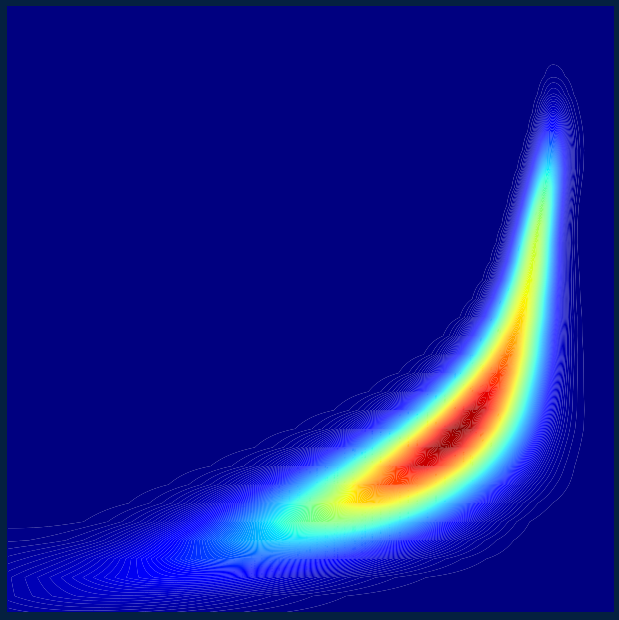
Gravitational-Wave Astronomy: Modelling, detection, and data analysis



Alejandro Torres Forné

Doctorat en Física
Departament d'Astronomia i Astrofísica
Universitat de València

TESIS DOCTORAL
2017

VNIVERSITAT DE VALÈNCIA

Gravitational-Wave Astronomy: Modelling, detection, and data analysis

Alejandro Torres Forné

Doctorat en Física
Departament d'Astronomia i Astrofísica
Universitat de València

TESIS DOCTORAL

Directores:
José Antonio Font
Pablo Cerdá Durán

Mayo 2017

A Ángel, a Rosario y a Laura.
A los petimetres y a los parolos.

Agradecimientos

Esta tesis doctoral nunca habría sido posible sin la ayuda, el apoyo y los consejos de muchas personas.

Primero quiero agradecer a mis dos directores Prof. Toni Font y Prof. Pablo Cerdá-Durán por haberme guiado durante todo el desarrollo de la tesis, por sus siempre acertados consejos y por su paciencia aguantando mis constantes visitas a sus respectivos despachos.

También quiero dar las gracias a Antonio Marquina y a José María Ibáñez. A Antonio por haber sido la persona clave en el desarrollo de la parte de data analysis y por lo mucho que he aprendido con sus explicaciones. A José María quiero agradecerle sus consejos, sus comentarios desde la experiencia y sobre todo por animarme siempre cuando encontrábamos alguna dificultad.

A Elena Cuoco quiero agradecerle su hospitalidad en mis múltiples visitas a Pisa. También darle las gracias por haberme incluido en sus colaboraciones y por haber sido en parte responsable de que el grupo de Valencia se uniera a la colaboración Virgo y por iniciar el grupo de Machine Learning dentro de la colaboración Virgo/LIGO.

Agradecerles a Jade Powell, Ik Siong Heng, Marco Cavaglià y Daniele Trifirò las facilidades que me dieron para unirme a su proyecto conjunto, resolviendo mis dudas y creando un entorno perfecto para trabajar.

También quiero darles las gracias a todos los miembros del Departamento de Astronomía y Astrofísica, en especial a Manel, a Feli, a Lupe y Enric, porque ellos el motor que hace que el departamento funcione. Afortunamente siempre están ellos atentos a los plazos, a los documentos que hay que entregar y sobre todo al bienestar general de todos.

A mis compañeros, Nico, Jose, Jesús, Vass, Nicolás, Tomasz, Isa, Adri y Jens porque sin ellos estos 4 años no habrían sido ni la mitad de divertidos. Siempre recordaré las risas, los spoilers inesperados, la radio de Baza, a Juan Antonio Chistoso y las mil y una anécdotas que hemos compartido.

También quiero dar las gracias a mis amigos los parolos, los petimetres y los locos porque todos y cada uno son personas increibles. En especial, y sin que sirva de démerito de los demás, quiero darles las gracias a Juan, a Rafa, a Carlos, a Buris, a Pablo, a María, a Claudia, a Marta, a Silvia, a Josete, a David y a Nadyl, porque a pesar de que a muchos los conozco desde hace más de 20 años, su amistad, su confianza y su apoyo son indispensables para mi.

Finalmente, quiero hacer mención especial a mi familia, a mis padres y mi hermana, por su apoyo y su cariño y porque sin ellos no sería la persona que soy ahora.

Esta investigación ha sido financiada por el MICINN (AYA 2010-21097-C03-01), MINECO (MTM2011-28043) and by the Generalitat Valenciana (PROMETEO-2009-103).

Acknowledgements

This doctoral thesis would never have been possible without the help, support and advice of many people.

First I want to thank my supervisors Prof. Toni Font and Prof. Pablo Cerdá-Durán for guiding me during all the development of the thesis, for their always successful advice and for their patience enduring my constant visits to their respective offices.

I also want to thank Antonio Marquina and José María Ibáñez. To Antonio for having been the key person in the development of the data analysis part and for how much I have learned with his explanations. To José María I want to thank for his advice, his comments from the experience and above all for always encouraging me when we find some difficulty.

To Elena Cuoco I want to thank for her hospitality on my multiple visits to Pisa. I also thank for including me in her collaborations and for having been partly responsible for the group of Valencia to join the collaboration Virgo and for start the Machine Learning group within the Virgo / LIGO collaboration.

I thank Jade Powell, Ik Siong Heng, Marco Cavaglià and Daniele Trifirò for the facilities they gave me to join their joint project, solving my doubts and creating a perfect working environment.

I also want to thank all the members of the Department of Astronomy and Astrophysics, especially Manel, Feli, Lupe and Enric, because they are the engine that makes the department work. Fortunately, they are always attentive to the deadlines, the documents to be delivered and, above all, the general welfare of all.

To my mates Nico, Jose, Jesús, Vass, Nicolas, Tomasz, Isa, Adri and Jens because without them these 4 years would not have been as funny as have been. Always remember the laughs, the unexpected spoilers, the radio of Baza, Juan Antonio Chistoso and the thousand and one stories we have shared.

I also want to thank my friends of the parolos, the petimetres and the locos, because they are the family that you choose and I am fortunate to have so many

incredible people nearby. In particular, and without being a demeanor of others, I want to thank Juan, Rafa, Carlos, Buris, Pablo, María, Claudia, Marta, Silvia, Josete, David and Nadyl, because in spite of the fact that I have known many of them for more than 20 years, our friendship, trust and mutual support are indispensable for me.

Finally, I want to make a special mention to my family, especially my parents and my sister, for their support of all the decisions I have made, for their unconditional love and because without them, I never became the person I am now.

This research have been funded by the Spanish MICINN (AYA 2010-21097-C03-01), MINECO (MTM2011-28043) and by the Generalitat Valenciana (PROMETEO-2009-103).

Abstract

During this thesis, I have worked on three main topics of research related with gravitational-wave astronomy, namely (i) on the numerical modelling of astrophysical sources of gravitational waves, (ii) on the detector characterization and classification of transients of noise, and (iii) on the development of new methods for data analysis.

In the context of numerical relativity I have simulated accretion processes onto neutron stars to study the so-called hidden magnetic field scenario. I have developed a 1D model and performed a parameter-space study aimed at determining under which conditions this scenario can be a viable model to explain the low magnetic fields observed in some central compact objects in supernova remnants.The cause of those low values of the magnetic field is still unclear. Previous numerical simulations suggested that the magnetic field can be compressed to the surface of the star due to the pressure exerted by the infalling fluid. Our results show that the accretion rate required to compress the magnetosphere of a typical pulsar is modest. Therefore, this scenario should not be regarded as particularly unusual. However, our results also show that it is fairly complicated to compress the magnetic field if it is stronger than 10^{14} G, which are the typical values found in magnetars, since the required accretion rate would cause the star to collapse to a black hole.

To this day, the physical mechanisms behind core-collapse supernovae explosions and their subsequent evolution are still not entirely known. The information contained in the gravitational-wave signal produced by this type of sources can greatly help determining the physics involved in the explosions. This thesis presents the first results of a project aimed at inferring some of those physical parameters from the study of gravitational-wave signals from core-collapse. I have studied the existing relationship between the modes of oscillation of the proto-neutron star that forms after the collapse of a massive star and the spectrum of the gravitational signal. The model handles the oscillations of the proto-neutron star as the perturbations of a system in equilibrium. We compare

the data from the gravitational-wave signal generated by the simulation of the collapse of a massive star with the time-frequency distribution of the different modes of oscillation, obtaining a remarkably close correspondence between them.

Detectors of gravitational waves are affected by many sources of noise due to the extreme sensitivity required to measure the small-amplitude variations caused in the distance between test masses by passing gravitational waves. The search for the sources of noise and their subsequent elimination is a fundamental task. This thesis presents results of a collaborative project to automatically classify and remove noise transients (glitches) produced in the advanced LIGO and Virgo detectors. Some of these transients may be particularly problematic because they can be misinterpreted as true gravitational-wave signals. The three methods employed in the project are able to correctly classify 95% of detected glitches. Since all three methods use different strategies to perform the classification, they are complementary, so that transients not classified by one of the methods may be classified by the other two.

In spite of the efforts to reduce the noise of interferometers, it is inevitable that part of the noise may affect and bury actual gravitational-wave signals. There are many data-analysis methods designed to extract signals from noisy backgrounds. During this thesis I have explored the performance of denoising algorithms based on the concept of Total Variation. These algorithms, which do not require any a priori information about the signal, have been shown to be highly efficient for noise suppression in the context of image processing. Our pioneering results for gravitational-wave signals show that the algorithms can remove enough noise to produce distinguishable signals in the two scenarios we have considered, signals mixed with Gaussian noise and with real detector noise. One of the most interesting future applications of this line of work is the combination of these methods with other common techniques of gravitational-wave analysis (e.g. Bayesian inference) to improve the results. Finally, in an attempt to bridge the gap between numerical modelling and data analysis, this thesis has also explored the use of dictionary-learning techniques with numerical-relativity waveform templates in order to reconstruct signals embedded in Gaussian noise. These techniques offer a number of possibilities, not only to extract signals from noise, but also to classify glitches or to extract physical parameters from detected signals.

Resumen

Introducción

La detección directa de la primera señal de ondas gravitatorias, el 14 de Septiembre de 2015, puede considerarse uno de los mayores hitos científicos de todos los tiempos. No solo porque supone la confirmación de la última de las predicciones de la Teoría de la Relatividad General de Albert Einstein, sino porque anticipa una autentica revolución en el campo de las astrofísica, comparable a la producida con la invención del telescopio por Galileo Galilei en 1609. Este descubrimiento ha inaugurado un nuevo tipo de astronomía, la astronomía de ondas gravitatorias. Se abre así una nueva ventana al universo que permitirá el estudio de procesos físicos producidos en regiones no accesibles al espectro electromagnético, con lo que se podrá obtener información clave sobre la estructura de los agujeros negros, la evolución estelar, la ecuación de estado de las estrellas de neutrones o el universo primitivo. En esta tesis he trabajado fundamentalmente en tres temas de investigación relacionados con las ondas gravitatorias, como son, (i) el modelado numérico de fuentes astrofísicas de radiación gravitatoria, (ii) la caracterización y clasificación del ruido presente en los detectores interferométricos, y (iii) el desarrollo de nuevas técnicas de análisis de datos de señales gravitatorias.

Objetivos

El objetivo global de la presente tesis es estudiar los diferentes aspectos que se incluyen en campo de la astronomía de ondas gravitatorias. Cada uno de los tres apartados, modelado, detección y análisis de datos, tienen sus objetivos generales que a su vez se pueden desarrollar en objetivos específicos dentro de cada estudio realizado.

El modelado de fuentes astrofísicas de radiación gravitatoria es clave para entender los procesos físicos que llevan a la generación de una determinada señal gravitatoria susceptible de ser detectada. Por lo tanto, el objetivo general de esta

sección es conocer de manera general los distintos tipos de fuentes astrofísicas de señal gravitatoria y las métodos de relatividad numérica empleados en las simulaciones. En este contexto se han estudiado dos escenarios particulares, el proceso de acreción sobre estrellas de neutrones magnetizadas y los modos de oscilación de protoestrellas de neutrones.

Más especificamente, se han estudiado los procesos de acreción sobre estrellas de neutrones que dan origen a un fenómeno conocido como "enterramiento" del campo magnético. El objetivo es averiguar bajo qué condiciones es posible que dicho fenómeno proporcione una explicación válida para los inusualmente bajos valores del campo magnético observados en algunas estrellas de neutrones que se encuentran en el centro de restos de supernovas. La explicación de dichos valores del campo magnético es todavía causa de controversia en la actualidad. Simulaciones numéricas previas sugieren que es posible que el campo magnético se comprima hasta la superficie de la estrella debido a la presión ejercida por el fluido acretante.

El estudio de los modos de oscilación de protoestrellas de neutrones procedentes del colapso gravitational de estrellas masivas se engloba dentro de un proyecto más amplio cuyo objetivo es obtener un modelo sencillo que permita obtener información sobre los parámetros físicos del progenitor de la supernova a partir de su impronta en la posible señal gravitatoria detectada. Este proyecto se puede dividir en tres pasos esenciales. El primero es estudiar los modos de oscilación de una protoestrella de neutrones a partir de los datos de una simulación numérica y comprobar la correspondencia de la distribución tiempo-frecuencia de dichos modos con el diagrama tiempo-frecuencia (espectrograma) de la señal gravitatoria obtenida de la simulación. El segundo paso consiste determinar la relación entre los parámetros físicos del progenitor y la distribución tiempo-frecuencia de los corresponientes modos de oscilación. Finalmente, el último paso consistirá en realizar un modelo sencillo que relacione los modos de oscilación con la onda gravitatoria e implementarlo en el software de inferencia Bayesiana LALInference desarrollado por la colaboración LIGO/Virgo.

La siguiente sección de la presente tesis doctoral se centra en el estudio de los detectores de ondas gravitatorias. El objetivo global consiste en conocer dicho detectores, su principio de detección y las diversas fuentes de ruido que limitan su sensibilidad. La detección y eliminación de transitorios de ruido es una prioridad dentro de la colaboración LIGO/Virgo. Este proceso, conocido como caracterización del detector, es clave para conseguir la sensibilidad necesaria para detectar señales gravitatorias, y, en el caso de se produzcan, que éstas tengan la suficiente significación estadística. Más concretamente, dentro de la colaboración

LIGO/Virgo se han puesto en marcha varios proyectos para clasificar y eliminar dichos transitorios de ruido. Así pues, los objetivos de esta parte de la tesis son colaborar en el desarrollo del código WDF-ML, diseñado por la Dr. Elena Cuoco dentro de la colaboración Virgo, y utilizar dicho código para detectar y clasificar transitorios de ruido usando datos de los detectores LIGO.

Finalmente, la última parte de la tesis está dedicada a desarrollar y testar métodos de análisis de datos y aplicarlos al contexto de la astronomía de ondas gravitatorias. El objetivo fundamental es desarrollar nuevas técnicas de eliminación de ruido que permitan mejorar las posibilidades de detección. Para ello se han estudiado dos técnicas diferentes. La primera está basada en el concepto de Variación Total. Desarrollado ampliamente en el contexto de tratamiento de imágenes, esta clase de métodos no se han utilizado hasta ahora para el tratamiento de señales gravitatorias. El objetivo de este trabajo consiste en adaptar dichas técnicas al escenario particular de la astronomía de ondas gravitatorias y comprobar su eficacia para eliminar ruido. La segunda técnica está basada en el uso de diccionarios para reconstruir una señal mezclada con ruido. Al igual que el primer método, esta técnica ha resultado ser muy eficiente cuando se ha usado con imágenes. Los objetivos principales son dos, por un lado generar diccionarios basados en señales de ondas gravitatorias de diferente tipología y, por otro lado, aplicarlos para extraer señal de un entorno ruidoso.

Metodología

Para estudiar los procesos de acreción sobre estrellas de neutrones magnetizadas se ha desarrollado un modelo 1D basado en la solución del problema de Riemann magnetohidrodinámico. Las condiciones particulares de nuestro escenario, campo magnético perpendicular a la velocidad del fluido acretante, permiten usar la descripción del problema de Riemann propuesto por Romero et al. [2005]. Además de los parámetros principales del escenario astrofísico, como son la tasa de acreción y el valor del campo magnético de la estrella, se ha estudiado la influencia de otros parámetros físicos, como la masa y la ecuación de estado de la estrella de neutrones, la composición química del fluido y la configuración geométrica del campo magnético. Con todos estos ingredientes, se ha obtenido una expresión sencilla de la relación entre tasa de acreción y el campo magnético que permite distinguir en qué casos este escenario es válido.

A día de hoy, los mecanismos físicos involucrados en la explosión de supernovas y su posterior evolución todavía no se conocen con seguridad. La información contenida en la señal de la radiación gravitatoria que producen este tipo de fuentes

es muy importante para determinar la física implicada en dichas explosiones. En esta línea de investigación, esta tesis incluye resultados iniciales de un proyecto destinado a obtener algunos de los parámetros físicos de los progenitores de supernovas a partir del estudio de la posible señal gravitatoria producida. Para ello, se ha estudiado la relación existente entre los modos de oscilación de la protoestrella de neutrones que se forma tras el colapso del núcleo de una estrella masiva y el espectro de la señal gravitatoria. Nuestro modelo aproxima las oscilaciones de la protoestrella de neutrones como perturbaciones de un sistema en equilibrio. A partir de los datos proporcionados por la simulación numérica del colapso de una estrella masiva, se ha comparado la distribución tiempo-frecuencia de los diferentes modos de oscilación con la señal gravitatoria generada por la simulación, obteniendo una correspondencia muy significativa entre ambas magnitudes físicas.

La extrema sensibilidad que requiere la medición de las minúsculas variaciones en la distancia que separa las masas prueba de un detector interferométrico al paso de una onda gravitatoria, hace que tales detectores se vean afectados por innumerables fuentes de ruido. La búsqueda de las fuentes de ruido y su posterior eliminación es, por tanto, una tarea fundamental para garantizar el éxito de la nueva astronomía de ondas gravitatorias. En este contexto, la presente tesis incluye resultados de un proyecto de colaboración en el cual se ha intentado clasificar y eliminar de manera automática los transitorios de ruido (comúnmente denominados 'glitches' en inglés) que aparecen en los detectores avanzados LIGO y Virgo. Algunos de estos ruidos transitorios son particularmente dañinos pues pueden ser muy semejantes a una señal gravitatoria real. El primer paso ha consistido en estudiar los diferentes tipos de ruido presentes en los detectores de ondas gravitatorias. El código WDF-ML es un generador de eventos (busca excesos de potencia en la trama de datos y genera un aviso cada vez que encuentra uno) basado en transformaciones wavelet. Con la lista de eventos, se aplican técnicas de aprendizaje automático para buscar correlaciones entre los eventos y dividirlos en diferentes grupos según sus características. Esta clasificación se compara con las obtenidas por otros dos métodos (desarrollados dentro del la colaboración LIGO/Virgo) demostrando que los tres son capaces de clasificar correctamente el 95% de los transitorios detectados.

A pesar de los enormes esfuerzos realizados para reducir el ruido de los detectores interferométricos, es inevitable que parte del mismo se mezcle con la señal gravitatoria. En la actualidad, existen muchos métodos destinados a extraer la señal del fondo de ruido, muchos de ellos basados en técnicas de estadística bayesiana. Durante una buena parte de esta tesis, he explorado la

posibilidades de utilizar algoritmos para la eliminación de ruido basados en el concepto de variación total. Estos algoritmos, que no requieren ninguna información a priori sobre la señal, han demostrado ser altamente eficientes para la eliminación de ruido en el contexto de tratamiento de imágenes. Nuestros resultados, aplicados de forma pionera en esta tesis en el análisis de señales de radiación gravitatoria, muestran que los algoritmos de variación total son capaces de eliminar suficiente ruido como para producir señales gravitatorias distinguibles, tanto si éstas han sido inicialmente mezcladas con ruido gausiano como si lo han sido con ruido real del detector LIGO. Uno de los aspectos más interesantes de esta línea de investigación es la posible combinación de estos métodos con otros técnicas comunes del análisis de ondas gravitatorias, pues puede conducir a mejorar notablemente los resultados. La metodología utilizada ha sido la siguiente: primero se ha modificado el algoritmo basado en variación total para adaptarlo a las características propias de las señales gravitatorias. El primer paso para testar cualquier algoritmo de eliminación de ruido es analizar sus capacidades en un entorno controlado, usando ruido simulado con un espectro de frecuencias similar al de los detectores reales, pero que no contenga transitorios de ruido, para una vez determinados los parámetros del algoritmo que producen los mejores resultados, iniciar el estudio con datos reales procedentes de los detectores LIGO.

Por último, en un intento de acortar la brecha entre el modelado numérico y el análisis de datos, he explorado también en esta tesis el uso de técnicas de aprendizaje automático de diccionarios basados en plantillas de radiación gravitatoria proporcionadas por simulaciones de relatividad numérica para reconstruir formas de onda mezcladas con ruido gaussiano. Estas técnicas ofrecen muchas posibilidades, no solo para extraer señales, sino también para clasificar transitorios de ruido o para extraer parámetros físicos de señales detectadas. Primero se han generado dos diccionarios diferentes basados en distintas tipologías de señal gravitatoria, como son señales procedentes del colapso de estrellas compactas y señales generadas por binarias de agujeros negros. Seguidamente se han estudiado diversos casos típicos para determinar en que casos la extracción de señales del fondo de ruido se realiza con éxito.

Conclusiones

Acreción sobre estrella de neutrones y el escenario de campo magnético enterrado.

Hemos estudiado el proceso de submergencia del campo magnético en una estrella de neutrones recién nacida durante la etapa de acreción hipercrítica posterior a la explosión de una supernova de colapso gravitacional. Este es uno de los posibles escenarios propuestos para explicar el campo dipolar externo aparentemente bajo de los CCOs. Nuestro enfoque se ha basado en soluciones 1D del problema de Riemann relativista, que proporcionan la ubicación de la frontera esférica (magnetopausa) que coincide con una solución de acreción externa no magnetizada con una solución de potencial de campo magnético interno. Para una masa y una intensidad de campo magnético determinadas, la magnetopausa sigue moviéndose hacia la estrella si la presión total del fluido acretante excede la presión magnética debajo de la magnetopausa. Explorando una amplia gama de masas acumuladas y las intensidades de campo, hemos encontrado las condiciones para las que la magnetopausa alcanza el punto de equilibrio por debajo de la superficie de la estrella de neutrones, lo que implica el enterramiento del campo magnético. Nuestro estudio ha considerado varios modelos con diferente entropía específica, composición, y masas de la estrellas de neutrones, y ha demostrado que estos parámetros no tienen un impacto importante sobre los resultados.

Nuestra principal conclusión ha sido que, dada la modesta masa acumulada requerida para enterrar los campos magnéticos típicos encontrados en estrellas de neutrones, el escenario CCO no es inusual. Por el contrario, la masa acumulada requerida para enterrar el campo magnético de un magnetar es tan grande, que es más probable que la estrella de neutrones colapse a un agujero negro. Sin embargo, nuestro enfoque 1D, aunque suficiente para obtener una buena aproximación al problema, no tiene en cuenta otros efectos, como la convección, que podrían modificar los resultados.

La extensión natural de nuestro estudio inicial del escenario del campo magnético enterrado implica la realización de simulaciones numéricas 2D. Sin embargo, esto no es una tarea fácil debido a una serie de razones: (a) la gran diferencia entre los valores de la presión magnética y la presión térmica en algunos de los regímenes extremos de este escenario, puede dar lugar a inexactitudes numéricas o incluso a la terminación normal de la ejecución del código; (b) el acoplamiento entre la magnetosfera y el fluido caliente debe ser manejado con cuidado para evitar posibles efectos de sobrecalentamiento. Aún así, las

simulaciones 2D son esenciales para mejorar nuestra comprensión del proceso de enterramiento del campo magnético tras una explosión de supernova.

Oscilaciones lineales del espectro de protoestrellas de neutrones.

El estudio de la física involucrada en el colapso del núcleo de estrellas masivas y la posterior emisión de ondas gravitatorias no es bien conocida. Además, el modelado numérico de este tipo de sistemas es computacionalmente muy costoso; por ejemplo, las simulaciones en 3D pueden llevar varios meses en los actuales superordenadores. Por lo tanto, no es posible inferir la física de las posibles señales gravitatorias detectadas, como se hace en el caso de las señales producidas por binarias de agujeros negros. Estos últimos son sistemas significativamente más simples que las supernovas de colapso gravitacional y en su análisis pueden emplearse múltiples formas de onda basadas en relatividad general (usando enfoques postnewtonianos) e incluso la relatividad numérica para la fase de fusión, para hacer coincidir el filtro con formas de onda gravitacionales reales. En esta sección de la tesis nuestra meta ha sido diseñar un modelo sencillo pero robusto, basado en la teoría de los modos normales de oscilación de un sistema esférico, que permita inferir parámetros físicos del progenitor a partir de una señal de colapso gravitacional dada.

Como primer paso hemos determinado si los modos de oscilación de la protoestrella de neutrones pueden estar relacionados con el espectro de ondas gravitatorias del progenitor. Para ello, se ha utilizado un modelo 1D para obtener la función propia de los modos de oscilación a partir de los datos de una simulación numérica existente y de vanguardia de este tipo de escenario. Una vez identificado el origen de los modos, hemos separado los distintos tipos. Nuestros resultados muestran que existe una clara correspondencia entre el espectro de los modos y el espectro de las ondas gravitatorias. Por lo tanto, parece posible analizar una señal gravitatoria procedente del colapso del núcleo de estrellas masivas en términos de los modos de oscilación de la protoestrella de neutrones.

Habiendo cumplido esta necesaria prueba de concepto, planeamos dar el siguiente paso en un futuro próximo. La idea es realizar varias simulaciones numéricas unidimensionales cambiando los parámetros del progenitor para estudiar la dependencia de la distribución tiempo-frecuencia (y posiblemente relajando la aproximación de Cowling empleada en el trabajo actual). El objetivo final de este estudio es desarrollar un modelo que relacione los parámetros del progenitor del colapso con los correspondientes modos de oscilación de la protoestrella de neutrones (y, por lo tanto, con el espectro de ondas gravitatorias) que nos

permita inferir los parámetros directamente del espectro de ondas gravitatorias sin necesidad de realizar costosas simulaciones numéricas.

Clasificación de transitorios de ruido.

La presencia de transitorios de ruido (glitches) en los canales de señal de los detectores avanzados LIGO/Virgo es un efecto inevitablemente ligado a la extrema sensibilidad requerida para detectar este tipo de señales producidas a distancias cosmológicas. Existen dos estrategias posibles para eliminarlas, o al menos limitar su impacto. Una de ellas es tratar de determinar el origen y tomar las medidas necesarias para resolver el problema que causa el transitorio de ruido. Si esto no es posible o no se puede determinar el origen de los mismos, la estrategia consiste en vetar los datos y no utilizarlos para ciencia. Las técnicas automáticas de clasificación de transitorios de ruido que se presentan en esta tesis pueden contribuir a mejorar ambas estrategias. Por un lado, pueden relacionar la información proporcionada por los canales auxiliares de detección con los datos del canal de señal gravitatoria y ayudar a determinar la causa de un tipo determinado de transitorio. Por otro lado, la determinación del tipo de glitch automáticamente puede ayudar a los algoritmos de veto.

Nuestros resultados han demostrado que los tres métodos numéricos empleados pueden clasificar los transitorios de ruido en datos de LIGO avanzado con una precisión del 95%. Además, el uso de tres métodos sobre los mismos datos es una estrategia ganadora ya que los glitches no detectados por uno de los métodos pueden ser clasificados por los otros, aumentando la eficiencia general. Además, el método WDF-ML incluye un generador de eventos diferente al utilizado en los detectores LIGO/Virgo y, por lo tanto, puede utilizarse para comprobar los eventos generados por este último. Nuestros tres métodos son ejemplos de cómo los métodos de aprendizaje automático pueden aplicarse a la astronomía de ondas gravitatorias. Desde su uso en los subsistemas de detector y diseño de hardware hasta su uso en el ámbito del análisis de datos, existen una gran variedad de métodos de aprendizaje automático para mejorar las posibilidades de detección de señales de ondas gravitatorias.

Hay algunas mejoras interesantes que planeamos aplicar a los métodos de clasificación de transitorios. Para empezar, el procedimiento de blanqueamiento de los datos realizado por el algoritmo WDF-ML se mejorará usando una técnica conocida como blanqueamiento adaptativo [Cuoco, Cella, and Guidi 2004]. También se pueden realizar mejoras mediante el uso de un conjunto de formas de onda preclasificadas como entrenamiento o explorando el uso de algoritmos de aprendizaje con diccionarios, como el presentado en esta tesis, para el problema

específico de clasificación de glitches. También vale la pena mencionar el proyecto Gravity Spy, financiado por la NSF, que tiene como objetivo la creación de grupos de formas de onda clasificadas a través de un programa de ciudadano-ciencia (Zooniverse) [Zevin et al. 2016, Simpson, Page, and De Roure 2014]. El potencial de las clasificaciones de glitches se maximiza si los conjuntos de datos pueden ser empleados en tiempo real. Para lograr este objetivo, la eficiencia computacional de nuestros algoritmos tendrá que mejorar. El proyecto para construir un código más eficiente, añadiendo capacidades de paralelización e incluso capaz de funcionar con GPUs, ha comenzado en el momento de redactar esta tesis y continuará en un futuro cercano en estrechas colaboraciones con la Dra. Elena Cuoco (EGO) Y Massimiliano Razzano (Universidad de Pisa).

Métodos TV para la eliminación de ruido en ondas gravitatorias.

Nuestro estudio de los métodos de variación total en el contexto de la astronomía de ondas gravitatorias, introducido en esta tesis, comienza realizando pruebas en condiciones de ruido sencillas, como es el caso del ruido gaussiano. Este paso inicial ha sido necesario para entender el comportamiento de estos métodos y explorar su espacio de parámetros y determinar los valores que producen los mejores resultados. Este estudio inicial ha conducido a resultados interesantes. El más importante es que los dos algoritmos que hemos utilizado para resolver el problema ROF son capaces de eliminar el ruido con éxito y recuperar una señal gravitatoria reconocible. Hemos encontrado que la elección del parámetro de regularización λ es crucial para obtener resultados adecuados. Por lo tanto, es importante idear una estrategia para hallar el intervalo de los valores adecuado de λ. Como la varianza del ruido en un detector de ondas gravitatorias es desconocida, no es posible utilizar el principio de discrepancia para determinar el parámetro de regularización. Como resultado, realizar una búsqueda heurística basada en algún estimador de calidad (como el PSNR o el SSIM) parece ser la mejor solución.

El caso gaussiano nos ha permitido comprender el efecto que la elección de λ tiene en las diferentes escalas del problema variacional. Se ha hecho necesario utilizar diferentes valores de λ para recuperar diferentes partes de la señal descritas por diferentes escalas (frecuencial o temporal). En el caso de las señales de binarias de agujeros negros, este efecto es más evidente que para las señales de colapso gravitacional, ya que son significativamente más largas. En particular, hemos visto que la parte de bajas frecuencias de la señal se puede recuperar con mayor precisión utilizando un valor de λ menor que el requerido para recuperar la parte correspondiente a la fusión y la oscilación del agujero negro resultante.

Con las lecciones aprendidas con el caso gaussiano, hemos considerado un escenario realista, utilizando datos reales del detector LIGO en su configuración inicial. Hemos empleado solamente el método rROF, porque es el método que da los mejores resultados en el dominio del tiempo, evitando los problemas con los bordes que aparecen generalmente en el dominio de la frecuencia. También hemos modificado el algoritmo para utilizarlo iterativamente. Este cambio produce una dependencia más suave con λ, aumentando el intervalo de valores apropiados. Sin embargo, el coste computacional también aumenta ya que el algoritmo tiene que ser ejecutado más veces. También hemos observado que la componente de baja frecuencia del ruido presente en los detectores es difícil de eliminar utilizando sólo los métodos de variación total. Esto se relaciona con el hecho de que las propiedades de eliminación de ruido de estos métodos dependen de la escala. Como los algoritmos rROF reducen la variación total de la señal de entrada, las frecuencias más altas se eliminarán más rápido que las inferiores. Sin embargo, se puede usar rROF con un valor bajo de λ para aislar las componentes de baja frecuencia y luego eliminarlas. Además, hemos combinado el algoritmo rROF con dos métodos adicionales, un preprocesamiento basado en un filtrado sencillo y un método de blanqueamiento. Nuestro estudio ha demostrado que los métodos basados en variación total pueden mejorar los resultados de ambos algoritmos, conduciendo a señales identificables. Incluso con el ruido real del detector, los métodos de variación total son una herramienta interesante que se puede combinar con otros enfoques comunes dentro del contexto del análisis de datos de ondas gravitatorias.

Los algoritmos para eliminar el ruido de una señal gravitatoria presentados en esta tesis constituyen una línea de investigación muy prometedora para continuar en un futuro próximo. Hasta ahora, sólo hemos explorado dos algoritmos basados en la liberalización de la ecuación de Euler-Lagrange para resolver el problema ROF. Sería interesante implementar el algoritmo dual-primal y comparar los tres métodos en las mismas condiciones, tratando de determinar cuál es el mejor algoritmo para usar en cada caso. También vale la pena explorar aún más la relación del parámetro de regularización λ con las diferentes escalas, con el fin de diseñar un algoritmo con un λ ajustable que produzca los mejores resultados para cada condición de ruido. En este contexto, el algoritmo basado en parches como non-local means [Buades, Coll, and Morel 2011] puede ser una herramienta muy útil. Además, para comprobar las posibles mejoras introducidas al utilizar un método de eliminación de ruido basado en variación total como paso previo a los algoritmos de detección de señales y estimación de parámetros, se hace necesario realizar un estudio combinado. Por último, también se mencionan

los algoritmos basados en grafos, que han recibido gran atención en los últimos años. Aplicarlos a los datos de ondas gravitatorias puede ser un proyecto muy interesante.

Métodos de eliminación de ruido en ondas gravitatorias basadas en diccionarios.

También hemos estudiado las capacidades del aprendizaje basado en diccionarios para recuperar señales de ondas gravitatorias de un fondo dominado por el ruido. Nuestro algoritmo LASSO ha sido probado usando señales de dos fuentes principales, estallidos de ondas gravitatorias generados por el colapso del núcleo de estrellas en rotación y chirps procedentes de la fusión de dos agujeros negros. Para obtener los respectivos diccionarios, hemos utilizado el 80% de las formas de onda para realizar el entrenamiento, el 15% para la validación, es decir, para obtener el mejor conjunto de parámetros que produce los mejores resultados, y el último 5% de las formas de onda para evaluar el método. Una característica interesante de LASSO es que, para la mayoría de las realizaciones de ruido gaussianas que hemos considerado, devuelve cero si la señal de entrada no puede ser reconstruida por los elementos del diccionario. Como resultado, el método puede proporcionar una reconstrucción de señal bastante clara. Por otra parte, una limitación intrínseca del método es que los resultados dependen fuertemente de la selección del parámetro de regularización λ, cuyo valor óptimo no puede establecerse a priori, y debe ser estimado mediante estudios de validación.

Existe una gran variedad de técnicas de aprendizaje automático en la literatura. En esta tesis sólo hemos considerado una de ellas, pero en un futuro próximo planeamos implementar métodos adicionales para realizar los cálculos del algoritmo LASSO de manera más eficiente. El cálculo de la solución para un vector de 256 muestras tarda típicamente unas pocas decenas de ms en un ordenador Apple iMac con procesador Intel Core i7 y 16 Gb de Ram. El mayor coste computacional se asocia con la tarea de aprendizaje. Reducir el tiempo necesario en esta parte del método es una cuestión clave para aplicar el método en tiempo real a los datos reales generados por los detectores. En los próximos meses, las versiones avanzadas de LIGO y Virgo iniciarán de nuevo las observaciones con una sensibilidad mejorada, aumentando el número de detecciones. El desarrollo de técnicas de análisis de datos sofisticadas que mejoren las oportunidades de detección, especialmente para eventos con baja relación señal a ruido, es, por tanto, un esfuerzo crucial.

Aplazamos para un estudio futuro el análisis de la tasa de falsa alarma con glitches simulados (o reales) usando métodos de aprendizaje basados en

diccionarios. Creemos que este análisis, junto con los estudios de validación utilizando ruido real, son obligatorios antes de usar diccionarios en los algoritmos de un detector. Para evitar los falsos positivos causados por transitorios de ruido, es posible emplear diccionarios construidos a partir de una colección de transitorios conocidos, lo que permitiría utilizar LASSO como un clasificador de transitorios de ruido. Por otra parte, como todos los métodos discutidos en este trabajo se han desarrollado originalmente para la reconstrucción de imágenes (datos 2D), tenemos la intención de aplicarlos en espectrogramas de ondas gravitatorias.

Publications

The following peer-reviewed publications and proceedings reports have been published during the duration of the doctoral fellowship:

- Powell, J., **Torres-Forné, A.**, Lynch, R., Trifirò, D., Cuoco, E., Cavaglià, M., Heng, I.S., and Font, J.A., "Classification methods for noise transients in advanced gravitational-wave detectors II: performance tests on Advanced LIGO data" , Classical and Quantum Gravity, **34**, 034002 (2017).

- **Torres-Forné**, A., Marquina, A., Font, J. A., and Ibáñez, J. M. "Denoising of gravitational wave signals via dictionary learning algorithms". Physical Review D, **94**, 124040 (2016).

- **Torres-Forné**, A., Cerdá-Durán, P., Pons, J. A., and Font, J. A. "Are pulsars born with a hidden magnetic field?" Monthly Notices of the Royal Astronomical Society, **456**(4), 3813-3826 (2016).

- **Torres, A.**, Cerdá-Durán, P., and Font, J. A."Fallback accretion onto magnetized neutron stars and the hidden magnetic field model". In Journal of Physics: Conference Series (Vol. 600, No. 1, p. 012057). IOP Publishing (2015).

- **Torres, A.**, Marquina, A., Font, J. A., and Ibáñez J. M. "Split Bregman Method for Gravitational Wave Denoising". In Gravitational Wave Astrophysics (pp. 289-294). Springer International Publishing (2015).

- **Torres**, A., Marquina, A., Font, J. A., and Ibáñez, J. M., "Total-variation-based methods for gravitational wave denoising", Physical Review D, **90**, 084029 (2014).

Contents

List of Figures

List of Tables

Nomenclature

Acronyms / Abbreviations

BH	black hole
BBH	binary black hole
BHNS	black hole–neutron star
BNS	binary neutron star
CBC	compact binary coalescence
CCO	Central compact objects
EOB	Effective-one-body formalism
EoS	Equation of state
GRB	gamma ray burst
GR	General Relativity
GRHD	general relativistic hydrodynamics
GRMHD	general relativistic magnetohydrodynamics
GW	gravitational wave
LIGO	Laser Interferometer Gravitational-Wave Observatory
NR	numerical relativity
NS	neutron star
PNS	Protoneutron star
ROF	Rudin-Osher-Fatemi

SB	Split-Bregman
SNR	Signal-to-noise ratio
TV	Total variation

Part I

Introduction

Chapter 1

Introduction

During the course of the investigation that has led to the present thesis, I have worked on three main lines of research, namely (a) Numerical modelling of astrophysical systems such as neutron stars (NS), (b) classification and characterization of instrumental noise in gravitational-wave (GW) detectors, and (c) GW data analysis with techniques that had not been employed within such a context before. My research has been carried out within the Relativistic Astrophysics Group of the University of Valencia (UV), a group which plays a significant role in the field of numerical relativity, with relevant contributions regarding formulations of the equations of relativistic hydrodynamics (GRHD) [Martí, Ibáñez, and Miralles 1991, Banyuls et al. 1997] and magneto-hydrodynamics (GRMHD) [Antón et al. 2006] as well as simulating different scenarios of relativistic astrophysics. Given the group's background, the natural starting point of this thesis involved numerical modelling of astrophysical systems using codes and methods developed by the group (and elsewhere) throughout the years.

Halfway through my doctoral studies, however, the first direct detections ever of GWs by the Laser Interferometer Gravitational-Wave Observatory (LIGO) interferometers, took place [Abbott et al. 2016b, Abbott et al. 2016d]. These exceptional discoveries affected the research plans of this thesis which had to be adjusted to make room for the exciting new prospects of analyzing actual GW signals from real astrophysical systems. Part of the Relativistic Astrophysics Group at UV – me included – joined the Virgo Collaboration on the 1st of July 2016, eager to play an active role within the growing GW international community. By joining the Virgo Collaboration, I initiated new research projects with a number of coworkers from the LIGO Scientific Collaboration and the Virgo Collaboration. Part of the work done within those collaborations on

the topic of detector characterization is discussed in the present thesis. Other projects are still ongoing at the time of writing, but the results will be published in the near future.

The detection of GWs also boosted the research I had started doing (pre-detection) on data-analysis techniques based on Total-Variation (TV) methods. These algorithms, which do not need any a priori information about signals, have been originally developed and fully tested in the context of image processing. We adapted them and pioneered their use in the specific context of GW astronomy, applying them, in particular, to denoise two paradigmatic examples of numerically-simulated GW signals, namely *bursts* produced from the core collapse of rotating stars and waveforms from binary-black-hole mergers. These techniques can be combined with other standard techniques in data analysis and may be potentially implemented in a detector's pipeline.

Further work I carried out on the GW data-analysis front had to do with machine-learning algorithms, which have experienced a huge development in recent years. Machine-learning techniques are driving a revolution in data-analysis science, particularly in the context of massive datasets (i.e. Big Data). In this line of research, we have developed and tested a specific method based on learned dictionaries. The dictionaries are built from so-called "atoms" obtained out of numerical relativity waveform templates, and, after a learning step, they can nicely dig out a signal buried in a noisy background.

The link between numerical relativity and GW astronomy is obviously strong. As new detections arrive, especially involving astrophysical sources containing matter, more simulations will be required to infer the physical parameters of the progenitor sources. This thesis also discusses the first results of a project on parameter estimation I have recently started and which relates both topics. In particular, the project aims at inferring physical parameters of a progenitor star undergoing rotational core-collapse by studying the effects of the oscillating modes of the resulting proto-neutron star (PNS) on the GW spectrum.

I turn next to provide further content on these three main lines of research carried out in this work: numerical modelling, detection, and data analysis.

1.1 Numerical modelling

Regarding numerical modelling, I have worked on two topics, (a) matter accretion onto magnetized NSs, and (b) oscillation modes of PNSs formed dynamically following the gravitational collapse of a rotating progenitor.

In the former topic, my investigation has focused on the study of the so-called hidden magnetic-field scenario [Young and Chanmugam 1995], an intriguing situation in which the magnetic field of a PNS might be temporarily buried as a result of accretion of matter behind the supernova shock. This problem was investigated using axisymmetric numerical simulations with the `CoCoNuT` code. However, the simulations were hampered by a number of numerical issues not yet entirely addressed. The complexity of the simulations has to do with the coupling of the supersonically infalling unmagnetized fluid with the force-free magnetosphere and with the high-density magnetized region describing the NS. Due to the numerical difficulties encountered, we had to make do with simulations of the much simpler spherically symmetric case. This situation, however, proved sufficient to obtain the limiting values of the accretion rates and magnetic-field strengths that characterize the hidden magnetic-field scenario. Our approach included as many physical ingredients as possible, namely different realistic equations of state and initial models with varying magnetic-field configurations and neutron-star masses.

The second topic on numerical modelling has dealt with the study of the oscillation spectra of PNSs. For this work I have used data from the simulation performed by Cerdá-Durán et al. [2013] which explored the dynamics and GW signature of the collapse of a low-metallicity rapidly-rotating progenitor to a PNS, the subsequent long phase of accretion that ensues, and the final formation of a rotating black hole (BH). This project has tried to identify and classify the different oscillating modes resulting from the linear-perturbation analysis. The time-frequency distribution of those modes has been compared with the spectrogram of the corresponding GW signal obtained from the numerical simulation. A very clear and revealing correspondence between the two types of modes has been found, which may provide a procedure to infer progenitor parameters out of the GW spectrum. This is a stimulating possibility which will be fully investigated in the near future.

1.2 Gravitational-wave detection and detector characterization

The epoch-making detections of the transient GW signals GW150914 and GW151226 during the first observing run of the two Advanced LIGO (aLIGO) interferometers [Abbott et al. 2016b, Abbott et al. 2016d] has marked the start of GW astronomy. GW150914, detected with unexpectedly high signal-to-noise

(SNR) ratio (SNR∼24) and with a statistical significance greater than 5.1σ, is in excellent agreement with numerical-relativity waveforms [Mroué et al. 2013, Campanelli et al. 2006] for the final few cycles (chirp), merger (burst) and subsequent ringdown of the coalescence of two stellar-origin BHs in a binary system. GW151226, also the result of a binary-black-hole merger, was recovered with similar statistical significance but with a SNR∼13. Its initial black-hole masses, $14.2M_{\odot}$ and $7.5M_{\odot}$, are lower than in the case of GW150914, $36M_{\odot}$ and $29M_{\odot}$ [1]. As a result, GW151226 spent almost 1 s in the LIGO frequency band, increasing in frequency and amplitude from 35 to 450 Hz over about 55 cycles. Contrary to GW150914, matched filtering with waveform templates from general relativity was essential to detect GW151226, jointly with unmodeled searches, due to the smaller strain amplitude and the longer time interval [Abbott et al. 2016d].

At present, the two aLIGO interferometers are well within the second observing run (O2), which is expected to finish towards the end of August 2017. At the same time, the commissioning of the European detector Advanced Virgo [Accadia et al. 2012] is well underway, aiming at start observing in the second half of 2017, while the Japanese detector KAGRA [Aso et al. 2013] is still under construction. Simultaneous observational campaigns of these four detectors, five with the later addition of the recently-approved LIGO India, will also increase considerably the rate of detections along with their statistical significance and the accuracy of the sky location of each event [Abbott et al. 2016b].

The aLIGO detectors are two 4 km interferometers at Hanford, Washington (H1) and Livingston, Louisiana (L1) [Harry and Collaboration 2010, Aasi et al. 2015b]. The European 3 km advanced interferometer Virgo is located in Cascina, Italy [Acernese et al. 2008]. The duty cycle of the detectors and their sensitivity to astrophysical signals is determined by noise sources caused by the instruments themselves and by their environment. In particular, as the detector noise is non-Gaussian and non-stationary, short-duration transients significantly limit the sensitivity of searches for transient astrophysical sources such as compact binary coalescences [Flanagan and Hughes 1998]. GW detectors contain many environmental and instrumental sensors, which produce auxiliary channels of data to monitor the detector behavior and track the causes of short-duration noise artifacts. Auxiliary channels that are not sensitive to gravitational waves can be used to identify noise transients, also known as “glitches”, in the detector

[1] There is a third candidate event, labelled LVT151012 [Abbott et al. 2016a], which corresponds to the merger of a BBH with a SNR ~ 9.7. However, its low statistical significance (1.7σ) does not to allow to claim it as a detection.

output and veto those events [Smith et al. 2011, Ajith et al. 2014, Aasi et al. 2015a]. As the network of advanced detectors approaches design sensitivity, the number of detections is expected to increase. Adding further detectors to the network will increase the number of possible sources of noise and the time involved to identify their origin. Transients which occur in any one detector will limit the joint analysis time for the network. Understanding the sources of noise transients in the detectors will become increasingly more important as the sensitivity of the detectors increase.

Regarding this topic of research, the thesis presents results on the development and testing of three methods to automatically detect and classify glitches in the data of the 7th engineering run (ER7) of advanced LIGO. This has two interesting applications. Firstly, the automatic classification of known glitches can be combined with veto techniques to remove corrupted segments of data, hence reducing the noise background. Secondly, classification methods may discard GW signal candidates as actual glitches, hence increasing the significance of the detections.

1.3 Gravitational-wave data analysis

Despite the recent discoveries, noise removal remains one of the most challenging problems in GW data analysis. There exist a number of noise sources that limit the possibilities of detection [Martynov et al. 2016]. The most limiting source of noise for frequencies below a few tens of Hz is gravity-gradient noise. Thermal noise due to Brownian motion is dominant at intermediate frequencies, while shot noise, produced by quantum fluctuations of the laser, becomes prominent at frequencies above ~ 150 Hz, difficulting detection above 2 kHz. Nevertheless, searches for GW bursts up to frequencies of 5 kHz have been performed [Abadie et al. 2012]. To add more complexity, transient spurious noise signals (glitches) due to instrumental or environmental sources, may potentially disturb astrophysical signals. Glitches might mimic GW signals increasing the false-alarm rate and producing a decrease in the detectors' duty cycles. A huge effort in commissioning and detector characterization [Christensen et al. 2010] has been done to reduce the effect of glitches. Improving glitch identification and classification [Powell et al. 2015, Powell et al. 2017] would improve detection efficiency but there will always be a chance for false positives in the detectors.

GW interferometers are designed to be sensitive to waveforms produced by different astrophysical mechanisms. Sources can be separated in groups depending on how well-known and modeled their waveforms are. Specific

data-analysis techniques have been developed for each type of signal (for a review see Jaranowski and Kròlak [2005] and references therein). Transient gavitational-wave signals from compact binary coalescence (CBC), either from binary neutron stars (BNS) or binary black holes (BBH), are well studied and the corresponding waveforms can be calculated with high accuracy. These systems are nowadays modeled using the effective-one-body formalism (EOB) [Buonanno and Damour 2000] which combines post-Newtonian methods [Blanchet, Damour, and Iyer 1995, Blanchet et al. 2004] with numerical relativity and perturbation theory [Nagar and Rezzolla 2005]. This technique allows to generate template banks efficiently. This is the main reason matched-filtering is the most common method for CBC detection [Owen and Sathyaprakash 1999, Sathyaprakash and Schutz 2009, Abbott et al. 2016a] in which filters correlate signals with templates. A trigger associated with a specific template is generated when the filter output excess a certain threshold. In addition to EOB waveforms, other waveform families are needed to cover as much parameter space as possible (see Abbott et al. [2016c] and references therein).

Matched-filtering becomes however unpractical for well-modeled but continuous sources, like spinning NS, due to the large computational resources it would require. Nevertheless, as such signals are very stable and have long duration, a coherent integration can be performed. In addition, the data from all detectors can be compared, which increases the SNR of this type of events. Roughly speaking there are two main methods to fulfill this comparison, cross-correlation and coherent methods [Bose, Pai, and Dhurandhar 2000]. The former directly compare the data streams from a pair of detectors to search for a common signal within uncorrelated noise while the latter generalize the concepts of excess power and cross-correlation to take full advantage of having three or more data streams. The duration and the sky coverage (all-sky or targeted) [Palomba 2012] can vary depending of the type of source which is sought for.

In contrast with CBC and continuos sources, the non-spherical gravitational collapse of massive stars produces a short ($\sim$ms) duration (prompt) signal (but see e.g. Cerdá-Durán et al. [2013] for the case of collapsars where the duration of the signal is dominated by the accretion timescale, considerably longer) with a significant power in the kHz frequency band. In addition to core-collapse supernova, other astrophysical sources as cosmic string cusps [Damour and Vilenkin 2005], NS glitches [Van Eysden and Melatos 2008, Sidery, Passamonti, and Andersson 2010], magnetar flares [Ciolfi et al. 2011, Zink, Lasky, and Kokkotas 2012], BBH, BNS, and NSBH mergers, and BH accretion disks [Mewes et al. 2016] can also produce GW transients or "bursts". Such signals, in particular

core-collapse bursts, can only be modeled imperfectly, and the computational requirements for obtaining the corresponding waveforms from numerical relativity simulations and the parameter space are much larger than in the case of binary systems. Therefore, a bank of templates cannot be built with sufficient accuracy to meet the requirements of matched-filtering. For burst signals, the time-frequency analysis of the signal in all the detectors, related to each other with cross-correlation and coherent methods, is the best option. With initial detectors, several complete all-sky, all-time burst searches were performed [Abadie et al. 2012] and another one has also been carried out during advanced LIGO's first observing run (O1) [Abbott et al. 2016c]. These techniques, used in tandem with electromagnetic observations, can increase the possibilities of identifying a GW burst [Abbott et al. 2009].

The detection confidence of unmodeled astrophysical sources has significantly improved in recent years. In particular, coherent approaches over a network of GW detectors have proven to be very effective [Thrane and Coughlin 2015, Klimenko et al. 2016], increasing the detection confidence of long-duration (above several seconds) burst signals which are less sensitive to the presence of most noise transients. In contrast, short-duration bursts are more affected by detector glitches and specific pipelines based on Bayesian inference have been developed to differentiate between signals and noise transients, namely `BayesWave` [Littenberg et al. 2016], either alone or in combination with `coherentWaveBurst` [Kanner et al. 2016], and `oLIB` [Lynch et al. 2015]. Other approaches, like those of Röver et al. [2009] and Engels, Frey, and Ott [2014], have proven to be effective for estimating physical parameters and for the reconstruction of burst signal waveforms from (Gaussian) noisy environments.

In this thesis we deviate from the standard approaches used in the GW data analysis community by assessing a method based on TV-norm algorithms [Rudin, Osher, and Fatemi 1992] for denoising and detection of gravitational waves. This assessment is carried out both, for signals embedded in additive Gaussian noise as for signals embedded in actual detector noise, namely that of LIGO's S6 run in its initial configuration. Arguably, the main advantage of this technique is that no a priori information about the astrophysical source or about the signal morphology is required to perform the denoising. This feature allows us to obtain satisfactory results for data from two different numerically-relativity-generated catalogs of gravitational waveforms comprising signals with very different structure. Moreover, we also adapt dictionary-learning algorithms, which have been extensively developed in the last few years and successfully

applied in the context of image processing, to GW data. By building dictionaries from numerical relativity templates of both, BBH mergers and bursts of rotational core collapse, we show how machine-learning algorithms based on dictionaries can be also successfully applied for GW denoising.

1.4 Organization of the thesis

The rest of the thesis is organized in five parts:

Part II discusses numerical modelling and contains three chapters. Chapter 2 presents a brief summary of the general relativistic MHD equations and the Riemann problem, with emphasis on the specific configuration of the magnetic field used in Chapter 3, which presents the results of the one-dimensional parametric study of accretion onto NSs and the hidden-magnetic-field scenario. This part closes with Chapter 4 which deals with the study of the oscillation spectra of PNS.

In Part III we focus on detector characterization. The first chapter in this part, Chapter 5, introduces the problem of noise in GW detectors, describing the main sources and gives some examples of the noise transients (glitches) that limit the detection. Chapter 6 reports the results of the work done in the glitch-classification project.

Part IV discusses the data-analysis techniques we have developed during this thesis along with the results accomplished. We present a brief theoretical introduction to these methods and algorithms in Chapter 7, focusing on the most important methods developed to denoise images and signals. The results of the application of TV methods to both Gaussian noise and real noise from LIGO data are presented in Chapter 8 and Chapter 9, respectively. Finally, the last chapter of this part, Chapter 10, presents the results of applying dictionary-learning algorithms to extract GW signals embedded in Gaussian noise.

A summary of the results reported in the thesis and the outlook to possible future extensions of this work are presented in Part V.

Finally, Part VI contains two appendices. On the one hand, Appendix A describes the specific formulation of the Discrete Fourier Transform (DFT) we employ in this work and the method we use to generate Gaussian noise. On the other hand, Appendix B gives a succinct description of the two GW catalogs we use to perform the tests in the data-analysis part.

1.5 Conventions

Throughout this thesis we use a spacelike metric signature (-,+,+,+) and $c = G = 1$ (geometrized) units, where c stands for the speed of light and G is Newton's gravitational constant. We absorb the factor $1/\sqrt{4\pi}$ appearing in the GRMHD equations in the definition of the magnetic field B^i, i.e. the units of the magnetic field are $\sqrt{4\pi}$ Gauss. This system of units is known as Heaviside-Lorentz [Jackson 1962]. As customary, Greek indices run from 0 to 3, Latin indices from 1 to 3, and we use Einstein's summation convention for repeated indices. Vector and tensor variables are indicated in boldface.

Part II

Numerical modelling

Chapter 2

Numerical relativity and MHD

General Relativity (GR) is the theory of gravity that best describes the strong-field scenarios distinctive of the central engines responsible of the most energetic processes observed in the Universe, either in high-energy astrophysics (involving electromagnetic radiation) or in the newborn GW astronomy. High-energy radiation from NSs and accreting BHs, the production of relativistic radio jets in active galactic nuclei, the kHz quasi-periodic oscillations in low-mass X-ray binaries, the coalescence of binary systems of NSs and BHs, and events related with the gravitational collapse of massive stars, including the generation of gamma-ray bursts (GRB) and the emission of gravitational waves, are examples of present-day research in astrophysics where GR is required.

In general, understanding the complex dynamics of those scenarios is beyond the scope of analytic approaches (except when idealised physical conditions and unrealistic symmetries are assumed) and demands multidimensional, time-dependent, and highly accurate numerical simulations. Numerical relativity (NR) simulations in particular[1] play a fundamental role in the field of GW astronomy. The study of the astrophysical scenarios where detectable gravitational radiation is generated, either in vacuum (as in the the binary BH hole problem) or in those which involve matter, badly need simulated gravitational waveforms to infer the underlying physics and dynamics.

