

Proyecto Cosmología 3

Uriel Chávez Flores

20 de junio de 2022

Resumen

En este proyecto usamos cadenas de Montecarlo Markov (MCMC) para ajustar datos experimentales de mediciones de módulo de distancia en función de corrimiento al rojo de supernovas. Se consideraron diferentes variaciones del modelo $\Lambda - CDM$ para ajustes de diferentes parámetros cosmológicos. El primero ajusta el parámetro de densidad de materia Ω_m y el parámetro de densidad de energía oscura Ω_Λ asumiendo un universo con curvatura nula y usando la densidad de energía de la radiación Ω_{CMB} como el parámetro libre para asegurar la curvatura nula. El segundo ajusta Ω_m y Ω_Λ sin asumir curvatura plana y fijando $\Omega_{CMB} = 0$. El tercero ajusta la constante de Hubble H_0 y Ω_m asumiendo curvatura plana y fijando $\Omega_{CMB} = 0$. Y por último, el cuarto ajusta Ω_Λ y los parámetros ω_0 y ω_1 de la ecuación de estado de la energía oscura con forma funcional $\omega = \omega_0 + (1 - a)\omega_1$. Los resultados de los ajustes se muestran en gráficas triangulares y se hace un análisis cualitativo intentando encontrar las razones de los resultados.

1. Introducción

La investigación científica usa métodos estadísticos repetidamente en una iteración alternada de obtención controlada de datos y el análisis de dichos datos. Mientras que el análisis de los datos es una subiteración en la que la inferencia de un modelo provisionalmente introducido se alterna con la con la crítica de la inferencia mediante la inspección de sus diferencias con los datos y otros medios. [2]

1.1. Inferencia Bayesiana

La inferencia Bayesiana es uno de estos métodos de análisis que parece ofrecer la posibilidad de una flexibilidad suficiente para permitir la reacción a la complejidad científica sin el impedimento de una limitación puramente técnica. [2]

La inferencia Bayesiana involucra probabilidades condicionales, que las calcula usando el teorema de Bayes.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Con este teorema se calcula la probabilidad de un evento dada información a priori. En la ciencia, esta información a priori suele ser el valor aproximado de un parámetro del que queremos conocer su valor más probable dado un conjunto de datos experimental. La forma en la que se nombra a cada componente del teorema de Bayes es la siguiente:

$$\text{Posterior} = \frac{\text{Likelihood} * \text{Prior}}{\text{Evidence}}$$

El posterior es la medida de qué tan probable es un modelo dados los datos experimentales, entonces, lo que nos interesa es maximizarlo en función de los parámetros que determinan el modelo. El modelo se tiene que ajustar a los datos, no al revés.

Después, tenemos el likelihood, que es la probabilidad de obtener los datos experimentales dado un modelo, lo cual es calculable asumiendo la forma en la que los datos se distribuyen. En muchos fenómenos medibles se asume una distribución normal (aunque existen otras más que describen otros fenómenos de diferentes naturalezas).

El prior representa lo que conocemos o desconocemos de los parámetros, en términos prácticos es una distribución donde suponemos de antemano, puede estar el valor real del parámetro. Una distribución usada comúnmente (y que usamos en este proyecto) es el prior plano, que asigna probabilidad uniforme a un rango de valores de algún parámetro.

Y por último, la evidencia es la probabilidad de que se midan los datos experimentales, lo cual, en principio no conocemos, pero termina siendo irrelevante pues lo que nos interesa es maximizar el posterior en función de los parámetros, y la evidencia no depende de los parámetros. Por ello función que nos interesa maximizar es solamente la multiplicación de likelihood y el prior.

$$\text{Posterior} \propto \text{Likelihood} * \text{Prior}$$

El likelihood asumiendo distribución normal tiene la siguiente forma.

$$L = \prod_i^n \frac{1}{2\pi\sigma_i^2} \exp\left(-\frac{(y_i - \mu(x_i, \theta))^2}{2\sigma_i^2}\right)$$

Donde x_i, y_i son los puntos experimentales, σ_i sus desviaciones, θ los parámetros que determinan el modelo y μ el modelo que se ajusta.

