text{min}}}^{\infty} dM n(M, z) b(M, z)}{\int_{M_{\text{min}}}^{\infty} dM n(M, z)}, \quad (3.43)$$

<sup>6</sup>Esquema de asignación de masa que trata cada partícula como una nube cúbica de densidad uniforme.

siendo  $M_{min}$  la masa mínima de los halos (para nuestro caso, masa mínima en la que los halos de materia oscura pueden albergar galaxias con líneas de emisión) que se considera para el cálculo del sesgo efectivo,  $M$  la masa del halo,  $n(M, z)$  es la fmh (número de halos por unidad de volumen y masa) y  $b(M, z)$  el sesgo (es decir, como se relaciona el sesgo de los halos con la densidad subyacente de materia), ambos en función de la masa  $M$  y el corrimiento al rojo  $z$ .

Esta ecuación se utiliza para calcular el sesgo efectivo de los halos de materia oscura que se encuentran por encima de un cierto corte de masa, y por lo tanto, es útil para estudiar la relación entre el sesgo de los halos y la masa de estos (por ejemplo, el estudio de cómo las galaxias en cúmulos masivos difieren en su sesgo de las galaxias en cúmulos menos masivos).

Para comparar el sesgo calculado a partir de los espectros, se utilizarán unos valores del sesgo efectivo de referencia, los cuales, se calcularán a través de los resultados obtenidos por Tinker [v.g. 43] para la función de masa de los halos.

## Capítulo 4

# Análisis del espectro de halos de materia oscura

*This section is divided into two parts. The first one focuses on BAM, and develops topics such as the problem it solves, how it works, the inputs it requires and its products. The second part contains the results obtained after analyzing the power spectrum of halos produced by BAM. In this part the mean value, variance, correlation matrix and higher moments such as kurtosis and skewness of the power spectra are measured, and histograms of these are generated. These measurements are made for different values of the redshift  $z$ , as well as for different scales. The results obtained are compared, as the case may be, with the theoretical results or with a Gaussian distribution with mean value and variance equal to that of the spectra, in order to obtain conclusions on the quality of the reproduction of the two-point halo statistics of this catalog of spectra. Finally, the bias of the halo spectra is analyzed as a function of redshift and as a function of  $k$  modes and this bias is also compared with reference values to contrast the results obtained.*

### 4.1 Preámbulo

Con el fin de generar modelos precisos de agrupación de galaxias y la obtención de las matrices de covarianza se han llevado a cabo diferentes métodos entre los cuales se encuentran las simulaciones de  $N$ -cuerpos, cuyo principal desventaja es el coste elevado de tiempo y/o almacenamiento. Con el fin de disminuir dichos costes se han desarrollado métodos aproximados y métodos calibrados. En este último conjunto se encuentran métodos como PATCHY [40], EzMOCKS [48] y BAM [37].

En este trabajo vamos a analizar la estadística de dos puntos en el espacio de Fourier a partir de ensambles estadísticos de distribuciones de halos obtenidos con el método BAM. En primer lugar se presenta el método [v.g. 37, 49, 1], cómo funciona y qué entradas requiere, en segundo lugar se describe la simulación de referencia utilizada por BAM y por último se presentan sus resultados [v.g. 2, 50, 51].

### 4.2 BAM: Método de asignación de sesgo

- **Entradas del método:**

- una realización de las condiciones iniciales (CI a partir de ahora) de una simulación de  $N$ -cuerpos de alta resolución

- el catálogo de halos correspondiente a esa realización particular de la simulación de referencia obtenido a partir de un solucionador de gravedad aproximado
- **Cómo funciona el método:** BAM se centra en la idea de mapear la distribución del número de halos en una malla sobre un campo de densidad de materia oscura a través de un esquema de sesgo extraído de la distribución de halos de materia oscura de una simulación de N-cuerpos y un CDM obtenido a partir de la evolución de condiciones iniciales (CI) cosmológicas de baja resolución con métodos aproximados de evolución gravitacional. Dicho mapeo puede aplicarse a cualquier configuración de CI conservando los mismos parámetros cosmológicos y la configuración numérica. En adición, el enfoque de BAM incluye un procedimiento de muestreo iterativo que tiene en cuenta las correlaciones cruzadas y otras dependencias que van a mejorar los resultados a pequeña escala.

De esta forma, los pasos a seguir por BAM se pueden resumir en:

1. **Etapa I:** Calibración: se trata de un proceso de aprendizaje en el que se calibran el halo-bias y el BAM-kernel utilizando los estadísticos de dos puntos de la referencia como función de costo.
2. **Etapa II:** Producción simulada de halo: utilizando el sesgo del halo y a través del muestreo de campos de densidad de materia oscura independientes, se generan campos de conteo de número de halo independientes.
3. **Etapa III:** Coordenadas y propiedades del espacio de fase: asignación de posición, velocidades y propiedades intrínsecas de los halos de materia oscura
4. **Etapa IV:** Catálogos de galaxias: En adición, como último paso, una vez generados los catálogos de halos y haberles asignado coordenadas y propiedades, BAM implementa un modelo de distribución de la ocupación del halo (DOH, a partir de ahora) para poblar halos de materia oscura con galaxias.

- **Simulación de referencia:**

Con el fin de evaluar el rendimiento de BAM se ha utilizado la simulación UNIT [38]. Se trata de un conjunto de dos realizaciones de par fijo en un volumen de  $1(Gpch^{-1})^3$ , constituido por  $4096^3$  partículas de materia oscura y una resolución de masa de  $1.2 \times 10^{11} M_{\odot} h^{-1}$  y catálogos de halos con una masa (virial) mínima de  $3,6 \times 10^{11} M_{\odot} h^{-1}$ . En concreto, BAM utilizó un número de instantáneas con valores de redshift entre  $0 \leq z \leq 3$ , con el fin de generar el mismo número de núcleos y halo-bias tanto para la realización normal y la de fase invertida. Los parámetros cosmológicos utilizados por esta simulación fueron los valores determinados por la misión Planck [41].

- **Productos:**

En la figura 4.1 se pueden observar cortes a través de los diferentes campos de densidad de referencia implicados en la etapa de calibración de BAM, los cuales se han seleccionado al azar. El panel inferior muestra los recuentos del número de halos en una malla

reconstruida utilizando tres modelos de halo-bias (TkWEB <sup>1</sup>, IkWEB <sup>2</sup>, TIWEB <sup>3</sup>) [2]. En la columna de la derecha se representan, en la primera fila, el campo de densidad de galaxias de referencia y en la segunda, el campo de densidad de galaxias de BAM, el cual se ha contruido poblando los catálogos de halo con un modelo de distribución de la ocupación del halo.

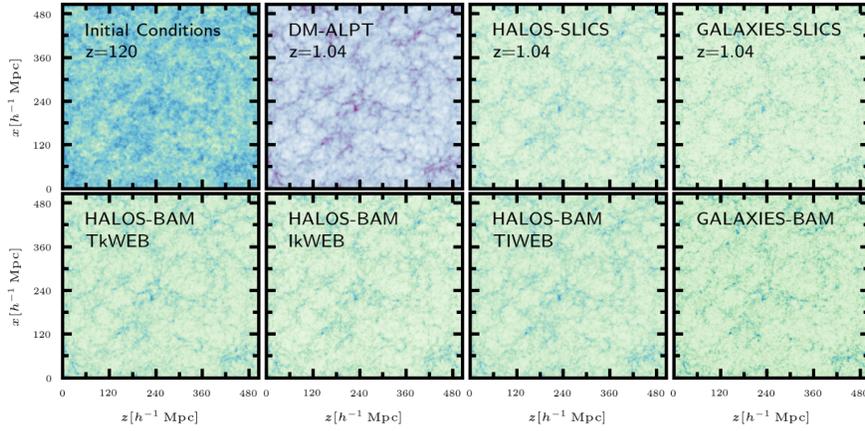


Figure 4.1: Representación de las rebanas de  $25 \text{ Mpc } h^{-1}$  de grosor a través de los diferentes campos de densidad implicados en la calibración de BAM. La segunda fila representa la reconstrucción del campo de densidad numérica de halos para diferentes modelos de halo-bias, mientras que la columna de la derecha muestra en la gráfica superior el campo de densidad de galaxias de referencia y en la inferior el campo de densidad de galaxias de BAM. [Gráfica tomada de 2]

### 4.3 Análisis del espectro de potencia de halos en BAM

Para analizar la calidad de los espectros de potencia de BAM se generó un catálogo de 990 realizaciones a diferentes valores de corrimiento al rojo, desde  $z = 0[h^{-1}\text{Mpc}]$  a  $z = 3.037[h^{-1}\text{Mpc}]$ , y para un rango de bins de  $k$  de 128.

#### 4.3.1 Valores medios

En la figura 4.2 observamos el espectro de potencias medio producido por BAM con su correspondiente varianza, obtenido para distintos valores del corrimiento al rojo. Este espectro va a ser el objeto de estudio de nuestro trabajo y a lo largo de las siguientes secciones se analizará su calidad a la hora de reproducir la estadística de 2 puntos a gran escala.

<sup>1</sup>TkWEB: Modelo halo-bias que utiliza la densidad local, los tipos de redes cósmicas y la masa de las regiones en colapso. Este tipo de modelo es el que mejor funciona ya que tiene en cuenta el sesgo de ensamblaje de halos y galaxias, el cual incluye tanto la agrupación de halos en función de sus propiedades intrínsecas como en función de su entorno. Además, al tener en cuenta la masa de las regiones en colapso, este tipo de modelos aporta el hecho de incluir el corto alcance del sesgo no local centrado en regiones con un estado dinámico distinto.

<sup>2</sup>IkWEB: Modelo halo-bias que utiliza la masa de las regiones en colapso y los invariantes del campo de marea.

<sup>3</sup>TIWEB: Modelo halo-bias que utiliza la clasificación de la red cósmica y un invariante del campo de marea.

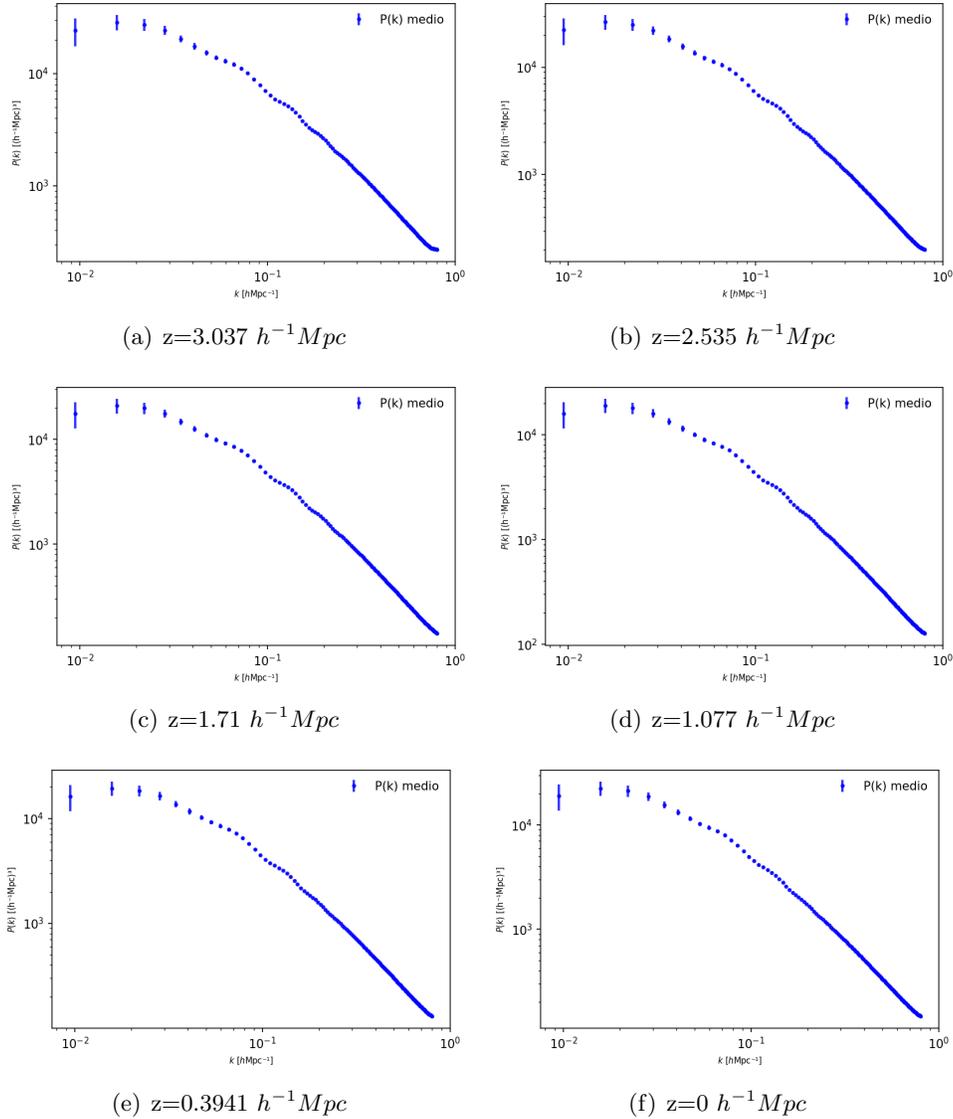


Figure 4.2: Resultados de la medida de la media del espectro de potencia con sus correspondientes barras de error para distintos valores del corrimiento al rojo  $z$

En la figura 4.3 se representa a la izquierda el valor medio de los espectros de los halos de materia producidos por BAM, y a su derecha el espectro de potencias lineal calculado a través de la herramienta `nbodykit`, utilizando los parámetros cosmológicos de la simulación UNIT [38] y evaluando la función de transferencia de densidad total al desplazamiento al rojo especificado, que en este caso, para ambos espectros es  $z = 0 [h^{-1} Mpc]$ . Vemos que para los valores de la escala a la que representamos el espectro medio, la tendencia de ambos espectros es muy similar.