[1] Numerical relativity is the field of research of GR devoted to solving numerically Einstein's equations for the dynamics of the gravitational field on supercomputers [Alcubierre 2008, Baumgarte and Shapiro 2010, Shibata 2015].

The goal of is this chapter is to present a brief overview of the NR equations and numerical methods which are specifically employed for the simulations presented in this work. The field of NR is much too wide to be properly covered in this manuscript. The interested reader is addressed to recent text-books and review articles for further details [Alcubierre 2008, Font 2008, Baumgarte and Shapiro 2010, Rezzolla and Zanotti 2013, Shibata 2015].

2.1 General relativistic magneto-hydrodynamics

In relativistic astrophysics, there are situations where, along with the gravitational field, the magnetic field also plays an important role in the dynamics and may affect the evolution of the system. For example, the internal structure of NS stars is determined by their intense magnetic field, of the order of $10^{12} - 10^{13}$ G or even larger, $\sim 10^{14} - 10^{15}$ G in the case of magnetars. In such a circumstance, the dynamics of the system is described by a coupled system of time-dependent, partial differential equations, comprising the general relativistic MHD equations and the Einstein gravitational field equations. During the last 20 years, many efficient numerical implementations of those equations have been developed, yielding considerable progress in the field of NR (see in particular Font [2008] and references therein for an overview of state-of-the-art approaches to solve the GRMHD equations).

We adopt the 3+1 formalism [Lichnerowicz 1944] to foliate the spacetime into a set of spacelike hypersurfaces. In this formulation, the line element can be written as,

$$ds^2 = -(\alpha^2 - \beta_i\beta^i)dt^2 + 2\beta_i dx^i dt + \gamma_{ij} dx^i dx^j, \tag{2.1}$$

where α (lapse function), β^i (shift vector) and γ_{ij} (spatial metric) are functions of the coordinates t, x^i. The *Eulerian observer* is the one whose four-velocity $\boldsymbol{n}$,

$$n^\mu = \frac{1}{\alpha}(1, -\beta^i), \tag{2.2}$$

is perpendicular to the hypersurfaces of constant coordinate time t.

When matter is present in the spacetime, another observer can be naturally introduced, the so-called *comoving observer*. This observer moves with the fluid with four-velocity $\boldsymbol{u}$. The three-velocity of the fluid measured by an Eulerian observer reads,

$$v^i = \frac{u^i}{\alpha u^t} + \frac{\beta^i}{\alpha}\ , \tag{2.3}$$

while $v_i = u_i/W$ and $W = \alpha u^0$ is the Lorentz factor.

The electromagnetic field in GR is described by the (Faraday) electromagnetic tensor $F^{\mu\nu}$. It is worth to define the dual of the electromagnetic tensor ${}^*F^{\mu\nu}$

$$ {}^*F^{\mu\nu} = \frac{1}{2}\eta^{\mu\nu\lambda\delta}F_{\lambda\delta}, \tag{2.4} $$

with $\eta^{\mu\nu\lambda\delta}$ being the volume element. The Maxwell equations can be written in terms of the electromagnetic tensor as,

$$ \nabla_\nu {}^*F^{\mu\nu} = 0, \tag{2.5} $$

$$ \nabla_\nu F^{\mu\nu} = 4\pi\mathcal{J}^\mu, \tag{2.6} $$

where ∇_ν stands for the covariant derivative and $\mathcal{J}^\mu$ is the electric four-current. According to Ohm's law, the latter can be in general expressed as

$$ \mathcal{J}^\mu = \rho_q u^\mu + \sigma F^{\mu\nu}u_\nu, \tag{2.7} $$

where ρ_q is the proper charge density measured by the comoving observer and σ is the electric conductivity.

In the so-called ideal MHD limit, the conductivity of the fluid is infinite (perfect conductor) and the temporal component of the electric field vanishes for an Eulerian observer, $E^\mu = (0, -\epsilon_{ijk}v^j B^k)$, where ϵ_{ijk} is the permutation tensor and B^k is the magnetic field. In this case, Maxwell's equations reduce to the divergence-free condition plus the induction equation for the evolution of the magnetic field.

$$ \vec{\nabla}\cdot\vec{B} = 0 \tag{2.8} $$

$$ \frac{1}{\sqrt{\gamma}}\frac{\partial}{\partial t}\left(\sqrt{\gamma}\vec{B}\right) = \vec{\nabla}\times\left[\left(\alpha\vec{v}-\vec{\beta}\right)\times\vec{B}\right]. \tag{2.9} $$

where $\gamma := \sqrt{\gamma_{ij}}$ and $\vec{\nabla}$ is the covariant derivative associated with the spatial 3-metric.

The evolution of a magnetised fluid (a plasma) is governed by the conservation laws of energy momentum $\nabla_\nu T^{\mu\nu} = 0$ and baryon number (conservation equation) $\nabla_\nu J^\nu = 0$, where the rest-mass current is given by $J^\mu = \rho u^\mu$, and ρ denotes the rest-mass density. For a magnetised fluid, the energy-momentum tensor is the combination of the energy-momentum tensor of the fluid and of the electromagnetic tensor,

$$ T^{\mu\nu} = T^{\mu\nu}_{\text{Fluid}} + T^{\mu\nu}_{\text{EM}}. \tag{2.10} $$

The fluid component is given by,

$$ T^{\mu\nu}_{\text{Fluid}} = \rho h u^\mu u^\nu + p g^{\mu\nu}, \tag{2.11} $$

where $h = 1 + \epsilon + P/\rho$ is the relativistic enthalpy, ϵ is the specific internal energy, p the fluid pressure, u^μ is the four-velocity and $g^{\mu\nu}$ is the metric. An equation of state (EoS) relates all these thermodynamical quantities.

In the ideal MHD limit, the energy-momentum tensor of the electromagnetic field can be defined in terms of the magnetic field b^μ measured by a comoving observer,

$$T^{\mu\nu}_{\rm EM} = \left(u^\mu u^\nu + \frac{1}{2} g^{\mu\nu} \right) b^2 - b^\mu b^\nu, \tag{2.12}$$

where $b^2 = b^\mu b_\mu$. Joining Eq. (2.11) and Eq. (2.12), the total energy-momentum tensor reads as,

$$T^{\mu\nu} = \left(\rho h + \frac{b^2}{2} \right) u^\mu u^\nu + \left(p + \frac{b^2}{2} \right) g^{\mu\nu} - b^\mu b^\nu. \tag{2.13}$$

As shown by Antón et al. [2006], the equations of GRMHD, together with the induction equation, can be cast as a first-order, flux-conservative hyperbolic system, in a similar way as it was done by Banyuls et al. [1997] in the purely hydrodynamical case. This system of equations is nowadays known as the *Valencia formulation* of ideal GRMHD. The set of conserved quantities of the system is given by

$$D = \rho W, \tag{2.14}$$

$$S_i = (\rho h + b^2) W^2 v_i - \alpha b_i b^0, \tag{2.15}$$

$$\tau = (\rho h + b^2) W^2 - \left(p + \frac{b^2}{2} \right) - \alpha^2 (b^0)^2 - D, \tag{2.16}$$

The explicit form of the flux-conservative, hyperbolic system of GRMHD equations reads,

$$\frac{1}{\sqrt{-g}} \left(\frac{\partial \sqrt{\gamma} \mathbf{U}}{\partial t} + \frac{\partial \sqrt{-g} \mathbf{F}^i}{\partial x^i} \right) = \mathbf{S}, \tag{2.17}$$

where $\sqrt{-g} = \alpha \sqrt{\gamma}$ and the state vector and flux vector are given by,

$$\mathbf{U} = \begin{bmatrix} D \\ S_j \\ \tau \\ B^k \end{bmatrix}, \tag{2.18}$$

$$\mathbf{F}^i = \begin{bmatrix} D\tilde{v}^i \\ S_j\tilde{v}^i + \left(p + \frac{b^2}{2}\right)\delta^i_j - b_j B^i/W \\ \tau\tilde{v}^i + \left(p + \frac{b^2}{2}\right)v^i - \alpha b^0 B^i/W \\ \tilde{v}^i B^k - \tilde{v}^k B^i \end{bmatrix} \tag{2.19}$$

with $\tilde{v}^i = v^i - \frac{\beta^i}{\alpha}$. The corresponding sources $\mathbf{S}$ are given by

$$\mathbf{S} = \begin{bmatrix} 0 \\ T^{\mu\nu}\left(\frac{\partial g_{\nu j}}{\partial x^\mu} - \Gamma^\delta_{\nu\mu} g_{\delta j}\right) \\ \alpha\left(T^{\mu 0}\frac{\partial \ln\alpha}{\partial x^\mu} - T^{\mu\nu}\Gamma^0_{\nu\mu}\right) \\ 0^k \end{bmatrix}, \tag{2.20}$$

where $0^k \equiv (0,0,0)^T$, δ^i_j is the Kronecker delta, and $\Gamma^\delta_{\nu\mu}$ are the Christoffel symbols associated with the four-metric. Note that the above system of equations contain components of the magnetic field measured by both a comoving observer and an Eulerian observer. They are related by

$$b^0 = \frac{W B^i v_i}{\alpha} \tag{2.21}$$

$$b^i = \frac{B^i + \alpha b^0 u^i}{W}\,. \tag{2.22}$$

The state of the fluid is uniquely described by the *primitive variables* which conform a vector $\boldsymbol{w} = [\rho, u_i, \epsilon]$ or by the vector of *conserved variables* $\boldsymbol{U} = [D, S_i, \tau]$. Note that one can be obtained from the other following Eq. (2.16).

2.2 The Riemann Problem

The Riemann problem[2] is an initial value problem with discontinuous data, given a linear (or non-linear) hyperbolic system of conservation laws. For example, consider the 1D linear problem,

$$U_t + U_x = 0, \tag{2.23}$$

where $U_t := \partial U/\partial t$ and $U_x := \partial U/\partial x$. The initial conditions at the *left* and *right* states of the initial discontinuity

$$U(x,0) = \begin{cases} U_{\rm L} & \text{if } x < 0 \\ U_{\rm R} & \text{if } x > 0 \end{cases} \tag{2.24}$$

[2]The analytic solution of the Riemann problem in relativistic MHD will be necessary when discussing the results of Chapter 3.

completely determine the evolution of the Riemann problem and how the discontinuity evolves with time.

Many numerical codes use finite-difference or finite-volume methods to compute the solution of a conservative system of equations, after a discretization of the computational domain. In the process of discretization (say, using piecewise constant data), at the interface of each pair of adjacent numerical cells a Riemann problem arises. The evolution of the system can thus be obtained through the solution of a sequence of Riemann problems along the computational grid, as Godunov [1959] first showed in 1959 for the equations of classical hydrodynamics. For this reason, the Riemann problem has gained an enormous importance in the field of numerical relativistic hydrodynamics and MHD.

In general, the solution of the Riemann problem cannot be obtained analytically (e.g. for microphysical EoS or for the GRMHD equations) and numerical approaches are required. Those *approximate* solutions derive from the solution of the Riemann problem for a reformulated system of equations, i.e. cast in a locally linear form (see Toro [2013] for details). The exact solution of the Riemann problem in one-dimensional special relativistic hydrodynamics was found by Martí and Müller [1994]. This seminal work has been extended to account for more general situations by Pons, Marti, and Müller [2000] and Rezzolla, Zanotti, and Pons [2003] (including transverse speeds) and by Giacomazzo and Rezzolla [2006] (for 1D special relativistic MHD). The analytic solution of the Riemann problem is not only important to understand fundamental issues of wave propagation in relativistic hydrodynamics and MHD but also to test modern numerical codes written in conservation form and based on *approximate Riemann solvers*, i.e. the so-called high-resolution shock-capturing (HRSC) schemes [LeVeque 1992]. These methods rely on the *shock-capturing* property which ensures that the numerical scheme converges to the correct Rankine-Hugoniot conditions across discontinuities. HRSC methods guarantee that the wave propagation of hyperbolic systems of conservation laws (such as the relativistic hydrodynamics and MHD systems) is accurate and correctly captured.

2.2.1 The Riemann problem in GRMHD

When the magnetic field of the initial states is tangential to the discontinuity and orthogonal to the fluid velocity, the exact solution of the Riemann problem was derived by Romero et al. [2005]. The general, exact solution of the Riemann problem in relativistic MHD was obtained by Giacomazzo and Rezzolla [2006]. It comprises a set of seven nonlinear waves: two *fast magnetosonic waves*, two

slow magnetosonic waves, two *Alfvèn waves* and a *contact discontinuity* at which only the density may be discontinuous. We discuss next the exact solution of the Riemann problem in GRMHD that will be employed in the results reported in Chapter 3.

For an ideal magneto-fluid, the energy-momentum tensor $T^{\mu\nu}$ and Maxwell dual tensor $F^{*\mu\nu}$ can we rewritten as

$$T^{\mu\nu} = \rho\hat{h}u^{\mu}u^{\nu} + g^{\mu\nu}\hat{p} - b^{\mu}b^{\nu}\,, \tag{2.25}$$

$$F^{*\mu\nu} = u^{\mu}b^{\nu} - u^{\nu}b^{\mu}\,, \tag{2.26}$$

where $\hat{h} = 1 + \varepsilon + p/\rho + b^2/\rho$ is the specific enthalpy including the contribution of the magnetic field and $\hat{p} = p + b^2/2$ is the total pressure.

The particular configuration of the Riemann problem discussed in Chapter 3 is $u^{\mu} = W(1, v^x, 0, v^z)$ and $b^{\mu} = (0, 0, b, 0)$. Therefore, the term $\nabla_{\mu}b^{\mu}b^{\nu}$ in the conservation of the stress-energy tensor, vanishes. Hence, the conservation equations reduce to the purely hydrodynamical case with the only contributions from the magnetic field appearing in the total pressure and specific enthalpy, and an additional continuity equation for the evolution of the transverse magnetic field. We have extended the previous work using a more general EoS instead of the polytropic EoS used in the original work.

The Riemann problem in this particular configuration is described in terms of three characteristics, one entropy wave and two fast magneto-sonic waves. The initial problem with two states L (left) and R (right) breaks up into four states,

$$L\mathcal{W}_{\leftarrow}L_{*}\mathcal{C}R_{*}\mathcal{W}_{\rightarrow}R,$$

where $\mathcal{W}$ indicates a fast magnetosonic shock wave or a rarefaction wave and $\mathcal{C}$ indicates the contact discontinuity.

Solving the Riemann problem entails finding the intermediate states (L_{*}, R_{*}) and the position of the waves which separates the four states (see Romero et al. [2005]). The functions $\mathcal{W}_{\rightarrow}$ and $\mathcal{W}_{\leftarrow}$ determine the functions $v^x_{R_{*}}(\hat{p})$ and $v^x_{L_{*}}(\hat{p})$ respectively. The condition,

$$v^x_{R_{*}}(\hat{p_{*}}) = v^x_{L_{*}}(\hat{p_{*}}) = v^x_{*}, \tag{2.27}$$

determines the pressure $\hat{p}_{*}$ and the flow velocity v^x_{*} at the intermediate states. The functions $v^x_{S*}(\hat{p})$ are defined by

$$v^x_{S*}(\hat{p}) = \begin{cases} \hat{\mathcal{R}}^S(\hat{p}) & \text{if } \hat{p} \leq \hat{p}_S \\ \hat{\mathcal{S}}^S(\hat{p}) & \text{if } \hat{p} > \hat{p}_S \end{cases}, \tag{2.28}$$

where $\hat{\mathcal{R}}^S(\hat{p})$ ($\hat{\mathcal{S}}^S(\hat{p})$) denotes the family of all states which can be connected through a rarefaction (shock) with a given state S (L, R) ahead of the wave. Once $\hat{p}_*$ and v^x_* have been obtained the remaining quantities can be computed. The goal is, therefore, to find the relation between the pressures of states ahead (a) and behind (b) the waves (rarefaction or shock) with the normal velocity of the state behind v_b.

If $\hat{p} \leq \hat{p}_*$ the wave is a rarefaction wave (a self-similar continuous solution), otherwise the solution is a shock wave. In our case [see Romero et al. 2005] the ordinary differential equation that allows to obtain the solution for a rarefaction wave is given by

$$\frac{dv^x}{1-(v^x)^2} = \pm \frac{(1+b^2/(\rho h c_s))\sqrt{\hat{h}+\hat{\mathcal{A}}^2(1-w^2)}}{\hat{h}^2+\hat{\mathcal{A}}^2}\frac{dp}{\rho w}, \tag{2.29}$$

where $\hat{\mathcal{A}} = \hat{h} W v^z$, $w = c_s^2 + v_A^2 - c_s^2 v_A^2$, $v_A = b^2/\rho\hat{h}$ is the Alfvén velocity and $c_s = \sqrt{\frac{1}{h}\left.\frac{\partial p}{\partial \rho}\right|_s}$ is the sound speed. The integration of Eq. (2.29) allows to connect the states ahead and behind the rarefaction wave. Rarefaction waves conserve entropy, hence, all the thermodynamical variables and the differential of p must be calculated at the same entropy of the initial state. From this equation, the normal velocity behind the rarefaction can be obtained directly,

$$v^x_b = \tanh \hat{\mathcal{C}}, \tag{2.30}$$

with

$$\begin{aligned} \hat{\mathcal{C}} &= \frac{1}{2}\log\left(\frac{1+v^x_a}{1-v^x_a}\right) \\ &\pm \int_{\hat{p}_a}^{\hat{p}_b} \frac{(1+b^2/(\rho h c_s))\sqrt{\hat{h}+\hat{\mathcal{A}}^2(1-w^2)}}{\hat{h}^2+\hat{\mathcal{A}}^2}\frac{dp}{\rho w}. \end{aligned} \tag{2.31}$$

In the same way, the velocity inside the rarefaction can be obtained replacing the thermal pressure p by the total pressure $\hat{p}$.

Shock waves should fulfill the so-called Rankine-Hugoniot conditions [Lichnerowicz 1967, Anile 1989]:

$$[\rho u^\mu]\, n_\mu = 0\,, \tag{2.32}$$

$$[T^{\mu\nu}]\, n_\nu = 0\,, \tag{2.33}$$

$$[F^{*\mu\nu}]\, n_\nu = 0\,, \tag{2.34}$$

where n_μ is the unit normal to a given surface and $[H] \equiv H_a - H_b$ being H_a and H_b the boundary values.

The invariant j can be directly obtained from these equations,

$$j \equiv W_s D_a (V_s - v_a^x) = W_s D_b (V_s - v_b^x)\,, \tag{2.35}$$

where V_s is interpreted as the coordinate velocity of the surface that defines the position of the shock wave and W_s is the Lorentz factor of the shock.

The normal flow speed in the post-shock state, v_b^x, can be extracted from the Rankine-Hugoniot equations [see Romero et al. 2005, for a detailed discussion],

$$v_b^x = \left(\hat{h}_a W_a v_a^x + \frac{W_s(\hat{p}_b - \hat{p}_a)}{j} \right) \tag{2.36}$$

$$\times \left(\hat{h}_a W_a + (\hat{p}_b - \hat{p}_a) \left(\frac{W_s v_a^x}{j} + \frac{1}{\rho_a W_a} \right) \right)^{-1} . \tag{2.37}$$

Using the definition of mass flux the shock velocity V_s can be expressed as,

$$V_s^{\pm} = \frac{\rho_a^2 W_a^2 v_a^x \pm |j| \sqrt{j^2 + \rho_a^2 W_a^2 (1 - v_a^x)^2}}{\rho_a^2 W_a^2 + j^2}\,. \tag{2.38}$$

These expressions, together with the Lichnerowicz adiabat for this particular case [see Pons, Marti, and Müller 2000, for details],

$$[\hat{h}^2] = \left(\frac{\hat{h}_b}{\rho_b} + \frac{\hat{h}_a}{\rho_a} \right), \tag{2.39}$$

allows to write ρ_b as a function of p_b and the preshock state a, thought the expression

$$j^2 = \frac{-[\hat{p}]}{\left[\hat{h}/\rho\right]}. \tag{2.40}$$

Using the positive (negative) root of j^2 for shock waves propagating towards the right (left), equation (2.40) allows one to obtain the desired relation between the post-shock normal velocity v_b^x and the post-shock pressure p_b.

Chapter 3

Accretion onto neutron stars and the hidden magnetic field scenario

The results of this chapter have been originally published in:
Torres-Forné, A., Cerdá-Durán, P., Pons, J. A., and Font, J. A. "Are pulsars born with a hidden magnetic field?" Monthly Notices of the Royal Astronomical Society, **456**(4), 3813-3826 (2016).

3.1 Astrophysical motivation

Central Compact Objects (CCOs) are isolated, young NSs which show no radio emission and are located near the centre of young supernova remnants (SNRs). Three such NSs, PSR E1207.4-5209, PSR J0821.0-4300, and PSR J1852.3-0040, show an inferred magnetic field significantly lower than the standard values for radio-pulsars (i.e. 10^{12} G). The main properties of these sources are summarised in Table 3.1. In all cases, the difference between the characteristic age of the NS $\tau_c = P/\dot{P}$ and the age of the SNR indicates that these NSs were born spinning at nearly their present periods ($P \sim 0.1 - 0.4$ s). This discovery has challenged theoretical models of magnetic field generation, that need to be modified to account for their peculiar properties.

The first possible explanation for the unusual magnetic field found in these objects simply assumes that these NSs are born with a magnetic field much lower than that of their classmates. This value can be amplified by turbulent

Table 3.1 Central Compact Objects in Supernova Remnants. From left to right the columns indicate the name of the CCO, the age, the distance d, the period P, the inferred surface magnetic field, B_s, the bolometric luminosity in X-rays, $L_{x,bol}$, the name of the remnant, the characteristic age, and bibliographical references.

CCO	Age (kyr)	d (kpc)	P (s)	B_s 10^{11}G	$L_{x,bol}$ (erg s^{-1})	SNR	τ_c (Myr)	References
J0822.0-4300	3.7	2.2	0.112	0.65	6.5×10^{33}	Puppis A	190	1, 2
1E 1207.4-5209	7	2.2	0.424	2	2.5×10^{33}	PKS 1209-51/52	310	2, 3, 4, 5, 6, 7
J185238.6+004020	7	7	0.105	0.61	5.3×10^{33}	Kes 79	190	8, 9, 10, 11

References: (1) Hui and Becker 2006, (2) Gotthelf, Halpern, and Alford 2013, (3) Zavlin et al. 2000, (4) Mereghetti et al. 2002, (5) Bignami et al. 2003, (6) De Luca et al. 2004, (7) Gotthelf and Halpern 2007, (8) Seward et al. 2003, (9) Gotthelf, Halpern, and Seward 2005, (10) Halpern et al. 2007, (11) Halpern and Gotthelf 2010

dynamo action during the PNS phase [Thompson and Duncan 1993, Bonanno, Urpin, and Belvedere 2005] . In this model, the final low values of the magnetic field would reflect the fact that the slow rotation of the NS at birth does not suffice to effectively amplify the magnetic field through dynamo effects. However, recent studies have shown that, even in the absence of rapid rotation, magnetic fields in PNS can be amplified by other mechanisms such as convection and the standing accretion shock instability (SASI) [Endeve et al. 2012, Obergaulinger, Janka, and Aloy 2014].

An alternative explanation is the hidden magnetic field scenario [Young and Chanmugam 1995, Muslimov and Page 1995, Geppert, Page, and Zannias 1999, Shabaltas and Lai 2012]. Following the supernova explosion and the NS birth, the supernova shock travels outwards through the external layers of the star. When this shock crosses a discontinuity in density, it is partially reflected and moves backwards (reverse shock). The total mass accreted by the reverse shock in this process is in the range from $\sim 10^{-4} M_\odot$ to a few solar masses on a typical timescale of hours to days [Ugliano et al. 2012]. Such a high accretion rate can compress the magnetic field of the NS which can eventually be buried into the NS crust. As a result, the value of the external magnetic field would be significantly lower than the internal 'hidden' magnetic field. Bernal, Lee, and Page [2010] performed 1D and 2D numerical simulations of a single column of material falling onto a magnetised NS and showed how the magnetic field can be buried into the NS crust.

Once the accretion process stops, the magnetic field might eventually reemerge. The initial studies investigated the process of reemergence using simplified 1D models and dipolar fields [Young and Chanmugam 1995, Muslimov and Page 1995, Geppert, Page, and Zannias 1999] and established that the timescale for the magnetic field reemergence is $\sim 1 - 10^7$ kyr, critically depending on the depth at which the magnetic field is buried. More recent investigations have

confirmed this result. Ho [2011] observed similar timescales for the reemergence using a 1D cooling code. Viganò and Pons [2012] carried out simulations of the evolution of the interior magnetic field during the accretion phase and the magnetic field submergence phase.

In the present chapter we study the feasibility of the hidden magnetic field scenario using a novel numerical approach based on the solutions of 1D Riemann problems (discontinuous initial value problems) to model the compression of the magnetic field of the NS. The two initial states for the Riemann problem are defined by the magnetosphere and by the accreting fluid, at either sides of a moving, discontinuous interface. Following the notation defined in Michel [1977], the NS magnetosphere refers to the area surrounding the star where the magnetic pressure dominates over the thermal pressure of the accreting fluid. The magnetopause is the interface between the magnetically dominated area and the thermally dominated area. The equilibrium point is defined as the radius at which the velocity of the contact discontinuity is zero.

The chapter is organized as follows. In Sections 3.2 to 3.4.1 we present the model we use to perform our study. We describe in these sections the equation of state (EoS) of the accreting fluid, the spherically symmetric Michel solution characterising the accreting fluid, and all the expressions needed to compute the potential solution for the magnetic field in the magnetosphere. Section 3.5 contains the main results of this work. After establishing a reference model, we vary the remaining parameters, namely entropy, composition and the initial distribution of the magnetic field, and study their influence on the fate of the magnetic field. Finally, in Section 3.6 we discuss our findings and summarise the main results of our study.

3.2 The reverse shock and the fallback scenario.

At the end of their lives, massive stars ($M_{\rm star} \gtrsim 8 M_\odot$) possess an onion-shell structure as a result of successive stages of nuclear burning. An inner core, typically formed by iron, with a mass of $\sim 1.4 M_\odot$ and ~ 1000 km radius develops at the centre, balancing gravity through the pressure generated by a relativistic, degenerate, $\gamma = 4/3$, fermion gas. The iron core is unstable due to photodisintegration of nuclei and electron captures, which result in a deleptonization of the core and a significant pressure reduction ($\gamma < 4/3$). As a result, the core shrinks and collapses gravitationally to nuclear matter densities on dynamical timescales (~ 100 ms). As the centre of the star reaches nuclear saturation density ($\sim 2 \times 10^{14}$ g cm^{-3}), the EoS stiffens and an outward moving

(prompt) shock is produced. As it propagates out the shock suffers severe energy losses dissociating Fe nuclei into free nucleons ($\sim 1.7 \times 10^{51}$ erg/$0.1 M_\odot$), consuming its entire kinetic energy inside the iron core (it stalls at $\sim 100 - 200$ km), becoming a standing accretion shock in a few ms. There is still debate about the exact mechanism and conditions for a successful explosion, but it is commonly accepted that the standing shock has to be revived on a timescale of $\lesssim 1$ s by the energy deposition of neutrinos streaming out of the innermost regions, and some form of convective transport for the shock to carry sufficient energy to disrupt the whole star [see Janka et al. 2007, for a review on the topic].

Even if the shock is sufficiently strong to power the supernova, part of the material between the nascent NS and the propagating shock may fall back into the NS [Colgate 1971, Chevalier 1989]. Determining the amount of fallback material depends not only on the energy of the shock but also on the radial structure of the progenitor star [Fryer 2006]. Most of the fallback accretion is the result of the formation of an inward moving reverse shock produced as the main supernova-driving shock crosses the discontinuity between the helium shell and the hydrogen envelope [Chevalier 1989]. For typical supernova progenitors ($10 - 30\, M_\odot$) the base of the hydrogen envelope is at $r_{\rm H} \sim 10^{11}$ cm to 3×10^{12} cm [Woosley, Heger, and Weaver 2002], which is reached by the main shock on a timescale of a few hours. The reverse shock travels inwards carrying mass that accretes onto the NS. It reaches the vicinity of the NS on a timescale of hours, about the same time at which the main supernova shock reaches the surface of the star [Chevalier 1989]. By the time the reverse shock reaches the NS, the initially hot PNS has cooled down significantly. In its first minute of life the PNS contracts, cools down to $T < 10^{10}$ K and becomes transparent to neutrinos [Burrows and Lattimer 1986, Pons et al. 1999]. In the next few hours the inner crust ($\rho \in [2 \times 10^{11}, 2 \times 10^{14}]$ g cm^{-3}) solidifies but the low density envelope ($\rho < 2 \times 10^{11}$ g cm^{-3}), which will form the outer crust on a timescale of $1 - 100$ yr, remains fluid [Page et al. 2004, Aguilera, Pons, and Miralles 2008].

Understanding the processes generating the magnetic field observed in NSs, in the range from $\sim 10^{10}$ G to $\sim 10^{15}$ G, is still a open issue. Most likely, convection, rotation and turbulence during the PNS phase play a crucial role in field amplification [Thompson and Duncan 1993]. However, at the time in the evolution that we are considering (hours after birth), none of these processes can be active anymore and the electric current distribution generating the magnetic field will be frozen in the interior of the NS. These currents evolve now on the characteristic Hall and Ohmic timescales of 10^4-10^6 yr [Pons and Geppert 2007, Pons, Miralles, and Geppert 2009, Viganò et al. 2013], much longer than the

timescale $t_{\rm acc}$ during which fallback is significant, which can be estimated as the free-fall time from the base of the hydrogen envelope

$$t_{\rm acc} \sim \frac{1}{2}\left(\frac{r_{\rm H}^3}{GM}\right)^{1/2} . \tag{3.1}$$

This ranges from 30 minutes to several days for the typical values of $r_{\rm H}$ and a $M = 1.4 M_\odot$.

The total mass accreted during this phase is more uncertain. Detailed 1D numerical simulations of the shock propagation and fallback estimate that typical values range from $10^{-4} M_\odot$ to a few solar masses [Woosley and Weaver 1995, Zhang, Woosley, and Heger 2008, Ugliano et al. 2012]. If more than a solar mass is accreted, the final outcome would be the delayed formation of a BH, hours to days after core bounce. Chevalier [1989] and Zhang, Woosley, and Heger [2008] showed that the accretion rate is expected to be maximum when the reverse shock reaches the NS and decreases as $t^{-5/3}$ at later times. Therefore, the total amount of accreted mass is dominated by the fallback during the first few hours. Given the theoretical uncertainties, we assume for the rest of this work that a total mass of $\delta M \in [10^{-5} M_\odot, \delta M_{\rm max}]$ is accreted during a typical timescale of $t_{\rm acc} \in [10^3, 10^4]$ s, being $\delta M_{\rm max} \sim 1 M_\odot$ the amount of mass necessary to add to the NS to form a BH hole. Therefore, the typical accretion rate during fallback is $\dot{M} \in [10^{-9}, 10^{-3}] M_\odot/{\rm s}$, which, for practical purposes, we assume to stay constant during the accretion phase. This accretion rate, even at its lowest value, exceeds by far the Eddington luminosity

$$\frac{\dot{M} c^2}{L_{\rm Edd}} = 5 \times 10^6 \left(\frac{\dot{M}}{10^{-9} M_\odot/{\rm s}}\right), \tag{3.2}$$

with $L_{\rm Edd} = 3.5 \times 10^{38}$ erg s^{-1} the Eddington luminosity for electron scattering.

In the hypercritical accretion regime, the optical depth is so large that photons are advected inwards with the flow faster than they can diffuse outwards [Blondin 1986, Chevalier 1989, Houck and Chevalier 1991]. As a result the accreting material cannot cool down resulting in an adiabatic compression of the fluid. The dominant process cooling down the accreting fluid and releasing the energy stored in the infalling fluid is neutrino emission [Houck and Chevalier 1991]. At temperatures above the pair creation threshold, $T_{\rm pair} \approx 10^{10}$ K, pair annihilation can produce neutrino-antineutrino pairs, for which the infalling material is essentially transparent and are able to cool down very efficiently the material as it is decelerated at the surface of the NS or at the magnetopause.

Therefore, the specific entropy, s, of the fallback material remains constant all through the accretion phase until it decelerates in the vicinity of the NS.

The value for s is set at the time of the reverse shock formation. Detailed 2D numerical simulations of the propagation of the shock through the star [Scheck et al. 2006, Kifonidis et al. 2003, Kifonidis et al. 2006] show that typical values of $s \sim 20\ k_{\rm B}$/nuc are found at the reverse shock. At this stage of the explosion the flow is highly anisotropic due to the Rayleigh-Taylor instability present in the expanding material and the Richtmyer-Meshkov instability at the He/H interface. Those instabilities generate substantial mixing between hydrogen and helium and even clumps of high-entropy heavier elements (from C to Ni) rising from the innermost parts of the star. Therefore, the fallback material has entropy in the range $s \sim 1 - 100\ k_{\rm B}$/nuc and its composition, although it is mostly helium, can contain almost any element present in the explosion. 3D simulations show qualitatively similar results regarding the entropy values and mixing [Hammer, Janka, and Müller 2010, Joggerst, Almgren, and Woosley 2010, Wongwathanarat, Müller, and Janka 2015].

Outside the NS, the expanding supernova explosion leaves behind a low density rarefaction wave which is rapidly filled by the NS magnetic field, forming the magnetosphere. For the small magnetospheric densities, the inertia of the fluid can be neglected, and the magnetosphere can be considered force-free. The fallback reverse shock propagates inwards compressing this magnetosphere. The boundary between the unmagnetised material falling back and the force-free magnetosphere, i.e. the magnetopause, can be easily compressed at long distances ($r \gtrsim 10^8$ cm) due to the large difference of the pressure of the infalling material with respect to the magnetic pressure. The dynamical effect of the magnetosphere only plays a role at $r \lesssim 10^8$ cm, i.e. inside the light cylinder for most cases. The precise radius where the magnetic field becomes dynamically relevant is estimated later in Section 3.4.2. Only in the case of magnetar-like magnetic fields and fast initial spin ($P \lesssim 10$ ms) this consideration is not valid, although this is not the case for CCOs.

To conclude this scenario overview, we note that the magnetospheric torques will spin-down the NS on a characteristic timescale [Shapiro and Teukolsky 1983] given by

$$\tau_c = \frac{P}{2\dot{P}} \sim 180 \left(\frac{B_p}{10^{15}\mathrm{G}}\right)^{-2} \left(\frac{P}{1\mathrm{s}}\right)^{2} \mathrm{yr}, \tag{3.3}$$

for a typical NS with radius 10 km and mass $1.4 M_\odot$. B_p is the value of the magnetic field at the pole of the NS. The value of the moment of inertia is

1.4×10^{45} g cm^2. At birth, the spin period of a NS is limited by the mass-shedding limit to be $P > 1$ ms [Goussard, Haensel, and Zdunik 1998]. If all NSs were born with millisecond periods, purely magneto-dipolar spin-down would limit the observed period of young NSs (10^4 yr) to

$$P_{\mathrm{obs},10^4\mathrm{yr}} \lesssim 5.5 \left(\frac{B}{10^{15}\mathrm{G}} \right)^2 \mathrm{s}. \tag{3.4}$$

For magnetic fields $B \lesssim 1.4 \times 10^{13}$ G this criterion fails for the vast majority of pulsars and all CCOs ($P \gtrsim 0.1$ s) and therefore the measured spin period must be now very close to that hours after the onset of the supernova explosion. Detailed population synthesis studies of the radio-pulsar population clearly favour a broad initial period distribution in the range 0.1-0.5 s [Faucher-Giguère and Kaspi 2006, Gullón et al. 2014], rather than fast millisecond pulsars. Therefore, from observational constraints, it is reasonable to assume that progenitors of pulsars (including CCOs) have spin periods of $P \sim 0.1 - 0.5$ s at the moment of fallback. For such low rotation rates, the NS can be safely considered as a spherically symmetric body and its structure can thus be computed by solving the Tolman-Oppenheimer-Volkoff (TOV) equation.

3.3 Stationary spherical accretion

We model the fallback of the reverse shock as the spherically symmetric accretion of an unmagnetised relativistic fluid. The stationary solutions for this system were first obtained by Michel [1972] for the case of a polytropic EoS. Here, we extend this work to account for a general (microphysically motivated) EoS. The equations that describe the motion of matter captured by a compact object, i.e. a NS or BH, can be derived directly from the equations of relativistic hydrodynamics , namely the conservation of rest mass,

$$\nabla_\mu J^\mu = 0\,, \tag{3.5}$$

and the conservation of energy-momentum,

$$\nabla_\mu T^{\mu\nu} = 0\,, \tag{3.6}$$

where we use the notation ∇_μ for the covariant derivative and the density current J^μ and the (perfect fluid) energy-momentum tensor $T^{\mu\nu}$ are given by

$$J^\mu \quad = \quad \rho u^\mu\,, \tag{3.7}$$

$$T_{\mu\nu} = \rho h u_\mu u_\nu + p g_{\mu\nu} \,. \tag{3.8}$$

In the above equations ρ is the rest-mass density, p is the pressure and h is the specific enthalpy, defined by $h = 1 + \varepsilon + p/\rho$, where ε is the specific internal energy, u^μ is the four-velocity of the fluid and $g_{\mu\nu}$ defines the metric of the general spacetime where the fluid evolves. Assuming spherical symmetry and a steady state we have

$$\frac{d}{dr}(J^1\sqrt{-g}) = 0 \,, \tag{3.9}$$

$$\frac{d}{dr}(T^1_0\sqrt{-g}) = 0 \,, \tag{3.10}$$

where $g \equiv \det(g_{\mu\nu})$. The exterior metric of a non-rotating compact object is given by the Schwarzschild metric

$$\begin{aligned} ds^2 &= -\left(1 - \frac{2M}{r}\right) dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 \\ &+ r^2(d\theta^2 + \sin^2\theta\, d\varphi^2) \,. \end{aligned} \tag{3.11}$$

In Schwarzschild coordinates Eqs. (3.9) and (3.10) can be easily integrated to obtain [cf. Michel 1972]

$$\rho\, u\, r^2 = C_1 \,, \tag{3.12}$$

$$h^2\left(1 - \frac{2M}{r} + u^2\right) = C_2 \tag{3.13}$$

where C_1 and C_2 are integration constants and $u \equiv u^r$. To obtain an adiabatic solution for the accreting fluid, we differentiate Eqs. (3.12) and (3.13) at constant entropy and eliminate $d\rho$

$$\begin{aligned} &\frac{du}{u}\left[V^2 - u^2\left(1 - \frac{2M}{r} + u^2\right)^{-1}\right] \\ &+ \frac{dr}{r}\left[2V^2 - \frac{M}{r}\left(1 - \frac{2M}{r} + u^2\right)^{-1}\right] = 0 \,, \end{aligned} \tag{3.14}$$

where

$$V^2 \equiv \frac{\rho}{h}\left.\frac{\partial h}{\partial \rho}\right|_s \,. \tag{3.15}$$

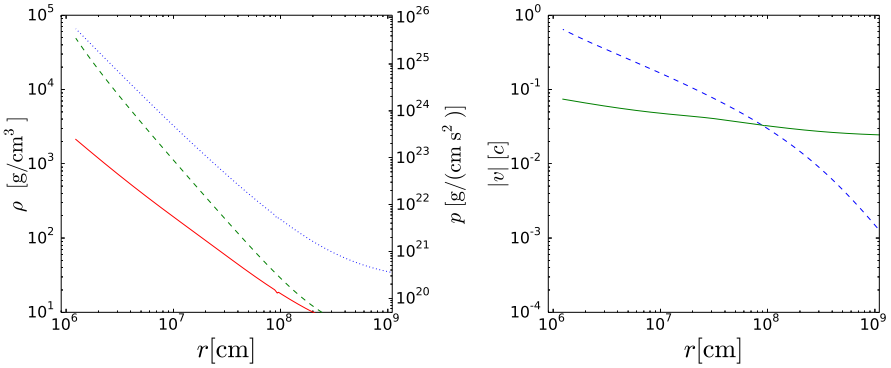


Figure 3.1 Illustrative accretion solution for an accretion rate $\dot{M} = 10^{-5} M_\odot/s$ and entropy per baryon $s = 80\ k_\mathrm{B}$/nuc. The left panel shows the density (green-dashed line, left axis), pressure (red-solid line, right axis) and ram pressure (blue-dotted line, right axis). The right panel shows the absolute value of the fluid velocity (blue-dashed line) and the sound speed (green-solid line). The two lines cross at the critical point.

The solutions of this equation are those passing through a critical point where both terms in brackets in equation (3.14) are zero, i.e. those fulfilling

$$\begin{aligned} 2u_c^2 &= \frac{M}{r_c}\,, \\ V_c^2 &= u_c^2(1-3u_c^2)^{-1}\,, \end{aligned} \tag{3.16}$$

where sub-index c indicates quantities evaluated at the critical point. The critical point can be identified as the sonic point, i.e. the point where the velocity of the fluid equals its own sound speed. After some algebra, it can be shown that the constant C_1 in Eq. (3.12) is related to the accretion rate $\dot{M}$ by

$$\dot{M} = -4\pi C_1. \tag{3.17}$$

Thereby we can obtain the accretion solution by simply selecting the mass accretion rate and the specific entropy of the fluid, which fixes the two constants C_1 and C_2. We note that, for each pair of values, the system (3.16) has two solutions, although only one represents a physical accretion solution ($|u| \to 0$ at $r \to \infty$). In this case the fluid is supersonic for radii below the critical radius and subsonic above. Fig. 3.1 displays one illustrative accretion solution for a mass accretion rate $\dot{M} = 10^{-5} M_\odot$/s and entropy per baryon $s = 80 k_\mathrm{B}$/nuc. For the accreting material, we use the tabulated Helmholtz EoS [Timmes and Swesty 2000], which is an accurate interpolation of the Helmholtz free-energy of the Timmes EoS [Timmes and Arnett 1999]. Timmes EoS, and Helmholtz EoS by extension, include the contributions from ionised nuclei, electrons, positrons

and radiation. By default, Timmes EoS uses the rest mass density ρ [g/cm^3], temperature T [K] and composition as input. For convenience, we have developed a search algorithm that allows to call the EoS with different thermodynamical variables as input (e.g. ρ, s and composition as inputs for the adiabatic flow of accreting material). Helmholtz EoS also requires the mean mass number $\bar{A}$ and the mean atomic number $\bar{Z}$.

At low densities, $\rho < 6\times10^7$ g cm^{-3}, and temperatures, $T \lesssim 2\times10^9$ K, nuclear reactions proceed much slower than the accretion timescale and the composition remains frozen during the accretion. We fix the composition to that at the reverse shock formation point. Given the uncertainties, we consider two possibilities in this regime, either pure Helium or pure Carbon. At temperatures $T \gtrsim 2\times10^9$ K nuclear burning becomes fast enough to change the composition. For $T \gtrsim 4\times10^9$ K the fluid reaches nuclear statistical equilibrium (NSE) on a significantly shorter timescale than the accretion timescale [see e.g. Woosley, Heger, and Weaver 2002]. To deal with the high temperature regime, $T \geq 2\times10^9$ K, we have tried three different approaches: 1) unchanged composition of the accreting material, 2) compute the NSE composition at a given temperature and density using a thermonuclear reaction network with 47 isotopes [Timmes 1999, Seitenzahl et al. 2008] and 3) simplified burning with four transitions: ^{4}He for $T \leq 2\times10^9$ K, ^{56}Ni for $2\times10^9 > T \geq 5\times10^9$ K, ^{4}He for $5\times10^9 > T \geq 2\times10^{10}$ K and protons and neutrons for $T > 2\times10^{10}$ K. We use the publicly available routines of the Hemlholtz EoS and the NSE equilibrium kindly provided by the authors[1].

3.4 Non-magnetised accretion and pile-up

Before considering the case of magnetised accretion onto a NS, we study the case of non-magnetised accretion. For the span of accretion rates considered in this work, the sonic point of the accreted fluid is located at $r > 23500$ km at entropy $s = 10\ k_{\rm B}$/nuc, and hence the fallback material falls supersonically onto the NS. Inevitably an accretion shock forms at the surface of the star, which propagates outwards. The accreted fluid crossing the shock will heat up, increasing its specific entropy and will fall subsonically. The high entropy of this material ($s_{shock} \in [70-300]\ k_{\rm B}$/nuc) and the compression that experiments as it flows inwards raises the temperature beyond the pair creation threshold, $T_{\rm pair} \approx 10^{10}$ K, and the fluid will cool efficiently via neutrino-antineutrino annihilation. Therefore, the kinetic energy of the supersonically accreting fluid

[1] http://cococubed.asu.edu/code_pages/codes.shtml

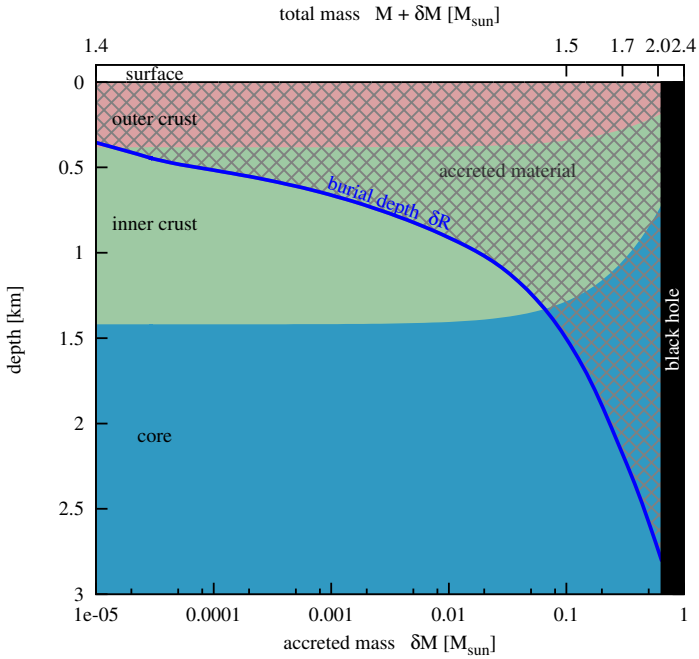


Figure 3.2 Dependence of the burial depth (blue solid lines), δR, with the accreted mass, δM (bottom axis), for a $M = 1.4 M_\odot$ NS using APRDH EoS. Note, that positive values of the depth increase downwards. Regions occupied by the outer crust, inner crust and core appear with different colours and labeled. The region occupied by the accreted material is plotted with a gray crosshatch pattern. The top axis shows the total mass of the NS after accretion, $M + \delta M$. Above $M = 2.25\ M_\odot$ the configuration is unstable and the object will collapse to a BH.

is mostly transformed into thermal energy as it crosses the accretion shock and then is dissipated to neutrinos close to the NS surface. Chevalier [1989] showed that the accretion shock will eventually stall at a certain radius as an energy balance is found. The radius of the stalled shock depends only on the accretion rate $\dot{M}$ and is located at about $R_{\rm shock} \sim 10^7 - 10^8$ km. In some estimates below in this work we use the values provided in Table 1 in Houck and Chevalier [1991], based in a more realistic treatment of the accretion and neutrino cooling.

The final fate of the neutrino-cooled material falling steadily onto the NS surface is to pile up on top of the original NS material forming a layer of new material. In order to study the effect of the pile up we consider a NS of mass M and radius R. If we add a mass δM to the equilibrium model, the new NS will have a new radius $R_{\rm new}$ smaller than the original one. The original surface of the star, will now be buried at a depth δR, i.e. the new surface will

be located at a distance δR over the old surface. Although trivial, the last statement is important because most of the discussion below in this work is carried out in terms of δR and in terms of distances with respect to the original NS surface. Therefore it makes sense to try to compute what is the dependence of the burial depth, δR, with the total accreted mass, δM. In order to compute this dependence we use the TOV equations to solve a sequence of NS equilibrium models starting with M and progressively increasing to $M + \delta M$ for different values of δM. For each model in the sequence we compute δR as the distance between the radius enclosing a mass M and the surface of the star, i.e. the radius enclosing $M + \delta M$. Given the small values of δM, we integrate the TOV equations using a simple forward Euler method, with a step limited to relative variations of density of 10^{-5} and a maximum step of 10 cm. We have computed the relation between δR and δM for four different NS masses, $M = 1.2$, 1.4, 1.6, 1.8 and $2.0 M_\odot$. We have used several realistic EoS in tabulated form, namely four different combinations using either EoS APR [Akmal, Pandharipande, and Ravenhall 1998] or EoS L [Pandharipande and Smith 1975] for the core and EoS NV [Negele and Vautherin 1973] or EoS DH [Douchin and Haensel 2001] for the crust. For each case we compute the sequence up to the maximum mass; beyond that mass, the equilibrium model is unstable and it will collapse to a BH in dynamical timescales. All EoS allow for equilibrium solutions with maximum mass consistent with recent observations of a NS with mass close to $2 M_\odot$ [Demorest et al. 2010, Antoniadis et al. 2013]. The blue solid line in Fig. 3.2 shows the dependence of δR with δM for a $1.4 M_\odot$ NS with the APRDH EoS. All other EoS and NS masses show similar behaviour. For all EoS, any amount of accreted mass larger than $\sim 10^{-4} M_\odot$ will sink the original NS surface to the inner crust, and for $\delta M \sim 0.1 M_\odot$ the entire crust is formed by newly accreted material. The bottom line is that, if the accreted material is able to compress the magnetosphere and deposit itself on top of the NS, the magnetic field trapped with the fluid may be buried into the NS crust, and depending on the conditions (accreted mass and magnetic field strength), the burial depth could be as deep as the inner crust. We study next the impact of magnetic fields in the vicinity of the NS, namely the magnetosphere, in the burial process.

3.4.1 Potential magnetospheric solution

For simplicity in the following discussion we use a *reference model* with the APRDH EoS and $M = 1.4 M_\odot$. This model results in a a NS with coordinate radius $R = 12.25$ km. The effect of the EoS and the NS mass are discussed

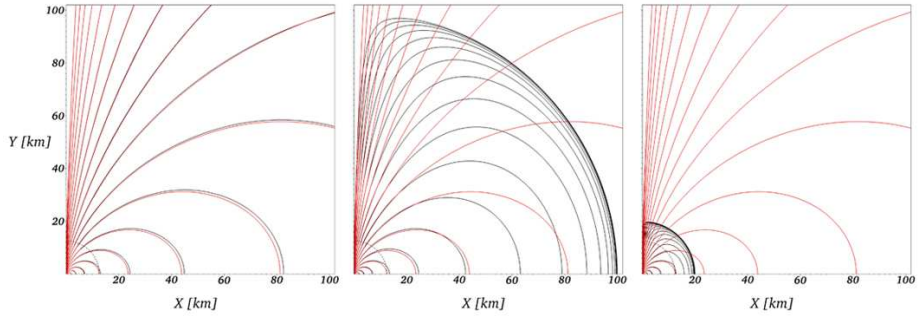


Figure 3.3 Magnetic field lines (black lines) for three different positions of the magnetopause for the same initial distribution of the magnetic field (red lines) with $B_p = 10^{13}$ G. The dashed line represents the NS surface.

later in the text. Given that both the magnetosphere and the accreted material involve low energy densities compared with those inside the NS, the spacetime outside the NS can be regarded as non-self-gravitating and approximated by the Schwarzschild exterior solution.

The magnetosphere extends between the NS surface and the magnetopause, which will be assumed to be a spherically symmetric surface at the location of the infaling reverse shock. We model this region using the force-free magnetic field approximation, $\mathbf{J} \times \mathbf{B} = 0$, $\mathbf{J}$ being the electric current and $\mathbf{B}$ the magnetic field. We neglect the currents resulting from the rotation of the star. Consequently the magnetic field has a potential solution, solution of the relativistic Grad-Shafranov equation. In spherical coordinates, the magnetic field vector components are related to the vector potential $\mathbf{A}$ as,

$$\hat{B}_r = \frac{1}{r^2 \sin\theta} \partial_\theta A_\phi \,, \tag{3.18}$$

$$\hat{B}_\theta = \frac{-1}{r^2 \sin\theta} \partial_r A_\phi \,, \tag{3.19}$$

$$\hat{B}_\phi = 0 \,, \tag{3.20}$$

where $\hat{B}_i = \sqrt{\gamma} B_i$ and γ is the determinant of the spatial metric. If we assume axisymmetry, the unique nonzero component of the electric current is the ϕ component,

$$J_\phi = \sin\theta \left[\partial_r (r \hat{B}_\theta) - \partial_\theta \hat{B}_r \right] . \tag{3.21}$$

Imposing the force-free condition, we obtain,

$$-J_\phi \hat{B}_\theta = 0 \,, \tag{3.22}$$

$$J_\phi \hat{B}_r = 0\,. \tag{3.23}$$

Since $\hat{B}_r, \hat{B}_\theta \neq 0$, the only possible solution is $J_\phi = 0$. As we want an expression that only depends on the vector potential, we replace Eqs. (3.18) and (3.19) in equation (3.21) resulting in

$$\begin{aligned} J_\phi &= \sin\theta \left[\frac{-1}{\sin\theta} \partial_r(\partial_r) A_\phi - \frac{1}{r^2} \partial_\theta \left(\frac{\partial_\theta A_\phi}{\sin\theta} \right) \right] \\ &= -\partial_{rr} A_\phi - \frac{1}{r^2} \partial_{\theta\theta} A_\phi + \frac{\cot\theta}{r^2} \partial_\theta A_\phi = 0\,. \end{aligned} \tag{3.24}$$

We discretise this expression using second order finite differences and solve the resulting linear system of equations using a cyclic reduction algorithm [Swarztrauber 1974]. We impose Dirichlet boundary conditions on A_ϕ at the surface of the NS to match with the interior value of the radial component of the magnetic field. Our aim is to describe a magnetosphere, which is confined within a certain radius, $R_{\rm mp}$, defining the magnetopause. Magnetic field lines at the magnetopause are parallel to this interface and they enter the NS along the axis. Therefore, they correspond to lines with $A_\phi = 0$, which we use as Dirichlet boundary condition at $R_{\rm mp}$ to solve the Grad-Shafranov equation. We can obtain the field distribution after the compression by simply changing the radius where the boundary conditions are imposed. The evolution of the magnetic field geometry before and after compression is shown in Fig. 3.3 for three illustrative cases.

For the interior magnetic field, which determines the boundary conditions at the surface of the star, we use two different magnetic field distributions, a dipolar magnetic field (*dipole* hereafter) and a poloidal field generated by a circular loop of radius $r = 4 \times 10^5$ cm [Jackson 1962] (*loop current* hereafter). Following Gabler et al. [2013], it is useful to introduce the *equivalent magnetic field*, B^*, which we define as the magnetic field strength at the surface of a Newtonian, uniformly magnetised sphere with radius 10 km having the same dipole magnetic moment as the configuration we want to describe. It spans the range $B^* \in [10^{10} - 10^{16}]$ G.

3.4.2 Magnetosphere compression

In the case of a fluid accreting onto a force-free magnetosphere, the magnetopause will remain spherical and will move inwards as long as the total pressure of the unmagnetised fluid, $p_{\rm tot} = p + p_{\rm ram}$, exceeds that of the magnetic pressure, $p_{\rm mag}$, of the magnetosphere. If we approximate the magnetopause as a spherical

boundary between the spherically symmetric accreting solution described in Section 3.3 and the potential solution computed in Section 3.4.1, its properties can be described as the solution of a Riemann problem at the magnetopause. Since the magnetic field of the initial state is tangential to the magnetopause, we can use the exact solution of the Riemann problem developed by Romero et al. [2005]. A succinct summary of the details of the implementation of the Riemann solver can be found in Chapter 2.

For illustrative purposes the left panel of Fig. 3.4 shows the solution of the Riemann problem for a supersonic fluid accreting from the right into a magnetically dominated region (magnetosphere) on the left. The figure displays both the density (left axis, solid lines) and the fluid velocity (right axis, dashed lines). The initial discontinuity is located at $x = 0$. The right constant state of the Riemann problem corresponds to the accreting fluid with an entropy of $s = 10\ k_{\rm B}$/nuc and accretion rate of $\dot{M} = 10^{-7}\ {\rm M}_{\odot}$/s. The left constant state corresponds a state with magnetic pressure $B^2/2$. The figure plots the corresponding solutions for different values of B around the equilibrium (indicated in the legend).

Looking at the upper panel of Fig. 3.4 from left to right, the first jump in density corresponds to the contact discontinuity, point at which, as expected, the velocity remains continuos. The next discontinuity is a shock wave, where both the density and velocity are discontinuous, and both decrease. For low magnetic fields, $B \leq 10^{10}G$, the low magnetic pressure on the left state cannot counteract the total pressure of the accreting fluid and the contact discontinuity advances to the left at a velocity equal to that of the accreting fluid; a shock front is practically nonexistent. As the magnetic field is increased the velocity of the contact discontinuity decreases and it becomes zero at about $B = 10^{13}$ G. We identify this point as the *equilibrium point*, since no net flux of matter crosses $x = 0$. Around this equilibrium point, an accretion shock appears, which heats and decelerates matter coming from the right. The equilibrium point corresponds to a solution in which the matter crossing the shock has zero velocity, i.e. it piles up on top of the left state as the shock progresses to the right.

The actual accretion of matter onto a magnetically dominated magnetosphere is expected to behave in a similar way as the described Riemann problem. At large distances (low B) the magnetopause (contact discontinuity) is compressed at the speed of the fluid. As the magnetosphere is compressed, the magnetic field strength raises and at some point an equilibrium point is found, beyond which the magnetosphere impedes the accretion of the fluid.

