



**Universitat de les
Illes Balears**

Facultat de Ciències

Memòria del Treball de Fi de Grau

Medición de ondas gravitacionales de sistemas binarios de agujeros negros con masas intermedias

Alicia Calafat Jaso

Grau de Física

Any acadèmic 2018-19

Treball tutelat per Sascha Husa
Departament de Física

S'autoritza la Universitat a incloure aquest treball en el Repositori Institucional per a la seva consulta en accés obert i difusió en línia, amb finalitats exclusivament acadèmiques i d'investigació	Autor		Tutor	
	Sí	No	Sí	No
	X		X	

Paraules clau del treball:
Ondas gravitacionales, análisis de datos, muestreo anidado

Motivación.

En los últimos años las ondas gravitacionales han ocupado gran parte del interés científico y mediático. Estas ondulaciones en el tejido espacio-temporal nos aportan información sobre sus dramáticos orígenes y sobre la naturaleza de la gravedad que no podría obtenerse de otra manera. Respecto a la primera detección directa de ondas gravitacionales producida por la fusión de dos agujeros negros a unos 1300 millones de años luz de la Tierra, GW150914, el profesor emérito de física en el MIT Rainer Weiss se pronunciaba de la siguiente manera: “La descripción de esta observación se encuentra bellamente descrita en la teoría de la relatividad general de Einstein formulada hace 100 años y comprende la primera prueba de la teoría de la interacción gravitacional fuerte. Hubiera sido maravilloso poder ver la cara de Einstein si hubiéramos tenido la oportunidad de decírselo.” [39]

Einstein publicó en 1915 la teoría de la relatividad general presentándonos las ecuaciones de campo de Einstein, que relacionan la presencia de materia con la curvatura del espacio-tiempo; en total constituyen 10 ecuaciones diferenciales parciales acopladas. Poco más de un mes después, Schwarzschild se apresuró a escribirle una carta: “A fin de volverme más versado en su teoría gravitacional, me ocupé más de cerca del problema planteado en su artículo sobre el perihelio de Mercurio. [...] Tomé mis riesgos y tuve la buena suerte de encontrar una solución. Un cálculo no demasiado extenso me entregó el siguiente resultado” [12], la métrica de Schwarzschild; una solución a las ecuaciones de campo de Einstein para una distribución esférica de materia, en el vacío, estática e invariante bajo inversión temporal. Esta solución describía lo que se conocería 50 años más tarde como un agujero negro. En 1916, Einstein desarrolló los fundamentos de la teoría linealizada de la gravedad al aplicar sus ecuaciones a situaciones de campo débil. Al estudiar esta teoría linealizada, se percató de la posible existencia de las ondas gravitacionales.

Durante más de dos siglos, la mecánica newtoniana había dominado el escenario científico compuesto por el espacio y el tiempo como dos entidades separadas, en las cuales la materia y la energía interactuaban como los actores en un teatro. Con la teoría de la relatividad general, el escenario espacio-temporal se pasaba al elenco de actores conformando una única identidad inseparable cuya doctrina ya no era independiente del resto de la física. La revolución científica einsteniana no había hecho más que empezar.

Abstract.

Los agujeros negros de masa intermedia (IMBH) son cruciales para comprender la evolución de los agujeros negros supermasivos (SMBHs) y para poder establecer vínculos con otros fenómenos observados como los sistemas de fuente X ultraluminosa. Aunque se cree que existen, su evidencia directa ha estado ausente por mucho tiempo. Motivados por las exitosas detecciones llevadas a cabo por LIGO y Virgo así como por el tercer periodo de observación O3 en curso, en este trabajo se estudian hipotéticas detecciones de los IMBH más ligeros, los LIMBH de masas $\gtrsim 100M_{\odot}$ cuya probabilidad de detección es mucho mayor que los IMBH. En consecuencia, se han inyectado varias señales razonables procedentes de binarias de agujeros negros LIMBH en el ruido de los dos detectores simulados de LIGO y se ha comprobado, para cada señal, la recuperación de los parámetros a través de una estimación utilizando el algoritmo de muestreo anidado. La perspectiva de establecer la presencia de LIMBHs en el tercer periodo de observación O3 en un futuro cercano es bastante prometedora, implicando otro gran avance en la astronomía de las ondas gravitacionales.

Índice general

Motivación.	III
Abstract.	III
1. Introducción a las ondas gravitacionales.	1
1.1. Ecuaciones de campo de Einstein.	1
1.2. Agujeros negros.	1
1.2.1. La solución de Schwarzschild.	2
1.2.2. Tipos de agujeros negros.	2
1.2.3. Agujeros negros de masa intermedia.	2
1.3. Ondas gravitacionales.	3
1.3.1. Categorías de ondas gravitacionales.	4
1.3.1.1. Ondas gravitacionales producidas por CBCs.	4
2. Análisis de datos de ondas gravitacionales para CBCs.	7
2.1. Modelos de ondas gravitacionales para CBCs.	7
2.1.1. El modelo IMRPhenomPv2.	8
2.2. Detección de ondas gravitacionales.	8
2.2.1. Señal de una onda gravitacional observada por el detector.	8
2.2.2. Los datos obtenidos del detector.	9
2.3. Análisis bayesiano.	10
2.3.1. Conceptos fundamentales.	10
2.3.1.1. El teorema de Bayes.	10
2.3.1.2. La función de verosimilitud.	11
2.3.1.3. La distribución de probabilidad a priori.	11
2.3.1.4. La distribución de probabilidad a posteriori.	12
2.3.1.5. La evidencia.	12
2.3.2. Muestreadores.	13
2.3.2.1. MCMC.	13
2.3.2.2. Muestreo anidado.	14
2.3.2.2.1. Dynesty.	16
3. Resultados.	19
3.1. Resultados.	19
3.1.1. El programa.	19
3.1.2. La evidencia, el número de muestras y las SNRs.	20
3.1.3. Análisis de las distribuciones de probabilidad a posteriori.	21
4. Conclusiones.	27
Bibliografía	29

Introducción a las ondas gravitacionales.

Sección 1.1. Ecuaciones de campo de Einstein.

Sección 1.2. Agujeros negros.

Sección 1.3. Ondas gravitacionales.

1.1. Ecuaciones de campo de Einstein.

Albert Einstein revolucionó nuestra comprensión del universo cuando en 1905 presentó la teoría especial de la relatividad y posteriormente en el año 1915, la teoría general de la relatividad.

Las transformaciones de Lorentz, introducidas para resolver inconsistencias entre la mecánica clásica y el electromagnetismo, junto con los trabajos de Henri Poincaré, permitieron a Einstein asentar las bases para su teoría de la relatividad. Einstein fundamentó la teoría de la relatividad especial en un principio de relatividad, el cual sugiere que no hay un estado de movimiento o marco referencial privilegiado y en la constancia de la velocidad de la luz en cualquier sistema de referencia inercial [18].

Esta teoría descarta explícitamente las acciones instantáneas a distancia, hecho incompatible con la teoría de la gravitación de Newton que explicaba la gravedad mediante acciones instantáneas a distancia. Además, aunque la teoría de la relatividad especial es generalizable a sistemas acelerados, Einstein la concibió como una teoría aplicable únicamente a sistemas de referencia inerciales. Ambos hechos le condujeron a la formulación de una nueva teoría que proporcionase las descripciones físicas adecuadas.

Einstein describió la gravedad como el efecto de la geometría del espacio-tiempo sobre la materia, en el cual las partículas adquieren sus órdenes de movimiento localmente siguiendo la trayectoria más recta posible: la de la geodésica [7]. Es decir, el espacio-tiempo actúa sobre la materia diciéndole cómo moverse. A su vez, la materia actúa sobre el espacio-tiempo diciéndole cómo curvarse [28]. En otras palabras, la dinámica del contenido del universo está gobernada por la estructura geométrica del espacio-tiempo; y a su vez este contenido es fuente de las deformaciones (como la curvatura, topología o la conexión) del espacio-tiempo [32]. La conexión entre materia/energía y geometría, es establecida por las ecuaciones de Einstein como

$$G^{\mu\nu} = \frac{8\pi G_N}{c^4} T^{\mu\nu}, \quad (1.1)$$

donde el lado izquierdo contiene las componentes $G^{\mu\nu}$ de un tensor construido a partir de la geometría del espacio-tiempo, el tensor de Einstein G , simétrico, de rango 2 y de divergencia nula; mientras que el derecho contiene la descripción del contenido que habita en el espacio-tiempo expresado por la geometría, siendo G_N la constante gravitacional de Newton, c la velocidad de la luz en el vacío y $T^{\mu\nu}$ ¹ las componentes del tensor de energía-momento T , simétrico, de rango 2 y de divergencia nula por la conservación de energía-momento [32].

1.2. Agujeros negros.

Los agujeros negros tienen su origen en el colapso de una estrella vieja. Cuando se agotan todas las fuentes termonucleares de energía, si la estrella es suficientemente pesada colapsará [31]. Es decir, al llegar a la etapa final de la fusión nuclear del hidrógeno, debido a la escasez de protones libres, la presión producida por la fusión se vuelve insuficiente para soportar la fuerza gravitacional. Esto provoca una explosión violenta de la capa exterior de la estrella, dejando un cuerpo celeste denso (de entre 10^3 y 10^{11} kg/cm^3) el cual, dependiendo de la masa de la estrella original, puede ser una enana blanca o una estrella de neutrones. Las estrellas de neutrones están compuestas por neutrones degenerados mientras que las enanas blancas contienen electrones degenerados. La degeneración cuántica impide, por el principio de exclusión de Pauli, el colapso total del material en $r = 0$ [32]. Esta estabilidad deja de ser sostenible cuando la masa total es mayor o bien al límite de Chandrasekhar, en el caso de enanas blancas, equivalente a $1.44M_\odot$ o bien al límite de Tolman-Oppenheimer-Volkoff, en el caso de estrellas de neutrones, equivalente a aproximadamente $2M_\odot$ [3]. Al sobrepasar estos límites, el destino final inevitable de la estrella es el colapso gravitacional catastrófico. Según la relatividad general, el colapso gravitacional completo de un cuerpo producirá una región de agujero negro de espacio-tiempo

¹El tensor de energía-momento es el encargado de establecer una conexión entre la fuente material de la gravedad y el espacio-tiempo con curvatura.

[43]. De esta manera, se define un agujero negro como una región del espacio-tiempo donde la gravedad es lo suficientemente fuerte como para que ninguna partícula o radiación electromagnética, como la luz, pueda escapar de ella [43].

1.2.1. La solución de Schwarzschild.

La solución de Schwarzschild es considerada la primera solución exacta, no trivial, de las ecuaciones de campo de Einstein. Esta solución describe un sistema con las siguientes propiedades: estático e invariante bajo inversión temporal ($t \rightarrow -t$), dotado de simetría esférica, y ocurre en el vacío [32].

La obtención de una solución a las ecuaciones de Einstein, implica conseguir una expresión más general de la métrica o, análogamente, del intervalo o elemento de línea que se ajuste a las cualidades buscadas. El intervalo de Schwarzschild buscado es, [32]

$$ds^2 = \left(1 - \frac{2M}{r}\right) dt^2 - \left(1 - \frac{2M}{r}\right)^{-1} dr^2 - r^2 d\Omega^2 \quad (1.2)$$

donde $d\Omega^2$ es el diferencial de ángulo sólido y M es la masa de un sistema gravitacional esféricamente simétrico $M = \int d^3x \rho_m = \int dr 4\pi r^2 \rho_m$, siendo ρ_m la densidad de energía equivalente a densidad de masa.

Esta métrica de Schwarzschild presenta dos singularidades radiales: $r = 0$, la singularidad; y $r = r_s \equiv 2M$, en unidades naturales $G_N = c = 1$, que es una singularidad aparente que desaparece con un cambio de coordenadas apropiado. También se le conoce como el radio de Schwarzschild que se corresponde con el horizonte de sucesos [32].

Los agujeros negros de Schwarzschild, descritos por esta métrica, son estructuras estáticas, esféricamente simétricas, sin carga eléctrica ni momento angular.

1.2.2. Tipos de agujeros negros.

Teniendo en cuenta el origen de los agujeros negros, podríamos pensar que existen tantos tipos como diversidad de estrellas en el universo, pero no es así. El teorema sin pelo o de no pelo limita el número de cantidades físicas necesarias para describirlos, el colapso gravitacional de una estrella realística (no esférica, con una carga neta pequeña pero no nula) produce un agujero negro diferente al agujero negro de Schwarzschild [28]. A raíz de varios teoremas enunciados por Stephen Hawking (1971, 1972), Werner Israel (1967, 1968) y Brandon Carter (1970), se llegó a la conclusión de que, tanto si el colapso tiene carga pequeña y pequeñas asimetrías, como si la carga es grande y tiene grandes asimetrías, los campos gravitacionales y electromagnéticos externos del agujero negro final quedan completamente determinados por la masa M , la carga Q y el momento angular intrínseco o espín J de la estrella colapsante [28]. Antes del colapso la estrella posee más información, como por ejemplo los números bariónico y leptónico, el espín cuántico o las cargas de color y sabor. Sin embargo esta información, llamada el “pelo” del agujero negro, desaparece dentro del horizonte de sucesos al colapsar la estrella y configurarse en un agujero negro. Para este observador externo, el agujero negro no tendrá pelo [32]. El físico John Archibald Wheeler expresó esta idea con la frase “black holes have no hair” dando nombre a dicho teorema [28].

Entonces, de acuerdo con este teorema sólo existen cuatro tensores métricos posibles para los agujeros negros en equilibrio clasificados según sus propiedades de carga Q y momento angular J : Schwarzschild de simetría esférica y $J = 0$, $Q = 0$; Reissner-Nordström de simetría esférica y $J = 0$, $Q \neq 0$; Kerr de simetría axial y $J \neq 0$, $Q = 0$; y Kerr-Newman de simetría axial y $J \neq 0$, $Q \neq 0$ [32].

1.2.3. Agujeros negros de masa intermedia.

Además de la clasificación anterior, los agujeros negros pueden dividirse en las siguientes categorías en función de sus masas:

- Agujeros negros supermasivos (SMBHs), con masas de $M \geq 10^5 M_\odot$ [14], que residen en los centros de las galaxias y cuya formación sigue siendo un campo abierto de investigación. El 10 de abril de 2019 fue publicada, la primera imagen real de la historia de un agujero negro de este tipo ubicado en el centro de la galaxia M87 tomada por Event Horizon Telescope [24].
- Agujeros negros de masa intermedia (IMBHs), con masas de $100 M_\odot \leq M \leq 10^5 M_\odot$ [14], su formación ha sido propuesta a través de varios mecanismos.

- Agujeros negros de masa estelar (SBHs), con masas típicas de $10M_{\odot} \leq M \leq 100M_{\odot}$, formados por el colapso gravitatorio de estrellas masivas [14] [44]. Hasta la fecha, LIGO y Virgo han observado 20 SBHs en la fusión de binarias compactas [14].
- Microagujeros negros, objetos hipotéticos lo suficientemente pequeños como para evaporarse mediante la emisión de radiación de Hawking inmediatamente después de su formación [42]. Si la escala de la gravedad cuántica se encuentra en el orden del TeV , el Gran Colisionador de Hadrones (LHC) podría generar un agujero negro microscópico cada segundo, aunque se evaporaría inmediatamente. De hecho, conociendo la energía disponible para crear el agujero negro en el LHC ($14TeV$) podemos obtener como tiempo de vida: $t_l = 1.310^{-84}s$ [42].