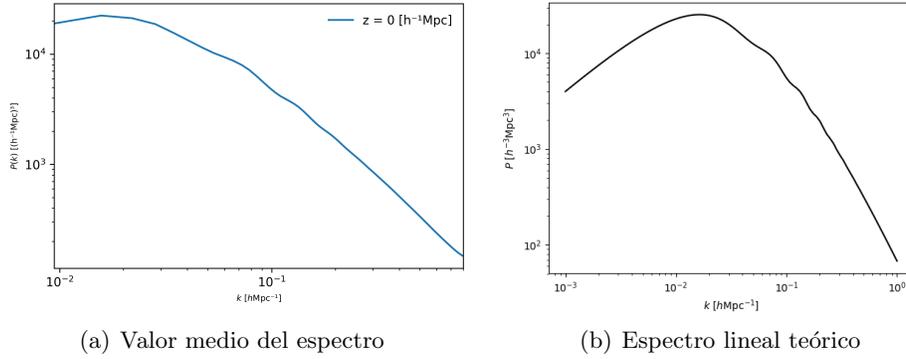


Figure 4.3: A la izquierda se muestra la media de los espectros de potencias de BAM, mientras que a la derecha se observa espectro de potencias lineal calculado para los parámetros cosmológicos de la misión de Planck [41], ambos calculados en  $z = 0 [h^{-1} Mpc]$

### 4.3.2 Varianza

La variancia es la medida de dispersión con respecto a la media del ensamble estadístico:

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{N} \quad (4.1)$$

Considerando ahora una distribución gaussiana, esta se puede definir completamente a través de dos parámetros: su *media*, que es el centro de la distribución y su *varianza*, que al ser una medida de dispersión con respecto de la media, determina el ancho de la distribución gaussiana.

En la figura 4.4 se puede observar la variancia calculada para distintos valores del corrimiento al rojo.

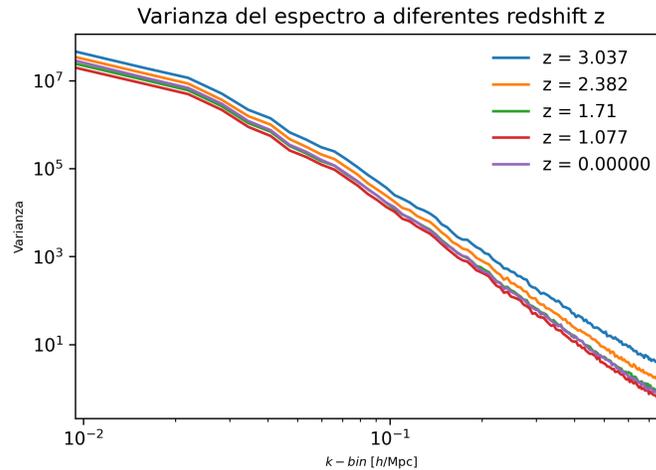


Figure 4.4: Varianza del espectro de potencia de BAM a diferentes redshift  $z$  calculada hasta  $k = 0.2$

Una forma interesante para deducir si la variancia de los espectros de potencia simulados es coherente con la teoría es compararla con los valores obtenidos para la variancia teórica que

se pueden obtener a través de la ecuación 4.2[v.g. 33], a partir de los valores del espectro de potencias  $P(k)$ , el volumen de la cáscara esférica  $V_k$  y la densidad numérica media  $\bar{n}$ :

$$\sigma_P^2(k) = \frac{2}{VV_k} \left[ P(k) + \frac{1}{\bar{n}} \right]^2, \quad (4.2)$$

donde  $V = L^3$ , siendo  $L$  la longitud de la caja, que en nuestro caso, es  $L = 1000 \text{ Mpc}/h$  y  $V_k \approx 4\pi k^2 \Delta k / (2\pi)^3$ . Para hallar la varianza teórica, en este caso, se ha obviado el segundo término  $\frac{1}{\bar{n}}$ , que corresponde al shot noise, por lo que finalmente, esta varianza, que aparece representada como una línea azul discontinua en las distintas gráficas de la figura 4.5, se ha obtenido a partir de la ecuación:

$$\sigma_P^2(k) = \frac{2}{VV_k} [P(k)]^2. \quad (4.3)$$

La expresión 4.2, cuyo derivación se encuentra detallada en [34, 33], es una medida de la variabilidad del espectro de potencias en diferentes escalas. Esta ecuación proporciona una herramienta teórica importante para comprender la distribución de la materia en el universo y su evolución. Además, dado que solo tenemos un solo Universo, es decir, una sola referencia, esta varianza se utiliza para analizar datos observacionales y compararlos con las predicciones teóricas de la teoría de perturbaciones lineales, permitiendo así mejorar nuestra comprensión del universo. En particular, para este trabajo es de gran utilidad ya que permite poner a prueba los espectros del catálogo contruido.

Esta expresión se deriva a partir de la teoría de perturbaciones lineales, por lo que el supuesto de homogeneidad e isotropía debe cumplirse, si no la ecuación deja de ser precisa. Además, la ecuación no tiene en cuenta el sesgo, por lo que esto puede introducir distorsiones en la distribución de objetos y por lo tanto afectar al espectro de potencias. Por último, esta ecuación no incluye todas las correcciones cosmológicas (presencia de FCM o energía oscura) para describir el espectro de potencias con alta precisión.

Debido a la limitación de esta ecuación al régimen lineal, en este trabajo, se van a tener en cuenta únicamente los primeros modos (hasta  $k = 0.2$ ), con el fin de asegurar la precisión de los resultados obtenidos.

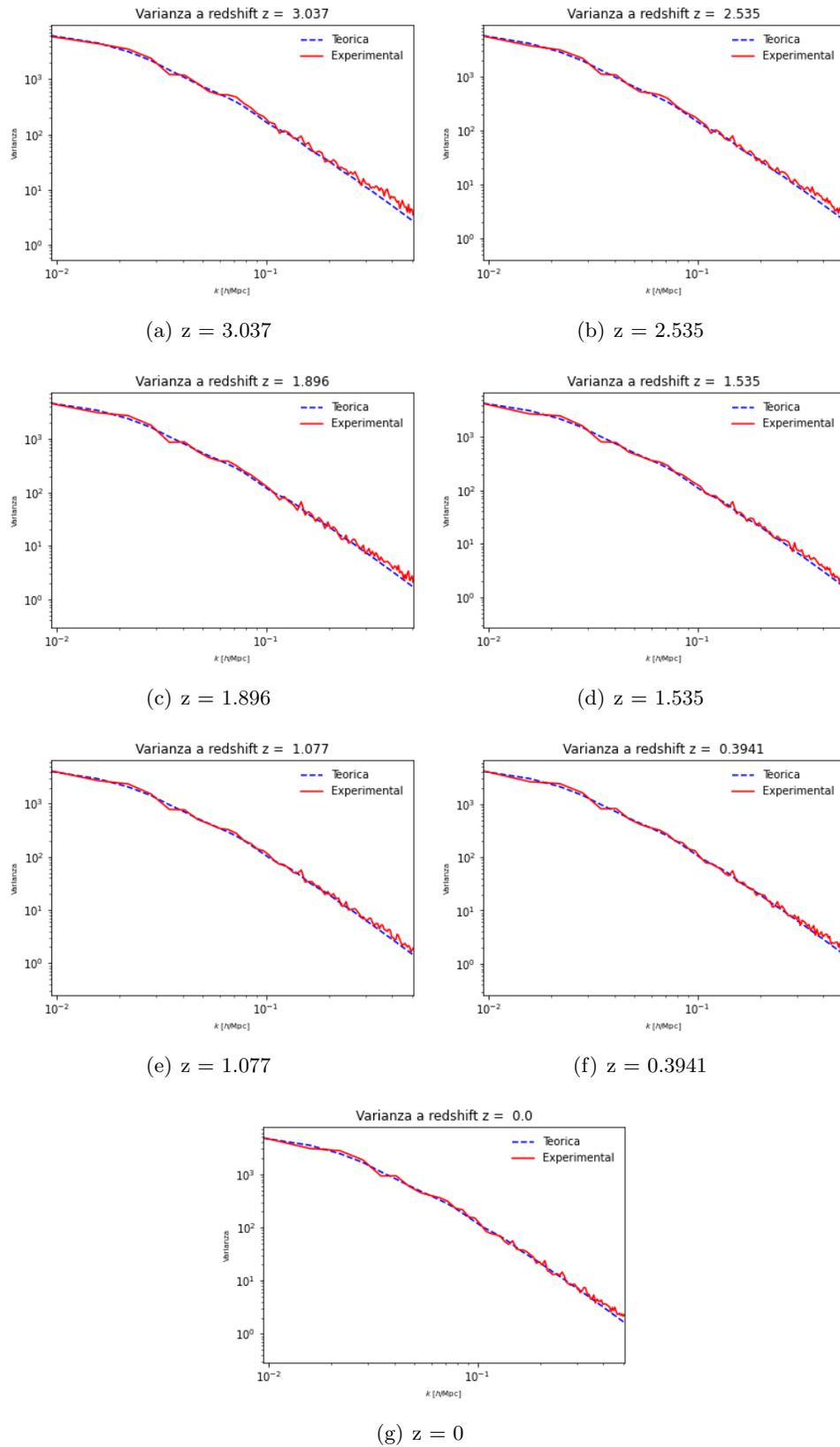


Figure 4.5: Varianza teórica vs varianza medida para distintos valores del corrimiento al rojo  $z$  calculadas hasta  $k = 0.2$

En la figura 4.5 observamos que a gran escala ( $k$  pequeños) la varianza del espectro de los halos se ajusta muy bien a la varianza teórica, sin embargo, a medida que nos vamos acercando a escalas más pequeñas ( $k$  mayores) empiezan a aparecer irregularidades y la varianza calculada del espectro comienza a desviarse de la curva teórica. Esto se debe a que, como se ha mencionado anteriormente, a medida que nos acercamos a escalas más pequeñas la no linealidades comienzan a surgir y por lo tanto ya no podemos analizar la distribución espacial de halos con la estadística de 2 puntos.

### 4.3.3 Matriz de correlación

Podemos definir la correlación como una medida estadística que expresa el nivel de relación lineal entre dos variables. De este modo, la matriz de correlación se define como la matriz cuadrada y simétrica que tiene la siguiente forma:

$$R = \begin{bmatrix} 1 & r_{12} & \cdots & r_{1k} \\ r_{21} & 1 & \cdots & r_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ r_{k1} & r_{k2} & \cdots & 1 \end{bmatrix} \quad (4.4)$$

donde se puede ver los elementos de la diagonal son unos y los no diagonales  $r_{ij}$  son los coeficientes de correlación entre las variables.

