



Centro de Investigación y estudios avanzados

Departamento de Física

Reconstrucción de la energía oscura con base en observaciones cosmológicas.

Luis Adrián Escamilla Torres

Tesis de maestría

Director de tesis: Dr. Tonatiuh Matos Chassin.

San Pedro Zacatenco, CDMX, 23 de noviembre de 2018

Agradecimientos.

A mis padres y hermanos que, a pesar de la distancia, siempre están para mi y me apoyan sin importar nada. A mis roomies los dos Jonas y mamá Wendy, por ayudarnos y cuidarnos mutuamente durante la estancia. A Osvaldo y Andrés, por dar su amistad a los regios. Al Dr. Tonatiuh Matos y al Dr. Alberto Vázquez por siempre estar dispuestos a apoyarme y ayudarme con mis dudas y por compartir su experiencia siempre de buen humor. A Malu y Rose por su ayuda con los procesos burocráticos. A mis amigos y compañeros de las clases de baile, por hacer más divertidos mis fines de semana (y a veces mis días entre semana). A Chris y LA por siempre llevarme a conocer lugares nuevos. Al CINVESTAV y Conacyt por facilitar las instalaciones y el material que me permitió realizar este trabajo. Y a mi abuela, por siempre apoyarme y mantener su buen humor, aún en sus últimos momentos.

Índice general

Introducción.	V
Resumen.	VII
Abstract.	IX
1. Un poco de cosmología	1
1.1. El universo homogéneo y su geometría	1
1.2. Dinámica	6
1.3. Menú cósmico	9
1.4. Energía oscura [3]	11
1.4.1. Supernovas tipo Ia	12
1.4.2. CMB	14
1.4.3. Oscilaciones acústicas de bariones (BAOs)	16
1.4.4. Parámetro de Hubble	17
1.4.5. Modelos de energía oscura	17
2. Reconstrucciones de la energía oscura	19
2.1. Reconstrucciones paramétricas	19
2.2. PCA	21
2.3. Reconstrucciones no paramétricas	23

3. Estadística Bayesiana y selección de modelos	29
3.1. Conceptos, definiciones y ejemplos	29
3.2. Estimación de parámetros	33
3.3. Prueba de hipótesis (selección de modelo)	37
3.4. Cadenas de Markov	38
3.5. Cadenas de Markov Monte Carlo (MCMC)	40
3.5.1. Ejemplo de MCMC con Λ CDM	41
3.6. AIC, AICc, BIC y DIC	44
4. Interpolaciones como reconstrucciones	47
4.1. Tipos de interpolaciones	47
5. Prueba	53
5.1. Reconstruyendo $H(z)$ con datos falsos	53
6. Resultados	63
6.1. Reconstrucción de $F_l(z)$.	63
6.2. Reconstrucción de $w(z)$.	66
6.3. Comparación con Λ CDM	70
6.4. Prueba de interpolaciones con datos reales.	72
7. Conclusiones y perspectivas.	75
7.1. Conclusión final	75
7.2. Perspectivas	76
Bibliografía	76

Introducción.

Aún quedan muchos misterios en el universo, y uno de ellos es la naturaleza de dos sustancias que se descubrieron apenas el siglo pasado: la materia oscura y la energía oscura. La materia oscura se descubrió luego de observar el movimiento de las galaxias, los astrónomos se dieron cuenta de que para que las galaxias giraran de la forma en la que lo hacían debían contener mucha más materia que la que observaban, aquí nació la idea de la materia oscura.

La energía oscura, en cambio, surgió debido al descubrimiento de la expansión acelerada del universo. Como no se sabía el *porqué* de dicha expansión se recurrió a la idea de la energía oscura, un tipo de energía (que a veces se le llama “energía del vacío”) que provoca un incremento en la velocidad de expansión del universo.

El objetivo de este trabajo es utilizar los datos observacionales de un tipo especial de supernova, la radiación cósmica de fondo y las oscilaciones acústicas de bariones para reconstruir de manera no paramétrica algunas propiedades de la energía oscura y poder ver si debe tener dinámica o debe ser una constante cosmológica como indica el modelo Λ CDM.

Resumen.

Este trabajo intenta demostrar que, dadas observaciones recientes, es necesario tener un modelo de energía oscura dinámica, en lugar de considerarlo una constante cosmológica. En el capítulo 1 se dará un breve repaso de los conceptos importantes de cosmología utilizados en este trabajo. En el capítulo 2 se habla sobre los tipos de reconstrucciones que se utilizan normalmente. El capítulo 3 trata de estadística Bayesiana. El capítulo 4 habla de la reconstrucción que se utilizará en este trabajo y el capítulo 5 muestra un ejemplo de su aplicación. El capítulo 6 es la aplicación de dicha reconstrucción con datos reales y su comparación con el modelo de constante cosmológica, y el capítulo 7 contiene las conclusiones y perspectivas.

Abstract.

This work tries to show that, given recent observations and probes, it is necessary to have a dynamical dark energy model, instead of a cosmological constant one. In chapter 1 it is given a brief review of the important concepts of cosmology used in this work. In chapter 2 the types of reconstruction are discussed. In chapter 3 an introduction to Bayesian statistics is presented. In chapter 4 the focus is the reconstruction used in this work. In chapter 5 an example of said reconstruction is shown. In chapter 6 this reconstruction is used with real data and in chapter 7 the conclusions and future plans are discussed.

Capítulo 1

Un poco de cosmología

Este capítulo pretende dar una introducción a los conceptos de cosmología más relevantes para este trabajo. Gran parte de este capítulo está basado en [1] y [2].

1.1. El universo homogéneo y su geometría

Al estudiar cosmología siempre se parte del mismo punto: el principio cosmológico. Este principio nos dice que el universo es homogéneo e isotrópico al irnos a grandes escalas (alrededor de 100 Megaparsecs), lo que significa que no ocupamos un lugar privilegiado en el espacio. Este principio sirve como base para poder definir la geometría y, por lo tanto, la dinámica de los objetos que se encuentran en el universo.

Partiendo de la idea de un universo homogéneo e isotrópico podemos definirlo como una serie de superficies tridimensionales ordenadas temporalmente con cierta curvatura, siendo cada una de estas superficies homogénea e isotrópica. Esta curvatura puede tomar 3 valores: curvatura cero, curvatura positiva y curvatura negativa (figura 1.1).

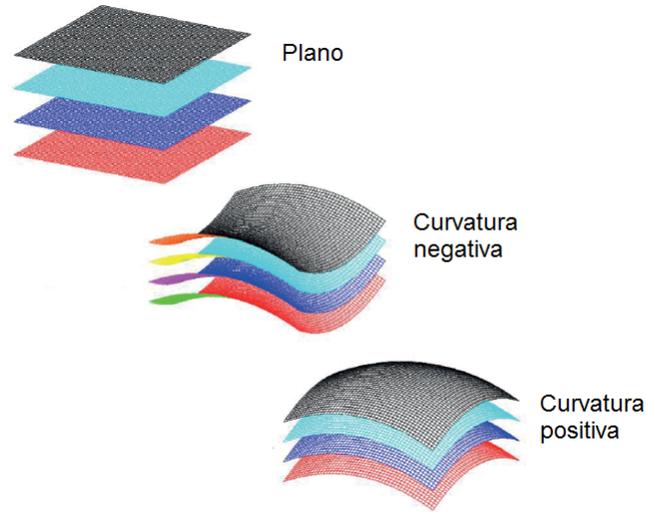


Figura 1.1: Representación visual de los tipos de curvatura.

- Curvatura cero: también llamado “espacio plano”, consiste en el espacio Euclideo donde el elemento de línea está dado por:

$$dl^2 = dx^2 = \delta_{ij} dx^i dx^j, \quad (1.1)$$

y es invariante ante traslaciones y rotaciones.

- Curvatura positiva: también llamada “geometría esférica”, es el espacio representado por una 3-esfera donde el elemento de línea viene dado por:

$$dl^2 = dx^2 + du^2, \quad x^2 + u^2 = a^2, \quad (1.2)$$

siendo a el radio de la 3-esfera. Esta geometría es homogénea e isotrópica bajo rotaciones 4-dimensionales.

- Curvatura negativa: también llamada “geometría hiperbólica”, se representa por un hiperboloide H^3 donde el elemento de línea está dado por:

$$dl^2 = dx^2 - du^2, \quad x^2 - u^2 = -a^2, \quad (1.3)$$

siendo a una constante arbitraria. La homogeneidad e isotropía vienen dadas por la simetría del elemento de línea bajo pseudo-rotaciones 4-dimensionales (transformadas de Lorentz siendo u el tiempo).

Si reescalamos $x \rightarrow xa$ y $u \rightarrow ua$ en las geometrías esférica e hiperbólica podemos entonces reescribirlas como:

$$dl^2 = a^2(dx^2 \pm du^2), \quad x^2 \pm u^2 = \pm 1, \quad (1.4)$$

y de aquí podemos ver que

$$udu = \mp \bar{x} \cdot d\bar{x}, \quad (1.5)$$

por lo que

$$dl^2 = a^2 \left(dx^2 \pm \frac{(\bar{x} \cdot d\bar{x})^2}{1 \mp x^2} \right), \quad (1.6)$$

y generalizando para el espacio plano obtenemos

$$dl^2 = a^2 \left(dx^2 + k \frac{(\bar{x} \cdot d\bar{x})^2}{1 - kx^2} \right), \quad (1.7)$$

donde $k = -1, 0, 1$. Si $k = -1$ tenemos geometría hiperbólica, si $k = 1$ es geometría esférica y con $k = 0$ es espacio plano. Debido a esto a k se le llama “curvatura”.

Reescribiéndolo en coordenadas esféricas con $dx^2 = dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)$ y por lo tanto $\bar{x} \cdot d\bar{x} = r dr$ obtenemos:

$$dl^2 = a^2 \left(\frac{(dr)^2}{1 - kr^2} + r^2 d\Omega^2 \right), \quad d\Omega^2 \equiv d\theta^2 + \sin^2\theta d\phi^2, \quad (1.8)$$

y de este resultado podemos obtener la famosa métrica de Friedmann-Lemaitre-Robertson-Walker (FLRW):

$$ds^2 = dt^2 - a^2(t) \left(\frac{(dr)^2}{1 - kr^2} + r^2 d\Omega^2 \right), \quad (1.9)$$

donde a ahora está en función del tiempo y se le llama “factor de escala” (figura 1.2). Este factor $a(t)$ nos dice qué tanto se expande (o contrae) el universo, por convención se define $a(t_0) = 1$ siendo t_0 hoy. Esta métrica nos dice cómo se comporta el elemento de línea en el espacio-tiempo. Esta métrica es importante pues es una solución exacta de las ecuaciones de Einstein de relatividad general.

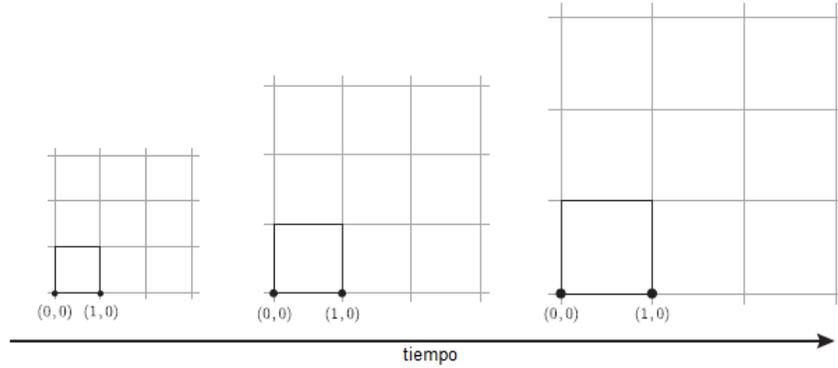


Figura 1.2: Factor de escala en una malla bidimensional. Se observa que el factor de escala aumenta mientras que la distancia comóvil (distancia entre los puntos $(0, 0)$ y $(1, 0)$) sigue siendo la misma.

De la métrica FLRW podemos empezar a definir algunas cantidades como:

- Coordenadas comóviles: $x^i \equiv [x^1, x^2, x^3]$,
- Coordenadas físicas: $x^i_{phys} = a(t)x^i$,
- Coordenadas dependientes del observador: $X^\mu = (t, x^i)$
- Parámetro de Hubble: $H \equiv \frac{\dot{a}}{a}$,
- Velocidad física:

$$v^i_{phys} \equiv \frac{dx^i_{phys}}{dt} = a(t) \frac{dx^i}{dt} + \frac{da}{dt} x^i \equiv v^i_{pec} + H x^i_{phys}, \quad (1.10)$$

- Velocidad peculiar: $v^i_{pec} \equiv a(t)\dot{x}^i$,

- Tiempo conforme: $d\tau = \frac{dt}{a(t)}$.

Ahora veremos cómo se mueven las partículas en esta métrica. Dado que se consideran en caída libre, la trayectoria que siguen es aquella descrita por una geodésica. Una geodésica es una curva que extremiza el tiempo propio $\Delta s/c$ entre dos puntos del espacio-tiempo (siendo Δs el elemento de línea). La ecuación de éstas curvas es:

$$\frac{dU^\mu}{ds} + \Gamma_{\alpha\beta}^\mu U^\alpha U^\beta = 0, \quad (1.11)$$

donde $U^\mu \equiv \frac{dX^\mu}{ds}$ es la cuadrivelocidad y $\Gamma_{\alpha\beta}^\mu$ son los símbolos de Christoffel

$$\Gamma_{\alpha\beta}^\mu \equiv \frac{1}{2}g^{\mu\lambda}(\partial_\alpha g_{\beta\lambda} + \partial_\beta g_{\alpha\lambda} - \partial_\lambda g_{\alpha\beta}), \quad \partial_\alpha \equiv \frac{\partial}{\partial X^\alpha}, \quad (1.12)$$

donde $g_{\alpha\beta}$ es la métrica, la cuál cumple $g^{\beta\lambda}g_{\lambda\alpha} = \delta_\beta^\alpha$. Si ahora utilizamos regla de la cadena

$$\frac{dU^\mu}{ds} = \frac{dX^\alpha}{ds} \frac{\partial U^\mu}{\partial X^\alpha} = U^\alpha \frac{\partial U^\mu}{\partial X^\alpha}, \quad (1.13)$$

obtenemos

$$U^\alpha \left(\frac{\partial U^\mu}{\partial X^\alpha} + \Gamma_{\alpha\beta}^\mu U^\beta \right) = 0, \quad (1.14)$$

y el término entre paréntesis es la derivada covariante de U^μ , por lo que podemos reescribir (1.14) como

$$U^\alpha \nabla_\alpha U^\mu = 0, \quad (1.15)$$

o utilizando la definición de cuadrimento $P^\mu = mU^\mu$, donde m es la masa, podemos reescribir (1.14) ahora como

$$P^\alpha \left(\frac{\partial P^\mu}{\partial X^\alpha} + \Gamma_{\alpha\beta}^\mu P^\beta \right) = 0. \quad (1.16)$$

Ahora utilizaremos la ecuación de las geodésicas en la métrica FLRW. Definiendo $\gamma_{ij} \equiv \delta_{ij} + k \frac{x_i x_j}{1 - k(x_k x^k)}$ entonces reescribimos la métrica como

$$ds^2 = dt^2 - a^2(t)\gamma_{ij}dx^i dx^j. \quad (1.17)$$

Los símbolos de Christoffel distintos de cero serán:

$$\Gamma_{ij}^0 = a\dot{a}\gamma_{ij}, \quad \Gamma_{0j}^i = \frac{\dot{a}}{a}\delta_j^i, \quad \Gamma_{jk}^i = \frac{1}{2}\gamma^{il}(\partial_j\gamma_{kl} + \partial_k\gamma_{jl} - \partial_l\gamma_{jk}), \quad (1.18)$$

y dado que FLRW supone un universo homogéneo tendremos $\partial_i P^\mu = 0$, por lo que la ecuación de la geodésica es

$$P^0 \frac{dP^\mu}{dt} = -\left(2\Gamma_{0j}^\mu P^0 + \Gamma_{ij}^\mu P^i\right) P^j. \quad (1.19)$$

1.2. Dinámica

La dinámica del universo está gobernada por la ecuación de Einstein

$$G_{\mu\nu} = 8\pi G T_{\mu\nu} \quad (1.20)$$

siendo $G_{\mu\nu}$ el tensor de Einstein que nos dice la curvatura del espacio tiempo y $T_{\mu\nu}$ el tensor de energía-impulso que nos dice el contenido de materia del universo.

Dado que el tensor de energía-impulso debe cumplir las condiciones de homogeneidad e isotropía se puede utilizar el tensor de un fluido perfecto, entonces

$$T_{\mu\nu} = (\rho + P)U_\mu U_\nu - P g_{\mu\nu} \quad (1.21)$$

$$T_{00} = \rho(t), \quad T_{i0} = 0, \quad T_{ij} = -P(t)g_{ij}(t, x) \quad (1.22)$$

$$T_\nu^\mu = g^{\mu\lambda} T_{\lambda\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & -P & 0 & 0 \\ 0 & 0 & -P & 0 \\ 0 & 0 & 0 & -P \end{pmatrix} \quad (1.23)$$

siendo ρ la densidad de energía, P la presión del fluido (no confundir con el cuadrimomento P^μ) y $g_{\alpha\beta}$ es la métrica. Pero la energía y el momento deben de conservarse, por lo que la densidad debe cumplir la ecuación de continuidad $\dot{\rho} = -\partial_i \pi^i$ (siendo $\pi^i \equiv T^{i0} = 0$) y la evolución de la densidad del momento debe obedecer la ecuación de Euler $\dot{\pi}_i = \partial_i P$. Estas dos ecuaciones de conservación pueden combinarse para obtener

$$\partial_\mu T_\nu^\mu = 0, \quad (1.24)$$

y en relatividad general se utiliza la derivada covariante, entonces

$$\nabla_\mu T_\nu^\mu = \partial_\mu T_\nu^\mu + \Gamma_{\mu\lambda}^\mu T_\nu^\lambda - \Gamma_{\mu\nu}^\lambda T_\lambda^\mu = 0, \quad (1.25)$$

la evolución de la densidad de energía corresponde a $\nu = 0$, por lo que tendremos

$$\partial_\mu T_0^\mu + \Gamma_{\mu\lambda}^\mu T_0^\lambda - \Gamma_{\mu 0}^\lambda T_\lambda^\mu = 0, \quad (1.26)$$

además tenemos, por isotropía, que $T_0^i = 0$, entonces

$$\frac{d\rho}{dt} + \Gamma_{\mu 0}^\mu \rho - \Gamma_{\mu 0}^\lambda T_\lambda^\mu = 0, \quad (1.27)$$

y usando (1.18) obtenemos la ecuación de continuidad

$$\dot{\rho} + 3\frac{\dot{a}}{a}(\rho + P) = 0. \quad (1.28)$$

Ahora debemos calcular el tensor de Einstein $G_{\mu\nu}$ de la ecuación de Einstein. De relatividad general tenemos

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}, \quad (1.29)$$

donde $R_{\mu\nu} \equiv \partial_\lambda \Gamma_{\mu\nu}^\lambda - \partial_\nu \Gamma_{\mu\lambda}^\lambda + \Gamma_{\lambda\rho}^\lambda \Gamma_{\mu\nu}^\rho - \Gamma_{\mu\lambda}^\rho \Gamma_{\nu\rho}^\lambda$ es el tensor de Ricci y $R = R^\mu_\mu = g^{\mu\nu} R_{\mu\nu}$ el escalar de Ricci. Debido a la isotropía de la métrica FLRW $R_{i0} = R_{0i}$ deben ser cero.

Entonces obtenemos

$$R_{00} = -3\frac{\ddot{a}}{a}, \quad R_{ij} = -\left(\frac{\ddot{a}}{a} + 2\left(\frac{\dot{a}}{a}\right)^2 + 2\frac{k}{a^2}\right)g_{ij}, \quad R = -6\left(\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2}\right), \quad (1.30)$$

y las componentes distintas de cero en el tensor de Einstein $G_\nu^\mu \equiv g^{\mu\lambda}G_{\lambda\nu}$

$$G_0^0 = 3\left(\left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2}\right), \quad G_j^i = \left(2\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2}\right)\delta_j^i, \quad (1.31)$$

utilizando (1.23) y (1.31) en la ecuación de Einstein obtenemos las **ecuaciones de Friedmann**

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2}, \quad (1.32)$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P). \quad (1.33)$$

Estas ecuaciones gobiernan la expansión del universo (homogéneo e isotrópico). La ecuación (1.32) por lo general se escribe en términos de $H \equiv \frac{\dot{a}}{a}$, y se toman ρ y P como las contribuciones totales de densidad y presión de todos los componentes del universo (sea materia, radiación, vacío, etc.). Se define la densidad crítica al día de hoy (subíndice 0) como $\rho_{crit,0} = \frac{3H_0^2}{8\pi G}$, y de aquí se definen los parámetros de densidad (adimensionales)

$$\Omega_{i,0} \equiv \frac{\rho_{i,0}}{\rho_{crit,0}}, \quad (1.34)$$

si le atribuimos un parámetro de densidad a la curvatura k como $\Omega_{k,0} \equiv \frac{-k}{(a_0 H_0)^2}$ podemos reescribir la primer ecuación de Friedmann como

$$H^2(a) = H_0^2 \left[\Omega_{total} + \Omega_k \left(\frac{a_0}{a}\right)^2 \right]. \quad (1.35)$$

Es momento de analizar de qué está hecho $\Omega_{total,0}$.