1.2. Cadenas de Montecarlo Markov

La maximización del posterior será de la dimensión de los parámetros que se estén ajustando, una forma aproximarse a esta maximización es mediante las cadenas de Montecarlo Markov (MCMC). Las MCMC son un método de exploración del espacio de parámetros que comienza con una propuesta del vector de parámetros para la cual se calcula el posterior, luego se van buscando nuevos vectores de parámetros candidatos (en la vecindad del vector anterior) a reemplazar la propuesta anterior. Los candidatos reemplazarán en función de si su posterior es mayor o menor al posterior anterior, añadiendo un factor de aleatoriedad al tomar la decisión de reemplazo en base a una probabilidad calculada como la razón entre el posterior actual y el del candidato.

Un ejemplo básico realizado con los programas creados en este proyecto es el de la línea recta. Se generaron puntos con distribución normal sobre una recta para tomarlos como los datos a los que se ajustará un modelo. El modelo que se ajusta es el de la línea recta, por lo que tiene dos parámetros libres, la pendiente y la ordenada al origen. Los resultados de este ajuste de ejemplo son los siguientes.

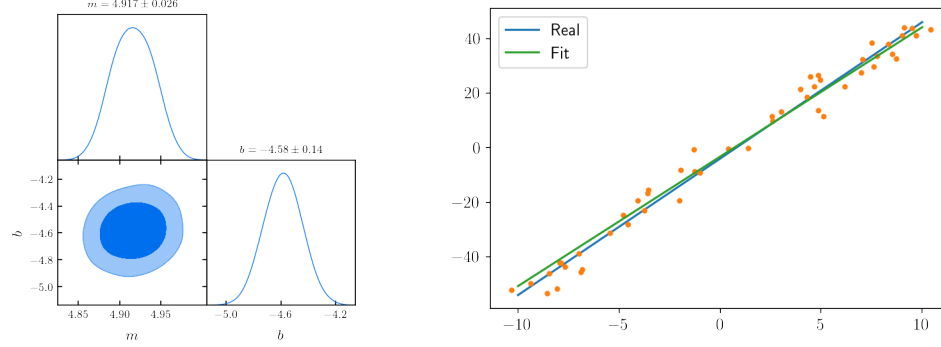


Figura 1: Gráfica triangular y comparación de modelo con datos experimentales de ajuste ejemplo.

Tras la exploración del espacio de parámetros se encuentra una zona en donde el posterior fue maximizado. Por la naturaleza de la exploración del espacio de parámetros se obtiene una distribución alrededor de el máximo. Para evitar que las cadenas se queden estancadas en máximos locales, se suelen ejecutar varias cadenas con posiciones iniciales diferentes, de forma que se explore el espacio de parámetros casi por completo.

1.3. Objetivo del proyecto

El objetivo del proyecto es usar un conjunto de datos de supernovas [3] y explorar qué parámetros cosmológicos podemos ajustar con ellos. En total se hicieron 4 ajustes diferentes de diferentes parámetros cosmológicos bajo diferentes condiciones del modelo $\Lambda - CDM$. La descripción de cada ajuste se explica detalladamente en la siguiente sección.

2. Métodos

El modelo al que se ajustaron los datos experimentales el módulo de distancia, para calcularlo en función de los parámetros cosmológicos se usó la ecuación de Friedmann, que despejando \dot{a}^2 tenemos la siguiente ecuación. [4]

$$\dot{a}^2 = H_0 \sum_{\omega} \Omega_{\omega,0} a^{-1-3\omega}$$

Donde el índice ω corresponde a cada tipo de energía en el universo. Para estos ajustes se consideraron solamente 3 componentes, energía oscura, materia (materia oscura fría y materia bariónica) y radiación del CMB.

$$\dot{a} = H_0 \sqrt{\Omega_{m,0} a^{-1} + \Omega_{CMB,0} a^{-2} + \Omega_{\Lambda,0} a^2}$$

Por lo tanto, el modelo que usado solo tiene como parámetros libres la constante de Hubble H_0 y los parámetros de densidad Ω_m , Ω_Λ y Ω_{CMB} .

Para obtener la función $\mu(z)$ que es el modelo que para los ajustes calculamos la distancia propia integrando numéricamente.

$$d_p(t_0) = c \int_{z'}^0 \frac{a(z)}{\dot{a}(z)} dz$$

Teniendo la distancia propia se puede calcular el módulo de distancia. [4]

$$\mu(z) = 5 \ln(d_p(t_0)(1+z)) + 25$$

Para la ejecución de las MCMC se configuraron los límites de prior plano dados los parámetros que se explorarían en cada ajuste. En todos los ajustes se crearon 10 cadenas de 10000 pasos con el MCMC programado para este proyecto y con la librería emcee, los resultados de emcee se muestran solamente cuando difieren significativamente de los obtenidos con el MCMC propio.