Son conocidos muchos ejemplos tanto de agujeros negros de masa estelar (SBHs) como de agujeros negros supermasivos (SMBHs) [26]. Evidencias de SMBHs incluyen las observaciones de cuásares por el Sloan Digital Sky Survey, otro cuásar identificado por el United Kingdom Infrared Telescope (UKIRT) Infrared Deep Sky Survey (UKIDSS) Eighth Data Release en 2010 y un reciente estudio con información combinada de Sloan Digital Sky Survey, los dos Micron All sky Survey y del Wide-field Infrared Survey Explorer [24]. El mecanismo de formación de tales sistemas todavía no se comprende al completo, aunque se postula que la formación de binarias de SMBHs podría ser a través de fusiones de galaxias [24]. Se espera que la brecha entre los SBHs y los SMBHs sea completada por los IMBHs; que aunque no hayan sido observados de forma directa, esto no significa que no existan [26] [24] [44]. De hecho, son conocidas algunas evidencias indirectas de su existencia, como una fuente X ultraluminosa ² en la galaxia espiral ESO 243-49 o las oscilaciones cuasi periódicas de rayos X provenientes de M82 X-1, que es la fuente de rayos X más brillante de la galaxia M82 [24]. Además se ha sugerido que los IMBHs podrían ubicarse en los centros de los cúmulos globulares [14] [26], grupos densos de alrededor de un millón de estrellas. No obstante, se desconocen sus mecanismos de formación. A continuación, resumiremos los posibles escenarios que podrían explicarlos: por colapso directo de estrellas masivas de la población III, correspondientes a la primera generación de estrellas formadas tras el Big Bang y de muy baja metalicidad ³; por fusiones de estrellas masivas de secuencia principal en cúmulos estelares densos; por acumulación de gas residual en los agujeros negros de masa estelar; o bien por evolución químicamente homogénea [25].

Posibles detecciones de los IMBHs, donde se midan parámetros importantes tales como la masa o el espín, nos permitirían hacer afirmaciones no solo sobre su formación y su trayectoria evolutiva respecto a ser las posibles semillas buscadas de SMBHs, sino también sobre su vínculo con otros fenómenos observados como la posible relación entre los IMBH y los sistemas de fuente X ultraluminosa.

En este trabajo nos hemos centrado en el estudio de los IMBHs más ligeros, los LIMBHs con masas $\gtrsim 100M_{\odot}$, que en general tienen muchas más posibilidades de ser detectados que los IMBHs. Los LIMBHs son especialmente difíciles de detectar electromagnéticamente, pero no lo son para LIGO que es sensible para masas dentro de ese rango [26]. Por ello, hemos considerado binarias de agujeros negros con masas totales de $100M_{\odot}$ y $200M_{\odot}$, cada una con 3 relaciones asimétricas de masa distintas: $q \equiv m_2/m_1 = 1, 3$ y 5 . Estos LIMBH pueden formarse o bien por el colapso directo de estrellas masivas con metalicidad muy baja o bien por la fusión de dos agujeros negros de masa estelar masiva. Se espera que esta primera opción sea silenciosa y difícil de detectar [44]. Por lo tanto, el segundo escenario puede que sea el único observable en un futuro cercano dentro del tercer periodo de observación O3 [44].

1.3. Ondas gravitacionales.

Un formalismo matemático que relacione la curvatura del espacio-tiempo con fuerzas gravitacionales nos sería de gran utilidad para introducir el concepto de onda gravitacional. Por este motivo, estudiaremos un escenario simple donde un campo gravitacional muy débil (como por ejemplo, nuestro planeta) pueda ser descrito en términos de un espacio-tiempo con muy poca curvatura. Consideremos las siguientes suposiciones sobre el espacio-tiempo: Las partículas se mueven lentamente en el espacio-tiempo; como el

²Una fuente X ultraluminosa (ULEX) es uno de los objetos más raros e intrigantes que existen cuya fuente de rayos X tiene una luminosidad por encima de los 10^{39}ergs/s (10^{32}W), excediendo el límite de Eddington (luminosidad máxima que un cuerpo puede alcanzar cuando hay un equilibrio hidrostático entre la fuerza gravitacional actuando hacia adentro y la fuerza de radiación actuando hacia afuera) de estrellas de neutrones y de agujeros negros. Estimados con una tasa de tan sólo 1 o 2 por galaxia, el origen de emisión es desconocida. Posibles explicaciones serían: la alta luminosidad de una distribución continua de fuentes típicas de rayos X como supernovas o binarias de rayos X, nuevas clases de objetos incluidos los IMBHs o hipernovas [10].

³La metalicidad, se utiliza para describir la abundancia de elementos más pesados que el hidrógeno o el helio presentes en un objeto.

espacio-tiempo es prácticamente plano, aproximamos la métrica como $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ con $|h_{\mu\nu}| \ll 1$, donde $h_{\mu\nu}$ es una pequeña perturbación de la métrica de Minkowski; suponemos que el campo débil es estático, con lo cual las componentes de la métrica no dependen del tiempo $g_{\mu\nu,0} = 0$ [32].

Con esta estructura es posible obtener una linealización de las ecuaciones de campo de Einstein [32]

$$-\bar{h}_{\mu\nu,\alpha}{}^\alpha = 16\pi T_{\mu\nu}, \quad (1.3)$$

donde $\bar{h}_{\mu\nu,\alpha}{}^\alpha$ es el operador barra definido para simplificar la notación como $\bar{h}_{\mu\nu} \equiv h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h$.

Reescribiendo esta ecuación, $\square\bar{h}_{\mu\nu} = \partial^\alpha\partial_\alpha\bar{h}_{\mu\nu} = -16\pi T_{\mu\nu}$. En el vacío se reduce a

$$\square\bar{h}_{\mu\nu} = 0. \quad (1.4)$$

Al ser el espacio plano, el d'Alembertiano tiene la forma $\square = \partial_t^2 - \nabla^2$. Es decir, la dinámica de las fluctuaciones del espacio-tiempo es ondulatoria; la ecuación de campo tiene la forma de ecuación de onda para $\bar{h}_{\mu\nu}$. Para esta ecuación diferencial, puede proponerse una solución de onda plana con una forma exponencial compleja como

$$\bar{h}_{\mu\nu} = A_{\mu\nu}e^{ik_\sigma x^\sigma}, \quad (1.5)$$

donde k^σ son las componentes de un vector de onda constante $(k^\sigma) = (w, k^1, k^2, k^3)^T$ y $A_{\mu\nu}$ es la amplitud de la onda plana que corresponde a las componentes de un tensor constante simétrico de rango (0,2) [32]

$$(A_{\mu\nu}) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & A_{11} & A_{12} & 0 \\ 0 & A_{12} & -A_{11} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (1.6)$$

Para que el ansatz 1.5 sea una solución válida, sustituyéndolo en 1.4, puede demostrarse que se requiere $k_\sigma k^\sigma = 0$. Esto significa que la onda plana 1.5 es solución de la ecuación linealizada siempre y cuando el vector de onda sea nulo o luminoide. Es decir, las perturbaciones de la métrica plana que están codificadas en h se propagan a la velocidad de la luz [32].

Fijémonos además que la frecuencia w junto a las componentes A_{11} y A_{12} caracterizan por completo la onda gravitacional. Para obtener sus efectos, consideremos el movimiento relativo de un anillo de partículas cercanas estacionario en el plano $x^1 - x^2$. Resulta que A_{11} y A_{12} corresponden a las dos polarizaciones de la onda: $h_+ = A_{11}$ y $h_\times = A_{12}$. A medida que la onda pasa con polarización h_+ , las partículas oscilarán vertical y horizontalmente en forma de “+”. Para la polarización h_\times , el anillo oscilará en forma de “×”. A estas perturbaciones ondulatorias que se desplazan mediante vectores de onda luminoides las llamamos ondas gravitacionales [32].

1.3.1. Categorías de ondas gravitacionales.

A continuación, describiremos las cuatro categorías de ondas gravitacionales que nos ayudan a comprender los diferentes tipos de ondas gravitacionales que pueden producirse: [6]

- Ondas gravitacionales continuas, producidas por cuerpos compactos no esféricos de rotación rápida que las emiten de forma continua, como las estrellas de neutrones;
- Ondas gravitacionales producidas al orbitar pares de objetos muy densos, compactos, tales como un sistema binario de agujeros negros, un sistema binario de estrellas de neutrones o bien un sistema binario de un agujero negro y una estrella de neutrones;
- Ondas gravitacionales asociadas a estallidos cortos, como supernovas o brotes de rayos gamma, que pueden producirse cuando un sistema binario de dos estrellas de neutrones o de una estrella de neutrones y un agujero negro se fusionan;
- Ondas gravitacionales estocásticas, que producen señales estocásticas procedentes de todas las direcciones llegando al detector desde diferentes fuentes, entre ellas el Big Bang.

1.3.1.1. Ondas gravitacionales producidas por CBCs.

Este trabajo se centra en el estudio de las ondas gravitacionales producidas por la coalescencia de binarias compactas (CBC), concretamente en aquellas producidas por la fusión de IMBHs. Este tipo de ondas se producen por el mecanismo llamado “inspiral”. Este mecanismo ocurre a lo largo de millones

de años a medida que los objetos de las binarias orbitan entre sí. A lo largo de eones, ambos cuerpos se acercan cada vez más, generando ondas gravitacionales cada vez más fuertes que harán que se vaya perdiendo más energía orbital [6]. El final inevitable es la fusión de ambos objetos, dando como resultado un agujero negro de Kerr perturbado que evolucionará a un agujero negro de Kerr durante la fase de relajamiento [4].

La primera observación directa de ondas gravitacionales se produjo el 14 de septiembre de 2015, dando nombre al evento GW150914, por las colaboraciones LIGO y Virgo. Esta señal provenía de dos agujeros negros de masa estelar de $29M_{\odot}$ y $36M_{\odot}$ [30]. La energía liberada por la colisión en forma de ondas gravitacionales fue equivalente a $3M_{\odot}$. Su potencia máxima, liberada en tan sólo una fracción de segundo, fue más poderosa que la luz combinada de todas las estrellas en todas las galaxias de todo el universo visible. Aún así lo que detectamos en la Tierra fueron dos espejos moviéndose de un lado a otro en una milésima del ancho de un protón [30]. “Por sus contribuciones decisivas al detector LIGO y por la observación de ondas gravitatorias” les fue concedido el premio Nobel de Física 2017 a Kip Thorne, Rainer Weiss y Barry C. Barish [1].

Las ondas gravitacionales se producen durante todo el proceso. A medida que la órbita de la binaria se va reduciendo, la amplitud y frecuencia de la onda aumenta, generando una señal que alcanza su máximo cuando ambos cuerpos se fusionan y que recibe el nombre de señal chirrido [21]. La señal detectada puede clasificarse en las siguientes tres fases: fase inspiral, referida a los segundos finales antes de que los cuerpos colapsen que es cuando las ondas gravitacionales son más fuertes; fase de fusión, referida a la colisión entre ambos objetos; fase de relajamiento, durante la cual el cuerpo resultante de la fusión se recupera del evento de su formación.

Durante la primera parte de esta coalescencia, cuando tenemos una frecuencia orbital pequeña, las ondas gravitacionales pueden ser modeladas analíticamente con precisión mediante expansiones analíticas post-newtonianas (PN) en las ecuaciones de campo de Einstein [22]. Sin embargo, se requieren de las soluciones no lineales completas de estas ecuaciones para describir las etapas finales de la binaria, cuando los cuerpos orbitan rápidamente y se fusionan. Debido a su alta no linealidad, las ecuaciones deben resolverse numéricamente en superordenadores. Esto es abordado en el contexto de la relatividad numérica (NR) [4] [22].

Son necesarios unos modelos de ondas gravitacionales que sean precisos para la búsqueda de estas señales. De lo contrario la detección no sería tan efectiva y podría identificarse erróneamente los parámetros de las fuentes.

Sección 2.1. Modelos de ondas gravitacionales para CBCs.

Sección 2.2. Detección de ondas gravitacionales.

Sección 2.3. Análisis bayesiano.

2.1. Modelos de ondas gravitacionales para CBCs.

Al conocer la señal de una binaria compacta, las búsquedas realizadas se basan en la técnica de filtrado adaptado consistente en un filtrado de los datos entrantes con plantillas de la radiación gravitatoria que esperamos observar. Los análisis basados en el filtrado adaptado son utilizados en la búsqueda y estimación de parámetros de estas señales, que requieren de modelos precisos y computacionalmente económicos, utilizándolos como plantillas [34]. Estas plantillas o modelos, conocidos como familias de formas de onda, difieren en su complejidad computacional, en la física que simulan y en su régimen de aplicabilidad [19]. En las últimas décadas han sido desarrolladas diferentes técnicas para el cálculo de dichos modelos, que pueden ser clasificadas según el tipo de procedimiento utilizado, como: expansiones analíticas computadas en el formalismo post-newtoniano (PN); teoría perturbativa; formalismo efectivo a un solo cuerpo (EOB); soluciones de relatividad numérica (NR); formas de onda híbrida, una combinación de los resultados del formalismo post-newtoniano y de la relatividad numérica (PN/NR); formalismo efectivo a un solo cuerpo combinado con la relatividad numérica (EOBNR); y modelos Phenom, que son familias analíticas de formas de onda Inspiral-Merger-Ringdown (IMR) obtenidas al ajustar modelos fenomenológicos (phenom) a un conjunto discreto dado de formas de onda híbridas [4]. También se utilizan modelos alternativos de la gravedad como la teoría tensor-escalar, que explica la gravedad mediante el uso conjunto de un campo tensorial (de orden $n > 1$) y un campo escalar de manera que no pueda ser explicada únicamente a través de un campo escalar ni únicamente a través de una métrica.

Cada familia de formas de onda es considerada como una función que toma como entrada un vector de parámetros $\vec{\theta} = \{\theta_1, \theta_2, \dots, \theta_N\}$ y produce una salida $\vec{h}_{+, \times}(\theta)$, ya sea como señal en el dominio frecuencial $h(\vec{\theta}; f)$ o en el dominio temporal $h(\vec{\theta}; t)$ [19]. Son necesarios como mínimo 9 parámetros para describir binarias cuyos objetos se suponen de masa puntual y sin espín. Al incluir los espines de las componentes en el modelo y considerar precesión, como es nuestro caso, el vector $\vec{\theta}$ deberá incluir 6 parámetros más adicionales para un total de 15. Estos modelos son descritos por 8 parámetros intrínsecos que describen las propiedades fundamentales de una binaria: la masa primaria m_1 , la masa secundaria m_2 ¹, los espines χ_1 y χ_2 , ángulos relacionados con los espines t_1 , t_2 y ϕ_{12} , y un ángulo azimutal del momento angular orbital ϕ_{JL} ; y por 7 parámetros extrínsecos que determinan la localización y orientación de la binaria respecto del observador: la ascensión recta α , declinación δ , distancia luminosa d_L , el ángulo de inclinación ι , el ángulo de polarización ψ , el tiempo de coalescencia t_c y la fase orbital ϕ_c . Describamos brevemente estos parámetros utilizados en el trabajo: [19]

- La masa chirrido, definida por las masas de cada componente m_1 y m_2 como: $\mathcal{M} = (m_1 m_2)^{3/5} (m_1 + m_2)^{-1/5}$. Es un parámetro importante porque, junto con la frecuencia de la señal y la distancia a la binaria, determina la amplitud adimensional de las polarizaciones de la onda gravitacional y el tiempo de coalescencia de la binaria a determinada frecuencia [21].
- El cociente de masa, definida con el convenio $m_1 \geq m_2$, como: $q = m_2/m_1$, para que estas variables sean más fáciles de muestrear.
- La distancia luminosa a la fuente d_L .
- La ascensión recta α y la declinación δ , para ubicar la fuente en la esfera celeste.
- El ángulo de inclinación ι , entre el momento angular orbital del sistema y la línea de visión.
- El ángulo de polarización ψ , que describe la orientación de la proyección del vector momento orbital de la binaria sobre el plano en el cielo.
- El tiempo de referencia que será el tiempo de coalescencia de la binaria t_c .