1.3. Menú cósmico

Empezamos con la **materia**. En cosmología se le llama materia a todo aquel componente cuyo presión sea mucho menor a su densidad, por lo que hacemos $P \approx 0$ en (1.28), y de esta forma obtenemos

$$\frac{\dot{\rho}}{\rho} = -3\frac{\dot{a}}{a}, \quad (1.36)$$

de aquí entonces

$$\rho \propto a^{-3}, \quad (1.37)$$

lo cual es esperado, pues al ser materia su densidad es inversamente proporcional al volumen del universo.

Esta densidad está principalmente compuesta de 2 tipos de materia:

- Materia bariónica: la materia observable (luminosa) como los núcleos y electrones.
- Materia oscura: un tipo de materia desconocida que no interactúa con la luz.

Veamos ahora el siguiente componente en el cosmos: **radiación**. Se le llama radiación, en cosmología, a aquello cuya presión sea un tercio de su densidad, entonces $P = \frac{1}{3}\rho$, por lo que (1.28) nos da

$$\frac{\dot{\rho}}{\rho} = -4\frac{\dot{a}}{a}, \quad (1.38)$$

entonces $\rho \propto a^{-4}$. Este es el caso para algún gas de partículas relativistas (el momento de las partículas es mucho mayor a su masa). La densidad disminuye con a^4 , esto se debe a el cambio del volumen (a^3) y a algo conocido como “corrimiento al rojo” (redshift en inglés), y para entender de donde viene la a hace falta explicar el corrimiento al rojo.

Partiendo de la **ley de Hubble**

$$v(t) = Hr(t), \quad (1.39)$$

que nos indica que los objetos más lejanos (con respecto a nosotros) se alejan con una velocidad proporcional a su distancia de nosotros, siendo $v(t)$ la velocidad de recesión de algún objeto (la velocidad con la que se aleja de nosotros), H el parámetro de Hubble y $r(t)$ la posición del objeto con respecto a nosotros. Ahora, si hacemos $dv = Hdr$, pero $H = \frac{\dot{a}}{a}$ entonces

$$dv = \frac{\dot{a}}{a} dr. \quad (1.40)$$

Pero si hablamos de la luz que nos llega del objeto entonces tenemos que tomar en cuenta el efecto Doppler, pues el objeto se está moviendo. De este efecto tenemos

$$\frac{\lambda_r - \lambda_e}{\lambda_e} = \frac{d\lambda}{\lambda_e} = \frac{dv}{c}, \quad (1.41)$$

donde c es la velocidad de la luz en el vacío, λ_r es la longitud de onda recibida por el observador y λ_e la emitida por el objeto. Ahora, dado que estamos hablando de luz, tenemos que $c = \frac{dr}{dt}$, y de todo esto obtenemos que

$$\frac{d\lambda}{\lambda_e} = \frac{\dot{a}}{a} dt = \frac{da}{a}, \quad (1.42)$$

esto implica $\lambda \propto a$. Entonces definimos **el corrimiento al rojo** z como

$$1 + z = \frac{\lambda_r}{\lambda_e} = \frac{a(t_r)}{a(t_e)} \quad (1.43)$$

y, usando la convención de que en el tiempo actual t_0 , tenemos $a(t_0) = 1$, por lo tanto

$$z = \frac{1}{a} - 1. \quad (1.44)$$

Como se observa en (1.42) tenemos que $\lambda \propto a$, y como la energía de una onda depende del inverso de su longitud de onda, al tener mayor longitud de onda la energía decrece proporcionalmente con a , lo que nos da el -4 de (1.38), 3 debido al espacio y 1 por el corrimiento al rojo. La radiación está hecha principalmente de:

- Fotones: radiación electromagnética (luz).
- Neutrinos: partículas con masas demasiado pequeñas pero con energía cinética muy grande.

Tenemos hasta ahora 2 componentes de Ω_{total} , pero existe un elemento más. Éste recibe el nombre de **energía oscura** y es el enfoque principal de este trabajo el ver cómo se comporta (o debería comportarse) la energía oscura con base en observaciones y experimentos recientes.

1.4. Energía oscura [3]

En 1917 Albert Einstein introdujo la famosa constante cosmológica Λ en la ecuación (1.20), obteniendo entonces (1.45). Esto lo hizo debido a su fuerte creencia en que el universo es estático, mientras que su ecuación original (antes de agregar Λ) predecía un universo que variaba en tamaño [4]. Era una propuesta interesante hasta que, en 1929, Edwin Hubble descubrió que el universo no se encuentra estático si no que se está expandiendo.

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu} \quad (1.45)$$

Aunque la inclusión de esta constante fue considerada como un error se seguía incluyendo en la ecuación pero con un valor de cero. Pero luego del descubrimiento de la expansión acelerada del universo en 1998 se encontró que Λ debe tener un valor positivo. Este podría considerarse como el nacimiento de la energía oscura en la cosmología [5].

A veces se le llama también “energía del vacío” en relatividad general pues, según la hipótesis de su existencia, permea todo el espacio de manera uniforme. Aunque tiene una densidad bastante pequeña (aproximadamente $7 \times 10^{-30} g/cm^3$), según observaciones recientes la energía oscura contribuye aproximadamente en un 68.3 % a la energía total del universo en el modelo Λ CDM [16]. Mientras que la materia oscura y la materia bariónica contribuyen con el 26.8 % y 4.9 % respectivamente [16]. Si tomamos en cuenta un parámetro

de densidad para la energía oscura tendremos entonces 3 componentes de Ω_{total} , lo que nos da una ecuación de Friedmann

$$H^2(a) = H_0^2 \left[\Omega_{r,0} \left(\frac{a_0}{a} \right)^4 + \Omega_{m,0} \left(\frac{a_0}{a} \right)^3 + \Omega_{DE}(a) + \Omega_{k,0} \left(\frac{a_0}{a} \right)^2 \right]. \quad (1.46)$$

siendo $\Omega_{r,0}$ el parámetro de densidad de radiación, $\Omega_{m,0}$ el de materia y $\Omega_{DE}(a)$ el de la energía oscura. En este caso tenemos

$$\Omega_{DE}(a) = \Omega_{DE,0} e^{-3 \int_1^a \frac{da'}{a'} (1+w)}, \quad (1.47)$$

donde w es el parámetro de la ecuación de estado (a veces simplemente llamado ecuación de estado) de la energía oscura, surge al considerar a la energía oscura como un fluido perfecto cuya densidad y presión están relacionadas como

$$w \equiv P_{DE} / \rho_{DE}, \quad (1.48)$$

siendo P_{DE} la presión de la energía oscura y ρ_{DE} su densidad, esta ecuación luego se introduce en (1.28) y de ahí obtenemos (1.47).

A pesar de que no se sabe qué es la energía oscura, existe evidencia bastante fuerte de su existencia. Debido a esto, los modelos cosmológicos actuales incluyen energía oscura y un intento por explicarla. Algunos experimentos en los que se observa la acción de la energía oscura son: observaciones de supernovas tipo 1a, medición del CMB, la medición de las oscilaciones acústicas de bariones (BAOs por sus siglas en inglés) y la medición del parámetro de Hubble ($H(z)$).

1.4.1. Supernovas tipo 1a

Las supernovas tipo 1a son un tipo especial de supernova que ocurre en un sistema binario entre una enana blanca y cualquier otro tipo de estrella [6]. La enana blanca es uno de los objetos más densos del universo, por lo que su gravedad es bastante fuerte. Debido a esto la enana blanca comienza a atraer material de la otra estrella, aumentando de esta

forma su masa. Al obtener aproximadamente 1.4 masas solares la enana blanca comenzará a sufrir una reacción nuclear en cadena en su interior, provocando de esta manera la explosión de la enana blanca. Esta reacción nuclear es bastante parecida entre todos los sistemas binarios de este tipo, por lo que la explosión resultante de la reacción nuclear es prácticamente la misma, y esto significa que este tipo de supernovas tienen una curva de luz igual (figura 1.3). Esto las convierte en un excelente candidato a ser una candela estándar en astronomía.

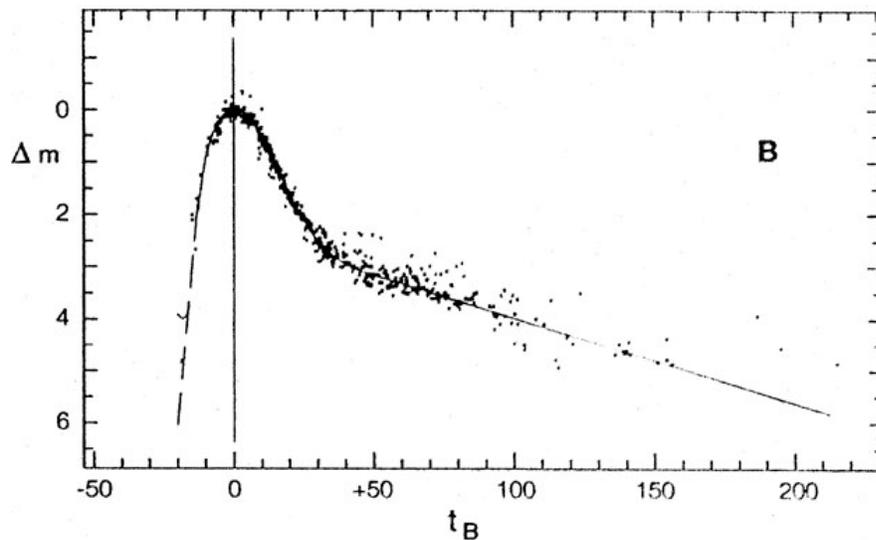


Figura 1.3: Curva de luz para 22 supernovas tipo 1a, siendo Δm el brillo y t_B el tiempo desde que ocurre la explosión, la B indica que se midió la luz azul. Fuente: https://ned.ipac.caltech.edu/level5/Branch2/Branch2_1.html.

Estas supernovas se utilizan para poder medir la velocidad y distancia de objetos a distancias del orden de megaparsecs debido a su enorme (y consistente) luminosidad. Para la velocidad todo lo que se necesita es comparar la luminosidad que se sabe tiene la supernova tipo 1a con la observada y calcular el corrimiento al rojo, y para la distancia se mide el brillo.

En 1998 se realizó el descubrimiento de la expansión acelerada del universo observando

supernovas tipo 1a [6]. Con este descubrimiento se inició el estudio de aquello que provoca esta aceleración, dando el inicio del estudio de la energía oscura.

1.4.2. CMB

Holmdel, Nueva Jersey, año 1964, Arno Penzias y Robert Wilson se encuentran en los laboratorios Bell experimentando con una antena de bocina de 6 metros diseñada para detectar ondas de radio provenientes del rebote de los satélites Echo de la NASA. Las ondas provenientes de estos satélites eran bastante débiles. Si querían detectarlas entonces debían remover toda la interferencia que pudiera afectar a la antena, como las ondas de radio provenientes de las estaciones de radio cercanas, de radares y hasta se eliminó la interferencia del mismo receptor de ondas enfriándolo con helio líquido. Incluso se removieron los nidos de palomas y se limpiaron las antenas de cualquier desperdicio originado por las mismas.

Aún después de todos sus esfuerzos para eliminar toda la interferencia existía un ruido que no podían eliminar, un ruido pequeño y constante que permanecía en el receptor. Este ruido estaba presente en todas las direcciones y a todas horas. Tanto Penzias como Wilson concluyeron que este ruido, correspondiente a una onda de 7.35 centímetros, no provenía de ninguna fuente terrestre, ni tampoco de algo dentro de nuestra galaxia, el ruido debía provenir de afuera de la Vía Láctea. Después se descubrió que este ruido era causado por la radiación de fondo de microondas (CMB por sus siglas en inglés), un tipo de radiación cuyo origen se le atribuye al Big Bang [9].

Esta radiación se encuentra en todo el espacio y es la radiación electromagnética más antigua del universo debido a que es un remanente del Big Bang. Es casi isotrópica con una temperatura de 2.72 grados Kelvin. El CMB es una de las evidencias más fuertes que se tiene sobre el Big Bang, pues de acuerdo a la teoría hubo un tiempo en el universo temprano conocido como “época de recombinación” en el que el universo pasó de ser un plasma de hidrógeno a un universo más frío, lo que permitió a los protones y electrones combinarse para formar hidrógeno y, debido a que estos átomos no absorbían la radiación

térmica, los fotones podían viajar a través del espacio, en lugar de ser dispersados por el plasma. Estos fotones siguen viajando y, dado que el universo está expandiéndose, tienen cada vez menos energía, pues su longitud de onda es cada vez mayor.

Si se miden las anisotropías en esta radiación de fondo se puede encontrar que la geometría del universo es casi plana. Para que el universo tenga una geometría plana su densidad de materia y energía debe ser la densidad crítica. Las mediciones actuales del CMB realizadas por el WMAP (Wilkinson Microwave Anisotropy Probe) y por el telescopio Planck (figura 1.4), al ser combinadas con las observaciones de supernovas tipo 1a realizadas por las colaboraciones de Supernova Cosmology Project y High-z Supernova Search Team nos indican que la materia tanto oscura como bariónica conforman aproximadamente 30 % de la densidad crítica, mientras que el otro 70 % se debe a la energía oscura [8] [16].

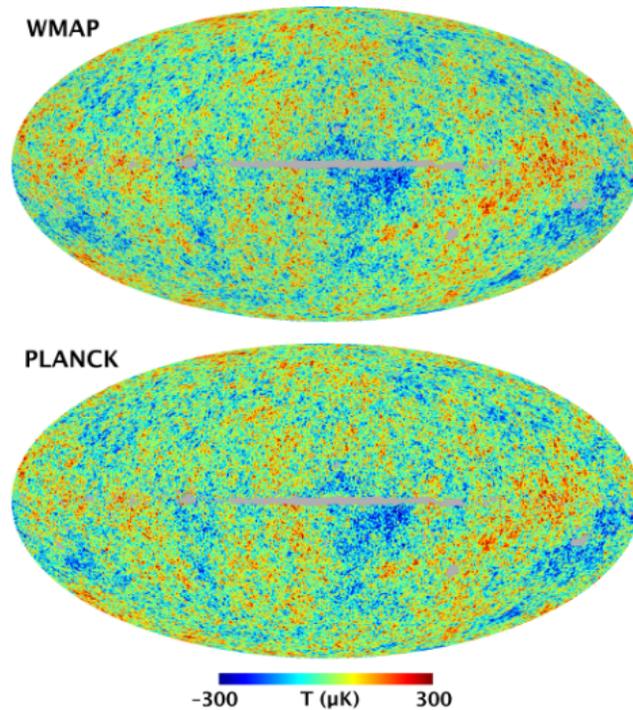


Figura 1.4: Comparación entre mapeos del CMB de WMAP año 9 y el telescopio Planck. Son prácticamente idénticos, por lo que se cree que son mediciones bastante confiables. Fuente: https://lambda.gsfc.nasa.gov/product/map/dr5/m_images.cfm.

1.4.3. Oscilaciones acústicas de bariones (BAOs)

Las BAOs son, como su nombre lo indica, oscilaciones que ocurrieron debido a perturbaciones de sobredensidades de materia en el universo temprano (antes de la época de recombinación [10]). Dichas densidades se debieron al colapso gravitacional de materia. Las perturbaciones en estas sobredensidades se deben a la presión ejercida por la radiación, por lo que tenemos dos fuerzas oponiéndose entre sí: la atracción gravitacional de las sobredensidades de materia y la ejercida por la presión de radiación. Estas dos fuerzas crearon oscilaciones de bariones y fotones. Luego de la época de recombinación los fotones se desacoplaron de los bariones y se liberó la presión ejercida por la radiación [10] [11], y esto dejó un cascarón de bariones a una cierta distancia de las sobredensidades de materia, a esta distancia se le conoce como “horizonte de sonido”.

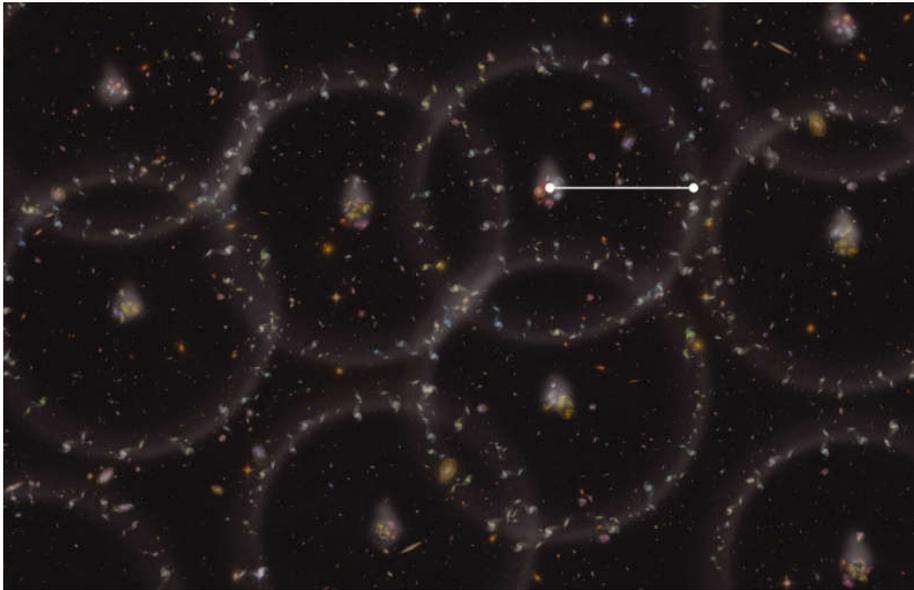


Figura 1.5: Representación de las BAOs. En realidad no se ven de esta manera, pero esta figura nos da una buena idea de qué son. Fuente: <http://www.astronet.ru/db/msg/1298862>

Las BAOs son utilizadas como “reglas estándar”, así como las supernovas son usadas como candelas estándar.

1.4.4. Parámetro de Hubble

El parámetro de Hubble ($H(z)$) nos dice la rapidez con la que se expande el universo y puede ponerse en función del corrimiento al rojo z . La medición de $H(z)$ nos dice la historia de la expansión cosmológica, por lo que es muy usado para examinar la expansión acelerada del universo y, por lo tanto, las propiedades de la energía oscura. Para la medición de este parámetro se utilizan los llamados *cronómetros cósmicos* [12] [13]. Estos cronómetros consisten en la observación de estrellas viejas formadas en sincronía y funcionan como “relojes estándar”.

Además, el parámetro de Hubble aparece en una gran cantidad de propiedades del universo, pero de esto se hará mención más adelante.

1.4.5. Modelos de energía oscura

Como ya se mencionó, la energía oscura es lo más abundante en el universo actualmente, ocupando más de dos terceras partes del mismo [16]. A pesar de esto, no se sabe exactamente qué es la energía oscura, pero se sabe que debe cumplir ciertas características. Debido a esto se han realizado muchos esfuerzos por crear un modelo que pueda describirla. Algunos de éstos son:

- Constante cosmológica (mejor conocido como Λ CDM)
- Quintesencia [18]
- Energía phantom [18]
- Quintom [19]
- Gravedad modificada [28].

Estos son sólo algunos de los modelos que se estudian e intentan explicar el universo. De todos ellos, el modelo Λ CDM es el considerado “modelo estándar” en cosmología.

El modelo Λ CDM

Se le conoce como Λ CDM pues se considera que la energía oscura es una constante cosmológica (Λ) y el CDM son siglas en inglés para *materia oscura fría*, esto significa que la materia oscura en este modelo sólo interactúa de manera gravitacional. Dado que la energía oscura es constante entonces su ecuación de estado debe ser constante de valor -1 , por lo que (1.48) sería:

$$w \equiv P_{DE}/\rho_{DE} = -1. \quad (1.49)$$

Esto es lo mismo que decir que la energía oscura tiene una presión negativa, lo que provoca la expansión del universo, y suele referirse a esta presión como la presión del vacío. Al utilizar $w = -1$ en (1.47) obtenemos la ecuación de Friedmann para Λ CDM

$$H^2(a) = H_0^2 \left[\Omega_{r,0} \left(\frac{a_0}{a} \right)^4 + \Omega_{m,0} \left(\frac{a_0}{a} \right)^3 + \Omega_{DE,0} + \Omega_{k,0} \left(\frac{a_0}{a} \right)^2 \right]. \quad (1.50)$$

Este modelo se considera el estándar pues puede predecir muchas observaciones y es el más sencillo (es el que tiene menos parámetros libres), pero recientemente ha tenido algunos problemas y se han presentado ciertas tensiones en las observaciones [20]. Los dos principales problemas son el **problema de la coincidencia** [21] [22] y el **problema del ajuste fino** [23]. El problema de la coincidencia a veces no es considerado como un problema físico, más bien filosófico, y plantea la pregunta: ¿Porqué el universo se está expandiendo de manera acelerada en este momento del tiempo?. El problema del ajuste fino tiene que ver con el valor preciso que deben tener los parámetros del modelo para que el universo pueda funcionar como lo hace. Además de los problemas ya mencionados también hay muchos problemas a escalas pequeñas [15].