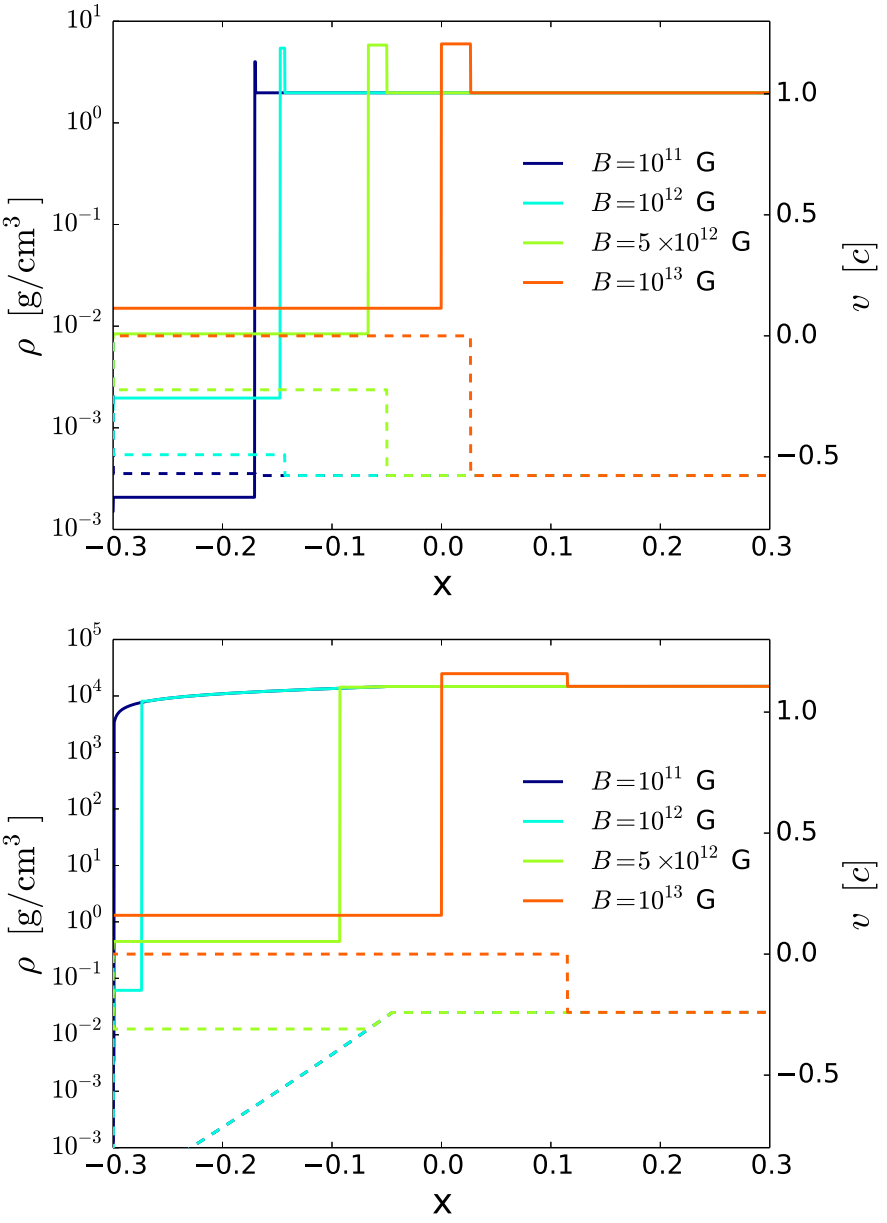


Figure 3.4 Density (solid lines, left axis) and velocity (dashed lines, right axis) profiles of the solution of the Riemann problem for several values of the magnetic field. Initially the discontinuity is set at $x = 0$, an accreting fluid at $x > 0$ and a magnetised fluid at $x < 0$, with constant magnetic field B. The upper panel shows the case of supersonic accretion of a fluid with specific entropy $s = 10k_{\rm B}$/nuc and $\dot{M} = 10^{-7}$ M$_\odot$/s at $t = 0.3$ s . The bottom panel shows the case of subsonic accretion of a fluid with $s = 2000k_{\rm B}$/nuc and $\dot{M} = 10^{-5}$ M$_\odot$/s at $t = 0.3$ s.

In the bottom panel of Fig. 3.4 we show for the sake of completeness the solution for a subsonic accreting fluid. In accreting NS this regime is probably unrealistic, since very large specific entropy is necessary ($s = 2000 k_{\rm B}$/nuc in the example plotted). In this case the solution is qualitatively different; instead of a shock, a rarefaction wave is formed for B below the equilibrium point. For larger values of B, an accretion shock is formed.

3.4.3 Setup

Our goal is to study the conditions under which the magnetic field of a new-born NS can be buried by fallback material during a supernova. We have spanned a large range of values for both, the magnetic field strength and the accretion rate, proceeding as follows. We obtain the distance from the NS surface where the magnetosphere and the accreting fluid are in balance, i.e. the radial point where the velocity of the contact discontinuity is zero. We reduce our 2D configuration to a 1D Riemann problem by restricting the evaluation of the equilibrium point to the equatorial plane of the NS, due to the fact that the magnetic pressure is maximum at the equator. Therefore, if the magnetic field can be buried in this latitude, it will be buried in all latitudes of the NS.

The code developed by Romero et al. [2005] requires as input the knowledge of the density, velocity, thermal pressure, and magnetic pressure at both left and right states of the initial discontinuity. In all cases we consider, the left state corresponds to the force-free magnetosphere while the right state is occupied by the accreting fluid. To obtain the magnetic pressure of the left state we find the solution of the Grad-Shafranov equation (see Section 3.4.1). This allows to locate the position of the magnetopause where the Riemann problem must be solved. Since the inertia of the fluid at the magnetosphere can be neglected in front of the magnetic pressure, the value of the density on the left state is set to yield an Alfvén velocity near to one , the thermal pressure is set to be at least six orders of magnitude lower than the magnetic pressure, and the velocity is set to zero. On the other hand, the values on the right state are fixed to the corresponding values of density, pressure and velocity of the stationary spherical accretion solution (see Section 3.3) and the magnetic pressure is set to zero.

A sketch of the different stages of the accretion process is shown in Fig. 3.5. The plots depict the location of the NS (including its core and inner and outer crust), the magnetosphere, the magnetopause, and part of the region where material is falling back. Each region is shaded in a different colour for a simple identification. Note that the scale ratio of the different regions is not preserved in

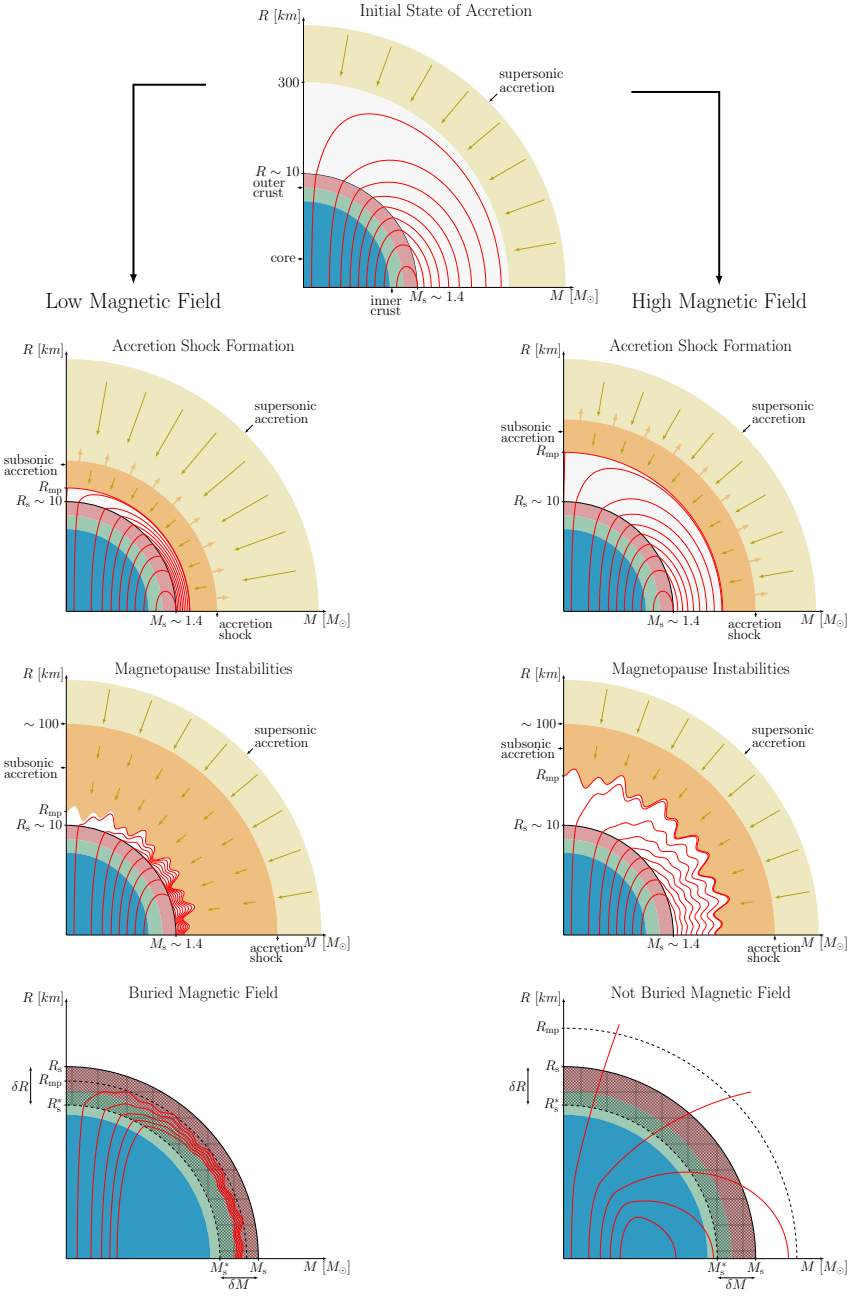


Figure 3.5 Sketch of the representative stages of the accretion process. The upper panel shows the initial state of the process. The left column shows the expected evolutionary path for a low magnetic field ($B \lesssim 10^{13}$ G) while the right column correspond to a typical high magnetic field case (e.g. $B \gtrsim 10^{13}$ G). A mass accretion rate of $10^{-5} M_\odot$/s is assumed. The scale ratio of the different regions is not preserved. See main text for details.

the figure. The upper panel in Fig. 3.5 shows the initial state of the process. The panels on the left column show the expected evolution for a low magnetic field case (e.g. $B \lesssim 10^{13}$ G) while those on the right column correspond to a typical high magnetic field case (e.g. $B \gtrsim 10^{13}$ G). In general, the value of B separating between the two regimes depends on the accretion rate. For this figure we have chosen a value of the magnetic field that corresponds to a representative example of our results (see Section 6), for which $\dot{M} = 10^{-5} M_{\odot}$/s. At the beginning of the evolution, the reverse shock falls over the magnetosphere. The magnetic field lines are confined inside the magnetosphere, which is shown in white on the diagram. Depending on the position of the sonic point, which in turn depends on the values of the specific entropy and the accretion rate, the motion of the reverse shock may be either supersonic or subsonic. We limit the qualitative description of the evolution below to the case of a supersonic reverse shock as in the subsonic case no accretion shock forms, as shown in Section 3.4.2.

The middle two panels in both evolutionary tracks show only qualitative differences in the size of the resulting magnetosphere after its compression and in the amplitude of the instabilities that may arise in the magnetopause (see below). Therefore, our description can be used for either path keeping this quantitative differences in mind. The evolution on the left column shows the case where the magnetic pressure is weak compared with the ram pressure of the fluid. In this case the magnetosphere shrinks significantly until the equilibrium point is reached ($R_{\rm mp}$; zero speed contact discontinuity) close to the NS surface at $R_{\rm s} \sim 10$ km. If the infall of the reverse shock is supersonic an accretion shock will appear simultaneously. The location of this accretion shock is shown on the horizontal axis of the four middle panels. As a result, the velocity of the reverse shock is reduced due to the presence of a region of subsonic accretion behind the accretion shock. Nevertheless, as through the accretion shock the momentum is conserved, the compression is not affected. The evolution on the right column, where the magnetic pressure is stronger, is qualitatively similar, only the accretion shock is located further away from the NS surface and the magnetosphere is not so deeply compressed.

As we will discuss below in more detail, the compression phase is unstable against the growth of Rayleigh-Taylor instabilities and the development of convection on the dynamic timescale. Therefore, the fluid and the magnetic field lines can mix, which provides a mechanism for the infalling fluid to actually reach the star. As the fluid reaches the NS, the mass of the star grows from $M_{\rm s}^*$ to $M_{\rm s}$ and its radius increases from $R_{\rm s}^*$ to $R_{\rm s}$, encompassing the twisted magnetic field lines a short distance away. The mass accreted δM forms part of the new

crust of the NS, whose final radius will depend on the total mass accreted during the process. The bottom panels of the diagram depict a magnified view of the NS to better visualise the rearrangement the mass of the star and the magnetic field undergo. If the radius $R_{\rm mp}$ of the equilibrium point is lower than the new radius $R_{\rm s}$, all the magnetic field lines will be frozen inside the NS new crust, as shown in the bottom-left plot of Fig. 3.5 which corresponds to the end of the accretion process for a low magnetic field evolution. On the contrary, if the magnetic field is high, as considered on the evolutionary path on the right, the equilibrium point $R_{\rm mp}$ is far from the surface of the NS. Although part of the infalling matter may still reach the star and form a new crust, the mechanism is not as efficient as in the low magnetic field case. This is depicted in the bottom-right panel of the figure.

In our approach, that we discuss in more detail in the section on results, we compare the distance obtained by the Riemann solver for the location of $R_{\rm mp}$ (zero speed in the contact discontinuity) with the increment of the radius of the NS, δR, due to the pile up of the accreting matter. If the radial location of the equilibrium point $R_{\rm mp}$ is lower than δR (as in the bottom-left panel of Fig. 3.5) we conclude that the magnetic field is completely buried into the NS crust. On the contrary, if $R_{\rm mp} > \delta R$, our approach does not allow us to draw any conclusion. In this case, multidimensional MHD numerical simulations must be performed to obtain the final state of the magnetic field.

3.5 Results

We turn next to describe the main results of our study. In order to be as comprehensive as possible, we cover a large number of cases which are obtained from varying the physical parameters of the model, namely the composition and entropy of the accreting fluid, the mass of the NS, and the initial magnetic field distribution. For all possible combinations of these parameters the outcome of the accretion process depends both on the magnetic field strength and on the mass accretion rate. This dependence is presented in the following sections in a series of representative figures. A summary of all the combinations considered and the description of the model parameters can be found in Table 3.2.

3.5.1 Reference model

We use as a reference model the one corresponding to an accreting fluid with $s = 10\ k_{\rm B}/{\rm nuc}$, and composed essentially by Helium. The nuclear reactions to

Table 3.2 Models considered in this study.

Model #	Composition	Entropy k_B/nuc	NS Mass $M_\odot$	MF distribution
Reference	He + NSE	10	1.4	loop current
1	He + NSE	100	1.4	loop current
2	He + NSE	1000	1.4	loop current
3	He + NSE	5000	1.4	loop current
4	He	10	1.4	loop current
5	He	100	1.4	loop current
6	He	1000	1.4	loop current
7	He	5000	1.4	loop current
8	C + NSE	10	1.4	loop current
9	C + NSE	100	1.4	loop current
10	C + NSE	1000	1.4	loop current
11	C + NSE	5000	1.4	loop current
12	He	10	1.4	dipole
13	He	1000	1.4	dipole
14	He	10	1.2	loop current
15	He	10	1.6	loop current
16	He	10	1.8	loop current
17	He	10	2.0	loop current

reach nuclear statistical equilibrium are also allowed in this model. The mass of the NS is 1.4 $M_\odot$ and the magnetic field is generated by a *loop current* in the NS. The results are shown in Fig. 3.6. The solid lines in this figure represent the distance δR of the equilibrium point (position of the magnetopause) above the NS surface as a function of the total accreted mass δM. The limit of the horizontal axis is given by the maximum mass that can be accreted without forming a BH. Each line corresponds to a different value of the initial magnetic field, indicated in the legend of the figure. The yellow area represents the region in which the accretion of the reverse shock is supersonic and the black dotted line shows the limit of the accretion shock. The dashed red line shows the radial location of the new surface of the star due to the accretion of the infalling matter. The lines which cross the dashed red line have the equilibrium point inside the crust of the NS and, therefore, the corresponding magnetic fields will be buried into the crust. However, for the lines that are in the white area, the equilibrium point is not close enough to the NS surface and the magnetic field can not be buried. Note that for initial values of the magnetic field $B \gtrsim 10^{15}$ G, the magnetic field is never buried for all mass accretion rates considered.

An alternative view of this result is shown in Fig. 3.7. The goal of this figure is to provide a clearer representation of the dependence of the equilibrium point

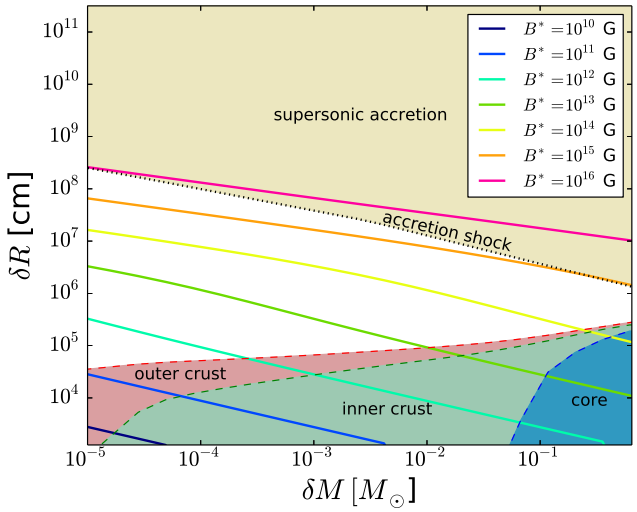


Figure 3.6 Distance above the star of the equilibrium point above the star as a function of the total mass accreted δM for each value of the magnetic field (solid lines) for the reference model. The yellow area indicates the region where the accretion flow is supersonic. The dotted line represents the limit of the accretion shock. The red area marks the outer crust of the NS after accretion, while the green and blue areas display the inner crust and the core respectively, as shown in Fig. 3.5.

with the span of values of the magnetic field and the total mass accreted we are considering. The figure shows the isocontours where the equilibrium point is equal to the increment of the radius of the NS, i.e. $R_{\rm mp} = \delta R$. The two lines plotted (dotted, $t = 10^4$ s, and solid, $t = 10^3$ s) correspond to the limits of the total accretion time, which relates the accretion rate $\dot{M}$ and the total mass increment δM. The black area indicates the values of the maximum mass of the NS beyond which it will form a BH. The dark orange region represents the span of values of δM and B^* where we cannot assure that the magnetic field could be buried completely. The light orange area, on the other hand, represents the cases where the magnetic field is totally buried. The results show that for low values of the magnetic field ($B^* < 10^{11} G$) the field can be buried even with the lowest accretion rates we have considered. As expected, as the accreted mass increases it is possible to bury the magnetic field for larger initial field values, up to a certain maximum. Indeed, for $B^* > 2 \times 10^{14}$ G we cannot find any accretion rate which can bury the magnetic field.

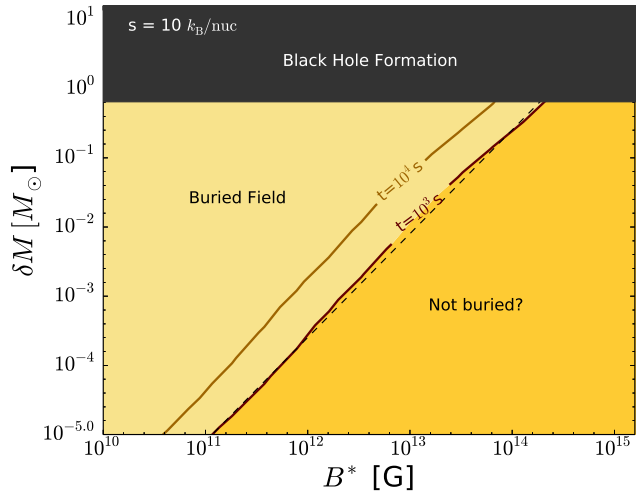


Figure 3.7 Outcome of the accretion depending on the total accreted mass (δM) and the initial magnetic field (B^*) for the reference model. For the two accretion times considered, $t = 10^3$ s (dark brown) and $t = 10^4$ s (light brown), the respective line splits the parameter space in a region where the magnetic field will be buried (left side) or not completely buried (right side). Above certain δM a BH will be formed. The dashed line represents the fit shown in Eq. (3.25).

3.5.2 Models with higher specific entropy

We turn next to analyse the behaviour of the magnetic field compression when the accreting fluid has higher specific entropy than in the reference model, keeping the same conditions for the composition, mass and magnetic field distribution (Models 1, 2 and 3 in Table 3.2). Fig. 3.8 shows the results for values of the specific entropy of $s = 100$ k_B/nuc, 1000 k_B/nuc and 5000 k_B/nuc compared with the reference model ($s = 10$ k_B/nuc). For the model with specific entropy 100 k_B/nuc, the results are very similar to the reference model as both lines almost perfectly overlap. For larger specific entropy the difference is more noticeable; for $s = 1000$ k_B/nuc and 5000 k_B/nuc, the burial/reemergence boundary of the parameter space is shifted toward larger magnetic fields, i.e. higher entropy material compress the magnetosphere more easily and it is possible to bury larger magnetic fields. This behaviour can be understood if one considers that the equilibrium point is a balance between the total pressure of the infall material, $p_{\rm tot} = p + p_{\rm ram} \approx p + \rho v^2$, and the magnetic pressure of the magnetosphere. For low specific entropy, the total pressure is dominated by the ram pressure and changes in s do not produce significant changes in the equilibrium point. Above a certain threshold, the thermal pressure p dominates the total pressure and increasing s induces a larger compression of the magnetosphere, shifting

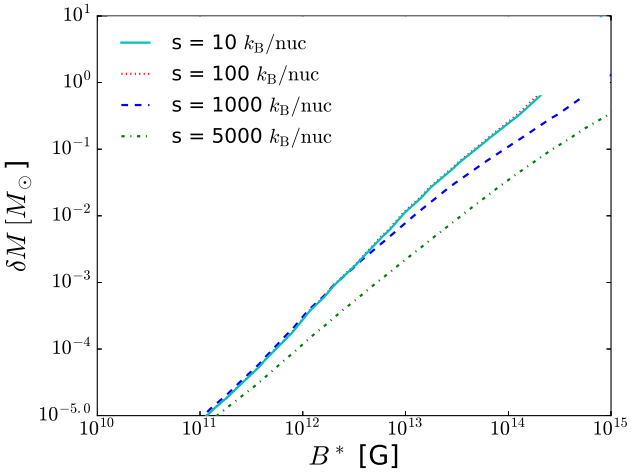


Figure 3.8 Similar to Fig. 3.7 but for the models with different specific entropy for the accreting fluid (namely, models 1 to 3 and reference). All cases are shown for a total accretion time of 10^3 s. Each line ends at the maximum mass of the corresponding model.

the equilibrium point downwards. For realistic values of the specific entropy in supernovae, $s \sim 10 - 100\ k_B$/nuc [Scheck et al. 2006, Kifonidis et al. 2003, Kifonidis et al. 2006], we expect the ram pressure to be dominant and hence the influence of s to be minimal. Even for an unrealistically large value of the specific entropy, 5000 k_B/nuc, the maximum magnetic field that can be buried increases one order of magnitude at most, and only for the largest mass accretion rates considered.

3.5.3 Models with different NS mass

We consider next the effect of the NS mass, within astrophysically relevant limits. According to observations [see Lattimer 2012, and references therein] the lower limit for the NS mass is around 1.2 $M_\odot$. The maximum achievable mass of a NS is strongly dependent on the equation of state [Lattimer and Prakash 2005]. Nowadays, there are a few observations that support the existence of pulsars and NS with masses greater than 1.5 $M_\odot$, in particular an observation of a $\sim 2\ M_\odot$ NS [Demorest et al. 2010, Antoniadis et al. 2013]. For this reason, we explore the results for several values of the NS mass between 1.2 $M_\odot$ and 2 $M_\odot$. The results are shown in Fig. 3.9, where each line corresponds to a model with different NS mass as indicated in the legend. The results for all masses are very similar. In general we observe that for more massive NS, a higher accreted mass is needed to bury the magnetic field. Our interpretation is that higher mass

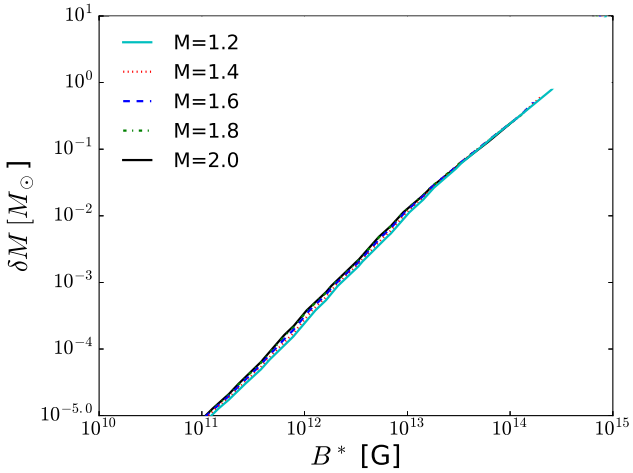


Figure 3.9 Similar to Fig. 3.7 for the models with different NS mass: 1.2 $M_\odot$, 1.6 $M_\odot$, 1.8 $M_\odot$, 2.0 $M_\odot$ (models 14, 15, 16 and 17 and reference). All cases are shown for a total accretion time of 10^3 s. Each line ends at the maximum mass of the corresponding model.

NS have lower radii and hence we have to compress more the magnetosphere to successfully bury it into the crust. Therefore, a higher accreted mass is needed to bury the field for NS with larger mass (smaller radius). Since the radius difference between a 1.2 and a 2 $M_\odot$ NS is small, the impact of the NS mass on the burial is minimal. The maximum value of the magnetic field which can be buried is $\sim 2 \times 10^{14}$ G in all cases. For smaller NS masses slightly larger values of the magnetic field can be buried due to the ability to support a larger accreted mass. We conclude that the burial of the magnetic field is not crucially sensitive to the NS mass.

3.5.4 Models with different EoS

Fig. 3.10 shows the comparison of the results for the reference model when using the four different equations of state described in section 3.4. For $M = 1.4\ M_\odot$, the coordinate radius of these NS models is 12.25 km for APRDH, 12.11 km for APRNV, 15.77 km for LDH and 15.37 km for LNV. Since the maximum mass is sensitive to the EoS, each line ends at different points in the δM vs B^* plot. The use of APRDH or APRNV EoSs leads to almost indistinguishable results (the two lines lay on top of each other). This is expected since the radius of this two models differs only by about 1%, because the EoS are very similar and only differ at low densities (at the crust). The LDH and LNV EoSs allow the burial of a larger magnetic field for a given accreted mass, in comparison with APRDH

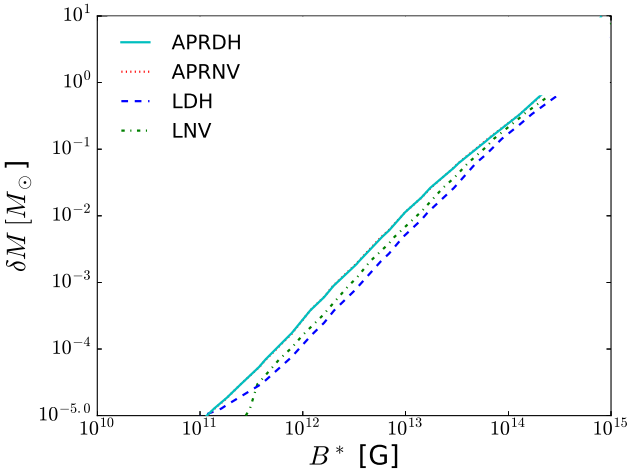


Figure 3.10 Similar to Fig. 3.7 for the models with different EoS, a NS of mass 1.4M$_\odot$ and specific entropy of the accreting fluid $10k_B$/nuc All cases are shown for a total accretion time of 10^3 s. Each line ends at the maximum mass of the corresponding model.

and APRNV. The maximum magnetic field that can be buried in the LDH and LNV models is $\sim 6 \times 10^{14}$ G and $\sim 5 \times 10^{14}$ G respectively, which is about a factor 2 larger than for the APRDH EoS. In general, for a $M = 1.4\ M_\odot$, EoS resulting in a larger NS radius allow to bury larger magnetic fields for a given δM. Given that the results of this work are meant to be an order-of-magnitude estimate of the location in the parameter space of the limit between burial and reemergence, a difference of a factor 2 due to the EoS, does not change the main conclusions of this work. For practical purposes the APRDH EoS can be taken as a good estimator for this limit.

3.5.5 Remaining models

We do not observe any significant differences with respect to the reference model in the results for the models with different initial composition of the reverse shock (models 8 to 11) or the ones using the NSE calculations (models 4 to 7). As a result we do not present additional figures for these models since the limiting lines overlap with those of the reference model. The observed lack of dependence is due to the fact that the EoS only depends on the electron fraction, Y_e. This value is obtained from the ratio between the mean atomic mass number ($\bar{A}$) and the mean atomic number ($\bar{Z}$). For both cases of pure Helium and pure Carbon, this ratio is equal to $Y_e = 0.5$ and, consequently, the values of pressure and density for the accreting fluid are almost identical, producing differences in the results

below the numerical error of our method [2]. In the case of the NSE calculation, the reason is similar. For low entropies ($s = 10 - 100\,k_{\rm B}$/nuc) the temperature is not sufficiently high to start the nuclear reactions and the composition remains constant throughout the accretion phase. For higher entropies, although the value of the electron fraction may differ from 0.5 during the accretion process, the differences produced in the thermodynamical variables lead to changes in the results of the Riemann problem still below the numerical error of the method.

Regarding the initial distribution of the magnetic field, we do not observe either any significant difference in the results in the two cases that we have considered, *loop current* and *dipole*. Given that we are comparing models with the same effective magnetic field, B^*, and thus the same magnetic dipolar moment, the magnetic field is virtually identical at long radial distances and the only differences appear close to the NS surface. In practice the magnetic field structure only changes the details of the burial in the cases in which the equilibrium point is close to the burial depth (the limiting line plotted in the Figs. 3.7 to 3.10), but it does not change the location of the limit itself in a sensitive way. As a conclusion, we can say that the dominant ingredient affecting the burial of the magnetic field is the presence of a dipolar component of the magnetic field but, for order-of-magnitude estimations, a multipolar structure of the field is mostly irrelevant.

3.6 Discussion

Assuming an accretion time of 1000s, our findings can be summarised by a general condition, rather independent on the model details, relating the required total accreted mass to bury the magnetic field with the field strength. An approximate fit is (see dashed line in Fig. 3.7)

$$\frac{\delta M}{M_\odot} \approx \left(\frac{B}{2.5\times 10^{14}}\right)^{2/3}. \tag{3.25}$$

The most important caveat in our approach is that we are restricted to a simplistic 1D spherical geometry, which does not allow us to consistently account for the effect of different MHD instabilities that can modify the results. We also note that our scenario is quite different from the extensively studied case of X-ray binaries, in which the NS accretes matter from a companion but at much lower rates (sub-Eddington) and matter is mostly transparent to radiation during

[2] The numerical error is dominated by the calculation of the equilibrium point, which is computed with a relative accuracy of 10^{-4}.

accretion. In that case, matter cools down through X-ray emission during the accretion process. Davidson and Ostriker [1973] and Lamb, Pethick, and Pines [1973] already noticed this fact and predicted that the accretion will most likely be channeled through the magnetic poles, in analogy to the Earth's magnetosphere. In the context of X-ray binaries, Arons and Lea [1976] and Michel [1977] were able to compute equilibrium solutions with a deformed magnetosphere and a cusp like accretion region at the magnetic poles. However, as the same authors pointed out, these systems are unstable to the interchange instability [Kruskal and Schwarzschild 1954], a Railegh-Taylor-like instabilitiy in which magnetic field flux tubes from the magnetosphere can raise, allowing the fluid to sink. This might allow for the formation of bubbles of material that fall through the magnetosphere down to the NS surface. In the case of a fluid deposited on top of a highly magnetised region, modes with any possible wavelength will be unstable [Kruskal and Schwarzschild 1954], however, in practice these instabilities are limited to the size of the magnetosphere ($\sim R_{\rm mp}$) in the angular direction. As the bubbles of accreted material sink, magnetic flux tubes raise, as long as their magnetic pressure equilibrates the ram pressure of the unmagnetised accreting fluid [Arons and Lea 1976]. Therefore, in a natural way, the equilibrium radius computed in Section 3.4.2 roughly determines the highest value at which the magnetic field can raise.

This accretion mechanism through instabilities has been shown to work in the case of X-ray binaries in global 3D numerical simulations [e.g. Kulkarni and Romanova 2008, Romanova, Kulkarni, and Lovelace 2008]. In the case of the hypercritical accretion present in the supernova fallback, Rayleigh-Taylor instabilities have been studied by Payne and Melatos [2004], Payne and Melatos [2007], Bernal, Lee, and Page [2010], Bernal, Page, and Lee [2013], Mukherjee, Bhattacharya, and Mignone [2013a], and Mukherjee, Bhattacharya, and Mignone [2013b]. The simulations of Bernal, Page, and Lee [2013] also show that the height of the unstable magnetic field over the NS surface decreases with increasing accretion rate, for fixed NS magnetic field strength, as expected. Using the method described in Section 3.4.2 we have estimated the equilibrium height over the NS surface for the 4 models presented in Fig. 9 of Bernal, Page, and Lee [2013], for their lower accretion rates ($\dot{M} \leq 10^{-6}$ $\mathrm{M}_{\odot}$/s). Our results predict correctly the order of magnitude of the extent of the unstable magnetic field over the NS surface. Therefore, our simple 1D model for the equilibrium radius serves as a good estimator of the radius confining the magnetic field during the accretion process, although details about the magnetic field structure cannot be predicted. Another important difference with the binary scenario is the duration

of the accretion process. In X-ray binaries, a low accretion rate is maintained over very long times, so that instabilities have always time to grow. In our case, hypercritical accretion can last only hundreds or thousands of seconds, and depending on the particular values of density and magnetic field, this may be too short for some instabilities to fully develop. This issue is out of the scope of this paper and deserves a more detailed study.

Our main conclusion is that a typical magnetic field of a few times 10^{12} G can in principle be buried by accreting only $10^{-3} - 10^{-2} M_{\odot}$, a relatively modest amount of mass. This estimate has interesting implications: since it is likely that most NSs can undergo such an accretion process, and the field would only reemerge after a few thousand years [Geppert, Page, and Zannias 1999, Viganò and Pons 2012], the CCO scenario is actually not peculiar at all and we expect that most very young NSs show actually an anomalously low value of the magnetic field. On the contrary, magnetar-like field strengths are much harder to screen and the required accreted mass is very large, in some cases so large that the NS would collapse to a BH. We also stress that the concept of *burial* of the magnetic field refers only to the large scale dipolar component, responsible for the magnetospheric torque spinning down the star. Small scale structures produced by instabilities can exist in the vicinity of the star surface, and this locally strong field is likely to have a visible imprint in the star thermal spectrum, as in Kes 79 [Shabaltas and Lai 2012], without modifying the spin-down torque. However, the high field burial scenario should not be very common because both, high field NSs are only a fraction to the entire population, and only a part of them would undergo the fallback episode with the right amount of matter. This is consistent with the recent results of Bogdanov, Ng, and Kaspi [2014] who searched for the hidden population of evolved CCOs among a sample of normal pulsars with old characteristic ages but close to a supernova remnant. None of the eight sources studied was found to have a luminosity higher than 10^{33} erg/s, which would have been an evidence of a hidden strong field. They all show X-ray luminosities in the 0.3-3 keV band of the order of 10^{31} erg/s (or similar upper limits), consistent with the properties of other low field NSs with $B \approx 10^{12}$ G. Thus, these sample of sources are not likely to be linked to the family of descendants of Kes 79-like objects, but there is no contradiction with these being pulsars with reemerged normal fields. Finally, we note that the slow reemergence process on timescales of kyrs mimics the increase of the magnetic field strength, and it is therefore consistent with a value of the braking index smaller than 3 [Espinoza et al. 2011], which should be common for all young pulsars in this scenario.

Chapter 4

Linear-oscillation spectrum of proto-neutron stars

4.1 Gravitational waves in the collapse of massive stars

The collapse of massive stars, i.e. those stars with a mass larger than about $8M_{\odot}$, is among the most important sources of gravitational waves. Collapsing stars produce rich and complex waveforms, which could provide ample information about the phenomenology of the scenario, specially when combined with observations of their electromagnetic emission and neutrino emission. The outcome of those events is either a neutron star or a black hole, typically (but not necessarily in the latter case) accompanied by a supernova explosion. The modelling of core collapse supernova requires a wide variety of physical ingredients, including general relativity, a nuclear-physics-motivated EoS, and a detailed description of neutrino interaction [see e.g. Janka et al. 2007, for a review]. For sufficiently compact stellar cores a black hole is likely to form instead of a neutron star [O'Connor and Ott 2013], although those cases not always coincide with the most massive stars [see e.g. Ugliano et al. 2012]. Recent studies have shown that, even if a black hole is formed, a successful explosion is still possible provided sufficient rotation is present in the core [Obergaulinger and Aloy 2017]. In addition, it is necessary to model complex multidimensional effects and instabilities, such as convection, the standing accretion shock instability (SASI) and turbulence, which are crucial for the development of a successful supernova explosion.

The numerical modelling of this scenario is computationally challenging and even today, with the use of the largest scientific supercomputing facilities available, we are only starting to understand the physics involved and we are probably still far away from having detailed waveforms. Unlike the BBH case, it is currently not possible to relate uniquely and unambiguously the properties of the progenitor stars (mass, rotation rate, metallicity, magnetic fields) with the resulting waveforms. The reasons are the complex non-linear dynamics associated with the evolution of a fluid interacting with neutrino radiation, the stochastic and chaotic behaviour of instabilities (both during and prior to the collapse of the star), the uncertainties in stellar evolution of massive stars (specially regarding the treatment of convection, magnetic fields and angular momentum transport) and the uncertainties in the nuclear and weak interactions necessary for the EoS at high densities and neutrino radiation, respectively.

Despite the difficulties, impressive progress has been made in the last decade regarding waveforms. The core bounce is the part of the waveform which is best understood [Dimmelmeier, Font, and Müller 2002]. Its frequency (at about 800 Hz) can be directly related to the rotational properties of the core [Dimmelmeier et al. 2008, Abdikamalov et al. 2014a, Richers et al. 2017]. However, fast-rotating progenitors are not common and their bounce signal will be probably difficult to observe in typical non-rotating galactic events, due to its high frequency and low amplitude. More interesting is the signal related to the post-bounce evolution of the newly formed proto-neutron star (PNS). The main sources of gravitational waves in that case are convection and the excitation of highly damped modes in the PNS by the accreting material and instabilities (SASI). A number of groups have identified features in the gravitational-wave signal as associated with g-modes in the PNS and SASI in either 2D simulations [Murphy, Ott, and Burrows 2009, Müller, Janka, and Marek 2013, Cerdá-Durán et al. 2013] or 3D simulations [Kuroda, Kotake, and Takiwaki 2016, Andresen et al. 2017]. Typically the waveforms last for about 500 ms until the supernova explodes [see e.g Müller, Janka, and Marek 2013] or, in the case of black hole formation, the typical duration is above 1 s [Cerdá-Durán et al. 2013]. Typical frequencies raise monotonically with time due to the contraction of the PNS, whose mass is steadily increasing. Characteristic frequencies of the PNS can be as low as ~ 100 Hz, specially those related to g-modes, which make them a perfect target for ground-based interferometers with the highest sensitivity at those frequencies. Two regions in the PNS appear to be susceptible to g-modes [Cerdá-Durán et al. 2013, Andresen et al. 2017, Kuroda, Kotake, and Takiwaki 2016], namely the surface of the PNS and the innermost cold core of the PNS. In those two regions,

the specific entropy increases with the radius, resulting in convectively stable regions in which buoyancy acts as a restoring force. Interestingly, quasi-radial modes in rotating cores produce a distinctive frequency-decreasing pattern in the waveforms that signals the formation of a black hole as its frequency approaches zero [Cerdá-Durán et al. 2013]. All these results imply that it may be possible to infer the properties of PNS based on the identification of mode frequencies in their waveforms, without the necessity of a complete understanding of the details about the physics involved in the core-collapse scenario.

The idea of identifying the properties of PNS based on the study of the frequencies of their normal modes of oscillation is not new. Studies of PNS asteroseismology have been reported in a number of works [see e.g. Reisenegger and Goldreich 1992, Ferrari et al. 2004, Passamonti et al. 2005, Camelio et al. 2017]. The common approach in those cases is to study the oscillations in PNS as linear perturbations of a spherical equilibrium star (which we call "background model" hereafter). This results in an eigenvalue problem, whose solutions are the normal oscillation modes of the PNS. In most of the previous work a fairly simplified description of the post-bounce evolution of the PNS has been considered. Recently, Sotani and Takiwaki [2016] Have performed a linear perturbation analysis of PNS, based on simple fits to realistic 1D core-collapse simulations, to study the evolution of the mode frequencies up to several hundreds of ms after bounce. Fuller et al. [2015] performed a Newtonian linear perturbation analysis using more realistic profiles from 2D simulations, but they restricted their analysis to the bounce signal of rapidly-rotating progenitors and did not consider the post-bounce evolution. So far, the presence of a standing shock above the PNS surface has not been taken into account and the oscillations have been limited to the interior of the PNS by imposing boundary conditions at the PNS surface[1].

In the work presented in this chapter, we have developed a method to perform the linear-perturbation analysis in general relativity of a background model, which is the result of multi-dimensional core-collapse simulations including the PNS and the hot-bubble region up to the shock location. This allows us to compare directly the frequencies of the modes obtained by our analysis with the actual frequencies in the gravitational-wave spectrum of the very same simulation. This analysis provides a proof of principle that asteroseismology is indeed possible in the core-collapse scenario, despite the complexities of the

[1]The exception being Fuller et al. [2015] that considered the PNS surrounded by a low-density accreting region and imposed outgoing sound-wave boundary conditions in the outer boundary.

system, and will serve as a basis for future work on PNS parameter inference based on gravitational-wave observations.

4.2 Black-hole-forming model 35OC

In this work we concentrate our efforts in applying our analysis to a single core-collapse model. The idea is to understand in detail the spectrum of eigenmodes of the coupled physical system formed by the PNS and the shock wave. To accomplish this goal, we have developed and tested a numerical code that computes and automatically classifies eigenmodes of a model we simulated previously (and hence can confidently use to text our approach), before applying this tool to larger sets of models. For this study we choose the 2D simulation performed by Cerdá-Durán et al. [2013] using the general-relativistic code `CoCoNuT` [Dimmelmeier, Font, and Müller 2002, Dimmelmeier et al. 2005]. The progenitor is a low-metallicity $35M_\odot$ star at zero-age main-sequence from Woosley and Heger [2006]. This progenitor has a high rotation rate and is usually regarded as a progenitor of long GRBs. The LS220 EoS [Lattimer and Swesty 1991] was used in the simulation to describe matter at high densities along with a simplified leakage scheme to approximate neutrino transport. The core of the progenitor collapses after 342.7 ms, forming a PNS and an accretion shock, after which an accretion phase ensues. The infalling matter crosses the stalled shock, heats up and falls through the hot bubble in which particles can dwell for some time before reaching the surface of the PNS, due to convection and the SASI. Finally, after 1.6 s, the PNS becomes unstable to radial perturbations and collapses to a black hole. During this time the highly-perturbed PNS is an efficient emitter of gravitational waves, as Fig. 4.1 shows.

Throughout all the accretion phase we compute the eigenmodes of the region including the PNS up to the shock location. The size of this region varies in time as the shock position changes. At post-bounce time, this region is approximately at hydrostatic equilibrium and flow velocities are small compared to the speed of sound (supersonically falling matter becomes subsonic as it crosses the shock). Therefore, we can study linear perturbations of a background, provided by the result of the simulation at a given time. This approach is possible as long as the typical evolution timescales of the background are much longer than the inverse of the frequency of the modes studied. In our case, the structure of the PNS varies in a timescale of ~ 100 ms, which limits the validity of our approach to frequencies larger than ~ 10 Hz. Moreover, despite its relatively high rotation frequency, centrifugal forces are not dominant when compared to

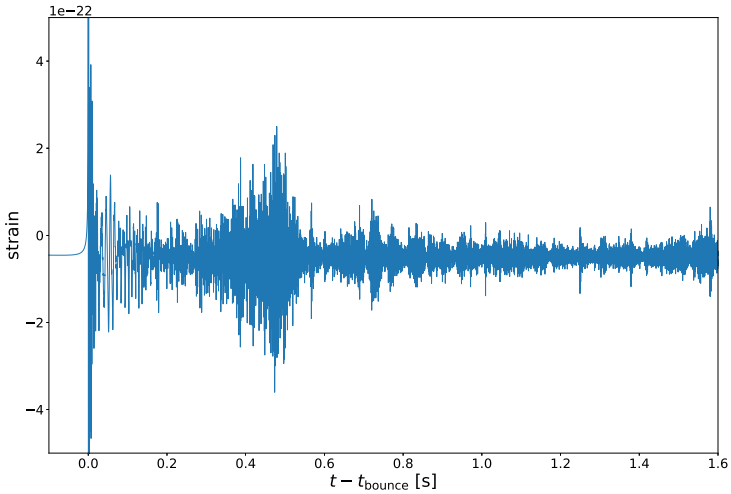


Figure 4.1 Waveform of the GW signal computed in the core-collapse simulation of Cerdá-Durán et al. [2013]. The progenitor is a low-metallicity 35 $M_\odot$ star at zero-age main-sequence from Woosley and Heger [2006].

gravitational forces [see Cerdá-Durán et al. 2013]. To simplify the analysis, we perform angular averages of all quantities to compute an effective 1D model, which will be used to perform the linear perturbation analysis. Since the shock is not completely spherical in the simulation but deformed by the presence of the SASI, angular averages have to be performed with care. We first compute the averaged shock location, then rescale the radial profiles of all quantities to the average shock radius and finally perform the angular average.

For our analysis it is interesting to define three different regions between the centre of the star and the position of the accretion shock. These regions are displayed in Fig. 4.2. The inner cold core (in blue) is the area between the centre and the radius at which the specific entropy is lower than 3 k_B per baryon, marked in the figure as r_cold. It is not easy to define the surface of the PNS. Above the cold inner core there is a broad hot mantle of $10 - 20$ km where the density decreases rapidly. The neutrinosphere, r_ν, is typically located just bellow the surface and could be regarded as a proxy for the PNS radius. However, it tends to underestimate the size of the PNS slightly. We found that the point where the density becomes 10^{11} g cm^{-3} is a much better proxy for the PNS surface, and we use its value, $r_{\rho_{11}}$, as a definition of the PNS surface location in this work. Finally, the position of the shock is labeled in Fig. 4.2 as r_shock. To summarise, the three regions are: region I, corresponding to the

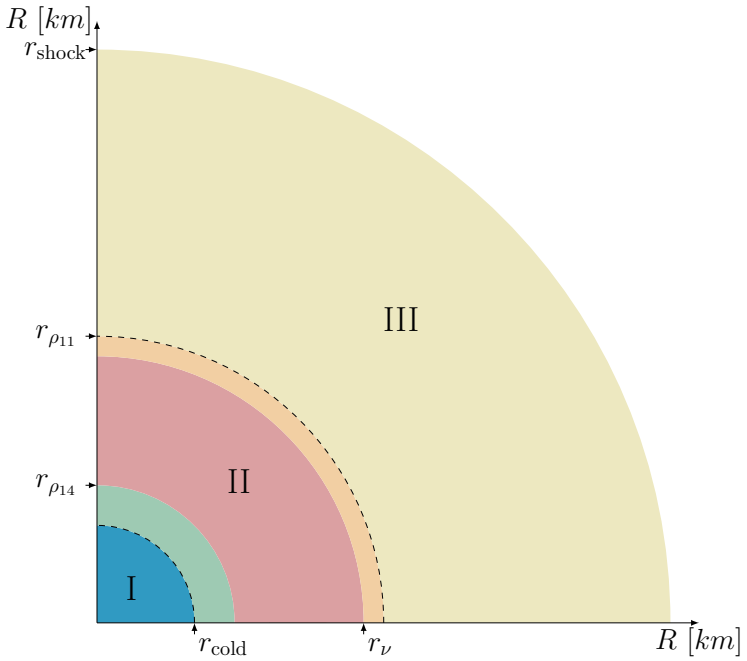


Figure 4.2 Representation of the different regions of the PNS considered. (See main text for details.)

inner core, region II, defined as the region between the core and the surface of the PNS, corresponding to the PNS mantle, and region III, which is the area between the surface and the shock, corresponding to the hot bubble.

4.3 Linear perturbations of a spherically-symmetric background

We start our analysis with the description of the perturbations of a spherically-symmetric, self-gravitating, equilibrium configuration. Classically, this analysis was performed in Schwarzshild coordinates [Thorne and Campolattaro 1967].The interested reader is addressed to Kokkotas and Schmidt [1999] and to Friedman and Stergioulas [2013] for detailed information on linear perturbations of compact stars and asteroseismology.. In our work we will use isotropic coordinates instead, which are closer to the gauge condition used in the numerical simulations being analysed (CFC approximation, Isenberg [2008] and Wilson, Mathews, and Marronetti [1996]). Moreover, the derivation of the equations in these

coordinates also bears resemblance with the equations in the Newtonian case [see Reisenegger and Goldreich 1992], which makes it easier to identify the role of the different terms in the equations and interpret the solutions. This choice of gauge also makes it straightforward to perform the mode analysis of Newtonian simulations.

Let us consider a $3+1$ foliation of spacetime in coordinates (t, x^i), in which the metric can be written as

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = (\beta^i\beta_i - \alpha^2)dt^2 + 2\beta_i dt dx^i + \gamma_{ij}dx^i dx^j, \tag{4.1}$$

where β^i is the shift 3-vector, α is the lapse function, and γ_{ij} is the spatial 3-metric. In isotropic coordinates, the background metric of a static and spherically-symmetric configuration can be written as

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = -\alpha^2 dt^2 + \psi^4 f_{ij}dx^i dx^j, \tag{4.2}$$

where ψ is the conformal factor and f_{ij} is the flat spatial 3-metric. In the spherically-symmetric and static limit, Einstein's equations for the CFC metric read

$$\Delta\psi = -2\pi\psi^5 E, \tag{4.3}$$

$$\Delta(\alpha\psi) = 2\pi\alpha\psi^5\left[E + 2S\right], \tag{4.4}$$

where Δ is the Laplacian operator with respect to the flat 3-metric. In this case $\beta_i = 0$ and $\gamma_{ij} = \psi^4 f_{ij}$. The energy-momentum content couples to the spacetime geometry through the projections of the energy-momentum tensor, $T_{\mu\nu}$, onto the $3+1$ foliation

$$E \equiv \alpha^2 T^{00}, \qquad S_i \equiv -(T_{0i} - T_{ij}\beta^j), \tag{4.5}$$

$$S_{ij} \equiv T_{ij}, \qquad S \equiv S_{ij}\gamma^{ij}. \tag{4.6}$$

Let us consider a perfect fluid. Its energy-momentum tensor is given by

$$T^{\mu\nu} = \rho h u^\mu u^\nu + P g^{\mu\nu} \tag{4.7}$$

where ρ is the rest mass density, P is the pressure, u^μ is the 4-velocity, $h \equiv 1 + \epsilon + P/\rho$ is the specific enthalpy, and ϵ is the specific internal energy. It is useful to define the energy density as $e \equiv \rho(1+\epsilon)$.

If we consider the Bianchi identities and the conservation of the number of baryons (continuity equation) the GRHD equations written in the coordinate basis read

$$\frac{1}{\sqrt{\gamma}}\partial_t\left[\sqrt{\gamma}D\right] + \frac{1}{\sqrt{\gamma}}\partial_i\left[\sqrt{\gamma}Dv^{*i}\right] = 0, \tag{4.8}$$

$$\frac{1}{\sqrt{\gamma}}\partial_t\left[\sqrt{\gamma}S_j\right]+\frac{1}{\sqrt{\gamma}}\partial_i\left[\sqrt{\gamma}S_j v^{*i}\right]+\alpha\partial_i P=\frac{\alpha\rho h}{2}u^\mu u^\nu\partial_j g_{\mu\nu}, \tag{4.9}$$

$$\frac{1}{\sqrt{\gamma}}\partial_t\left[\sqrt{\gamma}E\right]+\frac{1}{\sqrt{\gamma}}\nabla_i\left[\sqrt{\gamma}\left(Ev^{*i}+\alpha P v^i\right)\right]= \\ =\alpha^2\left(T^{\mu 0}\partial_\mu\ln\alpha-T^{\mu\nu}\Gamma^0_{\mu\nu}\right), \tag{4.10}$$

where γ is the determinant of the three-metric, $\Gamma^\lambda_{\mu\nu}$ is the Christoffel symbol associated with the four-metric, $W=1/\sqrt{1-v_i v^i}$ is the Lorentz factor and the conserved quantities are

$$D=\rho W, \qquad S_j=\rho h W^2 v_j, \qquad E=\rho h W^2-P. \tag{4.11}$$

The Eulerian and "advective" velocities are, respectively,

$$v^i=\frac{u^i}{W}+\frac{\beta^i}{\alpha}, \qquad\qquad v^{*i}=\frac{u^i}{u^0}=\alpha v^i-\beta^i. \tag{4.12}$$

Let us consider a solution of the hydrodynamics equations that is in equilibrium ($\partial_t=0$) and is static ($v^i=0$). In this case Eq. (4.9) reads

$$\frac{1}{\rho h}\partial_i P=-\partial_i\ln\alpha\equiv G_i, \tag{4.13}$$

where G_i is the gravitational acceleration, in the Newtonian limit, whose only non-zero component is $G_r\equiv G$. The solution of Eq. (4.13) corresponds to the unperturbed state or background solution.

Let us consider linear adiabatic perturbations of the hydrodynamics equations with respect to the background equilibrium configuration. We denote Eulerian perturbations of the different quantities with δ, e.g. for the rest-mass density ρ, the Eulerian perturbation is $\delta\rho$. The linearised equations are obtained by substituting $\rho\to\rho+\delta\rho$, and so forth, in Eqs. (4.8-4.10). Note that ρ, P, etc, correspond to the background value in the linearised equations. Furthermore, we consider the Cowling approximation, i.e. we do not take into account perturbations of the metric ($\delta\alpha=\delta\psi=\delta\beta^j=0$). We discuss below the impact of this approximation in our results.

We denote as ξ^i the Lagrangian displacement of a fluid element with respect to its position at rest. Its value is related to the advective velocity as

$$\partial_t\xi^i=\delta v^{*i}. \tag{4.14}$$

The Lagrangian perturbation of any quantity, e.g. ρ, is related to the Eulerian perturbations as

$$\Delta\rho=\delta\rho+\xi^i\partial_i\rho. \tag{4.15}$$

The linearised version of Eq. (4.8) and (4.9) are

$$\frac{\Delta\rho}{\rho} = -\left(\partial_i\xi^i + \xi^i\partial_i \ln\sqrt{f}\right), \tag{4.16}$$

$$\rho h\,\partial_t\delta v_i + \alpha\partial_i\delta P = -\delta\left(\rho h\right)\partial_i\alpha\,, \tag{4.17}$$

where $\sqrt{f}$ is the determinant of the flat 3-metric. We use spherical coordinates $\{r,\theta,\varphi\}$ in the following, in which $\sqrt{f} = r^2\sin\theta$. Since we are considering adiabatic perturbations, Eq. (4.10) does not add additional information to the problem.

The condition of adiabaticity of the perturbations implies that

$$\frac{\Delta P}{\Delta\rho} = \left.\frac{\partial P}{\partial\rho}\right|_{\rm adiabatic} = hc_{\rm s}^2 = \frac{P}{\rho}\Gamma_1, \tag{4.18}$$

where $c_{\rm s}$ is the relativistic speed of sound and Γ_1 is the adiabatic index. This allows us to write

$$\delta\left(\rho h\right) = \left(1+\frac{1}{c_s^2}\right)\delta P - \rho h\,\xi^i\mathcal{B}_i\,, \tag{4.19}$$

where

$$\mathcal{B}_i \equiv \frac{\partial_i e}{\rho h} - \frac{1}{\Gamma_1}\frac{\partial_i P}{P}, \tag{4.20}$$

is the relativistic version of the Schwarzschild discriminant. Since the background is spherically symmetric, the only non-zero component is $\mathcal{B}_r \equiv \mathcal{B}$.

The radial and angular parts of equation (4.17) are given by

$$\rho h\,\psi^4\alpha^{-2}\frac{\partial\xi^r}{\partial t^2} + \partial_r\delta P = \delta\left(\rho h\right)G \tag{4.21}$$

$$\rho h\,\psi^4\alpha^{-2}r^2\frac{\partial\xi^\theta}{\partial t^2} + \partial_\theta\delta P = 0, \tag{4.22}$$

$$\rho h\,\psi^4\alpha^{-2}r^2\sin^2\theta\frac{\partial\xi^\varphi}{\partial t^2} + \partial_\varphi\delta P = 0, \tag{4.23}$$

where we have used that, in the coordinate basis, the covariant components of the velocity are given by $\delta v_r = \psi^4\delta v^r$, $\delta v_\theta = r^2\psi^4\delta v^\theta$ and $\delta v_\varphi = r^2\sin^2\theta\psi^4\delta v^\varphi$.