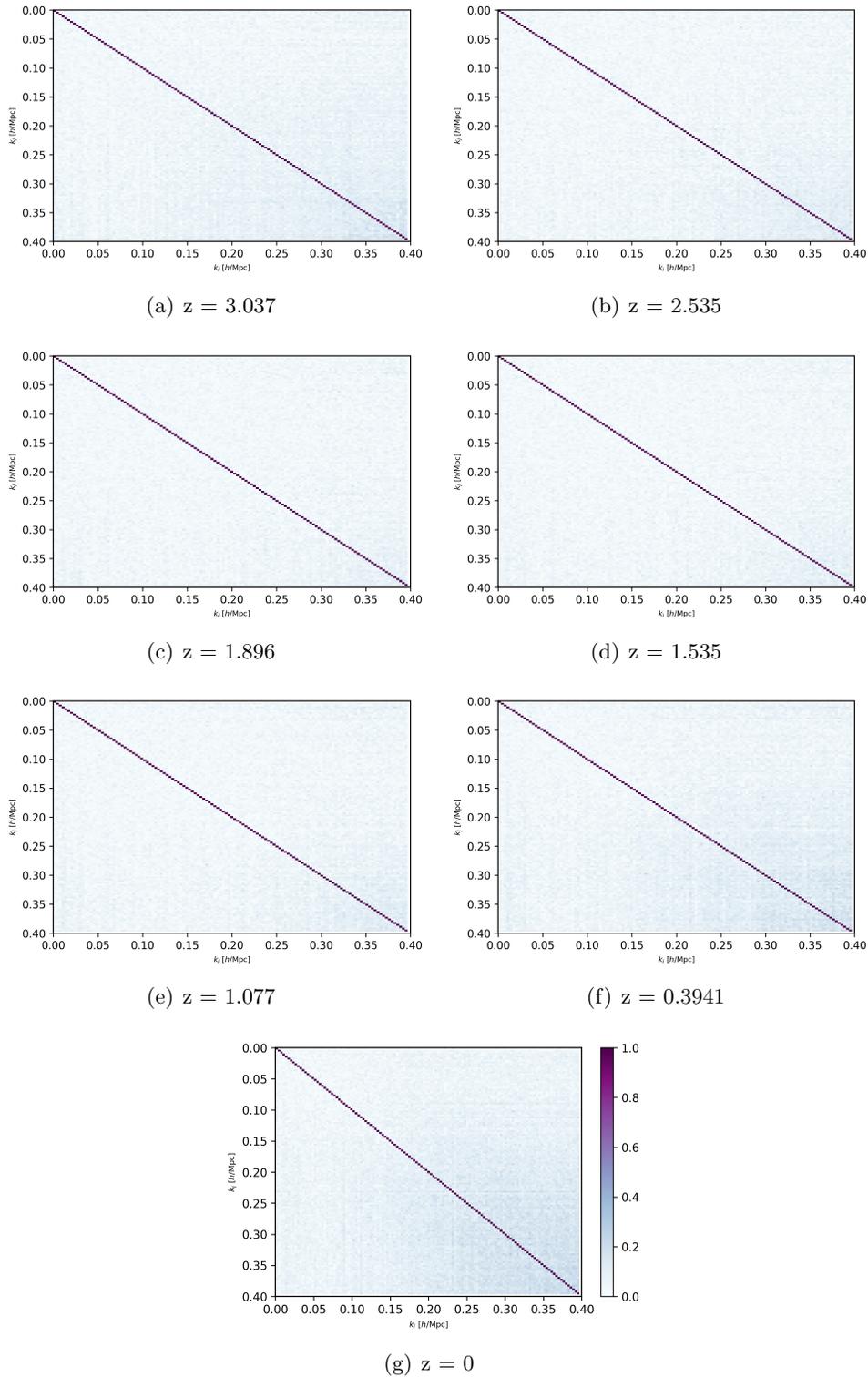


Figure 4.6: Matriz de correlación de los espectros de potencia de los halos de BAM para distintos valores del corrimiento al rojo  $z$  calculadas en función de los modos hasta el valor  $k = 0.4$ . poner barra de color y tamaños bien el eje y debe ir creciendo

En la figura 4.6 se puede observar los coeficientes de correlación de los espectros de potencia de los halos obtenido por BAM en función de los modos de  $k$  en los que se calcula. En este caso, se ha representado solo hasta el valor de  $k = 0.4$ , ya que a partir de  $k = 0.2$  comienzan

a surgir no linealidades, y esto se hace más visible a medida que nos acercamos a modos  $k$  mayores, traduciéndose esto en mayor acople entre modos, como se puede observar en estas figuras. La escala de colores es tal que cuando la acoplación entre modos es alta se representa de color magenta, mientras que cuando la acoplación entre modos es nula su representación sería de color blanco. Para estas gráficas la línea magenta representa la diagonal de la matriz de correlación, es decir 1, que sería el valor máximo de acople entre modos, mientras que los puntos de alrededor representan los coeficientes de correlación entre modos. En este caso, se ha representado solo hasta el valor de  $k = 0.4$ , ya que a partir de  $k = 0.2$  comienzan a surgir no linealidades, y esto se hace más visible a medida que nos acercamos a modos  $k$  mayores, traduciéndose esto en mayor acople entre modos, como se puede observar en estas figuras. De esto se concluye que hay una cierta no linealidad en la distribución del espectro de potencia, y que dicha no linealidad aumenta a medida que nos acercamos a escalas más pequeñas (modo  $k$  más grande), ya que el color a número de  $k$  grande es considerablemente más azul que a modos de  $k$  pequeño, en los cuales la no-linealidad es imperceptible. Estos resultados son coherentes con la teoría, ya que como se ha mencionado anteriormente, el Principio Cosmológico predice una distribución de densidades lineales a grandes escalas, mientras que a pequeñas escalas el régimen es no lineal a consecuencia de las subestructuras.

#### 4.3.4 Momentos más altos

Una distribución gaussiana se caracteriza por poseer los momentos más altos (a partir del tercer momento) nulos. Si una distribución de densidad tiene momentos más altos, como la asimetría (skewness), que es el tercer momento, o la curtosis que es el cuarto momento, el régimen ya no es completamente lineal y no podemos atribuirle una distribución gaussiana. Sin embargo, el valor de estos momentos puede ser muy pequeño o casi nulo, por lo que para analizar esto, hemos realizado el siguiente estudio.

Para este estudio, se han realizado distribuciones gaussianas con la misma media y varianza de nuestro espectro de potencias y con el mismo número de bins que para el histograma de espectros que veremos en la siguiente sección 4.3.5. Una vez obtenida esta distribución gaussiana, se ha calculado su curtosis y skewness de la misma forma que para el espectro de potencia.

Esta comparación se puede ver en la figura 4.7 y 4.8. Se observa que para la curtosis(4.7) se obtienen buenos resultados, ya que para modos pequeños, la curtosis del espectro es muy similar a la de la distribución gaussiana y además sus valores son  $\sim 0$ , por lo que podemos concluir que se está comportando como una distribución gaussiana, ajustándose así a lo que se esperaba obtener en el régimen lineal. También se puede observar que a medida que nos acercamos a  $k$  mayores (a partir de  $k > 0.2$ ) la curtosis del espectro comienza a desviarse cada vez más de 0, confirmando así la no linealidad a escalas más pequeñas. Sin embargo, los resultados para la skewness del espectro (4.8) no son los esperados, ya que precisamente, a  $k$  más pequeños (hasta  $k = 0.05$ ), donde se supone régimen lineal, se están obteniendo resultados que difieren de los resultados de la skewness para la distribución gaussiana y además estos se alejan del valor cercano a 0 esperado. Pero a  $k > 0.05$  se vuelven a obtener los resultados esperados y se confirma que el espectro se comporta como una distribución gaussiana. El resultado para modos de  $k < 0.5$  puede significar que nuestro espectro no es lineal a esta escala, lo cual difiere con la teoría del régimen lineal anteriormente establecida y por lo tanto, es posible que los espectros del catálogo de halos no sean del todo precisos en este rango de modos. Sin embargo, en modos entre  $0.5 < k < 0.2$  los resultados se ajustan a la teoría del régimen lineal a gran escala.

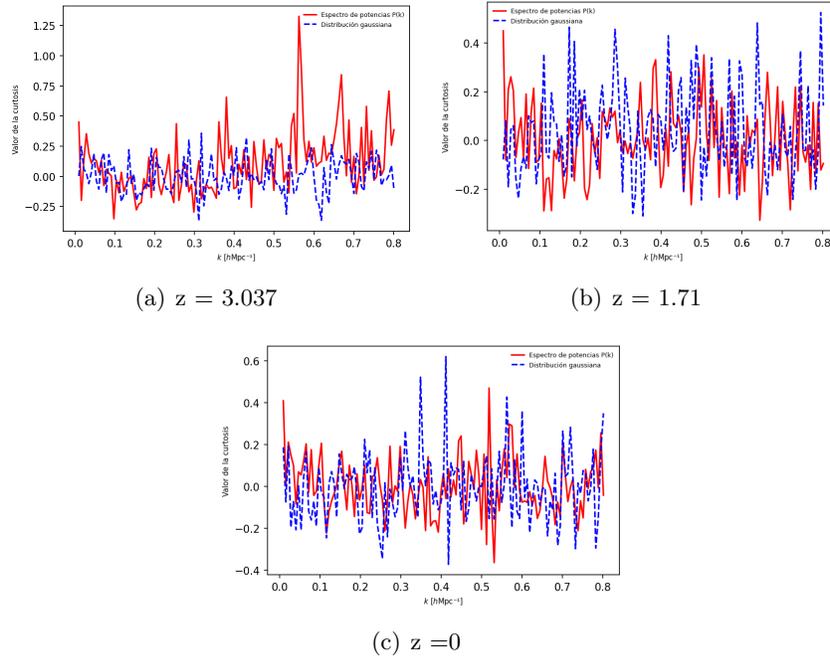


Figure 4.7: Kurtosis de los Espectros de Potencia vs Kurtosis Gaussiana para diferentes valores del corrimiento al rojo  $z$

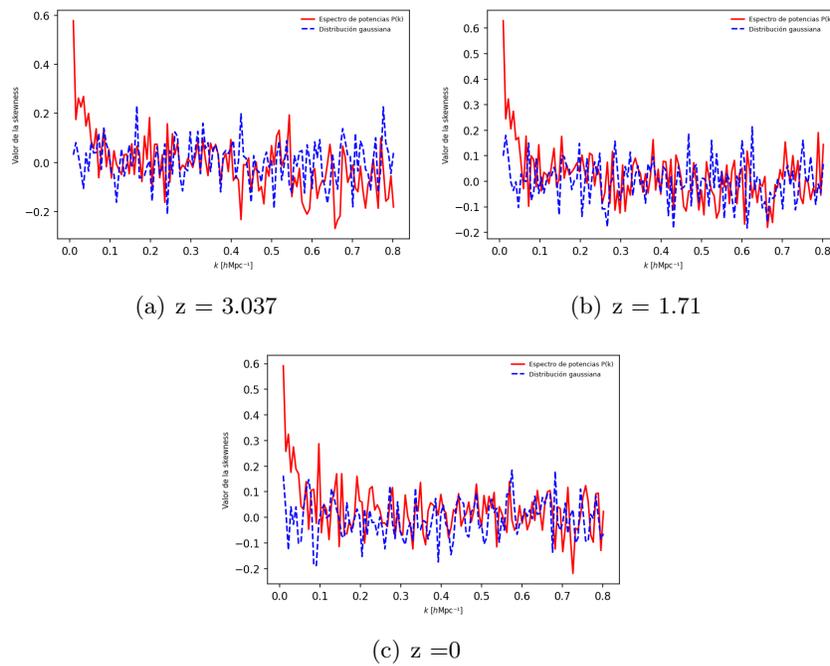


Figure 4.8: Skewness de los Espectros de Potencia vs Skewness Gaussiana para diferentes valores del corrimiento al rojo  $z$

### 4.3.5 Histograma de Espectros de Potencia $P(k)$ vs Histograma de Gaussianas

Para determinar si la distribución de probabilidad del espectro de potencias sigue una distribución gaussiana o no, dado que, a simple vista, observando los histogramas no podemos determinar con exactitud si esto es así, hemos realizado una prueba no paramétrica para determinar la bondad de ajuste entre una distribución gaussiana y la distribución del espectro de potencias de los halos, entre sí, en concreto, realizamos el test de Kolmogórov-Smirnov para 2 muestras (test K-S, a partir de ahora). La hipótesis nula de esta prueba es : ambas muestras provienen de la misma distribución. Este test nos da dos valores,  $D_\alpha$  y  $p$ value, los cuales vienen indicados en cada figura para cada valor del modo  $k$  y del redshift  $z$ . Para poder interpretar estos resultados vamos a utilizar el valor crítico  $D_\alpha$ , que es el valor absoluto de la distancia entre dos conjuntos de datos y donde  $\alpha$  representa el nivel de significancia, el cual, en general, se suele suponer igual a 0.05, lo cual indica asumir un riesgo del 5% de que las muestras no provienen de la misma distribución, cuando las muestras sí provienen de la misma distribución. Para calcular  $D_\alpha$ , se utilizará la siguiente ecuación:

$$D_\alpha = c(\alpha) \sqrt{\frac{n_1 + n_2}{n_1 n_2}} \quad (4.5)$$

siendo  $n_1$  y  $n_2$  el número de muestras de la distribución de los espectros de potencia y la distribución gaussiana respectivamente, cuyos valores son  $n_1 = 25$  y  $n_2 = 25$ . El coeficiente  $c(\alpha)$  tiene un valor asignado que para  $\alpha = 0.05$  es  $c(\alpha) = 1.36$ . De forma que sustituyendo todos estos valores obtenemos un resultado para el valor crítico de  $D_\alpha = 0.3847$ . El siguiente paso es comparar los valores de  $D_\alpha$  obtenidos, ya que si el valor  $D_\alpha$  es menor que el valor crítico de  $D_\alpha$ , podremos concluir que ambas muestras provienen de la misma distribución, pero si, por lo contrario,  $D_\alpha$  es mayor que el valor crítico  $D_\alpha$ , las dos muestras provendrían de distribuciones diferentes. Comparando ahora los valores que hemos obtenido en la prueba para  $D_\alpha$ , vemos que para todos los valores de  $k$  y del redshift  $z$ ,  $D_\alpha$  es menor que el valor crítico de  $D_\alpha$ , y por lo tanto, ambas muestras provienen de la misma distribución, y dado que la segunda muestra es una distribución gaussiana, podemos concluir que la primera, la distribución del espectro de potencias, proviene de una distribución gaussiana. Sin embargo, en el apartado anterior, para la skewness 4.8 los resultados para modos pequeños no han dado tan buenos resultados, por lo que seguramente el test K-S esté obviando algún detalle y en realidad, los espectros no se comporten como una distribución gaussiana o por lo menos no con tanta precisión.