Esta gran cantidad de problemas con Λ CDM han motivado a los cosmólogos a intentar buscar distintas alternativas para explicar la dinámica del universo, y para poner a prueba dichas alternativas se deben utilizar los experimentos y observaciones que se han realizado, aquí es donde entran las llamadas **reconstrucciones**.

Capítulo 2

Reconstrucciones de la energía oscura

No se tiene aún ningún modelo que explique de manera satisfactoria la expansión acelerada del universo, y por lo tanto no se tiene una descripción concluyente sobre la energía oscura. Entonces es bastante común intentar utilizar los datos y observaciones provenientes de varios experimentos y sondeos para intentar reconstruir, ya sea de forma paramétrica o no paramétrica, algunas cantidades o ecuaciones que nos den alguna pista sobre el comportamiento de la energía oscura (como su densidad o la ecuación de estado).

A continuación se hablará sobre algunas reconstrucciones.

2.1. Reconstrucciones paramétricas

Una parametrización bastante conocida consiste en, asumiendo una energía oscura que sea espacialmente suave, que la ecuación de estado (1.48) sea una constante [28]. Esta parametrización describe la energía del vacío ($w = -1$) y los defectos topológicos ($w = -N/3$, siendo N la dimensión del defecto, cuyo valor es 0 para monopolos, 1 para cuerdas cósmicas y 2 para texturas). Con w y Ω_{DE} tenemos entonces una descripción de la energía oscura que depende únicamente de dos parámetros. Pero esta parametrización no describe la evolución temporal de w , y aquí es donde entran otro tipo de modelos y parametrizaciones.

Si se propone que $w = w(z)$, entonces por (1.47) tenemos que $\Omega_{DE} = \Omega_{DE}(z)$, siendo z

el corrimiento al rojo (aunque también podría ser función del factor de escala a), se pueden proponer una gran cantidad de parametrizaciones de w , por ejemplo una de las primeras consistió en una evolución lineal en z ($w(z) = w_0 + w_1 z$) y daba una buena descripción de la energía oscura para corrimientos al rojo bajos, pero al incrementar z daba resultados inconsistentes. Entonces se propuso

$$w(a) = w_0 + w_a(1 - a) = w_0 + w_a \frac{z}{1 + z}, \quad (2.1)$$

siendo a el factor de escala, w_0 es la ecuación de estado evaluada al día de hoy y w_a es la primer derivada de $w(z)$ evaluada al día de hoy. Esta parametrización se conoce como CPL [26, 25] y evita el problema antes mencionado y se ajusta a algunos modelos de campo escalar y a algunas descripciones de la expansión en gravedad modificada. De esta parametrización surge una descripción bastante popular de la energía oscura, una que es descrita por tres parámetros Ω_{DE}, w_0, w_a . Lo que nos da una densidad de energía

$$\frac{\rho_{de}}{\rho_{crit,0}} = \Omega_{DE} a^{-3(1+w_0+w_a)} e^{-3w_a(1-a)}. \quad (2.2)$$

Esta densidad luego es sustituida en las ecs. de Friedmann y se comparan con datos observacionales para obtener los valores de los parámetros implicados en el modelo. La parametrización CPL corresponde a la serie de Taylor

$$w_a = \sum_{i=0}^N (1 - a)^i w_i, \quad (2.3)$$

para $N = 1$.

Otra parametrización popular ([24]) es una serie de Taylor para el parámetro de densidad como $\Omega_{DE} = A_1 + A_2 x + A_3 x^2$ con $x = z + 1$, por lo que obtenemos una ec. de Friedmann

$$H(x) = H_0 (\Omega_m x^3 + A_1 + A_2 x + A_3 x^2)^{\frac{1}{2}}, \quad (2.4)$$

y se pretende buscar el valor de los coeficientes de Taylor A_i .

También se puede expandir en series de Taylor el parámetro de la ecuación de estado ([24])

$$w(z) = \sum_{i=0}^N w_i z^i, \quad (2.5)$$

y se encuentra el valor de los coeficientes w_i .

Se tienen además algunos modelos de una ecuación de estado que decae ([32]), proponiendo una parametrización parecida a

$$w(z) = w_i + \frac{w_f - w_i}{1 + e^{\frac{z - z_t}{\Delta}}}, \quad (2.6)$$

donde w_i es el valor de la ecuación de estado a grandes z , z_t es el valor donde $w(z) = (w_i + w_f)/2$ y Δ es la tasa de cambio de $w(z)$.

Y también existen modelos donde la ecuación de estado puede oscilar ([31]). Un ejemplo de una parametrización para una ecuación de estado oscilatoria es

$$w(\log a) = w_0 + w_1 \cos\left(A \log \frac{a}{a_c}\right), \quad (2.7)$$

donde se tienen w_0 , w_1 , A y a_c como parámetros.

Así como todas las mencionadas en esta sección existe una cantidad enorme de posibles parametrizaciones, por lo que las reconstrucciones paramétricas no tienen un enfoque muy fijo pues en principio se puede proponer cualquier forma funcional de $w(z)$. Debido a esto el enfoque de este trabajo será en otro tipo de reconstrucciones.

2.2. PCA

El análisis de componentes principales (PCA por sus siglas en inglés) es una técnica que se utiliza para poder visualizar mejor los datos de algún experimento. Esta técnica consiste en usar una matriz de rotación que está conformada por los eigenvalores de la matriz de covarianza de los datos en cuestión para, de esta forma, escribir los datos en una

nueva “base”. Al representar los datos en esta nueva base podemos reducir la cantidad de dimensiones de la nueva gráfica al omitir los datos “menos relevantes” usando el criterio de los eigenvalores, aunque al hacer esto se puede perder información, pero por lo general es un precio bajo a pagar para poder reducir las dimensiones del problema.

Debido a que esta técnica utiliza la covarianza como criterio es necesario que los datos estén estandarizados, es decir: que tengan un promedio 0 y varianza 1. Esto es necesario pues se comparan covarianzas para obtener los llamados “componentes principales”, y dada la naturaleza de algunos datos se puede malinterpretar la covarianza entre ambas al compararlas (como comparar una covarianza entre la masa de un grupo de personas y su estatura, no tienen la misma relevancia una covarianza de 1 kilo contra una covarianza de 1 metro).

Ésta técnica es bastante útil pero tiene ciertas desventajas como:

- No es invariante de escala.
- Se asume que las direcciones de mayor varianza son las de mayor interés.
- Sólo considera transformaciones ortogonales de las variables.
- Se deben estandarizar los datos pues se asume una distribución normal, y no todas las distribuciones cumplen esto.
- La reducción dimensional ocurre cuando hay correlación entre variables, si no la hay entonces sólo se acomodan los datos de acuerdo a su varianza.

A pesar de las desventajas antes mencionadas esta técnica se puede utilizar para encontrar relaciones no tan obvias entre distintos elementos de una medición y en algunos casos hasta puede decirnos alguna dirección preferencial sobre la dinámica del sistema.

Si queremos utilizar esta técnica con la ecuación de estado [29] se tendría que encontrar los eigenvectores $e_i(z)$ de la matriz inversa de la matriz de covarianza para los parámetros w_i ($i = 1, \dots, N$) y cada eigenvalor correspondiente λ_i , de forma que la parametrización de $w(z)$ sería de la forma

$$w(z) + 1 = \sum_{i=1}^N \alpha_i e_i, \quad (2.8)$$

siendo e_i los componentes principales y los coeficientes α_i se obtienen de la relación de ortonormalidad

$$\alpha_i = \int (1 + w(z)) e_i(z) dz, \quad (2.9)$$

y tienen una exactitud $1/\sqrt{\lambda_i}$. Este método tiene varias ventajas, algunas de ellas son:

- 1.- El método no presupone modelos ni asume dependencia temporal a priori debido a que no hay parametrizaciones arbitrarias, básicamente los datos nos están diciendo “qué estamos midiendo” y “qué tan bien estamos midiéndolo”.
- 2.- Se puede utilizar este método para optimizar mediciones, por ejemplo diseñando sondeos que sean más sensibles al parámetro de la ecuación de estado de energía oscura en un cierto intervalo de z .
- 3.- Los PCs pueden hacer posible el cuantificar qué tantos parámetros individuales pueden ser medidos dada una combinación de sondeos cosmológicos.

2.3. Reconstrucciones no paramétricas

Como se vió anteriormente, las reconstrucciones paramétricas le asignan una forma funcional a la propiedad que se desea reconstruir (como el parámetro de densidad o la ecuación de estado), pero existe otro enfoque conocido como *no paramétrico* donde se utilizan otro tipo de métodos para intentar encontrar la forma que tendría la propiedad a reconstruir.

Un ejemplo de esto es la reconstrucción utilizada por Daly y Djorgovsky [34], que consiste en reconstruir la distancia comóvil $r_i(z) = r(z_i) + n_i$, donde n_i es el término del ruido y $r(z)$ está relacionado con el modulo μ_0 de las supernovas como $r(z)/1Mpc =$

$10^{\frac{\mu_0-25}{5}}/2997.5(1+z)$, en trozos utilizando una serie de funciones base, después se utiliza $y = H_0 r$ para reconstruir el parámetro de desaceleración utilizando

$$\frac{\ddot{a}a}{\dot{a}^2} = a + (1+z) \frac{d^2 y/dz^2}{dy/dz}. \quad (2.10)$$

También está la reconstrucción discutida por Shafieloo [35], aquí se construye una cantidad $D^s(x)$ utilizando un filtro F y una cantidad “cruda” $D(x')$, además se utiliza una escala de suavizado Δ

$$D^s(x, \Delta) = \int D(x') F(|x - x'|, \Delta) dx'. \quad (2.11)$$

La D podría ser alguna cantidad cosmológica como $D_L(z)$, $D_A(z)$ o $r(z)$. Algunos filtros utilizados comúnmente son el filtro de corte

$$F_{TH} \propto \Theta\left(1 - \frac{|x - x'|}{\Delta}\right), \quad (2.12)$$

siendo Θ la función escalón de Heaviside, y el filtro Gaussiano

$$F_G \propto e^{-\frac{|x-x'|^2}{2\Delta^2}}. \quad (2.13)$$

También está la llamada *reconstrucción directa* [28]. Se pueden utilizar los datos directamente de las mediciones de la distancia comóvil $r(z)$ (figura 1.2) en un cierto intervalo de corrimientos al rojo para poder obtener de forma directa ρ_{DE} y $w(z) = cte$ usando [28]

$$\rho_{DE}(z) = \rho_{DE,0}(1+z)^{3(1+w)}, \quad (2.14)$$

$$r(z) = \lim_{\Omega'_k \rightarrow \Omega_k} \frac{c}{H_0 \sqrt{\Omega_k}} \sinh\left(\sqrt{\Omega'_k} \int_0^z \frac{H_0}{H(z')} dz'\right), \quad (2.15)$$

y de aquí podemos recuperar la ecuación de estado

$$w(z) + 1 = \frac{1+z}{3} \frac{3H_0^2 \Omega_m (1+z)^2 + 2\left(\frac{d^2 r}{dz^2}\right)\left(\frac{dr}{dz}\right)^{-3}}{H_0^2 \Omega_m (1+z)^3 - \left(\frac{dr}{dz}\right)^{-2}}. \quad (2.16)$$

Pero $w(z)$ depende de la segunda derivada de $r(z)$, por lo que los errores al medir la distancia comóvil se propagaran bastante, por esto mismo se recurre regularmente a reconstruir ρ_{DE} pues depende únicamente de la primer derivada de $r(z)$

$$\rho_{DE} = \frac{3}{8\pi G} \left(\frac{1}{\left(\frac{dr}{dz}\right)^2} - \Omega_m H_0^2 (1+z)^3 \right). \quad (2.17)$$

Reconstruir ρ_{DE} , aunque es fácil, no nos da mucha información sobre la naturaleza de la energía oscura, por lo que es preferible tratar de reconstruir $w(z)$, pero, aún con mediciones de gran calidad, la reconstrucción directa no es lo suficientemente buena para darnos el comportamiento de $w(z)$, pero se han usado tratamientos estadísticos avanzados para obtener una idea de cómo se comporta.

Otra cantidad que es popular en reconstrucción es el parámetro de Hubble $H(z)$, pues al reconstruirlo se pueden recuperar una gran cantidad de propiedades físicas del universo. Éstas son:

- La edad del universo

$$t(z) = \int_z^\infty \frac{dz'}{(1+z')H(z')}, \quad (2.18)$$

- El parámetro de desaceleración

$$q \equiv \frac{h'(x)}{H(x)} x - 1, x = 1 + z, \quad (2.19)$$

- La profundidad óptica de dispersión de electrones para un corrimiento al rojo de reionización z_{reion}

$$\tau(z_{reion}) = c \int_0^{z_{reion}} \frac{n_e(z) \sigma_T dz}{(1+z)H(z)}, \quad (2.20)$$

donde n_e es la densidad de electrones y σ_T es la sección eficaz de Thomson que describe la dispersión entre electrones y fotones del CMB.

- El parámetro A asociado con el pico acústico de bariones

$$A = \frac{\sqrt{\Omega_{m0}}}{h(z_1)^{1/3}} \left[\frac{1}{z_1} \int_0^{z_1} \frac{dz}{h(z)} \right]^{2/3} = 0.469 \pm 0.017, \quad (2.21)$$

siendo $h(z) = H(z)/H_0$ y $z_1 = 0.35$.

- El parámetro R asociado con el pico acústico en las anisotropías del espectro de potencias del CMB, conocido como parametro *shift*

$$R = \sqrt{\Omega_{m0}} \int_0^{z_{lss}} \frac{dz}{h(z)} = 1.7 \pm 0.03. \quad (2.22)$$

- En quintesencia se puede reconstruir el potencial dado que se tienen las ecs. de Einstein

$$H^2 = \frac{8}{3}\pi G \left(\rho_m + \frac{1}{2} \left(\frac{d\phi}{dt} \right)^2 + V(\phi) \right), \quad (2.23)$$

$$\frac{dH}{dt} = -4\pi G \left(\rho_m + \frac{d\phi}{dt} \right)^2, \quad (2.24)$$

que pueden ser reescritas como [\[33\]](#)

$$\frac{8\pi G}{3H_0^2} V(x) = \frac{H^2}{H_0^2} - \frac{x}{6H_0^2} \frac{dH^2}{dx} - \frac{1}{2} \Omega_{m0} x^3, \quad (2.25)$$

$$\frac{8\pi G}{3H_0^2} \left(\frac{d\phi}{dx} \right)^2 = \frac{2}{3H_0^2 x} \frac{d \ln H}{dx} - \frac{\Omega_{m0} x}{H^2}, \quad (2.26)$$

con $x \equiv 1 + z$. De aquí podemos ver que si se integra (2.26) se puede obtener $\phi(z)$, y al obtener $\phi(z)$ se puede obtener $z(\phi)$, lo que podemos sustituir en (2.25) para determinar $V(\phi)$. Aquí se tiene un pequeño inconveniente debido a la presencia de Ω_{m0} , pues esto indica que se debe conocer bastante bien el valor de la densidad de materia para poder reconstruir de manera precisa $V(\phi)$.

Las reconstrucciones directas son (al menos conceptualmente) independientes de suposiciones de algún modelo debido a que no se hace ninguna suposición acerca de la forma funcional de la variación temporal. Pero existe un problema, aunque no se hagan suposiciones sobre algún modelo en específico se tiene una primera derivada de r en (2.17) si se quiere reconstruir ρ_{DE} (y potencialmente también la segunda derivada de r si se desea reconstruir $w(z)$ en (2.16)), por lo que se debe realizar un ajuste sobre los datos medidos

para obtener una función suave que pueda ser diferenciable. Este ajuste provoca que sí se tenga un “modelo”, pues se tendrá que elegir un ajuste de acuerdo a la preferencia de la persona. Entonces las reconstrucciones directas no están completamente libres de modelos.

Existen otros tipos de reconstrucciones no paramétricas pero lo importante es la diferencia entre las reconstrucciones paramétricas y las no paramétricas: las paramétricas asumen una forma funcional de la propiedad a reconstruir mientras que las no paramétricas utilizan los datos para intentar encontrar la “forma” de dicha propiedad.

En este trabajo se utilizó un tipo de reconstrucción no paramétrica para reconstruir la ecuación de estado de la energía oscura, pero antes de explicar esta reconstrucción se dará un breve repaso de estadística Bayesiana y del método MCMC.

Capítulo 3

Estadística Bayesiana y selección de modelos

Para entender cómo se hará la reconstrucción se necesita entender un poco de estadística Bayesiana. En específico se necesita saber sobre cadenas de Markov y sobre el método de Cadenas de Markov Monte Carlo (MCMC por sus siglas en inglés), por lo que se dará una breve introducción a la estadística Bayesiana y sus conceptos antes de discutir MCMC. Gran parte este capítulo está basado en las referencias [36] y [37].

3.1. Conceptos, definiciones y ejemplos

En estadística Bayesiana las probabilidades se encuentran en la mente, no en el mundo. Esta frase nos da una pequeña noción sobre el cómo se trabaja con estadística Bayesiana. Por ejemplo, supongamos que algún participante de un programa de televisión debe contestar una pregunta de opción múltiple, si hay 4 opciones, digamos que son incisos a, b, c y d, y la persona no tiene idea de cuál sea la respuesta correcta, entonces esa persona tiene $1/4$ de probabilidades de escoger la opción correcta si elige al azar. Ahora supongamos que se le permite pedir ayuda a otra persona para responder la pregunta, si esta persona le dice “creo que es el inciso a ” entonces el participante podría inclinarse en creer que el inciso

a es la respuesta correcta, lo que haría que la probabilidad de a aumente en la mente del participante. Si dejan que pida ayuda a una segunda persona y ésta le responde también que cree que es el inciso a entonces el participante estará aún más seguro de que a es la respuesta correcta.

Analizemos este ejemplo que acabamos de plantear. El participante tenía 4 opciones iniciales pero no estaba seguro de cuál era la correcta, al recibir ayuda de la primera persona recibió nueva información sobre el problema, lo que causó una inclinación hacia una de las 4 opciones, aumentando su probabilidad. Al hacerle la pregunta a la segunda persona y responder lo mismo que la primera entonces aumentó aún más la probabilidad de que a fuese la respuesta correcta. Este ejemplo nos da la noción de un concepto importante para entender la estadística bayesiana: **al recibir nueva información sobre algún fenómeno para el que se tienen varias hipótesis que intenten explicar su comportamiento, cada una de ellas con su propia probabilidad de ser cierta, se deben actualizar las probabilidades de que alguna de estas hipótesis sea la correcta con base en la información que se recibió.** En el ejemplo de arriba las hipótesis son las 4 posibles respuestas, el fenómeno es la pregunta y la nueva información recibida es la respuesta de las 2 personas a las que se les pidió ayuda. Al pedirles ayuda se actualizaron las probabilidades de cada una de las 4 posibles respuestas pues la “nueva información” indicaba que a tenía mayor probabilidad de ser correcta que las otras 3.

Ahora veamos otro ejemplo en el que emplearemos números para poder dejar en claro la base de toda la estadística Bayesiana: la **regla de Bayes**. Supongamos que tenemos una bolsa con 2 bolas dentro, una de ellas estamos completamente seguros de que es negra, pero la otra podría ser negra (N) o blanca (B). Esta situación nos presenta dos casos a los que llamaremos **hipótesis** NN (dos bolas negras) y NB (una bola negra y una blanca). En principio ambos casos son igualmente probables, por lo que su probabilidad inicial es 0.5. Si lo acomodamos en una **caja de Bayes** (cuadro 3.1) quedará mejor representado.

Nuestras hipótesis tienen una cierta probabilidad inicial (anterior) que puede ser escogida dependiendo de las circunstancias. Si ahora procedemos a sacar una bola de la bolsa y

Hipótesis	anterior	verosimilitud	h	posterior
NN	0.5			
NB	0.5			
Total:	1			

Cuadro 3.1: Caja de Bayes donde se observa la probabilidad anterior de ambas hipótesis. Se tiene que $h = anterior \times verosimilitud$. Por el momento sólo se conoce la probabilidad anterior de ambas hipótesis.

resulta ser negra podremos empezar a calcular la *verosimilitud* (likelihood en inglés), que es un tipo muy especial de probabilidad. La verosimilitud se define como: probabilidad de haber calculado los datos suponiendo que la hipótesis fuese verdad. En este caso los “datos calculados” es el haber sacado una bola negra de la bolsa, si NN fuese la hipótesis correcta entonces la probabilidad de sacar una bola negra es 1, mientras que en NB es 0.5, lo que nos permite obtener tanto la verosimilitud como h (cuadro 3.2).