Para la visualización de los resultados de cada ajuste se hicieron gráficas triangulares y una gráfica de comparación del modelo con los parámetros obtenidos con los datos experimentales [3] y con el modelo con los parámetros obtenidos en la colaboración Planck 2018 [1].

2.1. Ajuste 1

El modelo $\Lambda - CDM$ no impone algún valor de curvatura, pero las observaciones cosmológicas apuntan a que la curvatura es plana o muy cercana a plana. En este ajuste se asumen curvatura plana. Se ajustan los parámetros de densidad Ω_m y Ω_Λ , mientras que Ω_{CMB} queda variable para asegurar la curvatura plana, es decir, que se cumpla la siguiente ecuación.

$$\Omega_m + \Omega_\Lambda + \Omega_{CMB} = 1$$

El prior usado fue modificado para rechazar vectores de parámetros tales que $\Omega_m + \Omega_\Lambda > 1$.

2.2. Ajuste 2

A diferencia del primer ajuste, este no asume curvatura plana, y solo ajusta los parámetros Ω_m y Ω_Λ . Ya que los primeros resultados obtenidos no coincidían con los esperados, se repitió el mismo ajuste con funciones de astropy para obtener el modelo, por ello en la sección de resultados se muestran 4 ajustes en total y no 2 como en el resto.

2.3. Ajuste 3

Los parámetros que se ajustan son H_0 y Ω_Λ , la razón de esta elección es en el proyecto 2 se observó que estos dos parámetros son los que más afectan del grupo de 6 parámetros que se eligieron en dicho proyecto. Es razonable pues las mediciones de módulo de distancia de supernovas han mostrado la expansión acelerada del universo, y justamente H_0 es una medida del ritmo de expansión, y por otro lado, la materia oscura Ω_Λ es la encargada de la expansión acelerada.

2.4. Ajuste 4

En el modelo $\Lambda - CDM$ la ecuación de estado es proporcional a una potencia de el factor de escala a en donde está involucrado el parámetro ω que antes vimos en la ecuación de Friedmann.

$$\rho \propto a^{-3(1+\omega)}$$

Para este ajuste se propone una forma funcional específica para el parámetro ω de la energía oscura.

$$\omega = \omega_0 + (1 - a)\omega_1$$

Los dos parámetros de los que depende fueron ajustados junto al parámetro de densidad Ω_Λ .

2.5. Resumen de descripción de ajustes

En la tabla 1 se muestran los parámetros y los límites de prior plano usados para cada ajuste.

	Parámetros	Limites	Curvatura	Ω_{CMB}
Ajuste 1	Ω_m	0, 1	Plana	Variable
	Ω_Λ	0, 1		
Ajuste 2	Ω_m	0, 1.5	Variable	0
	Ω_Λ	0, 1.5		
Ajuste 3	H_0	50, 100	Plana	0
	Ω_Λ	0, 1		
Ajuste 4	Ω_Λ	0, 1	Plana	0
	ω_0	-3, 1		
	ω_1	-2, 2		

Tabla 1: Parámetros y límites de prior plano para cada ajuste.

Los códigos usados para la ejecución de las MCMC y las gráficas creadas se almacenan en un repositorio de GitHub. El enlace al repositorio es el siguiente:

<https://github.com/ChavezUriel/CosmoProj3>

3. Resultados y Discusión

En las figuras siguientes se mostraran las trayectorias de las cadenas, las cadenas recortadas sin las primeras iteraciones, la gráfica triangular y la comparación del ajuste con el modulo de distancia obtenido de los parámetros reportados por la colaboración Planck 2018 [1] de cada ajuste.

3.1. Ajuste 1

En los resultados del ajuste 1 con la MCMC propia podemos ver una clara correlación inversa entre Ω_m y Ω_Λ . La magnitud de Ω_Λ concuerda con la esperada, pero la magnitud de Ω_m , por la restricción de curvatura que se impuso, está fuera de lo esperado, pues el parámetro de radiación Ω_{CMB} es mucho más grande de lo que debería.