$z = 3.037$ , variando  $k$

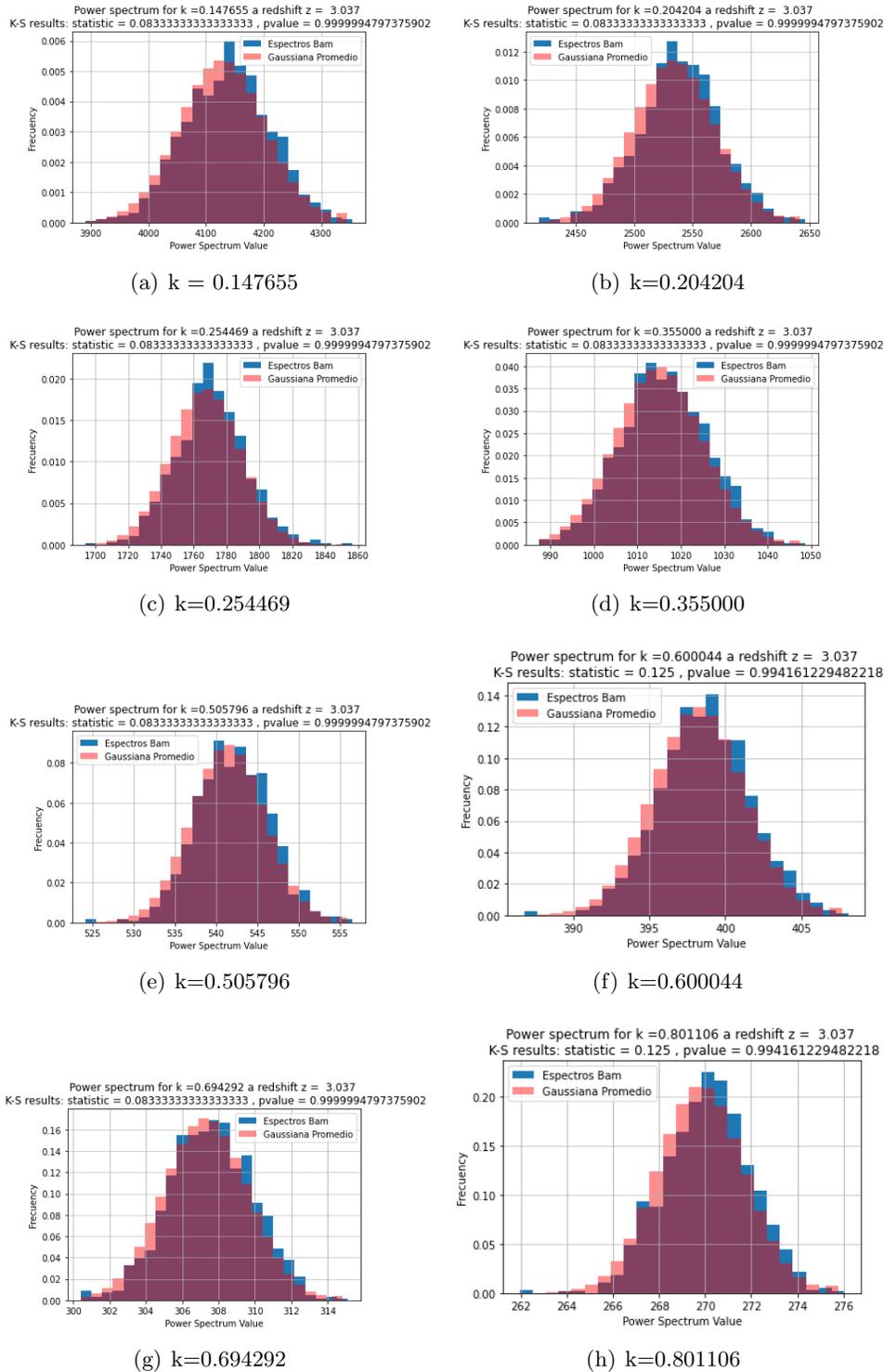


Figure 4.9: Histograma de espectros de potencias de BAM vs histogramas del espectro de potencias de una distribución gaussiana con misma media y varianza que la medida en el espectro de BAM para un valor fijo del corrimiento al rojo  $z = 3.037$  y distintos valores de  $k$

$k = 0.21$ , variando  $z$

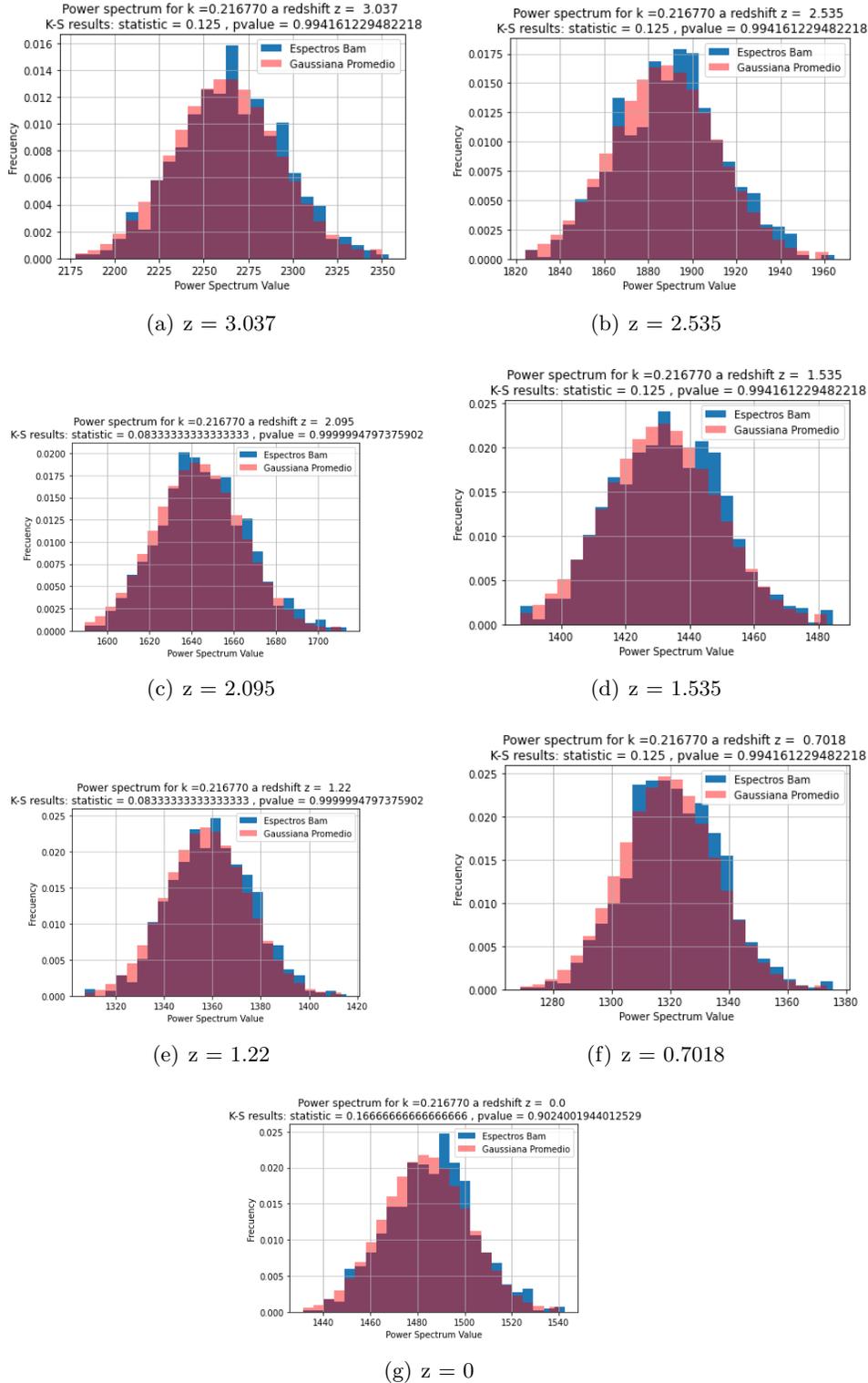


Figure 4.10: Histograma de espectros de potencias de BAM vs histogramas del espectro de potencias de una distribución gaussiana con misma media y varianza que la medida en el espectro de BAM para un valor fijo de  $k = 0.216770$  y distintos valores del corrimiento al rojo  $z$

## 4.4 Sesgo de los halos de materia oscura

En esta sección se analiza el sesgo de los halos de la simulación de BAM. Para ello, se calcula el sesgo promedio a partir de la ecuación 3.40.

El método que se aplicó para deducir hasta que valor de  $k$  es el sesgo constante fue hallar un valor medio a partir de los tres primeros valores del sesgo en función de  $k$ , obtenidos a partir de la ecuación 3.41, y generar una gráfica (ver figura 4.11) del sesgo con barras de error en función de  $k$ , de modo que se pueden considerar valores constante aquellos cuyas barras de error caigan dentro del valor constante del sesgo (promedio de los 3 primeros valores), que viene representado como una línea roja en la siguiente figura:

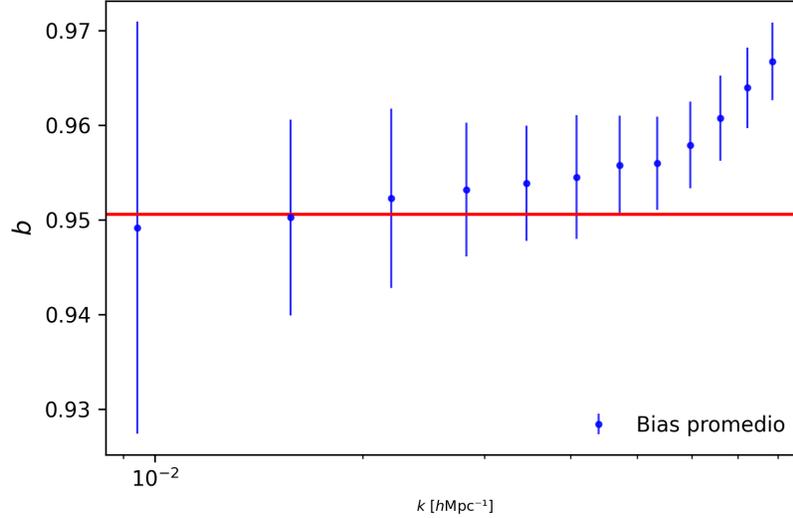


Figure 4.11: Sesgo para  $z = 0$  en función de  $k$ . Este sesgo se mantiene constante hasta  $k = 0.05hMpc^{-1}$

Como se puede observar en la figura 4.11 el sesgo se mantiene constante hasta  $k = 0.05h/Mpc$  y empieza a crecer a medida que nos acercamos a escalas más pequeñas. Este estudio se realizó para distintos valores de  $z$  con el fin de ver si había alguna dependencia del valor máximo del sesgo con el corrimiento al rojo. Finalmente se concluyó que, aunque dicho valor máximo varía un cierto rango, no aparece ninguna tendencia con  $z$ , por lo que se tomó  $k = 0.05h/Mpc$  para todos los valores del corrimiento al rojo  $z$ .

### 4.4.1 Sesgo en función de la escala:

A continuación, podemos observar el sesgo en función de los modos de  $k$  para diferentes valores del corrimiento al rojo. A medida que avanzamos hacia escalas más pequeñas, es decir, valores de  $k$  más grandes, se observa una disminución en la varianza del sesgo (las barras de error son más pequeñas). Estos resultados tienen sentido dado que si se compara con los resultados obtenidos para los espectros de potencia de los halos de materia oscura en la figura 4.2, estos espectros se comportan de la misma forma que el sesgo en términos de su varianza. Esto se debe a que hay mayor desviación respecto a la media en los valores de los espectros de potencia a gran escala que a escalas inferiores y este comportamiento se refleja en el sesgo debido a que este está relacionado con el espectro de los halos y el de la materia oscura, y este último, veremos más adelante que se comporta de forma muy similar al espectro de los halos. Esta disminución de la varianza a medida que disminuimos la escala puede verse de forma más clara en la figura 4.4, donde la varianza de los espectros de los halos es mayor a

gran escala que a escalas menores. En la práctica, esto significa que la varianza del sesgo de los halos de materia oscura es más grande a mayor escala debido a una mayor variación en la relación entre la distribución espacial de los halos y la distribución de la materia subyacente en diferentes regiones del universo.