Hipótesis	anterior	verosimilitud	h	posterior
NN	0.5	1	0.5	
NB	0.5	0.5	0.25	
Total:	1		0.75	

Cuadro 3.2: Caja de Bayes donde sólo falta llenar la columna con la *probabilidad posterior*. Dado que se tiene nueva información se pueden actualizar las probabilidades, por lo que podemos llenar la columna de *verosimilitud* y de h .

Pero aún nos falta calcular el *posterior*. El *posterior* es la “nueva probabilidad” dados los datos o información nueva y no es más que obtener la proporción de la *verosimilitud* de cada hipótesis con la suma de todas las verosimilitudes (básicamente lo normalizamos para obtener nuevas probabilidades que sumen 1). Entonces nuestra caja de Bayes está

terminada como se observa en el cuadro 3.3.

Hipótesis	anterior	verosimilitud	h	posterior
NN	0.5	1	0.5	0.667
NB	0.5	0.5	0.25	0.333
Total:	1		0.75	1

Cuadro 3.3: Caja de Bayes completa para dos hipótesis. Se observa que ahora NN tiene mayor probabilidad que NB pues hay mayor probabilidad de sacar una bola negra en NN en NB, haciendo que NN sea la hipótesis más probable luego de realizar la “medición”.

Este ejemplo nos ayuda a sentar las bases para definir la **regla de Bayes**:

$$P(H|D) = \frac{P(D|H)P(H)}{P(D)}, \quad (3.1)$$

siendo H la hipótesis, D los datos y:

- $P(H|D)$: probabilidad posterior. Qué tan probable es que H sea cierto dado que obtenemos D .
- $P(H)$: probabilidad anterior. Qué tan probable es que H sea cierto antes de haber observado D .
- $P(D|H)$: verosimilitud. Qué tan probable es que observemos D si H es cierto.
- $P(D)$: verosimilitud marginal. Probabilidad de observar D sin saber si H es verdad o no.

y si generalizamos para N hipótesis:

$$P(H_i|D) = \frac{P(D|H_i)P(H_i)}{P(D)}, \quad i = 1, 2, \dots, N \quad (3.2)$$

$$P(D) = \sum_{i=1}^N P(H_i)P(D|H_i), \quad (3.3)$$

y esto podemos ponerlo en una caja de Bayes como se observa en el cuadro 3.4.

Hipótesis	anterior	verosimilitud	h	posterior
H_1	$P(H_1)$	$P(D H_1)$	$P(H_1) \times P(D H_1)$	$P(H_1 D)$
H_2	$P(H_2)$	$P(D H_2)$	$P(H_2) \times P(D H_2)$	$P(H_2 D)$
...
Total:	1		$P(D)$	1

Cuadro 3.4: Caja de Bayes para N hipótesis.

3.2. Estimación de parámetros

La estadística Bayesiana también se puede utilizar para la estimación de parámetros de ciertas distribuciones. Para demostrarlo haremos un ejemplo.

Supongamos que un estudiante acaba de cambiarse de estado y está averiguando cómo ir a su nueva universidad tomando el transporte público. Al investigar un poco se da cuenta de que tiene que tomar un camión en cierta calle y éste lo llevará a su universidad. En la primer semana toma 5 camiones, uno por día, y 2 lo llevan a su universidad mientras que los otros 3 lo llevan más lejos y tiene que caminar más para poder llegar a la universidad. Este estudiante quiere entonces calcular el porcentaje de camiones que lo dejan cerca. Vamos a atacar este problema con estadística bayesiana.

Definamos a la proporción de camiones que lo dejan cerca como θ , dado que es una proporción entonces $0 \leq \theta \leq 1$. En principio θ puede tomar cualquier valor entre 0 y 1, pero por ahora lo restringiremos a tomar valores en intervalos de 0.1, esto nos da 11 hipótesis, cada una con probabilidad igual 1/11, escribiendo esto en una caja de Bayes obtenemos el cuadro 3.5.

Hipótesis de θ	anterior $p(\theta)$	verosimilitud $p(x \theta)$	h $p(\theta)p(x \theta)$	posterior $p(\theta x)$
0	0.0909			
0.1	0.0909			
0.2	0.0909			
0.3	0.0909			
0.4	0.0909			
0.5	0.0909			
0.6	0.0909			
0.7	0.0909			
0.8	0.0909			
0.9	0.0909			
1	0.0909			
Total:	1			

Cuadro 3.5: Caja de Bayes incompleta para las hipótesis en el ejemplo del camión. En esta caja se tienen únicamente las probabilidades anteriores de las hipótesis.

Definamos x como el número de éxitos (número de camiones que si dejan al estudiante cerca de su escuela) de un total N de repeticiones. Si hay N repeticiones de un experimento “aleatorio” con probabilidad θ de que ocurra x veces el evento buscado entonces podemos expresar su función de probabilidad como una función binomial:

$$p(x|\theta) = \frac{N!}{x!(N-x)!} \theta^x (1-\theta)^{N-x}. \quad (3.4)$$

Esta función nos dice la distribución de probabilidad de x dado θ . Para este ejemplo $N = 5$ y $x = 2$. Esta distribución nos da entonces la verosimilitud para cada θ , teniendo la verosimilitud podemos calcular tanto h como la probabilidad posterior, entonces la caja de Bayes queda completa (cuadro 3.6).

Hipótesis de θ	anterior $p(\theta)$	verosimilitud $p(x \theta)$	h $p(\theta)p(x \theta)$	posterior $p(\theta x)$
0	0.0909	0	0	0
0.1	0.0909	0.0729	0.0066	0.0437
0.2	0.0909	0.2048	0.0186	0.1229
0.3	0.0909	0.3087	0.0281	0.1852
0.4	0.0909	0.3456	0.0314	0.2074
0.5	0.0909	0.3125	0.0284	0.1875
0.6	0.0909	0.2304	0.0209	0.1383
0.7	0.0909	0.1323	0.0120	0.0794
0.8	0.0909	0.0512	0.0047	0.0307
0.9	0.0909	0.0081	0.0007	0.0049
1	0.0909	0	0	0
Total:	1		0.1515	1

Cuadro 3.6: Caja de Bayes completa para el ejemplo del camión. Se observa que la probabilidad posterior más alta está en $\theta = 0.4$ mientras que empieza a bajar al alejarse de 0.4 en ambos sentidos.

Analizemos este ejemplo ahora que tenemos la caja de Bayes completa (cuadro 3.6). Tenemos una función de densidad de probabilidad (3.4), y queremos saber el valor del parámetro θ dados los datos x y N . En la tabla observamos que la probabilidad posterior mayor corresponde a (sin mucha sorpresa) $\theta = 0.4$. Entonces es seguro decir que, dada la información proporcionada, el valor del parámetro $\theta = 0.4$ es el que mejor se ajusta a la información. Esto tiene sentido, pues de los 5 camiones que pasaron, únicamente 2 lo dejaron en la escuela, y $\frac{2}{5} = 0.4$.

De este ejemplo podemos ver otro concepto importante. A partir de la información de la caja de Bayes (cuadro 3.6) podemos reconstruir la **distribución de probabilidad para**

el **parámetro** θ , si graficamos los valores del posterior obtenemos la figura 3.1 se obtiene dicha distribución, que nos da información de los valores que puede tomar θ de acuerdo a los datos experimentales utilizados (número de camiones que dejan al alumno cerca de su universidad).

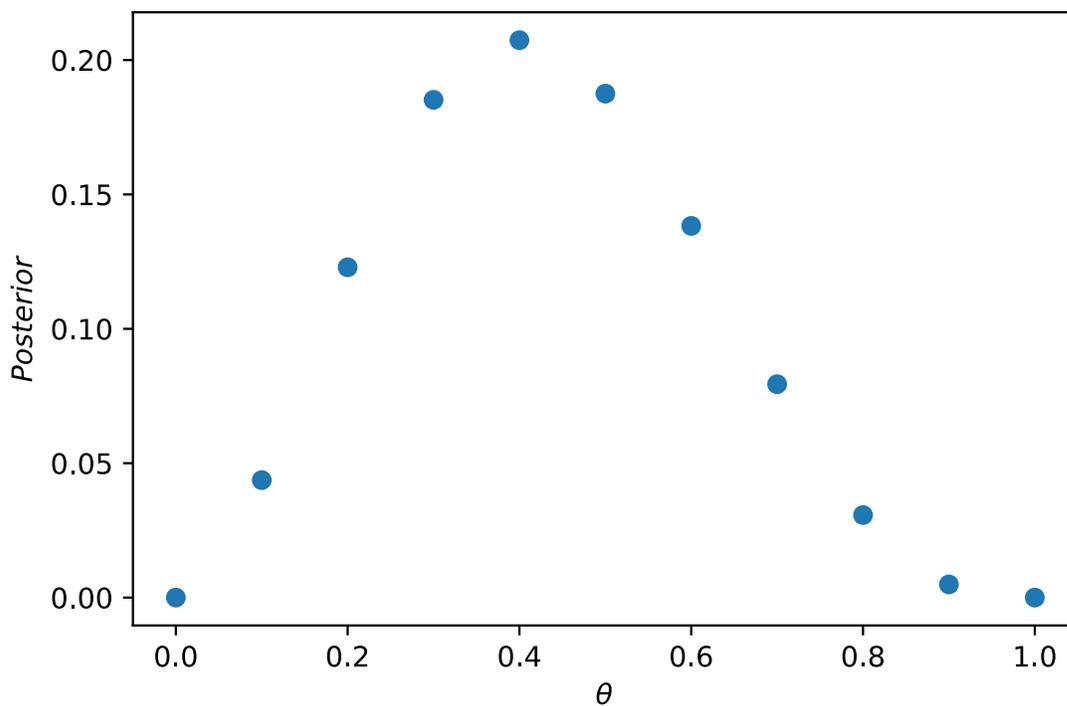


Figura 3.1: Distribución de probabilidad para el parámetro θ . Se observa claramente un comportamiento gaussiano.

Podemos reescribir la regla de Bayes para la estimación de parámetros. Como el término $P(D)$ es un término de normalización podemos retirarlo para reescribir (3.1) como:

$$p(\theta|x) \propto p(\theta)p(x|\theta), \tag{3.5}$$

o

$$posterior \propto anterior \times verosimilitud. \tag{3.6}$$

3.3. Prueba de hipótesis (selección de modelo)

Como vimos en la sección anterior, para realizar estimación de parámetros se necesitan las probabilidades posteriores de los valores posibles del parámetro. El valor del parámetro con una probabilidad posterior mayor se considera que es el valor que mejor se ajusta con los datos u observaciones.

Para realizar una prueba de hipótesis (o selección de modelo) en estadística Bayesiana se hace algo similar a una estimación de parámetros, pues los valores probables del parámetro son las hipótesis, pero un cierto intervalo de valores también pueden ser una hipótesis. Un ejemplo rápido de este caso sería, utilizando el ejemplo anterior de los camiones, decir que tenemos 2 hipótesis posibles, la nula H_0 y la alternativa H_1 . Las definimos como:

$$H_0 : \theta \leq 0.5,$$

$$H_1 : \theta > 0.5,$$

por lo que tenemos dos posibilidades, o el valor de θ está de 0.5 hacia abajo o arriba de 0.5. Para obtener la probabilidad sólo tenemos que sumar las probabilidades posteriores de los valores independientes de θ según se ajusten a alguna de las dos hipótesis. Entonces

$$P(H_0|x) = P(\theta \leq 0.5|x) = 0.7467,$$

$$P(H_1|x) = P(\theta > 0.5|x) = 0.2533.$$

En este ejemplo se cumple que $P(H_0|x) + P(H_1|x) = 1$ pues las hipótesis son mutuamente excluyentes (si ocurre H_0 es imposible que ocurra H_1).

La regla de Bayes se cumple para ambas hipótesis, entonces

$$P(H_0|x) = \frac{P(x|H_0)P(H_0)}{P(x)}, \quad (3.7)$$

$$P(H_1|x) = \frac{P(x|H_1)P(H_1)}{P(x)}, \quad (3.8)$$

y de aquí obtenemos la forma *odds* de la regla de Bayes

$$\frac{P(H_0|x)}{P(H_1|x)} = \frac{P(x|H_0)}{P(x|H_1)} \times \frac{P(H_0)}{P(H_1)}, \quad (3.9)$$

donde a $\frac{P(x|H_0)}{P(x|H_1)}$ se le llama **factor de Bayes**, el cuál nos dice qué tan probable es H_0 con respecto a H_1 .

Esto se puede generalizar para hipótesis con más de un parámetro. Supongamos que tenemos dos modelos H_0 y H_1 cada uno con un vector de parámetros θ_0 y θ_1 , entonces

$$\frac{P(x|H_0)}{P(x|H_1)} = \frac{\int P(\theta_0|H_0)P(x|\theta_0, H_0)d\theta_0}{\int P(\theta_1|H_1)P(x|\theta_1, H_1)d\theta_1}. \quad (3.10)$$

3.4. Cadenas de Markov

Para explicar la reconstrucción que se utilizará hace falta explicar un poco un proceso estocástico conocido como **Cadena de Markov**. Pero primero unas definiciones importantes:

- Proceso estocástico: proceso que pretende describir la evolución temporal de algún fenómeno aleatorio.
- Propiedad de Markov: propiedad de los procesos estocásticos que consiste en su ausencia de memoria, esto significa que la distribución de probabilidad de estados futuros del proceso depende únicamente del estado presente (ninguno de los estados anteriores al presente afectan esta distribución de eventos futuros, por esto se dice que tiene ausencia de memoria, pues no “recuerda” los pasos anteriores al actual o presente).

Si un proceso tiene esta propiedad se le llama *proceso de Markov*.

Usando estas definiciones ahora podemos definir una cadena de Markov.

- Cadena de Markov: proceso de Markov que describe una secuencia de posibles eventos representándolos como “pasos” o “eslabones”.

Ejemplos muy populares de procesos que pueden ser modelados por una cadena de Markov son los llamados *caminantes aleatorios*.

Dado que las cadenas de Markov son un proceso de Markov cada paso que da la cadena es indiferente a todos los anteriores excepto por el paso en el que se encuentre en ese momento.

Las cadenas de Markov pueden representarse como un diagrama de distintos estados. Pongamos un ejemplo, imaginemos que Pepito realiza alguna de las siguientes 3 actividades todos los días: correr, dormir o comer helado y por alguna razón un amigo suyo realizó una estadística de su comportamiento. El amigo descubrió que si Pepito dormía, al día siguiente tenía una probabilidad de 0.6 de correr, una probabilidad de 0.2 de comer helado y una de 0.2 de volver a dormir. Si Pepito comía helado entonces era 0.7 de correr, 0.1 de comer helado de nuevo y 0.2 de dormir, y si corría entonces era 0.6 de volver a correr al día siguiente, 0.1 de dormir y 0.3 de comer helado. Esta información se puede representar de manera más fácil en un diagrama de un proceso de Markov como el de la Figura 3.2.

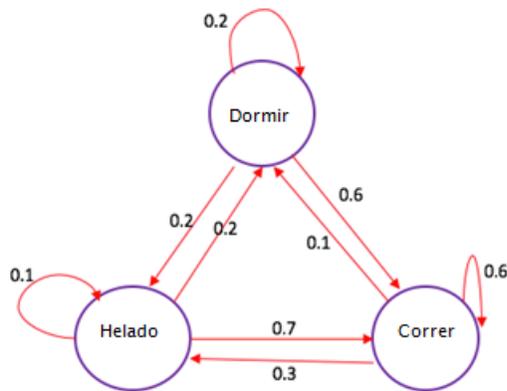


Figura 3.2: Diagrama que representa el proceso de Markov.

Si la cadena de Markov comenzase en el estado *dormir* podemos ver en el diagrama la probabilidad que tienen los 3 estados de ser el siguiente paso en la cadena.

3.5. Cadenas de Markov Monte Carlo (MCMC)

Se le llama *Monte Carlo* de manera general a alguna técnica computacional que utilice números aleatorios. Un tipo de Monte Carlo es MCMC, y un tipo de método MCMC es el algoritmo Metropolis-Hastings. Este algoritmo nos da una manera de construir una cadena donde los valores de algún parámetro con probabilidades posteriores mayores aparezcan de manera más frecuente en la cadena que los valores con probabilidades posteriores menores. Esto no hace que sea imposible que los valores con probabilidades menores ocurran, por lo que estarán presentes en la cadena pero con menos frecuencia.

De manera sencilla, el algoritmo funciona de la siguiente manera:

1. Se comienza en algún estado θ .
2. Se genera una propuesta θ' para el siguiente paso.
3. Si la $h' = \text{anterior} \times \text{verosimilitud}$ de θ' es mayor que la h del estado actual entonces se da el paso hacia θ' en la cadena y éste será el nuevo estado actual de la cadena, se reinicia el proceso con θ' como el nuevo estado inicial.
4. Si h' es menor entonces se propone una probabilidad mínima h'/h y se genera un número aleatorio entre 0 y 1, si el número generado es menor o igual a h'/h entonces se da el paso y θ' será el nuevo estado actual en la cadena, se reinicia el proceso con θ' como nuevo estado inicial, pero si el número generado es mayor que h'/h entonces se rechaza la propuesta θ' y se reinicia el proceso.

Veamos al paso 4 del algoritmo, se observa que h'/h es la ec. (3.9), por lo que el algoritmo Metropolis-Hastings es una forma de realizar comparaciones entre hipótesis. Cada hipótesis tiene una cierta probabilidad de ser cierta.

Entonces tendremos cadenas de Markov recorriendo el **espacio de parámetros** (figura 3.3) intentando buscar la combinación de valores con mejor verosimilitud de acuerdo a los datos.

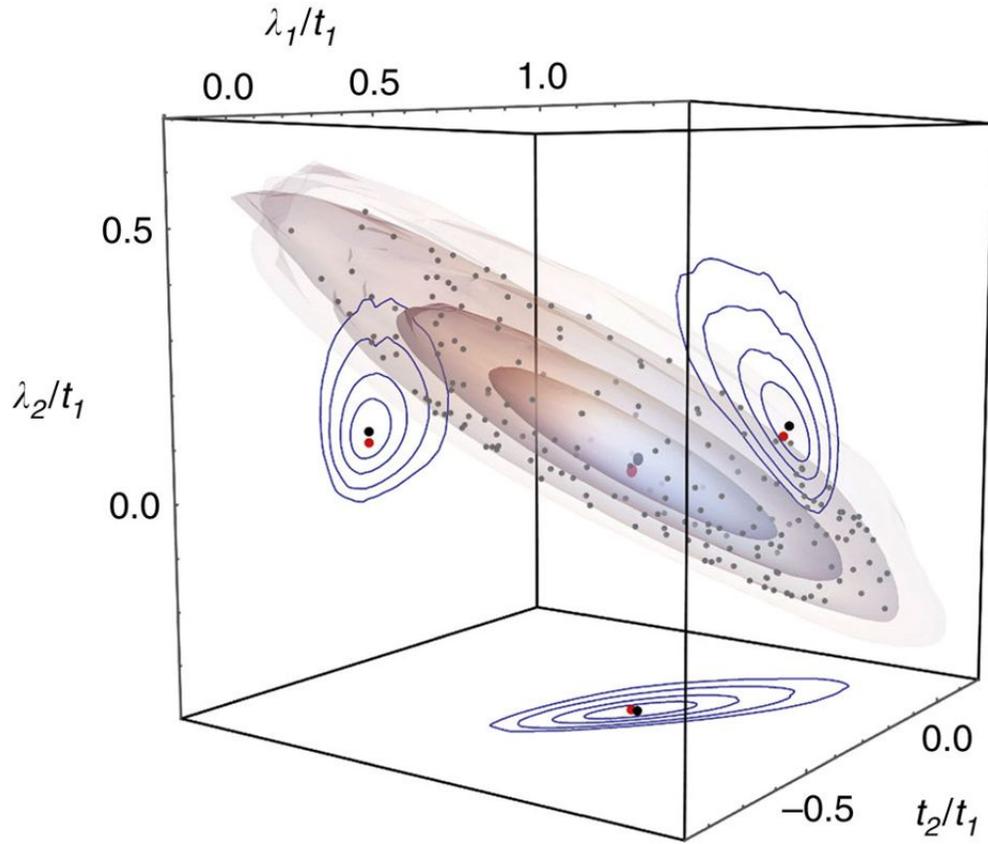


Figura 3.3: Representación de un espacio de parámetros en 3 dimensiones. Se observan las distintas curvas de confianza y sus proyecciones en cada uno de los 3 planos. Fuente: <https://www.nature.com/articles/ncomms9629/figures/10>.

3.5.1. Ejemplo de MCMC con Λ CDM

Haremos un ejemplo utilizando una versión muy sencilla del modelo Λ CDM como base donde tendremos 2 parámetros: el parámetro de densidad de materia Ω_m y el parámetro adimensional de Hubble h . En este ejemplo reconstruiremos

$$H(z) = 100 \frac{\text{km}}{\text{sMpc}} \times h \left(\Omega_m (1+z)^3 + (1 - \Omega_m) \right)^{\frac{1}{2}}, \quad (3.11)$$

usando mediciones del parámetro de Hubble. Es un modelo muy simple, pues no estamos tomando en cuenta nada que no sea materia ni energía oscura. Se desea utilizar el código SimpleMC [43] para encontrar los valores de ambos parámetros que mejor se ajusten a los datos y además reconstruir las distribuciones de los parámetros para intentar ver si existe correlación entre ellos. Las cadenas llegaron a los 10,000 pasos y dado que tratamos con Λ CDM esperamos obtener $h \approx 0.68$ y $\Omega_m \approx 0.28$.