We perform an expansion of the perturbations with a harmonic time dependence of frequency σ and a spherical-harmonic expansion for the angular dependence

$$\delta\hat{P} = \delta P\; Y_{lm}e^{-i\sigma t}, \tag{4.24}$$

$$\xi^r = \eta_r\; Y_{lm}e^{-i\sigma t}, \tag{4.25}$$

$$\xi^\theta = \eta_\perp\;\frac{1}{r^2}\partial_\theta Y_{lm}e^{-i\sigma t}, \tag{4.26}$$

$$\xi^\varphi = \eta_\perp\;\frac{1}{r^2\sin^2\theta}\partial_\varphi Y_{lm}e^{-i\sigma t}. \tag{4.27}$$

Here we only consider polar perturbations, because axial (torsional) oscillations are not possible in non-rotating stars. The quantities $\eta_r, \eta_\perp$ and the scalar perturbations with the hat, i.e. $\delta\hat{P}$ are only function of the radial coordinate. For $l \neq 0$, by inserting the spherical-harmonic expansion into equations (4.21)-(4.22) we obtain:

$$-\sigma^2 \rho h\, \psi^4 \alpha^{-2} \eta_r + \partial_r \delta\hat{P} = \delta\left(\rho h\right) G, \tag{4.28}$$

$$-\sigma^2 \rho h\, \psi^4 \alpha^{-2} \eta_\perp + \delta\hat{P} = 0. \tag{4.29}$$

From Eq. (4.29) we get

$$\delta\hat{P} = q\sigma^2 \eta_\perp\,, \tag{4.30}$$

where for simplicity we have defined $q \equiv \rho h\, \alpha^{-2}\psi^4$.

Using Eqs. (4.30) and (4.19) to simplify Eqs. (4.16) and (4.28) we obtain

$$\partial_r \eta_r + \left[\frac{2}{r} + \frac{1}{\Gamma_1}\frac{\partial_r P}{P} + 6\frac{\partial_r \psi}{\psi}\right]\eta_r + \frac{\psi^4}{\alpha^2 c_s^2}\left(\sigma^2 - \mathcal{L}^2\right)\eta_\perp = 0\,, \tag{4.31}$$

$$\partial_r \eta_\perp - \left(1 - \frac{\mathcal{N}^2}{\sigma^2}\right)\eta_r + \left[\partial_r \ln q - G\left(1 + \frac{1}{c_s^2}\right)\right]\eta_\perp = 0, \tag{4.32}$$

where $\mathcal{N}$ is the relativistic Brunt-Väisälä frequency defined as

$$\mathcal{N}^2 \equiv \frac{\alpha^2}{\psi^4} G^i \mathcal{B}_i = \frac{\alpha^2}{\psi^4}\mathcal{B}G\,, \tag{4.33}$$

and $\mathcal{L}$ is the relativistic Lamb frequency defined as

$$\mathcal{L}^2 \equiv \frac{\alpha^2}{\psi^4} c_s^2 \frac{l\,(l+1)}{r^2}\,. \tag{4.34}$$

To simplify the discussion we define the coefficients A, B, C and D in Eqs. (4.31) and (4.32) such that the equations take the form

$$\partial_r \eta_r = A\eta_r + B\eta_\perp, \tag{4.35}$$

$$\partial_r \eta_\perp = C\eta_r + D\eta_\perp. \tag{4.36}$$

4.3.1 Plane-wave limit

In the plane-wave approximation $\eta_r, \eta_\perp \sim e^{ikr}$, we can write the following dispersion relation

$$-k^2 - ik(A + D) + AD - BC = 0. \tag{4.37}$$

The real part of k is non zero, i.e. locally there exist plane-wave solutions, if and only if

$$4BC < (A - D)^2. \tag{4.38}$$

Therefore, a sufficient condition to have a real part of k is that $BC < 0$ and hence

$$(\sigma^2 - \mathcal{L}^2)(\sigma^2 - \mathcal{N}^2) > 0 \quad \rightarrow \quad Re(k) \neq 0. \tag{4.39}$$

In cold NSs, typically $\mathcal{L}^2 >> \mathcal{N}^2$, thus the solutions can be either $\sigma^2 > \mathcal{L}^2 > \mathcal{N}^2$ (acoustic modes) or $\sigma^2 < \mathcal{N}^2 < \mathcal{L}^2$ (g-modes). However, in the physical core-collapse situation considered in this chapter we deal with a new-born PNS. Therefore, the PNS is still hot and surrounded by an extended mantle and those simplifying assumptions do not hold.

4.3.2 G-modes limit

G-modes can appear in regions of the star where buoyancy acts as a restoring force. In those regions the Brunt-Väisälä frequency is such that $\mathcal{N}^2 > 0$, i.e. the regions are stable to convective instabilities. It is possible to obtain approximate solutions for the g-modes in a star by neglecting sound waves in the system. The calculation of g-modes in this approximation can be used to have a rough idea of what are the typical g-mode frequencies of PNSs. It also serves as a basis to identify which modes belong to the g-mode class when the full analysis without approximation is performed. It is possible to remove acoustic waves from the system by taking the limit of Eqs. (4.31) and (4.32) when the speed of sound tends to infinity, $c_s^2 \rightarrow \infty$. In this limit the equations read

$$\partial_r \eta_r + \left[\frac{2}{r} + 6\frac{\partial_r \psi}{\psi}\right] \eta_r - \frac{l(l+1)}{r^2}\eta_\perp = 0\,, \tag{4.40}$$

$$\partial_r \eta_\perp - \left(1 - \frac{\mathcal{N}^2}{\sigma^2}\right) \eta_r + \left[\partial_r \ln q - G\right] \eta_\perp = 0. \tag{4.41}$$

4.3.3 Acoustic-modes limit

Similarly to the previous case, it is also possible to remove buoyancy from the system by considering only modes supported by sound waves, i.e. p-modes. The acoustic limit is reached by setting $\mathcal{B} = 0$, and therefore, $\mathcal{N}^2 = 0$. The resulting system of equations reads,

$$\partial_r \eta_r + \left[\frac{2}{r} + \frac{\partial_r e}{\rho h} + 6\frac{\partial_r \psi}{\psi}\right] \eta_r + \frac{\psi^4}{\alpha^2 c_s^2}\left(\sigma^2 - \mathcal{L}^2\right) \eta_\perp = 0\,, \tag{4.42}$$

$$\partial_r \eta_\perp - \eta_r + \left[\partial_r \ln q - G\left(1 + \frac{1}{c_s^2}\right)\right] \eta_\perp = 0. \tag{4.43}$$

4.3.4 Boundary conditions

We impose boundary conditions at the shock location. The shock is a sonic point in which the flow decelerates from a supersonic regime (outside) to a subsonic regime (inside). Therefore, all the characteristic curves in the inner part of the shock, where we need to impose boundary conditions, are pointing inwards; a wave propagating outwards from the inside will stall when reaching the shock location. In other words, any (radial) perturbation at the shock location should have zero displacement

$$\xi_r|_{\text{shock}} = 0 \quad \rightarrow \quad \eta_r|_{\text{shock}} = 0. \tag{4.44}$$

At the origin ($r = 0$) it is sufficient to impose regularity [see Reisenegger and Goldreich 1992]

$$\eta_r|_{r=0} = l\,\eta_\perp|_{r_0} \propto r^{l-1}. \tag{4.45}$$

4.3.5 Eigenmode computation

The procedure to compute the eigenmodes ("modes" hereafter) entails the integration of Eqs. (4.31) and (4.32) from the centre of the PNS to the shock for different values of σ. Initial data at the origin is somewhat arbitrary as long as it fulfils Eq. (4.45). Those values of σ such that η_r vanishes at the shock correspond to eigenvalues of the system. Accurate eigenvalues can be obtained by finding the roots of $\eta_r|_{\text{shock}}(\sigma) = 0$ (e.g. using the bisection method), where $\eta_r|_{\text{shock}}(\omega)$ is the value of $\eta_r|_{\text{shock}}$ after integrating for a given value of σ. The integration of Eqs. (4.31) and (4.32) is performed by means of a backward Euler scheme. The same numerical procedure is employed when computing the approximate g-modes, Eqs. (4.40) and (4.41), and the approximate p-modes, Eqs. (4.42) and (4.43). Hereafter, we refer as "approximate modes" to the solutions of the latter two approximate systems, to distinguish them from the "complete modes" (usually just referred as "modes") resulting from the complete system of equations. Fig. 4.3 shows an example of the absolute value of both, the radial and the perpendicular components of the eigenfunction corresponding to a frequency of 2608 Hz calculated 1 s after core bounce. In this case both components have several nodes inside the domain. The number of nodes is used as a basis for the identification and classification of the different modes, as we show in the next sections.

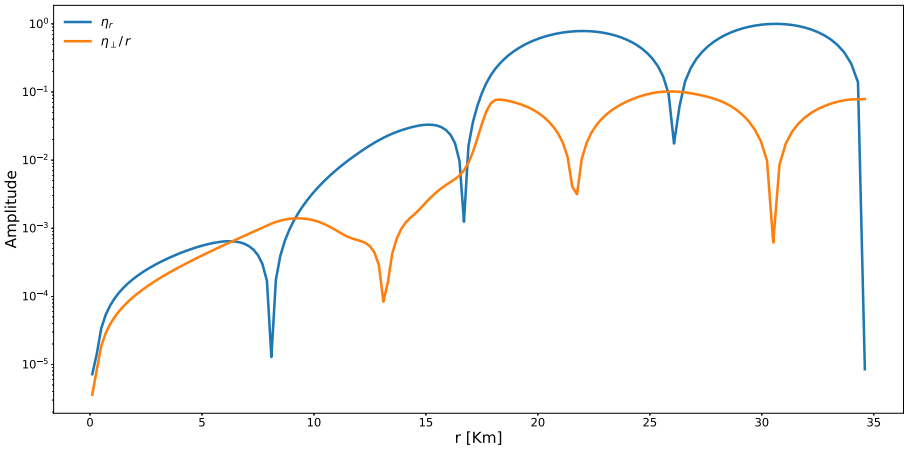


Figure 4.3 Absolute values of the radial component η_r (blue line) and of the perpendicular component $\eta_\perp$ (orange) of one example mode computed ≈ 1 s after core bounce. The mode frequency is 2608 Hz. The y-axis represents the normalised amplitude in logarithmic scale, while the x-axis is the radius in km which ends at the shock.

4.3.6 Energy and radiated power

In the Newtonian limit, the energy stored in a mode with a certain amplitude can be approximated as

$$E = \frac{\sigma^2}{2} \int_0^{r_{\rm shock}} \rho\, r^2 [\eta_r^2 + l(l+1)\frac{\eta_\perp^2}{r^2}] dr\,. \tag{4.46}$$

The integrand can be identified as the energy density $\mathcal{E}$. Following Thorne [1969] it is possible to compute the total radiated power[2] in gravitational waves by each mode,

$$P_g = \frac{1}{8\pi} \frac{(l+1)(l+2)}{(l-1)l} \left[\frac{4\pi\sigma^{l+1}}{(2l+1)!!} \int_0^{r_{\rm shock}} \delta\hat{\rho}\, r^{l+2} dr \right]^2, \tag{4.47}$$

where

$$\delta\hat{\rho} \approx \rho \left(\frac{\mathcal{N}^2}{G}\eta_r + \frac{\sigma^2}{c_s^2}\eta_\perp \right). \tag{4.48}$$

4.4 Mode classification

In addition to the simple classification of the modes as a function of the frequency, our goal is to classify the modes in a way which allows us to identify the contribution of each mode to the GW emission. In this section we focus on the

[2]Note the change in the expression due to the use of coordinate basis instead of orthonormal basis.

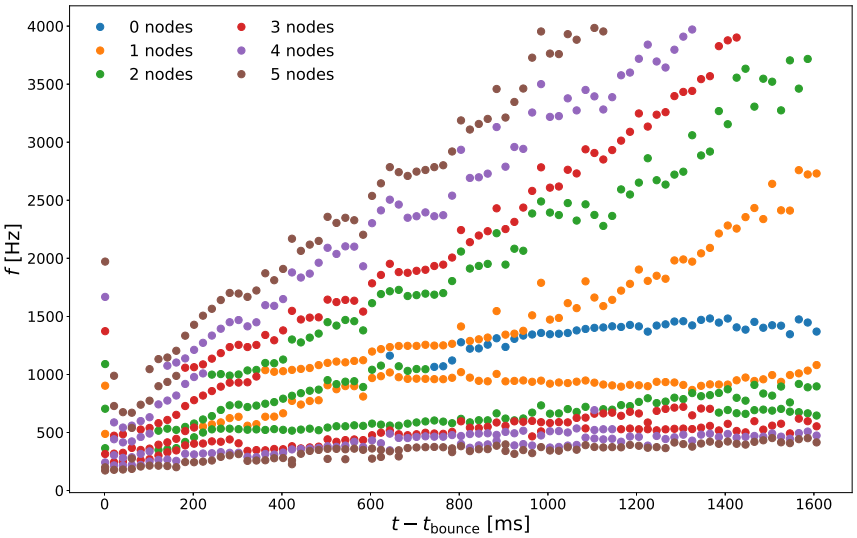


Figure 4.4 Post-bounce time evolution of the frequency of the eigenmodes of the system. Modes with the same number of nodes are represented with the same colour. Only modes with less than five nodes are represented.

$l = 2, m = 0$ modes. Modes for other values of l look qualitatively similar and will be discussed in the next section. Non-axisymmetric modes with $m \neq 0$, have the same frequency as modes with $m = 0$, for the same value of l, and are not considered.

4.4.1 Number of nodes

Our first attempt to classify the modes is according to their number of nodes. We define the number of nodes as the number of sign changes of the radial function η_r. However, in regions where the value of η_r is small, numerical discretisation errors of the equations can induce small fluctuations of the eigenfunction that may lead to a miscount of the number of nodes. Therefore, we try to minimise as much as possible this error in our counting algorithm; for example we do not count nodes as different unless they are more than a few numerical cells apart, to avoid possible high-frequency numerical noise. However, there could still be node miscounts in a few cases. While this issue could be solved by counting the number of nodes manually, our aim is to build a fully automatised, and reliable, algorithm for the mode classification.

Fig. 4.4 shows the post-bounce evolution of the mode frequencies, with modes classified in different colours according to their number of nodes. Some interesting conclusions can be extracted from this figure. Our simple classification in terms of the number of nodes allows to tell apart modes and to follow the temporal evolution of their frequencies during the simulation. Note that other classification methods are possible, as those based on the continuity of mode frequency or similar characteristics. After $\sim$ 200 ms, at which time the accretion rate drops [see Cerdá-Durán et al. 2013, for details] different types of modes start to split away, and it becomes clear that there is a class of modes with a rapidly-increasing frequency and another class in which the frequency barely changes with time or even decreases. As we show in the next sections, these two classes correspond approximately to p-modes and g-modes, respectively. The jumps in frequency between consecutive times are due to changes in the accretion-shock position, which is mainly due to SASI. We call fundamental mode (or f-mode), ${}^{l}f$, the mode with zero radial nodes. In Fig. 4.4 this mode is shown in blue and is clearly distinguishable starting from $\sim$ 800 ms, with a frequency of around 1000 Hz. Once the fundamental mode is identified it becomes apparent that it separates two classes of modes: one class above the f-mode, in which the number of nodes increases with increasing frequency and another class of modes below the f-mode in which the number of nodes increases with decreasing frequency. This is the expected behaviour of p-modes and g-modes, respectively. Before $\sim$ 800 ms there is no mode with zero radial nodes to be identified as f-mode. The two classes of modes seem to mix and cross. At each crossing there is a change on the number of nodes. This behaviour, the so-called avoided crossing of frequencies, is typical in linear analysis of oscillations, when modes of different nature (in this case p-modes and g-modes) have similar frequencies [see e.g. Stergioulas 2003]. At these frequencies, although not exclusively, we expect to also have hybrid modes, which are neither p-modes nor g-modes.

4.4.2 Mode identification

The previous analysis suggests that in the evolution of the new-born PNS there exist at least two types of modes. To check if these classes correspond to the theoretical separation in g-modes and acoustic-modes (p-modes), we compute them in the approximations presented in Sections 4.3.2 and 4.3.3 and calculate the number of nodes in each case. We label them as ${}^{l}g_n$ and ${}^{l}p_n$, respectively, with index n being the number of nodes. Fig. 4.5 shows the results of both approximate g-modes (crosses) and p-modes (stars). It is clear that for the

approximate g-modes the number of nodes increases as the frequency decreases. In contrast, for the approximate p-modes this relation is the opposite, increasing the number of nodes with increasing frequency. This behaviour supports the identification hypothesis we have done of the complete modes in the previous section.

We turn next to analyse the time evolution of the frequency of both classes of approximate modes. The fundamental approximate g-mode, labelled as ${}^2g_0$, has a frequency in the range $f \in [500, 1000]$ Hz, varying non-monotonically but slowly between the hydrodynamical bounce and the black hole formation at the end of the simulation. The rest of g-modes have similar evolution albeit at lower frequencies. This is the expected behaviour of g-modes living inside the PNS, whose frequencies trace the surface gravity of the region generating the modes. The fundamental approximate p-mode ${}^2p_0$ is the one with the lowest frequency. As the system evolves all approximate p-mode frequencies increase almost monotonically. This kind of modes correspond to standing sound waves trapped between the surface of the PNS and the shock. As the shock contracts during the evolution, the radial extent of this cavity decreases and the frequency of p-modes increases. The short time-scale variability of those modes is related to variations in the shock location due to the sloshing motions associated with the SASI. Comparing both types of approximate modes one observes several frequency crossings between modes of each kind. This fact explains the hybridisation of the complete modes that was described in the previous section. We check this hybridisation in detail below.

In view of these results, we have devised a method for classifying the complete modes as p-modes or g-modes: we define the n^{th} p-mode/g-mode, for a given time, as the mode with the largest/lowest frequency with n nodes. Those modes not falling in either category are classified as hybrid modes (h-modes, hereafter). The result of this classification is shown in Fig. 4.6. It is interesting to compare this figure with Fig. 4.5. There is clearly a good match between the classification procedure of the complete eigenvalue problem and the frequencies of the approximate g-modes and p-modes calculations, specially at frequencies well above or below the fundamental-mode frequency. This makes us confident that our procedure is accurately classifying different modes. We also note that a careful look to both figures shows a tendency for complete modes to have somewhat larger frequencies than their corresponding approximate modes. As expected, hybrid modes appear mainly at the crossing of g-modes and p-modes, although some may appear at different frequencies and may persist for the entire evolution (see Fig. 4.6).

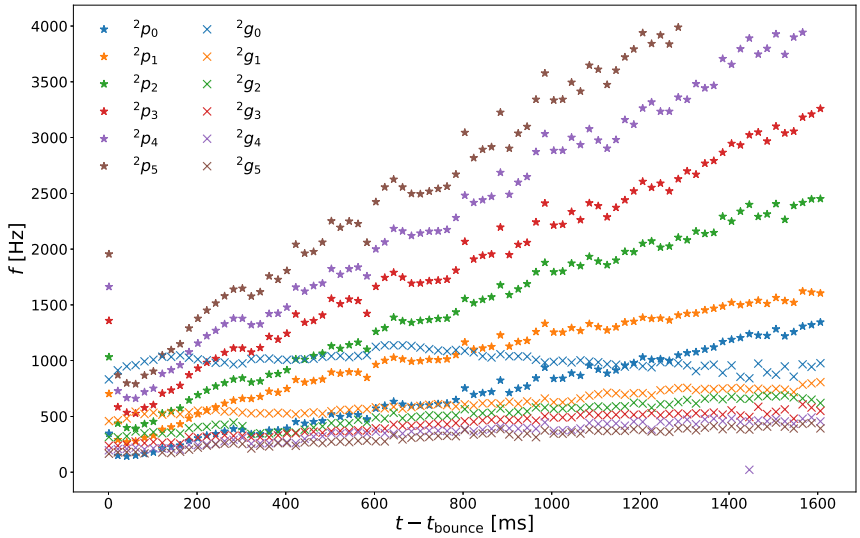


Figure 4.5 Time-frequency diagram of the approximate modes computed in the g-mode limit (crosses) and in the p-mode limit (stars). Modes with the same number of nodes are represented with the same colour. Only modes with less than five nodes are represented.

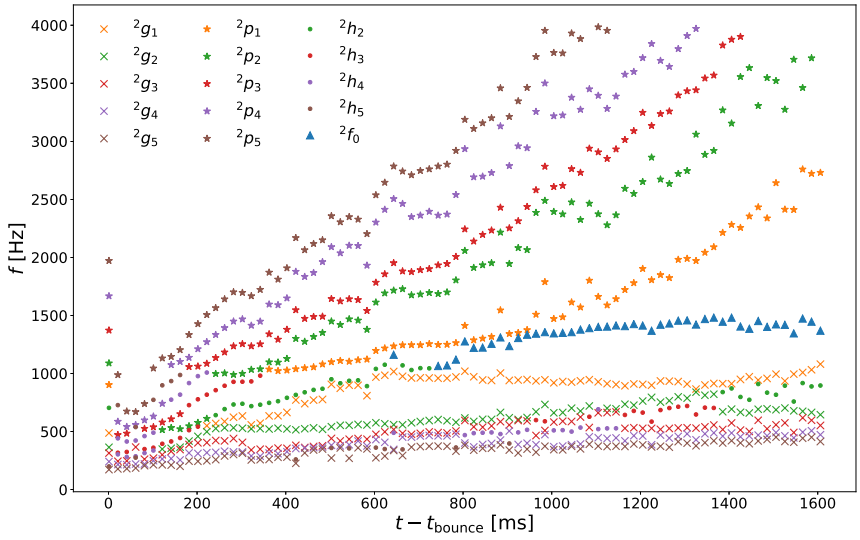


Figure 4.6 Time-frequency diagram of the modes of the complete problem separated in g-modes (crosses), p-modes (stars), f-mode (triangles) and h-modes (dots). Modes with the same number of nodes are represented with the same colour. Only modes with less than five nodes are represented.

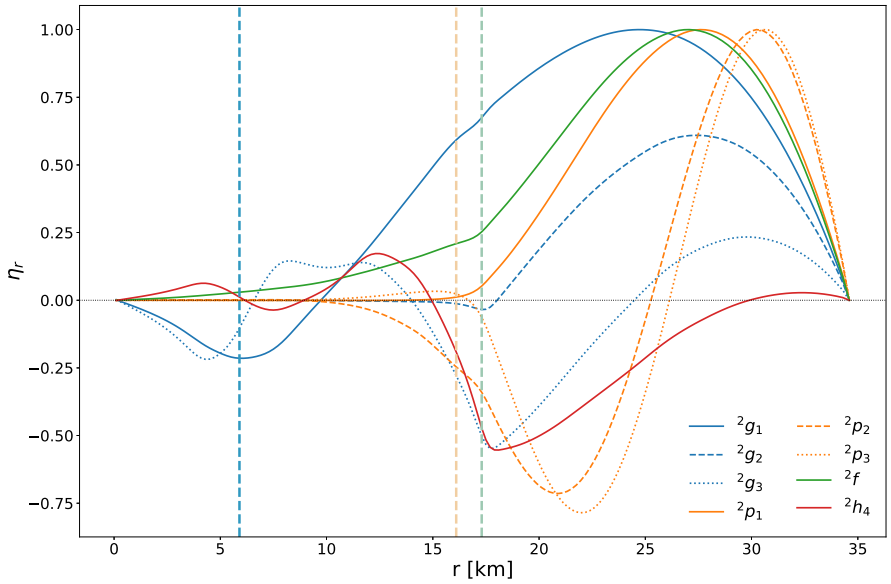


Figure 4.7 Radial component (η_r) of the first few modes of each class, 1 s after bounce. g-modes are represented in blue, p-modes in orange, the f-mode in green, and the h-modes in red. The vertical lines indicate the radial position of the three parts of the star as introduced in Fig. 4.2. (See main text for details.)

Once the modes have been identified it is possible to examine in detail their corresponding eigenfunctions. Figs. 4.7, 4.8 and 4.9 show, respectively, the radial component, the perpendicular component, and the energy density of a representative set of the complete modes. Attending to the shape of the modes, one can see that modes within a class are somewhat similar, but with different number of nodes. The vertical dashed lines in Figs. 4.7, 4.8 and 4.9 represent the radial position of the different parts of the star we are considering. The blue line is the limit of the iron core $r_{\rm cold}$. The orange line is the position of the neutrinosphere r_ν and limits the lower radius of the PNS surface. Finally, the green line represents the radius at which the density is lower than 10^{11} g cm^{-3} and marks the position of the upper radius of the PNS surface. The relative position of each radius at the time shown in these figures is the same as in Fig. 4.2.

The energy density shown in Fig. 4.9 offers a complementary view to the eigenfunctions, which is useful to further interpret the mode classification. All the energy density of the ^2f mode is located inside the star, between the core and the surface, which we labelled as region II in Fig. 4.2. The mode $^2h_4$ is

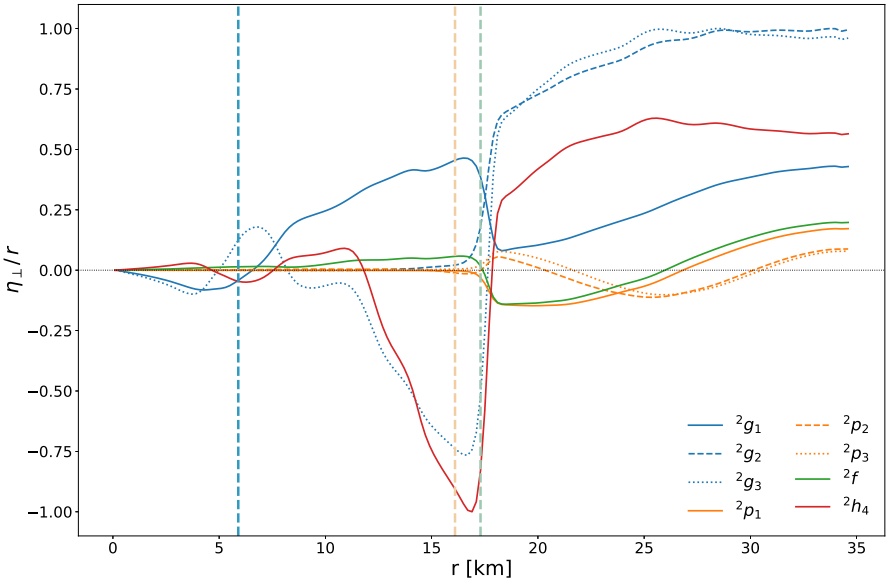


Figure 4.8 Same as Fig. 4.7 but for the perpendicular component ($\eta_\perp$).

also confined in the same region, differentiating itself from the g-modes and the p-modes. The g-modes with an odd number of nodes have their energy density confined inside the star, with part of it distributed in regions I and II, while the p-modes with an odd number of nodes have it distributed in region III. The g-mode $^2g_2$ displays a peculiar behaviour as its energy density is confined in region III while the rest of g-modes belong to region I. Also the p-mode $^2p_2$ differs from the others as its energy density is more similar to the h-mode $^2h_4$ than to the other p-modes. It is currently unclear what can be the explanation for the distribution of the energy density among the different regions. Understanding this issue deserves a future study as it could improve our classification procedure.

To conclude this section we show in Fig. 4.10 the 2D representation of both components of the modes $^2p_1$ and $^2g_1$. The black dashed lines in this figure indicate the zeros of the corresponding eigenfunction. The blue radial line and the orange radial line represent the limits of the PNS core and of the neutrinosphere, respectively. The larger values of the amplitude of each mode are shown in yellow tones, while the lower values are shown in dark red colour. The maximum amplitude of the p-mode (left panels) is concentrated in the exterior part of the PNS, which is consistent with the interpretation of p-modes being standing sound waves trapped between the PNS surface and the shock.

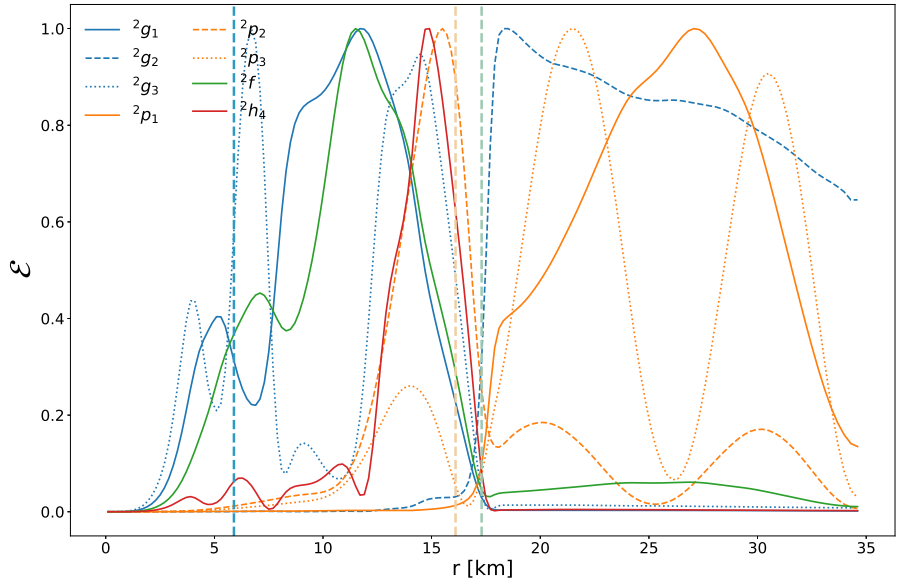


Figure 4.9 Same as Fig. 4.7 but for the energy density ($\mathcal{E}$).

However, there is an additional non-negligible component in the interior of the PNS, due to the coupling of the acoustic waves with buoyantly-responding layers in the interior. The right panels show the components of a g-mode that has its maximum at the PNS interior. At this frequency the convectively stable regions in the interior of the PNS allow for the formation of g-modes. However these g-modes extend outside of the PNS up to the shock. The coupling of the PNS surface with the shock is mediated by nodeless sound waves propagating in this region.

4.5 Gravitational-wave emission

4.5.1 Comparison in the frequency domain

In the previous sections we have computed the eigenmodes of the system formed by the PNS and the shock at any given time in the simulation. It is expected that, if we perturb the system, some of those modes will be excited and gravitational waves will be emitted at those particular frequencies. These perturbations are indeed present in the simulation: the region between the PNS surface and the shock is subject to competing instabilities, such as the SASI and convection, that break spherical symmetry. In this region down-flowing plumes of cooled

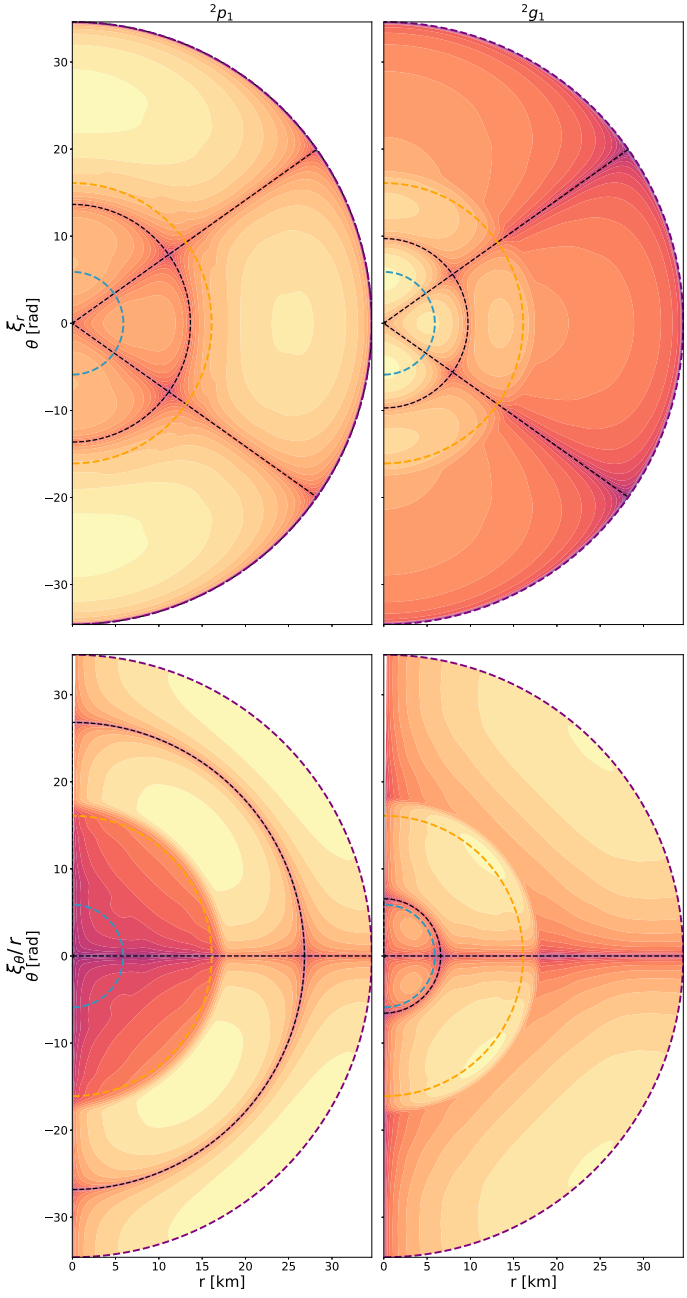


Figure 4.10 Polar-coordinates representation of both components of the modes $^2p_1$ and $^2g_1$.

matter hit the PNS surface exciting the eigenmodes of the system. All these perturbations translate into the gravitational-wave signal (see Fig. 4.1, whose frequency spectrum is expected to be correlated with the frequencies of the modes). In order to compare both we overplot the time-frequency distribution of the modes with the spectrogram of the GW signal obtained from the simulation. For this comparison we consider only modes with $l = 2$ and $l = 4$. The numerical simulation of Cerdá-Durán et al. [2013] was performed assuming symmetry with respect to the equatorial plane, therefore modes with odd l are not present, and are not considered here. Moreover, $l = 0$ modes were not considered in the analysis, and will be considered elsewhere. We do not take into account $l = 6$ or higher modes because those typically have less energy stored, and hence will produce weaker gravitational-wave signals, than modes with lower l.

The results are shown in the spectrograms of Fig. 4.11. As a general conclusion, all of the features in the spectrograms have a corresponding eigenmode tracing its evolution closely. But the inverse is not true, not every eigenmode computed has a trace in the spectrogram. This is expected, since not every mode has to be excited during the evolution or, even if excited, not every mode can have a sufficiently high amplitude to leave an imprint in the GW signal.

From Fig. 4.11 it seems clear that the GW emission cannot be explained with $l = 2$ modes alone and some $l = 4$ features are also needed. This is particularly true for the ${}^4f$ mode (lower-right panel, in blue colour). During the first half a second after the bounce, g-modes are the dominant GW emission mechanism (upper panels). This is consistent with previous core-collapse simulations, in which spectrogram features during the first half second (previous to the supernova explosion) were interpreted as g-modes [see e.g. Müller, Janka, and Marek 2013]. At later times, a component of the gravitational-wave signal with increasing frequency appears. This component is perfectly fitted by ${}^2p_n$ modes (upper-right panel) at frequencies larger than $\sim$ 1500 Hz. ${}^4p_n$ modes do not appear to contribute to the signal, and were probably not excited during the simulation. Those p-modes were identified as SASI modes by Cerdá-Durán et al. [2013] and Kuroda, Kotake, and Takiwaki [2016], because their frequency traces the sloshing motions of the shock. This is expected because the acoustic waves forming the p-modes play a main role in the advective-acoustic cycle responsible for the SASI [Foglizzo and Tagger 2000, Blondin, Mezzacappa, and DeMarino 2003]. Below $\sim$ 1500 Hz, the GW emission seems to be dominated by g-modes. The most relevant ones are g-modes with $n \leq 3$ and $l = 2$ (upper-left panel). Those with larger l do not appear to leave an imprint in the signal.

In general, it is quite reassuring that all lowest-order modes (in terms of l and n) leave an imprint in the spectrograms. This is consistent with the interpretation that perturbations excite mainly low-order modes and higher-order modes do not contribute to the signal. Furthermore, this observation also helps to decide which modes are relevant for the GW emission, irrespective of the waveforms computed in the simulations. The only feature in the spectrogram that cannot be explained by our eigenmode analysis is a component whose frequency decreases in time from ~ 1000 Hz at about 1.2 s after bounce to 0 Hz at the time of BH formation. This feature was interpreted in Cerdá-Durán et al. [2013] as a quasi-radial mode ($l = 0$), which is not possible to study within our approach. The Cowling approximation does not appear to introduce large errors in the eigenmode computation. Many of the modes show a small mismatch with respect to the spectrogram, typically a small shift to larger frequencies, which is probably caused by the Cowling approximation. In the near future we plan to relax this approximation to analyse whether a closer agreement is found and also to study the behaviour of quasi-radial modes.

4.5.2 Gravitational-wave radiation efficiency

In the previous section we have compared our eigenmode analysis with the GW signal only in terms of frequency evolution. Unfortunately, the eigenmode analysis does not allow us to predict what would be the amplitude of each mode in the GW signature. This amplitude depends on two factors: the energy stored in each mode and the efficiency of each mode to emit gravitational waves. In the previous section we have also found indications that most of the energy of the perturbations is stored in the lower-order modes. The energy E of a mode can be computed from Eq. (4.46) and the radiated power P is given by Eq. (4.47). Both quantities are proportional to the (unknown) eigenmode amplitude. The ratio between both quantities

$$\tau_{\rm GW} = \frac{E}{P}, \tag{4.49}$$

is the gravitational-wave timescale, which gives an idea of the characteristic time in which the energy of a mode will be radiated in GWs. Modes with smaller value of $\tau_{\rm GW}$ are expected to be more efficient emitters of GWs and will produce a stronger imprint in the signal. We can also define the GW emission efficiency as

$$(\text{GW efficiency}) = \frac{P}{Ef}, \tag{4.50}$$

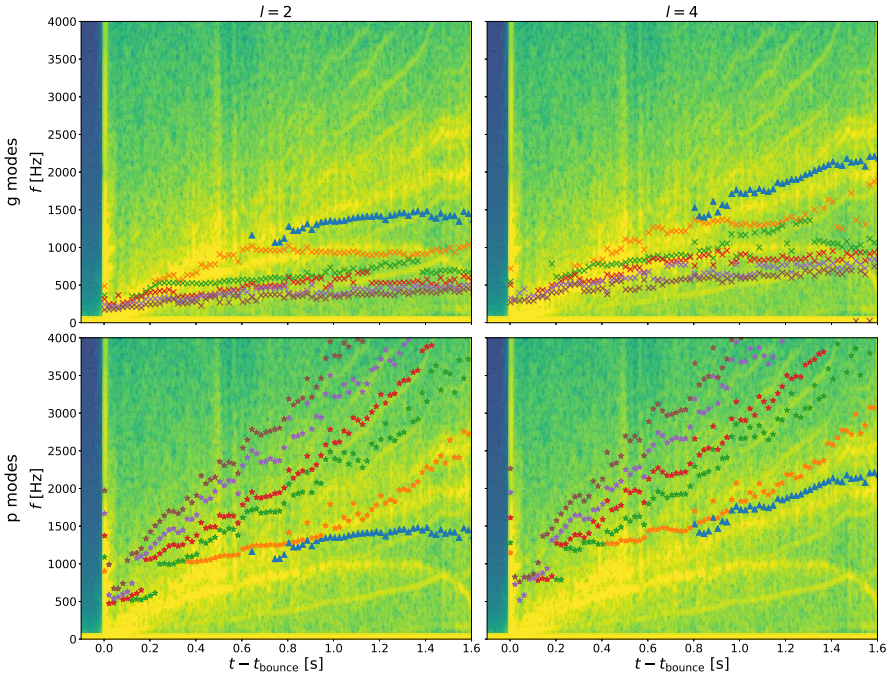


Figure 4.11 Comparison of the spectrogram from the numerical simulation of Cerdá-Durán et al. [2013] with the time-frequency distribution of the $l = 2$ and $l = 4$ modes. Colour coded is the GW power, with yellow indicating the highest values. The upper panels show the g-modes (crosses) and the lower panels the p-modes (stars). Left and right panels show the $l = 2$ and $l = 4$ modes, respectively. The number of nodes are represented in different colours: blue (0), orange (1), green (2), red (3), violet (4), and brown (5).

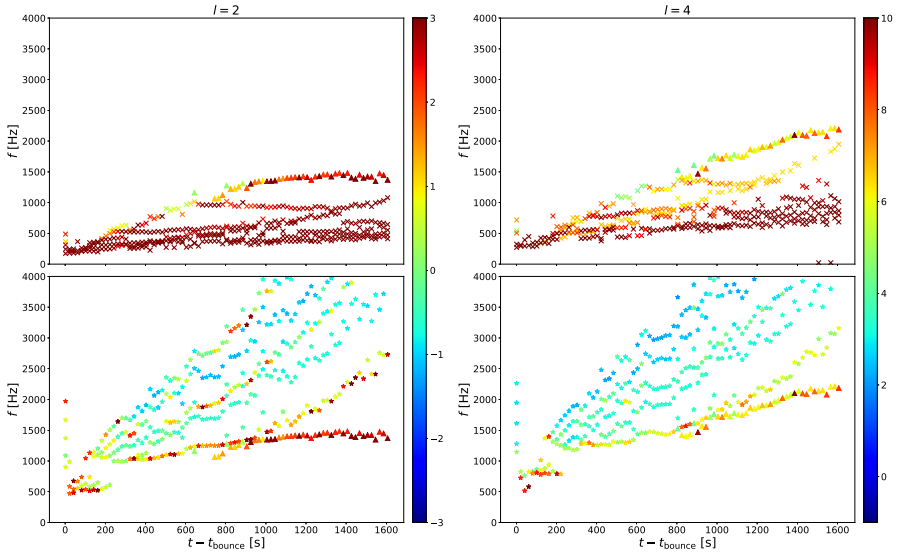


Figure 4.12 Time-frequency diagram of the modes of the complete problem separated by $\tau_{\rm GW}$. Modes with lower values of $\tau_{\rm GW}$ are coloured in blue and the values scale to the larger values in red. The upper row corresponds to g-modes (crosses), while the lower row corresponds to p-modes (stars). The fundamental mode is represented with triangles. The colour bar shows the logarithm of $\tau_{\rm GW}$, in seconds.

where f is the frequency of the mode. This equation gives an idea of the fraction of the mode energy radiated in GWs per oscillation cycle.

Fig. 4.12 shows the time-frequency diagram of $\tau_{\rm GW}$ for each mode. The modes with lower value of $\tau_{\rm GW}$ (colour coded) are stronger emitters of GWs and the most efficient emitters (see Fig. 4.13). Comparing this figure with Fig. 4.11, one can see that the ${}^2p_4$ mode (light-blue stars in Fig. 4.12) is the most efficient emitter. This mode also coincides with the most prominent feature in the spectrogram. Inspecting the definition of $\tau_{\rm GW}$, one realises that

$$\tau_{\rm GW} \propto \frac{(n(2l+1)!!)^2}{\sigma^{2l}}, \tag{4.51}$$

where the dependence with n is a crude approximation and a consequence of the different number of nodes in the integrands of Eqs. (4.46) and (4.47). This dependence produces a qualitatively different behaviour of p-modes and g-modes in terms of $\tau_{\rm GW}$. In both cases, the GW emission of high-order n modes is expected to be suppressed. However, for p-modes, $\tau_{\rm GW}$ decreases for increasing frequency. This compensates somewhat the suppression factor at high n and allows the appearance of high order p-modes up to $n = 5$ in the spectrogram. On the contrary, GW emission in high order g-modes is suppressed by both

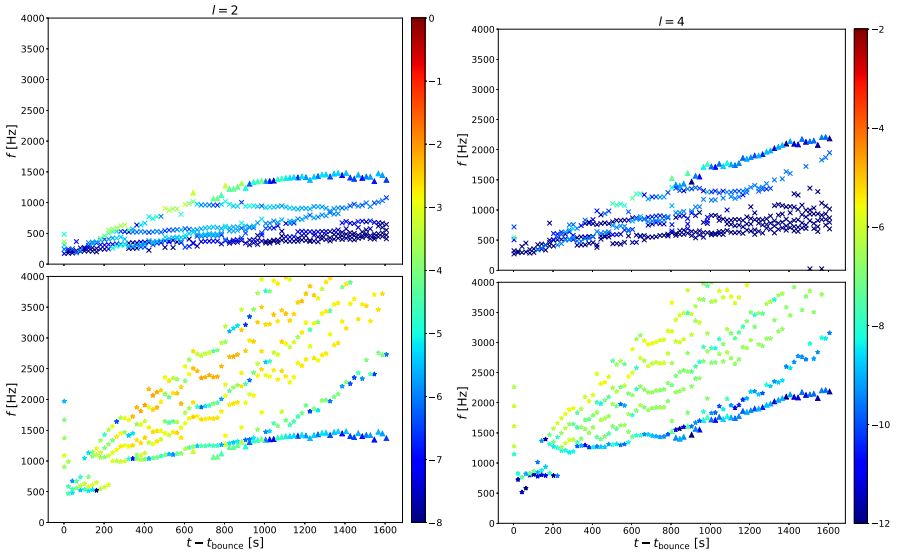


Figure 4.13 Time-frequency diagram of the modes of the complete problem separated their efficiency. Modes with low values are coloured in blue and their values scale to the larger values in red. The upper row corresponds to g-modes (crosses), while the lower row corresponds to p-modes (stars). The fundamental mode is represented with triangles. The colour bar shows the logarithm of the efficiency.

the high value of n and the low frequency, which results in no observed g-mode frequencies for $n > 3$. It seems clear that our classification allows us to match the features in the spectrogram with modes with low values of $\tau_{\rm GW}$, i.e. with high efficiency. Regarding $l = 4$ modes, the mode with lowest $\tau_{\rm GW}$ is the ${}^4p_4$, which explains the appearance of this feature in the spectrogram. Note that when comparing $l = 2$ and $l = 4$ modes, there is a difference in $\tau_{\rm GW}$ of several orders of magnitude. The reason for this difference is that, for perturbations of a spherical background, $l = 4$ modes do not contribute to the quadrupolar contribution of the GW emission, but exclusively to the octupolar component. This component is highly suppressed by a $1/c^2$ factor with respect to $l = 2$ modes[3]. However, in the simulation, the PNS is deformed by the presence of rotation and hence the $l = 4$ mode is able to contribute to the quadrupolar radiation. This enhances emission from these modes, which are observed in the GW spectrogram. Therefore, in order to properly capture the GW emission properties of modes with $l \neq 2$ in rotating cores, one should have to consider corrections due to to deformations of the PNS. These corrections are out of the scope of the present qualitative analysis and will be investigated in the future.

[3]This factor cannot be seen explicitly in Eq. (4.47) because of the use of units with $c = 1$.

We can also conclude that, for nonrotating cores, in which the background is spherically symmetric, modes with $l \neq 2$ cannot contribute significantly to the GW emission, and one could restrict the analysis to $l = 2$ modes.

4.6 Discussion

In this chapter we have reported results from the computation of the eigenmodes of a physical system formed by a PNS and an accretion shock obtained from a simulation of the core-collapse of a massive star. The goal has been to explore if the spectrum of the GW signal emitted during the collapse to a PNS and subsequent accretion phase leading to the formation of a black hole, can be related to the oscillating modes of the system. To this aim we have computed the number of nodes of each of the modes we have found. Mode classification in terms of the number of nodes has shown that there are two basic sets of modes, those with a number of nodes increasing with frequency and another group that follows the opposite trend. The two families are neatly separated by the so-called fundamental mode, a mode that has no radial nodes. The comparison with the g-modes and p-modes obtained with the analysis of the numerical simulation has allowed for an unambiguous identification of the two groups of modes found. We have further compared the modes obtained with the linear-oscillation approximation with the spectrogram of the GW signal obtained in the numerical simulation finding a remarkable correspondence between both results.

Part III

Gravitational-wave detector characterization

Chapter 5

Introduction to gravitational-wave detectors

The second generation laser interferometer GW observatories LIGO and Virgo have been commissioned in the last two years. The two American advanced LIGO detectors, located in Hanford (Washington) and Livingston (Louisiana), along with the European advanced Virgo detector, situated in Cascina (Italy), have been steadily upgraded to reach sensitivity levels that allowed to detect the first gravitational signals ever. The much-awaited discovery, a truly milestone in contemporary physics and astronomy, was finally achieved on September 14th 2015, when the two LIGO detectors captured the elusive, tiny spacetime vibrations associated with the very last cycles of the inspiral and merger of two stellar-origin BHs, and the resulting Kerr BH [Abbott et al. 2016b]. The signal, dubbed GW150914, confirmed a one-hundred-year-old prediction from the theory of general relativity, demonstrating for the umpteenth time that Einstein was indeed right. More importantly, that infinitesimally small vibration opened, lo and behold, the exciting prospect of studying the wonders of the Universe through the GW channel, complementing electromagnetic observations (i.e. what we know as traditional astronomy since the days of Galileo Galilei, more than four centuries ago) and neutrino detections of cosmic origin. Multi-messenger astronomy started to unfold on that 14th of September of 2015.

The detection of signal GW150914, followed shortly after by the detection of GW151226, have been both a magnificent demonstration of the capabilities of the state-of-the-art technology at work on advanced interferometric detectors. Those detections, however, would have not been possible without the vast amount of

effort done by a large number of scientists in the LIGO Scientific Collaboration and in the Virgo Collaboration in a challenging area of research known as detector characterisation. Understanding the different sources of noise of these detectors, their characterisation, as well as developing accurate techniques to remove corrupted data, have all played crucial roles to achieve the required sensitivity. In the present chapter we will briefly outline the basics of current GW detectors, we will illustrate their (multiple) sources of instrumental noise and the need to characterise the detectors. In particular, we will introduce the glitch-classification project, whose results will be presented in the next chapter.

5.1 Detectors of gravitational waves

The first gravitational wave detectors were developed during the 1960s by Joseph Weber. They consisted of massive ($M \sim 1000$ kg) cylindrical bars of aluminium. The principle of detection in these instruments (nowadays called Weber bars) relies on the mechanical vibration produced when a gravitational wave crosses the bar. The sensitivity response of the instrument is a narrow band centred on the resonance frequency of the cylinder. A transducer converts the mechanical energy of the bar into electrical energy, and an amplifier increases the electrical signal to record it. Due to their limited sensitivity ($\sim 10^{-21}$ in the most recent version), and its narrow-band bandwidth, these kind of detectors have been progressively abandoned in favour of detectors based in laser interferometry.

Laser interferometer detectors measure the stretching of spacetime and thus the relative distance between two free-falling masses placed at the end of the two arms of the interferometer. The changes in the space-time induced by a gravitational wave produce variations in this distance directly related with the amplitude of the incident wave. The principle of detection is based on the spacetime metric describing a plane gravitational wave traveling e.g. in the $+z$ direction of the so-called *transverse traceless* (TT) coordinate system [see Thorne, Misner, and Wheeler 2000]

$$ds^2 = -c^2dt + \left(1 + h_+\left(t - \frac{z}{c}\right)\right)dx^2 + \left(1 + h_-\left(t - \frac{z}{c}\right)\right)dy^2 + \\ + 2h_\times\left(t - \frac{z}{c}\right)dxdy + dz^2, \quad (5.1)$$

where the two independent polarisations of a gravitational wave, $h_+(t)$ and $h_\times(t)$, induce a perturbation of the flat Minkowski metric at the origin of the coordinate system; this is the so-called *weak-field* approximation.

Moreover, the interferometers also fulfil the *long-wavelength* approximation, which stands that if the reduced wavelength of the wave, $\tilde{\lambda} = \lambda/2\pi$ is much larger than the size of detectors arms L, the GW field can be treated as being uniform (but time-dependent) in the space region that covers the entire detector. With these two main approximations (for a detailed discussion see Jaranowski and Kròlak [2005]), it is possible to relate the changes in the phase of the interferometric pattern measured by a photodiode $\Delta\phi(t)$ with the detector response $h(t)$,

$$\Delta\phi(t) = 4\pi\nu_0 L h(t), \tag{5.2}$$

where ν_0 is the laser central frequency. The response function h is a linear combination of the two wave polarisations h_+ and $h_\times$,

$$h(t) = F_+(t)h_+\left(t - \frac{z_d(t)}{c}\right) + F_\times(t)h_\times\left(t - \frac{z_d(t)}{c}\right), \tag{5.3}$$

where $z_d(t)$ is the distance of the origin of the TT coordinate system with the corner station of the interferometer. Functions $F_+(t)$ and $F_\times(t)$ are the *beam patterns* of the interferometer. These functions relate the position on the sky of the source of gravitational waves, the polarisation angle, and the location of the detectors on Earth. The specific expression of the beam patterns can be found in Jaranowski and Królak [2009].

From Eq. (5.2) and Eq. (5.3) it follows that the measurable sensitivity of a gravitational wave depends directly on the length of the interferometer arms. The sensitivity can further be improved using Fabry-Perot cavities, which employ low-transmissivity mirrors to increase the number of trips that the light makes inside the cavity before it reaches the photodiode at the output of the detector, increasing the effective distance of the arms. The other important design factor for interferometers is the laser power. To improve the possibilities of GW signal detection, the interference patterns should be as sharp as possible. This is achieved by increasing the laser power. However, building a laser with the power required to operate at full sensitivity results practically impossible. To solve this issue, the LIGO and Virgo detectors also recycle the laser power using an additional mirror, called power-recycling mirror, at the entrance of the interferometer. The light reflected by the arms follows a path back to the recycling mirror, rather than to the photodetector. This light is reflected back into the interferometer, increasing the total power inside the detector. The advanced LIGO detector has incorporated an additional mirror, called signal-recycling mirror. This technology allows to "tune" the detector's response to the frequency band of the GW and was incorporated successfully in the GEO600

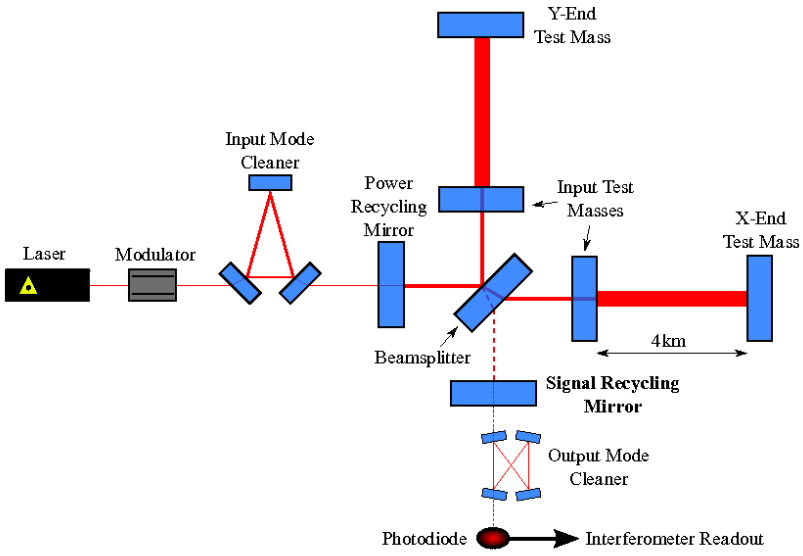


Figure 5.1 Simplified layout of the Advanced LIGO interferometers taken from Smith and Collaboration 2009.

detector [Grote and Collaboration 2010]. Fig. 5.1 illustrates the optical scheme of GW detectors.

5.2 Noise sources

Due to the minuscule amplitude of the spacetime vibrations caused by gravitational waves (when they hit the Earth), GW detectors based on interferometry have to be isolated from all significant non-gravitational-wave external phenomena, either seismic, electromagnetic, or acoustic (see Pitkin et al. [2011] for a detailed discussion). At the same time, they also have to be equipped with the most sophisticated technology regarding high-power lasers, vacuum systems (to avoid light dispersion), mirror coatings, suspension and seismic isolation subsystems, and optics. However, despite the astounding sensitivity of current advanced interferometers, the detection of gravitational waves is still terribly limited by noise.

Any of the required technologies can be a potential source of noise. While many of these can be controlled by design, there exist other sources of noise that have an external origin and require auxiliary channels and sensors to measure them and control them. The design of an interferometer is hampered by three

main sources of noise, namely seismic noise, quantum noise, and thermal noise. Other sources, such as laser frequency noise, photodetector dark noise, actuator noise, etc, are classified as *technical* noises.

Thermal noise is produced by the vibration of the mirrors and the suspension pendulums, designed to screen the ground vibrations, at the ends of the Fabry-Perot cavities. In the test masses, the Brownian noise produced by mechanical dissipation in the mirror coatings is the dominant source of thermal noise [Harry et al. 2002, Cumming et al. 2012]. In the suspension system, the thermal noise is due to the loss (violin modes) in the fused silica fibers [Hammond et al. 2012].

Quantum noise is an effect of the statistical fluctuations in detected photon arrival rate (shot noise) and radiation pressure due to photon number fluctuations. In particular, shot noise is inversely proportional to the square root of the number of photons collected by the photodiode, which is proportional to the power of the laser [Buonanno and Chen 2001]. Therefore, this type of noise can be reduced by increasing the laser power and with an efficient control of the light scattering. The former is accomplished through light-recycling techniques which use additional mirrors that reflect the wasted light back into the interferometer in the positions where light losses in the photodiode are largest (signal-recycling mirror) and in the place where the laser beam is injected into the interferometer (power-recycling mirror). The last main source of noise is due to gravity gradients. Seismic waves produce density perturbations in the ground close to the test masses, which in turn produce fluctuating gravitational forces on the masses. This source of noise limits the possibilities of detection in the mHz frequency range but is negligible above 11 Hz.