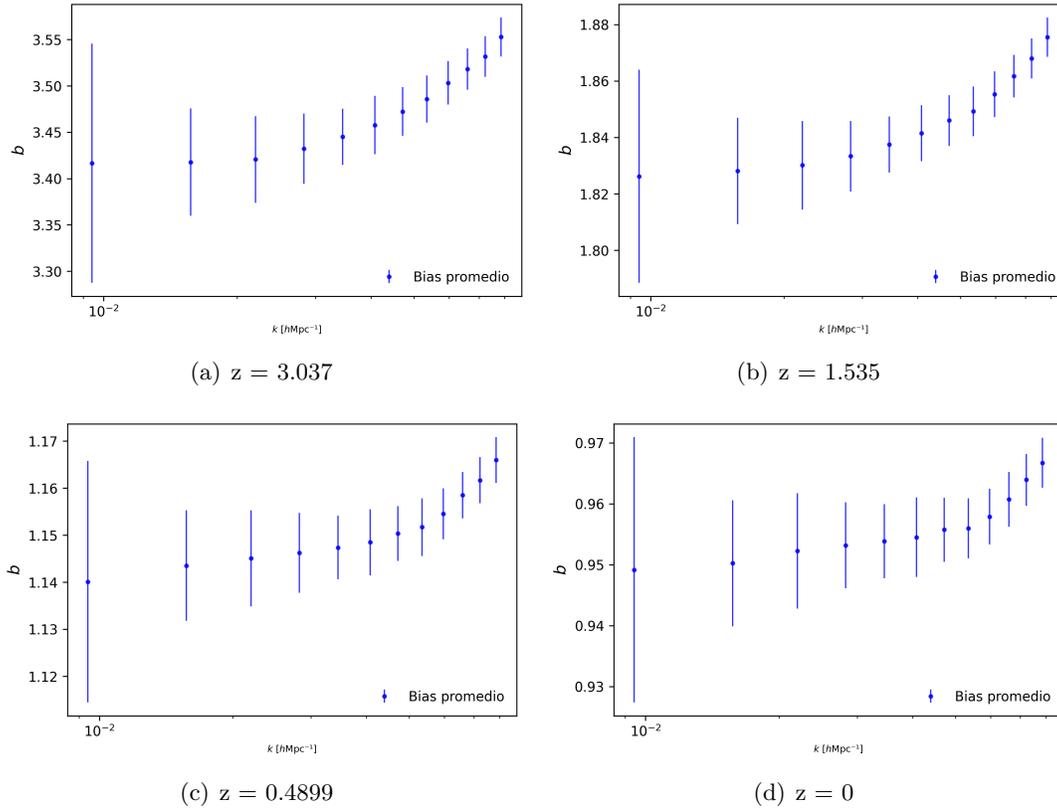


Figure 4.12: Sesgo de los espectros de los halos de materia oscura en función de la escala hasta el valor máximo de  $k = 0.05h/Mpc$ .

#### 4.4.2 Sesgo en función del corrimiento al rojo y la escala

En las siguientes figuras observamos el sesgo en función del corrimiento al rojo y de los modos  $k$ , para 3 valores máximos de  $k = [0.02, 0.05, 0.07][hMpc^{-1}]$  respectivamente. En las figuras también aparece representada una línea roja que son los valores de referencia para el sesgo efectivo, obtenidos a través de la fórmula de ajuste de la función de masa del halo desarrollada por Tinker [v.g. 43]. Es característico el tamaño relativamente pequeño de las barras de error y el hecho de que los valores del sesgo calculado no se encuentren más cerca de los valores del sesgo de referencia, si no un 10% por encima.

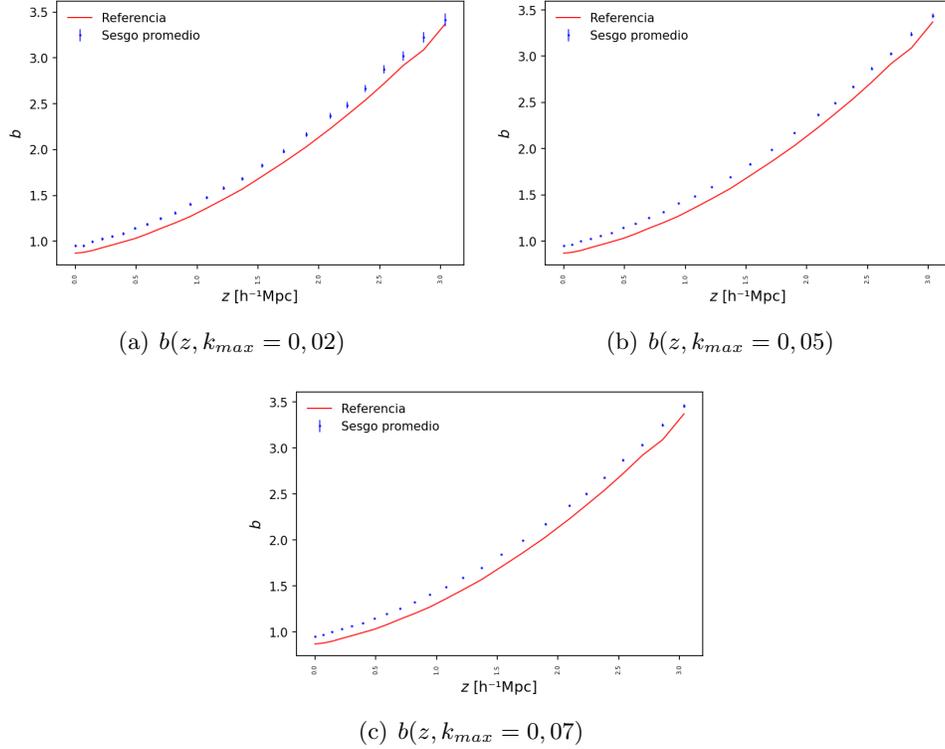


Figure 4.13: Sesgo en función del corrimiento al rojo  $z$  y la escala a distintos valores de  $k_{max}$ . La línea roja representa los valores del redshift efectivo de referencia obtenidos a través de la fórmula de ajuste para la función de masa desarrollada por Tinker [v.g. 43]

Sin embargo, se puede deducir que una explicación posible para las barras de error pequeñas, es que se deban a la gran similitud entre cada par de realización del espectro de la materia oscura y de los halos de materia oscura. No obstante, en la práctica, estas realizaciones no deberían de ser tan similares ya que en la formación de halos a partir de campos de densidad de materia oscura surgen diferencias entre los espectros de potencia de dichos halos y los espectros de la materia oscura original que los ha generado. Estas diferencias aportan estocasticidad a los espectros de los halos de materia oscura. Sin embargo, a la hora de calcular el sesgo representado en la figura 4.13, no estamos utilizando el campo de densidad de materia oscura original, sino uno con una resolución mucho menor, por lo que, la única estocasticidad a la que están sujetos nuestros espectros es a la varianza cósmica, la cual al hallar el sesgo, y ser este un cociente entre ambos espectros, se cancela, y de esta forma, se elimina en parte la estocasticidad asociada a las CI. Esto da lugar a un error estadístico casi nulo, lo cual no es consistente con la realidad. Por último, en cuanto a las barras de error respecta, para  $k = 0.02$  4.13(a), se observan unas barras de error más grandes. Esto se puede deber a que al reducir el valor máximo de  $k$  no tenemos tanto espectros para promediar y consecuentemente, al disminuir las muestras, el error estadístico aumenta. No obstante, esta mayor varianza también se puede deber a que nos encontramos a escalas mayores, donde la varianza del sesgo, como se ha observado en 4.12, es mayor.

Por otro lado, los valores obtenidos del sesgo  $b(z, k)$  no son consistentes con los valores de referencia, ya que los primeros son un 10% mayores que los segundos. Un aumento en el corte de masa, puede traducirse en una disminución del sesgo del halo. Aumentar el corte de masa conlleva un criterio de elección de halos a considerar más riguroso, por lo que se considerarían un menor número de halos de materia oscura, lo cual resultaría en un menor sesgo de los halos respecto a la materia oscura subyacente. Por lo tanto, estos resultados podrían modificar y

mejorarse estableciendo un corte de masa mayor a la hora de generar los catálogos de halos, es decir, cambiando el modelo. Sin embargo, la causa del error podría tener otra fuente, y es que el método no esté generando los espectros de la forma esperada y por lo tanto, los espectros del catálogo generado estén sesgados. Para deducir dónde se encuentra el error, se analizarán los espectros de los halos de referencia con la media de los espectros del catálogo de halos con el fin de observar si coinciden. Estos resultados se muestran en las figuras 4.14:

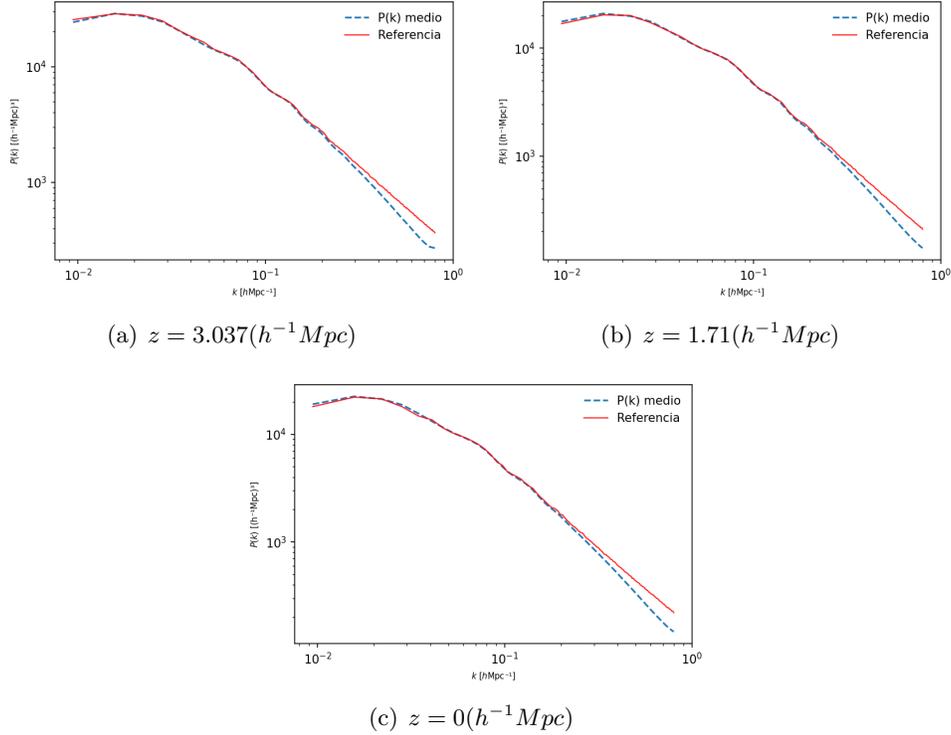


Figure 4.14: Valor medio de los espectros de potencia de los halos (línea discontinua azul) y el espectro de los halos de referencia (línea roja) en función de la escala

Observando los resultado de las gráficas 4.14 se puede concluir que el espectro de los halos de referencia y el valor medio de los espectros producidos por BAM coinciden, sobretodo a gran escala, que era la región de nuestro interés ya que el sesgo solo se calcula para escalas grandes hasta  $k \sim 0.05$ . Por lo tanto, se puede deducir que es el modelo lo que no es correcto, y la elección de otro corte de masa podría mejorar estos resultado.

#### 4.4.3 Espectro de los halos de materia oscura

En las siguientes figuras 4.15 podemos observar el espectro de la materia oscura multiplicado por el sesgo efectivo de referencia utilizado en las figuras anteriores (4.13).

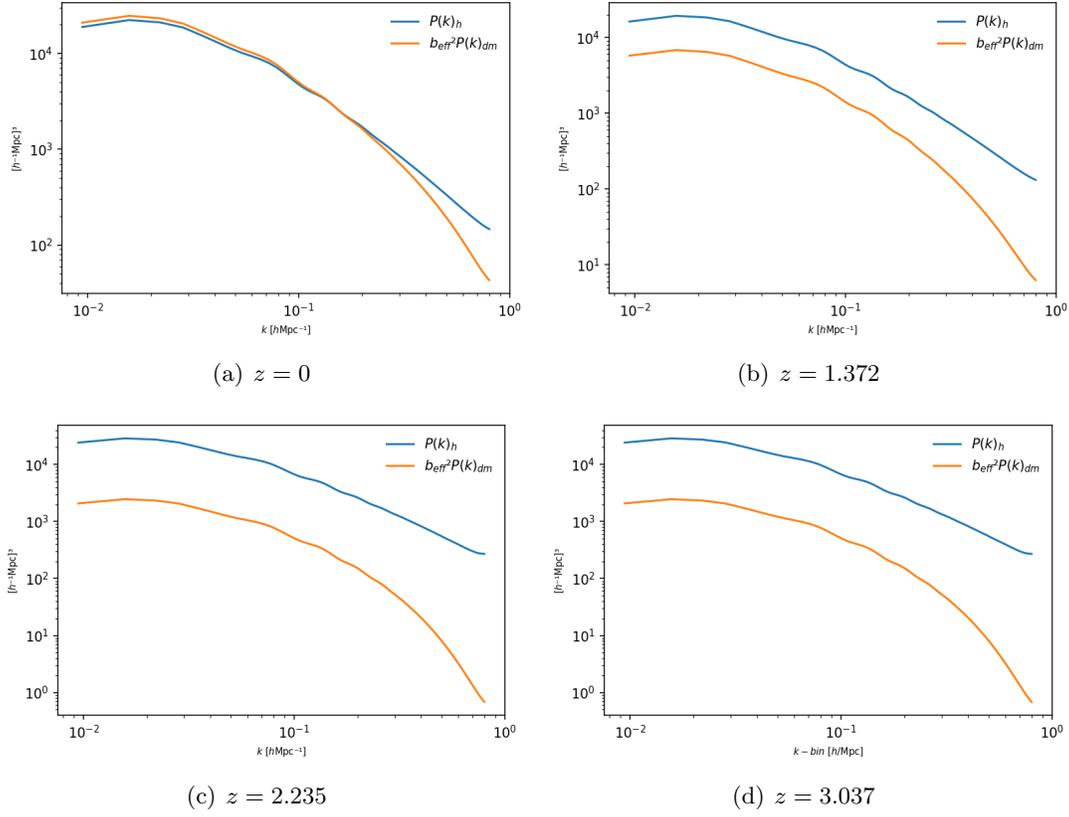


Figure 4.15: La línea naranja es el espectro de potencias de la materia oscura multiplicado por el valor del sesgo efectivo obtenido a partir de a fórmula de ajuste para la función de masa del halo de [43]. La línea azul representa el espectro de potencia de los halos de materia oscura de BAM.