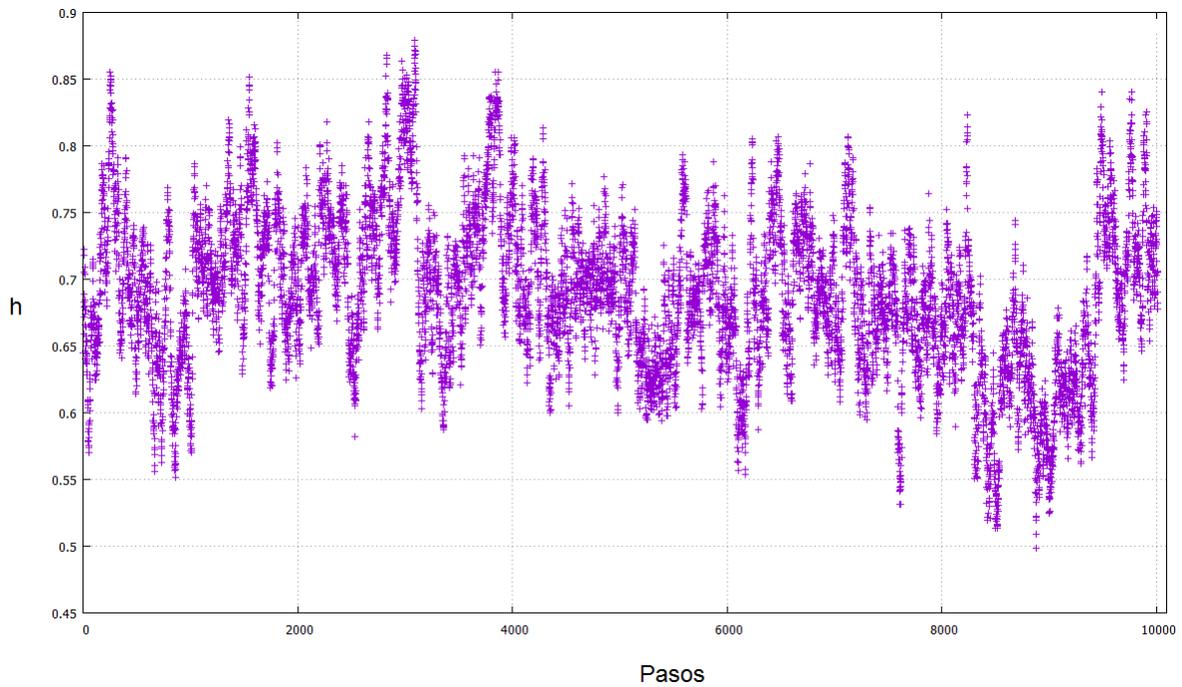


Figura 3.4: Cadena de Markov para el parámetro h

Como se observa en las figuras 3.4 y 3.5, ambos parámetros están oscilando entre los valores esperados. Se realizó un análisis de estas cadenas para encontrar el valor con mayor probabilidad posterior y se obtuvieron los valores de $h = 7.01 \pm 0.73$ y $\Omega_m = 2.64 \pm 0.59$, lo cual es bastante aceptable pues es una sola cadena y sólo se utilizaron mediciones del parámetro de Hubble para la reconstrucción. Lo interesante es que se observa cierta correlación en los parámetros pues la gráfica de la figura 3.6 se observa una inclinación

bastante pronunciada.

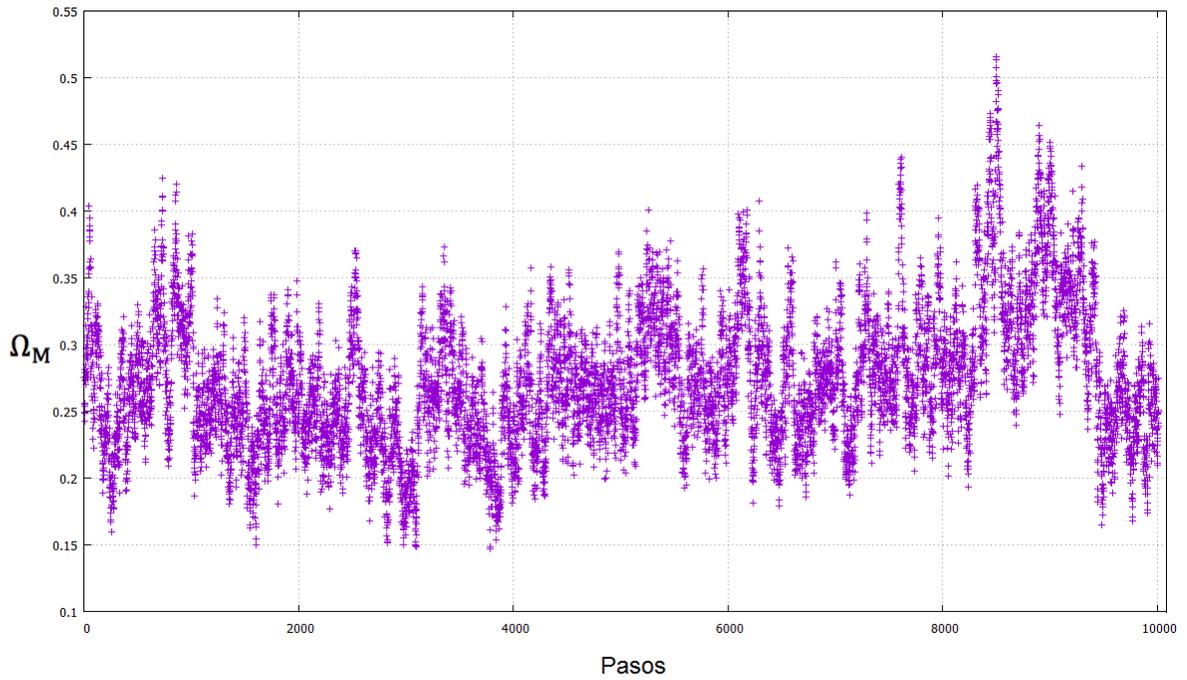


Figura 3.5: Cadena de Markov para el parámetro Ω_m

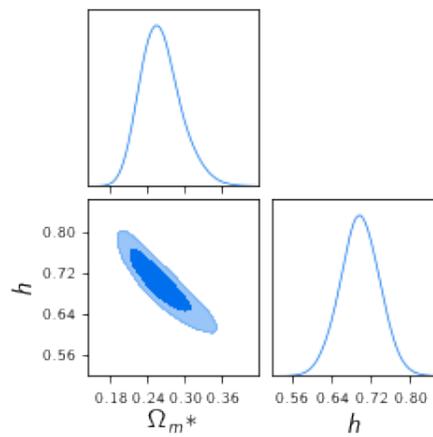


Figura 3.6: Grafica mostrando correlación entre h y Ω_m

3.6. AIC, AICc, BIC y DIC

Dado que en este trabajo se realiza selección de modelos se necesita algún criterio que nos diga si algún modelo es mejor que otro. Lo ideal sería utilizar el factor de Bayes (3.10), pero la integral es muy difícil de calcular y costosa en tiempo de computo, por lo que se recurre a utilizar aproximaciones a la misma u otro criterio.

Los criterios a utilizar para evaluar la calidad de los modelos recompensan un buen ajuste a los datos y dan una penalización de acuerdo al número de parámetros implicado. Esto se hace debido a un principio filosófico conocido como “La navaja de Occam”, el cuál nos dice que: **la solución más sencilla o simple a un problema debe ser la correcta.**

Uno de los criterios más conocidos es el *criterio de información de Akaike* (AIC por sus siglas en inglés) [39] donde se busca minimizar el valor de AIC dado por:

$$AIC = 2k - 2 \ln L, \quad (3.12)$$

donde k es el número de parámetros en el modelo y L es la verosimilitud máxima del modelo. La idea es que el modelo sea lo más sencillo posible, por lo que se “penaliza” al modelo mientras más parámetros tenga y además se le recompensa si tiene una verosimilitud muy buena. Este criterio surge tras minimizar una cantidad conocida como entropía relativa o divergencia de Kullback-Leibler, que nos dice la probabilidad de distintas distribuciones de probabilidad de ser ciertas de acuerdo a otra.

Existe un ajuste al criterio AIC si existen pocos valores (datos) en la muestra que se está modelando, a este criterio se le llama AICc [39] y se calcula con:

$$AICc = AIC + \frac{2k^2 + 2k}{n - k + 1}, \quad (3.13)$$

siendo n el número de datos. Se observa que para $n \gg 1$ el segundo término tiende a cero.

Ahora veremos el **criterio de información Bayesianana** (BIC por sus siglas en inglés) dado por:

$$BIC = k \ln n - 2 \ln L, \quad (3.14)$$

este criterio es inferior al AIC a menos que el modelo real (el que reproduce perfectamente los datos) se encuentre en los modelos posibles [41], pero en estadística siempre se asume que ningún modelo es el real. Este criterio surge de una aproximación al factor de Bayes.

Por último tenemos el **criterio de información de desviación** (DIC por sus siglas en inglés). Este criterio es parecido al AIC, pero en lugar de utilizar los parámetros libres utiliza algo llamado **complejidad bayesiana**, que nos dice el número efectivo de parámetros de un modelo. La complejidad bayesiana viene dada por [39] [40]

$$p_D = \overline{\chi^2(\theta)} - \chi^2(\bar{\theta}), \quad (3.15)$$

siendo el primer término del lado derecho es la chi cuadrada promedio del modelo y el segundo término es la chi cuadrada del valor estimado de los parámetros (se hablará un poco más de la función chi cuadrada en el capítulo siguiente), que en este caso es la chi cuadrada de mejor ajuste. También necesitamos definir la *desviación*

$$D(\theta) = -2 \ln L + C, \quad (3.16)$$

siendo L la verosimilitud máxima y C una constante relacionada con los datos que desaparecerá al hacer comparación de DIC entre modelos. De aquí definimos el DIC como

$$DIC = D(\bar{\theta}) + 2p_D = \chi^2(\bar{\theta}) + C + 2p_D. \quad (3.17)$$

En este trabajo se utilizarán los 4 criterios para evaluar la calidad de los modelos.

Capítulo 4

Interpolaciones como reconstrucciones

La reconstrucción que se utilizará en este trabajo consiste en utilizar **interpolaciones**. Una interpolación es un método para obtener nuevos datos a partir de datos conocidos. Existen varios tipos de interpolación, pero los que se utilizarán en este trabajo son la interpolación lineal y la interpolación spline cúbica.

4.1. Tipos de interpolaciones

La interpolación lineal consiste en unir los puntos de los datos a través de líneas rectas. Digamos que se tienen dos puntos (x_a, y_a) y (x_b, y_b) , la interpolación entre estos dos puntos está dada por la recta:

$$y - y_a = \frac{y_b - y_a}{x_b - x_a}(x - x_a), \quad (4.1)$$

y así sucesivamente para los puntos posteriores. Es un método bastante rápido pero impreciso, y además la función no es diferenciable, pero dado que en este trabajo no se diferenciará la función reconstruida entonces se puede utilizar.

La otra interpolación a utilizar es la spline cúbica que funciona similar a la lineal, pero la diferencia es que la lineal utiliza rectas para unir los datos mientras que la spline cúbica utiliza polinomios. Esta interpolación es más lenta que la lineal, pero dado que genera curvas suaves es más precisa. En la figura 4.1 se puede apreciar la diferencia entre ambas.

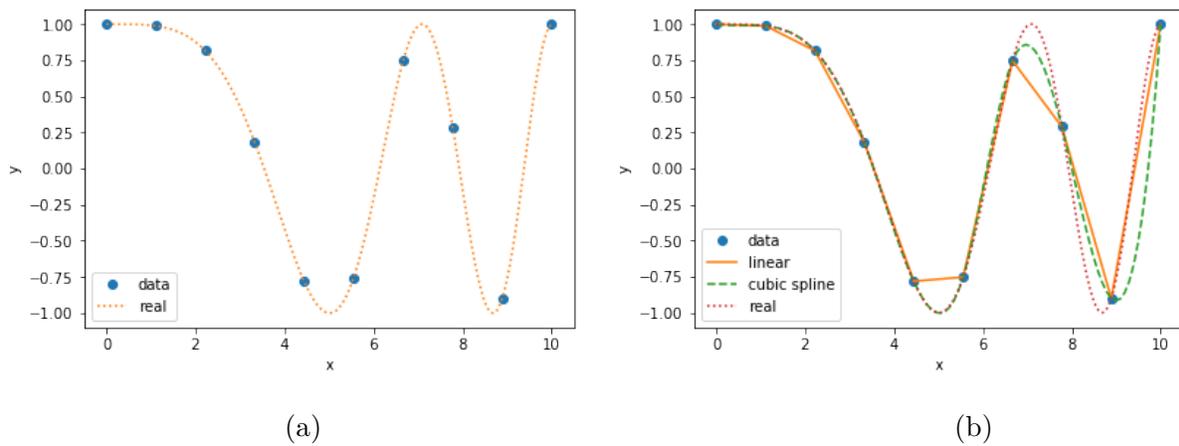


Figura 4.1: (a) En esta gráfica tenemos puntos de una cierta función que se utilizarán para interpolar. (b) Comparación entre las interpolaciones spline cúbica y lineal con respecto a la función original (la línea punteada).

Ahora veamos cómo se utilizarán las interpolaciones en este trabajo. Supongamos que tenemos 4 puntos $(z_i, F_i(z_i))$ con $i = 1, \dots, 4$ en el plano cartesiano y estos toman el valor de $z_1 = 0$, $z_2 = 1$, $z_3 = 2$ y $z_4 = 3$, mientras que $F_i(z_i) = 1$ para $i = 1, 2, 3, 4$, si interpolamos estos cuatro puntos obtendremos una gráfica que representa a $F(z)$ como constante (figura 4.2). Ahora supongamos que F_2 es un parámetro que puede tomar valores entre 0.5 y 1.5 dando pasos de tamaño 0.2, mientras que el resto de los F_i y z_i permanecen con un valor fijo. Esto nos generará una gráfica distinta para cada valor de F_2 como se observa en la figura 4.3. Esto ocurre si permitimos que varíe únicamente F_2 , pero si además permitimos que F_3 tome los mismos valores empezamos a ver que aumenta la cantidad de gráficas producto de hacer interpolación (figura 4.4).

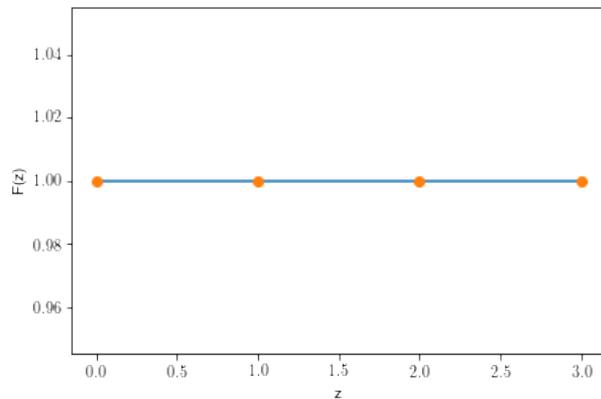


Figura 4.2: Interpolación asumiendo puntos en una recta correspondiente a $F(z) = cte$.

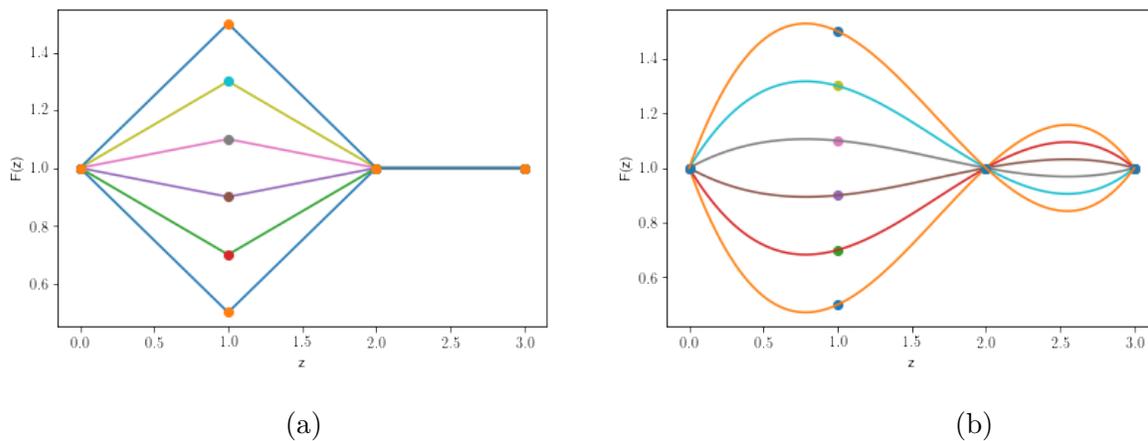


Figura 4.3: (a) Interpolación lineal con un punto variando en el eje y . (b) Interpolación cúbica con un punto variando en el eje y . Se puede apreciar que ambas figuras comparten cierta similitud en $0 < z < 2$ en el sentido de que se tiene un máximo (o mínimo), pero en $2 < z < 3$ la interpolación cúbica tiene un poco de “ruido” debido a la forma en la que se hace este tipo de interpolación. Esto afecta el ajuste de esta reconstrucción con los datos.

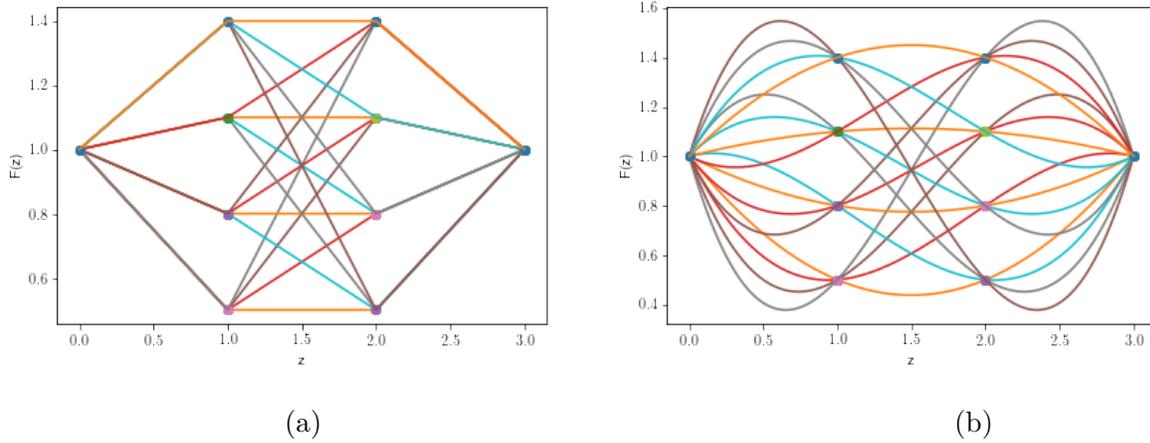


Figura 4.4: (a) Interpolación lineal con 2 puntos variando en el eje y . (b) Interpolación spline cúbica con 2 puntos variando en el eje y .

Por último supongamos que en lugar de F_3 hacemos que z_2 sea un parámetro, esto permitiría tener un punto variable en el plano, esto queda mejor representado (al menos visualmente) en la figura 4.5.

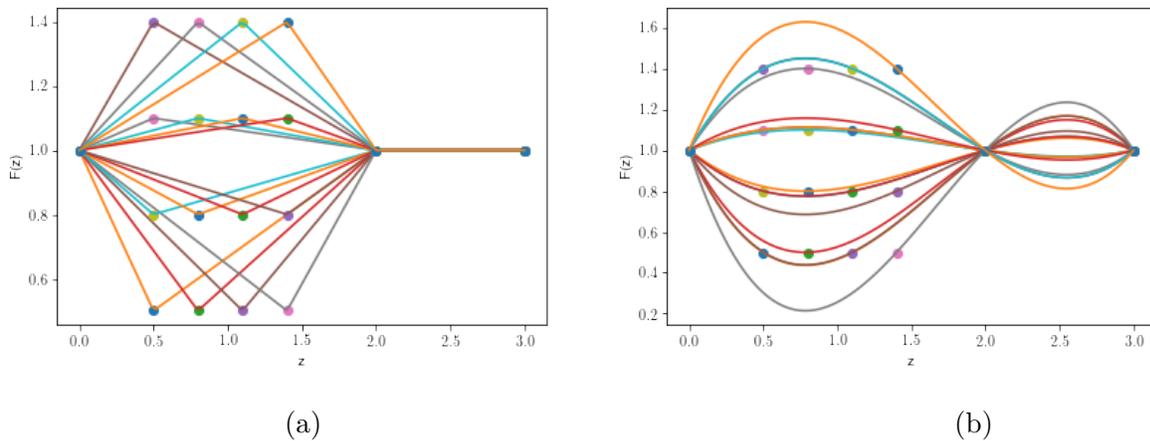


Figura 4.5: (a) Interpolación lineal con un punto variando en el eje y (F_2) y uno en el eje x (z_2). (b) Interpolación cúbica con un punto variando en el eje y (F_2) y uno en el eje x (z_2). Se observa el mismo efecto de “ruido” en $2 < z < 3$ que en la figura 4.3.