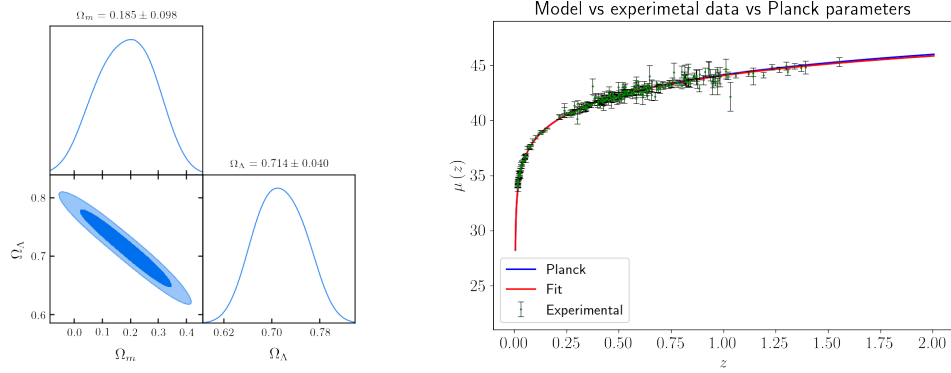


Figura 2: Gráfica triangular y comparación de modelo con datos experimentales de ajuste 1 con MCMC propia.

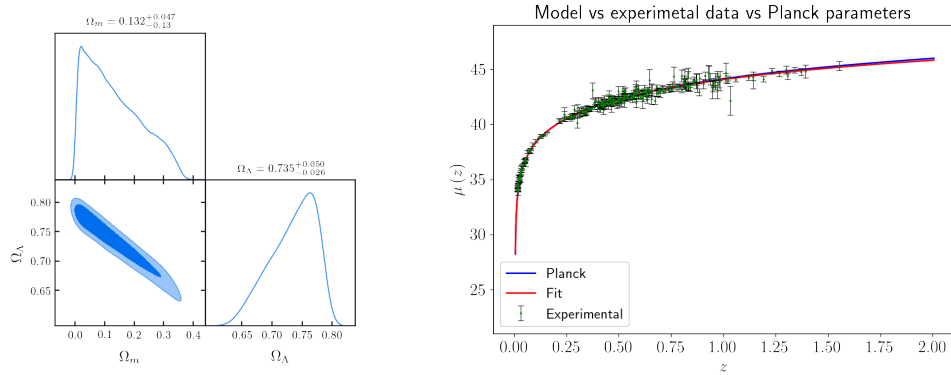


Figura 3: Gráfica triangular y comparación de modelo con datos experimentales de ajuste 1 con emcee.

Por otra parte, el resultado del ajuste con emcee también refleja correlación inversa, pero tiene una forma extraña, probablemente es porque el prior modificado implique una restricción muy fuerte para la forma en que funcionan las cadenas de emcee. Aparte el resultado de Ω_m es todavía menor que el ajuste con las MCMC propias. A pesar de la diferencia en las formas de los ajustes, ambas curvas ajustadas no difieren mucho de los datos experimentales y del modelo con los parámetros de Planck 2018.

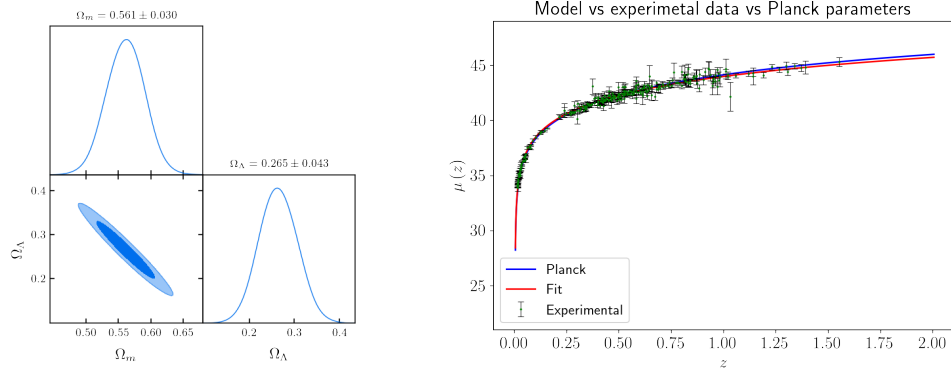


Figura 4: Gráfica triangular y comparación de modelo con datos experimentales de ajuste 2 con MCMC propia y función $\mu(z)$ propia.