Se observa que el espectro de potencias de los halos de BAM en las figuras 4.15 no se ajusta al producto del espectro de la materia oscura y el sesgo de referencia, y que a medida que aumenta el corrimiento al rojo, esta diferencia entre ambas distribuciones se hace cada vez más evidente. Sin embargo, esto tiene sentido ya que en la sección anterior (4.13), ya se ha comentado que el sesgo calculado es un 10% mayor que el sesgo de referencia y dado que a medida que el corrimiento al rojo aumenta, el sesgo de referencia también aumenta, esta diferencia entre ambas distribuciones es la esperada.

# Capítulo 5

## Conclusiones

Tras el análisis del espectro de potencias de los halos de BAM, podemos concluir que en general, se han obtenido resultados de calidad y que se ajustan con precisión a la teoría y referencias a gran escala  $k \sim 0.2hMpc^{-1}$ .

En primer lugar, la varianza experimental 4.5 se ha ajustado con precisión a la varianza teórica a pequeños modos, como era de esperar dado que a gran escala nos encontramos en el régimen lineal. Del mismo modo, la matriz de correlación 4.6 presenta baja correlación entre modos pequeños y a medida que aumenta  $k$  también lo han hecho los coeficientes de correlación, cumpliendo así lo esperado, ya que debido al acoplamiento entre modos, a medida que disminuimos la escala, comienzan a aparecer no linealidades (v.g. estructuras). Al analizar la curtosis de los espectros de potencia, hemos obtenido resultados consistentes con los valores para la curtosis de la distribución gaussiana a gran escala. Esto es lo esperado y significa que los espectros se están comportando como una distribución gaussiana, en respecto a la medida de la curtosis. Sin embargo, cabe destacar que al analizar la skewness del espectro, se han obtenido resultados que no corresponden a una distribución gaussiana a muy pequeños modos ( $k < 0.05hMpc^{-1}$ ). No obstante, cabe recordar que se considera que el espectro se puede analizar en el régimen lineal hasta  $k \sim 0.2hMpc^{-1}$ , y en el rango entre  $k = 0.05hMpc^{-1}$  y  $k = 0.2hMpc^{-1}$ , la skewness de la distribución gaussiana varía de forma similar a la de los espectros, por lo que en ese rango, podemos concluir que el espectro se comporta como una distribución gaussiana. Por otro lado, el test K-S ha dado buenos resultados para todos los valores de  $k$  y de  $z$  utilizados, es decir, el resultado de este test sugiere que los espectros de los halos se comportan como una distribución gaussiana. De esto se puede deducir que el test K-S está obviando algún aspecto o no es tan preciso como el análisis de la skewness en modos muy pequeños.

En cuanto al sesgo de los halos de BAM, siguen la tendencia esperada, pero sus valores obtenidos son un 10% mayores que los valores para el sesgo de referencia. Probablemente, esto se deba al corte de masa de los halos escogidos, por lo que para futuros experimentos, sería de interés generar un catálogo de halos cuyo modelo utilice un corte de masa superior, con el fin de observar si de esta forma se minimiza el sesgo del catálogo de halos obtenido. Por otro lado, las barras de error de este sesgo que se obtienen son muy pequeñas, por lo que en un futuro se podrían investigar diferentes rutas para no minimizar en tal medida las barras de error, dado que al estar utilizando un conjunto de espectros de materia oscura de baja resolución (siendo además estos muy similares a los espectros de los halos), y no el conjunto de espectros de materia oscura original, nuestra única fuente de estocasticidad es la varianza cósmica, la cual se elimina al calcular el sesgo, por ser este el cociente entre los dos conjuntos de espectros de potencia. Una alternativa para no perder esta estocasticidad podría ser utilizar un conjunto de espectros de potencia de la materia oscura de mayor resolución. Por último, los resultados del producto del espectro de la materia oscura y el sesgo efectivo concuerdan a la perfección

con los resultados esperados, teniendo en cuenta que el sesgo calculado en 4.13 era un 10% mayor que el sesgo efectivo de referencia.

En conclusión, la mayoría de comparaciones con referencias teóricas han dado buenos resultados y se han ajustado a lo que se esperaba, y aunque la skewness y los valores del sesgo no hayan sido los esperados, en el futuro podrán realizarse mejoras para ajustar mejor los productos teniendo en cuenta estos resultados. Por lo que, en gran medida, podemos concluir que los productos de BAM reproducen correctamente la estadística de 2 puntos de los halos a gran escala.

# Bibliografía

- [1] Marcos Pellejero-Ibañez, Andres Balaguera-Antolínez, Francisco-Shu Kitaura, Raúl E. Angulo, Gustavo Yepes, Chia-Hsun Chuang, Guillermo Reyes-Peraza, Mathieu Autefage, Mohammadjavad Vakili, and Cheng Zhao. The bias of dark matter tracers: assessing the accuracy of mapping techniques. , 493(1):586–593, March 2020.
- [2] Andrés Balaguera-Antolínez, Francisco-Shu Kitaura, Shadab Alam, Chia-Hsun Chuang, Yu Yu, Ginevra Favole, Cheng Zhao, David Brooks, Axel de la Macorra, Andreu Font-Ribera, Satya Gontcho A Gontcho, Klaus Honscheid, Robert Kehoe, Aron Meisner, Ramon Miquel, Gregory Tarlè, Mariana Vargas-Magaña, and Zhimin Zhou. DESI Mock Challenge VII: halo and galaxy catalogs with the Bias Assignment Method. *arXiv e-prints*, page arXiv:2211.10640, November 2022.
- [3] DESI Collaboration, Amir Aghamousa, Jessica Aguilar, Steve Ahlen, Shadab Alam, Lori E. Allen, Carlos Allende Prieto, James Annis, Stephen Bailey, Christophe Balland, Otger Ballester, Charles Baltay, Lucas Beaufore, Chris Bebek, Timothy C. Beers, Eric F. Bell, José Luis Bernal, Robert Besuner, Florian Beutler, Chris Blake, Hannes Bleuler, Michael Blomqvist, Robert Blum, Adam S. Bolton, Cesar Briceno, David Brooks, Joel R. Brownstein, Elizabeth Buckley-Geer, Angela Burden, Etienne Burtin, Nicolas G. Busca, Robert N. Cahn, Yan-Chuan Cai, Laia Cardiel-Sas, Raymond G. Carlberg, Pierre-Henri Carton, Ricard Casas, Francisco J. Castander, Jorge L. Cervantes-Cota, Todd M. Claybaugh, Madeline Close, Carl T. Coker, Shaun Cole, Johan Comparat, Andrew P. Cooper, M. C. Cousinou, Martin Crocce, Jean-Gabriel Cuby, Daniel P. Cunningham, Tamara M. Davis, Kyle S. Dawson, Axel de la Macorra, Juan De Vicente, Timothée Delubac, Mark Derwent, Arjun Dey, Govinda Dhungana, Zhejie Ding, Peter Doel, Yutong T. Duan, Anne Ealet, Jerry Edelstein, Sarah Eftekhazadeh, Daniel J. Eisenstein, Ann Elliott, Stéphanie Escoffier, Matthew Evatt, Parker Fagrelius, Xiaohui Fan, Kevin Fanning, Arya Farahi, Jay Farihi, Ginevra Favole, Yu Feng, Enrique Fernandez, Joseph R. Findlay, Douglas P. Finkbeiner, Michael J. Fitzpatrick, Brenna Flaugher, Samuel Flender, Andreu Font-Ribera, Jaime E. Forero-Romero, Pablo Fosalba, Carlos S. Frenk, Michele Fumagalli, Boris T. Gaensicke, Giuseppe Gallo, Juan Garcia-Bellido, Enrique Gaztanaga, Nicola Pietro Gentile Fusillo, Terry Gerard, Irena Gershkovich, Tommaso Giannantonio, Denis Gillet, Guillermo Gonzalez-de-Rivera, Violeta Gonzalez-Perez, Shelby Gott, Or Graur, Gaston Gutierrez, Julien Guy, Salman Habib, Henry Heetderks, Ian Heetderks, Katrin Heitmann, Wojciech A. Hellwing, David A. Herrera, Shirley Ho, Stephen Holland, Klaus Honscheid, Eric Huff, Timothy A. Hutchinson, Dragan Huterer, Ho Seong Hwang, Joseph Maria Illa Laguna, Yuzo Ishikawa, Dianna Jacobs, Niall Jeffrey, Patrick Jelsky, Elise Jennings, Linhua Jiang, Jorge Jimenez, Jennifer Johnson, Richard Joyce, Eric Jullo, Stéphanie Juneau, Sami Kama, Armin Karcher, Sonia Karkar, Robert Kehoe, Noble Kennamer, Stephen Kent, Martin Kilbinger, Alex G. Kim, David Kirkby, Theodore Kisner, Ellie Kitanidis, Jean-Paul Kneib, Sergey Koposov, Eve Kovacs, Kazuya Koyama, Anthony Kremin, Richard Kron, Luzius Kronig, Andrea Kueter-Young, Cedric G. Lacey,

- Robin Lafever, Ofer Lahav, Andrew Lambert, Michael Lampton, Martin Landriau, Dustin Lang, Tod R. Lauer, Jean-Marc Le Goff, Laurent Le Guillou, Auguste Le Van Suu, Jae Hyeon Lee, Su-Jeong Lee, Daniela Leitner, Michael Lesser, Michael E. Levi, Benjamin L’Huillier, Baojiu Li, Ming Liang, Huan Lin, Eric Linder, Sarah R. Loebman, Zarija Lukić, Jun Ma, Niall MacCrann, Christophe Magneville, Laleh Makarem, Marc Manera, Christopher J. Manser, Robert Marshall, Paul Martini, Richard Massey, Thomas Matheson, Jeremy McCauley, Patrick McDonald, Ian D. McGreer, Aaron Meisner, Nigel Metcalfe, Timothy N. Miller, Ramon Miquel, John Moustakas, Adam Myers, Milind Naik, Jeffrey A. Newman, Robert C. Nichol, Andrina Nicola, Luiz Nicolati da Costa, Jundan Nie, Gustavo Niz, Peder Norberg, Brian Nord, Dara Norman, Peter Nugent, Thomas O’Brien, Minji Oh, Knut A. G. Olsen, Cristobal Padilla, Hamsa Padmanabhan, Nikhil Padmanabhan, Nathalie Palanque-Delabrouille, Antonella Palmese, Daniel Pappalardo, Isabelle Pâris, Changbom Park, Anna Patej, John A. Peacock, Hiranya V. Peiris, Xiyan Peng, Will J. Percival, Sandrine Perruchot, Matthew M. Pieri, Richard Pogge, Jennifer E. Pollack, Claire Poppett, Francisco Prada, Abhishek Prakash, Ronald G. Probst, David Rabinowitz, Anand Raichoor, Chang Hee Ree, Alexandre Refregier, Xavier Regal, Beth Reid, Kevin Reil, Mehdi Rezaie, Constance M. Rockosi, Natalie Roe, Samuel Ronayette, Aaron Roodman, Ashley J. Ross, Nicholas P. Ross, Graziano Rossi, Eduardo Roza, Vanina Ruhlmann-Kleider, Eli S. Rykoff, Cristiano Sabiu, Lado Samushia, Eusebio Sanchez, Javier Sanchez, David J. Schlegel, Michael Schneider, Michael Schubnell, Aurélia Secroun, Uros Seljak, Hee-Jong Seo, Santiago Serrano, Arman Shafieloo, Huanyuan Shan, Ray Sharples, Michael J. Sholl, William V. Shourt, Joseph H. Silber, David R. Silva, Martin M. Sirk, Anze Slosar, Alex Smith, George F. Smoot, Debopam Som, Yong-Seon Song, David Sprayberry, Ryan Staten, Andy Stefanik, Gregory Tarle, Suk Sien Tie, Jeremy L. Tinker, Rita Tojeiro, Francisco Valdes, Octavio Valenzuela, Monica Valluri, Mariana Vargas-Magana, Licia Verde, Alistair R. Walker, Jiali Wang, Yuting Wang, Benjamin A. Weaver, Curtis Weaverdyck, Risa H. Wechsler, David H. Weinberg, Martin White, Qian Yang, Christophe Yeche, Tianmeng Zhang, Gong-Bo Zhao, Yi Zheng, Xu Zhou, Zhimin Zhou, Yaling Zhu, Hu Zou, and Ying Zu. The DESI Experiment Part I: Science, Targeting, and Survey Design. *arXiv e-prints*, page arXiv:1611.00036, October 2016.
- [4] Kyle S. Dawson, Jean-Paul Kneib, Will J. Percival, Shadab Alam, Franco D. Albareti, Scott F. Anderson, Eric Armengaud, Éric Aubourg, Stephen Bailey, Julian E. Bautista, Andreas A. Berlind, Matthew A. Bershad, Florian Beutler, Dmitry Bizyaev, Michael R. Blanton, Michael Blomqvist, Adam S. Bolton, Jo Bovy, W. N. Brandt, Jon Brinkmann, Joel R. Brownstein, Etienne Burtin, N. G. Busca, Zheng Cai, Chia-Hsun Chuang, Nicolas Clerc, Johan Comparat, Frances Cope, Rupert A. C. Croft, Irene Cruz-Gonzalez, Luiz N. da Costa, Marie-Claude Cousinou, Jeremy Darling, Axel de la Macorra, Sylvain de la Torre, Timothée Delubac, Héliion du Mas des Bourboux, Tom Dwelly, Anne Ealet, Daniel J. Eisenstein, Michael Eracleous, S. Escoffier, Xiaohui Fan, Alexis Finoguenov, Andreu Font-Ribera, Peter Frinchaboy, Patrick Gaulme, Antonis Georgakakis, Paul Green, Hong Guo, Julien Guy, Shirley Ho, Diana Holder, Joe Huehnerhoff, Timothy Hutchinson, Yipeng Jing, Eric Jullo, Vikrant Kamble, Karen Kinemuchi, David Kirkby, Francisco-Shu Kitaura, Mark A. Klaene, Russ R. Laher, Dustin Lang, Pierre Laurent, Jean-Marc Le Goff, Cheng Li, Yu Liang, Marcos Lima, Qiufan Lin, Weipeng Lin, Yen-Ting Lin, Daniel C. Long, Britt Lundgren, Nicholas MacDonald, Marcio Antonio Gemiba Maia, Elena Malanushenko, Viktor Malanushenko, Vivek Mariappan, Cameron K. McBride, Ian D. McGreer, Brice Ménard, Andrea Merloni, Andres Meza, Antonio D. Montero-Dorta, Demitri Muna, Adam D. Myers, Kirpal Nandra, Tracy Naugle, Jef-