Nuestra reconstrucción consiste en proponer parámetros que sirvan como puntos en una gráfica generada a partir de la interpolación de dichos parámetros, esta gráfica puede representar alguna propiedad o cantidad cosmológica (como la ecuación de estado $w(z)$ de la energía oscura) y se calculará qué tan bien se ajusta con las observaciones, aquí es donde entra SimpleMC (un código escrito en lenguaje Python 2 que utiliza el método MCMC para restringir los valores de parámetros de acuerdo a observaciones y experimentos cosmológicos [43]), cada parámetro tiene un cierto intervalo de valores que puede tomar, y cada uno de estos valores tiene una *probabilidad anterior* igual, SimpleMC varía el valor de los parámetros y procede a comparar la interpolación con las observaciones cosmológicas. La calidad del ajuste se mide utilizando la función χ^2 . Esta función nos dice qué tan cerca está el modelo de los datos reales. Cada combinación de parámetros (si suponemos que las distribuciones posteriores de probabilidad son gaussianas) tiene una *verosimilitud* dada por [38]

$$L_m = e^{-\frac{\chi_m^2}{2}}, \quad (4.2)$$

siendo m el subíndice de dicha combinación de parámetros, y debido a que tenemos muchos conjuntos de datos como: supernovas tipo 1a (SN), radiación de fondo (CMB), parámetro de Hubble (H), etc., supondremos que son independientes unos de otros, por lo que la χ_m^2 total estará dada como [38]

$$\chi_m^2 = \chi_{m,SN}^2 + \chi_{m,CMB}^2 + \chi_{m,H}^2 + \dots \quad (4.3)$$

También se observa de (4.3) que podemos reescribir (3.12) y (3.14) como

$$AIC = 2k + \chi_m^2, \quad (4.4)$$

$$BIC = k \ln n + \chi_m^2. \quad (4.5)$$

Capítulo 5

Prueba

5.1. Reconstruyendo $H(z)$ con datos falsos

Para revisar qué tan fiable es el código SimpleMC se realizó una prueba con datos creados artificialmente de $H(z)$ (figura 5.1). La función a partir de la cuál se generaron estos datos es desconocida para hacer esta prueba “a ciegas”.

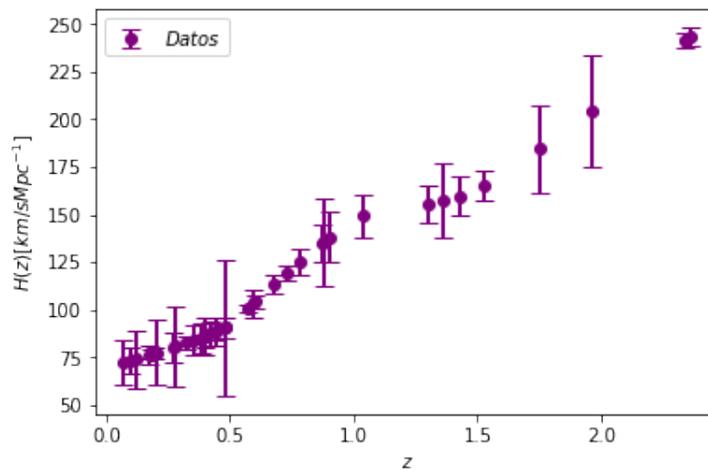


Figura 5.1: Datos artificiales de $H(z)$.

Se utilizaron reconstrucciones no paramétricas de interpolaciones lineal y cúbica. Si a

las interpolaciones las llamamos $F_l(z)$ entonces el $H(z)$ sería

$$H(z) = 100 \frac{\text{km}}{\text{sMpc}} \times h \left(\Omega_m (1+z)^3 + (1 - \Omega_m) F_l(z) \right)^{\frac{1}{2}} \quad (5.1)$$

y la idea es encontrar la forma de la gráfica de $F_l(z)$ que mejor se ajuste a los datos falsos. (5.1) es una forma “reducida” de la ecuación (1.46) donde se utilizó la convención $a_0 = 1$, se removieron los ceros de los parámetros de densidad, se asume un universo plano ($\Omega_k = 0$), se desprecia la aportación del parámetro de radiación a la densidad total ($\Omega_r \approx 0$) y se sustituye $\Omega_{DE}(z)$ por $(1 - \Omega_m)F_l(z)$.

Lo que hace SimpleMC es asignarle valores a los parámetros, luego se interpolan los puntos (de manera lineal o cúbica según se escoja), se sustituye la gráfica generada en (5.1) y se grafica $H(z)$, la gráfica luego es comparada con los datos y se evalúa qué tanto se acerca a los datos, se vuelve a comenzar el proceso con distintos valores de los parámetros y al final se utiliza el valor que mejor se ajuste a los datos para hacer comparación de modelos. Se realizó una prueba con cada interpolación y se obtuvieron los valores de mejor ajuste (valores con la mayor verosimilitud). Para realizar buena selección de modelos se realizó una comparación entre 7 distintas interpolaciones (figura 5.2) llamadas F_{li} con $i = 1, \dots, 7$. Las 7 fueron realizadas con interpolación lineal, mientras que las $i = 2, 4, 6, 7$ fueron hechas con interpolación spline cúbica. Las interpolaciones que se realizaron son las siguientes:

$$F_{l1} : F_l : (1.0, F_1, 1.0), \quad z : (0.0, 1.5, 3.0)$$

$$F_{l2} : F_l : (0.0, F_1, F_2, 1.0), \quad z : (0.0, 1.0, 2.0, 3.0)$$

$$F_{l3} : F_l : (F_1, F_2, F_3), \quad z : (0.0, 1.5, 3.0)$$

$$F_{l4} : F_l : (F_1, F_2, F_3, F_4), \quad z : (0.0, 1.0, 2.0, 3.0)$$

$$F_{l5} : F_l : (F_1, F_2, F_3), \quad z : (0.0, z_0, 3.0)$$

$$F_{l6} : F_l : (F_1, F_2, F_3, F_4, F_5), \quad z : (0.0, 0.8, 1.6, 2.4, 3.0)$$

$$F_{l7} : F_l : (F_1, F_2, F_3, F_4), \quad z : (0.0, z_1, z_2, 3.0)$$

siendo F_i y z_i los parámetros a variar. Los valores que pueden tomar dichos parámetros se encuentran en el cuadro 5.1. Las reconstrucciones de $F_l(z)$ con interpolación lineal para los valores con mejor verosimilitud se observan en la figura 5.3, las reconstrucciones con interpolación spline cúbica se encuentran en la figura 5.4. Se observan en todas las interpolaciones (en especial en las lineales) que se tiene un máximo alrededor de $z \approx 1$. Para hacer más clara la calidad de estas reconstrucciones se compararon con el modelo Λ CDM. Dado que los datos utilizados son falsos se espera que la función χ^2 de Λ CDM sea mayor que las asociadas a nuestras reconstrucciones. La reconstrucción de Λ CDM se realiza haciendo $F_l(z) = 1$ en (5.1), por lo que sólo se tienen dos parámetros: h y Ω_m (en el cuadro 5.2 se observan los valores obtenidos para ambos parámetros y la función chi cuadrada para la verosimilitud máxima).

Se graficaron las reconstrucciones F_{l7} lineal y F_{l6} spline cúbica (no se escogieron estas dos por nada en particular, sólo es para dar una representación visual del ajuste) con los datos falsos de $H(z)$ como se observa en la figura 5.5.

Parámetros	Intervalo de valores
Ω_m	[0.05, 1.5]
h	[0.4, 1.0]
F_1	[-2,4]
F_2	[-2,4]
F_3	[-2,4]
F_4	[-2,4]
F_5	[-2,4]
F_6	[-2,4]
F_7	[-2,4]
z_0	[0.2, 2.8]
z_1	[0.2, 1.3]
z_2	[1.7, 2.8]

Cuadro 5.1: Intervalo de valores de los parámetros para las 7 interpolaciones. Los intervalos de valores para los parámetros F_i son bastante grandes pues se desconocía la función original a partir de la cuál se generaron los datos artificiales de $H(z)$.

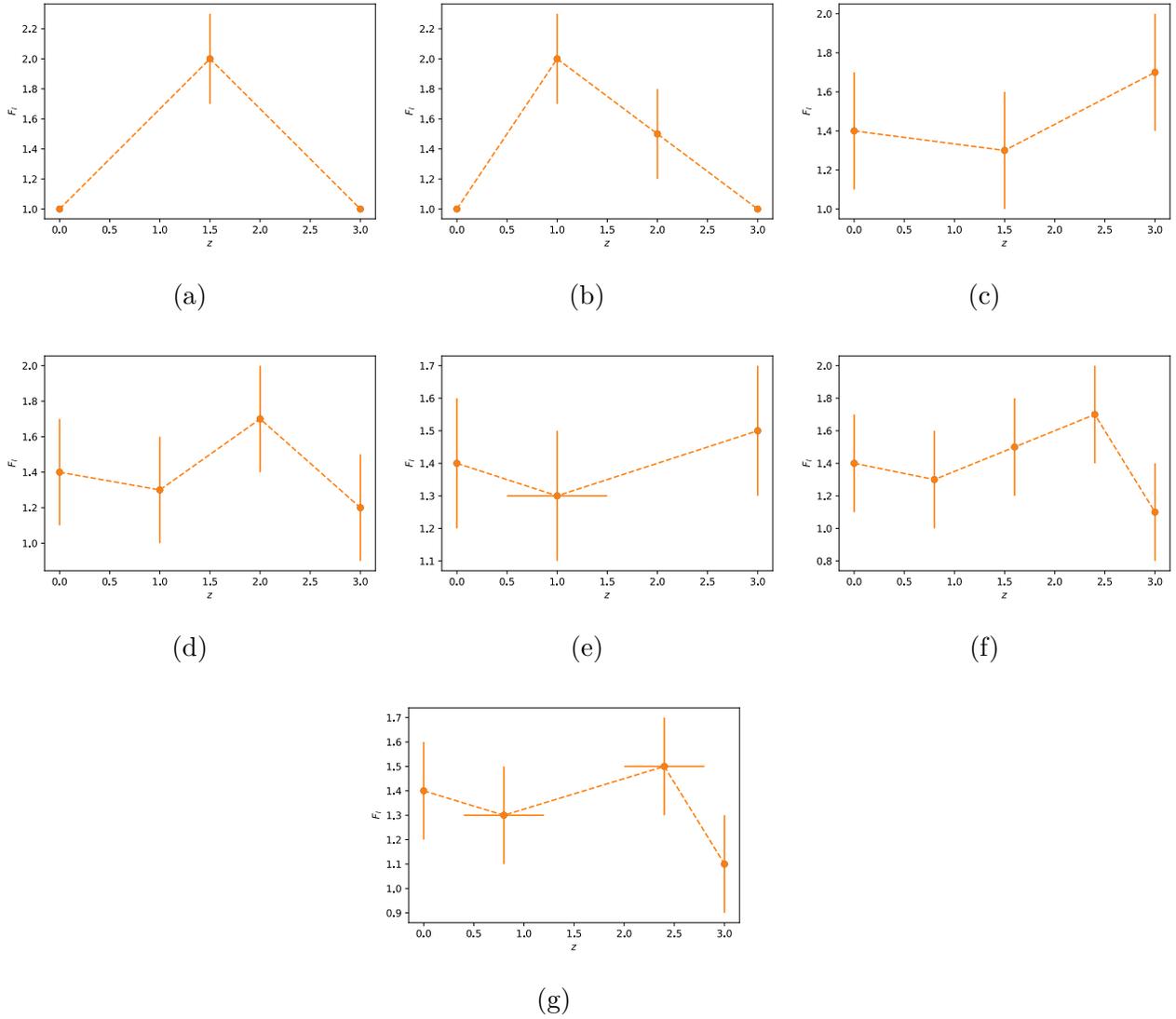


Figura 5.2: Interpolaciones a realizar. Las líneas en cada punto representan la dirección en la que pueden variar el punto. (a) F_{l1} (b) F_{l2} (c) F_{l3} (d) F_{l4} (e) F_{l5} (f) F_{l6} (g) F_{l7} .

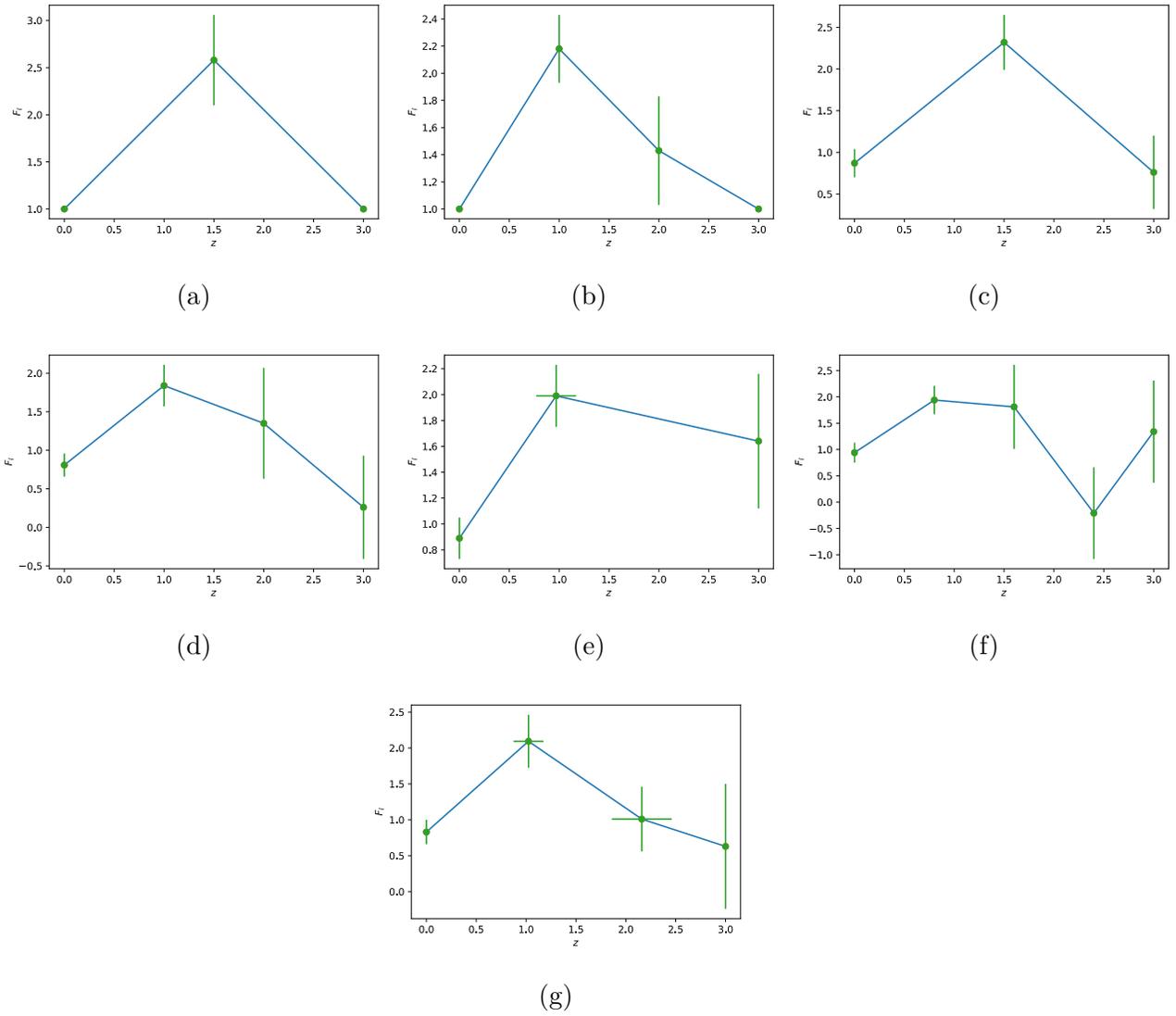
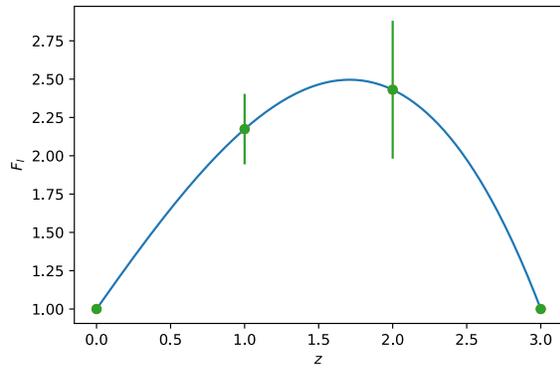
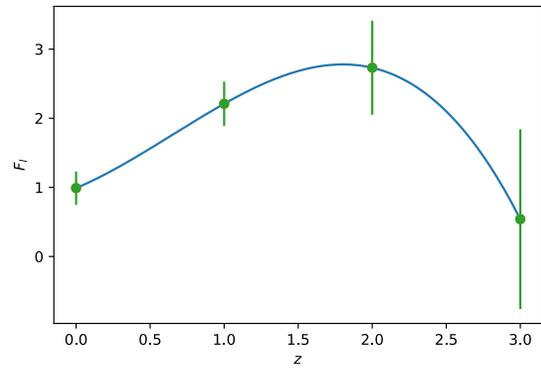


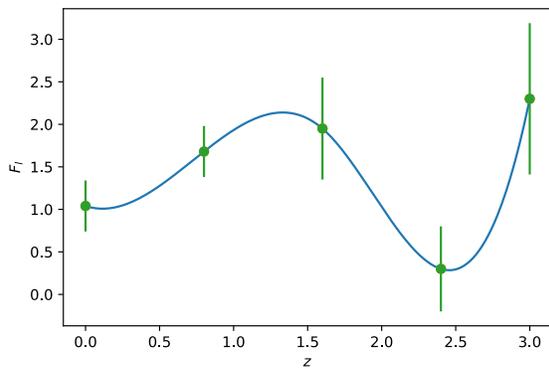
Figura 5.3: Reconstrucciones realizadas con los valores de mejor verosimilitud para las interpolaciones lineales. Las líneas en cada punto representan el error para cada parámetro. (a) F_{l1} (b) F_{l2} (c) F_{l3} (d) F_{l4} (e) F_{l5} (f) F_{l6} (g) F_{l7} .



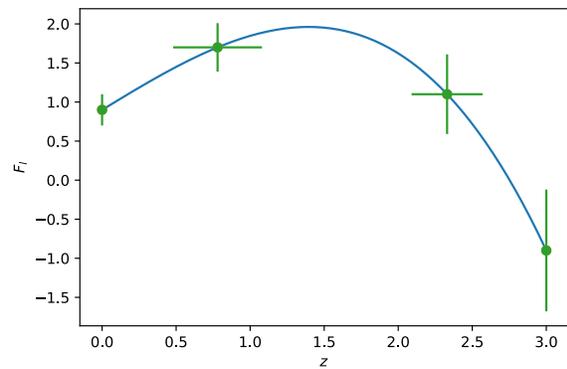
(a)



(b)



(c)



(d)

Figura 5.4: Reconstrucciones realizadas con los valores de mejor verosimilitud para las interpolaciones lineales. Las líneas en cada punto representan el error para cada parámetro. (a) F_{l2} (b) F_{l4} (c) F_{l6} (d) F_{l7} .

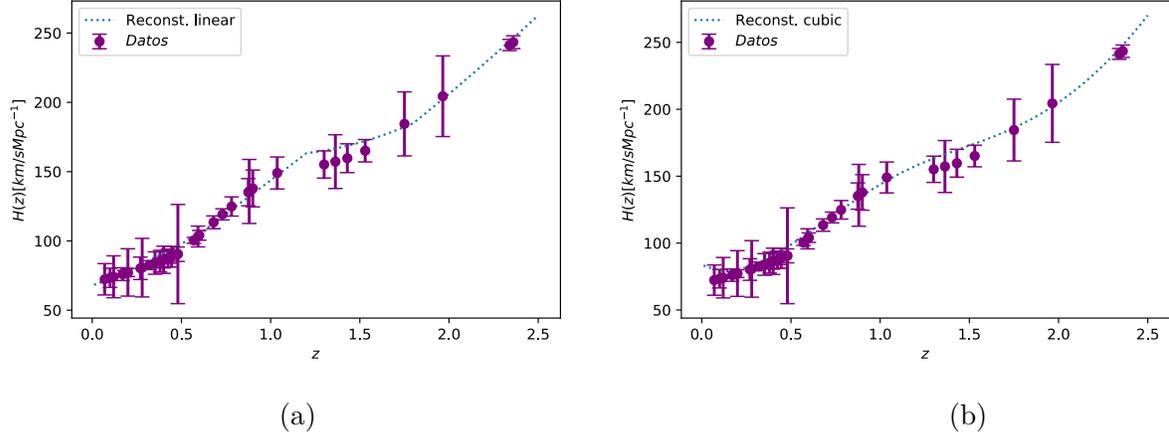


Figura 5.5: (a) Gráfica donde se observa el ajuste de los valores de la interpolación lineal F_{l7} con mejor verosimilitud con los valores artificiales de $H(z)$. (b) Gráfica donde se observa el ajuste de los valores de la interpolación spline cúbica F_{l6} con mejor verosimilitud con los valores artificiales de $H(z)$.