Each noise source has a different impact on the instrument depending on the frequency. The combination of all noise sources produces the so-called detector sensitivity curve, which illustrates the dependence of the detection capabilities with the frequency. Fig. 5.2 shows these curves for aLIGO (upper panel) and Advanced Virgo (bottom panel), corresponding to the theoretical design of the two instruments. The different noise sources are shown in different colours, white the total noise or sensitivity curve is shown with a solid black curve in both cases.

5.3 Transient noise sources

An important contribution to the noise budget of interferometers comes from non-Gaussian transients of noise, also known as “glitches”. The nature of this noise is purely instrumental or environmental. Transients constitute the sources

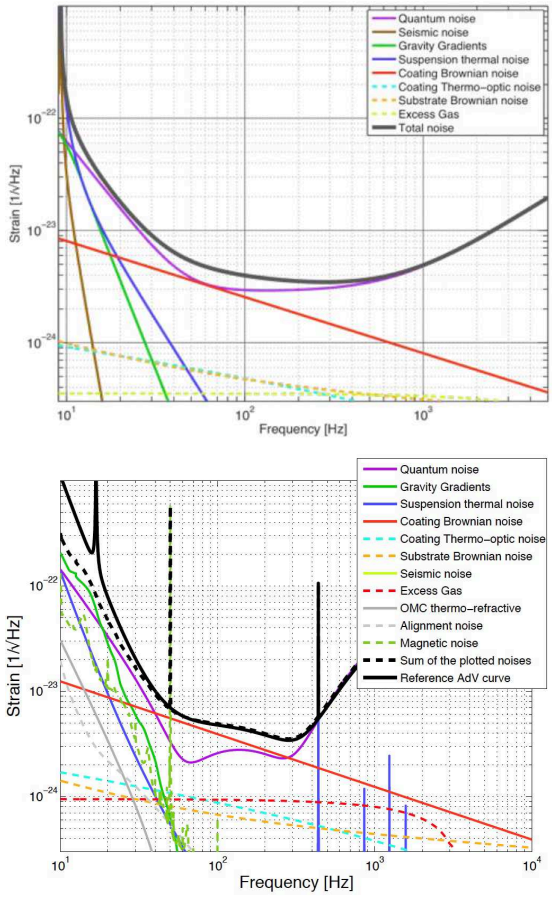


Figure 5.2 Spectral sensitivity curves of Advanced LIGO [Aasi et al. 2015b] (upper panel) and Advanced Virgo [Acernese et al. 2014] (bottom panel).

of noise that most limit the sensitivity of searches for short and poorly-modelled astrophysical sources of gravitational waves, like unmodelled bursts and high-mass binary coalescence signals. Advanced detectors contain many environmental and instrumental sensors, which produce auxiliary channels of data that can be used to monitor the detector behaviour, track the causes of short-duration noise artefacts, and ultimately veto noise transient events [Smith et al. 2011, Ajith et al. 2014, Aasi et al. 2015a].

These artefacts have a large variety of time-frequency-amplitude morphologies, with the undesirable particularity that many of them can mimic a true GW signal. To reduce the impact of glitches in the analysis of the collected data, the coincidence criteria is applied, i.e a transient should appear in more than one

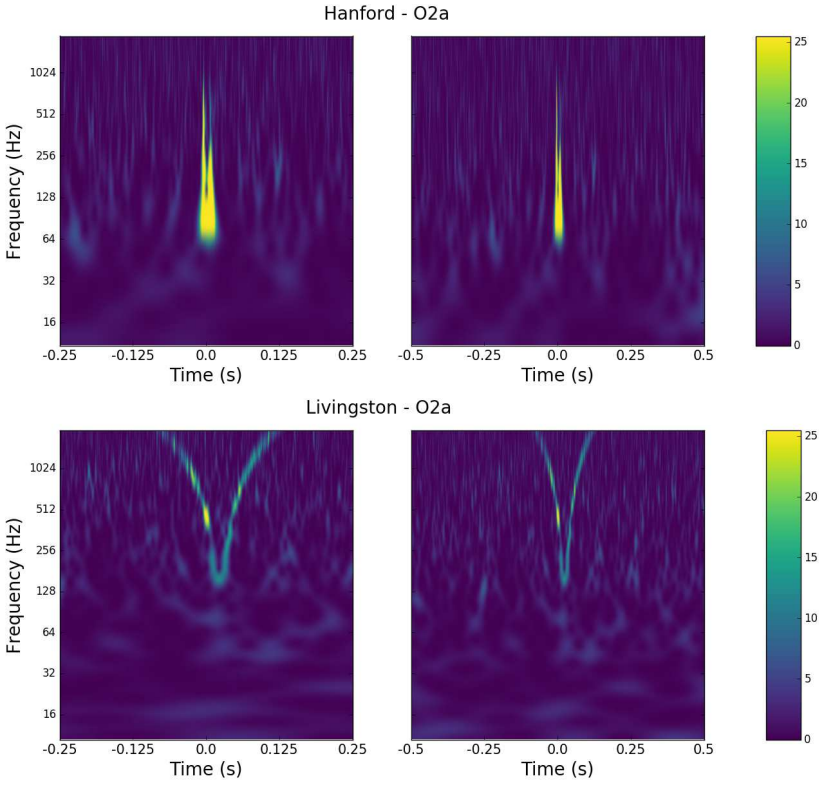


Figure 5.3 Time-frequency plots of two typical glitches commonly found in interferometers. An example of a "blip gliclh" is show in the upper panels while the lower panels show an example of a "whistle glitch".

detector with a time delay less than or equal to the light-travel-time between the observatories, to be worth of further consideration. If a signal only appears in one observatory during this time window, it is rejected. However, the rate of glitches at the advanced LIGO detectors is fairly large (see Abbott et al. [2016a]) and thus, accidental coincidence between the two detectors is non-negligible. The situation will obviously improve when the Advanced Virgo detector joins the observational campaigns of the two LIGO observatories. The impact of glitches is not negligible. They reduce the significance of candidate GW events because they increase the background loudness. If they occur sufficiently close to a detected signal, they affect the estimation of the physical parameters, producing broader uncertainties. Moreover, the veto techniques used to remove glitches also reduce the amount of usable data. Therefore, to maximise the GW detection rate, either the causes of the glitches must be identified and fixed within the

detectors (in the best case scenario) or glitches must be removed from the data set.

The origin of glitches is not completely clear. Some categories have known causes, while others have causes yet to be identified. Fig. 5.3 shows two examples of typical glitches found in interferometer data. The first row shows several time windows of a *blip* glitch, whose origin is unknown. The second row shows the type of glitch called *whistle*. Whistles are known to be caused by radio signals at MHz frequencies that beat with Voltage Controlled Oscillators in the interferometer control system, and have a typical morphology in the form of a "W" or a "V". As detectors are upgraded and evolve, new classes of glitches can be found. Therefore, it is necessary to develop robust methods to identify and characterise glitches. This is discussed in the following chapter.

Chapter 6

Glitch Classification

The results of this chapter have been originally published in:

Powell, J., Torres-Forné, A., Lynch, R., Trifirò, D., Cuoco, E., Cavaglià, M., Heng, I.S., and Font, J.A., "Classification methods for noise transients in advanced gravitational-wave detectors II: performance tests on Advanced LIGO data" , Classical and Quantum Gravity, **34**, 034002 (2017).

Classification and categorization of noise transients in GW detectors using individual channels of data may provide valuable clues for the identification of their sources, which can be of great help to eliminate them [Powell et al. 2015, Mukherjee, Obaid, and Matkarimov 2010]. So far transient classification has mainly been achieved by visual inspection of spectrograms. Automatic classification is however essential for future detections of astrophysical GW signals.

The goal of this chapter is to evaluate the performance of three methods for fast classification of transients which have been developed for the analysis of Advanced LIGO (aLIGO hereafter) and Virgo data, using glitches in actual data from aLIGO. Specifically, these methods are Principal Component Analysis for Transients (PCAT), Principal Component LALInference Burst (PC-LIB) and Wavelet Detection Filter with Machine Learning (WDF-ML). Previous work has shown that these methods can classify artificial data sets with an efficiency up to 95% [Powell et al. 2015].

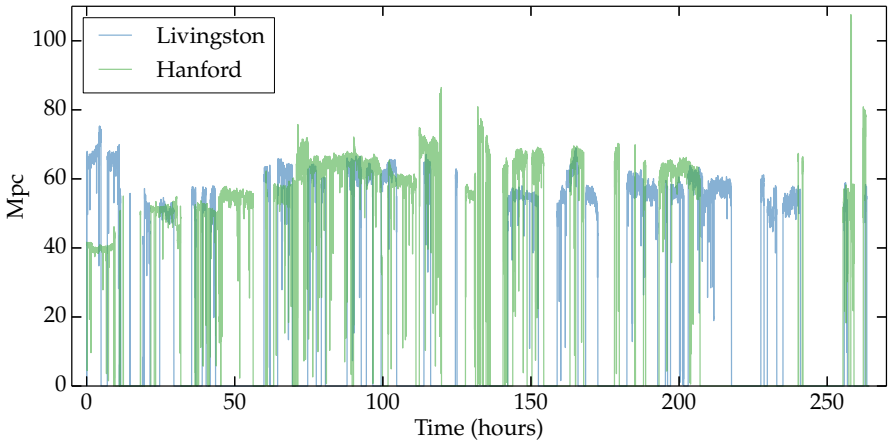


Figure 6.1 The mean binary NS star inspiral range for the two aLIGO detectors during ER7. The Hanford detector had a higher range but also a higher glitch rate. The average range was 50-60 Mpc.

6.1 The Data

In this study we use data from the 7th aLIGO engineering run (ER7), which began on the 3rd of June 2015 and finished on the 14th of June 2015. Engineering runs are stretches of time where the detectors are operated as through in an observing run, and allow the commissioning and the Detector Characterization groups to adjust the operational parameters of the interferometer and find and minimize many technical sources of noise, inproving the sensitivity for the Observational runs. The average binary NS star inspiral range i.e. the distance at which these sources can be detected, for both H1 and L1 detectors in data analysis mode during ER7 was $50-60$ Mpc [Areeda and Smith 2016]. The mean range for both detectors is shown in Fig. 6.1.

6.1.0.1 Livingston.

In the period analyzed, data from L1 consists of 48 segments where the interferometer was locked and in data analysis ready mode. These data segments vary in length from 1 second to $\sim$ 7 hours. We discard any segments of data that are less than a minute in duration as a longer segment of data is required to measure the power spectral density (PSD). The total discarded amount was 49 seconds of data. The total length of L1 data analyzed is $\sim$ 87 hours.

Glitches of different types are often recognised by their shape in a spectrogram such as those shown in Fig. 6.2. A description of the most common glitch types,

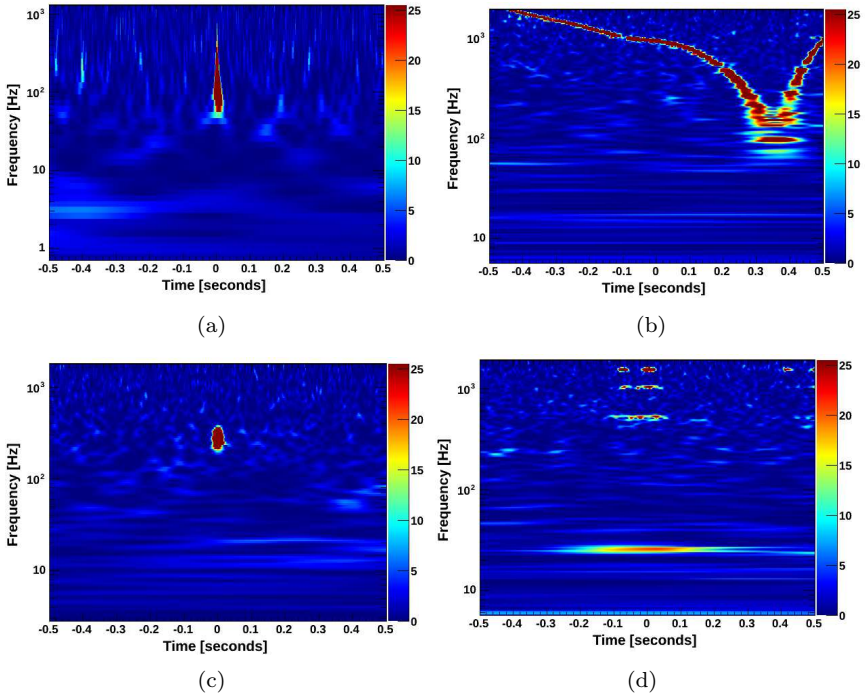


Figure 6.2 Spectrograms of typical transient types found in the aLIGO Livingston ER7 data. They are generated using the Omega scan tool in LigoDV-Web, which matches the data to sine Gaussians. (a) A transient characterised by a tear drop shape in the spectrogram. (b) A "whistle" glitch that often has a long duration and occurs at high frequencies. (c) A hardware injection. (d) A transient type characterised by high frequency lines and lower frequency features.

which have occurred in aLIGO data, are described in Smith [2016]. Fig. 6.2(a) shows glitches characterised by a tear drop shape. Fig. 6.2(b) shows longer duration transients known as "whistles", which are caused by radio-frequency beats [Smith 2016]. Only a small number of whistles ($\sim$ 11) were found in the frequency (30-2048 Hz) and SNR range used in this study. Some other glitches in the data that are not shown in Fig. 6.2 have energy content below 10 Hz or are due to scattered light. Glitches span the entire frequency range considered in this study. Some noise transients may have occurred due to the increased microseism created by tropical storm "Bill" in the Gulf of Mexico [Areeda and Smith 2016].

A number of hardware injections were also made during ER7. An example is shown in Fig. 6.2(c). Hardware injections are artificial signals simulated by

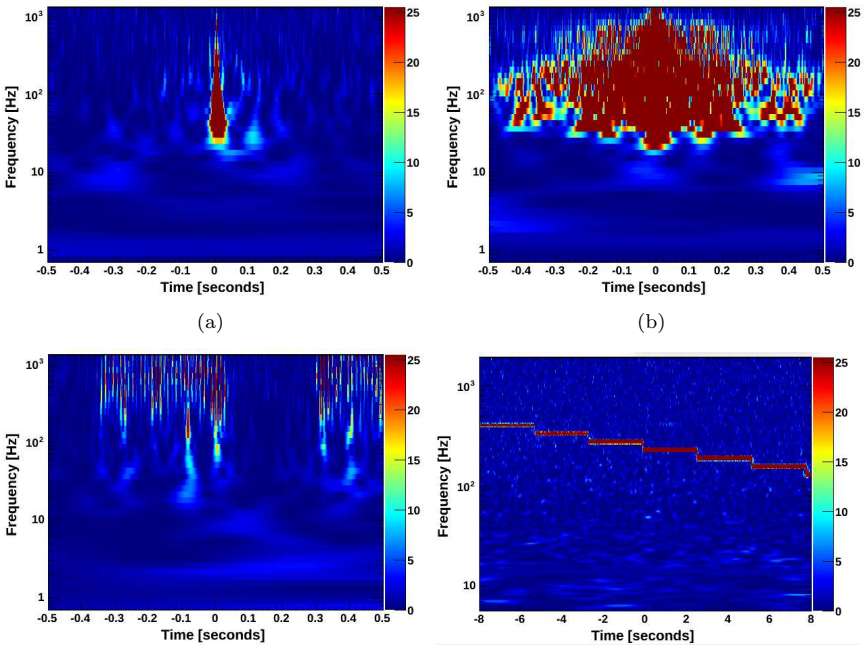


Figure 6.3 Examples of some of the most common transient types found in the Hanford ER7 data (a) A tear drop glitch. (b) Transients of this type have a large SNR and duration. They created significant drops in the detectors range. (c) A high frequency transient type. (d) A longer duration line occurring at the beginning of a number of data segments.

inducing a motion of the optics that can be used to test which auxiliary channels are sensitive to gravitational waves [Smith et al. 2011, Ajith et al. 2014].

6.1.0.2 Hanford.

In the period analysed, data from the H1 detector consists of 50 segments where the interferometer was locked and in data analysis ready mode. The data segments vary in length from 1 second to almost 14 hours. As with L1 we discard any segments of data that are less than a minute in duration, which was a total of 116 seconds of data. The total length of Hanford data analysed is $\sim$ 141 hours.

The H1 data is highly non-stationary and contains many more transients than the aLIGO L1 data. In particular, the H1 data contains many high SNR transients that caused a significant drop in the binary NS star inspiral range.

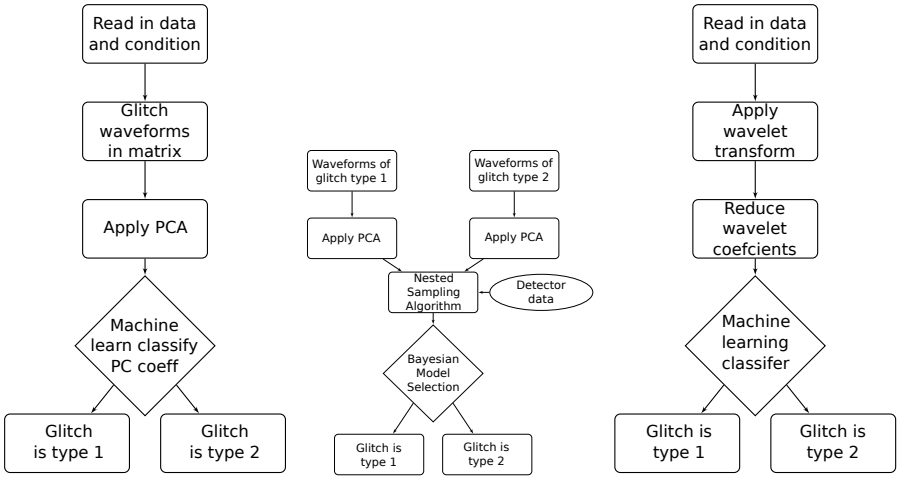


Figure 6.4 Classification procedures for the three different methods used in this study. (a) PCAT applies principal component analysis to all transients detected in a stretch of data and then applies a machine learning classifier to the principal component coefficients. (b) PC-LIB uses a combination of principal component analysis and Bayesian model selection to determine the glitch type. (c) WDF-ML applies a machine learning classifier to wavelet coefficients obtained by applying a wavelet transform to the transients in the data.

An example is shown in Fig. 6.3(b). It was suspected that these large transients were caused by cleaning of the beam tube. A few other examples of common transients found are shown in the other spectrograms displayed in Fig. 6.3. As with the L1 data, H1 data also contains a number of hardware injections.

6.2 Transient classifying algorithms

Three different classifying algorithms were developed for the fast classification of noise transients in the detectors. Most of the technical details have been described in Powell et al. [2015]. Here we give a brief outline of the three methods and describe any changes that have been made to improve their performance and latency. Fig. 6.4 outlines the classification procedures for all three methods. More details are given in the following subsections.

To find transients in the data we use event trigger generators (ETGs). ETGs typically search for excess power in individual interferometers and output the time, SNR, frequency, duration and other parameters of transients found in the data. PC-LIB uses Omicron, the main ETG used by the LIGO Scientific

Collaboration's (LSC) detector characterisation group [Chatterji et al. 2004, Robinet 2015]. WDF-ML and PCAT have their own internal ETGs.

6.2.1 PCAT

PCAT uses a technique called Principal Component Analysis (PCA) that allows for dimensional reduction of large data sets [Friedman, Hastie, and Tibshirani 2001a, Powell et al. 2015]. In the first stage of the PCAT analysis, the data are downsampled to 8192 Hz, whitened and high-pass filtered at 10 Hz. Then PCA is applied to all of the noise transients found by the ETG in all the spherically-symmetric segments of data. PCAT uses a 0.125 s window around each GPS time as glitches are typically of ms duration. This can lead to a loss of sensitivity to longer duration glitches. However, this effect can be safely neglected as longer duration glitches do not occur very often during observing runs, when the data is generally more stable than the ER data.

A projection of the original waveforms on to the Principal Components (PCs) allows for the calculation of scale factors for each PC called PC coefficients. Noise transients of different types are separated in the PC coefficient parameter space. This allows PCAT to classify the transients by applying a Gaussian Mixture Model (GMM) machine learning classifier to the PC coefficients [Pedregosa et al. 2011].

6.2.2 PC-LIB

LALInference Burst (LIB) is a Bayesian parameter estimation and model selection tool, which uses a sine-Gaussian signal model to estimate parameters of GW bursts [Essick et al. 2015]. It can also be combined with Omicron to be run as a search [Lynch et al. 2015]. PC-LIB adapts LIB for the classification of transients by replacing the LIB sine-Gaussian signal model with a new signal model created from a linear combination of PCs calculated from the waveforms of known transient types [Heng 2009, Logue et al. 2012]. These known transients may have been previously classified by examining spectrograms of the transients or by one of the other methods. Thus PC-LIB can only classify transients that have occurred in the data many times before. When transients of a new type start to appear in the data new signal models must be created.

In Powell et al. [2015] we created signal models using fifty transient waveforms. In this study we only use ten waveforms. This change will allow us to start classifying new transient types more quickly as they start to appear in O2 data, without any loss in sensitivity, as glitch waveforms for specific types do not

have much variance in shape. Bayesian model selection can then be used to determine what population of noise transient each new glitch belongs to [Sivia 1996, Veitch et al. 2015, Lynch et al. 2015]. First, one second of data around the trigger time is downsampled to 8192 Hz and a 10 Hz high pass filter is applied. Nested sampling is then used to calculate Bayes factors to determine the correct transient type [Sivia 1996].

6.2.3 WDF-ML

Wavelet detection filter (WDF) is an ETG that is part of the Noise Analysis Package (NAP), developed by the Virgo collaboration [Acernese et al. 2005, Acernese et al. 2007]. It is combined with a machine learning classifier for transient classification (WDF-ML).

In order to reduce the number of wavelet coefficients produced by WDF-ML, the data are downsampled before any data conditioning in the time domain to prevent border effects introduced by the Fast Fourier Transform (FFT). The downsampling is a new feature of WDF-ML that was not implemented in the version of the algorithm used in Powell et al. [2015]. The data are then whitened using parameters estimated at the beginning of each locked segment. After whitening, a wavelet-transform is applied, using a bank of wavelets, as described in Powell et al. [2015]. We use a window of 2048 samples, with an overlap of 1968 samples, which corresponds to a duration of 0.25 seconds, as transients are typically of a short (ms) duration.

The wavelet coefficients identified by the WDF-ML ETG are further cleaned using a wavelet de-noising procedure where only wavelet coefficients above the noise level are retained [Powell et al. 2015]. WDF-ML produces a list of wavelet coefficients, frequency, duration and SNR for each transient. The dimensions of the wavelet coefficients are then reduced by applying PCA and Spectral Embedding [Ng, Jordan, and Weiss 2001, Belkin and Niyogi 2003]. The transient classification is then performed by applying a machine learning GMM classifier to the reduced wavelet coefficients [Pedregosa et al. 2011].

6.3 Classification

In the following sections we show the classification results obtained by PCAT, PC-LIB and WDF-ML on aLIGO H1 and L1 data. All algorithms are run with the same configurations that we expect to use during O2 to better understand our

performance during the future observing runs. To determine if the glitches are classified correctly spectrograms of all glitches are made and visually inspected.

6.3.1 Livingston

To find the transients in the L1 data we look for triggers that are coincident within half a second in the outputs of all ETGs. The WDF-ML ETG was run with an SNR threshold of 10 at a sampling rate of 8192 Hz. Omicron was run with a lower SNR threshold of 5. We then look for transients that are coincident between both WDF-ML and Omicron, above SNR 20, and find a total of 426 coincident transients. As the PCAT cannot find the lower frequency (below 10 Hz) triggers and some longer duration triggers we still classify transients that are coincident between Omicron and WDF, but missed by PCAT, as those triggers would still be classified when running in low latency.

PCAT applies a threshold on the SNR of the transients of 4.5 and the maximum possible number of transient types was set to 10. The ideal number of PCs can be estimated through the analysis of the variance curve of the whole dataset. This gives a total number of 20 PCs.

PCAT classifies all the transients into 10 different classes. 90 triggers that were coincident between the Omicron and WDF-ML ETGs were missed by the PCAT ETG. Included in these missed triggers are all of the whistles, as their duration is longer than the PCAT analysis window, and 17 transients that are not visible in a spectrogram. 20 of the lower SNR hardware injections are also missed.

The data contains three main types of transients with examples shown in Fig. 6.2(a), (b) and (c). As PCAT does not detect any of the whistles shown in Fig. 6.2(b) the remaining glitches are classified into two main types, further split into sub-types. PCAT classes 1, 4 and 10 contain the transients which appear as a spike in the time series, as shown in the 6.5(a) and in the spectrogram in 6.2(a). Class 4 contains only 2 transients, class 1 contains 123 transients and class 10 contains 100 transients. Class 1 and 10 contain 11 and 20 hardware injections respectively. The three sub classes are characterised by different duration of the transients. Triggers in class 1 have the lowest (≤ 0.005 s) duration, class 10 have a larger (≤ 0.01 s) duration, and class 4 contains two longer (≥ 0.01 s) duration spikes. Two of the transients in class 10 were incorrectly classified.

Classes 3, 5, 6, 7 and 8 contain the transients with a time series waveform shown in Fig. 6.5(b) and a spectrogram shown in 6.2(d). Triggers in classes 5, 7 and 8 all have SNR values between 20 and 25 and durations of $\sim$ 0.01 s. Class 3

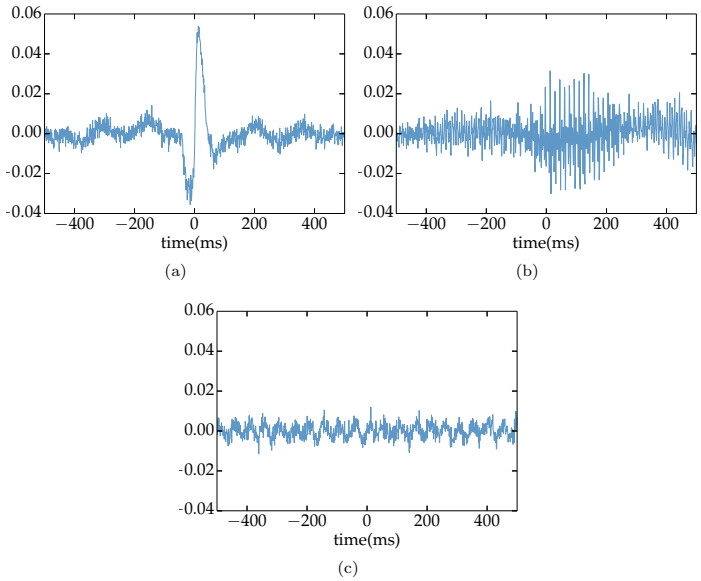


Figure 6.5 The typical high pass filtered and whitened time series waveforms for three of the most common transient types found in the Livingston detector. (a) A spike which appears as a tear drop in a spectrogram. (b) The time series waveform of the glitches shown in Fig. 6.2(d). (c) The time series of a whistle glitch.

contains triggers of the same transient type but with larger durations (≤ 0.02 s) and SNR values up to 50. Class 6 contains only one transient, also of the same type, but with an SNR value of 57 and a duration value of 0.005 s.

PCAT classes 2 and 9 contain 11 and 7 glitches respectively. As these transients are not visible in a spectrogram it is not possible to determine what their type is and if they are classified correctly. Overall 95% of the transients are correctly classified by PCAT.

PC-LIB classifies all transients into four different types. To create the signal models the first 5 PCs for each glitch type are used as determined thought the analysis of the variance curve of the whole dataset. Class 0 contains 33 transients that are not detected by PC-LIB and are thus classified into a noise class. Most of the noise class transients occur at frequencies lower than the 10 Hz cutoff used by PC-LIB.

Class 1 contains 249 transients that correspond to PCAT class 4 and 5 and appear as a spike in the time series. Two of the transients in this type are mis-classified. All of the hardware injections in the data are classified in this type. Class 2 contains 131 transients which correspond to PCAT sub-classes 3,

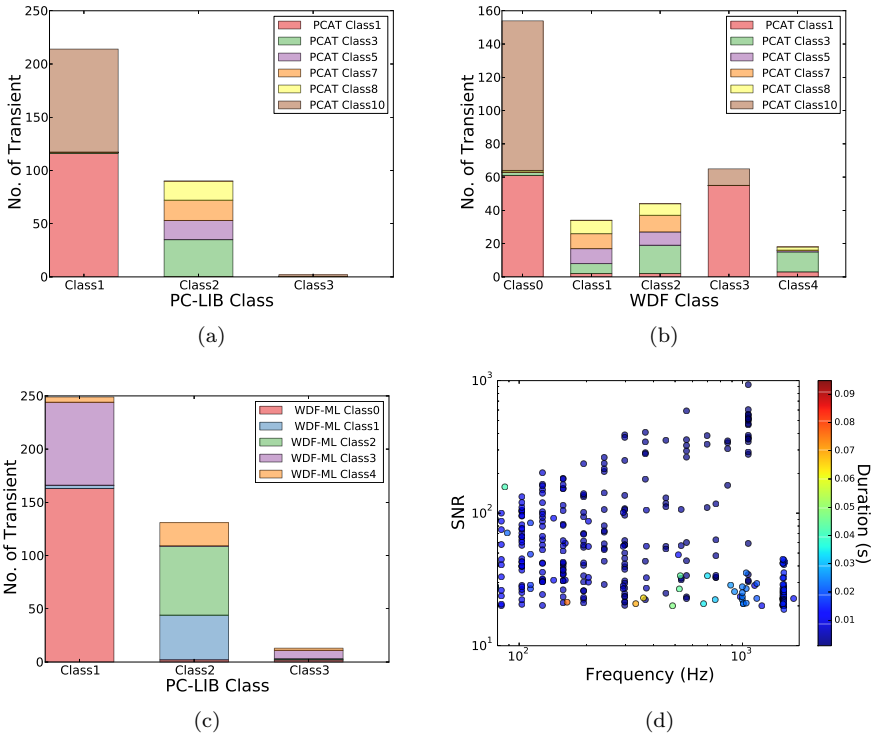


Figure 6.6 Classification comparisons for the three different methods on the data from LIGO Livingston. (a) Compares the classification results of PCAT and PC-LIB. PCAT class 2,4,6 and 9 are not shown as they contain less than 15 transients. (b) Compares the classification results of PCAT and WDF-ML. (c) Compares the classification results of PC-LIB and WDF-ML. (d) The SNR and frequency of all the transients classified in the data.

5, 6 and 7 and have a time series waveform shown in Fig. 6.5(b). There are no incorrectly classified transients in this class.

Finally, class 3 contains 13 transients. Most of the transients in this class are the whistle transients shown in 6.2(b) and in Fig. 6.5(c). Three of the transients in this class are mis-classified and should be in class 2. Overall PC-LIB classifies 98% of the detected transients correctly.

WDF-ML classifies all transients into five different classes. The 5 classes consist of two different types of transients as WDF-ML cannot accurately classify the longer duration whistles due to the short analysis window. Sub-classes are determined by the wavelet family of the transients rather than split by duration or SNR as for PCAT.

WDF-ML class 0 contains 195 transients and class 3 contains 86 transients that appear as a spike in the time series. These correspond to PCAT types 1, 4 and 10 and PC-LIB's class 1. The two sub-classes contain 29 hardware injections. They also contain 8 of the whistle glitches, as WDF-ML cannot accurately classify longer duration transients. Four of the class 0 and one of the class 3 transients are incorrectly classified.

The second main transient type found by WDF-ML corresponds to PC-LIB's class 2 and contains transients characterised by the typical spectrogram shown in 6.2(d). The transients were split into three sub-classes, namely class 1 that contains 46 transients, class 2 that contains 70 transients and class 4 that contains 29 transients. Class 1 contains three incorrectly classified transients and class 3 contains two of the whistles glitches. Class 4 contains 4 hardware injections that are mis-classified. Overall WDF-ML classifies 95% of the L1 transients correctly.

Comparison

Fig. 6.6 shows a comparison of the classifications made by all three methods. All methods are able to classify transients with a high level of accuracy in real non-stationary data. WDF-ML performs better at classifying very low frequency transients as it does not need to use a lower frequency cutoff. The Omicron SNR, duration and frequency of all 426 transients are shown in Fig. 6.6(d). The discrete frequency values are due to the Omicron's method for measuring frequency [Robinet 2015].

Only PC-LIB is able to separate the whistle transients into a separate class due to the longer 1 s time window used by this method. The efficiency in classifying these transients for the other algorithms could be improved by using a longer time window. However, this could lead to multiple shorter duration transients occurring in the same time window. As PC-LIB looks for specific known transient types it could be used to add labels to the classifications of the other methods so that it will make it easier to find out which class corresponds to which transient type, defined in [Smith 2016], and which classes are new types that have not occurred previously. As WDF-ML and PCAT can classify new transient types as soon as they appear in the data they can be used to provide waveforms for PC-LIB to use to create new signal models.

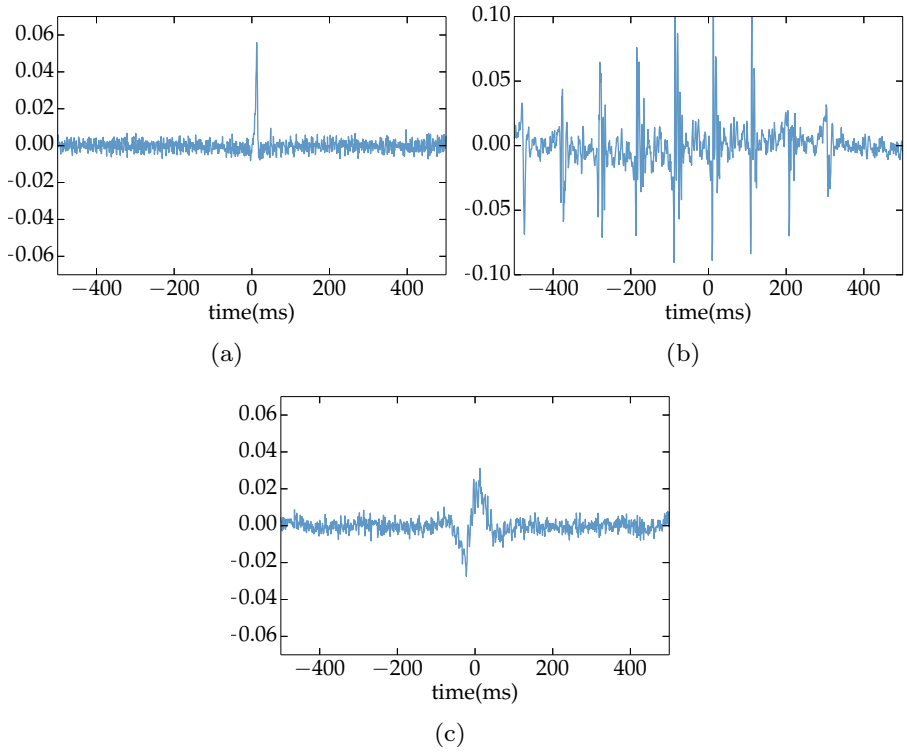


Figure 6.7 The typical whitened and high pass filtered time series waveforms for three of the most common transient types found in the Hanford detector. (a) A spike in the time series that appears as a tear drop in a spectrogram. (b) The time series of second most common glitch found in the data, as shown in 6.3(c). (c) The time series waveform of a hardware injection.

6.3.2 Hanford

As for the L1 data transients coincident within 0.5 s between all ETGs are classified. A higher SNR threshold of 30 is used for H1 as the data contains many more transients than the L1 data and is more non-stationary. A total of 1865 coincident transients are classified in H1.

PCAT uses 20 PCs to classify the transients into 7 different types. 120 of the transients coincident between WDF-ML and Omicron ETGs are not detected by the PCAT ETG. They are transients below 10 Hz or triggers from the long duration lines, shown in 6.3(d), which are not really glitches. The detected transients are split into 7 different classes.

The data contains two main types of transients. The first type is characterised by a typical spectrogram shown in 6.3(a) and a time series waveform shown in

Fig. 6.7(a). PCAT splits this type into 6 different sub-classes with 267, 603, 648 44, 1, and 64 transients respectively. Class 1 has 9 mis-classified transients. Classes 2, 3 and 5 all have one mis-classified transient. Class 2, 3 and 6 contain lower duration (~ 0.005 s), but with different Q and frequency ranges, where Q is defined as $Q = \text{duration} \times 2\pi \times \text{frequency}$. Class 1, 5 and 16 contain relatively longer duration waveforms (~ 0.01 s) which also have different Q and frequency ranges.

The second type of transient has a typical time-frequency morphology shown in 6.3(c) and time series waveforms shown in Fig. 6.7(b). This type is found in PCAT class 4 that contains 117 transients that are all classified correctly. Overall PCAT classifies 99% of the detected H1 transients correctly.

PC-LIB As with the L1 data we use 5 PCs to create signal models for the H1 transients. PC-LIB splits the transients into two different classes. A noise class contains the 6 transients shown in 6.3(d) as they cannot be detected.

Class 1 contains 1651 transients that correspond to a spike in the time series as in PCAT sub-classes 1, 2, 3, 5, 6 and 7. This class also contains 13 hardware injections. 23 transients are mis-classified and should be in class 2.

Class 2 contains 207 transients, which have a typical spectrogram shown in 6.3(c), and correspond to PCAT class 4. This class also includes 4 hardware injections that are more similar to a sine-Gaussian in shape than those classified into class 1. This class includes 61 transients that are mis-classified and should be in class 1. Overall PC-LIB classifies 95% of the detected H1 transients correctly.

WDF-ML splits the H1 data into three different classes. Class 1 contains 1358 transients, which appear as a spike in the time series, and correspond to PC-LIB class 1 and the 6 PCAT sub-classes. This class contains all hardware injections and all very low frequency transients that can not be detected by PCAT and PC-LIB. 10 of the transients in this class are mis-classified. WDF-ML class 2 contains 145 transients that are characterized by spikes in the time series, but have longer durations and lower SNR values than the transients in WDF-ML class 1.

WDF-ML class 0 contains 326 transients corresponding to PCAT class 4 and PC-LIB class 2. This class also contains 122 mis-classified transients. As before, this is because all of the transients in the class have a duration (~ 1 s) which is much longer than the time window used in the WDF-ML analysis. Overall WDF-ML classifies $\sim 92\%$ of the H1 transients correctly.

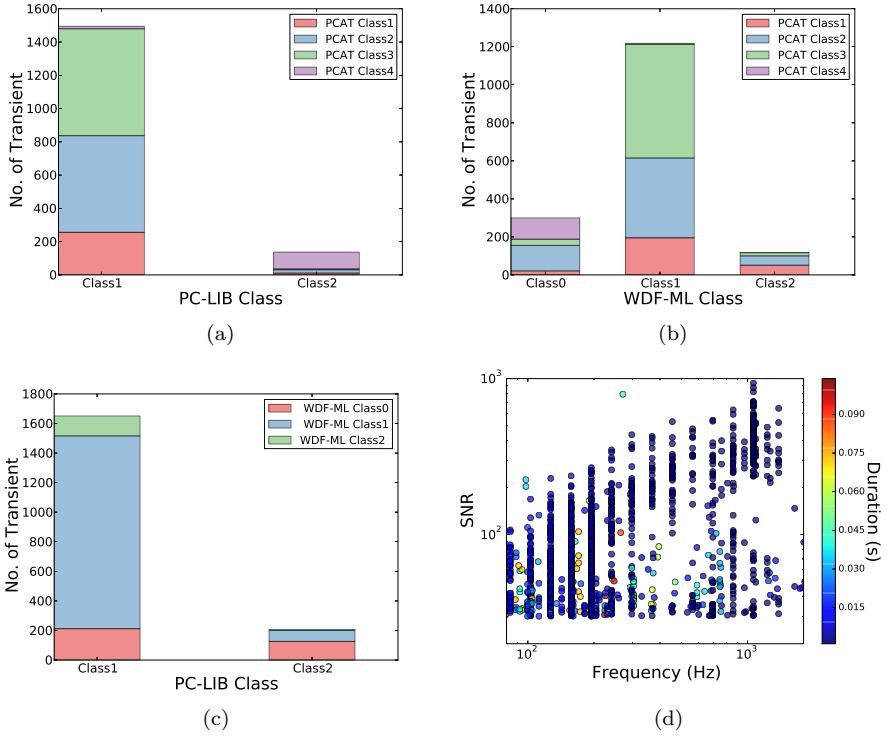


Figure 6.8 Classification comparisons for the three different methods for aLIGO Hanford data. (a) PC-LIB splits the transients into two classes. PCAT can split different types into sub-classes. (b) PCAT and WDF-ML comparison. WDF-ML has difficulty with transients which have a larger duration than their analysis window. (c) Comparison of PC-LIB and WDF-ML classifications. (d) The Omicron SNR, duration and frequency of all the transients classified in the data. The discreteness in frequency is due to Omicron.

Comparison

The results obtained by all three methods for the H1 data are compared in Fig. 6.8. The Omicron SNR, duration and frequency of the transients is shown in Fig. reffig:compareh(d). As WDF-ML uses a small time window of 0.25 s the efficiency of the classification is reduced when the data are highly non-stationary and contain many long (~ 1 s) duration transients. Even with 137 mis-classified transients the overall accuracy of the WDF-ML H1 results is $\sim 92\%$. WDF-ML estimates the PSD at the beginning of each locked segment. This may introduce errors towards the end of the segment if the data is highly non-stationary. Machine learning methods perform better when the data set

analysed is large. Therefore, the larger number of glitches in H1 may have improved the classification efficiency.

6.4 Discussion

Non-Gaussian noise in the aLIGO and Virgo detectors can potentially mimic a GW signal, reduce the duty cycle of the instruments and decrease the sensitivity of the detectors. Classification of different noise transient signals may help identify their origins and lead to a reduction in their number. We have developed three methods for noise classification and have previously demonstrated their performance on simulated transients in simulated Gaussian aLIGO noise [Powell et al. 2015]. However, as real noise from the advanced detectors is non-stationary and non-Gaussian, a better understanding of how our methods will perform during the upcoming observation runs of the advanced detectors is required. Although the detectors are typically more stable during observing runs than during ER7, we expect the types of glitches investigated in this work to be representative of the glitch classes in the upcoming observing runs.

In the ER7 data from the L1 detector PCAT missed 90 transients and classified 95% of the remaining transients correctly. PC-LIB missed 33 transients and classified 98% of the remaining transients correctly. WDF-ML classified all transients and 95% of them were correct. In the H1 data PCAT missed 120 transients and classified 99% of the remaining transients correctly. PC-LIB missed 6 transients and classified 95% of the remaining transients correctly. WDF-ML classified all transients and 92% of them were correct. We conclude that our methods have a high efficiency in real non-stationary and non-Gaussian detector noise.

The efficiency of the WDF-ML algorithm is reduced for the Hanford glitches because the duration of the transients becomes much larger than the analysis window, which reduces the efficiency of the overall classification. This could be prevented by applying a high duration cutoff to the transients found by the ETG before classification. Most high duration and SNR transients are removed by data quality vetoes. Conversely, short duration transients will be more important as they have a higher impact on the GW search backgrounds. Since they are rarely removed by vetoes their accurate classification is crucial to improve GW searches as an accurate categorisation will allow us to search for couplings within the detector [Abbott et al. 2016e, Abbott et al. 2016a]. These classification methods can help to discover new families of glitches.

Because of the different strengths and weaknesses of the different methods having multiple classifiers is a winning strategy. WDF-ML can classify lower frequency transients than the other two methods. PC-LIB is better able to classify longer duration transients due to its longer analysis window. PCAT can classify new types of transients as soon as they appear in the data and thus provide transient waveforms for PC-LIB's signal models.

Part IV

Gravitational-Wave Data Analysis

Chapter 7

Introduction to Denoising Methods

7.1 The noise problem

One of the most common problems in communications and signal processing is noise removal, either in 1D signals or in 2D images. Denoising is the process of removing noise from an image or signal preserving as many structures and details from the original object as possible. The denoising problem is a clear example of an inverse problem, where we start from a set of observations, the noisy data, and we want to recover the original signal. In general, the starting point for solving inverse problems is to consider a model that relates the original with the measured signal and a prior probability density which corresponds to the notion we have about the true signal or image.

Henceforth, we will use the so-called linear degradation model,

$$\boldsymbol{f} = \boldsymbol{A}\boldsymbol{u} + \boldsymbol{n} \ , \tag{7.1}$$

where $\boldsymbol{f}$ is the observed signal, $\boldsymbol{u} \in \mathbb{R}^N$ is the initial "perfect" signal (i.e. noise-free), and $\boldsymbol{A}$ is some linear operator which represents any linear transformation (e.g. sampling or blurring). For the sake of understanding of the adopted mathematical models we will assume that $\boldsymbol{n}$ is Gaussian white noise, (i.e. $\boldsymbol{n}$ is a square integrable function with zero mean). Moreover, throughout this section we will assume that the signals belong to a k-dimensional Euclidean space equipped either with discrete L_1 or L_2 norms.

The current state-of-the-art methods for denoising can be classified in three classes:

- Variational methods, i.e, those based in the Rudin-Osher-Fatemi model.
- Methods based in dictionaries.
- Patched-based methods.

In the present work, we have explored the feasibility of the first two methods to remove noise in the context of GW astronomy, leaving the third one for future analysis.

7.2 Bayesian approach

One of the most common methods to solve inverse problems is the Bayesian approach, where the model is related with the observations through the Bayes theorem,

$$\text{prob}(H|data, I) = \frac{\text{prob}(data|H, I) \times \text{prob}(H|I)}{\text{prob}(data|I)}, \tag{7.2}$$

where H represents the hypothesis one does about the model and I is the other information one assumes as true. The key point of the Bayes theorem is that it relates the probability that the hypothesis is true given the data, with other probabilities that are much simple to assign,e.g. the probability observing the data if the hypothesis is true. The last term on the r.h.s. of Eq. (7.2) is the *prior* probability and represents the knowledge one has about the truth of the hypothesis done in the model before analyzing any data. This probability is modified according to the data through the *likelihood function*, $\text{prob}(data|H, I)$ and yields the *posterior* probability $\text{prob}(H|data, I)$ which represents the knowledge of the hypothesis once the data has been analyzed. Note that the denominator in Eq (7.2) can be omitted as it works as a proportionality constant.

Following Chambolle et al. [2010] we can describe an example of the Bayesian approach in the case of the linear degradation model, Eq. (7.1). In this case, the hypothesis would be that $\boldsymbol{u}$ is the perfect original signal degraded by the noise. In Bayesian inference one should select the prior function which better represents the knowledge about the hypothesis. For convenience, we will suppose that the prior probability can be written as an exponential of a certain generic function $p(u)$,

$$P(u) \propto e^{-p(u)} du. \tag{7.3}$$

Applying the Bayes theorem, it is possible to compute the *a posteriori* probability, $P(u|f)$, which represents the probability that the signal $\boldsymbol{u}$ is present given the observed signal $\boldsymbol{f}$,

$$P(u|f) = \frac{P(f|u)P(u)}{P(f)} \tag{7.4}$$

where $P(f|u)$ is the probability of observing $\boldsymbol{f}$ (known as likelihood) given a real signal $\boldsymbol{u}$. From Eq. (7.1), the noise can be expressed as $\boldsymbol{n} = \boldsymbol{f} - \boldsymbol{Au}$. Since in our model the noise is defined as Gaussian, this probability reads as a Gaussian distribution of variance σ^2,

$$P(f|u) \propto e^{-\frac{1}{2\sigma^2}\sum_{i,j}||f_{i,j}-(Au)_{i,j}||^2} \ . \tag{7.5}$$

Note that this expression is a proportionality since we have omitted the normalization term of the Gaussian distribution. With these ingredients, it is possible to obtain an expression for the posterior probability $P(u|f)$,

$$P(u|f) = \frac{1}{Z}e^{-p(u)}e^{-\frac{1}{2\sigma^2}\sum_{i,j}||f_{i,j}-(Au)_{i,j}||^2}, \tag{7.6}$$

where Z is a normalization factor,

$$P(f) = Z = \int e^{-(p(u)+\frac{1}{2\sigma^2}\sum_{i,j}||f_{i,j}-(Au)_{i,j}||^2)}du \ . \tag{7.7}$$

The "maximum a posteriori" (MAP) signal reconstruction tries to find the signal which maximizes this probability or, in other words, which solves the minimization problem,

$$u = \min_u \left\{ p(u) + \frac{1}{2\sigma^2}\sum_{i,j}||f_{i,j} - (Au)_{i,j}||^2 \right\}. \tag{7.8}$$

The signal with the highest probability obtained by the resolution of Eq. (7.8) might be strange, even if the model is built properly (see Chambolle et al. [2010] for an example). As pointed out by Chambolle et al. [2010], a better approximation would be computing the *expectation* of u,

$$\mathrm{E}[u|f] = \frac{1}{Z}\int_u ue^{-p(u)}e^{-\frac{1}{2\sigma^2}\sum_{i,j}||f_{i,j}-(Au)_{i,j}||^2}du. \tag{7.9}$$

This calculation is hard to compute [Louchet 2008] and requires stochastic algorithms like Monte Carlo methods. However, the idea of minimizing an energy such as in the MAP approximation leads to the mathematical form which is used to formulate most of the inverse problems (see Chambolle et al. [2010] and references therein).

The Bayesian approach has two main applications in GW astronomy, parameter estimation [Balasubramanian, Sathyaprakash, and Dhurandhar 1996, Veitch et al. 2015] and model selection [Powell et al. 2016].

7.3 Total-variation-based methods

7.3.1 Variational models for denoising

Leaving aside the Bayesian approach, the problem of signal denoising can be expressed as the estimation of the true signal u as a function whose square of the L^2-distance to the observed noisy signal f is the variance of the noise σ^2, i.e.

$$||\boldsymbol{u} - \boldsymbol{f}||^2_{\mathrm{L}_2} = \sigma^2 \ , \tag{7.10}$$

where σ denotes the standard deviation of the noise. For simplicity reasons, here we have considered the linear operator A as the identity operator.

Classical models and algorithms for solving the denoising problem are based on least squares, Fourier series and other L_2-norm approximations. A least-squares problem can be solved by computing the solution of the associated normal equations, which is a linear system of equations where the unknowns are the coefficients of a linear combination of polynomials or a wavelet basis [Irani and Peleg 1993]. The main drawback of this technique is that the results are contaminated by Gibbs' phenomena (ringing) and/or smearing near the edges (see Marquina and Osher [2008a] and references therein). Moreover, the linear system to be solved is large (related to the size of the sample of the observed signal f) and ill-conditioned, i.e. close to singular, further hindering the resolution of the problem.

The usual approach to overcome these problems is to regularize the least-squares problem using an auxiliary energy $\mathcal{R}(\boldsymbol{u})$, called regularisation term. This energy corresponds to the a priori probability density $p(\boldsymbol{u})$, and the solution can be found by solving the following constrained variational problem [1]

$$\min_u \quad \mathcal{R}(\boldsymbol{u}) \text{ subject to } ||\boldsymbol{f} - \boldsymbol{u}||_2^2 = \sigma^2. \tag{7.11}$$

This general model can be applied to 1D signals, 2D images or multidimensional volume data.

The above variational problem has a unique solution when the energy $\mathcal{R}(\boldsymbol{u})$ is convex. We will assume from now on that $\mathcal{R}(\boldsymbol{u})$ is convex. The constrained variational problem can be formulated as an unconstrained variational problem

[1] To simplify the notation, we write $||\cdot||_2$ or simply $||\cdot||$ for the L_2 norm and $|\cdot|_1$ or simply $|\cdot|$ for the L_1 norm.

using the Tikhonov regularisation, which adds the constraint weighted by a positive Lagrange multiplier $\lambda > 0$ (also unknown) to the energy $\mathcal{R}(\boldsymbol{u})$

$$\boldsymbol{u} = \underset{\boldsymbol{u}}{\operatorname{argmin}} \left\{ \mathcal{R}(\boldsymbol{u}) + \frac{\lambda}{2}\mathcal{F} \right\} \ , \tag{7.12}$$

where $\mathcal{F}$ is the so-called fidelity term $||\boldsymbol{f} - \boldsymbol{u}||_2^2$.

There exists a unique value of $\lambda > 0$ such that the unique solution $\boldsymbol{u}$ matches the constraint. The Lagrange multiplier $\lambda > 0$ becomes a scale parameter in the sense that larger values of λ allow to recover finer scales in a scale space determined by the regulariser functional $R(\boldsymbol{u})$. This can be easily understood as λ controls the relative importance of the fidelity term.

The proper values of $\mathcal{R}$ should simultaneously ensure some spatial regularity, but also preserve the steep gradients. If $\mathcal{R}(\boldsymbol{u}) := \int |\nabla \boldsymbol{u}|_1^2$ where the integral is extended to the domain of the signal, either discrete or continuous, then model (7.12) becomes the so-called Wiener filter. In order to compute the solution in this case we solve the associated Euler-Lagrange equation

$$\Delta \boldsymbol{u} + \lambda(\boldsymbol{f} - \boldsymbol{u}) = 0 \ , \tag{7.13}$$

under homogeneous Neumann boundary conditions. In the previous equation, and in the definition of the energy $\mathcal{R}(\boldsymbol{u})$, Δ and ∇ stand, respectively, for the (discrete) Laplacian and gradient operators. Eq. 7.13 corresponds to a nondegenerate second-order, linear, elliptic differential equation, which is easy to solve due to differentiability and strict convexity of the energy term. We note that this equation satisfies the conditions that guarantee uniqueness of the solution (see [Taylor 1997, Evans 1997, Protter and Weinberger 2012, Cordero-Carrión et al. 2009] and references therein). Indeed, the equation can be efficiently solved by means of the Fast Fourier Transform (FFT). The choice of a quadratic energy for the regulariser function makes the variational problem more tractable. It encourages Fourier coefficients of the solution to decay towards zero, surviving the ones representing the processed signal $\boldsymbol{u}$. However this good behaviour is no longer valid when noise is present in the signal. Noise amplifies high frequencies and the recovered smooth solution $\boldsymbol{u}$ prescribed by the model contains spurious oscillations near steep gradients or edges (see Marquina and Osher [2000], Vogel and Oman [1998], and Osher et al. [2005]). The Wiener filter procedure reduces noise by shrinking Fourier coefficients of the signal towards zero but adds spurious oscillations due to the Gibbs' phenomenon.

In order to avoid the aforementioned problems arising by using quadratic variational models, Rudin, Osher, and Fatemi [1992] (ROF hereafter) proposed using the TV norm as regularizing functional for the variational model for

denoising Eq. (7.12)

$$\mathrm{TV}(\boldsymbol{u}) = \int |\nabla \boldsymbol{u}|_1 \tag{7.14}$$

where the integral is defined on the domain of the signal. The ROF model is based on solving the following variational problem for denoising:

$$\boldsymbol{u} = \underset{\boldsymbol{u}}{\operatorname{argmin}} \left\{ \mathrm{TV}(\boldsymbol{u}) + \frac{\lambda}{2} ||\boldsymbol{u} - \boldsymbol{f}||_2^2 \right\} \ . \tag{7.15}$$

The TV norm energy is essentially the L_1-norm of the gradient of the signal. Although many L_1-based norms have been usually avoided because of its lack of differentiability, the way the L_1-norm is used in the ROF model has provided a great success in denoising problems. The ROF model allows to recover edges of the original signal removing noise and avoiding ringing. The parameter $\lambda > 0$ runs a different scale space than the Wiener model. Since the energy is convex there is a unique optimal value of the Lagrange multiplier λ (scale) for which equation (7.10) is satisfied. When the standard deviation of the noise is unknown a heuristic estimation of λ is needed to find the optimal value. Indeed, if we choose a large value of λ the ROF model will remove very little noise, while finer scales will be destroyed if small values of λ are chosen instead.

The use of L_1-norm related energies for least-squares regularisation has been popularised following the pioneering contribution of Rudin, Osher and Fatemi. For example, soft thresholding is a denoising algorithm related to L_1-norm minimisation introduced by Donoho and Johnstone [1994] and Donoho [1995]. The L_1-norm regularisation selects a unique *sparse* solution, i.e., solutions with few nonzero elements. This property is essential for *compressive sensing* problems, (see Donoho [2006] and Candès, Romberg, and Tao [2006]). An earlier application of penalising (regularizing) with the L_1-norm is the LASSO regression (from Least Absolute Shrinkage and Selection Operator) proposed in the seminal paper of Tibshirani [1996].

Since the ROF model uses the TV-norm the solution is the only one with the *sparsest* gradient. Thus, the ROF model reduces noise by *sparsifying* the gradient of the signal and avoiding spurious oscillations (ringing).

Similar to the ROF model, the $\mathrm{TV} - \mathrm{L}_1$ norm [Chan and Esedoglu 2005] employs the L_1 norm also for the fidelity term $\mathcal{F}$,

$$\boldsymbol{u} = \underset{\boldsymbol{u}}{\operatorname{argmin}} \left\{ \mathrm{TV}(\boldsymbol{u}) + \lambda |\boldsymbol{u} - \boldsymbol{f}|_1 \right\} \ . \tag{7.16}$$

In contrast with the ROF model (7.15), the $\mathrm{TV} - \mathrm{L}_1$ model is not strictly convex. Therefore, there is no unique global minimiser. In general, however, this last formulation has some interesting advantages. It is more effective removing

impulse noise in images than ROF. It is also contrast invariant, i.e., if $\boldsymbol{u}$ is a solution of Eq. (7.16) for a given input image $\boldsymbol{f}$, the solution will be $c\boldsymbol{u}$ (c is a constant) for an input image $c\boldsymbol{f}$. If these advantages remain also for GW signals is something that we plan to study in the near future.