- frey A. Newman, Pasquier Noterdaeme, Peter Nugent, Ricardo Ogando, Matthew D. Olmstead, Audrey Oravetz, Daniel J. Oravetz, Nikhil Padmanabhan, Nathalie Palanque-Delabrouille, Kaike Pan, John K. Parejko, Isabelle Pâris, John A. Peacock, Patrick Petitjean, Matthew M. Pieri, Alice Pisani, Francisco Prada, Abhishek Prakash, Anand Raichoor, Beth Reid, James Rich, Jethro Ridl, Sergio Rodriguez-Torres, Aurelio Carnero Rosell, Ashley J. Ross, Graziano Rossi, John Ruan, Mara Salvato, Conor Sayres, Donald P. Schneider, David J. Schlegel, Uros Seljak, Hee-Jong Seo, Branimir Sesar, Sarah Shandera, Yiping Shu, Anže Slosar, Flavia Sobreira, Alina Streblyanska, Nao Suzuki, Donna Taylor, Charling Tao, Jeremy L. Tinker, Rita Tojeiro, Mariana Vargas-Magaña, Yuting Wang, Benjamin A. Weaver, David H. Weinberg, Martin White, W. M. Wood-Vasey, Christophe Yeche, Zhongxu Zhai, Cheng Zhao, Gong-bo Zhao, Zheng Zheng, Guangtun Ben Zhu, and Hu Zou. The SDSS-IV Extended Baryon Oscillation Spectroscopic Survey: Overview and Early Data. , 151(2):44, February 2016.
- [5] The Dark Energy Survey Collaboration. The Dark Energy Survey. *arXiv e-prints*, pages astro-ph/0510346, October 2005.
- [6] Luca Amendola, Stephen Appleby, Anastasios Avgoustidis, David Bacon, Tessa Baker, Marco Baldi, Nicola Bartolo, Alain Blanchard, Camille Bonvin, Stefano Borgani, Enzo Branchini, Clare Burrage, Stefano Camera, Carmelita Carbone, Luciano Casarini, Mark Cropper, Claudia de Rham, Jörg P. Dietrich, Cinzia Di Porto, Ruth Durrer, Anne Ealet, Pedro G. Ferreira, Fabio Finelli, Juan García-Bellido, Tommaso Giannantonio, Luigi Guzzo, Alan Heavens, Lavinia Heisenberg, Catherine Heymans, Henk Hoekstra, Lukas Hollenstein, Rory Holmes, Zhiqi Hwang, Knud Jahnke, Thomas D. Kitching, Tomi Koivisto, Martin Kunz, Giuseppe La Vacca, Eric Linder, Marisa March, Valerio Marra, Carlos Martins, Elisabetta Majerotto, Dida Markovic, David Marsh, Federico Marulli, Richard Massey, Yannick Mellier, Francesco Montanari, David F. Mota, Nelson J. Nunes, Will Percival, Valeria Pettorino, Cristiano Porciani, Claudia Quercellini, Justin Read, Massimiliano Rinaldi, Domenico Sapone, Ignacy Sawicki, Roberto Scaramella, Constantinos Skordis, Fergus Simpson, Andy Taylor, Shaun Thomas, Roberto Trotta, Licia Verde, Filippo Vernizzi, Adrian Vollmer, Yun Wang, Jochen Weller, and Tom Zlosnik. Cosmology and fundamental physics with the Euclid satellite. *Living Reviews in Relativity*, 21(1):2, April 2018.
- [7] Scott Dodelson and Michael D. Schneider. The effect of covariance estimator error on cosmological parameter constraints. , 88(6):063537, September 2013.
- [8] Andy Taylor, Benjamin Joachimi, and Thomas Kitching. Putting the precision in precision cosmology: How accurate should your data covariance matrix be? , 432(3):1928–1946, July 2013.
- [9] Will J. Percival, Ashley J. Ross, Ariel G. Sánchez, Lado Samushia, Angela Burden, Robert Crittenden, Antonio J. Cuesta, Mariana Vargas Magana, Marc Manera, Florian Beutler, Chia-Hsun Chuang, Daniel J. Eisenstein, Shirley Ho, Cameron K. McBride, Francesco Montesano, Nikhil Padmanabhan, Beth Reid, Shun Saito, Donald P. Schneider, Hee-Jong Seo, Rita Tojeiro, and Benjamin A. Weaver. The clustering of Galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: including covariance matrix errors. , 439(3):2531–2541, April 2014.
- [10] Dante J. Paz and Ariel G. Sánchez. Improving the precision matrix for precision cosmology. , 454(4):4326–4334, December 2015.

- [11] David W. Pearson and Lado Samushia. Estimating the power spectrum covariance matrix with fewer mock samples. , 457(1):993–999, March 2016.
- [12] Cullan Howlett and Will J. Percival. Galaxy two-point covariance matrix estimation for next generation surveys. , 472(4):4935–4952, December 2017.
- [13] Ross O’Connell and Daniel J. Eisenstein. Large covariance matrices: accurate models without mocks. , 487(2):2701–2717, August 2019.
- [14] Daniel J. Eisenstein and Wayne Hu. Baryonic Features in the Matter Transfer Function. , 496(2):605–614, March 1998.
- [15] Douglas H. Rudd, Andrew R. Zentner, and Andrey V. Kravtsov. Effects of Baryons and Dissipation on the Matter Power Spectrum. , 672(1):19–32, January 2008.
- [16] Nick Kaiser. Clustering in real space and in redshift space. , 227:1–21, July 1987.
- [17] Steven Weinberg. *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*. Wiley, New York, NY, 1972.
- [18] P. J. E. Peebles. *The large-scale structure of the universe*. 1980.
- [19] Houjun Mo, Frank C. van den Bosch, and Simon White. *Galaxy Formation and Evolution*. 2010.
- [20] Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, S. Basak, R. Battye, K. Benabed, J. P. Bernard, M. Bersanelli, P. Bielewicz, J. J. Bock, J. R. Bond, J. Borrill, F. R. Bouchet, F. Boulanger, M. Bucher, C. Burigana, R. C. Butler, E. Calabrese, J. F. Cardoso, J. Carron, A. Challinor, H. C. Chiang, J. Chluba, L. P. L. Colombo, C. Combet, D. Contreras, B. P. Crill, F. Cuttaia, P. de Bernardis, G. de Zotti, J. Delabrouille, J. M. Delouis, E. Di Valentino, J. M. Diego, O. Doré, M. Douspis, A. Ducout, X. Dupac, S. Dusini, G. Efstathiou, F. Elsner, T. A. Enßlin, H. K. Eriksen, Y. Fantaye, M. Farhang, J. Fergusson, R. Fernandez-Cobos, F. Finelli, F. Forastieri, M. Frailis, A. A. Fraisse, E. Franceschi, A. Frolov, S. Galeotta, S. Galli, K. Ganga, R. T. Génova-Santos, M. Gerbino, T. Ghosh, J. González-Nuevo, K. M. Górski, S. Gratton, A. Gruppuso, J. E. Gudmundsson, J. Hamann, W. Handley, F. K. Hansen, D. Herranz, S. R. Hildebrandt, E. Hivon, Z. Huang, A. H. Jaffe, W. C. Jones, A. Karakci, E. Keihänen, R. Keskitalo, K. Kiiveri, J. Kim, T. S. Kisner, L. Knox, N. Krachmalnicoff, M. Kunz, H. Kurki-Suonio, G. Lagache, J. M. Lamarre, A. Lasenby, M. Lattanzi, C. R. Lawrence, M. Le Jeune, P. Lemos, J. Lesgourgues, F. Levrier, A. Lewis, M. Liguori, P. B. Lilje, M. Lilley, V. Lindholm, M. López-Caniago, P. M. Lubin, Y. Z. Ma, J. F. Macías-Pérez, G. Maggio, D. Maino, N. Mandolesi, A. Mangilli, A. Marcos-Caballero, M. Maris, P. G. Martin, M. Martinelli, E. Martínez-González, S. Matarrese, N. Mauri, J. D. McEwen, P. R. Meinhold, A. Melchiorri, A. Mennella, M. Migliaccio, M. Millea, S. Mitra, M. A. Miville-Deschênes, D. Molinari, L. Montier, G. Morgante, A. Moss, P. Natoli, H. U. Nørgaard-Nielsen, L. Pagano, D. Paoletti, B. Partridge, G. Patanchon, H. V. Peiris, F. Perrotta, V. Pettorino, F. Piacentini, L. Polastri, G. Polenta, J. L. Puget, J. P. Rachen, M. Reinecke, M. Remazeilles, A. Renzi, G. Rocha, C. Rosset, G. Roudier, J. A. Rubiño-Martín, B. Ruiz-Granados, L. Salvati, M. Sandri, M. Savelainen, D. Scott, E. P. S. Shellard, C. Sirignano, G. Sirri, L. D. Spencer, R. Sunyaev, A. S. Suur-Uski, J. A. Tauber, D. Tavagnacco, M. Tenti, L. Toffolatti, M. Tomasi, T. Trombetti, L. Valenziano, J. Valiviita, B. Van Tent, L. Vibert, P. Vielva, F. Villa, N. Vittorio, B. D. Wandelt, I. K.

- Wehus, M. White, S. D. M. White, A. Zacchei, and A. Zonca. Planck 2018 results. VI. Cosmological parameters. , 641:A6, September 2020.
- [21] G. Hinshaw, D. Larson, E. Komatsu, D. N. Spergel, C. L. Bennett, J. Dunkley, M. R. Nolta, M. Halpern, R. S. Hill, N. Odegard, L. Page, K. M. Smith, J. L. Weiland, B. Gold, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, E. Wollack, and E. L. Wright. Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results. , 208(2):19, October 2013.
- [22] Shadab Alam, Metin Ata, Stephen Bailey, Florian Beutler, Dmitry Bizyaev, Jonathan A. Blazek, Adam S. Bolton, Joel R. Brownstein, Angela Burden, Chia-Hsun Chuang, Johan Comparat, Antonio J. Cuesta, Kyle S. Dawson, Daniel J. Eisenstein, Stephanie Escoffier, Héctor Gil-Marín, Jan Niklas Grieb, Nick Hand, Shirley Ho, Karen Kinemuchi, David Kirkby, Francisco Kitaura, Elena Malanushenko, Viktor Malanushenko, Claudia Maraston, Cameron K. McBride, Robert C. Nichol, Matthew D. Olmstead, Daniel Oravetz, Nikhil Padmanabhan, Nathalie Palanque-Delabrouille, Kaike Pan, Marcos Pellejero-Ibanez, Will J. Percival, Patrick Petitjean, Francisco Prada, Adrian M. Price-Whelan, Beth A. Reid, Sergio A. Rodríguez-Torres, Natalie A. Roe, Ashley J. Ross, Nicholas P. Ross, Graziano Rossi, Jose Alberto Rubiño-Martín, Shun Saito, Salvador Salazar-Albornoz, Lado Samushia, Ariel G. Sánchez, Siddharth Satpathy, David J. Schlegel, Donald P. Schneider, Claudia G. Scóccola, Hee-Jong Seo, Erin S. Sheldon, Audrey Simmons, Anže Slosar, Michael A. Strauss, Molly E. C. Swanson, Daniel Thomas, Jeremy L. Tinker, Rita Tojeiro, Mariana Vargas Magaña, Jose Alberto Vazquez, Licia Verde, David A. Wake, Yuting Wang, David H. Weinberg, Martin White, W. Michael Wood-Vasey, Christophe Yèche, Idit Zehavi, Zhongxu Zhai, and Gong-Bo Zhao. The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample. , 470(3):2617–2652, September 2017.
- [23] Michael Levi, Chris Bebek, Timothy Beers, Robert Blum, Robert Cahn, Daniel Eisenstein, Brenna Flaugher, Klaus Honscheid, Richard Kron, Ofer Lahav, Patrick McDonald, Natalie Roe, David Schlegel, and representing the DESI collaboration. The DESI Experiment, a whitepaper for Snowmass 2013. *arXiv e-prints*, page arXiv:1308.0847, August 2013.
- [24] Adam G. Riess, Alexei V. Filippenko, Peter Challis, Alejandro Clocchiatti, Alan Diercks, Peter M. Garnavich, Ron L. Gilliland, Craig J. Hogan, Saurabh Jha, Robert P. Kirshner, B. Leibundgut, M. M. Phillips, David Reiss, Brian P. Schmidt, Robert A. Schommer, R. Chris Smith, J. Spyromilio, Christopher Stubbs, Nicholas B. Suntzeff, and John Tonry. Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. , 116(3):1009–1038, September 1998.
- [25] Andrew R. Liddle and David H. Lyth. *Cosmological Inflation and Large-Scale Structure*. 2000.
- [26] Steven R. Furlanetto, S. Peng Oh, and Frank H. Briggs. Cosmology at low frequencies: The 21 cm transition and the high-redshift Universe. , 433(4-6):181–301, October 2006.
- [27] Judd D. Bowman, Alan E. E. Rogers, Raul A. Monsalve, Thomas J. Mozdzen, and Nivedita Mahesh. An absorption profile centred at 78 megahertz in the sky-averaged spectrum. , 555(7694):67–70, March 2018.