¹En una binaria interactuante, y en nuestro caso compacta, se le asigna a la componente más pesada el nombre de agujero negro primario que por convenio será la masa m_1 ; al más ligero de los dos se le llama secundario y le es asignado la masa m_2 . Por lo tanto: $m_1 \geq m_2$.

- La fase orbital de la binaria ϕ_c en el tiempo de referencia.
- Los espines adimensionales χ_i con $i = 1, 2$ dentro del rango $[0, 1]$, definidos como: $\chi_i \equiv |\vec{s}_i|/m_i^2$, donde \vec{s}_i es el vector de espín para el objeto i .
- Los ángulos de inclinación entre los espines de los objetos compactos y el momento angular orbital t_1 y t_2 .
- El ángulo azimutal que separa los vectores de espín ϕ_{12} .
- La ángulo azimutal del momento angular orbital ϕ_{JL} .

En este trabajo se ha utilizado el modelo IMRPhenomPv2 inyectado en el ruido de los detectores simulados localizados en Hanford Site (H1) y Livingston (L1) de LIGO, con la finalidad de generar varias hipotéticas ondas gravitacionales razonables procedentes de binarias compactas de IMBHs para recuperar sus respectivos parámetros a través de una estimación.

2.1.1. El modelo IMRPhenomPv2.

Uno de los problemas más relevantes en la astronomía de ondas gravitacionales es la construcción de un modelo de señal a partir de configuraciones genéricas de binarias compactas de agujeros negros con precesión durante toda la coalescencia: inspiral, merger y ringdown (IMR). Investigadores de la UIB, liderados por Sascha Husa, y de la universidad de Cardiff presentaron el primer modelo de este tipo en el dominio de frecuencia, el modelo “IMRPhenomP” [27].

El primer paso en la producción de estos modelos fenomenológicos es la generación de formas de onda híbridas PN/NR, que son el resultado de añadir al final del código de las formas de onda NR, las de PN para cubrir las primeras etapas de la coalescencia que las simulaciones de NR no alcanzan [4]. A continuación, varios parámetros ajustables de los modelos analíticos son calibrados para un conjunto finito de formas de onda híbridas. Aquellos parámetros que describen la amplitud y la fase durante las etapas inspiral, merger y ringdown son calibrados aparte para luego construir formas de onda IMR que se interpolarán en el espacio de parámetros dando como resultado una familia de formas de onda IMRPhenom [4]. Existen varios modelos Phenom que utilizan diferentes estrategias para la calibración, entre ellos IMRPhenomP que describe binarias compactas con precesión.

2.2. Detección de ondas gravitacionales.

De acuerdo con la teoría, las ondas gravitacionales pueden producirse por cualquier movimiento que sea acelerado (incluso un aplauso). Sin embargo, debido a que la gravedad es la fuerza más débil de la naturaleza, sólo aquellas ondas generadas por eventos cósmicos violentos, fuentes gravitatorias inmensamente masivas o energéticas como estrellas de neutrones o agujeros negros, son detectables desde la Tierra con sensores tan sensibles como para medir deformaciones del espacio de una milésima del ancho de un protón [30].

Los interferómetros de LIGO constan de dos brazos en forma de “L” de 4km cada uno capaces de detectar alteraciones de hasta 10^{-19} m de longitud. Las ondas gravitacionales contraen el espacio en una dirección mientras que lo expanden en la dirección perpendicular. En consecuencia, al atravesar un interferómetro uno de los brazos se alarga mientras que el otro se reduce provocando que los haces de luz emerjan desincronizados, produciendo una señal oscilatoria de frecuencia compatible con la de la onda que originó la deformación. Es decir, la onda gravitacional a su paso modifica la distancia L entre cualquier par de puntos (ΔL) deformando los brazos. Suele trabajarse en función de la deformación unitaria definida como $h = \Delta L/L$, de acuerdo con 1.4 [33].

2.2.1. Señal de una onda gravitacional observada por el detector.

La señal obtenida por el detector dependerá de la ubicación $(d_L, \bar{\theta}, \bar{\varphi})$ de la fuente en el cielo y de la polarización ψ de la onda gravitacional. Esta respuesta está codificada en unas funciones que describen la sensibilidad relativa del interferómetro en diferentes orientaciones celestes, llamadas patrones de antena,

definidas para cada polarización como [4] [19]

$$\begin{aligned} F_+ &= \frac{1 + \cos^2\bar{\theta}}{2} \cos 2\bar{\varphi} \cos 2\psi - \cos\bar{\theta} \sin 2\bar{\varphi} \sin 2\psi, \\ F_\times &= \frac{1 + \cos^2\bar{\theta}}{2} \cos 2\bar{\varphi} \sin 2\psi + \cos\bar{\theta} \sin 2\bar{\varphi} \cos 2\psi. \end{aligned} \quad (2.1)$$

De manera que la señal observada de las ondas gravitacionales es

$$h_d(\Xi; \theta, \varphi, \bar{\theta}, \bar{\varphi}, \psi; d_L, t) = \frac{1}{d_L} [F_+(\bar{\theta}, \bar{\varphi}, \psi) h_+(\Xi; \theta, \phi; t) + F_\times(\bar{\theta}, \bar{\varphi}, \psi) h_\times(\Xi; \theta, \phi; t)]. \quad (2.2)$$

donde $h_{+, \times}$ son las dos polarizaciones de la onda gravitacional mencionadas en la sección 1.3.

2.2.2. Los datos obtenidos del detector.

Como las señales de ondas gravitacionales son muy débiles y los datos obtenidos por el detector $d(t)$ son realizaciones de un proceso estocástico, se requieren de métodos estadísticos optimizados para extraer las señales e identificarlas dentro de un ruido de fondo [7] [21].

Estos datos $d(t)$ son modelados como la suma de la señal de la onda gravitacional $h(t)$ y el ruido de fondo $n(t)$, [19]

$$d(t) = n(t) + h(t). \quad (2.3)$$

La señal $h(t)$ se encuentra enterrada por este ruido de fondo $n(t)$ de varios órdenes de magnitud mayor, cuyos orígenes resumiremos a continuación: [4]

- Ruido sísmico, consistente en vibraciones del suelo en el entorno circundante. Para intentar inhibir sus efectos, los detectores utilizan un sistema cuádruple de suspensión en el que los espejos penden de cuatro péndulos. Esto amortigua los movimientos de altas frecuencias, por encima de unos 10 Hz.
- Ruido térmico, referido a vibraciones térmicas que afectan a los espejos, expandiéndolos, y a las suspensiones correspondientes debido al calentamiento causado por los láseres.
- Ruido de disparo, producido por la naturaleza estadística de la detección de los cuantos de luz. La incerteza de este proceso es reducida incrementando la potencia del láser emitido.
- Ruido de presión de radiación, ya que la naturaleza cuántica de los láseres ocasiona que cuanto mayor sea la potencia, mayor es la incertidumbre existente en el momento transferido a los espejos.
- Ruido del gradiente de la gravedad, debido a cambios en el campo gravitatorio newtoniano. De momento, esto no ha afectado a las mediciones realizadas hasta la fecha, pero a medida que los detectores puedan operar a frecuencias más bajas, este efecto se hará cada vez más notable.

Por simplicidad, se asume que el ruido $n(t)$ del detector es el resultado de una superposición de un ruido gaussiano estacionario de media cero, $S_n(f)$, y fallos técnicos de origen terrestre. Esta $S_n(f)$, conocida como densidad espectral de amplitud (ASD), es definida de manera que la integral de $P_n(f)$ venga descrita solamente sobre las frecuencias positivas, [5] [41]

$$\frac{1}{2} S_n(f) = P_n(f); \quad (2.4)$$

donde $P_n(f)$ es la densidad espectral de potencia (PSD) cuya finalidad es la de darnos una imagen de la sensibilidad y del ruido promedio presente en el detector,

$$P_n(f) = \lim_{T \rightarrow \infty} \left| \int_{-T/2}^{T/2} n(t) e^{-2\pi i f t} dt \right|^2. \quad (2.5)$$

La idea básica, detrás de la detección de señales, es que la presencia de la señal en los datos recopilados por un detector cambia las características estadísticas de los datos, en particular su distribución de probabilidad [21].

Al despreciar los fallos técnicos de origen terrestre presentes en el ruido, el mejor método para decidir si una señal $h(t)$ se encuentra enterrada en el ruido $n(t)$ es el filtrado adaptado. Su salida, es la relación

señal-ruido (SNR) expresada como [4]

$$\rho[d, h] = \frac{\langle d|h \rangle}{\sqrt{\langle h, h \rangle}}, \quad (2.6)$$

donde se ha definido el producto interno ²

$$\langle a, b \rangle \equiv 4\Delta f \sum_{j,l} \Re\left(\frac{a_{j,l}^* b_{j,l}}{P_{j,l}}\right). \quad (2.7)$$

En un ruido gaussiano estacionario de media cero como el nuestro, ρ se encuentra gaussianamente distribuido con media cero y desviación estándar uno, [4]

$$f(\rho) = \frac{1}{\sqrt{2\pi}} e^{-\rho^2/2}. \quad (2.8)$$

Esta f es la función de densidad de probabilidad para un ruido gaussiano puro que genera un valor de SNR igual o mayor a ρ al filtrarlo por una plantilla h . Si la SNR cruza algún umbral ρ_0 , considerado en torno a 8, se puede suponer que en los datos del detector hay algo más que ruido gaussiano [4]. En este caso se puede asumir que la señal $h(t)$ se encuentra enterrada en los datos. Es decir, la presencia de dicha señal en los datos modifica la distribución estadística gaussiana que tiene la SNR en un ruido gaussiano puro. La SNR adquiere entonces un significado estadístico claro cuando se supone que el ruido de fondo es gaussiano con media cero; se encuentra relacionado directamente con la probabilidad de encontrar una onda gravitacional en el ruido.

2.3. Análisis bayesiano.

Durante los aproximadamente 14 años de funcionamiento de la red de interferómetros de ondas gravitacionales, los esfuerzos en el análisis de CBCs se han centrado en el problema de la detección; concretamente en cómo identificar sin ambigüedad una fusión binaria dentro del ruido. Una vez aislado un candidato de detección, el siguiente paso es extraer una información completa sobre los parámetros del origen y la física subyacente [19]. Podemos dividir la tarea de realizar la inferencia sobre la fuente de onda gravitacional principalmente en dos retos: usar los datos observados para restringir o estimar los parámetros desconocidos de la fuente bajo un modelo fijo de la forma de onda (estimación de parámetros), y decidir cuál de entre varios modelos son más probables a la luz de los datos observados (selección de modelo) [19].

Al tener una teoría que describe cómo se emiten las ondas gravitacionales, existe un vínculo directo entre los datos y el modelo. Por esta razón, es esencial la aplicación de la inferencia bayesiana en la astronomía de ondas gravitacionales [40]. Los dos problemas mencionados anteriormente son abordados dentro de este formalismo, el cual describe el estado de conocimiento, bien sobre una hipótesis incierta H como una probabilidad, denotada como $P(H) \in [0, 1]$, o bien sobre un parámetro θ desconocido como una densidad de probabilidad, denotada como $p(\theta|H)$, donde $\int p(\theta|H)d\theta = 1$ [19].

En este trabajo nos centramos en la estimación de parámetros de ondas gravitacionales procedentes de binarias compactas de IMBHs bajo el modelo IMRPhenomPv2.

2.3.1. Conceptos fundamentales.

2.3.1.1. El teorema de Bayes.

El objetivo principal de la inferencia bayesiana es la construcción de una distribución de probabilidad a posteriori $p(\theta|d, H)$, donde d son los datos asociados a la medición y H es un modelo descrito por unos parámetros denotados como $\vec{\theta} = \{\theta_1, \theta_2, \dots, \theta_N\}$. Este conjunto de parámetros son los 15 parámetros que describen nuestra coalescencia definidos en la primera sección de este capítulo 2.1 y d son los datos obtenidos por el detector descritos en 2.2.2 [40]. La distribución de probabilidad a posteriori $p(\theta|d, H)$ es la función de densidad de probabilidad para la variable continua θ dados los datos d . La probabilidad de que el verdadero valor de θ se encuentre entre $(\theta', \theta' + d\theta)$ viene dado por $p(\theta'|d, H)d\theta'$ [40].

² Siguiendo la convención en la astronomía de ondas gravitacionales, el producto interno es real por construcción. Sin embargo, en ocasiones también es útil definir un producto interno complejo [40].

La distribución de probabilidad a posteriori es la que usaremos para construir intervalos de confianza ³ que nos estimen, por ejemplo, las masas de las componentes de uno de los eventos simulados. La estimación de parámetros se realiza utilizando el teorema de Bayes, donde se actualiza una distribución de probabilidad a priori $p(\theta|H)$ al recibir los nuevos datos $d(t)$ del experimento para dar una distribución de probabilidad a posteriori $p(\theta|d, H)$, [40]

$$p(\theta|d, H) = \frac{p(d|\theta, H)p(\theta|H)}{p(d|H)}; \quad (2.9)$$

escrito de otra manera como [40] ⁴

$$p(\theta|d) = \frac{\mathcal{L}(d|\theta)\pi(\theta)}{\mathcal{Z}}, \quad (2.10)$$

donde la distribución de probabilidad $\mathcal{L}(d|\theta)$ es conocida como función de verosimilitud o verosimilitud a través de la cual los datos $d(t)$ modifican el conocimiento inicial de θ [2] [19], $\pi(\theta)$ es la distribución de probabilidad a priori para θ y \mathcal{Z} es el factor de normalización llamado evidencia ⁵

$$\mathcal{Z} \equiv \int d\theta \mathcal{L}(d|\theta)\pi(\theta). \quad (2.11)$$

2.3.1.2. La función de verosimilitud.

La verosimilitud es una función que describe la medida y que por tanto podemos elegir. Al escribirla introducimos implícitamente un modelo del ruido. Como ya mencionamos en la sección 2.2.2, en la astronomía de ondas gravitacionales se asume normalmente una función de verosimilitud cuyo ruido es gaussiano ⁶ definido como, [40]

$$\mathcal{L}(d|\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2} \frac{(d - \mu(\theta))^2}{\sigma^2}\right), \quad (2.12)$$

donde π es el número pi, no la distribución de probabilidad a priori, $\mu(\theta)$ es una plantilla para la forma de onda dado θ que definiremos más adelante y σ es el ruido del detector.