A simple vista se observa que las reconstrucciones se ajustan bastante bien a los datos falsos, mientras que Λ CDM tiene problemas cerca de $z = 1$ (figura 5.6), que es el valor del corrimiento al rojo cerca del cuál todas nuestras reconstrucciones F_{li} tienen un máximo. Pero no se puede hacer una comparación de modelos utilizando argumentos visuales como las figuras antes mencionadas, por lo que recurrimos a los criterios de información mencionados en el capítulo 3 (cuadro 5.3).

Parámetros (Λ CDM)	Promedio $\pm 1-\sigma$	Intervalo de valores
Ω_m	0.311 ± 0.03	[0.05, 1.5]
h	0.69 ± 0.03	[0.4, 1.0]
χ^2	34.02	...

Cuadro 5.2: Tabla con los valores de mejor ajuste de los parámetros obtenidos para Λ CDM.

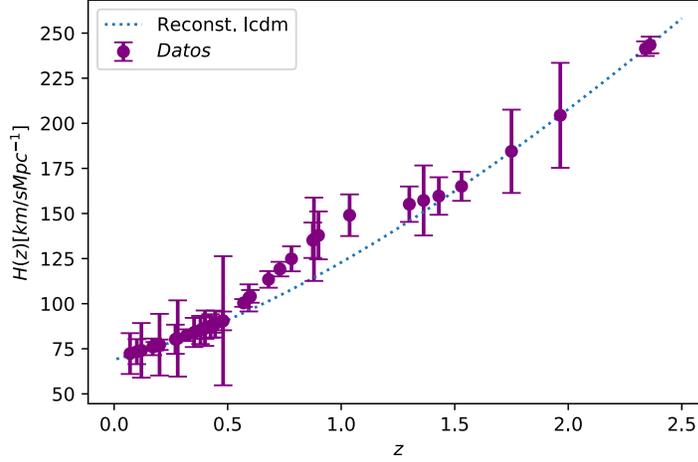


Figura 5.6: Ajuste de la reconstrucción de Λ CDM con los datos falsos de $H(z)$.

Modelo	$\chi^2(\bar{\theta})$	$\overline{\chi^2(\theta)}$	Núm. de params. (k)	p_D	ΔAIC	ΔAIC_c	ΔBIC	ΔDIC
Λ CDM	34.02	37.12	2	3.1	0	0	0	0
$F_{l_1}(z)$ lineal	16.01	19.67	3	3.66	-16.01	-15.66	-14.4	-16.89
$F_{l_2}(z)$ lineal	12.28	16.78	4	4.5	-17.74	-16.9	-14.52	-18.94
$F_{l_3}(z)$ lineal	15.76	20.06	5	4.3	-12.26	-10.78	-7.43	-15.86
$F_{l_4}(z)$ lineal	11.62	17.22	6	5.6	-14.4	-12.11	-7.96	-17.4
$F_{l_5}(z)$ lineal	13.14	18.04	6	4.9	-12.88	-10.59	-6.43	-17.28
$F_{l_6}(z)$ lineal	13.76	19.72	7	5.96	-10.26	-6.98	-2.21	-14.54
$F_{l_7}(z)$ lineal	12.16	16.94	8	4.78	-9.86	-5.39	-0.2	-18.5
$F_{l_2}(z)$ cúbica	14.28	19.36	4	5.08	-15.74	-14.9	-12.52	-15.78
$F_{l_4}(z)$ cúbica	13.71	19.18	6	5.47	-12.31	-10.02	-5.87	-15.57
$F_{l_6}(z)$ cúbica	12.01	17.11	7	5.1	-12.01	-8.73	-3.96	-18.01
$F_{l_7}(z)$ cúbica	13.42	19.32	8	5.9	-8.6	-4.13	1.06	-15

Cuadro 5.3: Comparación de modelos para el ajuste a los datos falsos de $H(z)$. Se prefieren los modelos con diferencia entre criterios negativa.

Veamos con cuidado estos resultados del cuadro 5.3. En la primer columna tenemos los modelos (Λ CDM y las interpolaciones $F_{l_i}, i = 1, \dots, 7$), en la segunda se tiene la chi

cuadrada de mejor ajuste, la tercer columna es la chi cuadrada promedio, en la cuarta tenemos k (número de parámetros del modelo), luego tenemos la complejidad bayesiana y en las últimas 4 tenemos la diferencia entre criterios de información con respecto al modelo Λ CDM. Estas últimas 4 columnas nos dice qué tan bien ajustan los modelos tomando Λ CDM como base. Dado que es una diferencia (por ejemplo $\Delta AIC = AIC - AIC_{\Lambda CDM}$) y se desea minimizar los criterios entonces se busca que el valor de esta diferencia sea negativo. Para analizar esta diferencia entre valores de criterios se utiliza la **escala de Jeffreys**:

- **Escala de Jeffreys:** para la diferencia entre criterios $\Delta IC < -5$ se dice que hay evidencia “fuerte” de que el modelo con el IC menor es mejor, y si $\Delta IC < -10$ la evidencia es decisiva [39].

En este caso IC puede ser cualquiera de los 4 criterios utilizados en este trabajo. En el cuadro se observa que todas nuestras reconstrucciones (excepto F_{l7} cúbica en ΔBIC) tienen valor negativo (como se esperaba pues tenemos datos falsos a los que Λ CDM no se ajusta bien) por lo que son mejores para explicar los datos falsos de $H(z)$.

Esta pequeña prueba es importante pues se demuestra que el código si funciona, que podemos comparar modelos distintos y que tenemos una forma de cuantificar (utilizando los criterios de información) qué tan buenos son nuestros modelos con respecto a un modelo base. El siguiente paso es utilizar datos reales en lugar de inventados. Esto lo veremos a detalle en la siguiente sección.

Capítulo 6

Resultados

6.1. Reconstrucción de $F_l(z)$.

Se hicieron muchas reconstrucciones y muchas cadenas, pero la mayoría fueron pruebas pues las cadenas se hicieron en una laptop únicamente para ver si el código y la reconstrucción de verdad servían. Las reconstrucciones que se presentarán en esta sección no son estas pruebas, si no unas hechas con cadenas de hasta 100,000 pasos hechas en un cluster. Se hicieron tanto reconstrucciones de $F_l(z)$ como de la ecuación de estado $w(z)$ de la energía oscura.

La reconstrucción de $F_l(z)$ usando interpolación spline cúbica como la utilizada en la sección anterior, se utilizaron 4 parámetros para F_l y se le permitió a los 2 valores centrales del eje z variar. Entonces se utilizaron 6 parámetros:

$$F_l : (F_1, F_2, F_3, F_4), \quad z : (0.0, z_1, z_2, 3.0). \quad (6.1)$$

y se utilizaron datos observacionales de SN1a, CMB y BBAOs. Como es común se utilizaron *probabilidades anteriores* planas (todas tienen la misma probabilidad de ocurrir). Se hizo la reconstrucción con 6 cadenas de markov, cada una de 100,000 pasos. Los resultados de los valores obtenidos para los parámetros se encuentran en el cuadro 6.1 y en la figura 6.1

se encuentran las distribuciones de probabilidad 1-D y 2-D.

Se procedió a utilizar estos resultados para reconstruir la distribución de probabilidad de $F_l(z)$, ésta se puede observar en la figura 6.2.

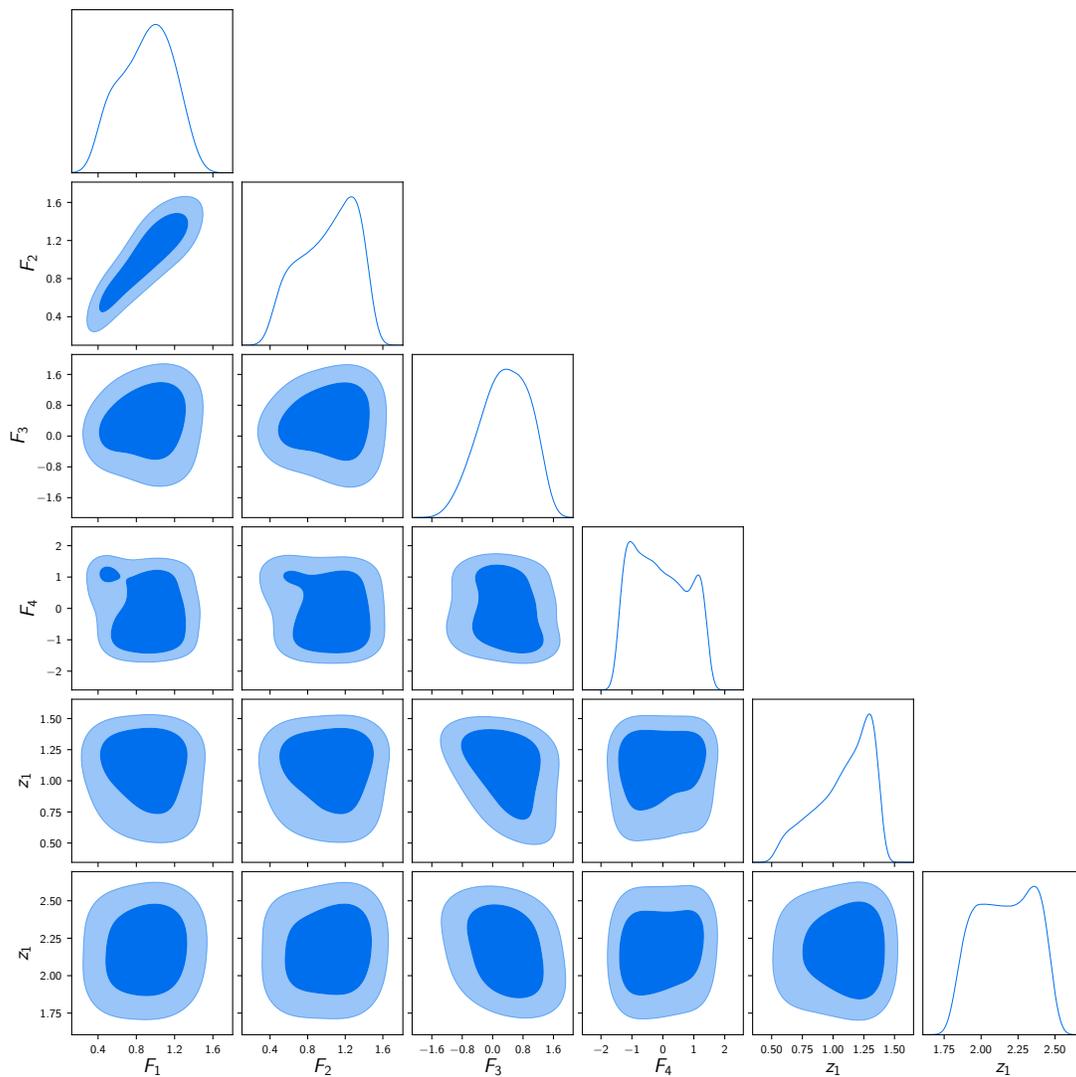


Figura 6.1: Distribuciones 1-D y 2-D de los parámetros para $F_l(z)$ con interpolación spline cúbica.

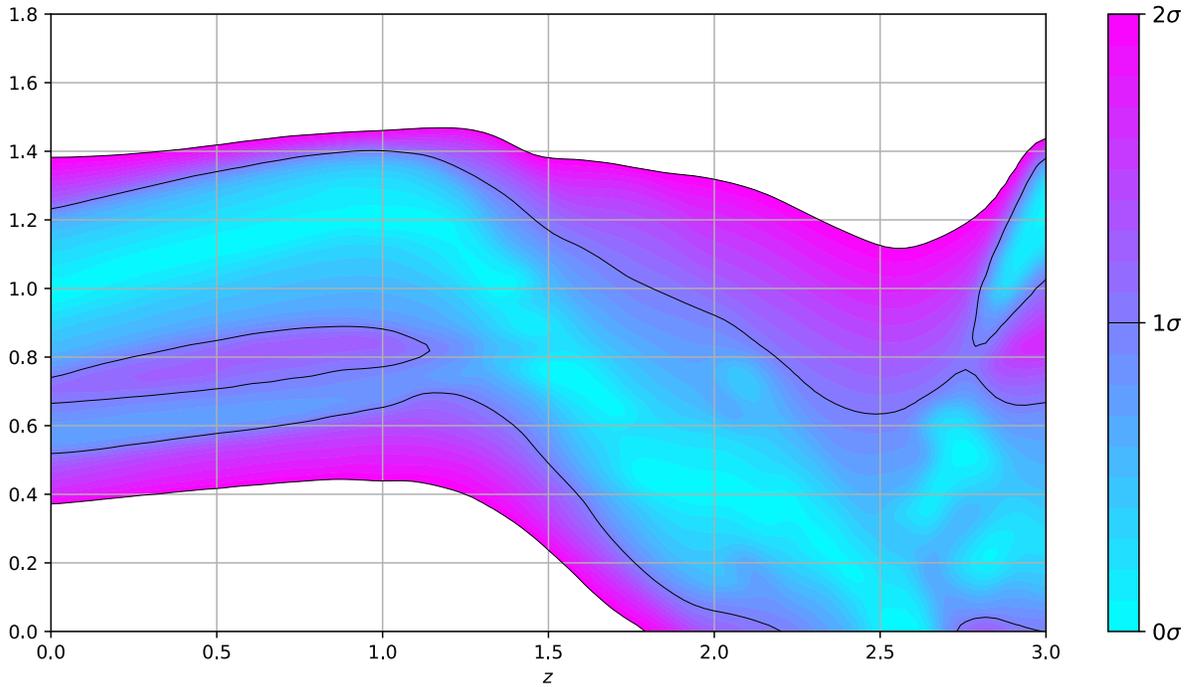


Figura 6.2: Distribución de probabilidad para la reconstrucción de $F_l(z)$ con interpolación spline cúbica. Los colores representan (como se observa en la barra de la derecha) en qué nivel de confianza se encuentra la reconstrucción, siendo el azul claro el mejor pues se encuentra cerca de $0 - \sigma$.

Después de esto se realizó otra reconstrucción, esta vez utilizando interpolación lineal y sin permitir que los puntos del eje z variaran. De manera que se tenía

$$F_l : (F_1, F_2, F_3, F_4, 1.0), \quad z : (0.0, 0.8, 1.6, 2.4, 3.0). \quad (6.2)$$

El último punto en el eje y se dejó fijo en $F_l(z) = 1$. Los valores de los parámetros se muestran en el cuadro 6.1, las distribuciones 1-D y 2-D se muestran en la figura 6.3 y la distribución de probabilidad de $F_l(z)$ se encuentra en la figura 6.4.

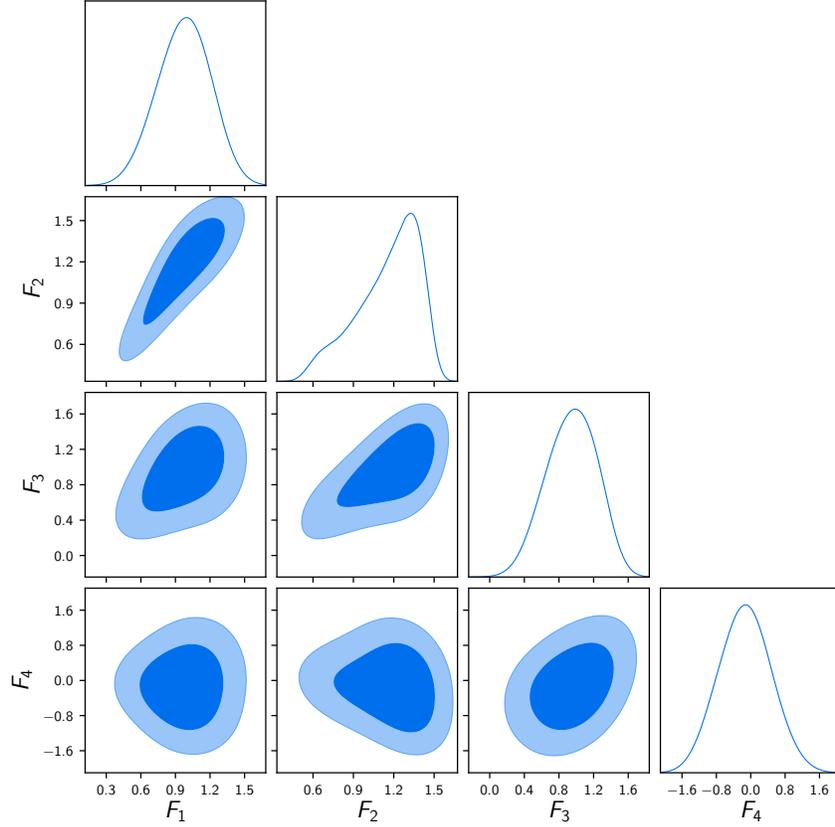


Figura 6.3: Distribuciones 1-D y 2-D de los parámetros para $F_l(z)$ con interpolación spline cúbica.

6.2. Reconstrucción de $w(z)$.

Se procedió a reconstruir la ecuación de estado de la energía oscura. Para esto se sustituyó $F_l(z)$ en (95) por $e^{-3 \int_1^a \frac{da'}{a'} (1+w(a'))} = e^{3 \int_0^z \frac{dz'}{1+z'} (1+w(z'))}$, por lo que tendremos la ecuación de Friedmann

$$H(z) = 100 \frac{\text{km}}{\text{sMpc}} \times h \left(\Omega_m (1+z)^3 + (1 - \Omega_m) e^{3 \int_0^z \frac{dz'}{1+z'} (1+w(z'))} \right)^{\frac{1}{2}}. \quad (6.3)$$

Esta reconstrucción de $w(z)$ se hizo con 5 parámetros en $w(z)$ y ninguno en z utilizando

interpolación spline cúbica, por lo que tenemos:

$$w(z) : (w_0, w_1, w_2, w_3, w_4), \quad z : (0.0, 0.8, 1.6, 2.4, 3.0). \quad (6.4)$$

En esta reconstrucción se utilizaron 6 cadenas que llegaron hasta 100,000 pasos cada una e igual que en la reconstrucción de $F_l(z)$ con interpolación spline cúbica se utilizaron datos observacionales de SN1a, CMB y BBAOs y se usaron probabilidades anteriores planas. Las distribuciones 1-D y 2-D de los parámetros w_i se encuentran en la figura 6.6, la distribución de probabilidad de $w(z)$ en la figura 6.5 y los valores de mejor ajuste de los parámetros en el cuadro 6.2.

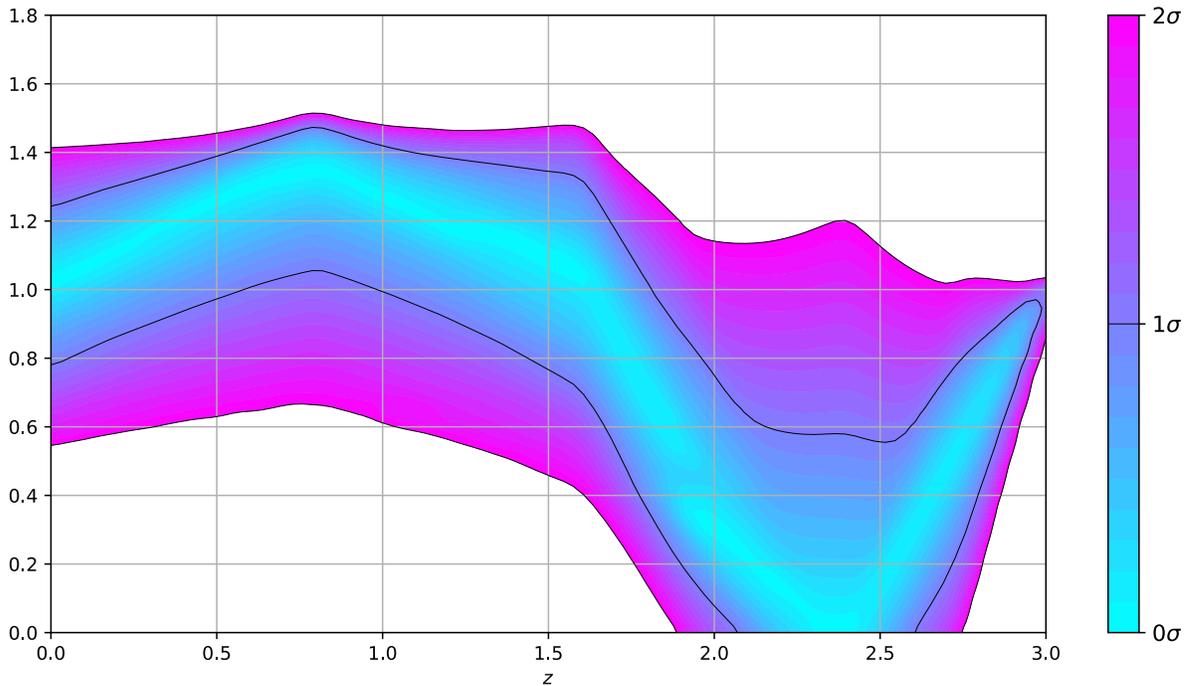


Figura 6.4: Distribución de probabilidad para la reconstrucción de $F_l(z)$ con interpolación lineal. Los colores representan (como se observa en la barra de la derecha) en qué nivel de confianza se encuentra la reconstrucción, siendo el azul claro el mejor pues se encuentra cerca de $0 - \sigma$.