7.3.2 Algorithms for TV-based denoising

In the literature there exists a large number of methods to solve the ROF and $\mathrm{TV} - \mathrm{L}_1$ models. As examples, we can cite the following:

- Explicit time marching.
- Linearization of the Euler-Lagrange equation.
- Nonlinear primal-dual method.
- Duality based methods.
- Non-linear multigrid methods.
- First-order schemes from convex optimization.
- Second-order cone programming.
- Graph cut methods.

For details about these methods see Pock et al. [2008] and references therein. In this work we focus our research in Variational Methods, i.e, methods based on the linearisation of the Euler-Lagrange equation, due to their adaptability to general problems and their computational efficiency, by the simplicity to parallelise them.

7.3.2.1 The regularized ROF model

The standard method to solve nonlinear smooth optimisation problems is to compute the solution of the associated Euler-Lagrange equation. The ROF model consists of a nonsmooth optimisation problem and the associated Euler-Lagrange equation can be expressed as

$$\nabla \cdot \frac{\nabla \boldsymbol{u}}{|\nabla \boldsymbol{u}|_1} + \lambda(\boldsymbol{f} - \boldsymbol{u}) = 0 \ , \tag{7.17}$$

where the differential operator becomes singular and has to be defined properly when $|\nabla \boldsymbol{u}|_1 = 0$ (see Meyer [2001]).

The first algorithm we use in this work is called regularised ROF algorithm (rROF hereafter) and was developed by Vogel and Oman [1996]. This algorithm computes an approximate solution of the ROF model by smoothing the total variation energy. Since the Euler-Lagrange derivative of the TV-norm is not well defined at points where $\nabla u = 0$, the TV functional is slightly perturbed as

$$\mathrm{TV}_\beta(\boldsymbol{u}) := \int \sqrt{|\nabla \boldsymbol{u}|_1^2 + \beta}\ , \tag{7.18}$$

where β is a small positive parameter. We will use the expression

$$\int |\nabla \boldsymbol{u}|_\beta \tag{7.19}$$

with the notation

$$|\boldsymbol{v}|_\beta = \sqrt{|\boldsymbol{v}|^2 + \beta} \tag{7.20}$$

for $\boldsymbol{v} \in \mathbb{R}^p$ where p is the dimension of the signal.

Therefore, the rROF model in terms of the small positive parameter $\beta > 0$ reads as

$$\boldsymbol{u} = \underset{\boldsymbol{u}}{\operatorname{argmin}} \left\{ \mathrm{TV}_\beta(\boldsymbol{u}) + \frac{\lambda}{2} ||\boldsymbol{u} - \boldsymbol{f}||_{\mathrm{L}_2}^2 \right\}\ , \tag{7.21}$$

and the associated Euler-Lagrange equation is

$$\nabla \cdot \frac{\nabla \boldsymbol{u}}{|\nabla \boldsymbol{u}|_\beta} + \lambda(\boldsymbol{f} - \boldsymbol{u}) = 0\ . \tag{7.22}$$

When the value of β is small, the problem turns nearly degenerated and the algorithm becomes slow in flat regions. In contrast, when β is large, the ROF model cannot preserve sharp discontinuities.

Assuming homogeneous Neumann boundary conditions, Eq. (7.22) becomes a nondegenerate second order nonlinear elliptic differential equation whose solution is smooth. In order to solve the above equation we use conservative second-order central differences for the differential operator and point values for the source term. The approximate solution will be obtained by means of a nonlinear Gauss-Seidel iterative procedure that uses as initial guess the observed signal f.

7.3.2.2 Primal-Dual Algorithm

The TV norm admits an alternative (dual) formulation,

$$|\nabla \boldsymbol{u}| = \max_p \{\boldsymbol{p} \cdot \nabla \boldsymbol{u} : ||\boldsymbol{p}|| \leq 1\}. \tag{7.23}$$

Based on this dual formulation, some authors [Chan, Golub, and Mulet 1999, Carter 2001, Chambolle 2004] proposed the so-called primal-dual formulation of

the ROF model,

$$\min_{\boldsymbol{u}} \max_{||\boldsymbol{p}||\leq 1} \left\{ \int \boldsymbol{p} \cdot \nabla \boldsymbol{u} + \frac{1}{2\lambda} ||\boldsymbol{f} - \boldsymbol{u}||^2 \right\}, \tag{7.24}$$

which is convex and has the optimality condition with respect to u given by,

$$\boldsymbol{u} = \boldsymbol{f} + \lambda \nabla \cdot \boldsymbol{p}. \tag{7.25}$$

Combining Eqs. (7.24) and (7.25) we obtain the dual ROF model,

$$\min_{||\boldsymbol{p}||\leq 1} \left\{ - \int \boldsymbol{p} \cdot \nabla \boldsymbol{f} + \frac{\lambda}{2} (\nabla \cdot \boldsymbol{p})^2 \right\}. \tag{7.26}$$

The associated Euler-Lagrange equation is,

$$-\nabla(\boldsymbol{f} + \lambda \nabla \cdot \boldsymbol{p}) = 0 \ , \ ||\boldsymbol{p}|| \leq 1. \tag{7.27}$$

This last equation does not become singular at any point, so it is continuously differentiable. However, the presence of the constrain $||p|| \leq 1$ requires sophisticated optimization techniques to find the solution. The most popular method due to its simplicity is the projected gradient descend, proposed by Chambolle [2005]. This algorithm has very interesting properties including robustness and fast convergence.

7.3.2.3 The Split Bregman Method

The next algorithm we shall use is the so-called "Split Bregman Method" (SB hereafter) proposed in Goldstein and Osher [2009]. SB is a flexible algorithm for solving non-differentiable convex minimisation problems and it is especially efficient for problems with L_1 or TV regularisation. The method is based on an iterative alternating procedure that splits the approximation of the minimiser into two steps: first, solving the least-squares minimisation and second, performing direct minimisation of the TV energy using the "shrinkage function" and freezing the fidelity term computed at the approximation obtained in the first step.

The splitting process is combined with the Bregman iterative refinement [Bregman 1967]. The Bregman iterative procedure can be applied to a general fidelity term. Let us assume that $\mathcal{E}(u)$ is a nonnegative convex energy and we wish to solve the following constrained variational problem:

$$\min_{\boldsymbol{u}} \quad \mathcal{E}(\boldsymbol{u}) \text{ subject to } \boldsymbol{u} = \boldsymbol{A}\boldsymbol{b} \tag{7.28}$$

where $\boldsymbol{b}$ is a vector and A is a linear operator as in Eq. (7.1) [2]. We rewrite the variational problem as an unconstrained optimisation by introducing a Lagrange

[2] Here we recover the operator $\boldsymbol{A}$ to account for the general formalism.

multiplier $\lambda > 0$ that weights the influence of the fidelity term as

$$\boldsymbol{u} = \underset{\boldsymbol{u}}{\operatorname{argmin}} \left\{ \mathcal{E}(\boldsymbol{u}) + \frac{\lambda}{2} ||\boldsymbol{A}\boldsymbol{b} - \boldsymbol{u}||_2^2 \right\} . \tag{7.29}$$

If we choose a small value of λ then the solution of the variational problem does not accurately enforce the constraint. The solution we need is to let λ be large. Alternatively, we shall use the following Bregman iterative refinement to enforce the constraint $Au = b$ accurately using, in contrast with the previous formulation (7.29), a fixed small value for λ:

$$\boldsymbol{u}^{k+1} = \underset{\boldsymbol{u}}{\operatorname{argmin}} \left\{ \mathcal{E}(\boldsymbol{u}) + \frac{\lambda}{2} ||\boldsymbol{b}^k - \boldsymbol{u}||_2^2 \right\} , \tag{7.30}$$

$$\boldsymbol{b}^{k+1} = \boldsymbol{b}^k + \boldsymbol{r}^k, \tag{7.31}$$

where $\boldsymbol{r}^k = \boldsymbol{b}^k - \boldsymbol{u}^{k+1}$ is the residual error and k is the index of the iteration, and again we set $\boldsymbol{A} = \boldsymbol{I}$ for simplicity reasons. Roughly speaking, the Bregman iteration involves adding the residual error (of the fidelity term) back to the constraint to solve a new variational problem in each iteration.

Next we shall sketch the SB method for the particular case of the ROF model applied to the one dimensional signals that we will use in our numerical experiments. The SB method to solve the ROF model combines the Bregman iterative procedure described by Eqs. (7.30) and (7.31) with the decoupling of the TV variational problem into L_1 and L_2 portions of the energy to be minimized. Each ROF problem appearing in every Bregman iteration is solved by splitting the L_1 and L_2 terms and minimizing them separately.

To start the procedure we solve each ROF problem in Eq. (7.30) by introducing a new variable $\boldsymbol{d}$ which represents the one-dimensional gradient $\nabla_x \boldsymbol{u}$. The idea of the SB method is to apply operator splitting and use Bregman iteration to solve the constrained minimisation problem,

$$\underset{\boldsymbol{d},\boldsymbol{u}}{\operatorname{argmin}} \left\{ |\boldsymbol{d}| + \frac{\lambda}{2} ||\boldsymbol{f} - \boldsymbol{u}||^2 \right\} \text{ subject to } \boldsymbol{d} = \nabla \boldsymbol{u} \tag{7.32}$$

We set the notation for the new constraint,

$$s(\boldsymbol{b}, \boldsymbol{u}, \boldsymbol{d}) := ||\boldsymbol{b} + (\nabla_x \boldsymbol{u}) - \boldsymbol{d}||_2^2 .$$

Then, we formulate the following Bregman iterative procedure applied on the new constraint (see Goldstein and Osher [2009]):

$$(\boldsymbol{u}^{k+1}, \boldsymbol{d}^{k+1}) = \underset{\boldsymbol{u}^k, \boldsymbol{d}^k}{\operatorname{argmin}} \left\{ |\boldsymbol{d}^k| + \frac{\lambda}{2} ||\boldsymbol{f} - \boldsymbol{u}^k||_{\mathrm{L}_2}^2 + \frac{\mu}{2} s(\boldsymbol{b}^k, \boldsymbol{u}^k, \boldsymbol{d}^k) \right\} \tag{7.33}$$

$$\boldsymbol{b}^{k+1} = \boldsymbol{b}^k + (\nabla_x \boldsymbol{u}^{k+1}) - \boldsymbol{d}^{k+1} , \tag{7.34}$$

where we set $s(\boldsymbol{b}^k, \boldsymbol{u}^k, \boldsymbol{d}^k) := \boldsymbol{b}^k + (\nabla_x \boldsymbol{u}^k) - \boldsymbol{d}^k$ and where we have included the constrain weighted by another Lagrange multiplier μ.

Thus, we can iteratively minimise with respect to u and d separately and the SB iterative procedure will read as follows:

$$\boldsymbol{u}^{k+1} = \underset{\boldsymbol{u}^k}{\operatorname{argmin}} \left\{ \frac{\lambda}{2} ||\boldsymbol{f} - \boldsymbol{u}^k||_2^2 + \frac{\mu}{2} s(\boldsymbol{b}^k, \boldsymbol{u}^k, \boldsymbol{d}^k) \right\} \tag{7.35}$$

$$\boldsymbol{d}^{k+1} = \underset{\boldsymbol{d}^k}{\operatorname{argmin}} \left\{ |\boldsymbol{d}^k| + \frac{\mu}{2} ||\boldsymbol{b}^k + (\nabla_x \boldsymbol{u}^{k+1}) - \boldsymbol{d}^k||_2^2 \right\} \tag{7.36}$$

$$\boldsymbol{b}^{k+1} = \boldsymbol{b}^k + (\nabla_x \boldsymbol{u}^{k+1}) - \boldsymbol{d}^{k+1} . \tag{7.37}$$

Since the two parts are decoupled, they can be solved independently. The energy of the first step is smooth (differentiable) and it can be solved using common techniques such as the Gauss-Seidel method. On the other hand, $\boldsymbol{d}$ can be solved using shrinkage operators

$$\boldsymbol{d}^{k+1} = \text{shrink}(\boldsymbol{b}^k + (\nabla_x \boldsymbol{u}^{k+1}), 1/\lambda) , \tag{7.38}$$

$$\text{shrink}(x, \gamma) \equiv \frac{x}{|x|} * \max(|x| - \gamma, 0) . \tag{7.39}$$

We only use one iteration for the splitting steps and the final algorithm only consists of just one loop (see Goldstein and Osher [2009] for a detailed discussion).

The SB algorithm we will use in this work is written as follows:

- Initial guess: $\boldsymbol{u}^0 = f$, $\boldsymbol{d}^0 = 0$ and $\boldsymbol{b}^0 = 0$
- while $||\boldsymbol{u}^k - \boldsymbol{u}^{k-1}||_2 >$ tolerance

  $\boldsymbol{u}^{k+1} = G^k$

  $\boldsymbol{d}^{k+1} = \text{shrink}(\boldsymbol{b}^k + (\nabla_x \boldsymbol{u}^{k+1}), 1/\mu)$

  $\boldsymbol{b}^{k+1} = \boldsymbol{b}^k + (\nabla_x \boldsymbol{u}^{k+1}) - \boldsymbol{d}^{k+1}$
- end

where the Gauss-Seidel step can be expressed as the loop

- for j

  $$G_j^k = \frac{\mu}{\lambda + 2\mu}(u_{j+1}^k + u_{j-1}^k - (d_j^k - d_{j-1}^k) + (b_j^k - b_{j-1}^k)) + \frac{\lambda}{\lambda + 2\mu} f_j$$
- end j

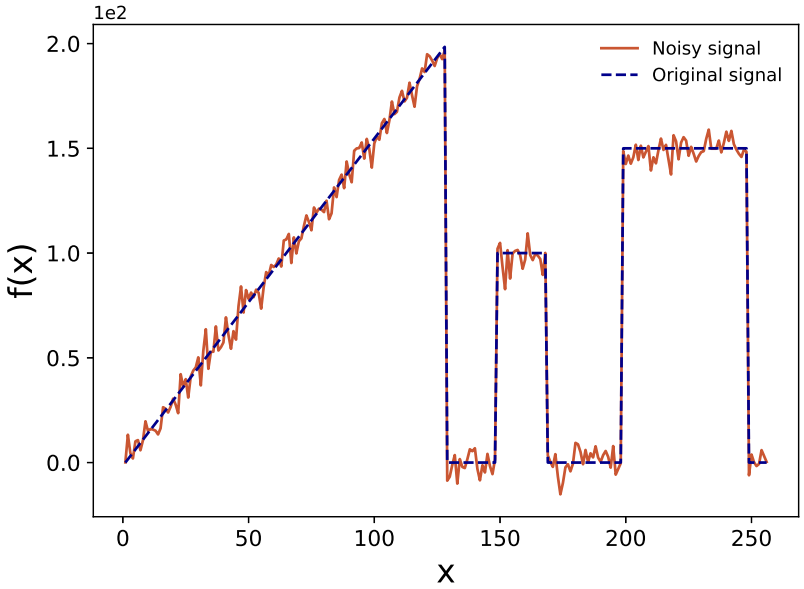


Figure 7.1 Test signal used to assess the rROF and SB algorithms. The original signal is shown in blue while the orange line depicts the noisy signal with $\sigma = 5$. Note the different size of the gradients with which this function has been designed.

and $(\nabla_x u^{k+1})_j = u_{j+1}^{k+1} - u_j^{k+1}$ where j runs the component positions of the discretisation. A tolerance of 10^{-2} should be enough to obtain appropriate results in most cases. As discussed in Getreuer [2012], the value of the Lagrangian multiplier μ should be selected carefully. The best choice is where both subproblems (for $\boldsymbol{u}$ and $\boldsymbol{d}$) converge quickly and are well-conditioned.

7.3.3 Numerical examples

In this section we show the results of some simple numerical tests to study the properties of the rROF and SB methods before their actual application for GW denoising in Chapters 8 and 9. For this thesis we have focused on these two algorithms due to the simplicity of their computational implementation, as compared to the more involved primal-dual algorithm. We leave for future work the assessment of the latter in the context of GW denoising.

To perform the tests we select the signal shown in Fig. 7.1. This type of signal has been used in the literature before [Marquina and Osher 2000]. We add Gaussian noise to the signal with a given variance σ^2.

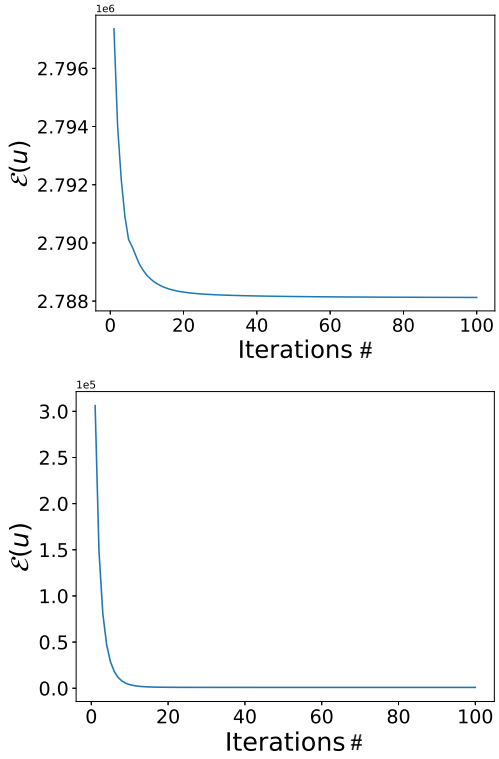


Figure 7.2 Evolution of the energy with the number of iterations for the rROF algorithm (top panel) and for the SB method (bottom panel). The algorithms converge after a few tens of iterations.

First of all we test the convergence at a given value of λ. Since the energy used in both methods is different, the values of λ in each case are not comparable. However, both methods should show the similar behaviour with variations in λ. Fig. 7.2 shows how the energy of the ROF model, $\mathcal{E}(\boldsymbol{u}) = \mathcal{R}(\boldsymbol{u}) + \lambda\mathcal{F}(\boldsymbol{u})$, decreases with the number of iterations. At a certain number of iterations the energy converges. The convergence is slightly faster for the rROF method. We turn next to analyse the dependence with the regularisation parameter λ. Fig. 7.3 displays the denoising results for two different values of λ. As expected, when λ is small ($\lambda = 0.01$ in this case) the denoised signal is very regular but some features of the signal are lost. On the contrary, for large values ($\lambda = 100$ in the plot), the fidelity term dominates and practically there is no noise removal. In the case of the SB method in particular, shown in the bottom panel of Fig. 7.3, the excess of regularisation is much more noticeable and, in addition, the SB algorithm also introduces the staircase effect [Chan, Marquina, and Mulet 2000].

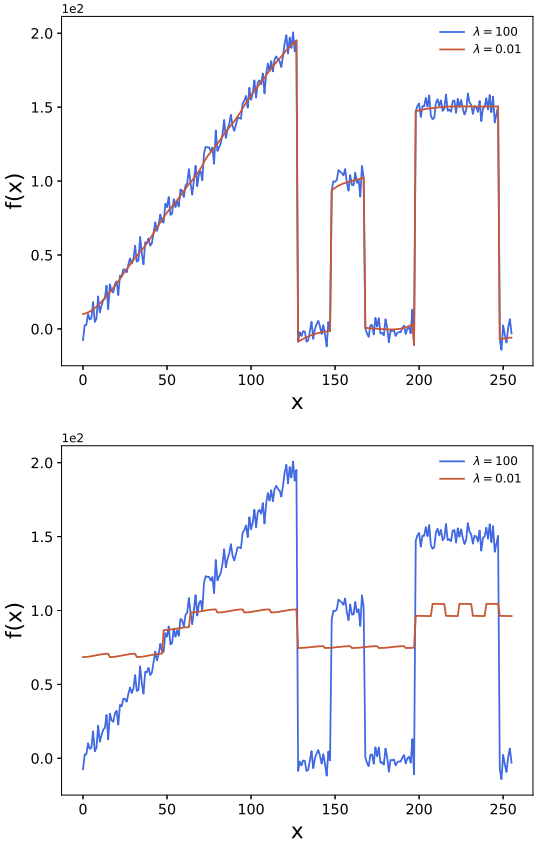


Figure 7.3 Denoising results for two values of λ, namely 0.01 (orange lines) and 100 (blue lines) The top panel corresponds to the rROF method and the bottom panel to the SB method. Note the staircase effect in the case of the SB method.

7.3.4 Estimation of the regularisation parameter

The denoised signal resulting from solving the ROF problem depends on the selected values for the Lagrange multiplier λ. This parameter controls the trade-off between the regularised solution and the fit that it provides to the given data f. Therefore, it is worth to find a strategy to determine the *optimal* value λ that produces the best denoising results. Basically there exist two kind of methods to achieve this goal, those that require some knowledge about the noise and those that do not. The *discrepancy principle* is a straightforward concept for parameter tuning, according to which λ is selected to match the noise variance σ^2. Unfortunately, the relation between λ and σ^2 is not direct, so it is not possible to derive a closed-form formula to calculate the optimal

value of λ for a given variance of noise. Chambolle and Lions [1997] proposed an iterative algorithm that converges to the unique λ such that $||f-u||_2^2 = \sigma^2$. The main idea is to solve the constrained variational problem to obtain λ,

$$\underset{\boldsymbol{u}}{\operatorname{argmin}} \left\{ \mathrm{TV}(\boldsymbol{u}) \text{ subject to } ||\boldsymbol{u}-\boldsymbol{f}||^2 = \sigma^2 \right\}. \tag{7.40}$$

For example, using the dual-primal method, this algorithm reads as follows:

$$\begin{aligned} \text{start with } \lambda &= \lambda_0 \\ \boldsymbol{u} &= \underline{\mathrm{f}} + \lambda \nabla \cdot \boldsymbol{p} \\ \lambda &= \frac{||\boldsymbol{u}-\boldsymbol{f}||}{N\sigma^2}\lambda \end{aligned} \tag{7.41}$$

where λ_0 is some preselected positive value. This method is fairly accurate and converges in a few iterations. However, it fails when the variance of noise is not well determined or not determined. The algorithm works with the other methods we have discussed previously. In the algorithm one simply has to change the method employed to calculate the solution for u in Eq. (7.41). A more detailed study about the performance of some algorithms based in this principle can be found in Galatsanos and Katsaggelos [1992].

If the variance of the noise is not accessible, it is still possible to find the optimal value of the regularisation parameter. One of the methods to do so that has gained attention in recent years is the *L-curve* method. As described in Hansen [1992], "The L-curve is a log-log plot between the squared norm of the regularised solution and the squared norm of the regularised residual for a range of values of regularisation parameter". The plot forms an "L" and hence the name. The optimal value of λ is the one that is placed at the vertex of the curve, i.e. the point where the curve achieves its maximum curvature. Therefore, the L-curve criterium tries to determine the position of this point. A practical way to determine the vertex was given by Hansen, Jensen, and Rodriguez [2007] who proposed to first construct a series of L-curves, removing and increasing the number of points. The optimal point is selected by taking from this list the last point before the part of the L-curve where the curve becomes flat is reached. Altought this criterium can be used in general because it does not rely on any a priori information, the optimal point can be difficult to determine and the computational cost of solving the ROF problem for a range of values of λ can be significant.

The last kind of methods are the "heuristic methods". In this case, the optimal value of the parameter is defined as the value that minimises (or maximises) a given quality function, as the mean squared error (MSE). The drawback of these

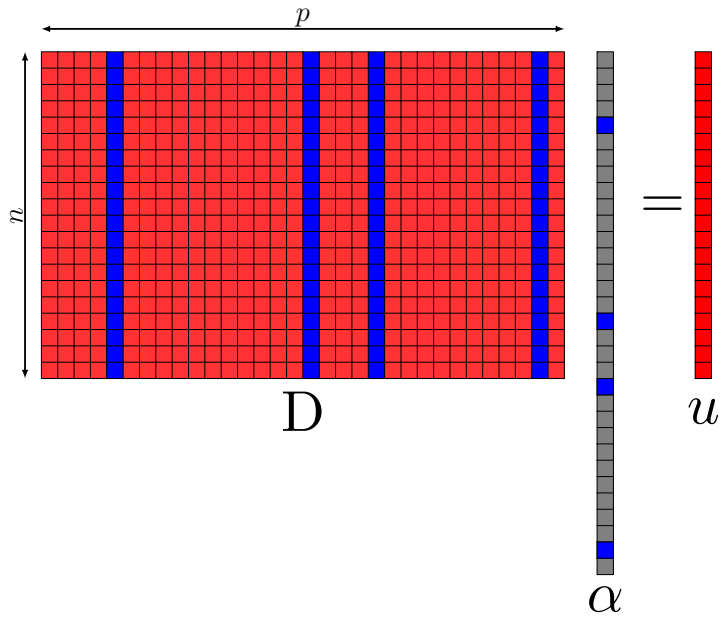


Figure 7.4 Graphical representation of the sparse problem, in which only a few columns of the matrix D are selected by the vector $\boldsymbol{\alpha}$ to produce the reconstructed signal u.

methods is that a "learning" step using known signals is required to determine the range of optimal values, which could be computationally expensive. Once the range is selected no further study is needed. However, the results will depend of how accurate the "learning" step is done.

7.4 The dictionary-learning problem

7.4.1 Sparse Reconstruction over a fixed dictionary

Let us consider again the linear degradation model 7.1 and the energy minimisation given by the MAP formalism. In dictionary-based methods, the denoising is performed assuming that the true signal u can be represented as a linear combination of the columns of a matrix $\mathbf{D}$ called dictionary. The columns of $\mathbf{D}$ are called basis vectors or *atoms*.

If signal $\mathbf{u}$ is a vector $\mathbf{u} \in \mathbb{R}^n$ one can say that the dictionary $\mathbf{D} = [d_1, \cdots, d_p] \in \mathbb{R}^{n \times p}$ is adapted to $\mathbf{u}$ if it can be represented with a *few* columns of $\mathbf{D}$,i.e., a few*atoms* In other words, there exists a "sparse vector" $\boldsymbol{\alpha}$ in $\mathbb{R}^p$ such that $\mathbf{u} \sim \mathbf{D}\boldsymbol{\alpha}$. Therefore, the sparse decomposition problem reads as follows,

$$\boldsymbol{\alpha} = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \left\{ \frac{1}{2} ||f - \mathbf{D}\boldsymbol{\alpha}||_2^2 \right\} \quad \text{subject to } \psi(\boldsymbol{\alpha}), \tag{7.42}$$

where ψ is a prior that induces sparsity. The classical dictionary learning techniques [Olshausen and Field 1997, Aharon, Elad, and Bruckstein 2006] try to solve the variational problem associated with Eq. (7.42) given by,

$$\boldsymbol{\alpha} = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \left\{ \frac{1}{2} ||f - \mathbf{D}\boldsymbol{\alpha}||_2^2 \right\} \quad \text{subject to } \ ||\boldsymbol{\alpha}||_0 \leq \mathrm{L}, \tag{7.43}$$

where $||\cdot||_0$ is the L_0-norm, which is chosen to ensure that we have the solution with the fewest number of nonzero coefficients. The L_0-norm is simply the number of nonzero components of the vector, which should be lower than a given value L. This constrained variational problem can be formulated as an unconstrained variational problem adding the L_0–norm term as a penalty term weighted by a Lagrangian multiplier λ,

$$\boldsymbol{\alpha} = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \left\{ ||\mathbf{D}\boldsymbol{\alpha} - \mathbf{f}||_2^2 + \lambda ||\boldsymbol{\alpha}||_0 \right\}. \tag{7.44}$$

This problem is not convex and is NP-hard [3] so, in practice, it cannot be solved in linear time [Marcellin et al. 2000]. Algorithms that produce and approximate solutions to this problem have been proposed in the past. The simplest one is matching pursuit (MP) [Mallat and Zhang 1993]. The algorithm selects the atom with maximum correlation with the residual, then updates the residual and the coefficients. It proceeds iteratively while the L_0–norm term is lower than the parameter L. Similar to MP, in the orthogonal matching pursuit (OMP) method [Pati, Rezaiifar, and Krishnaprasad 1993], an atom can only be selected once. It is, however, more difficult to implement efficiently than MP (see Chen, Donoho, and Saunders [2001] and references therein).

The variational problem defined by Eq. (7.44) can be reformulated into a convex variational formulation by substituting the L_0-norm by the nondifferentiable convex L_1–norm in the total energy. The regularization in the L_1–norm promotes zeros in the components of the vector coefficient $\boldsymbol{\alpha}$. This problem can be solved in linear time and the solution found is the sparsest one in most cases. The variational problem thus stands as,

$$\boldsymbol{\alpha} = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \left\{ ||\mathbf{D}\boldsymbol{\alpha} - \mathbf{f}||_2^2 + \lambda ||\boldsymbol{\alpha}||_1 \right\}, \tag{7.45}$$

which is known as *basis pursuit* [Chen, Donoho, and Saunders 2001] or *LASSO* [Tibshirani 1996].

[3] A problem is in the NP class if it can be solved in non-deterministic polynomial-time.

In general the solution of this problem has not a closed form, except when $\mathbf{D}$ is orthonormal. The derivation of the optimality conditions for the solution is as follows (note that the name of variables is slightly different than Eq. (7.45))

$$\begin{aligned} f(\alpha) &= ||\boldsymbol{y} - \mathbf{D}\boldsymbol{\alpha}||_2^2 + \lambda|\boldsymbol{\alpha}|_1 && (7.46)\\ &= \boldsymbol{y}\boldsymbol{y}^\intercal - 2\boldsymbol{\alpha}^\intercal\mathbf{D}^\intercal\boldsymbol{y} + \boldsymbol{\alpha}^\intercal\mathbf{D}^\intercal\mathbf{D}\boldsymbol{\alpha} + \lambda|\boldsymbol{\alpha}|_1. && (7.47)\end{aligned}$$

Taking the gradient we obtain,

$$\nabla f(\boldsymbol{\alpha}) = -2\mathbf{D}^\intercal\boldsymbol{y} + 2\mathbf{D}^\intercal\mathbf{D}\boldsymbol{\alpha} + \nabla(\lambda|\boldsymbol{\alpha}|_1). \tag{7.48}$$

At this point it is easier to look for the j-th component.

$$\frac{\partial f}{\partial \alpha_j} = -2D_j^\intercal y + 2(D^\intercal D)_j\boldsymbol{\alpha} + \frac{\partial}{\partial \alpha_j}(\lambda|\boldsymbol{\alpha}|_1). \tag{7.49}$$

With $\frac{\partial f}{\partial \alpha_j} = 0$,

$$2(D^\intercal D)_j\alpha = 2D_j^\intercal y - \frac{\partial}{\partial \alpha_j}(\lambda|\alpha_j|_1). \tag{7.50}$$

The subdifferential of the convex L_1-norm is a product of intervals,

$$\partial(|\alpha|) = J_1 \times \cdots \times J_n, \quad J_k(\alpha) = \begin{cases} \text{sign } \alpha & \text{if } \alpha \leq 0 \\ [-1, 1] & \text{if } \alpha = 0 \end{cases} \tag{7.51}$$

Therefore, α^* is the solution of Eq. (7.45) if and only if for all j in $\{1, \cdots, p\}$,

$$\begin{cases} D_j^\intercal(\boldsymbol{y} - \mathbf{D}\boldsymbol{\alpha}^*) = \text{sign } (\alpha_j^*) & \text{if } \alpha_j^* \leq 0 \\ |D_j^\intercal(\boldsymbol{y} - \mathbf{D}\boldsymbol{\alpha}^*)| \leq \lambda & \text{if } \alpha_j^* = 0 \end{cases} \tag{7.52}$$

Finally, it is worth pointing out the existence of an alternative formulation known as *elastic-net* [Zou and Hastie 2005],

$$\boldsymbol{\alpha} = \underset{\boldsymbol{\alpha}}{\text{argmin}} \left\{ ||\mathbf{D}\boldsymbol{\alpha} - \mathbf{f}||_2^2 + \lambda_1||\boldsymbol{\alpha}||_1 + \frac{\lambda_2}{2}||\boldsymbol{\alpha}||_2^2 \right\}. \tag{7.53}$$

Elastic-net adds a L_2–norm penalty for stability reasons, i.e., the calculated approximate representation depends on the data as a Lipschitz function.

7.4.1.1 Algorithms to solve the LASSO

We turn now to describe briefly some of the different algorithms developed in recent years to solve the LASSO. As indicated before, the sparse problem with L_0 norm is solved using the so-called greedy algorithms, such as, for example Matching Pursuit and Orthogonal Matching Pursuit.

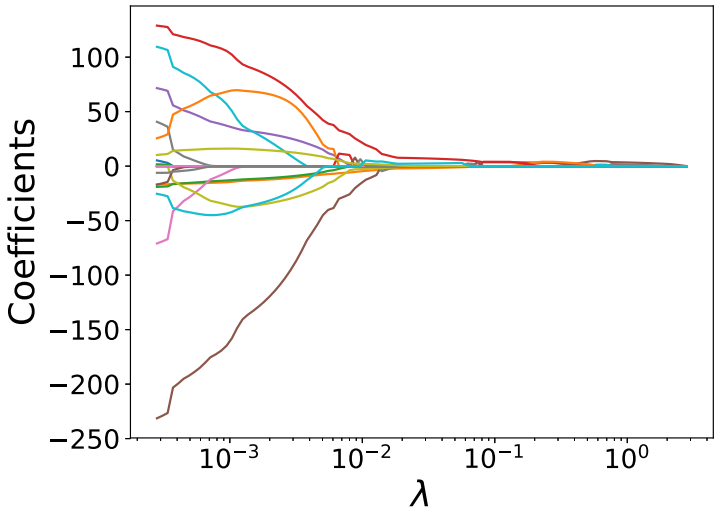


Figure 7.5 Values of the coefficients of each solution of LASSO computed with different values of λ. Each color represents a coefficient and its evolution as λ changes. This graph defines the LASSO path.

The solution of the problem, i.e. the coefficients $\boldsymbol{\alpha}$, depends on the selected value of λ. The dependence of the coefficients with the regularization parameter λ is the LASSO path $\boldsymbol{\alpha}(\lambda)$. The LASSO solution is piecewise linear, i.e. it is a function whose graph is composed of straight-line sections, with respect to λ (see Friedman, Hastie, and Tibshirani [2001b] and Tibshirani et al. [2011] for further details). An example of LASSO path is shown in Fig. 7.5.

Homotopy methods compute the full LASSO path varying the regularization parameter λ from large values to small ones. In a few cases the path can be shown to be piecewise linear as λ changes. These methods follow the piecewise-linear path by computing the direction of the current linear segment and the points where the direction changes to find the next value of λ.

The **least-angle regression** (LARS) [Efron et al. 2004] is also based in the computation of the LASSO path. It solves the problem for all values $\lambda \in [0, \infty)$ and, therefore, it is useful when the solution is desired for various values of λ. The LARS method fully characterises the trade-off between goodness-of-fit and sparsity in the LASSO solution (this is controlled by λ), and hence yields interesting statistical insights into the problem. However, when the problem is large the solution path tends to be very large and consequently the computational cost is large.

In this work we propose a decomposition based on the SB algorithm [Goldstein and Osher 2009], that we call **SB-LASSO**. As in the case of the ROF model, we introduce a new independent unknown vector $\mathbf{d}$ to split the minimization with respect to the L_1-norm. Applying this splitting the problem reads as follows

$$(\boldsymbol{\alpha}, \mathbf{d}) = \underset{\boldsymbol{\alpha},\mathbf{d}}{\operatorname{argmin}} \left\{ ||\mathbf{D}\boldsymbol{\alpha} - \mathbf{f}||_2^2 + \lambda||\mathbf{d}||_1 + \mu||\boldsymbol{\alpha} - \mathbf{d}||_2^2 \right\} . \tag{7.54}$$

By iteratively minimizing with respect to $\boldsymbol{\alpha}$ and $\mathbf{d}$ separately, the SB iterative procedure reads:

$$\boldsymbol{\alpha}^{k+1} = \underset{\boldsymbol{\alpha}^k}{\operatorname{argmin}} \left\{ \frac{\lambda}{2}||\mathbf{f} - \mathbf{D}\boldsymbol{\alpha}^k||_2^2 + \frac{\mu}{2}||\mathbf{b}^k + \boldsymbol{\alpha}^k - \mathbf{d}^k||_2^2 \right\} \tag{7.55}$$

$$\mathbf{d}^{k+1} = \underset{\mathbf{d}^k}{\operatorname{argmin}} \left\{ \|\mathbf{d}^k| + \frac{\mu}{2}||\mathbf{b}^k + \boldsymbol{\alpha}^{k+1} - \mathbf{d}^k||_2^2 \right\} , \tag{7.56}$$

$$\mathbf{b}^{k+1} = \mathbf{b}^k + \boldsymbol{\alpha}^{k+1} - \mathbf{d}^{k+1} , \tag{7.57}$$

starting with $\mathbf{b}^0 = 0$. The role of the auxiliary vector $\mathbf{b}$ is to enforce the unknowns $\mathbf{d}$ and $\boldsymbol{\alpha}$ be equal when convergence is reached. The iteration process uses a small positive fixed value of λ and it runs over a scale-space that reconstructs the signal as a linear combination of few elements of the dictionary. Again, since the two parts are decoupled, they can be solved independently. The first part with a classical Gauss-Seidel step and the second using the shinkrage function Eq. (7.39).

To illustrate the behaviour of the LASSO algorithm we perform next a very simple example. Using a 20th grade polynomial we interpolate a set of 20 points inside the interval $x \in [0, 1]$. The points are obtained from the addition of the values of the polynomial $p(x) = -2x^8 + 5x^5 + x^3$ at x and we use random noise of variance $\sigma^2 = 9$. To do the interpolation we use the classical least-squares method and SB-LASSO with $\lambda = 0.5$. The results are presented in Fig. 7.6. The least-squares solution is more oscillatory than the one obtained using SB-LASSO and develops Gibbs phenomenon (see upper panel in Fig. 7.6). Moreover, as LASSO induces sparsity, only three coefficients have non-zero value.

7.4.2 Dictionary-Learning Problem

Up to this point we have considered that the dictionary $\mathbf{D}$ is fixed and we only have to solve the problem of representation. Traditionally, predefined dictionaries based on various types of wavelets, curvelets, etc, have been used. However, the signal reconstruction can be dramatically improved by learning the dictionary instead of using a predefined one [Elad and Aharon 2006].

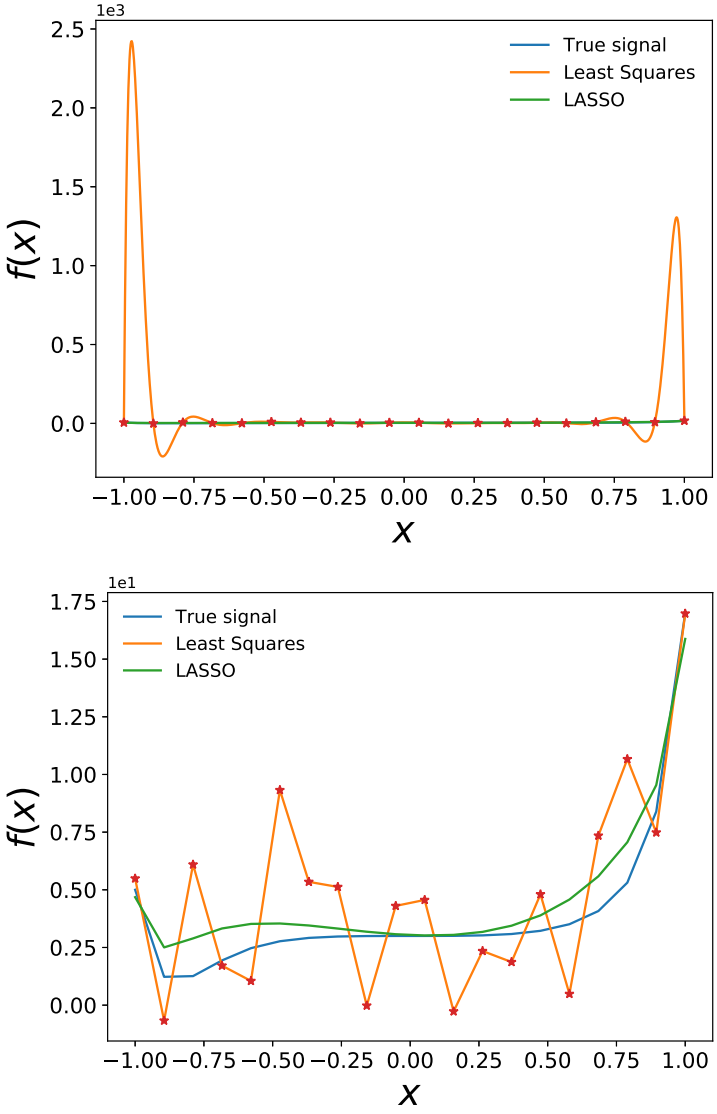


Figure 7.6 Comparison of the interpolation solution of the polynomial $p(x) = -2x^8 + 5x^5 + x^3$ using least squares and LASSO. The top panel shows the solutions using 1000 points, while the bottom panel shows the solution for the 20 points used to make the interpolation.

Classical dictionary-learning techniques [Olshausen and Field 1997, Aharon, Elad, and Bruckstein 2006] start considering a set of training signals $\mathbf{X} \in \mathbb{R}^{m \times n}$ and define a cost function which should be optimized,

$$f_n = \frac{1}{n} \sum_{i=1}^{n} l(\mathbf{x}_i, \mathbf{D}), \tag{7.58}$$

where $\mathbf{x}_i$ are the columns of the training set matrix $\mathbf{X}$, $\mathbf{D}$ is the dictionary, and l represent some function whose value is small when $\mathbf{D}$ is a "good" representation of the training signal $\mathbf{x}$. The training signals are commonly patches obtained from splitting an image or signal. Therefore, matrix $\mathbf{X}$ is formed by m patches of length n. In most common problems, the number of training patches m is large compared with the length of each patch, $n \ll m$. In general, the number of atoms in the dictionary is lower than the number of patches used for training, $p \ll m$, because each signal only uses a few elements in $\mathbf{D}$ for the representation.

When the cost function is the LASSO, the learning problem reads,

$$\boldsymbol{\alpha} = \underset{\boldsymbol{\alpha}, \mathbf{D}}{\operatorname{argmin}} \left\{ \frac{1}{n} \sum_{i=1}^{m} ||\mathbf{D}\boldsymbol{\alpha}_i - \mathbf{u}_i||_2^2 + \lambda ||\boldsymbol{\alpha}_i||_1 \right\}. \tag{7.59}$$

$\mathbf{D}$ is constrained by

$$\mathcal{C} = \left\{ \mathbf{D} \in \mathbb{R}^{n \times p} \text{ subject to } (\mathbf{d}_\mathrm{i}^\mathrm{T} \mathbf{d}_\mathrm{i}) \leq 1 \ \ \forall \mathrm{i} = 1, \ldots, \mathrm{p} \right\}. \tag{7.60}$$

The whole problem (7.59) is not jointly convex, but convex with respect to either of the two variables, $\boldsymbol{\alpha}, \mathbf{D}$, keeping the other one fixed. Classical optimisation, denominated batch algorithms, alternates between $\mathbf{D}$ and $\boldsymbol{\alpha}$ to solve the whole problem, solving for one variable while the other remains fixed. These algorithms produce good results, but they can be very slow. The computational cost is dominated by the computation of $\boldsymbol{\alpha}$. Therefore, it is possible to use a second-order optimisation to estimate $\mathbf{D}$ with good accuracy.

Bousquet and Bottou [2008] proposed an alternative formulation based in the computation of the expected cost,

$$\underset{\mathbf{D} \in \mathcal{C}}{\operatorname{argmin}} \left\{ f(\mathbf{D}) = \mathbb{E}_x[l(\mathbf{x}_i, \mathbf{D})] \approx \lim_{n \to +\infty} \frac{1}{n} \sum_{i=1}^{m} l(\mathbf{x}_i, \mathbf{D}) \right\}. \tag{7.61}$$

To solve this, the authors use stochastic gradient algorithms. This kind of methods are called on-line learning algorithms. It can be proven that, in certain settings, they can handle potentially infinite or dynamic datasets and can be dramatically faster than batch algorithms.

Classical methods for solving the dictionary-learning problem use classical projected first-order stochastic gradient descent [Aharon, Elad, and Bruckstein

2006] which is based on a sequence of updates of $\mathbf{D}$,

$$\mathbf{D}_t = \prod_{\mathcal{C}} \left[\mathbf{D}_{t-1} - \frac{\rho}{t} \nabla_{\mathbf{D}} l(\mathbf{x}_t, \mathbf{D}_{t-1}) \right], \tag{7.62}$$

where ρ is the gradient step or learning rate, $\prod_{\mathcal{C}}$ is the orthogonal projector on $\mathcal{C}$, and $\mathbf{x}$ is the training set. If the gradient step is selected appropriately, this algorithm can be competitive with batch methods for large training sets.

Mairal et al. [2009] proposed a similar approach, based on stochastic approximations, processing one sample at a time and taking advantage of the problem structure to efficiently solve it. For each element in the training set, the algorithm alternates a classical sparse coding step to solve $\boldsymbol{\alpha}$ using a dictionary $\mathbf{D}_t$ obtained in the previous iteration, with a dictionary update step, where the new dictionary is calculated with recently calculated values of $\boldsymbol{\alpha}$,

$$\boldsymbol{\alpha}^{k+1} = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \left\{ \frac{1}{n} \sum_{i=1}^{m} ||\mathbf{D}^k \boldsymbol{\alpha}_i - \mathbf{u}_i||_2^2 + \lambda ||\boldsymbol{\alpha}_i||_1 \right\} \tag{7.63}$$

$$\mathbf{D}^{k+1} = \underset{\mathbf{D}}{\operatorname{argmin}} \left\{ \frac{1}{n} \sum_{i=1}^{m} ||\mathbf{D} \boldsymbol{\alpha}_i^{k+1} - \mathbf{u}_i||_2^2 + \lambda ||\boldsymbol{\alpha}_i||_1 \right\} \tag{7.64}$$

The main advantage of this implementation is that it is parameter-free and does not require any learning rate.

Chapter 8

Total-Variation methods for gravitational-wave denoising: The Gaussian case

The results of this chapter have been originally published in:
Torres, A., Marquina, A., Font, J. A., & Ibáñez, J. M., "Total-variation-based methods for gravitational wave denoising", Physical Review D, **90**, 084029 (2014).

In this chapter we apply two of the TV-denoising techniques discussed in the previous chapter in the context of GW denoising. Arguably, the main advantage of these techniques is that no a priori information about the astrophysical source or the signal morphology is required to perform the denoising. As we illustrate below, this main feature allows us to obtain satisfactory results for two different catalogs of gravitational waveforms comprising signals with very different structure. The first catalog corresponds the "burst" signals from the collapse of rotating massive star [Dimmelmeier et al. 2008]. The second one is formed by waveforms obtained by the simulation of BBH systems [Mroué et al. 2013]. Both catalogs are described in detail in Appendix B. Our aim is to find optimal values of the parameters of the ROF model that can assure a proper noise removal. In order to do so we modify the ROF problem to take into consideration the sensitivity curve from the Advanced LIGO detector. We

restrict the formulation of the problem to 1D, since the available GW catalogs we employ are one-dimensional, even though the algorithm can be easily extended to higher dimensions. We emphasize that there are no restrictions about the data, and in this way the denoising can be performed in both the time or the frequency domain.

8.1 Regularization parameter estimation

Denoising results are strongly dependent on the value of the regularisation parameter λ. As discussed in the previous chapter, the optimal value of λ that produces the best results cannot be set up a priori, and must be defined empirically. In the following, we perform an heuristic search for the optimal value of the regularisation parameter to denoise a signal from the core collapse catalog of Dimmelmeier et al. [2008]. Since the procedure is the same regardless of the catalog, for the case of the BBH catalog we only give at the end of this section the results of the corresponding search. The goal is to find a small span of values of λ that provide a recovered (denoised) signal for all test signals under different SNR conditions. We shall apply the rROF algorithm for the time domain and the SB method for the frequency domain.

Standard algorithms assume stationary additive white Gaussian noise. However, as the noise of actual interferometric detectors is non-white, we have to adapt the denoising algorithms to take this fact into account. The weight distribution of the noise, w, is calculated as the inverse of the noise PSD according to the sensitivity curve of Advanced LIGO [Sathyaprakash and Schutz 2009], which is shown in Fig. 8.1.

On the one hand, in the time domain we do not make any assumption about the noise, i.e., we use the rROF algorithm and we filter out the obtained result below the lower cut-off frequency of the sensitivity curve, according to the weight distribution. On the other hand, in the frequency domain we proceed as follows: Given the observed signal g,

$$f = x + n\,, \tag{8.1}$$

where x is the signal from the catalog and n is the noise, we compute the Fourier transform of the mirror extension of f, f_{ext}. The idea of the mirror extension is to expand the data at the boundaries, attenuating the border effect introduced by the Fourier transform. Then, we solve the following TV-denoising model

$$v_{\text{opt}} = \underset{v}{\operatorname{argmin}} \int |\nabla v| + \frac{\lambda}{2}||v - f_{\text{ext}}||_w^2\ , \tag{8.2}$$

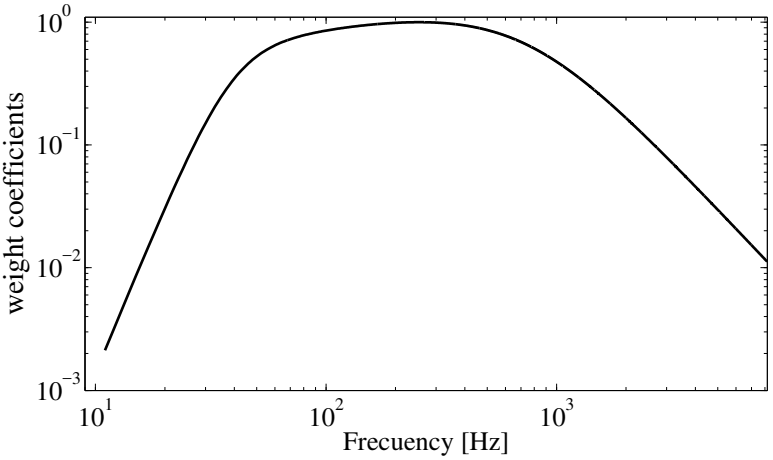


Figure 8.1 Frequency distribution of noise weight coefficients obtained from the Advanced LIGO sensitivity curve.

where $||v||_w^2 := \int w \cdot |v|^2$ is the weighted (by w) L_2-norm of v in the frequency domain, by using the SB method for complex functions of real variable. Finally we compute the inverse Fourier transform of v_{opt}, and after restricting its values to the appropriate time domain, we obtain the denoised signal u_{opt}. We remark that due to its appropriate border treatment and computational efficiency, we use the matrix formulation of the SB method developed by Micchelli, Shen, and Xu [2011] when we address intensive real-time calculation.

All signals of the core collapse catalog have been resampled to the LIGO/Virgo sampling rate of $16,384$ Hz and zero padded to be of equal length. For the SB algorithm in the frequency domain, signals have been Hanning-windowed and mirror-extended to avoid border effects in the Fourier Transform. We add non-white Gaussian noise to the signals generated as explained in Appendix A. To ensure the best conditions for the convergence of the algorithms and to avoid round-off errors, we also scale the amplitude of the test signals g of both catalogs to vary between -1 and 1. The values of λ we discuss in this section are hence determined by this normalisation. As we use the same noise frame in all the experiments for comparison reasons, the amplitude of all signals is scaled to the same values and differences are only given by the SNR which for a wave strain h is defined as [Jaranowski and Kròlak 2005]

$$\text{SNR} = \sqrt{4\Delta t^2 \Delta f \sum_{k=1}^{N_f} \frac{|\tilde{h}(f_k)|^2}{S(f_k)}} \,, \tag{8.3}$$

Table 8.1 Values of the fidelity term and of the optimal value of λ for several time windows for the core collapse signals A, B, and C (see Appendix B for details). The values are obtained after applying the SB algorithm in the frequency domain.

Δ_t(ms)	$\|\|g-x\|\|^2_{L_2}$	$\lambda_{\rm opt}$		
		A	B	C
1000	0.059	0.29	0.28	0.22
500	0.059	0.45	0.49	0.36
250	0.059	0.87	0.98	1.06
125	0.055	1.11	0.98	0.61
62.5	0.055	1.45	1.66	2.80
31.25	0.055	2.60	3.05	3.94

where $\tilde{h}$ indicates the discrete Fourier transform (see Appendix A) of signal h and $S(f_k)$ is the power spectral density.

The optimal value of the regularisation parameter, $\lambda_{\rm opt}$, is defined to be the one which gives the best results according to a suitable metric function applied to the denoised signal and the original one, measuring the quality of the recovered signal. In our case, we choose the peak signal-to-noise ratio, PSNR, based on the fidelity term, Eq. (7.10),

$$\text{PSNR(dB)} = 10\log_{10}\left(\frac{N}{\text{MSE}}\right), \tag{8.4}$$

$$\text{MSE} = \frac{||x-u||^2_{L_2}}{N}, \tag{8.5}$$

where x is the original signal from the catalog, u is the processed signal after applying the algorithms, and N is the number of samples.

First of all, we have to find the appropriate time window to perform this comparison. Since bursts have short duration (a few ms), if the time window is long the signal to compare with is going to be composed of mainly zeros, while, if it is short, some parts of the signal can be lost. To study this dependence, we seek for the value of λ in core collapse signals for which the fidelity term of the denoised signal matches the fidelity term of the original signal, that is, we seek for $\lambda_{\rm opt}$ subject to $||g-x||^2_{L_2} \approx ||g-u||^2_{L_2}$.

Results of this study for the three representative signals from the core collapse catalog are shown in Table 8.1. The SB method has been explicitly developed for discrete signals and it is well known that there exists a correlation between the number of samples and the value of $\lambda_{\rm opt}$. Therefore, different (time or frequency) scales of the same signal cannot be recovered using the same value of λ. Indeed, we find that the values of $\lambda_{\rm opt}$ we obtain show certain dependence on the number of samples, which is roughly equal to $\frac{1}{\sqrt{2}}$. Both, differences in

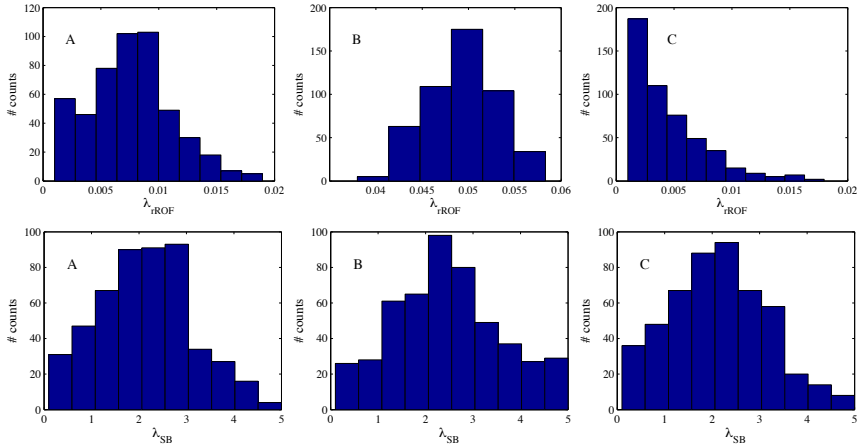


Figure 8.2 Histograms of the values of λ_{opt} for 500 noise generations for the three representative core collapse signals. The upper panels show the values of λ_{opt} for the rROF method in the time domain while the corresponding results for the SB method applied in the frequency domain are shown in the bottom panel. A SNR=15 is assumed.

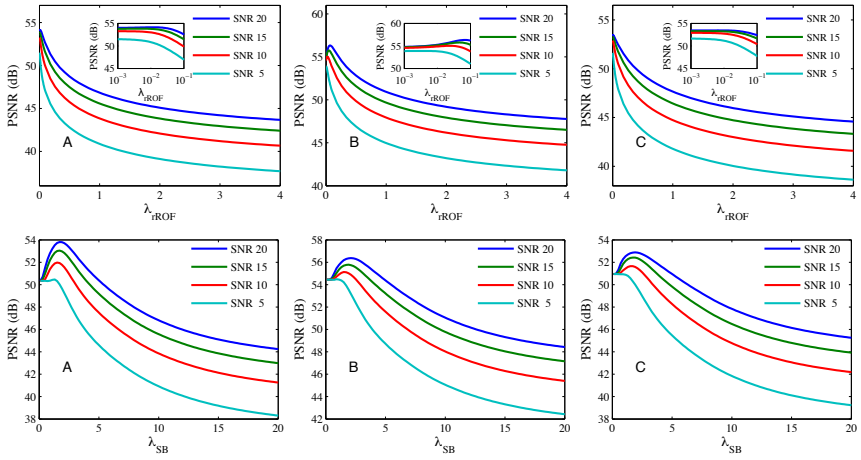


Figure 8.3 Dependence of the PSNR for different signal-to-noise ratios for the three representative core collapse signals. The upper panels show the values of PSNR for the rROF method in the time domain while the corresponding results for the SB method applied in the frequency domain are shown in the bottom panel. The insets in the top panels magnify the areas where the variations in the curves are largest.

$||g - x||^2_{L_2}$ and deviations from the previous ratio are due to the weight that the significant features of the GW signals have relative to the number of zeros. The rROF algorithm, on the other hand, reduces the staircase effect associated to the shrinkage operator in SB. The results of Table 8.1 for the SB method allow us to adjust the time window of the rROF method. From this comparison we choose a time window of 62.5 ms for the core collapse signals as it yields a complete representation of the waveforms without losing any significant feature.