Antes de la recopilación de datos, la función de verosimilitud describe la posibilidad de obtener datos d . Es una función de densidad de probabilidad con unidades de los datos inversos. La integral sobre toda d posible es la unidad. Una vez obtenidos los datos reales, d es obviamente fijada [40].

2.3.1.3. La distribución de probabilidad a priori.

Al igual que la función de verosimilitud, también podemos elegir la distribución de probabilidad a priori $\pi(\theta)$. Esta probabilidad incorpora nuestra creencia sobre θ antes de la realización de la medida. En algunos casos, se tiene una intuición bastante obvia de su elección. Por ejemplo, al considerar la ubicación en el cielo de nuestro evento, es razonable elegir una $\pi(\theta)$ isotrópica que pondere cada zona en el cielo cómo igualmente probable. Sin embargo, en otras situaciones la elección de esta distribución no es tan obvia. Como por ejemplo, antes de la primera detección de ondas gravitacionales, ¿cuál hubiera sido la opción más adecuada para la distribución de probabilidad a priori de la masa de la primaria compacta $\pi(m_1)$? En estos casos, cuando desconocemos θ expresamos esta ignorancia al elegir o bien una distribución uniforme o una distribución logarítmica uniforme [40] ⁷. La elección de una $\pi(\theta)$ uniforme se ha considerado en nuestros eventos para todos los parámetros, exceptuando los ángulos polares t_1, t_2 de inclinación entre los espines de los objetos compactos y el momento angular orbital, y el ángulo de inclinación ι , los 3 con una distribución sinusoidal; y la declinación, con una distribución coseno.

³En el apéndice A del paper [40] se explica claramente el método utilizado para construir intervalos de confianza llamado: método del intervalo más alto de la densidad a posteriori (HPDI).

⁴Al referirnos a parámetros θ de un modelo H , reconocemos implícitamente que tenemos un modelo. Por eso simplificamos esta notación innecesaria, puesto que no hace falta especificar siempre que debe asumirse un modelo [40].

⁵Aunque la evidencia parece ser “solo” un factor de normalización, veremos que juega un papel importante en la comparación y selección de modelos y que puede entenderse como una verosimilitud marginal [40].

⁶La suposición de ruido gaussiano es un buen punto de partida para describir el ruido de los detectores. El efecto combinado de muchos ruidos aleatorios tiende a producir aproximadamente una gaussiana. La descripción del ruido, puede generalizarse también para incluir fallos no gaussianos.

⁷Se utiliza una distribución logarítmica uniforme cuando no se conoce el orden de la magnitud de una cantidad.

2.3.1.4. La distribución de probabilidad a posteriori.

La distribución de probabilidad a posteriori para una coalescencia binaria de agujeros negros es una función de 15 dimensiones ⁸. Con el fin de poder visualizar la distribución de probabilidad a posteriori de cada parámetro θ , se marginan (integramos) los parámetros que no nos interesan, llamados parámetros molestos, para obtener una distribución de probabilidad a posteriori marginalizada [40]

$$p(\theta_i|d) = \int \left(\prod_{k \neq i} d\theta_k \right) p(\theta|d) = \frac{\mathcal{L}(d|\theta_i)\pi(\theta_i)}{\mathcal{Z}}, \quad (2.13)$$

donde $\mathcal{L}(d|\theta_i)$ es conocida como la función de verosimilitud marginalizada [40]

$$\mathcal{L}(d|\theta_i) = \int \left(\prod_{k \neq i} d\theta_k \right) \pi(\theta_k) \mathcal{L}(d|\theta). \quad (2.14)$$

Al marginalizar sobre un parámetro θ_a para obtener una distribución de probabilidad a posteriori en θ_b , estamos calculando nuestra mejor estimación para θ_b dada una incertidumbre en θ_a . Es decir, si θ_a y θ_b son covariantes, entonces al marginalizar sobre θ_a se inyecta una incerteza en la distribución de probabilidad a posteriori de θ_b . Cuando sucede esto, la distribución de probabilidad a posteriori $p(\theta_b|d)$ se hace más ancha que la distribución de probabilidad a posteriori condicional $p(\theta_b|d\theta_a)$ [40].

2.3.1.5. La evidencia.

La evidencia bayesiana, definida en 2.11, es un número que por lo general no significa nada por sí mismo; sin embargo se vuelve de gran utilidad al comparar una evidencia con otra. Concretamente, la evidencia es una función de verosimilitud marginalizada. Además de servir como factor de normalización para la distribución de probabilidad a posteriori $p(\theta|d)$, se utiliza también para la selección de modelos, respondiendo a la pregunta: ¿qué modelo es estadísticamente preferido por los datos y por cuánto? [40]. Podemos comparar, por ejemplo, un modelo de señal en el que suponemos que hay una señal de una binaria compacta presente en los datos con una distribución de probabilidad a priori $\pi(\theta)$, frente a un modelo de ruido en el que suponemos que no tenemos una señal de una binaria compacta. Mientras que el modelo de señal se describe a través de los 15 parámetros de la binaria compacta $\vec{\theta}$, el modelo de ruido es descrito sin ningún parámetro. Por lo tanto, podemos definir una evidencia de la señal \mathcal{Z}_S y una evidencia del ruido \mathcal{Z}_N , llamada también verosimilitud nula, como [40]

$$\mathcal{Z}_S \equiv \int d\theta \mathcal{L}(d|\theta) \pi(\theta) \quad (2.15)$$

$$\mathcal{Z}_N \equiv \mathcal{L}(d|0) \equiv \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-1}{2} \frac{h^2}{\sigma^2}\right). \quad (2.16)$$

La proporción de la evidencia entre dos modelos diferentes se conoce como el factor de Bayes. En este caso, el factor de Bayes señal/ruido es [40]

$$BF_N^S \equiv \frac{\mathcal{Z}_S}{\mathcal{Z}_N} \quad (2.17)$$

A menudo es conveniente trabajar con el logaritmo del factor de Bayes [40]

$$\log BF_N^S \equiv \log(\mathcal{Z}_S) - \log(\mathcal{Z}_N). \quad (2.18)$$

Cuando el valor absoluto de $\log BF$ es grande, decimos que se prefiere un modelo sobre el otro. El signo de $\log BF$ nos dice qué modelo es el preferido. Se suele utilizar un umbral de $|\log BF| = 8$ como nivel de “evidencia fuerte” a favor de una hipótesis sobre la otra [40].

La evidencia bayesiana codifica dos datos relevantes: la función de verosimilitud nos dice cómo de bueno se ajusta nuestro modelo a los datos; el acto de marginalizar nos da información sobre el tamaño del volumen del espacio de parámetros que utilizamos para realizar un ajuste [40].

Queremos obtener el mejor ajuste posible, es decir queremos una alta verosimilitud, pero con un volumen de la distribución de probabilidad a priori mínima. Un modelo con un ajuste decente y un volumen de la distribución de probabilidad a priori pequeña a menudo produce una mayor evidencia, que un modelo

⁸Recordar los 15 parámetros que describen nuestro evento, definidos en la sección 2.1.

con un ajuste excelente y un gran volumen de la distribución de probabilidad a priori [40]. En estos casos, el factor de Bayes penaliza al modelo más complicado por ser demasiado complicado. Esta penalización se denomina factor de Occam ⁹.

2.3.2. Muestreadores.

El desarrollo de formas de onda gravitacionales descritas en la sección 2.1, también llamadas aproximantes, han permitido hacer predicciones sobre cómo deben ser los datos d dados ciertos parámetros θ . Este problema es hacia adelante. El cálculo de la distribución de probabilidad a posteriori, la probabilidad de los parámetros θ dados los datos, como en la ecuación 2.10, es un problema inverso. En general, los problemas inversos presentan un fuerte desafío computacional en comparación con los problemas hacia adelante.

La solución es utilizar un muestreador estocástico. Los algoritmos de muestreo más comúnmente utilizados se pueden dividir en dos grandes categorías: método Monte Carlo basado en cadenas de Márkov (MCMC) [29] [16], y el algoritmo de muestreo anidado [36] [37].

Estos algoritmos generan una lista de muestras $\{\theta\}$ de la distribución de probabilidad a posteriori, de modo que el número de muestras se encuentran en el intervalo $(\theta, \theta + \Delta \theta) \propto p(\theta)$ [20]. En la figura 2.1, observamos una representación esquemática de ambos algoritmos y sus diferencias:

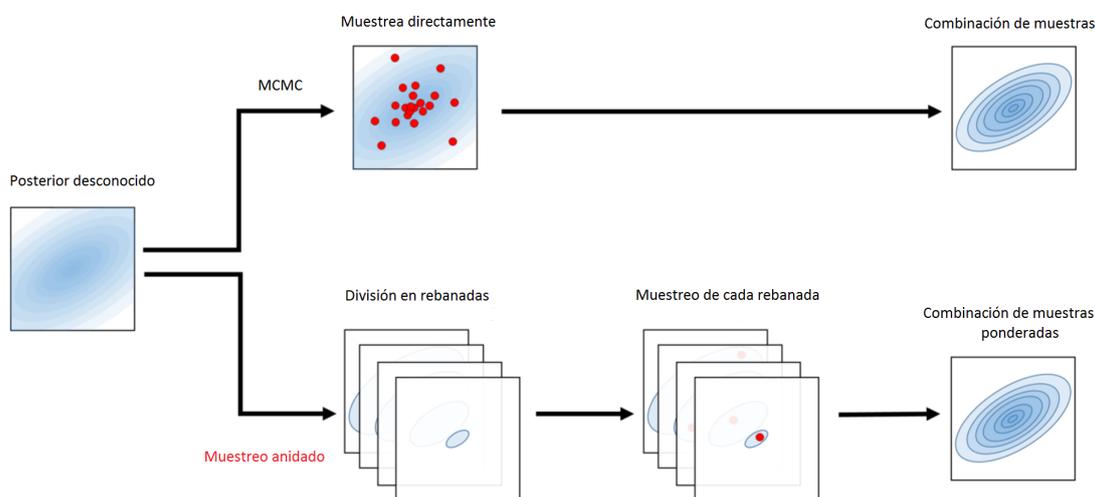


FIGURA 2.1: Los métodos MCMC generan muestras directamente desde la distribución de probabilidad a posteriori, mientras que el algoritmo de muestreo anidado divide esta distribución en rebanadas generando muestras de cada una, para luego volver a combinarlas reconstruyendo la distribución original escogiendo unos pesos apropiados w . Fuente: [38].

Podemos visualizar las muestras como una hoja de cálculo, donde cada columna es un parámetro diferente (tendremos 15 columnas) y cada fila representa una muestra de la distribución de probabilidad a posteriori diferente. Estas muestras presentan dos propiedades útiles: [40]

- Pueden utilizarse para calcular los valores esperados de ciertas cantidades $f(x)$ de interés [17]
- Al obtener muestras de un espacio N-dimensional, podemos generar la probabilidad marginal de cualquier subconjunto de parámetros seleccionando las columnas correspondientes en nuestra hoja de cálculo. Esta propiedad se utiliza para visualizar la salida de estos muestreadores mediante la construcción de “diagramas de esquina”, que muestran las distribuciones de probabilidad marginal a posteriori en una y dos dimensiones para cada parámetro.

2.3.2.1. MCMC.

MCMC es una clase de métodos utilizados para muestrear una distribución de probabilidad a posteriori utilizando una cadena de Márkov [35]. Las partículas son sometidas a un recorrido aleatorio a través

⁹Es una formulación matemática de la afirmación de que todo lo que es igual a una explicación simple es más probable que una explicación complicada.

de la distribución de probabilidad a posteriori, donde la probabilidad de moverse a un punto viene determinado por la probabilidad de transición de la cadena de Márkov. Al observar la posición de las partículas, o “caminantes”, en cada iteración, se va generando una imagen desde la distribución de probabilidad a posteriori [40].

Cuando se utilizan muestreadores MCMC, deben tenerse en consideración algunas sutilezas. La primera de ellas es que el comportamiento inicial de los caminantes MCMC depende en gran medida de las condiciones iniciales. Por lo tanto, es necesario incluir una fase de “quemado” para asegurarse de que el caminante se haya estabilizado antes de empezar a acumular muestras de la distribución de probabilidad a posteriori. Una vez que el caminante ha alcanzado un estado estable, el algoritmo puede continuar indefinidamente y, por lo tanto, es necesario que el usuario defina una condición de terminación. Esta condición suele elegirse cuando se han adquirido las muestras necesarias para que el usuario crea que ha obtenido una representación precisa de la distribución de probabilidad a posteriori [40].

Además, las posiciones de un caminante en una cadena suelen estar autocorrelacionadas; hecho que deriva en un muestreo poco fiel de la distribución de probabilidad a posteriori. Si no se aplicase ningún remedio, se subestimaría el ancho de esta distribución. Por lo tanto, es necesario adelgazar la cadena seleccionando muestras separadas por la longitud de autocorrelación de la cadena [40].

Los caminantes de MCMC también pueden fallar en encontrar modos múltiples de una distribución de probabilidad a posteriori si hay regiones de baja probabilidad a posteriori entre los modos. Sin embargo, esto puede ser mitigado ejecutando muchos caminantes que empiecen a explorar el espacio en diferentes puntos. Esto demuestra una forma de paralelizar los cálculos de MCMC para generar rápidamente muchas muestras. Han sido propuestas muchas variantes de MCMC para mejorar el rendimiento de estos algoritmos. El código MCMC más usado en astronomía es emcee [9] [40].

2.3.2.2. Muestreo anidado.

A diferencia de los métodos MCMC diseñados para extraer muestras de la distribución de probabilidad a posteriori, el muestreo anidado está diseñado para el cálculo de la evidencia 2.11. La generación de muestras a partir de la distribución de probabilidad a posteriori es un subproducto del cálculo de la evidencia del algoritmo. Al ponderar cada una de las muestras utilizadas para calcular la evidencia a través de la probabilidad a posteriori de la muestra, las muestras anidadas se convierten en muestras de esta probabilidad a posteriori [40].

Para el cálculo de la evidencia, se reduce la integral multidimensional 2.11 a una integral unidimensional, en términos del volumen de la distribución de probabilidad a priori contenido en un contorno de isoverosimilitud que cubre todos los valores tales que $\mathcal{L}(\theta) > \mathcal{L}^*$, definida como [37]

$$X(\mathcal{L}^*) \equiv \int_{\mathcal{L}(\theta) > \mathcal{L}^*} \pi(\theta) d\theta. \quad (2.19)$$

A medida que \mathcal{L}^* aumenta, la masa a priori encerrada $X(\mathcal{L}^*)$ va decreciendo desde 1 hasta 0. Escribiendo la función inversa, la evidencia podrá expresarse de la siguiente forma, [37]

$$\mathcal{Z} = \int_0^1 \mathcal{L}(X) dX. \quad (2.20)$$

El procedimiento seguido por el muestreo anidado independientemente de las dimensiones que tenga 2.11, es el siguiente:

- Rellenamos el espacio de parámetros con un conjunto de N puntos $(\theta_1, \dots, \theta_N)$ de la distribución de probabilidad a priori contenidos en el intervalo $(0, 1)$, a los que llamaremos “puntos vivos”. A cada punto vivo asociamos su correspondiente función de verosimilitud $(\mathcal{L}(\theta_1), \dots, \mathcal{L}(\theta_N))$. En la siguiente figura 2.2 se puede observar en concreto los vínculos establecidos entre los puntos de la distribución de probabilidad a priori con sus correspondientes contornos de isoverosimilitud que delimitan el volumen de la masa X .
- Empezando por $X_0 = 1$ y $\mathcal{Z} = 0$, para cada iteración $i = 1, \dots, j$:
 - Se escoge el peor punto vivo de todos, que es el que tiene un mayor X_i y menor \mathcal{L}_i , y se guarda su \mathcal{L}_i .
 - Se define $X_i = e^{-i/N}$.