- [28] Saurabh Singh, Nambissan T. Jishnu, Ravi Subrahmanyam, N. Udaya Shankar, B. S. Girish, A. Raghunathan, R. Somashekar, K. S. Srivani, and Mayuri Sathyanarayana Rao. On the detection of a cosmic dawn signal in the radio background. *Nature Astronomy*, 6:607–617, February 2022.
- [29] R. Barkana and A. Loeb. In the beginning: the first sources of light and the reionization of the universe. , 349(2):125–238, July 2001.
- [30] Asantha Cooray and Ravi Sheth. Halo models of large scale structure. , 372(1):1–129, December 2002.
- [31] John A. Peacock, 2dF Galaxy Redshift Survey Team, Matthew Colless, Ivan Baldry, Carlton Baugh, Joss Bland-Hawthorn, Terry Bridges, Russell Cannon, Shaun Cole, Chris Collins, Warrick Couch, Gavin Dalton, Robert De Propris, Simon Driver, George Efsthathiou, Richard Ellis, Carlos Frenk, Karl Glazebrook, Carole Jackson, Ofer Lahav, Ian Lewis, Stuart Lumsden, Steve Maddox, Darren Madgwick, Peder Norberg, Will Percival, Bruce Peterson, Will Sutherland, and Keith Taylor. Studying Large-scale Structure with the 2dF Galaxy Redshift Survey. In Nigel Metcalfe and Tom Shanks, editors, *A New Era in Cosmology*, volume 283 of *Astronomical Society of the Pacific Conference Series*, page 19, January 2002.
- [32] P. J. E. Peebles. Statistics of the distribution of galaxies. In *Ninth Texas Symposium on Relativistic Astrophysics*, volume 336, pages 161–171, February 1980.
- [33] John A. Peacock. *Cosmological Physics*. 1999.
- [34] Scott Dodelson. *Modern Cosmology*. 2003.
- [35] Lev Davidovich Landau and E. M. Lifshitz. *Statistical physics. Pt.1*. 1969.
- [36] Hume A. Feldman, Nick Kaiser, and John A. Peacock. Power-Spectrum Analysis of Three-dimensional Redshift Surveys. , 426:23, May 1994.
- [37] A. Balaguera-Antolínez, Francisco-Shu Kitaura, Marcos Pellejero-Ibáñez, Cheng Zhao, and Tom Abel. BAM: bias assignment method to generate mock catalogues. , 483(1):L58–L63, February 2019.
- [38] Chia-Hsun Chuang, Gustavo Yepes, Francisco-Shu Kitaura, Marcos Pellejero-Ibanez, Sergio Rodríguez-Torres, Yu Feng, Robert Benton Metcalf, Risa H. Wechsler, Cheng Zhao, Chun-Hao To, Shadab Alam, Arka Banerjee, Joseph DeRose, Carlo Giocoli, Alexander Knebe, and Guillermo Reyes. UNIT project: Universe N-body simulations for the Investigation of Theoretical models from galaxy surveys. , 487(1):48–59, July 2019.
- [39] Marc Manera, Roman Scoccimarro, Will J. Percival, Lado Samushia, Cameron K. McBride, Ashley J. Ross, Ravi K. Sheth, Martin White, Beth A. Reid, Ariel G. Sánchez, Roland de Putter, Xiaoying Xu, Andreas A. Berlind, Jonathan Brinkmann, Claudia Maraston, Bob Nichol, Francesco Montesano, Nikhil Padmanabhan, Ramin A. Skibba, Rita Tojeiro, and Benjamin A. Weaver. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: a large sample of mock galaxy catalogues. , 428(2):1036–1054, January 2013.
- [40] Francisco-Shu Kitaura, Sergio Rodríguez-Torres, Chia-Hsun Chuang, Cheng Zhao, Francisco Prada, Héctor Gil-Marín, Hong Guo, Gustavo Yepes, Anatoly Klypin, Claudia G. Scóccola, Jeremy Tinker, Cameron McBride, Beth Reid, Ariel G. Sánchez, Salvador

- Salazar-Albornoz, Jan Niklas Grieb, Mariana Vargas-Magana, Antonio J. Cuesta, Mark Neyrinck, Florian Beutler, Johan Comparat, Will J. Percival, and Ashley Ross. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: mock galaxy catalogues for the BOSS Final Data Release. , 456(4):4156–4173, March 2016.
- [41] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, N. Bartolo, E. Battaner, R. Battye, K. Benabed, A. Benoît, A. Benoit-Lévy, J. P. Bernard, M. Bersanelli, P. Bielewicz, J. J. Bock, A. Bonaldi, L. Bonavera, J. R. Bond, J. Borrill, F. R. Bouchet, F. Boulanger, M. Bucher, C. Burigana, R. C. Butler, E. Calabrese, J. F. Cardoso, A. Catalano, A. Challinor, A. Chamballu, R. R. Chary, H. C. Chiang, J. Chluba, P. R. Christensen, S. Church, D. L. Clements, S. Colombi, L. P. L. Colombo, C. Combet, A. Coulais, B. P. Crill, A. Curto, F. Cuttaia, L. Danese, R. D. Davies, R. J. Davis, P. de Bernardis, A. de Rosa, G. de Zotti, J. Delabrouille, F. X. Désert, E. Di Valentino, C. Dickinson, J. M. Diego, K. Dolag, H. Dole, S. Donzelli, O. Doré, M. Douspis, A. Ducout, J. Dunkley, X. Dupac, G. Efstathiou, F. Elsner, T. A. Enßlin, H. K. Eriksen, M. Farhang, J. Fergusson, F. Finelli, O. Forni, M. Frailis, A. A. Fraisse, E. Franceschi, A. Frejsel, S. Galeotta, S. Galli, K. Ganga, C. Gauthier, M. Gerbino, T. Ghosh, M. Giard, Y. Giraud-Héraud, E. Giusarma, E. Gjerløw, J. González-Nuevo, K. M. Górski, S. Gratton, A. Gregorio, A. Gruppuso, J. E. Gudmundsson, J. Hamann, F. K. Hansen, D. Hanson, D. L. Harrison, G. Helou, S. Henrot-Versillé, C. Hernández-Monteagudo, D. Herranz, S. R. Hildebrandt, E. Hivon, M. Hobson, W. A. Holmes, A. Hornstrup, W. Hovest, Z. Huang, K. M. Huffenberger, G. Hurier, A. H. Jaffe, T. R. Jaffe, W. C. Jones, M. Juvela, E. Keihänen, R. Kesitalo, T. S. Kisner, R. Kneissl, J. Knoche, L. Knox, M. Kunz, H. Kurki-Suonio, G. Lagache, A. Lähteenmäki, J. M. Lamarre, A. Lasenby, M. Lattanzi, C. R. Lawrence, J. P. Leahy, R. Leonardi, J. Lesgourgues, F. Levrier, A. Lewis, M. Liguori, P. B. Lilje, M. Linden-Vørnle, M. López-Caniego, P. M. Lubin, J. F. Macías-Pérez, G. Maggio, D. Maino, N. Mandolesi, A. Mangilli, A. Marchini, M. Maris, P. G. Martin, M. Martinelli, E. Martínez-González, S. Masi, S. Matarrese, P. McGehee, P. R. Meinhold, A. Melchiorri, J. B. Melin, L. Mendes, A. Mennella, M. Migliaccio, M. Millea, S. Mitra, M. A. Miville-Deschênes, A. Moneti, L. Montier, G. Morgante, D. Mortlock, A. Moss, D. Munshi, J. A. Murphy, P. Naselsky, F. Nati, P. Natoli, C. B. Netterfield, H. U. Nørgaard-Nielsen, F. Noviello, D. Novikov, I. Novikov, C. A. Oxborrow, F. Paci, L. Pagano, F. Pajot, R. Paladini, D. Paoletti, B. Partridge, F. Pasian, G. Patanchon, T. J. Pearson, O. Perdereau, L. Perotto, F. Perrotta, V. Pettorino, F. Piacentini, M. Piat, E. Pierpaoli, D. Pietrobon, S. Plaszczynski, E. Pointecouteau, G. Polenta, L. Popa, G. W. Pratt, G. Prézeau, S. Prunet, J. L. Puget, J. P. Rachen, W. T. Reach, R. Rebolo, M. Reinecke, M. Remazeilles, C. Renault, A. Renzi, I. Ristorcelli, G. Rocha, C. Rosset, M. Rossetti, G. Roudier, B. Rouillé d’Orfeuil, M. Rowan-Robinson, J. A. Rubiño-Martín, B. Rusholme, N. Said, V. Salvatelli, L. Salvati, M. Sandri, D. Santos, M. Savelainen, G. Savini, D. Scott, M. D. Seiffert, P. Serra, E. P. S. Shellard, L. D. Spencer, M. Spinelli, V. Stolyarov, R. Stompor, R. Sudiwala, R. Sunyaev, D. Sutton, A. S. Suur-Uski, J. F. Sygnet, J. A. Tauber, L. Terenzi, L. Toffolatti, M. Tomasi, M. Tristram, T. Trombetti, M. Tucci, J. Tuovinen, M. Türler, G. Umama, L. Valenziano, J. Valiviita, F. Van Tent, P. Vielva, F. Villa, L. A. Wade, B. D. Wandelt, I. K. Wehus, M. White, S. D. M. White, A. Wilkinson, D. Yvon, A. Zacchei, and A. Zonca. Planck 2015 results. XIII. Cosmological parameters. , 594:A13, September 2016.
- [42] Uroš Seljak and Michael S. Warren. Large-scale bias and stochasticity of haloes and dark matter. , 355(1):129–136, November 2004.

- [43] Jeremy Tinker, Andrey V. Kravtsov, Anatoly Klypin, Kevork Abazajian, Michael Warren, Gustavo Yepes, Stefan Gottlöber, and Daniel E. Holz. Toward a Halo Mass Function for Precision Cosmology: The Limits of Universality. , 688(2):709–728, December 2008.
- [44] William H. Press and Paul Schechter. Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. , 187:425–438, February 1974.
- [45] P. Valageas. Mass functions and bias of dark matter halos. , 508(1):93–106, December 2009.
- [46] Jeremy L. Tinker, Brant E. Robertson, Andrey V. Kravtsov, Anatoly Klypin, Michael S. Warren, Gustavo Yepes, and Stefan Gottlöber. The Large-scale Bias of Dark Matter Halos: Numerical Calibration and Model Tests. , 724(2):878–886, December 2010.
- [47] Ravi K. Sheth and Gerard Lemson. Biasing and the distribution of dark matter haloes. , 304(4):767–792, April 1999.
- [48] Chia-Hsun Chuang, Francisco-Shu Kitaura, Francisco Prada, Cheng Zhao, and Gustavo Yepes. EZmocks: extending the Zel’dovich approximation to generate mock galaxy catalogues with accurate clustering statistics. , 446(3):2621–2628, January 2015.
- [49] A. Balaguera-Antolínez, Francisco-Shu Kitaura, M. Pellejero-Ibáñez, Martha Lippich, Cheng Zhao, Ariel G. Sánchez, Claudio Dalla Vecchia, Raúl E. Angulo, and Martín Crocce. One simulation to have them all: performance of the Bias Assignment Method against N-body simulations. , 491(2):2565–2575, January 2020.
- [50] Francesco Sinigaglia, Francisco-Shu Kitaura, A. Balaguera-Antolínez, Kentaro Nagamine, Metin Ata, Ikkoh Shimizu, and M. Sánchez-Benavente. The Bias from Hydrodynamic Simulations: Mapping Baryon Physics onto Dark Matter Fields. , 921(1):66, November 2021.
- [51] F. S. Kitaura, A. Balaguera-Antolínez, F. Sinigaglia, and M. Pellejero-Ibáñez. The cosmic web connection to the dark matter halo distribution through gravity. , 512(2):2245–2265, May 2022.