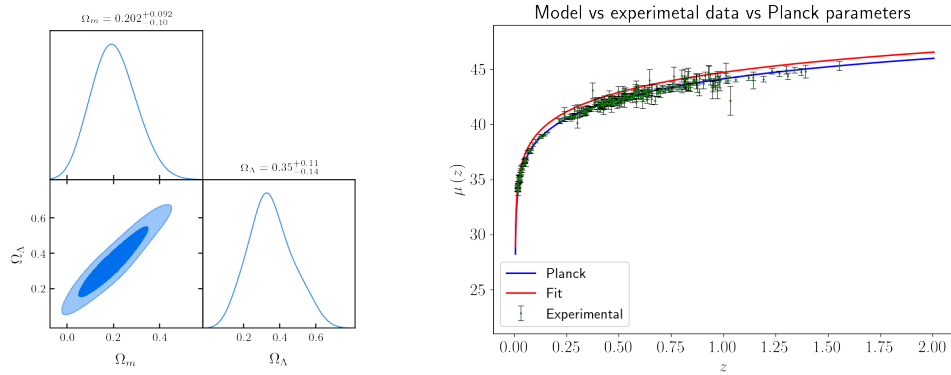


Figura 5: Gráfica triangular y comparación de modelo con datos experimentales de ajuste 2 con MCMC propia y función $\mu(z)$ astropy.

3.2. Ajuste 2

En el ajuste 2 con las funciones de modulo de distancia $\mu(z)$ propias obtenemos un resultado muy lejos de lo esperado (tanto con las MCMC propias como con emcee), las magnitudes de Ω_m y Ω_Λ están muy lejanas a los resultados de Planck 2018. Por ello se decidió probar con las funciones de astropy, aunque no con resultados mucho más convincentes, pues las magnitudes también difieren pero al menos Ω_Λ es mayor a Ω_m . La gran diferencia es que con funciones propias de $\mu(z)$ hay correlación inversa, y con las funciones de astropy hay correlación.

A pesar de la diferencia de resultados, las gráficas comparativas entre los modelos y los datos experimentales no son tan diferentes, probablemente sea debido a que esta observable es poco sensible a las variaciones de estos parámetros considerando la calidad de ajuste que estamos realizando.

3.3. Ajuste 3

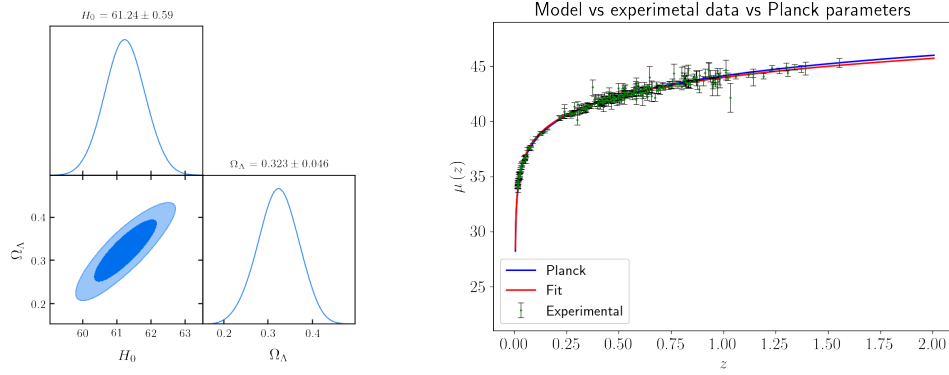


Figura 6: Gráfica triangular y comparación de modelo con datos experimentales de ajuste 3 con MCMC propia.

En el ajuste 3 se obtiene un resultado relativamente cercano al esperado en el parámetro H_0 , pero lejano en el parámetro Ω_Λ . Pero como se ha visto en el resto de ajustes, el modelo no difiere mucho de los puntos experimentales.

3.4. Ajuste 4

Por último, en el ajuste 4 vemos que los resultados con las MCMC propias y emcee son bastante parecidos tanto en forma como en magnitudes, sin embargo no se obtienen resultados parecidos a los ajustados en Planck 2018, en donde el parámetro $\omega_0 \approx 1$ y $\omega_1 \approx 0$.

Una posible razón de esta diferencia tan grande puede ser que el conjunto de datos al que se ajusta en Planck 2018, en el que incluyen datos de BAO (oscilaciones acústicas de bariones), sea mucho más restrictivo que el conjunto de datos aquí usado.