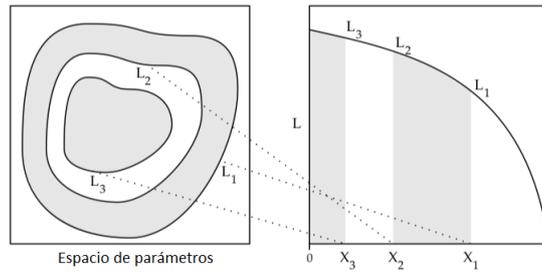


FIGURA 2.2: Contornos de la función de verosimilitud que contienen el volumen de la distribución de probabilidad a priori. Fuente: [37].

- Se establece un ancho del volumen de la distribución de probabilidad a priori definida como $w_i = X_{i-1} - X_i$ (caso simple) o bien $w_i = (X_{i_1} - X_{i+1})/2$ (caso de la regla de los trapecios).
- Se incrementa \mathcal{Z} con $\mathcal{L}_i w_i$. En este punto hacemos notar que la densidad de la distribución de probabilidad a posteriori se encuentra contenida en una pequeña fracción de la distribución de probabilidad a priori e^{-I} , donde I es la información definida como:

$$I = \int \log(dP/dX)dP. \quad (2.21)$$

- Se reemplaza el peor punto vivo escogido, al que ahora llamaremos “punto muerto”, por otro nuevo punto vivo extraído desde la distribución de probabilidad a priori sujeto a la restricción $\mathcal{L}(\theta) > \mathcal{L}_i$. El θ de este nuevo punto vivo se puede buscar por cualquier método que haga este proceso, como por ejemplo:
 - Método I: Reemplazando el punto muerto por uno nuevo a través de MCMC con una copia aleatoria de otro punto de los $N - 1$ restantes.
 - Método II: Generando un nuevo punto vivo a partir de una combinación de las coordenadas de los $N - 1$ puntos restantes.

Tanto si se hace por un método como por el otro, lo importante es que al acabar toda la iteración i se obtenga N muestras independientes.

- Finalmente, al haber comprimido el dominio de la distribución de probabilidad a priori a X_j , en el último paso se incrementa \mathcal{Z} rellenando el intervalo $0 < X < X_j$ con $w = N^{-1}X_j$ para cada punto muerto.

De esta manera, en cada iteración, el punto de verosimilitud más bajo es eliminado del conjunto de puntos vivos y las nuevas muestras son extraídas de la distribución de probabilidad a priori hasta que se encuentra un punto con mayor verosimilitud que el punto eliminado.

Como el algoritmo de muestreo anidado se mueve continuamente hacia regiones de mayor verosimilitud, es posible estimar un límite superior de la evidencia en cada iteración. Esto se hace imaginando que todo el volumen de la distribución de probabilidad a priori restante tiene una verosimilitud igual a la del punto vivo de mayor verosimilitud; se utiliza para informar a la condición de terminación del algoritmo. El algoritmo se detiene cuando la estimación actual de la evidencia está por encima de una cierta fracción del límite superior estimado ¹⁰.

Iterando hasta la condición de terminación, se genera una lista de puntos eliminados, a los que hemos llamado puntos muertos, utilizados para estimar la evidencia y hacer las inferencias de la distribución de probabilidad a posteriori. La evidencia es evaluada al asignar a cada punto muerto un volumen de la distribución de probabilidad a priori $w_i(\vec{t})$ y calculando la suma de la verosimilitud \mathcal{L}_i multiplicada por este volumen para cada muestra. Es decir, el algoritmo va aproximando la integral unidimensional 2.22 a una suma sobre los puntos muertos, [37]

$$\mathcal{Z}(\vec{t}) \sim \sum_{i \in \text{muerto}} \mathcal{L}_i w_i(\vec{t}). \quad (2.22)$$

¹⁰Se expresa como la diferencia entre la evidencia logarítmica calculada y el límite superior del logaritmo de la evidencia [40].

donde $\vec{t} = t_1, t_2, \dots, t_{N_{muerto}} \in (0, 1)$ es un conjunto desconocido de variables aleatorias para las N_{muerto} iteraciones.

Al ponderar con $w_i(\vec{t})$ cada una de las muestras utilizadas para el cálculo de la evidencia $\mathcal{Z}(\vec{t})$, éstas se convierten en muestras para la distribución de probabilidad a posteriori; que para la iteración i [37]:

$$p_i(\vec{t}) = \frac{w_i(\vec{t})\mathcal{L}_i}{\sum_i \mathcal{L}_i w_i(\vec{t})} = \frac{w_i(\vec{t})\mathcal{L}_i}{\mathcal{Z}(\vec{t})} \quad (2.23)$$

Entonces, dado un conjunto de puntos muertos sus correspondientes volúmenes de la distribución de probabilidad a priori $X_i(\vec{t}) = \prod_{k=0}^i t_k$ son desconocidos. Sin embargo, podemos cuantificar nuestro conocimiento de \mathcal{Z} simulando estadísticamente este conjunto \vec{t} de la siguiente forma [37]:

$$X_0 = 1, \quad X_i = t_i X_{i-1}, \quad X_i < X_{i-1}, \quad Pr(t_i) = N t_i^{N-1} \in (0, 1). \quad (2.24)$$

donde t_i es el mayor del conjunto \vec{t} y N el número total de puntos vivos.

De esta modo, se trabaja con un conjunto de evidencias resultantes $\mathcal{Z}_{\vec{t}}$ de $\mathcal{Z}(\vec{t})$. El valor medio y el error como desviación estándar de $\log(t)$ es [37]:

$$E[\log(t_i)] = \frac{1}{N}, \quad Var[\log(t_i)] = \frac{1}{N^2}. \quad (2.25)$$

Se espera que después de i iteraciones, el volumen de la distribución de probabilidad a priori se reduzca como $\log X_i \sim -(i \pm \sqrt{i})/N$. Por lo tanto para reducir al máximo la densidad de la distribución de probabilidad a posteriori se necesitan $NI \pm \sqrt{NI}$ pasos.

2.3.2.2.1 Dynesty.

Dynesty es el paquete de Python utilizado en este trabajo, para la estimación de las distribuciones de probabilidad a posteriori y las evidencias utilizando el algoritmo de muestreo anidado.

Aborda la generación de muestras a partir de la distribución de probabilidad a priori restringida, separando el problema en dos partes: primero construye distribuciones de contorno que abarquen el volumen de la distribución de probabilidad a priori; y a continuación, propone nuevos puntos vivos generando muestras condicionadas a encontrarse dentro de estos límites [38].

Dynesty utiliza la distribución del conjunto actual de puntos vivos para tratar de obtener una idea aproximada de la forma y el tamaño de las distintas regiones del volumen de la distribución de probabilidad a priori que se están muestreando. Con esta información intenta condicionar diferentes métodos de muestreo para tratar de mejorar la eficiencia. Actualmente, tiene 5 métodos para delimitar el volumen, mostrados en la figura 2.3:

- Sin límites, es decir, un cubo unitario que debe ser empleado como último recurso;
- un sólo elipsoide, actúa razonablemente bien al estimar estructuras cuya verosimilitud es aproximadamente gaussiana y unimodal;
- elipsoides múltiples, utilizado en nuestro programa, en casos más complejos multimodales es más útil la descomposición de los puntos vivos en grupos separados con sus propios elipsoides;
- muchas bolas superpuestas y
- muchos cubos superpuestos, ambos casos suelen aplicarse a dimensiones bajas ($dim \leq 5$).

Una vez construida una distribución de contorno, dynesty genera muestras condicionadas por estos límites, siguiendo la estrategia:

$$f(sC_b, \theta) \rightarrow \theta', \quad (2.26)$$

donde C_b es la covarianza asociada a un contorno delimitante en particular b (como por ejemplo, un elipsoide), θ es la posición inicial, θ' es la posición final propuesta y $s \sim 1$ es un factor de escala que se adapta en el transcurso del programa para asegurar que las tasas de aceptación son óptimas.

Para generar muestras dynesty tiene implementados 4 enfoques diferentes:

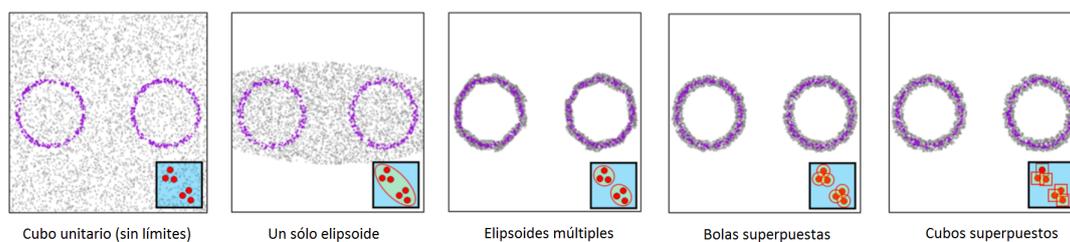


FIGURA 2.3: Distribuciones de contorno implementadas en dynesty. El conjunto actual de puntos vivos es mostrado en púrpura, mientras que la distintas distribuciones de contorno se observan en gris. En la esquina derecha de cada panel podemos ver una representación esquemática de cada distribución. [37].

- Muestreo uniforme, eficiente en dimensiones bajas ($dim < 10$) donde la distribución de contorno puede aproximarse mejor al volumen de la distribución de probabilidad a priori;
- caminatas aleatorias, utilizados en el programa, son menos sensibles al tamaño de la distribución de contorno y por lo tanto, tienden a funcionar mejor que el muestreo uniforme en espacios de dimensiones moderadas ($10 \leq dim \leq 20$), aún así luchan en espacios de alta dimensión debido a la cantidad de volumen exponencialmente creciente que necesitan explorar;
- muestreo de porciones multivariadas, y el
- muestreo hamiltoniano en porciones que a menudo ambos se desempeñan mejor en regímenes de alta dimensión ($dim > 20$).

A excepción del muestreo uniforme, que por construcción sólo puede muestrear directamente de la distribución de contorno, los otros métodos se basan en MCMC; generan muestras al evolucionar un punto vivo actual a una nueva posición.

| Sección 3.1. Resultados.

3.1. Resultados.

El principal objetivo en la astronomía de ondas gravitacionales es la comprensión del universo a través de las observaciones de radiación gravitacional. Además de requerir la detección de las ondas gravitacionales, deben inferirse las propiedades de sus fuentes. Con la finalidad de poder estudiar prometedoras detecciones de ondas gravitacionales procedentes de LIMBHs en el tercer periodo de observación actual O3, se han inyectado 48 señales hipotéticas de estos eventos en el ruido simulado de los dos detectores de LIGO: Hanford (H1) y Livingston (L1). El producto de este trabajo es traducido en un estudio del comportamiento del modelo IMRPhenomPv2 frente a la inyección de este tipo de señales, así como de las estimaciones de parámetros de dichos eventos, propia de los análisis de datos en esta nueva rama emergente de la astronomía.

Concretamente se inyectaron 24 señales con 1000 puntos vivos y 60 caminantes, y exactamente las mismas 24 señales con 14000 puntos vivos y 60 caminantes. En dimensiones moderadas ($dim \sim 10 - 20$) como la nuestra, $dim = 15$, suele recomendarse 50 caminantes o los necesarios para mantener las muestras independientes [11]. Puesto que la suma sobre los puntos muertos de la evidencia 2.22 en general tiene $m = i + N$ términos, es decir, el número de puntos vivos N más el número de iteraciones i del algoritmo, siendo $i \gg N$ [37]; era necesario encontrar una relación adecuada entre el número de caminantes y el número de puntos vivos. Desafortunadamente no hay una respuesta general y depende en gran medida de cada problema en particular. En nuestros 24 casos, tras múltiples pruebas con un número de caminantes cercano a 50, se determinó que para producir unos histogramas de las distribuciones de probabilidad a posteriori con la forma deseada, además de obtener una estimación de parámetros aceptable en general para todas las señales, se requerían 60 caminantes para cada i y aproximadamente alrededor de $m = 500000$ muestras. Para mantener la independencia de estas muestras con 60 caminantes, después de algunas pruebas, se estableció que como máximo debíamos fijar el número de puntos vivos a 14000. Por este motivo, en este trabajo, estudiamos señales para 60 caminantes con 14000 puntos vivos, y con 1000 puntos vivos con la finalidad de mostrar la diferencia entre ambos. Es sabido que podría conseguirse unos mejores resultados con un mayor número de muestras y de caminantes, sin embargo, atendiendo al tiempo de computación y a unos resultados razonablemente aceptables, han sido consideradas como válidas las características aquí expuestas.

En estas 24 señales, 12 de ellas se han destinado para una masa total de $100M_{\odot}$ y las otras 12 para una masa total de $200M_{\odot}$. Para cada una se han estudiado 3 casos relacionados con el cociente de masas: $q = 1, 3$ y 5 . En cada caso de q han sido considerados dos posibles valores del espín para el agujero negro primario $\chi_1 = 0.1$ y 0.7 , fijando a cero el espín del agujero negro secundario $\chi_2 = 0$. Además para cada χ_1 se ha tenido en cuenta dos posibles valores del ángulo de inclinación: $\iota = 0$ y $\pi/3$. Como resultado tenemos 24 binarias compactas procedentes de LIMBHs que junto con otras 24 para el estudio del número de puntos vivos, obtenemos un total de 48 señales.

A continuación, procederemos a describir de forma general el código utilizado en el trabajo, además de un análisis del mismo. Por razones de espacio, se ha decidido estudiar tres señales distintas para 1000 y 14000 puntos vivos, la señal más débil: $M_T = 100, q = 5, \chi_1 = 0.1, \iota = \pi/3$, la más fuerte: $M_T = 200, q = 1, \chi_1 = 0.7, \iota = 0$ y la señal correspondiente a $M_T = 100, q = 3, \chi_1 = 0.7, \iota = \pi/3$, puesto que posee una completa serie de características interesantes a comentar. Concretamente, para la descripción del programa en sí, se ha analizado la señal más débil y la más fuerte de entre las 24 para 14000 puntos vivos. En el estudio de las SNRs, del factor de Bayes y de las muestras han sido analizadas la señal más débil y más fuerte para 1000 y 14000 puntos vivos. Por último, en el análisis de las distribuciones de probabilidad a posteriori se ha elegido el tercer caso para 1000 puntos vivos y 14000 puntos vivos. En los tres casos, todas las disertaciones aquí mencionadas son igualmente extrapolables al resto de señales.