Once we have selected the time window, we must find the appropriate value of λ based on the PSNR value. First we seek the optimal value of the regularisation parameter for several realisations of noise. The corresponding histograms with the optimal value of λ for both algorithms, SB and rROF, are shown in Fig. 8.2 for the three representative burst signals, assuming a SNR value of 15. We note that both distributions have the expected Gaussian shape. The variance of each distribution gives us a window of variability around the mean value of λ to estimate the optimal one. The half-Gaussian in signal C for rROF is apparent and the whole Gaussian can be seen by log-scaling the range of λ.

Having found the mean value of λ for a given signal through noise variations, we extend the analysis to consider different signal-to-noise ratios. We re-scale the amplitude of the three signals to fix the value of the SNR. The results are displayed in Fig. 8.3 for both algorithms. This figure shows that for all SNR values considered, the PSNR values peak around the optimal value of the regularisation parameter. The span of values of λ to ensure a proper denoising is $1.5 - 3$ for the SB method and $0.001 - 0.015$ for the rROF algorithm. For very noisy signals (low SNR) the recovered ones are very oscillatory and cannot be distinguished from noise. In this case, it might be possible to apply other data analysis techniques to improve the results. From our analysis we observe that in some cases there is an interval of optimal values of λ giving very similar regularised result, probably due to the lack of resolution in the data.

Finally, we analyze the optimal value for *all* signals of the core collapse catalog assuming SNR=15. As shown in Fig. 8.4, the optimal value of λ is quite different across the entire catalog, particularly for the rROF algorithm (left panel). The values of $\lambda_{\rm SB}$ span the interval $1.5 - 2$ and show a Gaussian profile, while for the case of the rROF algorithm they span the interval $0.001 - 0.05$ and do not show an obvious trend. We note that although the range of values of $\lambda_{\rm opt}$ is large, it is still possible to perform the denoising procedure with acceptable results, by choosing the mean value for all the catalog signals and then tuning it up to find the value of λ that provides the best results.

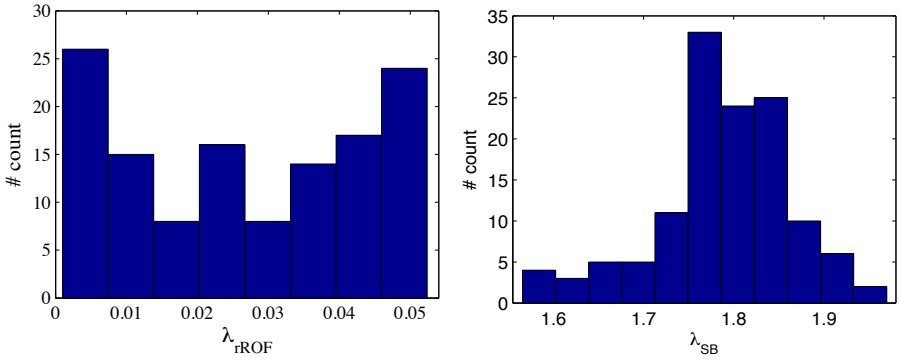


Figure 8.4 Histograms of $\lambda_{\rm opt}$ for all signals of the core collapse catalog with SNR = 15. The left panel shows the span of values of $\lambda_{\rm opt}$ for the rROF method while the right panel displays the corresponding values for the SB method.

As mentioned at the beginning of this section we have also performed the same analysis with the entire catalog of BBH signals [Mroué et al. 2013]. In this case we choose a window that contains the last $12-14$ cycles before the merger ($\Delta t \sim 62.5$ ms) for illustrative reasons. The results obtained are similar to those reported for the core collapse catalog, the optimal interval for the rROF algorithm being $0.001-0.01$ and $1-3$ for the case of the SB algorithm.

We can conclude that the appropriate values of λ for both algorithms are restricted to a sufficiently small interval, which remains approximately constant for all signals of the catalogs and for different SNR. We stress that the concrete values of $\lambda_{\rm opt}$ reported here are mainly used as a rough guide to apply the denoising procedures. If the properties of the signals change, such as the sampling frequency, the number of samples, or the noise distribution, the values of $\lambda_{\rm opt}$ can also change, and it would become necessary to recompute them. Nevertheless, as we will show next, it is indeed possible to obtain acceptable results for all signals using a generic value of λ within the intervals discussed here which can then be fine-tuned to improve the final outcome.

8.2 Results

8.2.1 Signal Denoising

We start applying both TV denoising methods to signals from both catalogs in a high SNR scenario, namely SNR=20. Our aim is to show how the two algorithms perform the denoising irrespective of the nature of the gravitational waveform considered. We assume that there is a signal in the dataset obtained from a list of

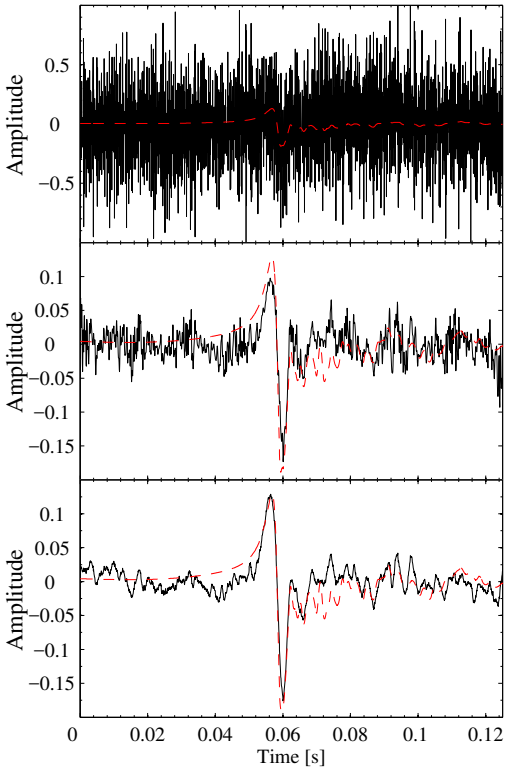


Figure 8.5 Denoising of the core collapse waveform signal C with SNR = 20. Top panel: original signal (red dashed line) and non-white Gaussian noise (black solid line). Middle panel: Original and denoised (black solid line) signals for the SB method in the frequency domain with $\lambda = 2.0$. Bottom panel: Original and denoised signals for the rROF method in the time domain with $\lambda = 0.09$.

candidate triggers, and that all glitches have been removed. This simple situation allows us to test our techniques as a denoising tool to extract the actual signal waveform from a noisy background. We apply the proposed methods, rROF in the time domain and SB in the frequency domain, independently. The results from applying our denoising procedure to a signal from the core collapse catalog is shown in the three panels of Fig. 8.5. For the sake of illustration we focus on signal C, since the results are similar for the other two types of signals. The most salient features of this signal are the two large positive and negative peaks around $t \sim 0.06$ s associated with the hydrodynamical bounce that follows the collapse of the iron core, and the subsequent series of small amplitude oscillations associated with the pulsations of the nascent PNS. In the top panel we plot the original signal (red dashed line) embedded in additive non-white Gaussian noise

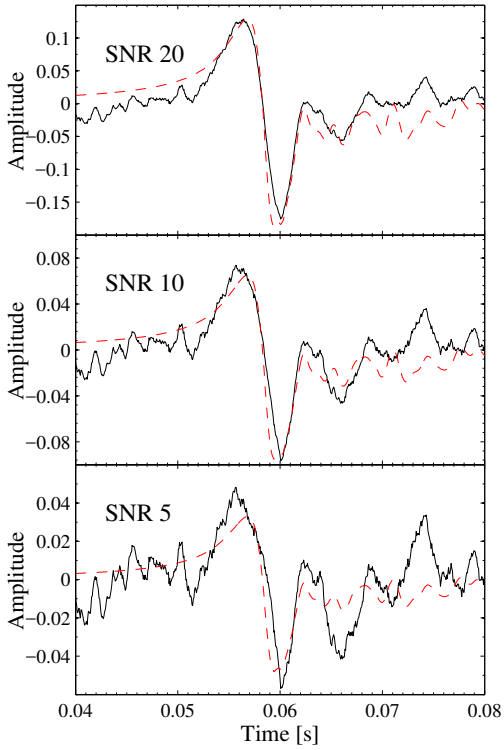


Figure 8.6 Denoising of the core collapse waveform signal C for the rROF method with $\lambda = 0.09$ for three values of the SNR, 20 (top), 10 (middle), and 5 (bottom).

(black solid line). The middle panel shows the result of the denoising procedure after applying the SB method in the frequency domain with $\lambda = 2.0$, and the bottom panel shows the corresponding result after applying the rROF algorithm in the time domain with $\lambda = 0.09$. The two large peaks are properly captured and denoised, most notably the main negative peak. This is expected due to the large amplitude of these two peaks, as TV denoising methods work best for signals with a large gradient. In turn, those parts of the signal with small gradients cannot be recovered as nicely, as seen in the damped pulsations that follow the burst and which have amplitudes much smaller than the noise. We note that both algorithms attenuate positive and negatives peaks due to noise effects. If desired, it would be possible to recover the actual amplitude of the main two peaks of the signal accurately by using a larger value of λ. However, this would introduce a more oscillatory signal in the part of the waveform with small gradients. Such oscillations are consistently more common for the SB

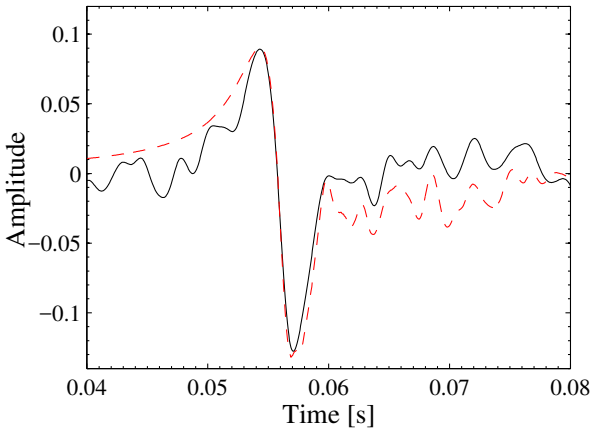


Figure 8.7 Denoising of the core collapse waveform signal C with SNR = 20 after applying both algorithms sequentially, rROF first with $\lambda = 0.05$ followed by SB with $\lambda = 8$.

method than for the rROF method, as can be seen from the middle and bottom panels of Fig. 8.5.

The effect of varying the SNR on the denoising procedure is shown in Fig. 8.6 for the same core collapse waveform C. This figure magnifies the signal around the late collapse and early post-bounce phase, i.e. $0.04\text{s} < t < 0.08\text{s}$. The three panels show, from top to bottom, the comparison between the denoised and the original signal for SNR=20, 10, and 5, respectively. Only the results for the rROF algorithm with $\lambda = 0.09$ are shown, as the results and the trend found for the SB method are similar (only more oscillatory in the small gradient part of the signal for the latter). This figure shows that, as the SNR decreases, the denoised signal recovers the original signal worse, as expected. While the amplitude of the oscillations of the denoised signal increases in the part with small gradients, it is nevertheless noticeable the correctness of the method to recover the amplitude of the largest negative peak of the signal even for SNR=5. We stress that all three signals have been denoised applying the same value of λ and recall that the study of the dependence of λ on the SNR (Section 8.1) predicts a lower value of λ_{opt} as the SNR decreases to obtain the best results. Therefore, the effect of using a value of λ greater than the optimal one is also noticeable in the middle and lower panels of Fig. 8.6.

The amount and amplitude of the oscillations of the denoised signals in small gradient regions can be somehow made less severe when both algorithms are applied sequentially. This is shown in Fig. 8.7 for the GW burst signal C with SNR=20. The denoised signal in black in this figure is the result of applying

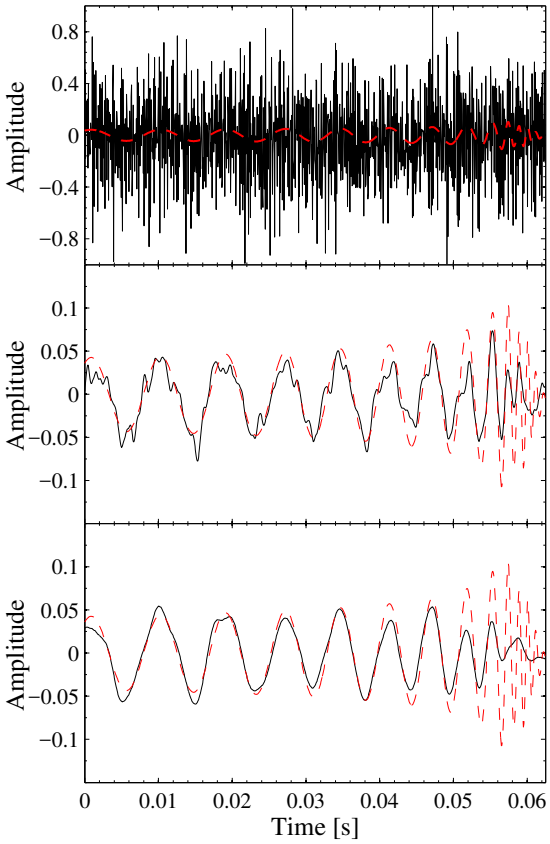


Figure 8.8 Denoising of the BBH waveform signal "0001" with SNR = 20. Top panel: original signal (red dashed line) and non-white Gaussian noise (black solid line). Middle panel: Original and denoised (black solid line) signals for the SB method in the frequency domain with $\lambda = 1.6$. Bottom panel: Original and denoised signals for the rROF method in the time domain with $\lambda = 0.0026$.

an initial denoising with the rROF algorithm in a 1 s window followed by a second step with the SB method in the frequency domain using a 125 ms time window. The values of the regularization parameters employed in this case are $\lambda = 0.05$ for the rROF algorithm and $\lambda = 8$ for the case of the SB method. We note that, in general, larger values of the regularisation parameters are required when applying both methods sequentially because regularisation is accumulative. Indeed, the output of the first step contains less noise and, therefore, the second step needs larger values of λ, i.e. less regularisation.

We turn next to apply the denoising procedure to the BBH catalog. For illustrative purposes we focus on a single signal of such catalog, signal "0001", as

this suffices to reveal the general trends. This signal corresponds to a binary of two black holes with similar mass ($\sim 10M_{\odot}$). The results are shown in Fig. 8.8. As above, the top panel shows the original signal (red dashed line) embedded in additive non-white Gaussian noise (black solid line). The middle panel shows the result of the denoising procedure after applying the SB method in the frequency domain with $\lambda = 1.6$, and the bottom panel shows the corresponding result after applying the rROF algorithm in the time domain with $\lambda = 0.0026$. A value of SNR=20 is assumed. To better visualise the results we display only the last few cycles up until the two BH holes merge.

As BBH waveforms have longer durations than bursts and the characteristics of the signal change over time, scales between the beginning and the end of the signal are significantly different. This becomes clear in Fig. 8.8 where as a result of the scale variations in frequency and amplitude during the late inspiral and merge, the signal cannot be properly denoised throughout using the same value of λ. Typically we find that using a comparatively large value of λ helps to accurately recover larger amplitudes and high frequencies than lower frequencies and amplitudes, and vice versa. The reason for this is again due to the fact that the methods we employ are gradient dependent, preserving the large gradients and removing the small ones. On the one hand, choosing a low value of λ to recover low frequency cycles makes the merger signal to be treated as high frequency noise. On the other hand, choosing a high value of λ to recover the merger part produces high oscillations in the rest of the signal. We have checked that it is nevertheless possible to obtain good results for the entire BBH waveform train using different values of λ for different intervals of the waveform.

8.2.2 Signal Detection

From the previous analysis, it becomes manifest that in a low SNR situation our denoising algorithms alone cannot remove sufficient noise to produce detectable signal, i.e. a signal clearly distinguishable from the noise background. In order to improve our results in such a situation, we can combine our techniques with the use of spectrograms. Such an approach is usually employed in GW data analysis, in a more complex pipeline, to seek for transient power peaks in the data that could correspond to actual GW wave signals, assuming that all known transients have been removed from the data [Ajith et al. 2014, Di Credico and LIGO Scientific Collaboration 2005, Anderson and Balasubramanian 1999].

First, we have to check if the information we can obtain from the spectrogram would be modified by the application of our techniques. To do this we compare

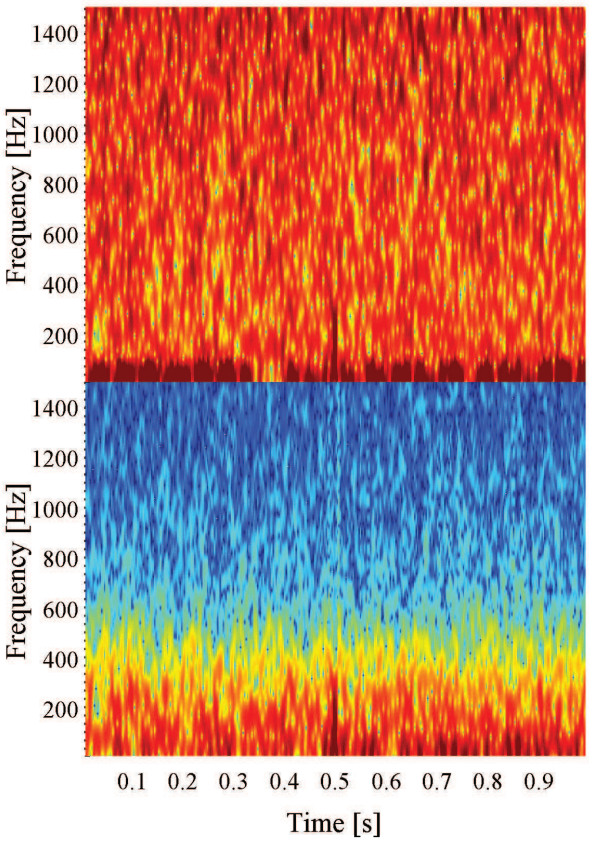


Figure 8.9 Spectrograms of the original noisy signal (top panel) and denoised signal (bottom panel) for the core collapse signal C and SNR =10. The higher values of the spectral power density are shown in red, while the lower power is represented in blue.

the spectrogram from an original noisy signal with SNR=10 with the spectrogram of the corresponding denoised signal. The results are shown in Fig. 8.9 for the core collapse signal C. The spectrogram of the original signal is shown in the top panel and the denoised spectrogram is shown in the bottom panel. Red colour represents high spectral power while lower power is represented in blue. In order to produce Fig. 8.9 we first apply the rROF algorithm in the time domain followed by the SB method in the frequency domain, both for a 1s time window, and then we calculate the spectrogram. The power peak around 0.5s corresponds to the GW signal which is clearly distinguishable in both spectrograms, and its structure remains similar after the denoising procedure.

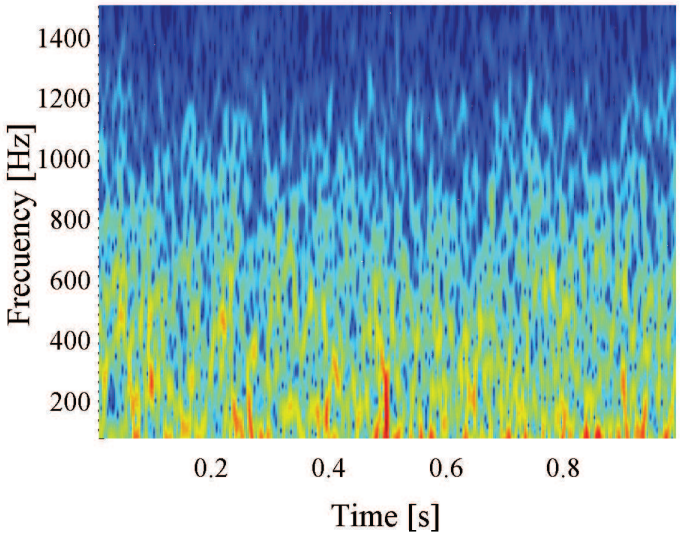


Figure 8.10 Spectrogram of the core collapse signal C for SNR=5. The excess power around 0.5 s is supposed to be produced by the GW signal.

An example of the application of the spectrogram together with our algorithms in a low SNR situation where the spectrogram alone would not reveal any high power peak, is shown in Fig. 8.10. This figure displays the denoised spectrogram of the same core collapse signal C originally embedded in non-white Gaussian noise but now with SNR=5. In this case we have a dataset which contains 1s of data from the detector. The exact arrival time of the signal is unknown and is what we want to determine by computing the spectrogram. In order to find the time of arrival of the signal we integrate the power of the first 2000 Hz for each temporal channel and look for the channel that contains the maximum power. After selecting this channel we perform the denoising procedure only in this channel, using first the SB method as a filter and then applying the rROF algorithm in order to obtain the signal waveform. Fig. 8.10 shows that the power peak of the signal (red colour) is distinguishable from the noisy background (green-blue colours). These findings give us confidence to use our algorithms as a denoising tool jointly with other data analysis techniques. However, to prove if the denoising techniques could improve the detection potential of the current LIGO-Virgo community algorithms, we need to perform an end-to-end pipeline comparison.

8.3 Discussion

The methods we use are based on L_1 norm minimisation and have been originally developed and fully tested in the context of image processing where they have been shown to be the best approach to solve the so-called Rudin-Osher-Fatemi denoising model. We have applied these algorithms to two different types of numerically-simulated gravitational wave signals, namely bursts produced from the core collapse of rotating stars and chirp waveforms from binary black hole mergers. The algorithms have been applied in both the time and the frequency domain. Both of our methods, SB and rROF, reduce the variation of the signal, assuming that due to its randomness the larger variations are due to the noise. We have performed an heuristic search to find the set of values best suited for denoising gravitational wave signals and have applied the methods to detect signals in a low signal-to-noise ratio scenario without any a priori information on the waveform. In particular, the rROF algorithm in the time domain has led to satisfactory results without any assumption about the noise distribution. On the other hand, in order to apply the SB method in the frequency domain, we have selected a particular weight distribution. This distribution can be chosen freely so as to adjust it to the specific spectral characteristics of the noise or of the detector sensitivity curve. Overall, we conclude that the techniques we have presented in this paper may be used along with other common techniques in gravitational wave data analysis to help increase the chances of detection. Likewise, these methods should also be useful to improve the results of other data analysis approaches such as Bayesian inference or matched filtering when used as a noise removal initial step that might induce more accurate results for the aforementioned traditional methods.

Chapter 9

Total-Variation methods for gravitational-wave denoising: LIGO data

In the current chapter we take a further step in the assessment of TV-methods for GW astronomy, using actual noise from the detectors instead of the idealized non-white Gaussian noise employed in previous chapters. More precisely, here we will inject numerically generated signals into the data collected by the initial LIGO detector during its 6th scientific run (S6) that extended from July 2009 to October 2010. As in the last chapter, we use two catalogs to test the methods, the first one from core-collapse simulations [Dimmelmeier et al. 2008] and the second one from BBH mergers [Mroué et al. 2013] (see Appendix B for a complete description).

9.1 Iterative rROF

In the results discussed in this chapter we will use an iterative procedure to solve denoising problems based on the so-called regularized ROF algorithm. The rROF method runs a new scale space from finer to coarser scales, with a termination criterion given by a discrepancy principle. Usually, this criterion is enforced by matching the square of the L_2-norm of the residual with the variance of the noise, if the latter is known. Since this is not the case for GW data, our iterative procedure is terminated as soon as the reconstructed signal starts to loose amplitude around local extrema.

Roughly speaking, we first choose the regularization parameter λ equal to a constant value λ_0, which is larger than the optimal value needed to obtain a denoised signal by application of the rROF algorithm. The value of λ is kept fixed through all the iterations. Next, we compute u_1 by solving

$$u_1 = \underset{u}{\operatorname{argmin}} \left\{ \mathrm{TV}(u) + \frac{\lambda_0}{2} ||u - f||^2_{\mathrm{L}_2} \right\} , \tag{9.1}$$

$$f = u_1 + v_1 , \tag{9.2}$$

where v_1 is the residual. Then, we apply again the rROF algorithm using the same λ_0 and taking as input signal u_1 to obtain u_2. We thus have

$$u_1 = u_2 + v_2 . \tag{9.3}$$

Applying this procedure for an arbitrary number of times n we obtain a sequence of signals u_n for $n = 1, \cdots$ such that

$$u_{n-1} = u_n + v_n , \tag{9.4}$$

$$f = u_n + \sum_{i=1}^{n} v_i . \tag{9.5}$$

The iteration stops when some discrepancy principle is satisfied, i.e. when the square of the L_2-norm of the residual matches the variance of the noise. In practice, however, the variance of the noise is not available and we have to resort to some other termination criterion. Therefore, the iterative procedure is stopped at some n_0 which is selected to make it coincide with the appearance of the denoised signal before its local extrema start loosing total variation. We regard signal u_{n_0} as the denoised signal. Note that the scale space defined by this iteration is different from the one observed when the parameter λ runs from small to larger values, as it is done in the Bregman refinement iterative algorithm (see Bregman [1967] and Osher et al. [2005]). The advantage of using the iterative procedure proposed here is that local extrema and edges are significantly better preserved when the scale space is run in the opposite direction (towards smaller values of λ).

9.2 Algorithm pipeline and data conditioning

The goal of this work is to test the rROF algorithm with the real conditions of a GW wave detector. We use data from the initial LIGO detector, collected during the 6th scientific run (S6) that spanned from July 2009 to October 2010

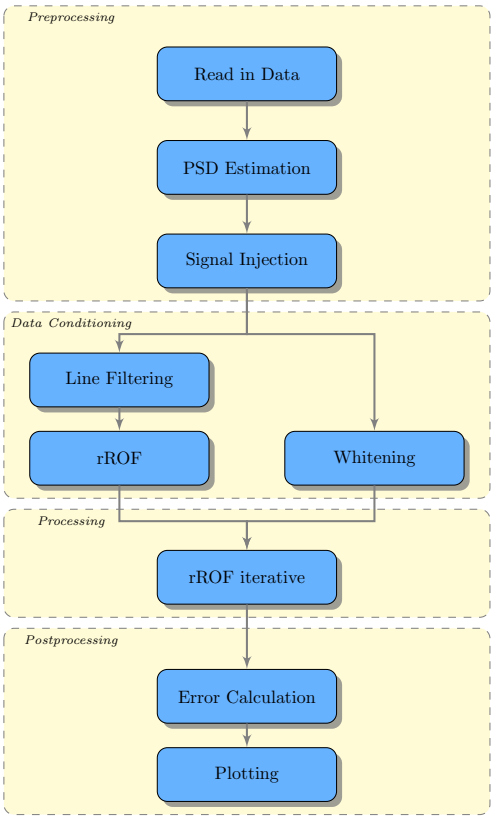


Figure 9.1 Algorithm workflow.

and which is publicly available at *LIGO Open Science Center*[1]. We test the method after injecting numerical relativity catalog signals into ten different data segments obtained randomly from 24h of data. The workflow is shown in Figure 9.1.

After reading the data at a given GPS time, we estimate the power spectral density (PSD) using 1024 s before the GPS time of the signal injection. Once the PSD has been estimated we use it to calculate the SNR, defined in Eq. (8.3). The amplitude of the signals is modified to reach the desired SNR given at the injection step.

In Chapter 8 we showed that the rROF algorithm leads to satisfactory results when the signal is embedded in Gaussian noise, at least in the case of core

[1]https://losc.ligo.org/about/

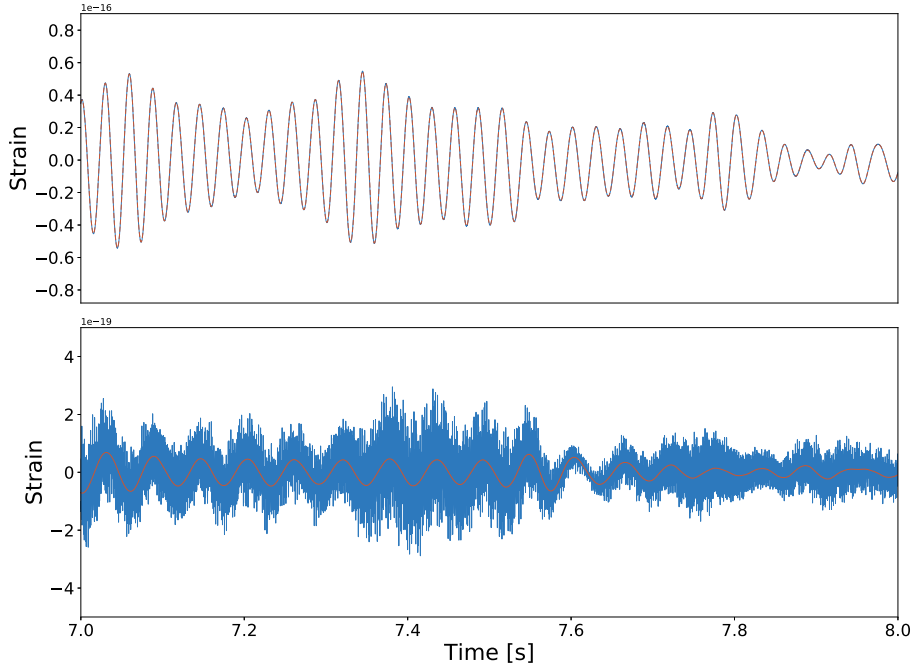


Figure 9.2 Result of applying the rROF algorithm as a low-frequency filter. Upper panel: Original signal (blue) and resulting signal after applying rROF (red) with a very low value of λ. Lower panel: Residual of the first iteration (blue) and resulting signal (red) from rROF.

collapse signals. As it is well known, the noise of GW detectors is non-Gaussian and non-stationary. For example, there exist well-known, modelled sources of narrow-band noise, such as the electric power (at 60 Hz and higher harmonics for the (American) LIGO detectors), mirror suspension resonances or calibration lines (see e.g. Fig. 3 of Abbott et al. 2016b). For this reason the data must first be pre-conditioned to remove these lines and produce a noise spectrum as flat as possible. To do so we employ two different methods, trying to make as few assumptions as possible.

The first method we use is the simple filtering of the data. We highpass the time series above 30 Hz to remove seismic noise and, following the approach described in Abbott et al. [2016b], we also filter out all spectral lines. While there are sophisticated procedures to detect and remove instrumental lines (see Abbott et al. [2008] and references therein), for our purpose it suffices to select the frequency of those lines by direct inspection. After the filtering, some low frequencies and lines still remain. To remove them we exploit one of the features of the rROF method. Since with a very low value of λ the resulting signal will

be very regular and will only contain low frequency features, we can subtract this signal to the original data and apply the rROF algorithm with the same value of λ iteratively. It can be proved that this procedure converges, so that at each iteration we subtract less low-frequency signal. As a result, the final signal has a much flatter spectrum than the original one. Fig. 9.2 illustrates this process. The upper panel displays the noisy signal after filtering (in blue) and the result of applying rROF with $\lambda = 0.001$. Visually both signals seem identical but they are not. The residual of the two signals is shown in the lower panel of Fig. 9.2 (blue line); note the difference scale in amplitude. The result of the next iteration, which contains only the lower frequencies, is also shown (red line).

Our second method to precondition the data is the whitening procedure developed by Elena Cuoco and coworkers [Cuoco, Cella, and Guidi 2004] which is included in the Noise Analysis Package (NAP), developed by the Virgo Collaboration [Acernese et al. 2005, Acernese et al. 2007]. This procedure uses an autoregressive (AR) filter to transform the coloured non-stationary noise into a white noise (see Cuoco, Cella, and Guidi [2004] for details). First we need to obtain the 3000 coefficients of the AR filter required by the whitening procedure using 500 s of data before the signal injection. The whitening is applied in the time domain, to avoid the border problems associated with the transformations in the frequency domain.

9.3 Estimation of the regularization parameter

As already discussed in the previous chapter the denoising results are strongly dependent on the value of the regularization parameter λ. If this value is too large the fidelity term in Eq. (7.15) dominates and the denoised signal is comparable to the original noisy signal y. On the contrary, if the value of λ is very small, it is the regularisation term in Eq. (7.15) the dominant one and the amplitude of the resulting signal tends to zero. We also mention that the existence of an *optimal* value of λ can be proven theoretically. However, this unique value is not equally appropriate for all possible cases one may consider (involving differences in the noise and/or in the signals) and, in practice, must therefore be set empirically.

In this section we carry out an heuristic search with the goal of determining the interval of values of λ where satisfactory results are expected. The optimal value of the regularisation parameter, $\lambda_{\rm opt}$, is the one that gives the best results according to some suitable metric function applied to the denoised signal and to

the original one. This function is used to measure the quality of the recovered signal. In the previous chapter we chose the peak signal-to-noise ratio, PSNR, as our quality estimator. In the present chapter we assess the results of the iterative rROF algorithm using two different methods, the Improvement factor in the PeakSignal to Noise Ratio (IPSNR), and an additional estimator, the so-called structural similarity (SSIM) index [Wang et al. 2004]. We have decided to change from using the simple PSNR estimator to IPSNR because the latter brings a measure of the improvement achieved when applying the rROF algorithm.

The SSIM index deviates from the traditional measures of error, which are based on the calculation of the absolute error, because it considers the structural information, taking into account the mean and the variance of both signals to be compared. The SSIM index varies between 0 (minimum similarity) and 1 (maximum similarity) and is defined as,

$$\mathrm{SSIM(x,y)} = \frac{(2\mu_{\mathrm{x}}\mu_{\mathrm{y}} + \mathrm{c}_1)(2\sigma_{\mathrm{xy}} + \mathrm{c}_2)}{(\mu_{\mathrm{x}}^2 + \mu_{\mathrm{y}}^2 + \mathrm{c}_1)(\sigma_{\mathrm{x}}^2 + \sigma_{\mathrm{y}}^2 + \mathrm{c}_2)}, \tag{9.6}$$

where c_1 and c_2 are constants, μ_x (μ_y) is the average of x (y), σ_x^2 (σ_y^2) the variance of x (y) and σ_{xy} the covariance of x and y.

Our second quality estimator is the IPSNR, defined as,

$$\mathrm{IPSNR(x,y,s)} = 10\log_{10}\frac{||\mathrm{y}-\mathrm{s}||}{||\mathrm{x}-\mathrm{s}||}, \tag{9.7}$$

where y is the original noisy signal, x is the resulting signal after applying one of our denoising algorithms and s is the original waveform. This indicator measures the improvement in the measure of the Mean Square Error (MSE) that is achieved.

These two estimators allow us to determine the optimal value of the regularisation parameter λ in each case.

9.3.1 Template error minimization

We start by computing the IPSNR of the output signal from rROF using 3 different signals from the core collapse catalog under 4 independent SNR conditions. This signals correspond to signals labeled as 60, 68 and 96 from the core collapse catalog (see Appendix B). The tolerance of the iterative rROF algorithm is set to 10^{-3} and the number of Bregman iterations is set to only 2. The results for the two different data-conditioning procedures are shown in Fig. 9.3. All the plots in this figure show a similar behavior. When the value of λ is small, the resulting signal is very regular and the IPSRN tends to the same value that would correspond to a zero signal. In contrast, for larger values

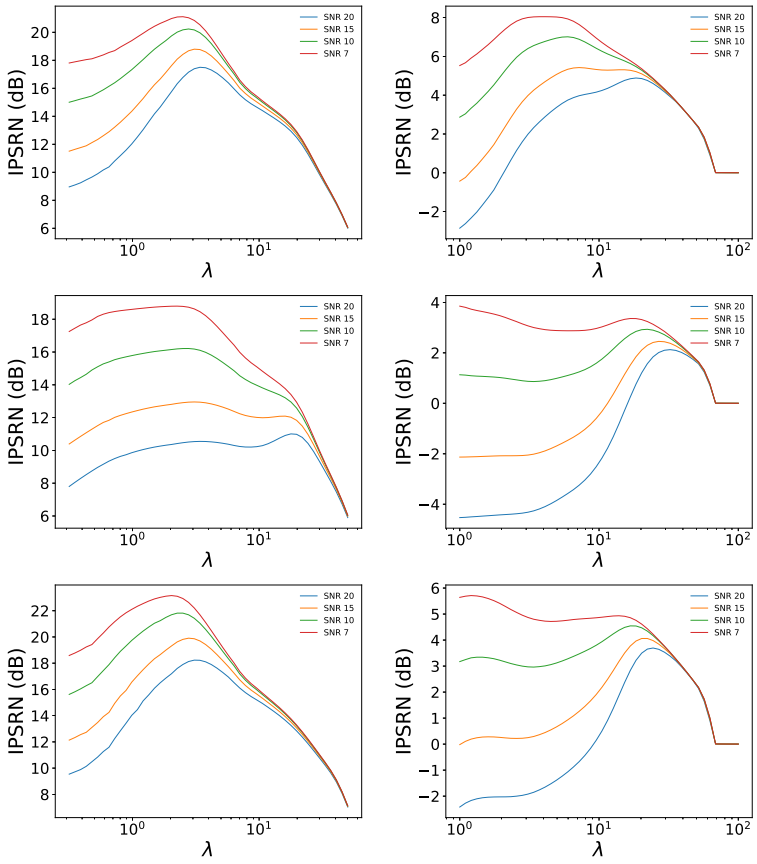


Figure 9.3 Evolution of the IPSNR with the variation of the regularisation parameter λ in 4 SNR scenarios (20,15,10,7). The left column of panels show the results for the filtering+rROF data conditioning while the right column of panels exhibit the results after the whitening. From top to bottom we show the results corresponding to signals 60, 68 and 96.

of λ the signal is increasingly more similar to the original one (before applying rROF) and the IPSNR tends to zero. We therefore consider as the optimal value of λ the value that produces the maximum value of the IPSNR. The behaviour of the SSIM index with λ is similar than for the case of IPSNR, reaching its maximum at roughly the same values.

In general, as the SNR decreases the values of IPSRN increase. To explain this we must recall that the IPSNR measures the improvement reached by the method in terms of the absolute error. For this reason, the improvement is better in low SNR conditions although the denoised signal will be more poorly recovered due to the noise conditions. The first column of Fig. 9.3 shows the

evolution of the IPSNR with λ for the three test signals preprocessed with line filtering. The value of λ_{opt} is similar for all the SNR conditions for a given signal, except for few cases as for SNR=20 in signal 68 (left column, central panel), where the maximum is reached in a slightly different point than for the other SNRs. However, in this case, the difference is not very pronounced and the denoising results do not change significantly. By comparing the three signals, one can also see that the values of λ_{opt} are in a similar range for the three signals.

The second column of Fig. 9.3 corresponds to results using the whitening procedure before applying the rROF algorithm. In this case the IPSNR values are lower, which is due to the fact that the whitening procedure performs the data conditioning much better than the line filtering. As a result, the extent of improvement that the rROF algorithm may reach is not as large. However, as we show below, the cleaning of the signal is better when conditioning the data with the whitening procedure. It is worth to point out that in the case of low SNR the plot does not show a clear maximum. In this case, when λ is small the method produces a very smooth signal for both numerical and noisy signals. Although the IPSNR is higher, the main features of the signal are lost. As for the line filtering, the range of values of the optimal regularisation parameter remains also similar for the same whitened signal under different SNR conditions. We note, however, that the values of λ_{opt} are different from those obtained with the line filtering method of data conditioning. The variation of λ seems to be restricted to an interval $\lambda \in [1, 20]$ in the case of filtering the lines and to an interval $\lambda \in [10, 18]$ for the case of whitening the data. Notwithstanding that performing the denoising with a general value of λ inside these intervals may not produce the best results, it is still worth to obtain such initial denoising to then vary the value of λ to obtain the best results.

In order to have a more complete study of the values of λ_{opt}, we have repeated the calculation of the IPSNR for other signals at varying GPS times. The mean value of λ_{opt} derived from this study is $\bar{\lambda} = 6$ for line filtering and $\bar{\lambda} = 16$ for the whitening procedure. These two values are inside the corresponding intervals indicated before.

We also perform the same analysis using three signals for the BBH catalog selected randomly, specifically these signals are labeled as BBH1 (mass ratio =1, no spin), BBH47 (mass ratio = 3, no spin), BBH94 (mass ratio = 1.5, $\text{spin}_1 = 0.5$, $\text{spin}_2 = 0$). The corresponding intervals are $\lambda \in [1, 10]$ with $\bar{\lambda} = 5$ for line filtering and $\lambda \in [2, 11]$ with $\bar{\lambda} = 5$ for the whitening approach.

One last possibility we have explored is the possibility of solving the ROF problem using multiple values of λ chosen randomly inside the intervals shown before. This allows to generalise the applicability of the method, covering the span of values of the regularisation parameter that produce good results for all signals (core collapse and BBH) under different conditions of noise. Once the two denoising procedures have been applied, we compute their mean to produce the final signal. We have explored several simple ways to join the results and the simple mean seems to be the best strategy.

9.4 Results

To test the technique with the two catalogs we inject the signals in 2 s of noise. We have used 10 different times over 24 h of LIGO S6 run. These times have been selected because the corresponding date have passed all the quality vetoes and there are no high SNR glitches near the injection times.

9.4.1 Core-collapse signals

First we test our method with the same three previous signals from the core-collapse catalog with SNR=20 using the optimal value of λ for each case. The results for both data-conditioning techniques are shown in Fig. 9.4 and Fig. 9.5, respectively. Comparing the results from both data-conditioning methods (first row in both figures), we conclude that AR-whitening performs better than line-filtering, in terms of removing noise. This was somewhat to be expected because AR-whitening takes into account many more aspects of the noise than simple filtering. In addition, since the AR-whitening is performed in the time domain, it avoids all border effects that may appear with frequency-domain transformations. However, this procedure also modifies the shape of the signal notably, producing more oscillatory signals than the line-filtering approach.

Regarding the denoised signal, core-collapse signals 60 and 96 (left and right panels in Fig. 9.4) are well recovered. Distinctive features of the signals, such as the amplitude and the phase of the main positive and negative peaks, as well as of the first secondary peak in both signals, are nicely captured. In contrast, the damping oscillations with lower amplitude are lost. This is because these peaks have lower amplitude compared with the main peak and, as a result, they are more affected by the noise. In addition, the rROF model preserves the larger gradients at the expense of the smallest ones. The denoised signal shown in the middle panel of Fig. 9.4 is also in good agreement with the processed

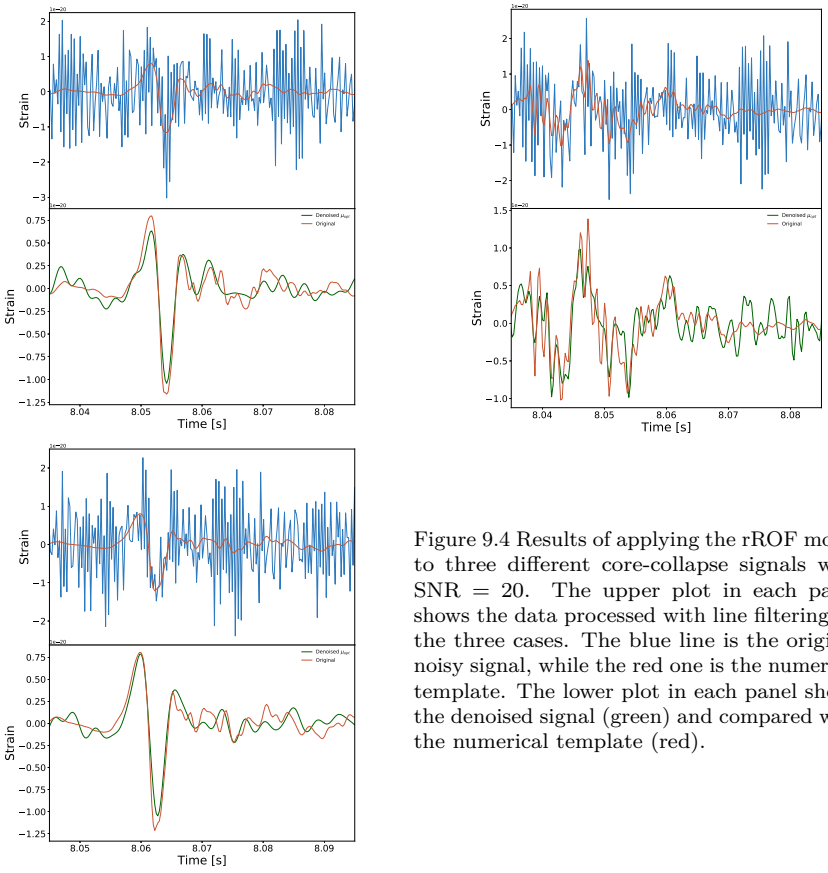


Figure 9.4 Results of applying the rROF model to three different core-collapse signals with SNR = 20. The upper plot in each panel shows the data processed with line filtering for the three cases. The blue line is the original noisy signal, while the red one is the numerical template. The lower plot in each panel shows the denoised signal (green) and compared with the numerical template (red).

template. As this signal is more oscillatory than the other two, a higher value of λ is required to properly recover the peaks. This results in the presence of more noise than in the other two cases. This fact has to be taken into account in a realistic situation where the actual signal is unknown, because the small oscillations can be lost if the strategy is to favour regular signals and significant noise removal. However, it is always possible to apply a larger value of λ and try to recover these parts of the signal.

The results of the optimal denoising for signals pre-processed with the AR-whitening procedure are shown in Fig. 9.5. After the data conditioning, the signal used as input for the rROF method is much more similar to the injected one than in the case of line-filtering (see plots in the upper row). Nevertheless, the rROF model is still able to remove part of the remaining noise and produce a more accurate result (lower row). The final gain in the IPSNR is not as large in this case than when applying the rROF algorithm to line-filtered data.

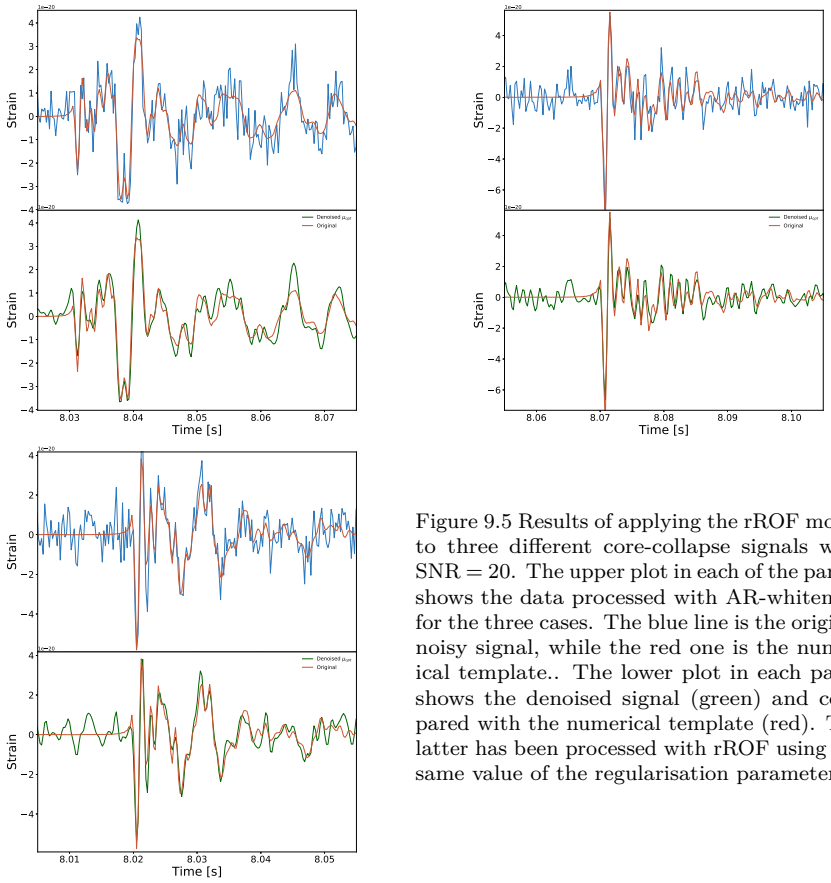


Figure 9.5 Results of applying the rROF model to three different core-collapse signals with SNR = 20. The upper plot in each of the panels shows the data processed with AR-whitening for the three cases. The blue line is the original noisy signal, while the red one is the numerical template.. The lower plot in each panel shows the denoised signal (green) and compared with the numerical template (red). The latter has been processed with rROF using the same value of the regularisation parameter λ.

We complete the analysis by performing the denoising for the same previous six cases in two new situations, (i) using the mean value of the regularisation parameter and, (ii) using multiple values of λ and computing their mean value. The resulting values of the SSIM and IPSNR estimators for the line-filtering preconditioning are shown in Table 9.1. The corresponding results for AR-whitening are reported in Table 9.2. To analyse the results of both tables, we must bear in mind that IPSNR measures the improvement in the MSE between the original noisy signal and the denoised one, while SSIM is a direct comparison between the denoised signal and the numerical template (processed with the same methods than the noisy signal). Moreover, given that IPSNR is computed in logarithmic scale, a difference of a few dBs does not imply noticeable differences in the results.

Comparing the results for both tables one can see that, even though in general the denoising is worse using the mean value, the results are not very

Table 9.1 Comparative denoising results when using λ_{opt} (optimal value), $\bar{\lambda}$ (mean value), and λ_{multi} (mean of signals with multiple values of λ). Results are for signals pre-conditioned with line filtering.

		λ_{opt}		$\bar{\lambda}$		λ_{multi}	
SNR	#	SSIM	IPSNR	SSIM	IPSNR	SSIM	IPSNR
	60	0.53	17.5	0.39	15.4	0.23	12.3
20	68	0.35	11	0.23	10.2	0.32	10.7
	96	0.42	18.2	0.31	15.9	0.22	13.2
	60	0.36	20.2	0.14	16.2	0.11	12.7
10	68	0.25	16.2	0.14	14.5	0.15	12.3
	96	0.37	21.8	0.18	16.98	0.1	13.6
20	BBH1	0.47	19.4	0.29	19.03	0.13	15.5
10	BBH1	0.28	23.5	0.11	20.8	0.1	15.6

different to the optimal one for both estimators. It seems therefore possible to use the mean value for all signals and still achieve satisfactory results. On the other hand, in the case of multiple values of λ no general trend is found, and the results do depend on the specific values of λ chosen to perform the denoising. Nevertheless, this approach seems to work better when pre-conditioning the data with AR-whitening than with line filtering, due to the shorter interval of values of λ needed in the former.

The values of the IPSNR increase when the SNR decreases for both types of pre-preprocessing methods, while the values of the SSIM index decreases. The reason is because as the SNR decreases, the rROF method is able to remove more noise (higher value of the IPSNR), but the resulting signal is more corrupted by noise (lower value of the SSIM index).

Finally, it is worth to compare the results of both branches of the pipeline. In the optimal case, both combinations of algorithms (pre-conditioning + rROF) produces similar results, according to the values of the SSIM index (first column in Tables 9.1 and 9.2). However, in a realistic case, the optimal value is not available and in such a case, employing $\bar{\lambda}$ or multiple denoising, pre-processing with AR-whitening produces better results.

Table 9.2 Comparative denoising results when using $\lambda_{\rm opt}$ (optimal value), $\bar{\lambda}$ (mean value), and $\lambda_{\rm multi}$ (mean of signals with multiple values of λ). Results are for signals pre-conditioned with AR-whitening.

		$\lambda_{\rm opt}$		$\bar{\lambda}$		$\lambda_{\rm multi}$	
SNR	#	SSIM	IPSNR	SSIM	IPSNR	SSIM	IPSNR
20	60	0.54	4.9	0.53	4.8	0.56	5.4
	68	0.34	2.1	0.31	1.1	0.31	0.8
	96	0.49	3.7	0.48	3.4	0.50	3.8
10	60	0.39	7.0	0.34	5.3	0.36	6.3
	68	0.21	2.9	0.20	2.9	0.22	3.1
	96	0.28	4.5	0.29	4.4	0.32	5.2
20	BBH1	0.77	10.3	0.42	5.1	0.46	5.9
10	BBH1	0.48	10.5	0.20	5.1	0.21	5.9

9.4.2 Binary black hole signals

We turn now to perform the same analysis to the signals from the BBH merger catalog. As we have already pointed out, these signals are much longer than core-collapse signals and show a distinctive chirp morphology, where the amplitude and the frequency increases with time up to the merger. The results are shown in Fig. 9.6 and reported in the last two rows of Table 9.1 and Table 9.2. As in the case of core-collapse signals, the values of IPSNR are higher for lower values of SNR since the algorithm is able to remove more noise. However, contrary to core-collapse signals, for BBH signals using AR-whitening as a pre-processing method works much better (in terms of SSIM index values) than just filtering the lines. This effect is clearly visible in Fig. 9.6, where the left panel corresponds to line-filtering and the right panel corresponds to AR-whitening. The denoised signal for the latter is smoother than for the former, and both phase and amplitude of the last cycles before merger are well recovered. We note that although the shape of the signal has changed due to the whitening (this is simply a numerical artefact), especially the merger and the ring-down parts, it is always possible to apply the reverse filter and recover the original signal, now with less noise.

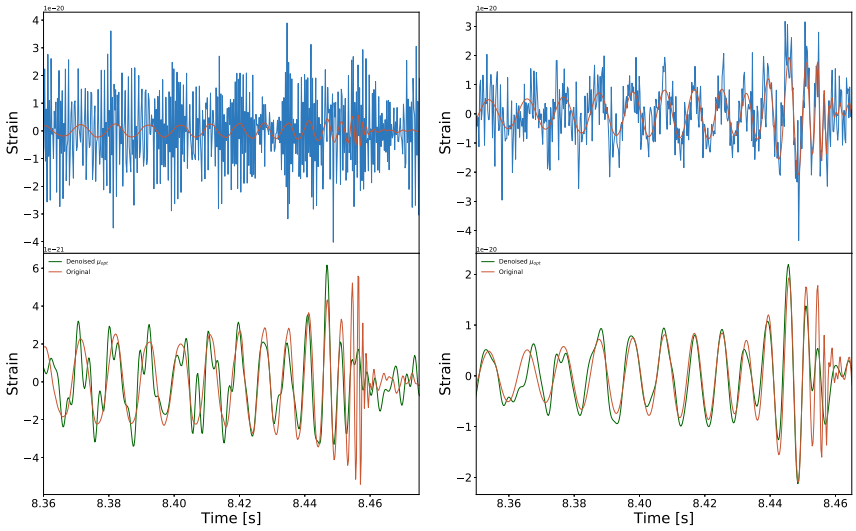


Figure 9.6 Results of applying the rROF model to signal 1 of the BBH catalog with SNR = 20. The upper row of panels shows the data pre-conditioned with line-filtering (left panel) and AR-whitening (right panel), respectively. The lower row shows the denoised signal (green) and the comparison with the numerical template (orange). The latter has been processed with rROF using the same value of the regularization parameter λ.

9.4.3 Low SNR case

In this section we test the combination of whitening and rROF in a more realistic scenario, using a lower SNR value than that considered previously. We set SNR=7 and consider that the arrival time of the signal is randomly placed inside a data segment spanning 8 s. The procedure starts with the AR-whitening of the data and then we apply rROF with the mean value of the regularisation coefficient $\bar{\lambda}$. Finally, we compute an spectrogram and the time window to plot the results is defined at 0.1 s around the time of the maximum power of the signal. Fig. 9.7 displays the results of this test for a core-collapse signal. Even with SNR = 7, the whitening pre-processing of the data reduces the noise significantly. Next, the subsequent use of the rROF method is able to reduce even more noise, and the spectrogram (shown in the middle panel) reproduces a very clear signal.

9.5 GW150914

In the last part of this chapter we test the rROF method using data from the first GW detection, GW150914 [Abbott et al. 2016b]. As the analysis available

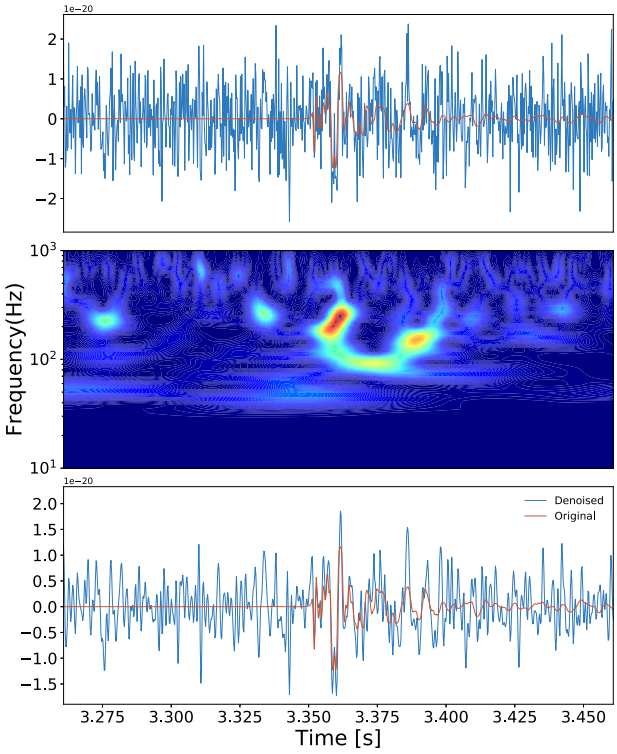


Figure 9.7 Results of applying the combination of AR-whitening and rROF to a core-collapse signal with SNR=7. The upper panel shows the signal (blue) and the original template (red) after the whitening. The middle panel shows the spectrogram, computed to find the signal. The lower panel shows the results after applying the rROF method to the signal of the upper panel with λ.

at *LIGO Open Science Center* uses filters and simple whitening, we decided to carry out the analysis with the rROF method after conditioning the data with line-filtering only. As we want to make as few assumptions about the noise as possible, we only filter the signal below 30 Hz, and remove the spectral lines as before.