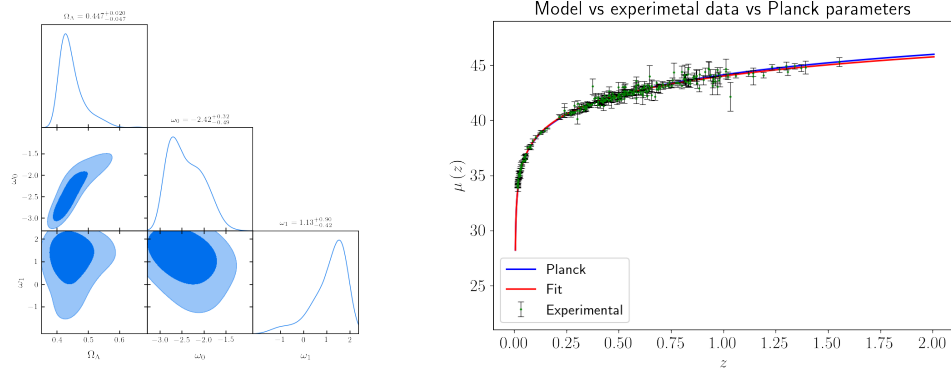


Figura 7: Gráfica triangular y comparación de modelo con datos experimentales de ajuste 4 con MCMC propia.

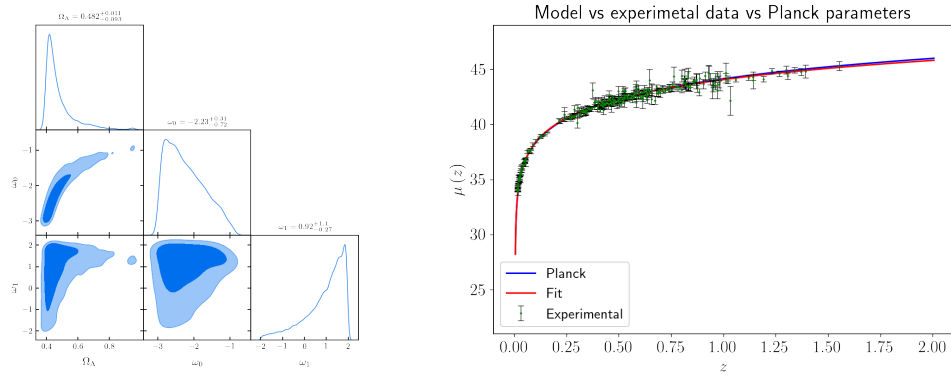


Figura 8: Gráfica triangular y comparación de modelo con datos experimentales de ajuste 4 con emcee.

4. Conclusiones

En los tres primeros ajustes, las gráficas triangulares muestran distribuciones muy cercanas a gaussianas, pero vemos que no hay compatibilidad entre la correlación inversa que se encuentra en el primero y segundo ajuste con las funciones propias de $\mu(z)$, y la correlación del ajuste 2 con funciones de astropy. Habiendo hecho varias pruebas con las que se intentó entender el efecto de la modificación al prior y modificaciones del modelo para ajustar el valor del parámetro de densidad del CMB, pienso que debe haber alguna diferencia importante en las funciones de $\mu(z)$ propias y de astropy.

En los resultados preeliminares noté que algunas distribuciones cambiaban mucho cada vez que se ejecutaba el programa, por ello decidí rehacer cada ajuste con la cantidad de iteraciones suficientes para no depender de las posiciones iniciales de las cadenas. Quizá esto mismo pueda ser una razón de las distribuciones tan irregulares que se obtuvieron en el ajuste 4, probablemente los límites del prior sean o muy grandes o muy pequeños para hacer el ajuste óptimo.

Referencias

- [1] N. Aghanim et al. Planck 2018 results. VI. Cosmological parameters. *Astron. Astrophys.*, 641:A6, 2020. [Erratum: *Astron.Astrophys.* 652, C4 (2021)].
- [2] G. E. P. Box and G. C. Tiao. *Bayesian inference in Statistical Analysis*. Wiley, 2009.
- [3] M. Kowalski et al. Improved Cosmological Constraints from New, Old and Combined Supernova Datasets. *Astrophys. J.*, 686:749–778, 2008.
- [4] B. Ryden. *Introduction to cosmology*. Cambridge University Press, 1970.