3.1.1. El programa.

En la construcción de cada señal, el primer paso es definir las polarizaciones de la onda gravitacional 2.2 comunes para ambos detectores, valiéndose del modelo. Como ejemplo ilustrativo en la siguiente

figura 3.1 podemos observar dichas polarizaciones para la señal más débil y la más fuerte entre las 24 simuladas para 14000 puntos vivos.

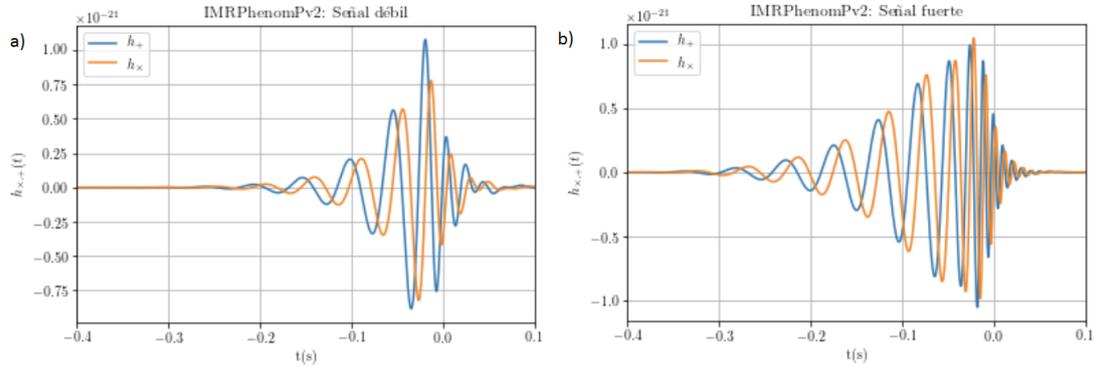


FIGURA 3.1: Modelo IMRPhenomPv2 para: a) la señal más débil y b) la más fuerte.

Fijémonos que, como era de esperar, al ser la señal b) más fuerte obtenemos un mayor número de ciclos que en la señal menos fuerte a).

El siguiente paso consiste en añadir el ruido de ambos detectores $n(t)$ 2.3 a las polarizaciones. Finalmente se inyectan las señales en cada detector, introduciendo los patrones de antena 2.1, para obtener los datos adquiridos por cada uno 2.3.

Una vez obtenidos los datos de la señal generada por cada detector, se definió una plantilla utilizando el mismo aproximante IMRPhenomPv2 para proceder a la estimación de parámetros. Se definieron 15 distribuciones de probabilidad a priori para cada uno de los parámetros que querían estimarse, descritos en la sección 2.1. Se eligieron $\pi(\theta)$ uniforme para todos los parámetros exceptuando los ya claramente conocidos con una distribución sinusoidal t_1 , t_2 y ι y con una distribución coseno δ . A continuación se definió la función de verosimilitud descrita en la sección 2.3.1.2. Finalmente se construyó un diagrama de esquina, mostrando las distribuciones de probabilidad a posteriori en una y dos dimensiones para cada parámetro. Además se calculó la SNR para los dos detectores de LIGO (H1 y L1) simulados, el logaritmo de la evidencia de la señal $\log(\mathcal{Z}_S)$, el logaritmo de la evidencia del ruido $\log(\mathcal{Z}_N)$, el logaritmo del factor de bayes $\log(BF)$ definidos en la sección 2.3.1.5 y el número de muestras del algoritmo n_s .

En la realización del código de este trabajo, además de haberse utilizado los paquetes más comunes de python como numpy o matplotlib, entre otros, ha sido posible gracias principalmente a los siguientes paquetes, propios del análisis bayesiano y del análisis de ondas gravitacionales:

- Bilby: una biblioteca de inferencia bayesiana, surgida recientemente en noviembre de 2018, para la astronomía de ondas gravitacionales, cuyo objetivo principal es el de proporcionar una interfaz para la estimación precisa y confiable de parámetros de eventos proporcionados por datos reales y de libre acceso de LIGO/Virgo o de datos simulados, mediante una sintaxis directa [15].
- Dynesty: un paquete utilizado para estimar distribuciones de probabilidad a posteriori y evidencias bayesianas empleando el algoritmo de muestreo anidado [38].

3.1.2. La evidencia, el número de muestras y las SNRs.

En la siguiente tabla 3.2 se muestran los cálculos realizados para la SNR de los dos detectores de LIGO (H1 y L1) simulados, el logaritmo de la evidencia de la señal $\log(\mathcal{Z}_S)$, el logaritmo de la evidencia del ruido $\log(\mathcal{Z}_N)$, el logaritmo del factor de bayes $\log(BF)$ definidos en la ecuación 2.17 y el número de muestras del algoritmo n_s para la señal más débil y la más fuerte de 1000 y 14000 puntos vivos.

Durante las pruebas realizadas, se advirtió, como puede observarse en esta misma tabla, que las SNRs para las señales con $\iota = 0$ eran mayores que las de los mismos eventos pero con $\iota = \pi/3$. Como la SNR es inversamente proporcional a la distancia luminosa d_L se definió, para cada χ_1 de las 24 señales, una d_L tal que se obtuvieran unas SNRs en las que, para los casos de $\iota = \pi/3$, se pasase el umbral ρ_0 descrito en la sección 2.2.2. De esta manera, automáticamente se obtendrían para los casos de $\iota = 0$ unos valores de SNRs aceptables. Las otras 24 señales para el otro punto vivo, tienen las mismas distancias luminosas. Es decir, a base de pruebas se definieron unas distancias luminosas apropiadas para que las SNR más bajas correspondientes a las de $\iota = \pi/3$ estuvieran por encima del umbral ρ_0 , de manera que para las 48 señales se considerase que en los datos del detector tenemos enterrada una señal de onda gravitacional

	1000 puntos vivos		14000 puntos vivos	
	Débil	Fuerte	Débil	Fuerte
SNR(H1 - L1)	14.05 - 11.41	27.87 - 22.57	14.05 - 11.41	27.87 - 22.57
log(\mathcal{Z}_N)	-8223.499	-8722.052	-8223.499	-8722.052
log(\mathcal{Z}_S)	-16281.036 \pm 0.102	-16789.241 \pm 0.154	-16281.608 \pm 0.026	-16782.899 \pm 0.034
log(BF)	-8057.537 \pm 0.102	-8067.189 \pm 0.154	-8058.109 \pm 0.026	-8060.846 \pm 0.034
n_s	28555	39570	424558	462804

FIGURA 3.2: Cálculo de la SNR para los dos detectores de LIGO (H1 y L1) simulados, $\log(\mathcal{Z}_S)$, $\log(\mathcal{Z}_N)$, $\log(BF)$ y n_s para la señal más débil y la más fuerte de 1000 y 14000 puntos vivos.

$h(t)$. De esta forma se obtuvieron señales para $\iota = \pi/3$ con unas SNR entorno a 8 – 20; frente a las señales para $\iota = 0$ con un rango de SNR más amplio entre 15 – 36, más propias del tercer período de observación en curso O3. Tal y como predice el artículo [8], los eventos de SNR más altos ofrecen unas mejores restricciones sobre los parámetros. En la tabla podemos verificar este hecho fijándonos en que, para los eventos de SNR más altos, tenemos un mayor número de muestras n_s , obteniendo una mejor resolución de las distribuciones de probabilidad a posteriori y por lo tanto restringiendo los parámetros. Además, como era de esperar, observamos que tenemos más n_s cuantos más puntos vivos tengamos. Finalmente obtenemos unos valores cuantitativamente esperados de $\log(\mathcal{Z}_N)$, $\log(\mathcal{Z}_S)$ y $\log(BF)$ para las señales débiles y fuertes para 1000 y 14000 puntos vivos. Resultados similares han sido adquiridos en los 44 casos restantes, en valor absoluto para: $\log(\mathcal{Z}_N)$ entre 8000 – 9000, $\log(\mathcal{Z}_S)$ entre 16000 – 17000 y $\log(BF)$ entre 8000 – 9000. El hecho de que el valor absoluto de $\log(BF)$ en nuestros eventos sea grande, pasándolo el umbral, indica una evidencia fuerte a favor de la hipótesis basada en la obtención de una señal enterrada en el ruido, frente a la hipótesis basada en obtener un modelo de sólo ruido.

3.1.3. Análisis de las distribuciones de probabilidad a posteriori.

En esta sección, se reprodujeron distribuciones de probabilidad a posteriori en una, dos y tres dimensiones, para dos señales idénticas ($M_T = 100$, $q = 3$, $\chi_1 = 0.7$, $\iota = \pi/3$) cuya única diferencia se encuentra en el número de puntos vivos (1000 y 14000).

A continuación presentamos la figura 3.3, en la cual se representan las distribuciones para las señales descritas, en una dimensión a través de un histograma y en dos dimensiones a través de un diagrama de dispersión, para los parámetros de la ascensión recta α y la declinación δ , y los dos espines adimensionales χ_1 y χ_2 .

Los 8 contornos mostrados en las gráficas de dos dimensiones, representan la iso-verosimilitud que contiene los volúmenes de la distribución de probabilidad a priori; y los puntos indican la distribución de las muestras del algoritmo. En escala degradada de rojo a lila, de más a menos, los contornos en rojo indican una mayor densidad de las muestras. Podemos visualizarlo mejor en la figura 3.4, donde se representan estas mismas gráficas en 3 dimensiones.

Observamos claramente, en la misma escala degradada de rojo a lila, que en la zona roja tenemos una mayor densidad de muestras. Además nos fijamos cómo el algoritmo ha dividido en rebanadas, tal y como comentábamos en la sección 2.3.2, la distribución de probabilidad a posteriori para generar muestras de cada una.

Comparando las gráficas 3.3 en una y dos dimensiones para cada parámetro, observamos que, obviamente, la densidad de las muestras es mayor donde la densidad de probabilidad a posteriori es mayor. Al tener un mayor número de puntos vivos en las gráficas *c*) y *d*) que en las de *a*) y *b*), se obtienen un mayor número de muestras. Además las distribuciones indican una fuerte correlación entre los espines adimensionales, así como una correlación más débil entre la ascensión recta y la declinación.

Fijándonos en las gráficas *b*) y *d*) (recordamos que son el mismo evento y por lo tanto tienen la misma SNR, por lo que la afirmación de la sección anterior apoyada por el artículo [8] sigue siendo válida), en concreto en las distribuciones unidimensionales, podría pensarse erróneamente que, por lo general, siempre que se aumenta el número de muestras, mejor es la estimación de parámetros puesto que al obtener unas distribuciones más acentuadas se obtienen unos errores menores; sin embargo, no siempre es así. Como veremos en las figuras 3.5 y 3.6, donde se grafican las líneas naranjas que indican los parámetros inyectados, y como vemos también en estas gráficas *a*) y *c*), al aumentar el número de puntos vivos la distribución se expande para el parámetro α y disminuye para el parámetro δ , cambiando la covarianza entre ambos y con ello estimando ambos parámetros (esto último se observa en la figura 3.6).

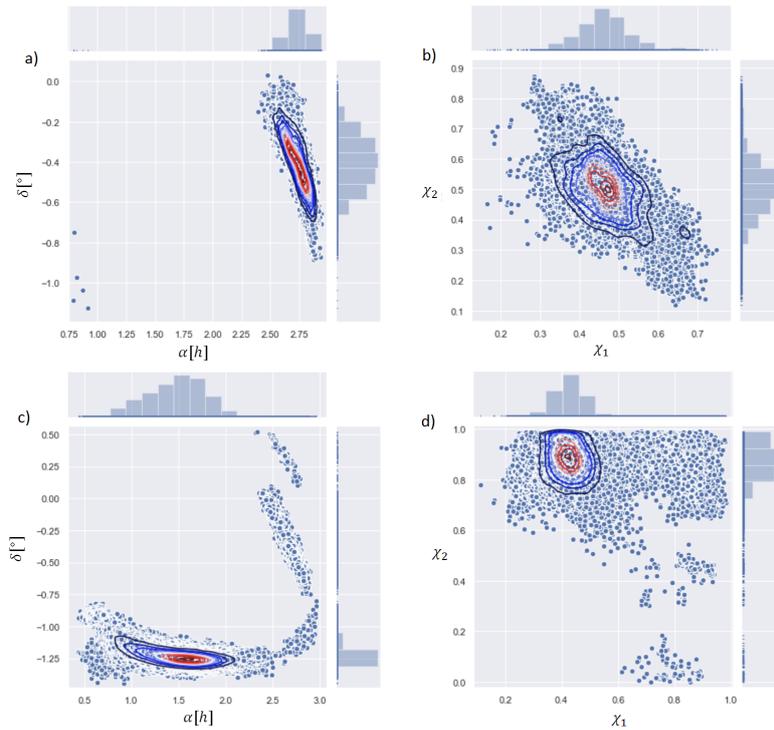


FIGURA 3.3: Distribución de probabilidad a posteriori en 1 y 2 dimensiones de los parámetros: a) α y δ para 1000 puntos vivos, b) χ_1 y χ_2 para 1000 puntos vivos, c) α y δ para 14000 puntos vivos, y d) χ_1 y χ_2 para 14000 puntos vivos.

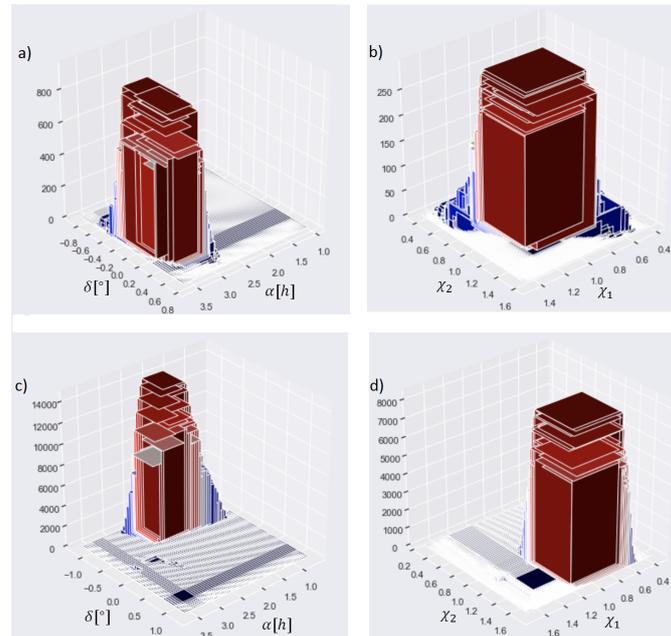


FIGURA 3.4: Distribución de probabilidad a posteriori en 3 dimensiones de los parámetros: a) α y δ para 1000 puntos vivos, b) χ_1 y χ_2 para 1000 puntos vivos, c) α y δ para 14000 puntos vivos, y d) χ_1 y χ_2 para 14000 puntos vivos.

En otras palabras, en las 24 señales simuladas, con SNRs diferentes para cada una, tendremos más o menos muestras dependiendo del problema en concreto. En aquellas con más muestras, la resolución de las distribuciones será mejor y con ello las restricciones sobre el parámetro. Si decidimos incrementar aún más la resolución aumentando el número de puntos vivos para un parámetro en concreto, por lo general, conseguiremos este aumento acompañado con una acertada restricción y estimación, pero, como acabamos de ver, no siempre se cumple.