Parámetros (spline cúbica)	Promedio \pm 1- σ	Intervalo de valores
Ω_m	0.28 ± 0.06	[0.05, 1.5]
h	0.72 ± 0.09	[0.4, 1.0]
F_1	0.89 ± 0.27	[-1.5, 1.5]
F_2	1.01 ± 0.3	[-1.5, 1.5]
F_3	0.35 ± 0.64	[-1.5, 1.5]
F_4	0.0 ± 0.8	[-1.5, 1.5]
z_1	1.09 ± 0.2	[0.2, 1.3]
z_2	2.16 ± 0.2	[1.7, 2.8]
χ^2	42.72	...
Parámetros (lineal)	Promedio \pm 1- σ	Intervalo de valores
Ω_m	0.3 ± 0.04	[0.05, 1.5]
h	0.69 ± 0.06	[0.4, 1.0]
F_1	0.98 ± 0.2	[-1.5, 1.5]
F_2	1.17 ± 0.2	[-1.5, 1.5]
F_3	0.96 ± 0.3	[-1.5, 1.5]
F_4	-0.1 ± 0.6	[-1.5, 1.5]
χ^2	42.91	...

Cuadro 6.1: Tabla con los parámetros utilizados en ambas reconstrucciones y sus valores de mejor ajuste. Al final se observa el valor de χ^2 , que nos dice qué tan bien se ajusta el modelo a los datos experimentales.

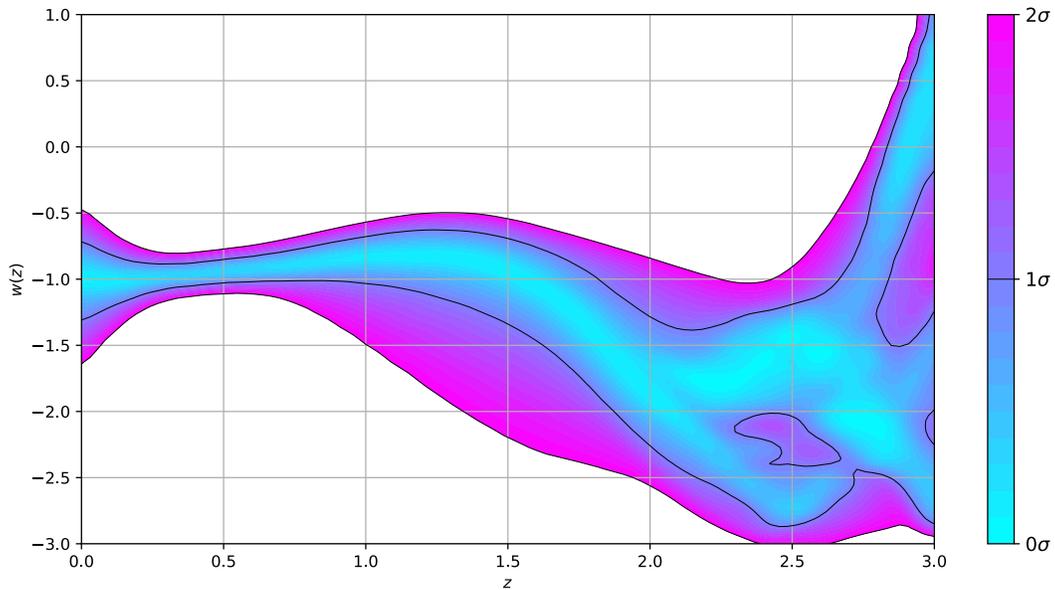


Figura 6.5: Distribución de probabilidad para la reconstrucción de $w(z)$ con interpolación spline cúbica.

Parámetros (spline cúbica)	Promedio \pm 1- σ	Intervalo de valores
Ω_m	0.3 ± 0.01	[0.05, 1.5]
h	0.67 ± 0.01	[0.4, 1.0]
w_0	-1.04 ± 0.3	[-3.0, 1.0]
w_1	-0.91 ± 0.1	[-3.0, 1.0]
w_2	-1.26 ± 0.4	[-3.0, 1.0]
w_3	-1.97 ± 0.6	[-3.0, 1.0]
w_4	-0.75 ± 1.3	[-3.0, 1.0]
χ^2	42.88	...

Cuadro 6.2: Tabla con los parámetros utilizados en la reconstrucción de $w(z)$ y sus valores de mejor ajuste.

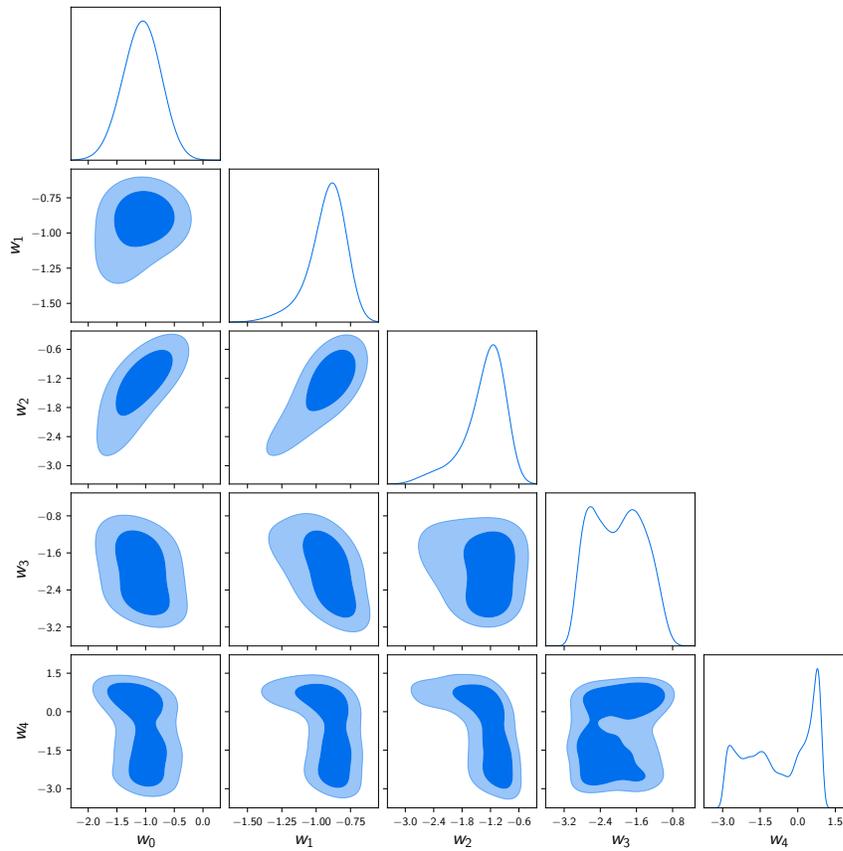


Figura 6.6: Distribuciones 1-D y 2-D de los parámetros para $w(z)$ con interpolación spline cúbica.

6.3. Comparación con Λ CDM

Para realizar una comparación contra Λ CDM se hizo una reconstrucción tomando una versión muy sencilla (como en las secciones anteriores) de este modelo dada por la ecuación de Friedmann

$$H(z) = 100 \frac{\text{km}}{\text{sMpc}} \times h \left(\Omega_m (1+z)^3 + (1 - \Omega_m) \right)^{\frac{1}{2}}. \quad (6.5)$$

Para este modelo se muestra el valor de mejor ajuste de los parámetros en el cuadro 6.3.

Parámetros (Λ CDM)	Promedio \pm 1- σ	Intervalo de valores
Ω_m	0.3 ± 0.01	[0.05, 1.5]
h	0.68 ± 0.01	[0.4, 1.0]
χ^2	47.01	...

Cuadro 6.3: Parámetros usados para Λ CDM.

Y ahora haremos una comparación entre las reconstrucciones y Λ CDM utilizando los criterios de manera similar a la sección anterior. En el cuadro 6.4 están los valores relevantes y la diferencia con cada criterio con respecto a Λ CDM.

Modelo	$\chi^2(\bar{\theta})$	$\overline{\chi^2(\theta)}$	Número de parámetros (k)	p_D	ΔAIC	ΔAICc	ΔBIC	ΔDIC
Λ CDM	47.01	49.94	2	2.94	0	0	0	0
$F_{luis(z)}$ cúbica	42.73	49.64	8	6.91	7.72	12.04	17.58	3.66
$F_{luis(z)}$ lineal	42.91	48.618	6	5.71	3.9	6.12	10.45	1.44
$w(z)$	42.98	50.06	7	7.08	5.97	9.15	14.16	4.25

Cuadro 6.4: Comparación de los modelos reconstruidos y el modelo Λ CDM. Se comparan de manera similar a la sección anterior con los 4 criterios antes mencionados.

Si utilizamos el criterio usual de la *escala de Jeffreys* entonces podemos ver en el cuadro 6.4 que dos de los criterios (AIC no se toma en cuenta pues sólo se utilizó para obtener AICc) tienen evidencia decisiva sobre la superioridad del modelo Λ CDM. En cambio, si

vemos el criterio DIC observamos que todos están debajo de $\Delta\text{IC} = -5$. Veamos cada reconstrucción por separado.

- $F_l(z)$ cúbica: Esta reconstrucción fue la peor de acuerdo a los criterios AICc y BIC. Este hecho no es sorpresa, pues ambos criterios castigan de manera muy fuerte el uso de muchos parámetros y esta reconstrucción es la que más tiene. Pero a pesar de esto es la que tiene la mejor χ^2 de todas las reconstrucciones, aunque también es esperado pues tiene una mayor libertad de ajuste debido al gran número de parámetros, pero acuerdo a DIC es buena opción, eso se debe a que DIC utiliza la complejidad bayesiana, y esto “disminuye” el número de parámetros libres en el modelo, lo que le ayuda mucho a esta reconstrucción (al menos en el DIC).
- $F_l(z)$ lineal: Esta reconstrucción fue la mejor entre las 3 de acuerdo a todos los criterios. Esto se debe a que es la que tiene menos parámetros (tiene 6, mientras las otras reconstrucciones tienen 7 y 8) y por lo mismo no sufrió tanto la penalización de los criterios. Pero es importante notar que ambas reconstrucciones de F_l tienen una gráfica similar (figura 6.2 y figura 6.4), ambas tienen un máximo cerca de $z = 1$ y un mínimo cerca de $z = 2.5$. Por lo que ambas sugieren un comportamiento similar de la energía oscura. En ambas, además, se puede ver que el comportamiento de constante cosmológica (en $F_l = 1$) queda en la región de confianza $2 - \sigma$ a partir de $z \approx 1.7$, lo que puede sugerir que si hay dinámica en la energía oscura o que se necesitan más observaciones en $z > 1.7$.
- $w(z)$: La reconstrucción de la ecuación de estado $w(z)$ es mala alternativa ($\Delta\text{IC} > 0$) a ΛCDM si usamos AICc y BIC, lo cual se espera pues tiene 7 parámetros libres y estos criterios castigan fuertemente el uso de muchos parámetros, y de acuerdo con DIC se observa que le falta poco para que haya evidencia fuerte ($\Delta\text{IC} < -5$) de que la ecuación de estado de ΛCDM (que corresponde a $w = -1$) sea mejor modelo para explicar las observaciones. Aún con estos valores de los criterios, se aprecia que la gráfica reconstruida de nuestra ecuación de estado (figura 6.5), no solo concuerda

con las dos gráficas anteriores en que a partir de $z \approx 1.7$ el modelo Λ CDM pasa a la región $2 - \sigma$, si no que en $z \approx 2.3$ sale de esta región.

6.4. Prueba de interpolaciones con datos reales.

Por último se repitió una comparación similar a la del ejemplo anterior pero utilizando, además de las observaciones ya mencionadas, datos de $H(z)$ obtenidas con cronómetros cósmicos [12, 13]. Por lo que tenemos datos de supernovas, BAOs, CMB y $H(z)$. Se realizaron las mismas 7 interpolaciones lineales y las mismas 4 interpolaciones spline cúbicas y, además, se utilizó la parametrización CPL. Los resultados de la comparación de modelos se muestran en el cuadro 6.5.

Modelo	$\chi^2(\bar{\theta})$	$\overline{\chi^2(\theta)}$	Núm. de params. (k)	p_D	ΔAIC	$\Delta AICc$	ΔBIC	ΔDIC
Λ CDM	71.58	75.62	2	4.04	0	0	0	0
CPL	71.18	78.26	4	5.08	3.6	3.96	8.36	1.68
$F_{l1}(z)$ lineal	72.9	76.83	3	3.93	3.32	3.47	5.7	1.1
$F_{l2}(z)$ lineal	72.58	77.14	4	4.56	5	5.36	9.77	2.04
$F_{l3}(z)$ lineal	65.84	73.5	5	7.66	0.26	0.89	7.41	1.5
$F_{l4}(z)$ lineal	67.08	75.42	6	8.34	3.5	4.47	13.03	4.1
$F_{l5}(z)$ lineal	66.52	73.66	6	7.14	2.94	3.91	12.47	1.14
$F_{l6}(z)$ lineal	61.96	67.66	7	5.7	0.38	1.74	12.29	-6.3
$F_{l7}(z)$ lineal	67.12	74.48	8	7.36	7.54	9.36	21.83	2.24
$F_{l2}(z)$ cúbica	72.58	77.14	4	4.56	5	5.36	9.26	2.04
$F_{l4}(z)$ cúbica	67.08	75.42	6	8.34	3.5	4.47	13.03	4.1
$F_{l6}(z)$ cúbica	61.96	68.24	7	6.28	0.38	1.74	12.29	-5.14
$F_{l7}(z)$ cúbica	67.12	75.04	8	7.92	7.54	9.36	21.83	2.3

Cuadro 6.5: Comparación de modelos para el ajuste a los datos reales de SN1a, CMB, BAOs y $H(z)$. Se prefieren los modelos con diferencia entre criterios negativa.

Se observa en el cuadro 6.5 que casi todas las reconstrucciones obtuvieron una mejor χ^2 que Λ CDM, siendo $F_{l6}(z)$ lineal y $F_{l6}(z)$ cúbica las mejores. Estas dos obtuvieron los únicos valores negativos (en ΔDIC), y si usamos la escala de Jeffreys ambas tienen evidencia

fuerte de ser mejores alternativas a Λ CDM (pues su Δ DIC está por debajo de -5) dadas las observaciones utilizadas.

Capítulo 7

Conclusiones y perspectivas.

7.1. Conclusión final

Nuestras tres primeras reconstrucciones con datos reales tienen evidencia fuerte de ser inferiores a Λ CDM en dos de los tres criterios (AICc y BIC), pero en DIC están en el margen de ser buenas alternativas o al menos similares. Aún dado esto todas tienen fuertes indicios de que hace falta un modelo de energía oscura que contenga dinámica, en especial a partir de $z \approx 1.7$. Además las reconstrucciones obtenidas en este trabajo se podrían utilizar para descartar modelos paramétricos que se salgan de las curvas de confianza reconstruidas, pues estas reconstrucciones fueron hechas dándoles libertad de moverse sin ningún modelo en específico y llegaron a resultados similares.

En la prueba final con las 7 interpolaciones lineales y 4 spline cúbicas se encontró que 2 de ellas son buenas alternativas (al menos de acuerdo a los datos que se utilizaron) al modelo Λ CDM. Esto aunado a las 3 reconstrucciones anteriores nos da una fuerte evidencia de que es necesario un modelo que considere energía oscura dinámica (dependencia de z) en lugar de una constante cosmológica.

7.2. Perspectivas

Se planea mejorar el código SimpleMC, permitiéndole usar perturbaciones (implementar CLASS), también se quieren actualizar los datos y observaciones a los más recientes de Planck 2018 y a los futuros de DESI. Se pretende utilizar también PCA para aprovechar las correlaciones vistas en algunos de los parámetros y de esa manera intentar reducir la dimensionalidad de las reconstrucciones.

Bibliografía

- [1] Baumann D., Cosmology Part III: Mathematical Tripos.
- [2] Ferguson J., Lecture Notes Part III: Cosmology, 2016.
- [3] Paul J. Steinhardt; A quintessential introduction to dark energy, Septiembre 2003.
- [4] <https://www.britannica.com/science/dark-energy>
- [5] https://ned.ipac.caltech.edu/level5/March01/Carroll/Carroll_contents.html
- [6] M. Pruzhinskaya, S. Lisakov, How supernovae became the basis of observational cosmology, 2016.
- [7] The Supernova Cosmology Project, SPECTRA AND HST LIGHT CURVES OF SIX TYPE IA SUPERNOVAE AT $0.511 < Z < 1.12$ AND THE Union2 COMPILATION, 2010.
- [8] O. Lahav, A. Liddle, The Cosmological Parameters 2014, 2014.
- [9] pdg.lbl.gov/2015/reviews/rpp2015-rev-cosmic-microwave-background.pdf
- [10] J. Eisenstein, et. al., Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies, 2005.
- [11] K. Dawson, et. al., The Baryon Oscillation Spectroscopic Survey of SDSS-III, 2011.

- [12] F. Melia, M. Yennapureddy, Model Selection Using Cosmic Chronometers with Gaussian Processes, 2018.
- [13] M. Moresco, et. al., Improved constraints on the expansion rate of the Universe up to $z \sim 1.1$ from the spectroscopic evolution of cosmic chronometers, 2013.
- [14] P. Bull, et. al., Beyond Λ CDM: Problems, solutions, and the road ahead, 2016.
- [15] A. Del Popolo, M. Le Delliou, Small Scale Problems of the Λ CDM Model: A Short Review, 2017.
- [16] Planck Collaboration, Planck 2018 results. VI. Cosmological parameters, 2018.
- [17] R. Caldwell, R. Dave, P. Steinhardt, Cosmological imprint of an energy component with general equation-of-state, 1998.
- [18] R. Caldwell, A phantom menace? Cosmological consequences of a dark energy component with super-negative equation of state, 2002.
- [19] W. Hu, Crossing the phantom divide: Dark energy internal degrees of freedom, 2005.
- [20] G. Zhao, et. al., Dynamical dark energy in light of the latest observations, 2017.
- [21] H. Velten, R. vom Marttens, W. Zimdahl, Aspects of the cosmological “coincidence problem”, 2014.
- [22] J. Lee, Dark energy, inflation and the cosmic coincidence problem, 2008.
- [23] S. Hossenfelder, Screams for Explanation: Finetuning and Naturalness in the Foundations of Physics, 2018.
- [24] V. Sahni, A. Starobinsky, Reconstructing Dark Energy, 2006.
- [25] A. Sangwan, A Mukherjee, H. Jassal, Reconstructing the dark energy potential, 2017.

- [26] J. Dai, Y. Yang, J. Xia, Reconstruction of the Dark Energy Equation of State from the Latest Observations, 2018.
- [27] E. Linder, RECONSTRUCTING AND DECONSTRUCTING DARK ENERGY, 2004.
- [28] D. Huterer, D. Shafer, Dark energy two decades after: observables, probes, consistency tests, 2017.
- [29] D. Huterer, G. Starkman, Parametrization of Dark-Energy properties: A Principal-Component Approach, 2003.
- [30] T. Holsclaw, et. al., Nonparametric Reconstruction of the Dark Energy Equation of State, 2010.
- [31] B. Feng, et. al., Oscillating Quintom and the Recurrent Universe, 2006.
- [32] B. Bassett, et. al., A late-time transition in the cosmic dark energy?, 2002.
- [33] A. Starobinsky, How to determine an effective potential for a variable cosmological term, 1998.
- [34] R. Daly, S. Djorgovski, A Model-independent Determination of the Expansion and Acceleration Rates of the Universe as a Function of Redshift and Constraints on Dark Energy, 2003.
- [35] A. Shafieloo, et. al., Smoothing Supernova Data to Reconstruct the Expansion History of the Universe and its Age, 2007.
- [36] H. Jeffreys, Theory of probability, 3era edición, 1998.
- [37] J. Bernardo and A. Smith, Bayesian Theory, 1994.
- [38] R. Trotta, Bayes in the sky: Bayesian inference and model selection in cosmology, 2008.

- [39] A. Liddle, Information criteria for astrophysical model selection, 2007.
- [40] D. Spiegelhalter, et. al., Bayesian measures of model complexity and fit, 2002.
- [41] K. Burnham, D. Anderson, Multimodel inference: understanding AIC and BIC in model selection, 2004.
- [42] [https : //getdist.readthedocs.io/en/latest/](https://getdist.readthedocs.io/en/latest/)
- [43] [https : //github.com/ja - vazquez/SimpleMC](https://github.com/ja-vazquez/SimpleMC)