Analicemos mejor este hecho comparando los diagramas de esquina 3.5 y 3.6 de las mismas señales para

los 15 parámetros, donde las líneas naranjas representan los valores inyectados, junto con un gráfico del rastro que dejan las muestras en el transcurso del programa.

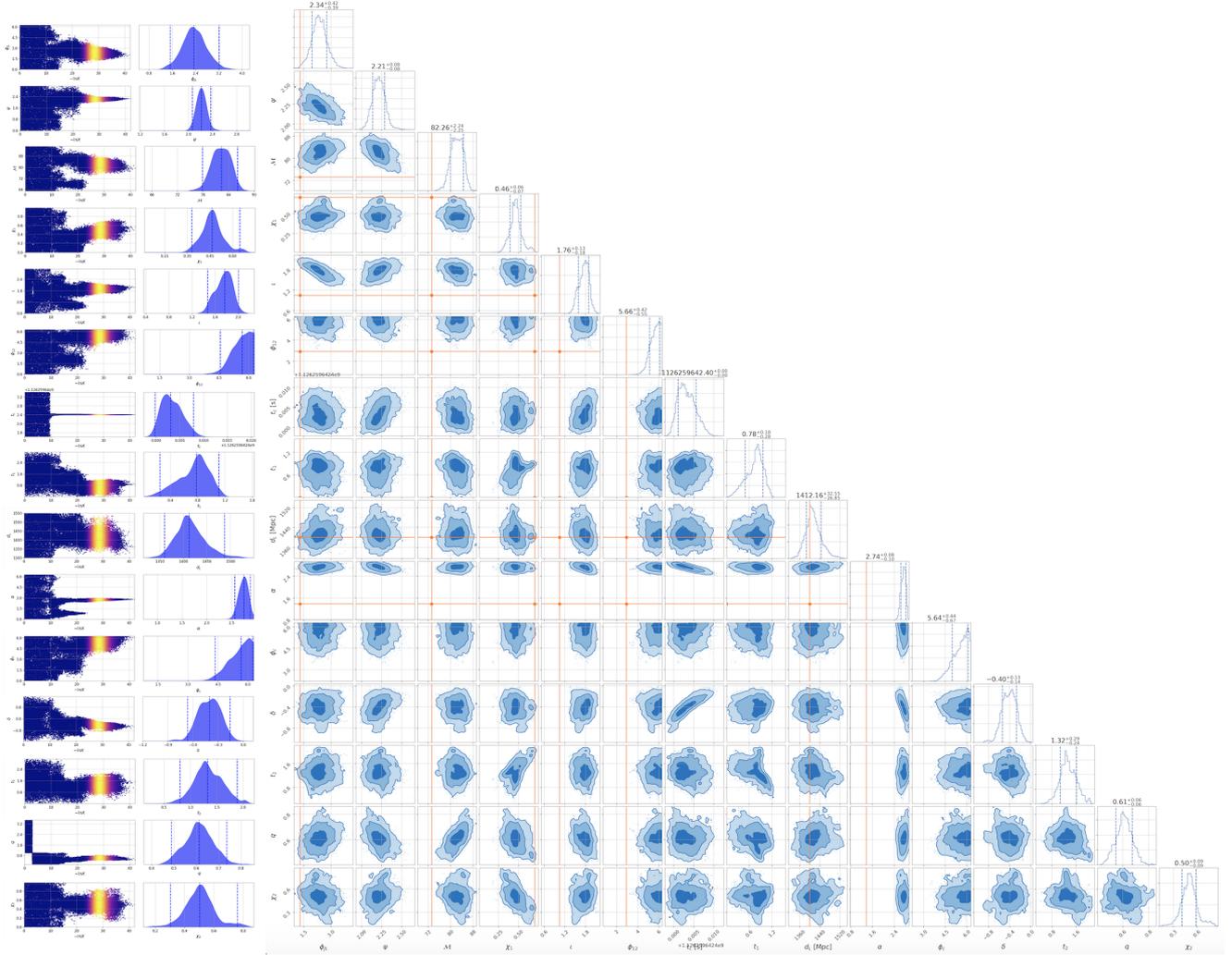


FIGURA 3.5: Ilustración de la estimación de parámetros para 1000 puntos vivos. Izquierda: Gráfico del rastro que dejan las posiciones, en una dimensión, de las muestras (los puntos muertos) en el transcurso de una ejecución de programa, coloreadas con la función de densidad del peso de [2.22](#). Derecha: Gráfica de esquina que muestra las distribuciones de probabilidad a posteriori en 1 y 2 dimensiones en azul y los valores inyectados en naranja, de los 15 parámetros.

Como decíamos, no hay que pensar que si tenemos un parámetro que para 1000 puntos vivos se obtiene una buena estimación, entonces para 14000 puntos vivos tendremos una estimación igual o mejor. Porque, comparando las gráficas [3.5](#) y [3.6](#), fijándonos en d_L , si bien es cierto que su resolución aumenta y con ello disminuye la restricción, su estimación empeora, sale de las barras de error. Tampoco hay que pensar que siempre que se aumenta el número de muestras, el número de puntos vivos, la resolución aumenta, disminuyendo los errores porque, comparando las gráficas [3.5](#) y [3.6](#) observamos que para ϕ_{JL} , ϕ_{12} , ψ y α la resolución disminuye, aumentando las restricciones, los errores. Como ya mencionamos, esta expansión de la distribución resulta beneficiosa para α y ϕ_{12} porque son estimados correctamente y para ψ porque lo estima mejor; sin embargo, empeora para ϕ_{JL} .

A pesar de ello, como podemos comprobar observando ambos diagramas de esquina, lo que sí es cierto es que, por lo general para los 15 parámetros, a medida que el número de muestras aumenta:

- las distribuciones adquieren una mejor resolución, traducida en unos histogramas mas acentuados, y con ello en una correlación entre los parámetros más débil; y que,
- en algunas señales, la estructura subyacente de la distribución empieza a emerger en más de dos etapas, en cuyos casos aparecen distribuciones cada vez más multimodales.

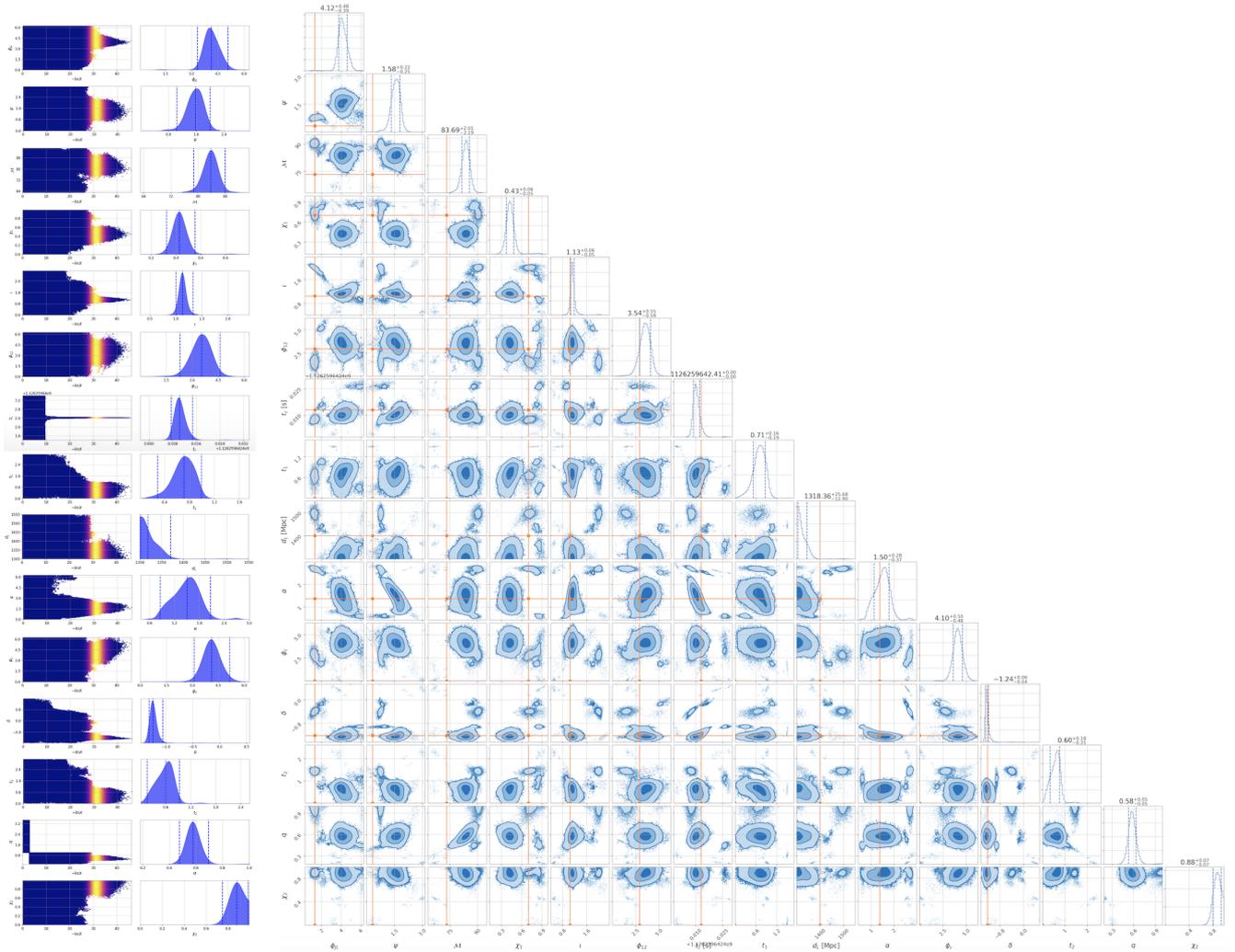


FIGURA 3.6: Ilustración de la estimación de parámetros para 14000 puntos vivos. Izquierda: Gráfico del rastro que dejan las posiciones, en una dimensión, de las muestras (los puntos muertos) en el transcurso de una ejecución de programa, coloreadas con la función de densidad de probabilidad del peso de 2.22. Derecha: Gráfica de esquina que muestra las distribuciones de probabilidad a posteriori en 1 y 2 dimensiones en azul y los valores inyectados en naranja, de los 15 parámetros.

En las cuatro figuras de 3.3, aunque se representen tan sólo 8 niveles de iso-verosimilitud, observamos que tenemos distribuciones multimodales. Fijándonos en la imagen 3.7, concretamente, en los histogramas *b*), *c*), *d*) y *e*), podemos observar esta multimodalidad; en la de χ_1 de *c*) se presenta un poco más acentuada. Estas curvas pronunciadas, pueden causar problemas para los métodos MCMC puesto que un caminante tardaría mucho tiempo en atravesar de un pico a otro. El mejor diagnóstico es empezar muchas cadenas en paralelo con diferentes inicializaciones y verificar si conducen a unas inferencias de las distribuciones de probabilidad a posteriori muy similares [17]. El algoritmo de muestreo anidado es un método de Monte Carlo que en un principio, era altamente ineficiente para las distribuciones de probabilidad a posteriori multimodales puesto que si la función de verosimilitud contenía alguna multimodalidad, separada sobre una buena porción del espacio de parámetros, solía escogerse una distribución de contorno de un sólo elipsoide sobre un pico. En 2007, el artículo [23] introdujo las distribuciones de contorno de elipsoides múltiples, utilizadas en este trabajo, en las cuales se detectan y aíslan los múltiples picos para después construir contornos elipsoidales alrededor de cada modo mejorando de esta manera la eficiencia del muestreo. En este mismo artículo [23] también se determina el valor esperado y la variancia de la evidencia final en una sola ejecución del algoritmo, proporcionando un aumento adicional a esta eficiencia. Esta multimodalidad la podemos observar con más claridad en la figura 3.8, donde se representan los mismos parámetros que en la gráfica 3.3.

Observamos claramente la multimodalidad de las distribuciones representadas por los elipsoides. En

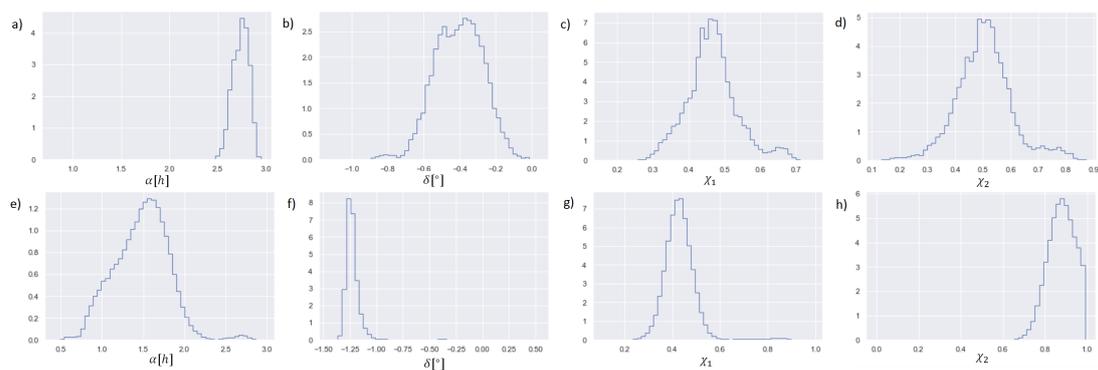


FIGURA 3.7: Distribución de probabilidad a posteriori en una dimensión de los parámetros mencionados para 1000 puntos vivos ($a - d$) y para 14000 puntos vivos ($e - h$).

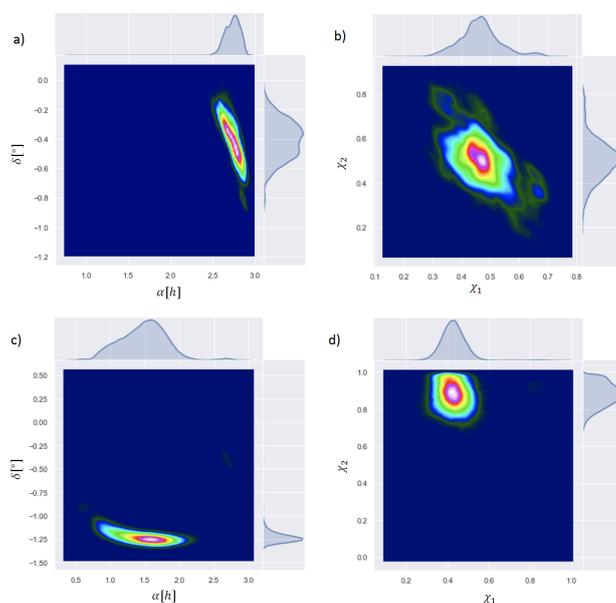


FIGURA 3.8: Distribución de probabilidad a posteriori en 1 y 2 dimensiones de los parámetros: a) α y δ para 1000 puntos vivos, b) χ_1 y χ_2 para 1000 puntos vivos, c) α y δ para 14000 puntos vivos, y d) χ_1 y χ_2 para 14000 puntos vivos.

c) observamos un pico en torno a $(2.50, -0.30)$ y otro en torno a $(0.60, -0.90)$. En d) lo tenemos en $(0.85, 0.90)$.

Para acabar, como podemos ver en la figura 3.5 y 3.6, concretamente en los gráficos del rastro de las muestras, gracias a la distribución de contorno de elipsoides múltiples escogida, el algoritmo identifica con confianza los modos finales convergiendo aproximadamente a los parámetros del modelo; obteniendo así en la gráfica 3.6 una buena estimación de 7 parámetros.

Conclusiones.

La detección de señales procedentes de LIMBHs, sería la primera evidencia directa de su existencia implicando otro gran avance en la astronomía de las ondas gravitacionales. Infiriendo las propiedades de sus fuentes, a través de una estimación de parámetros, podríamos hacer afirmaciones sobre su formación y trayectoria evolutiva respecto a ser las semillas buscadas de SMBHs, además de poder establecer vínculos con otros fenómenos como su posible relación con sistemas de fuente X ultraluminosa [24]. Es bastante prometedora la perspectiva de establecer su presencia en el tercer periodo de observación O3 en curso, siendo muy probable que su formación se deba a la fusión de dos agujeros negros de masa estelar masiva [44]. Por este motivo se han inyectado varias señales procedentes de LIMBHs en el ruido de los dos detectores de LIGO y se han recuperado sus parámetros a través de una estimación, con la finalidad de poder observar qué tipo de comportamiento esperamos.

Utilizando el modelo IMRPhenomPv2, se generaron varias señales de LIMBH que cubrieran el rango de masa interesado, escogiendo que la masa total de la binaria estuviera entre $100M_{\odot}$ y $200M_{\odot}$. Para cada una se estudiaron tres cocientes de masas: $q = 1, 3$ y 5 . En cada q , 2 valores de los espines primarios $\chi_1 = 0.1$ y 0.7 , fijando a 0 el espin secundario χ_2 . Para cada caso de χ_1 se estudiaron además dos posibles valores de la inclinación $\iota = 0$ y $\pi/3$. Además para cada inyección se le asignó una posición en el cielo. Finalmente, como describimos en la sección 3.1.2, se seleccionó la distancia a cada fuente d_L para obtener una relación señal-ruido (SNR) apropiada. Para la estimación de parámetros se utilizó el algoritmo de muestreo anidado.

En la realización de un análisis bayesiano de datos astronómicos, en general, surgen principalmente dos problemas difíciles de resolver:

- En la estimación de parámetros de algunos modelos para los datos, las resultantes distribuciones de probabilidad a priori pueden ser multimodales o exhibir curvas pronunciadas causando problemas en los métodos MCMC. Puesto que si la separación entre los picos de la probabilidad a posteriori es pronunciada, entonces la probabilidad de transición de uno a otro será muy pequeña. De manera que, la cadena parecerá que ha convergido después de moverse por uno de los picos; pero después de un tiempo finito acabará saltando repentinamente a otro pico, revelando que las indicaciones iniciales de convergencia eran incorrectas.
- El cálculo de la evidencia para, en nuestro caso, un modelo de señal y para un modelo de ruido, es computacionalmente costoso.

El algoritmo de muestreo anidado original presentado en 2004 [36] reduce en gran medida el gasto computacional del cálculo de evidencias y produce inferencias sobre la distribución de probabilidad a posteriori como subproducto. Sin embargo, su implementación era sólo eficiente para distribuciones unimodales y sin curvas muy pronunciadas. En 2007 se presentó un algoritmo de muestreo anidado [23] significativamente más eficiente en el muestreo de distribuciones de probabilidad a posteriori multimodales, introduciendo las distribuciones de contorno de elipsoides múltiples utilizadas en este trabajo; además el algoritmo proporcionó una estimación de la varianza y el valor esperado de la evidencia en una sola ejecución del programa. En 2018 se introdujo el algoritmo de muestreo anidado dinámico [13], una generalización del algoritmo de muestreo anidado en el cual el número de puntos vivos varía para distribuir las muestras de manera más eficiente. Se ha probado que, en comparación con el muestreo anidado estático con el mismo número de muestras, este nuevo método mejora significativamente la precisión del cálculo acelerándolo en un factor ~ 72 para la estimación de parámetros y en un factor ~ 7 para los cálculos de la evidencia [13].

En conclusión, generalmente para una buena estimación de parámetros la elección de los caminantes y del número de puntos vivos dependerá mucho del problema de estudio. En nuestro caso, después de algunas pruebas, para la obtención de una estimación óptima de los 15 parámetros, se fijó a 60 el número de caminantes y se observó que, por lo general, a mayor número de puntos vivos las distribuciones adquirirían una mejor resolución, una menor correlación entre los parámetros y una mejor estimación de los mismos.

Para mejorar los resultados obtenidos, podría implementarse el algoritmo de muestreo anidado dinámico y realizar pruebas de diagnóstico para evaluar la confiabilidad tanto de la estimación de parámetros como de los cálculos de la evidencia. Como el algoritmo de muestreo anidado es una técnica utilizada cada vez más en problemas multimodales de alta dimensionalidad, sin una configuración del programa adecuado,

el algoritmo podría, por ejemplo, no explorar las distribuciones a posteriori correctamente produciendo muestras correlacionadas o pasar por alto modos importantes en distribuciones multimodales. En este último caso, nos referimos a aquellos posibles casos en los que el algoritmo no detecte el modo y como resultado las muestras no se extraigan de todo el volumen de la distribución de probabilidad a priori especificado por los contornos de iso-verosimilitud.

Quizás, si se demostrase que podemos despreciar estos efectos podríamos aumentar, con más confianza, el número de puntos vivos para obtener resultados aún más precisos.

Bibliografía

- [1] Nobel Media AB 2019. «Press release: The Nobel Prize in Physics 2017». En: *Press release: The Nobel Prize in Physics 2017. NobelPrize.org. Nobel Media AB 2019. Tue. 18 Jun 2019.* (2017).
- [2] M. Carmen Díaz Batanero. «Viabilidad de la enseñanza de la inferencia bayesiana en el análisis de datos en psicología». Tesis doct. Universidad de Granada, 2007.
- [3] Ben Margalit y Brian D. Metzger. «Constraining the maximum mass of neutron stars from multi-messenger observations of GW170817». En: *arXiv:1710.05938v2* (2017).
- [4] Juan Calderón Bustillo. «Sub-dominant modes of the Gravitational Radiation from Compact Binary Coalescences: Construction of hybrid waveforms and Impact on Gravitational Wave searches». Tesis doct. Universitat de les Illes Balears, 2015.
- [5] R H Cole C J Moore y C P L Berry. «Gravitational-wave sensitivity curves». En: *arXiv:1408.0740v2* (2014).
- [6] LIGO Caltech. *Sources and Types of Gravitational Waves*. Disponible en <https://www.ligo.caltech.edu/page/gw-sources>.
- [7] Jolien D. E. Creighton y Warren G. Anderson. *Gravitational-Wave Physics and Astronomy. An Introduction to Theory, Experiment and Data Analysis*. WILEY-VCH Verlag GmbH & Co. KGaA, 2011.
- [8] Hsin-Yu Chen y Daniel E. Holz. «The Loudest Gravitational Wave Events». En: *arXiv:1409.0522v2* (2014).
- [9] Dustin Lang Daniel Foreman-Mackey David W. Hogg y Jonathan Goodman. «emcee: The MCMC Hammer». En: *arXiv:1202.3665v4* (2013).
- [10] Allyn F. Tennant Douglas A. Swartz Kajal K. Ghosh y Kinwah Wu. «The Ultra-Luminous X-ray source population from the Chandra archive of galaxies». En: *arXiv:astro-ph/0405498v2* (2004).
- [11] Joshua S. Speagle y colaboradores de la revisión eb3b313b. «Preguntas más frecuentes». En: *GitHub* (2017).
- [12] José Edelstein y Andrés Gomberoff. *Einstein para perplejos*. Editorial Debate, 2018.
- [13] Michael Hobson y Anthony Lasenby Edward Higson Will Handley. «Dynamic nested sampling: an improved algorithm for parameter estimation and evidence calculation». En: *arXiv:1704.03459v4* (2018).
- [14] Giacomo Fragione y Omer Bromberg. «Eccentric binary black hole mergers in globular clusters hosting intermediate-mass black holes». En: *arXiv:1903.09659v1* (2019).
- [15] Paul D. Lasky Colm Talbot Kendall Ackley Sylvia Biscoveanu Qi Chu Atul Divarkala Paul J. Easter Boris Goncharov Francisco Hernandez Vivanco Jan Harms Marcus E. Lower Grant D. Meadors Denyz Melchor Ethan Payne Matthew D. Pitkin Jade Powell Nikhil Sarin Rory J. E. Smith Gregory Ashton Moritz Hübner y Eric Thrane. «Bilby: A user-friendly Bayesian inference library for gravitational-wave astronomy». En: *arXiv:1811.02042v1* (2018).
- [16] W. K. Hastings. «Monte Carlo Sampling Methods Using Markov Chains and Their Applications». En: *Biometrika, Vol. 57, No. 1. (Apr., 1970), pp. 97-109* (1970).
- [17] David W. Hogg y Daniel Foreman-Mackey. «Data Analysis Recipes: Using Markov Chain Monte Carlo». En: *The Astrophysical Journal Supplement Series, 236:11 (18pp), 2018 May; doi: 10.3847/1538-4365/aab76e* (2018).
- [18] H. Pérez y J. Solbes. «Implicaciones de la evolución histórica de algunos conceptos en la enseñanza de la relatividad». En: *Rev. Eureka. Enseñ. Divul. Cien., 2006, 3(3), pp. 409-431* (2006).
- [19] B. Farr W. Farr P. Graff S. Vitale B. Aylott K. Blackburn N. Christensen M. Coughlin W. Del Pozzo F. Feroz J. Gair C.-J. Haster V. Kalogera T. Littenberg I. Mandel R. O'Shaughnessy M. Pitkin C. Rodriguez C. Röver T. Sidery R. Smith M. Van Der Sluys A. Vecchio W. Vousden J. Veitch V. Raymond y L. Wade. «Parameter estimation for compact binaries with ground-based gravitational-wave observations using LALInference». En: *arXiv:1409.7215v2* (2015).
- [20] B. Farr W. Farr P. Graff S. Vitale B. Aylott K. Blackburn N. Christensen M. Coughlin W. Del Pozzo F. Feroz J. Gair C.-J. Haster V. Kalogera T. Littenberg I. Mandel R. O'Shaughnessy M. Pitkin C. Rodriguez C. Röver T. Sidery R. Smith M. Van Der Sluys A. Vecchio W. Vousden J. Veitch V. Raymond y L. Wade. «Parameter estimation for compact binaries with ground-based gravitational-wave observations using LALInference». En: *Phys.Rev. D91 (2015) no.4, 042003* (2015).
- [21] Piotr Jaranowski y Andrzej Królak. *Analysis of Gravitational-Wave Data*. Cambridge University Press, 2009.

- [22] Sean T. McWilliams Joan Centrella John G. Baker James R. van Meter y Bernard J. Kelly. «Consistency of post-Newtonian waveforms with numerical relativity». En: *arXiv:gr-qc/0612024v2* (2007).
- [23] M. Bridges y M.P. Hobson J.R. Shaw. «Efficient Bayesian inference for multimodal problems in cosmology». En: *arXiv:astro-ph/0701867v2* (2007).
- [24] N. V. Krishnendu y Anjali B. Yelikar. «Testing the Kerr nature of intermediate-mass and supermassive black hole binaries using spin-induced multipole moment measurements». En: *arXiv:1904.12712v1* (2019).
- [25] B. P. Abbot et al. (LIGO Scientific Collaboration y Virgo Collaboration). «Search for intermediate mass black hole binaries in the first observing run of Advanced LIGO». En: *arXiv:1704.04628v4* (2017).
- [26] Ilya Mandel y Ben Farr. «LIGO Magazine: Black Holes in Advanced LIGO the observational payoff!» En: *LIGO Scientific Collaboration* (2013).
- [27] Alejandro Bohé Leïla Haegel Sascha Husa Frank Ohme Geraint Pratten y Michael Pürrer Mark Hannam Patricia Schmidt. «A simple model of complete precessing black-hole-binary gravitational waveforms». En: *arXiv:1308.3271v2* (2014).
- [28] Charles W. Misner, Kip S. Thorne y John Archibald Wheeler. *Gravitation*. W. H. Freeman y Princeton University Press, 1973.
- [29] Marshall N. Rosenbluth Augusta H. Teller Nicholas Metropolis Arianna W. Rosenbluth y Edward Teller. «Equation of State Calculations by Fast Computing Machines». En: *J. Chem. Phys.* 21, 1087 (1953); doi: 10.1063/1.1699114 (1953).
- [30] Chris North. *GW animations*.
- [31] J. R. Oppenheimer y H. Snyder. «On Continued Gravitational Contraction». En: *Physical Review* (1939).
- [32] Saúl Ramos-Sánchez. *Relatividad para futuros físicos*. CopIt Arxiv publishing Open Access with an Open Mind, 2018.
- [33] for the LIGO Scientific Collaboration S. J. Waldman. «The Advanced LIGO Gravitational Wave Detector». En: ().
- [34] Mark Hannam y Frank Ohme Sebastian Khan Katerina Chatziioannou. «Phenomenological model for the gravitational-wave signal from precessing binary black holes with two-spin effects». En: *arXiv:1809.10113v1* (2018).
- [35] Sanjib Sharma. «Markov Chain Monte Carlo Methods for Bayesian Data Analysis in Astronomy». En: *Annu. Rev. Astron. Astrophys.* 2017. 55:213–59; doi: 10.1146/annurev-astro-082214-122339 (2017).
- [36] John Skilling. «Nested Sampling». En: *AIP Conference Proceedings* 735, 395 (2004); doi: 10.1063/1.1835238 (2004).
- [37] John Skilling. «Nested Sampling for General Bayesian Computation». En: *Bayesian Anal.* 1 (2006), no. 4, 833–859; doi:10.1214/06-BA127 (2006).
- [38] Joshua S. Speagle. «dynesty: A Dynamic Nested Sampling Package for Estimating Bayesian Posteriors and Evidences». En: *arXiv:1904.02180v1* (2019).
- [39] Kathy Svitol y col. «Gravitational waves detected 100 years after Einstein’s prediction». En: *LIGO Scientific Collaboration News* (2016).
- [40] Eric Thrane y Colm Talbot. «An introduction to Bayesian inference in gravitational-wave astronomy: parameter estimation, model selection, and hierarchical models». En: *arXiv:1809.02293v4* (2019).
- [41] Neil J. Cornish Travis Robson y Chang Liu. «The construction and use of LISA sensitivity curves». En: *arXiv:1803.01944v2* (2018).
- [42] Xabier Cid Vidal y Ramón Cid Manzano. «¿Es la Tierra un lugar seguro para el LHC (CERN)?» En: *Real Sociedad Española de Física* (2008).
- [43] Robert M. Wald. *General Relativity*. The University of Chicago Press, 1984.
- [44] HaoWang XiangLi Yi-Ming Hu Zhi-Ping Jin Yi-Zhong Fan En-Wei Liang Yun-FengLiang Yuan-ZhuWang y Da-Ming Wei. «Hearing the birth of the lightest intermediate mass black holes with the second generation gravitational wave detectors». En: *arXiv:1705.01881v1* (2017).