The denoised waveforms for both the Hanford and Livingston advanced LIGO detectors are shown in Fig. 9.8 (red and blue curves, respectively). For both detectors the waveforms look remarkably similar, especially in the last few cycles of the signal. Comparing these denoised waveforms with the filtered waveforms reported in Abbott et al. [2016b] we notice that the latter show high-frequency features visible in some of the final cycles before the merger that do not appear in our results. By reducing the number of iterations in the rROF algorithm to about 3-5 and decreasing the value of λ, our method does not entirely smooth

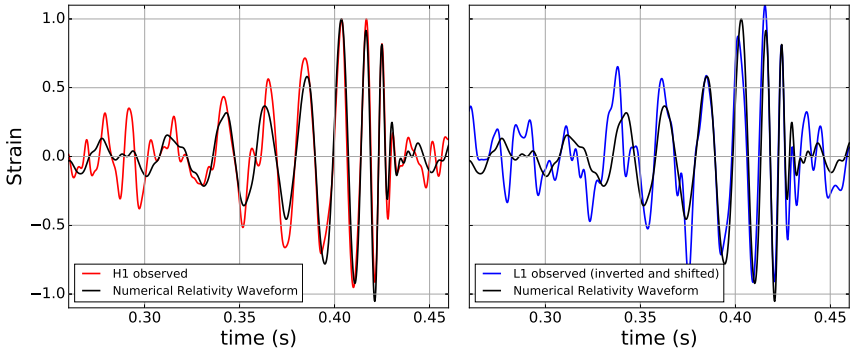


Figure 9.8 Comparison of our denoised waveforms (red and blue colours) with the filtered BBH template from numerical relativity [Abbott et al. 2016b] (black curve). The rROF algorithm has been used with $\lambda_0 = 0.2$ and 10 iterations. The times shown in the x-axis are as in Abbott et al. [2016b].

out higher frequencies locally which results in a somewhat closer similitude between both sets of waveforms.

In order to make a fair comparison, we have applied the rROF method to both, the observed data and to the same numerical relativity waveform employed by Abbott et al. [2016b]. This comparison is also shown in Fig. 9.8 (black lines in both plots). The shape of the denoised waveform agrees with the shape of the processed numerical relativity waveform template of Abbott et al. [2016b] for the last 4 cycles and the ringdown part of the signal [2]. To have a quantitative measure of the quality of our results, we calculate the Mean Square Error (MSE) between our denoised waveform and the numerical relativity waveform in the time interval shown in Fig. 9.8. We also perform the same calculation with the data resulting from applying a simple whitening process to the original H1 and L1 data and filtering them with a 35-350 Hz bandpass filter. For the Hanford detector our results yield a MSE value of 0.0195 while for the whitened and filtered data the MSE value is 0.0489. These low values of the MSE show that both methods remove noise successfully. As in Abbott et al. [2016b] we have also computed the spectrograms of the waveforms for our denoised signals. These are shown in Fig. 9.9 for the signals of both detectors. The superimposed isocontours appearing in both spectrograms correspond to the numerical relativity waveform. These lie on top of the time-frequency diagrams of the denoised signals. The distinctive increase in frequency during the final cycles of the chirp signal is

[2]Note that since we do not reconstruct the waveform but just denoise it, we cannot present the same comparison reported in the middle panel of Fig. 1 in Abbott et al. [2016b] between the waveform reconstructed through a Bayesian approach and the reconstructed numerical relativity template.

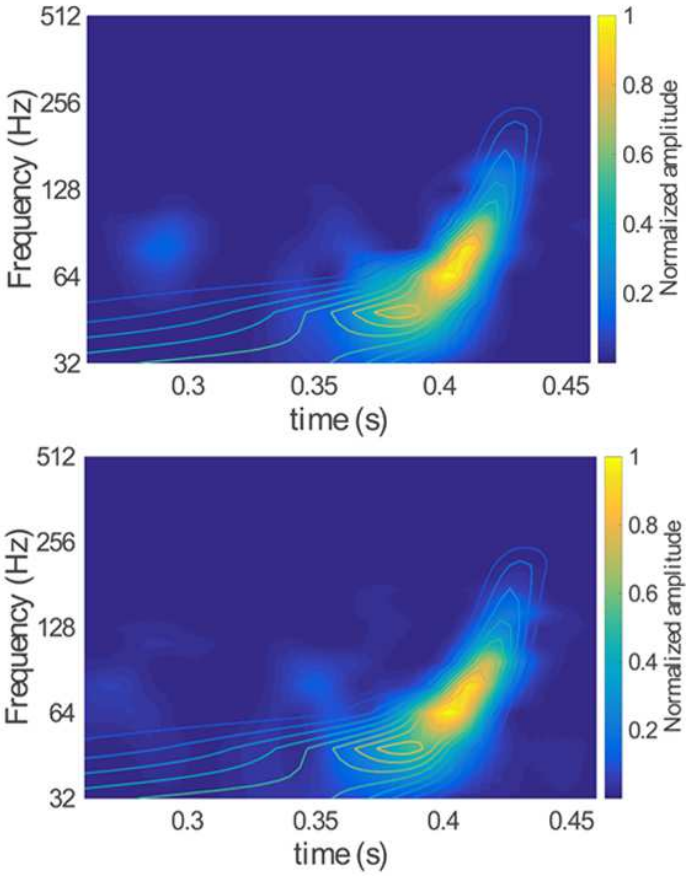


Figure 9.9 Spectrograms of our denoised waveforms after applying the rROF algorithm to the data from Hanford (upper panel) and Livingston (bottom panel). The superimposed isocontours correspond to the numerical relativity waveform template.

clearly recognisable in both spectrograms. Our results are qualitatively similar to the results reported by Abbott et al. [2016b].

9.6 Discussion

We have applied the ROF algorithm to denoise signals embedded in real noise from initial LIGO and to denoise the transient gravitational-wave signal GW150914 detected on September 14, 2015 by the two Advanced-LIGO interferometers. Our goal has been to assess if the rROF algorithm applied in the time domain to actual gravitational-wave data can successfully remove noise without any a priori information about the signal. The results reported in this

chapter show that the non-Gaussian, non-stationary noise from an actual observation such as the gravitational-wave event GW150914 can also be successfully removed with TV-denoising methods. That is indeed the case and the quality of the denoised waveform is qualitatively comparable to that obtained with the Bayesian approach used in the discovery paper. However, this signal has a high SNR, so the test is simple. Further analysis with other signals will be performed in the future.

Chapter 10

Gravitational-wave denoising via dictionary learning

The results of this chapter have been originally published in:
Torres-Forné, A., Marquina, A., Font, J. A., and Ibáñez, J. M. "Denoising of gravitational wave signals via dictionary learning algorithms". Physical Review D, **94**, 124040 (2016).

The sparse reconstruction of signals over trained dictionaries we discuss in this chapter are obtained for the same kind of GW trained signals that we did in the previous chapter, namely burst signals from a catalog of rotational stellar core collapse [Dimmelmeier et al. 2008] and chirp-burst-ringdown signals from a catalog of BBH mergers [Mroué et al. 2013] (see Appendix B for details). Once the dictionaries are set, we demonstrate their utility for the denoising of GW signals embedded in Gaussian noise only. The assessment of the approach when using actual detector noise is deferred for a future investigation.

10.1 Dictionary generation and parameter evaluation

10.1.1 Dictionary generation

We start describing the process to generate a learned dictionary from the waveforms of the catalogs. The goal is to find the best set of dictionary parameters that produce the best denoising results. As in the previous chapters we find that the results depend critically on the value of the regularization parameter λ selected. The way we build the dictionaries for the burst and BBH catalogs is similar. We divide randomly in three groups both the 128 burst waveform signals of Dimmelmeier et al. [2008] and the first 100 BBH waveforms of Mroué et al. [2013]. Since the BBH signals of Mroué et al. [2013] are quite large, we do not use the entire BBH catalog in order to save computational resources. Specifically, the BBH catalog covers binaries with total mass $20M_{\odot}$ and mass ratios up to 1:8, and so does our dictionary. We then use in either case 80% of the waveforms for training the dictionary, 15% for validation of the method, i.e. to search the best set of parameters, and the remaining 5% to test the algorithm in different situations.

The numerically generated signals are embedded in non-white, Gaussian noise corresponding to Advanced LIGO proposed broadband configuration, provided by the LSC Algorithm Library Suite (LAL) ["LAL software documentation (2007)"]. The frequency ranges from 10 Hz to 8192 Hz (one-sided spectrum). First of all, we resample the waveforms of both catalogs to the Advanced LIGO/Virgo sampling rate of 16384 Hz, zero-padded to have the same length. The corresponding signals are also shifted to be aligned with either the minimum peak in the case of bursts or with the maximum peak in the merger part for BBH signals. We select 2048 samples around the corresponding alignment points to train the dictionary. With this length, the waveforms of the burst catalog fit completely in the window, while only the last cycles of the inspiral, merger and ringdown of the BBH waveforms are taken into account to perform the denoising. This late part of the BBH signal is arguably the most interesting part, hence deserving to be denoised best. Below we comment on the reason for this choice and on existing alternatives to also reconstruct the early inspiral part of the signal. To ensure the best conditions for the convergence of the algorithms and to avoid round-off errors, we also scale the amplitude of the validation signals of both catalogs so that their maximum value is set to unity. The values of the regularization parameter λ we discuss in this section are hence determined by

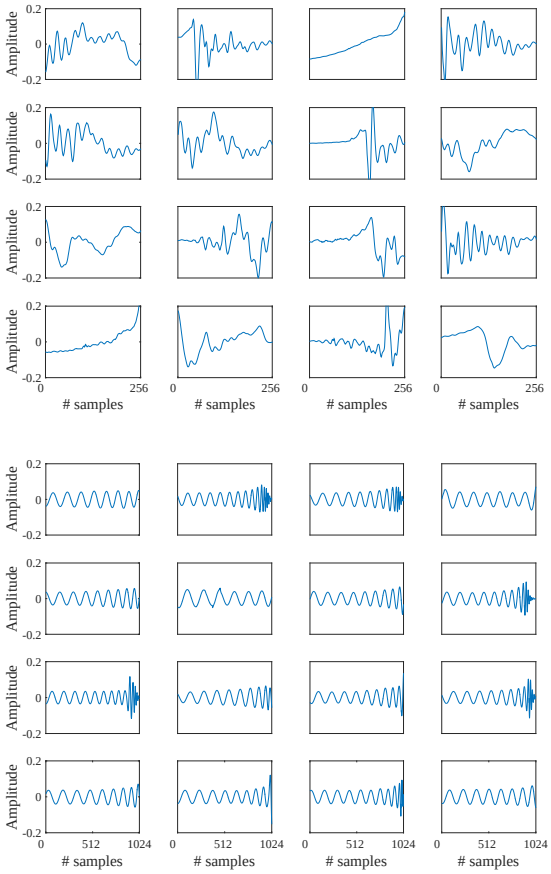


Figure 10.1 Random examples of the atoms of both dictionaries, for burst waveforms (upper panel) and for BBH signals (lower panel). The number of samples is shown in the horizontal axis while the normalized amplitude is shown in the vertical axis.

this normalization. Moreover, we scale each signal to achieve a specified value of the SNR, computed as in Eq. (8.3).

As discussed in preceding chapters, the optimal value of the regularization parameter, $\lambda_{\rm opt}$, is defined to be the one which gives the best results according to a suitable metric function applied to the denoised signal and the original one, measuring the quality of the recovered signal. In our case we choose two estimators, namely the Mean Squared Error,

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\hat{Y}_i - Y_i)^2 \ , \tag{10.1}$$

where $\hat{Y}$ and Y are the reconstructed and original signals, respectively, and n is the number of samples, and the structural similarity (SSIM) index [Wang et al. 2004], which deviates from the traditional measures of error because it takes into account the structural information. The SSIM index varies between 0 (minimum similarity) and 1 (maximum similarity) and is defined in Eq. (9.6)

As mentioned before we use 80% of the signals of each catalog to produce one dictionary per type of signal. To do the learning, we select 3×10^4 random patches (the starting sample is random) of a selected length, which is a parameter to be estimated. The patches are selected uniformly from all the learning waveforms of each catalog. Then, we select the p patches with the highest energy, defined as the square of the L^2-norm of each patch. After that, we solve problem (7.59) using a block-coordinate descent method. This step is done modifying the code developed by Peyré [2011]. Fig. 10.1 shows a small representation of the atoms of both dictionaries.

In addition to the search of λ_{opt} we must decide the best values for the size of the dictionary, i.e. the number of atoms and their length. To this aim we calculate the MSE for the reconstructed signals obtained using dictionaries of different sizes. In each case, the value λ_{opt} will be the corresponding value that minimizes the MSE. For this task we use the validation set of signals of the dictionaries and set the SNR to 20. As the length of the atoms is always shorter than the length of the validation signals, we do the denoising with a sliding window with an overlap of $n-2$ samples, where n is the length of the window, which agrees with the length of the atoms. With this overlap, there are many samples that are repeated on different windows. These samples must be averaged to obtain the final reconstructed signals. Our initial tests show that the best reconstruction is achieved using TV-averaging (see Marquina and Osher [2008b]),

$$s = \frac{\sum_{j=1}^{p}(f_j \mathrm{TV}_j)}{\sum_{j=1}^{p} \mathrm{TV}_j} \ , \tag{10.2}$$

where f_j corresponds to the current patch and $\mathrm{TV}_j = \sum |\nabla f_j|$ is the TV norm of that patch.

10.1.2 Parameter evaluation

We calculate the MSE for the validation set of burst signals with window lengths $l = 64$, 128, 256, and 512. The results are shown in Fig. 10.2. Each vertical bar represents the value of the MSE for each atom length. The figure shows that the largest value of the MSE is achieved for a length of 64. This is due to the fact

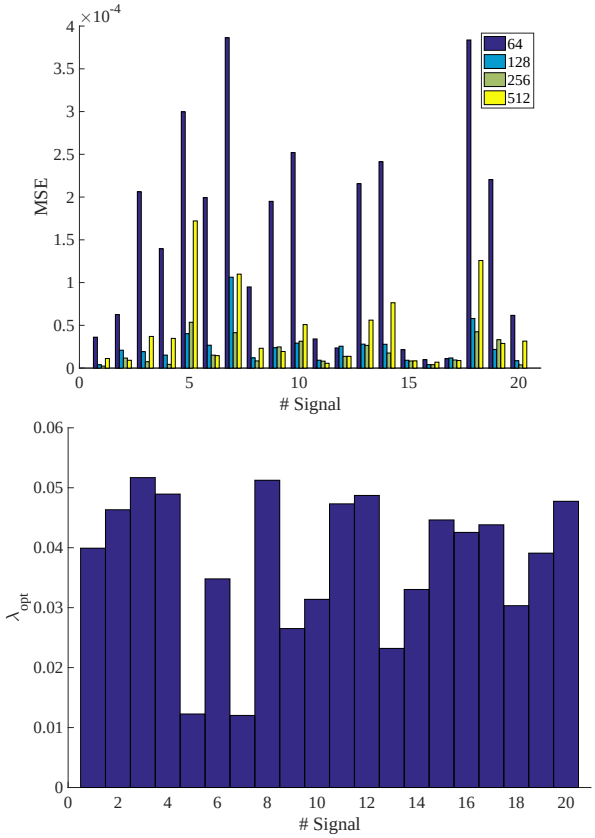


Figure 10.2 *Upper*: Bar diagram of the MSE for all burst validation signals. Each color represents a different window length as indicated in the legend. *Lower*: Bar diagram of the optimal value of λ for all burst validation signals and a window length $n = 256$. The mean value of λ_{opt} is 0.038 ± 0.018.

that if the atoms are too short the reconstructed signal is more oscillatory due to the noise. This effect can be corrected using larger lengths. However, the larger the length of the atoms the more difficult to recover the smallest oscillations of the original signal. This is the reason why the MSE actually grows for the larger window length analyzed (512 samples). While this is a generic trend, it is nevertheless still possible that the longest window may work better for specific signals (e.g. signals #6 or #11 in Fig. 10.2). However, in general the best results correspond to a length of 256 samples.

The lower panel of Fig. 10.2 displays the values of λ_{opt}, i.e. the value of the regularization parameter that minimizes the MSE value. It has been obtained for a fixed window length $l = 256$. This figure reveals that the values of λ

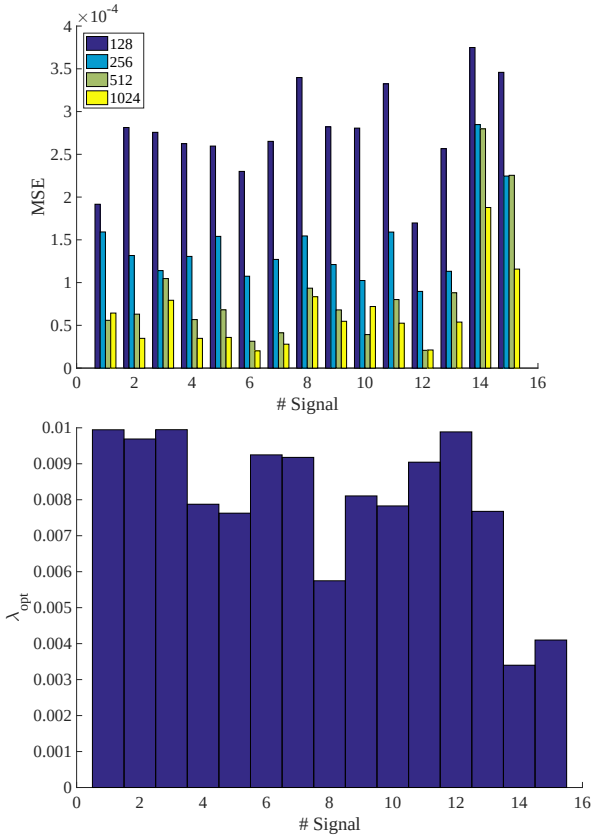


Figure 10.3 *Upper*: Bar diagram of the MSE for all the BBH validation signals. Each color represents a different window length as shown in the legend. *Lower*: Bar diagram of the optimal value of λ for all the BBH validation signals and a window of length 1024. The mean value of λ_{opt} is 0.008 ± 0.002.

are bounded between 0.01 and 0.06. Therefore, not all values of λ are possible and selecting the mean value $\lambda_{\text{opt}} = 0.03$ will produce, on average, a good reconstruction for all burst signals. Nevertheless, fine-tuning this parameter can improve the results in specific cases. We next carry out the same analysis for the case of BBH signals. As BBH waveforms are totally different to burst signals, the choices just discussed for bursts would not lead to satisfactory results if applied blindly to the BBH catalog. Contrary to burst signals, BBH waveforms are significantly longer, therefore we need to increase the length of the atoms. The values obtained for the MSE for the BBH catalog are shown in Fig. 10.3 and correspond to atom lengths that comprise from 128 to 1024 samples. The case of 64 samples is not shown in the figure because the corresponding value of

the MSE is much larger. As for the case of burst signals, Fig. 10.3 shows that the MSE decreases with the window length in most cases. Therefore, to denoise BBH waveform signals we select the length of 1024 samples as it produces the best results.

The corresponding results for the value of $\lambda_{\rm opt}$ for BBH are shown in the bottom panel of Fig. 10.3. Again, the values are restricted to a small interval between 0.002 and 0.012. As we show below, using the mean value, $\lambda_{\rm opt} = 0.008$, yields to satisfactory denoising results in most cases.

A similar study is required to determine how the results depend on the *number* of atoms of the dictionary p. In general, the larger the dictionary the better the results, but at a higher computational cost, ~ 1 hour on a desktop computer. Therefore, setting the size of the dictionary is often a trade-off between results quality and efficiency. To evaluate an optimal value for the number of atoms we carry out tests with the two catalogs using values from $p \in \{300, 500, 1000\}$ in the case of bursts with $n = 256$ atom length and from $p \in \{1100, 2000, 2500\}$ in the case of BBH with $n = 1024$ atom length. We find that using 500 and 1100 atoms for bursts and BBH, respectively, is a valid compromise as it produces good results at a reasonable computational cost. However, if computational resources are not an issue, there is no reason not to use larger dictionaries. For the two catalogs, the value of $\lambda_{\rm opt}$ for $p = 500$ and $p = 1100$ atoms are bounded in a similar interval than shown before.

10.2 Tests and results

10.2.1 No signal

The first test consists in studying the performance of the method when there is no signal inside the data set. The goal of this test is to check if in the absence of signal the dictionary produces spurious signals due to noise. The result of this test is shown in Fig. 10.4. A stream of 0.5 s of pure non-white Gaussian noise (upper panel) is denoised using the generic value of $\lambda_{\rm opt}$ corresponding to burst signals, i.e. $\lambda_{\rm opt} = 0.03$. One can see that the resulting signal has zero amplitude throughout the frame (lower panel) for this specific value of λ. This is the ideal behavior of the algorithm in order to avoid false detections due to noise.

We next repeat this test for 200 independent realizations of noise (following the procedure outlined in Appendix A of Torres et al. [2014]) to check if this behavior remains the same irrespective of the noise realization. For our specific value of λ we find 26 false reconstructions due to noise fluctuations. We note

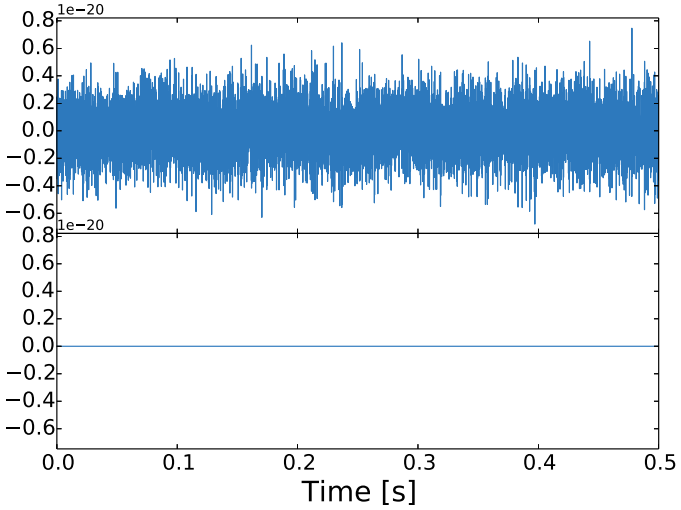


Figure 10.4 Denoising with no signal embedded into Gaussian noise with SNR=20. The upper panel shows the original noisy signal while the lower panel shows the ideal result of the denosing, i.e. a zero amplitude signal.

however that the smaller the λ the more coefficients of the representation become nonzero and more structures due to noise may appear. In contrast, a large value of λ will reduce the ratio of false reconstructions, even though a true GW signal with low SNR could be missed. For instance, for $\lambda = 0.045$ we only obtain one false reconstruction.

The results reported in this section are illustrative of the typical response of the LASSO algorithm on λ. A comprehensive statistical study of the dependence of the number of false reconstructions and signal misses on the parameters of the method, i.e. value of λ, type of signal injection, SNR, and noise realization, deserves further analysis. We also note that this is a fairly simple test because the noise is purely Gaussian. In a more realistic scenario, the presence of instrumental glitches in the detector data [Powell et al. 2015, Powell et al. 2017] could produce false reconstructions.

10.2.2 Signals from the catalogs

Next we study how the method works when applied to the eight test waveform signals of the burst catalog in a long data frame. In the figures for this and the following tests, we use the same noise realisation to compare the results on an equal footing. Correspondingly, in the tables reported in this section, we present results obtained with 20 different noise realisations to find out how the

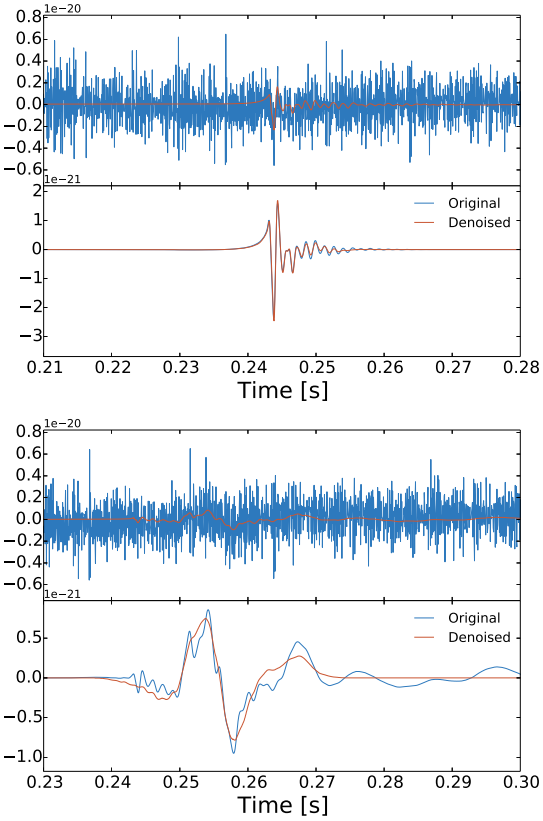


Figure 10.5 Denoising of signal #1 (Upper) and #6 (Lower) from the group of test signals of the burst catalog. The time of arrival is random and the SNR is 20. Upper panels: noisy signals (blue) superimposed with the original numerical relativity waveforms (red). Lower panels: comparison between the denoised signals (red) with the original ones (blue). The MSE and SSIM values are 0.018×10^{-3} and 0.98 for the signal on the upper panel and 0.271×10^{-3} and 0.67 for the signal on the lower panel.

reconstruction is affected by noise fluctuations. The signals are embedded in Gaussian noise with a SNR of 20. The time of arrival is fixed and it is the same for all the signals. The value of the regularisation parameter is set to $\lambda = 0.03$ and remains the same value for all the tests of this section. Although this value is not the optimal one, i.e. the one which produces the best results for a given signal, our goal is to determine if it is possible to recover the signal with a generic value of λ. This approach may be closer to what occurs in a realistic situation, where no information on the signal is available a priori.

The quality of the denoising is measured using the MSE and the SSIM metric functions, and is reported in Table 10.1 for all test signals. This table shows the

maximum and the minimum values for both MSE and SSIM for 20 independent noise realisations. We recall that the results depend on the value of λ. Each signal embedded in different noise realisations is a new scenario and the best results will be obtained with the optimal value of λ for each case. With SNR 20 and $\lambda = 0.03$ the relative variations are not too large (the highest variation in SSIM is 14% for signal #6). Therefore, at this SNR, the reconstruction is not very affected by noise fluctuations. Fig. 10.5 shows the results for only two signals of the catalog, namely those which yield the best (signal #1; top panel) and the worst (signal #6; bottom panel) denoising results, respectively (for the chosen value of λ and noise realisation). The figure displays the comparison of the two original noisy signals (upper subpanels) with the recovered ones (lower subpanels). Concerning the signal on the top panel our method can accurately recover the distinctive positive and negative peaks associated with the hydrodynamical bounce that follows the collapse of the inner iron core of the star once the equation of state stiffens and the central density exceeds nuclear matter density. This is particularly clear for the peaks with the larger amplitudes, which are recovered properly. However, when the amplitude decreases (i.e. in the part of the temporal evolution associated with the quasi-radial oscillations of the newly formed neutron star) the signal becomes weaker than the noise and, as a result, the method returns a zero amplitude signal. It is also worth mentioning that in the part of the time series where the data are purely noise (no numerical relativity signal injected) the method returns a zero signal, as it should. The same behaviour is seen for the signal displayed on the bottom panel of Fig. 10.5, the dampened oscillations are weaker than the noise and the method sets their amplitude to zero. We note that signal #6 is somewhat different from the common features of the dictionary. As a result, while the broad morphology is still captured to some extent, the overall result is poorer than for the signal on the top panel. Even so, we note that the results can be improved by changing slightly the value of λ by adding more atoms to the dictionary. (We have checked that for $\lambda = 0.026$ the MSE is 0.1×10^{-3} and the SSIM is 0.77.)

To find out the dependence of the procedure on the SNR we reduce its value from 20 to 10, keeping the same value of λ. The results are displayed in Fig. 10.6 for the same signals #1 and #6 of the burst catalog. The results for all test signals and the corresponding maximum and minimum measures of the MSE and SSIM are also reported in Table 10.1. Fig. 10.6 shows that for SNR 10 signal #1 is still very well recovered and its most significant features can be reconstructed with relatively high accuracy. The MSE for this signal increases

Table 10.1 Values (maximum - minimum) of the MSE and SSIM error estimators for the eight burst signals we use as test signals and 20 noise realisations. Values are reported for both SNR 20 and 10.

Signal	SNR 20		SNR10	
	MSE ($\times 10^{-3}$)	SSIM	MSE ($\times 10^{-3}$)	SSIM
#1	[0.033 - 0.015]	[0.97 - 0.93]	[1.389 - 0.021]	[0.96 - 0.74]
#2	[0.124 - 0.030]	[0.95 - 0.85]	[1.264 - 0.125]	[0.89 - 0.60]
#3	[0.066 - 0.040]	[0.93 - 0.88]	[0.691- 0.073]	[0.89 - 0.74]
#4	[0.068 - 0.007]	[0.97 - 0.88]	[0.684 - 0.014]	[0.95 - 0.74]
#5	[0.052 - 0.022]	[0.94 - 0.89]	[1.335 - 0.041]	[0.90 - 0.53]
#6	[0.210 - 0.084]	[0.83 - 0.72]	[0.861 - 0.205]	[0.72 - 0.51]
#7	[0.130 - 0.083]	[0.90 - 0.84]	[2.350 - 0.103]	[0.88 - 0.43]
#8	[0.042 - 0.016]	[0.93 - 0.85]	[0.594 - 0.026]	[0.91 - 0.74]

an order of magnitude and the SSIM decreases from 0.98 to 0.91, still reasonably high. For the worst possible case of the test waveforms, signal #6, Fig. 10.6 shows that it can still be distinguished from the noise.

Comparing the results for all burst signals reported in Table 10.1 for SNR 20 and SNR 10, we see that, in general, the values of the MSE (SSIM) increase (decrease) if the SNR decreases. As the SNR decreases, the reconstruction is more affected by noise fluctuations and the difference between the maximum and minimum values of the quality indicators increases. As we have mentioned before, other values of λ could improve the results in each case.

We turn next to test the results from the BBH catalog. As these signals are much longer than burst signals, the total segment of data has a length of 2 s, in order to allow to change the time of arrival. In this case, it is set as the time where the merger is produced, randomly. The denoising results are reported in Table 10.2 for all BBH signals and displayed in Fig. 10.7 for a representative signal (#2). As the figure shows, the signal is properly denoised during its three distinctive parts, the inspiral, the merger, and the ringdown. In particular, the phase of the signal is well captured and the main, yet small, differences between the original and the denoised signal appear in the amplitude. We note that the actual signal is significantly longer than the zoom shown in this figure. The initial part of the signal, the inspiral phase with low frequencies, is not recovered because, as mentioned before, the dictionary is specifically designed to recover the merger part. The most striking incorrect feature of Fig. 10.7 is the presence of spurious oscillations visible after the ringdown. This is due again to the selection of λ. While using a larger value would remove these oscillations it is also possible that the amplitudes of the merger and ringdown parts of the signal

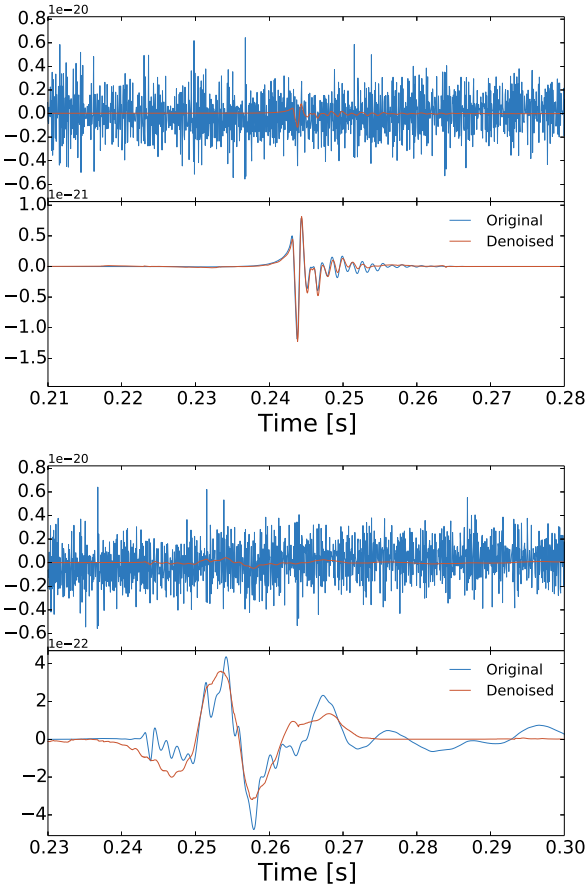


Figure 10.6 Same as Fig. 10.5 but with SNR 10. The MSE and SSIM values are 0.081×10^{-3} and 0.91 for the signal on the upper panel and 0.20×10^{-3} and 0.71 for the signal on the lower panel.

could be cut down. The corresponding MSE and SSIM measures are reported in Table 10.2 for both SNR 20 and 10. As for the case of burst waveforms, for BBH signals the values of the MSE (SSIM) also increase (decrease) as the SNR decreases, as expected.

10.2.3 Signals not included in the catalogs

In a realistic scenario, the GW signal will be unknown and it will be contaminated by several sources of noise. To test the code in a more unidealised setting we select in this section signals with similar broad morphology to those of the dictionaries but generated in a different way (e.g. employing different numerical

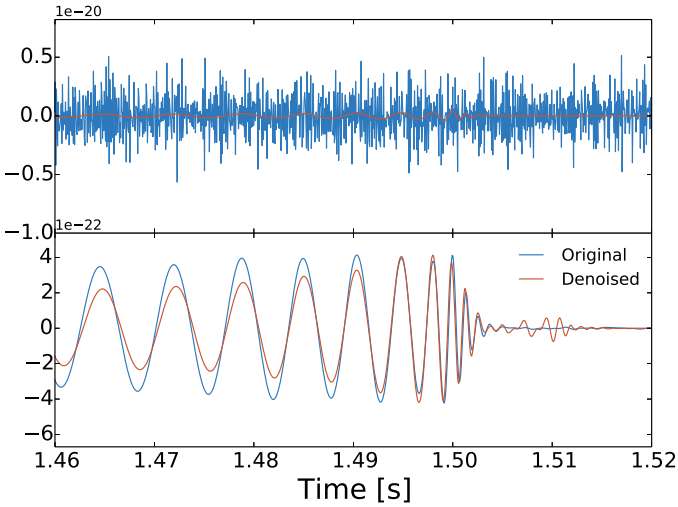


Figure 10.7 Denoising of the test signal #2 taken from the BBH catalog. The SNR is set to 20 in a 2 s frame. The value of λ used is 0.01 with TV averaging. The values of MSE and SSIM are 0.031×10^{-3} and 0.86 respectively.

Table 10.2 Values (maximum - minimum) of the MSE and SSIM error estimators for the four BBH signals we use as test signals and 20 noise realisations. Values are reported for both SNR 20 and 10.

Signal	SNR 20		SNR 10	
	MSE ($\times 10^{-3}$)	SSIM	MSE ($\times 10^{-3}$)	SSIM
#1	[0.025 - 0.019]	[0.89 - 0.86]	[0.084 - 0.027]	[0.87 - 0.76]
#2	[0.060 - 0.027]	[0.86 - 0.79]	[0.104 - 0.039]	[0.83 - 0.66]
#3	[0.029 - 0.020]	[0.88 - 0.87]	[0.101 - 0.032]	[0.86 - 0.74]
#4	[0.034 - 0.019]	[0.89 - 0.86]	[0.101 - 0.025]	[0.87 - 0.74]

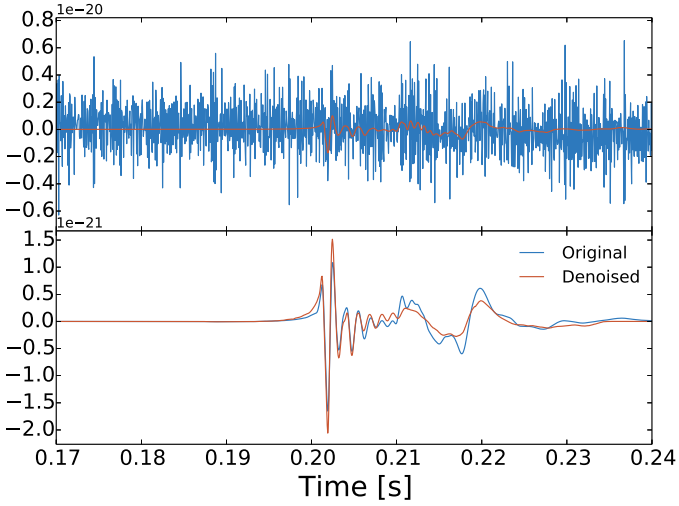


Figure 10.8 Denoising of a burst signal from the core collapse catalog of Abdikamalov et al. [2014b] using a dictionary generated from a different catalog [Dimmelmeier et al. 2008]. The arrival time is random, the SNR is 20 and $\lambda = 0.03$.

codes or input physics). While this analysis is still simple, since it involves simulated Gaussian noise without glitches, it is nonetheless more realistic because, contrary to the cases analysed before, the signals are now different from those of the catalogs from which the dictionaries are generated.

We first consider a burst signal from a core collapse catalog generated by Abdikamalov et al. [2014b]. We select signal #1 from this catalog, embed it into Gaussian noise with a SNR 20, and proceed to denoise it employing our burst dictionary. The results of the denoising are displayed in Fig. 10.8. This figure shows that the positive and negative peaks associated with core bounce are well recovered. The values of the MSE and SSIM error estimates are respectively 3.78×10^{-5} and 0.96. Notwithstanding some characteristics of the signal are lost, the signal can nevertheless be clearly distinguished from the noise and the main features are well recovered.

We can carry out a similar study for our BBH dictionary. As mentioned before, we select the BBH signal 'R1' from Baker et al. [2007]. The result of the denoising is displayed in Fig. 10.9 which shows that the reconstruction is much less accurate than in the case of the BBH test signals discussed before. We obtain MSE $= 5.79 \times 10^{-4}$ and SSIM $= 0.55$, values which indicate a poor reconstruction. The merger is not correctly recovered and the reconstruction introduces a phase shift. We must recall once again that we are using a generic

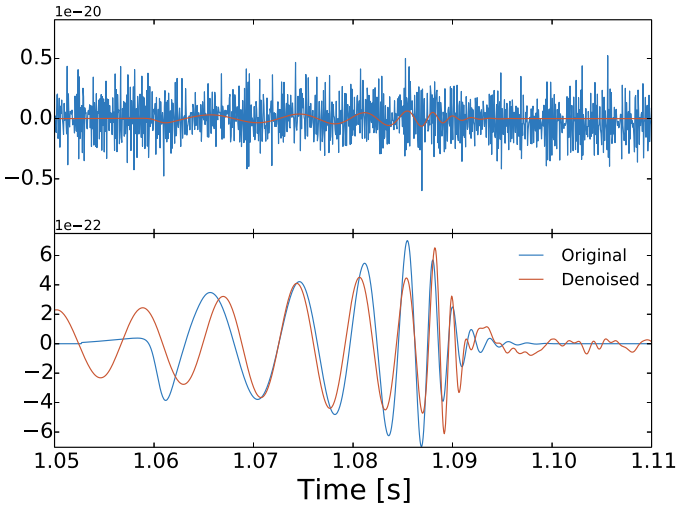


Figure 10.9 Denoising of the 'R1' BBH signal computed by the GSFC group [Baker et al. 2007] using a dictionary trained with signals from a different catalog [Mroué et al. 2013]. The SNR is set to 20 and the time of arrival is random in a 2 s frame. The value of λ used is 0.09 with TV averaging.

[

value of λ and therefore the result could be significantly improved by choosing a more suitable value. However, the goal of this test is not to obtain the best result possible but to assess our procedure in a scenario where the incoming signal is unknown and differs from those used to train the dictionary.

10.3 Complementary Tests

10.3.1 Iterative denoising

The next situation we consider involves a simple direct extension of the method, namely using the denoising procedure in an iterative way. By removing noise iteratively we find that the small amplitude oscillations of the signals are recovered better than using one single iteration. In this approach we use a generic, low value of λ, which only removes a small amount of noise in every iteration. We apply this iterative approach to burst signal #6 with SNR 20. The results are shown in Fig. 10.10 (to be compared with the bottom panel of Fig. 10.5). For this SNR we find that typically only 2 or 3 iterations suffice to recover the small amplitude oscillations of the signal and improve the results. For this test

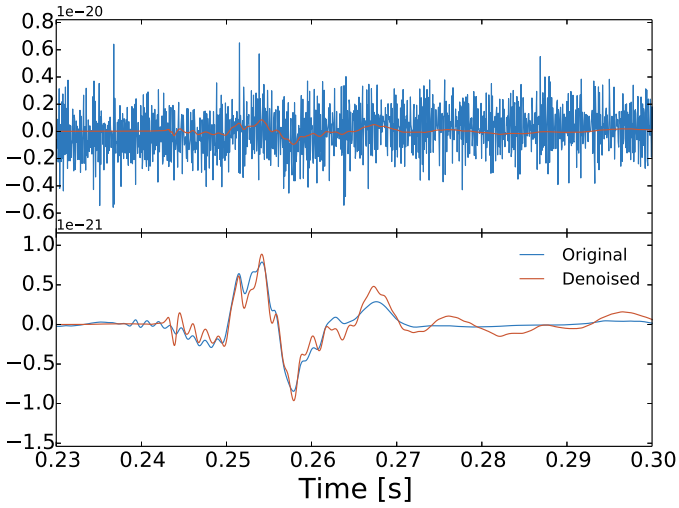


Figure 10.10 Denoising of signal #6 of the burst catalog employing an iterative procedure with only 2 iterations. We choose $\lambda = 0.01$ and SNR 20.

we obtain MSE $= 0.079 \times 10^{-3}$ and SSIM $= 0.81$. These values are considerable better than those reported in Table 10.1 for the case of a single iteration.

10.3.2 Combination of signals

For our next test we assess our procedure when the signal to denoise is a combination of two different signals. The goal of this test if to check the performance of our dictionaries when dealing with signals different from the type they are designed for. To do that, we build a test signal which is a combination of a burst and a BBH, both with SNR 20. We apply the algorithm using both dictionaries independently. The results of this test are shown in Fig. 10.11. The upper panel shows the original test signal (red line) embedded in Gaussian noise. The burst is located around $t \sim 1.34$ s while the merger of the BBH signal is visible at $t \sim 1.50$ s. The middle panel shows the reconstruction (red line) using only the burst dictionary with a value of $\lambda = 0.03$. The inset of this panel zooms around the time of the burst and subsequent oscillations of the PNS. Correspondingly, the lower panel displays the results of the denoising using only the BBH dictionary with a value of $\lambda = 0.01$. Clearly each dictionary discriminates well between the type of signal it has been designed to search for, despite both signals overlap in time. When using the burst dictionary the method returns no BBH signal, as can be seen in the middle panel. Likewise, when using the BBH dictionary, no burst signal is visible in the lower panel,

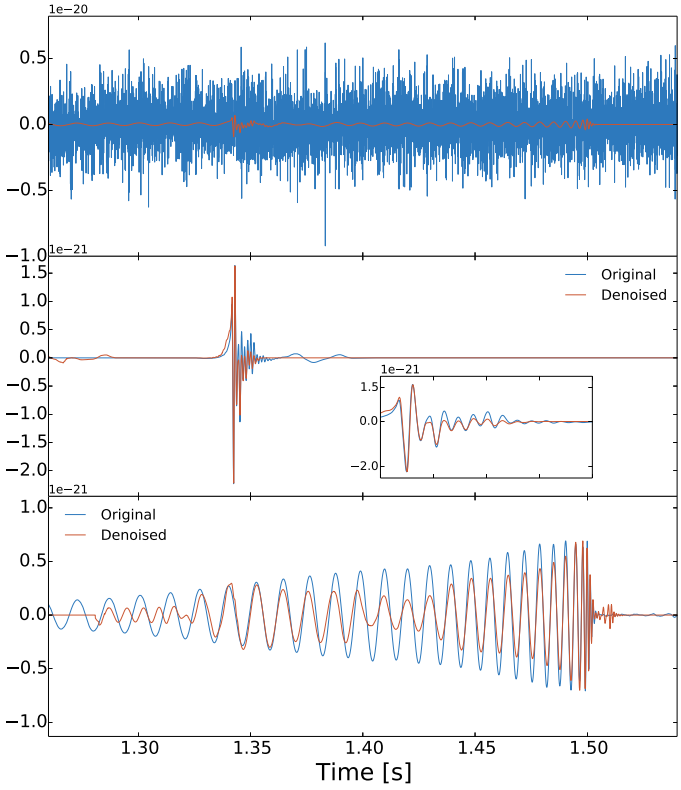


Figure 10.11 Denoising of a test signal composed by a combination of burst signal #5 and BBH signal #2. The individual signals are independently recovered when using the appropriate dictionary in a standalone way, as shown in the middle and bottom panels.

and the late inspiral and merger parts of the BBH signal are recovered properly. The discrepancies in the early inspiral and the spurious oscillations after the ringdown are to be expected, as we have explained before.

10.3.3 Low SNR scenario

The following test we consider is a low SNR scenario, namely SNR 6. Our strategy to denoise the signal in such a challenging situation consists in using the dictionary in combination with spectrograms. This test has two main goals. On the one hand, it allows us to check if our denoising procedure can improve the results of the spectrogram. On the other hand, we can test if the dictionary can recover the signal with acceptable accuracy in a low SNR scenario, once the time of arrival is known thanks to the spectrogram. We proceed as follows. First we apply the dictionary denoising with a generic value of λ, namely $\lambda = 0.02$.

Its value should be lower than that for SNR 20 to allow to recover the signal and also part of the noise. Then we calculate the spectrogram and select a window around the time of the maximum power (integrated over all frequencies). This step is a simple version of the event trigger generator implemented on the detectors [Abbott et al. 2016a]. Next, we apply the iterative denoising procedure to this small window. In this case, we select the number of iterations that minimizes (maximizes) the MSE (SSIM) values. We have observed that the dependence of these values with the number of iterations does not follow a convex distribution. Therefore, it is unfortunately difficult to find a general rule which gives information on what is the best number of iterations.

The results are shown in Fig. 10.12 for signal #1 of the burst catalog. The middle panel of this figure displays the spectrogram. The denoised signal shown in the lower panel after applying our two-step procedure clearly indicates the benefits of using the combined approach. The values of the MSE (0.2×10^{-3}) and SSIM (0.88) indicate that the reconstruction is quite accurate even with this low SNR. These values are similar to those for SNR 10 reported in Table 10.1. Therefore, the results of this test fulfill our two objectives. A key issue in this example is how to find the correct arrival time. We note that when using an even lower value of the SNR the arrival time obtained by the power integration does not always correspond to the arrival time of the signal. This issue could be solved by applying the iterative denoising directly to the list of candidate triggers.

10.3.4 LASSO selection

Up to now our experiments have focused on the goal of obtaining the best signal reconstruction for a generic value of the regularization parameter λ. In this section we show how the LASSO algorithm can also be used to infer some basic physical parameters of the sources from the denoised signals. We should point out that this analysis is just a numerical experiment and not a proper parameter-estimation procedure. We plan to further develop this approach towards parameter estimation in the near future.

We use the entire catalog as dictionary, except the test signals, without the learning procedure, and use the LASSO to select the signals closest to the one we use as test. The dictionary is not normalized in order to maintain the relative amplitude between the different waveforms that compose the catalog. Ideally, if a signal captured by the detector is inside the catalog, LASSO will select that signal. However, as the reconstruction is not perfect due to noise and the signal

is not inside the catalog, LASSO will return the combination of signals that are more similar to the denoised one. To investigate if the selection can indeed be used to extract the physical parameters of the original signal, we devise the following procedure: firstly, we perform the denoising of the signal with a random time of arrival. We employ SNR 20 and $\lambda = 0.03$. Secondly, once we have a clean signal, we use LASSO with the catalog and obtain the corresponding coefficients of each signal. It is possible to reconstruct the waveform using these coefficients and the catalog. Therefore, the value of λ in the selection is the one that minimizes the error between the denoised signal and the reconstructed one.

As an example, the results of the parameter estimation of test signals #1 and #7 of the burst catalog are shown in Table 10.3. The actual numbers of these two signals inside the catalog are #26 and #123 respectively. In this table the signal listed above the horizontal line is the test signal, and the next three lines indicate the corresponding three signals of the catalog with the highest LASSO coefficients (employing the original numbering of signals of the catalog). This table shows that the physical parameters of the collapse progenitors are reasonably identified, especially for test signal #123 where the only discrepancy is on the EOS. (This is however to be expected as the catalog only uses two EOS, Shen and LS, and the signal for the LS EOS is the actual test signal and is hence *removed* from the catalog.) The estimations for signal #26 are not as good but at least the differential rotation A of the progenitor is well obtained.

We turn next to estimate the parameters of test signal #2 from the BBH catalog using the same procedure as for burst waveforms. Table 10.4 reports the correspondence between the test signal and the three signals from the BBH catalog with the highest LASSO coefficients. Again, we find a good overall agreement, being the physical parameters of the test signal and the selected ones in a similar range. The main discrepancy is found in the BH masses, M_1 and M_2.

The mismatch observed in the parameter estimation is produced for two main reasons. On the one hand, the LASSO capabilities for parameter estimation obviously depend on how dense is the catalog. In the case of the two catalogs we employ in this work (and of most catalogs for that matter), the physical parameters are not sampled with detail due to the large computational cost of the simulations. It is likely that our results might improve when more complete catalogs become available. The second reason is due to the fact that we are using denoised waveforms as input, which includes the errors from the dictionary reconstruction. The more accurate the deinoising results, the more precise the parameter estimation. However, since in a real world application the original

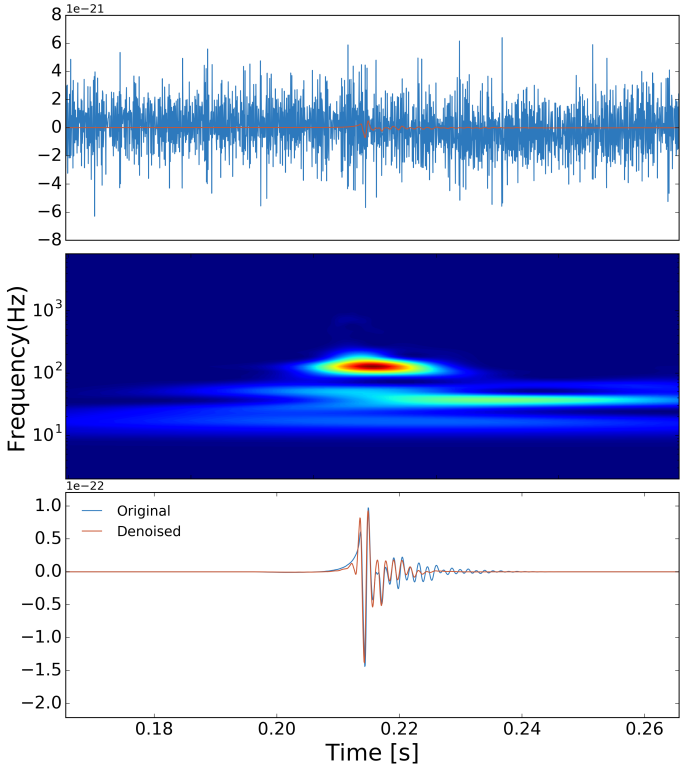


Figure 10.12 Denoising of signal #1 of the burst catalog employing iterative denosing and spectrograms. The middle panel shows the spectrogram of the 0.5 s denoising with a general value of λ. The results shown at the lower panel were calculated using $\lambda = 0.0095$ with 12 iterations, and SNR 6.

signal is unknown, the LASSO classification may still be regarded as a useful tool to complement existing parameter estimation techniques (see e.g. Jaranowski and Kròlak [2005], Abbott et al. [2016d], Logue et al. [2012], and Cornish and Littenberg [2015]).

10.4 GW150914

For our final test we assess our algorithms with the real GW data of the discovery signal GW150914. We use as few assumptions about instrumental noise as possible, due to the fact that the detector noise is non-Gaussian and non-stationary. However, a minimum noise preprocessing is required due to two main reasons. On the one hand, there are well-known modeled sources of narrow-band noise (see Fig. 3 of Abbott et al. [2016b]). On the other hand,

Table 10.3 Parameter estimation: comparison between the physical parameters of test signals #1 and #7 of the burst catalog. From left to right the columns report: number of catalog signal, model name, progenitor mass M, degree of differential rotation A, precollapse angular velocity at the center $\Omega_{c,i}$, precollapse rotation rate β_i, and equation of state.

Signal	Model name	M $[M_\odot]$	A $[10^8$ cm$]$	$\Omega_{c,i}$ [rad s^{-1}]	β_i [%]	EOS
#123 (#7)	s40A3O12	40	0.5	10.65	1.84	LS
#124	s40A3O12	40	0.5	10.65	1.84	Shen
#114	s40A2O13	40	1.0	6.45	2.60	Shen
#59	s15A3O12	15	0.5	10.65	1.60	LS
#26 (#1)	s11A3O09	11	0.5	8.99	0.72	Shen
#24	s11A3O07	11	0.5	5.95	0.40	Shen
#54	s15A3O05	15	0.5	4.21	0.25	Shen
#25	s11A3O09	11	0.5	8.99	0.72	LS

Table 10.4 Parameter estimation: comparison between the physical parameters of BBH test signal #2. From left to right the columns report: number of catalog signal, initial BH separation r_0 (in units of mass), initial orbital frequency $M\omega_0$, initial expansion factor $\dot{a}_0 = \dot{r}_0/r_0$, Christodoulou masses of the two BHs at $t = 0$, orbital eccentricity ϵ, number of orbits between $t = 0$ and common horizon time, and mass of final BH (remnant) M.

Signal	r_0	$M\omega_0$	$\dot{a}_0$	M_1	M_2	ϵ	Orbits	M
#47	14	0.017	-0.00028	0.75	0.25	0.00047	22.7	0.96
#56	15	0.015	-0.00028	0.84	0.16	0.00049	28.8	0.98
#29	16	0.014	-0.00033	0.60	0.40	0.00044	21.6	0.95
#28	16	0.014	-0.00026	0.60	0.40	0.00016	23.8	0.94

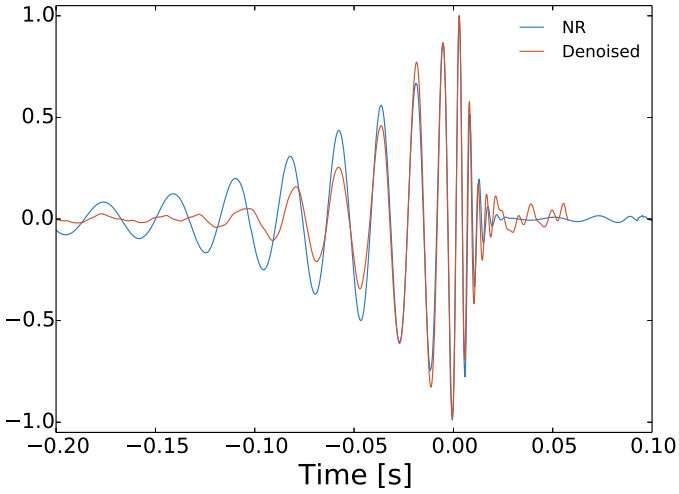


Figure 10.13 Denoising of signal GW150914 detected by Advanced LIGO Hanford interferometer. We choose $\lambda = 0.004$. The blue line indicates the NR template and the red curve corresponds to the actual signal. The amplitude of both signals has been rescaled to lie in the interval $[-1, 1]$.

ground-based detectors such as LIGO are not sensitive to low frequencies because of seismic noise. Therefore, we highpass the time series above 30 Hz to remove seismic noise and, following Abbott et al. [2016b], we also filter out all spectral lines.

The results of applying the dictionary denoising procedure to the GW150914 Hanford signal are shown in Fig. 10.13. The red curve in this figure displays the denoised signal obtained after applying the procedure to the real data. The implementation of the same approach to the best fit numerical relativity waveform [Abbott et al. 2016e] is shown with the blue curve. Even though the two signals are sampled at 4096 Hz, we use the high resolution dictionary as it leads to better results. This is common practice in the case of image denoising, where high resolution dictionaries are built to perform the denoising of low resolution images [Elad and Aharon 2006]. Fig. 10.13 shows that the last cycles of the inspiral signal, the merger part, and the ringdown agree well with the NR waveform, which corresponds to the best match template as in Abbott et al. [2016b]. Comparing the two signals of Fig. 10.13 at ± 0.15 s from the minimum of the numerical relativity signal yields MSE $= 0.0075$ and SSIM $= 0.4901$. These quality measures show that while visually the comparison between both signals seems satisfactory, the reconstruction is not very accurate. A full

parameter study to find the optimal values of the dictionary algorithms for signals embedded in real noise will be presented elsewhere.

10.5 Discussion

Our results demonstrate that it is possible to extract GW signals embedded in Gaussian noise with good accuracy, using a generic dictionary with a common value of λ for each type of waveforms. Overall, our results show that the denoising procedure works better for bursts than for BBH signals. This can be due to two main reasons. First, the morphology and the duration of both signals are very different. In particular, in the case of BBH signals, significantly longer than bursts, we need more atoms to cover all the signal duration. The second reason is related to the trained dictionary itself. The atoms on the burst dictionary are quite different between them which allows to cover more signal morphologies with a combination of only a few of them (sparse representation). In contrast, the atoms on the BBH dictionary are much more homogeneous since the signals used to train the dictionary are similar in the inspiral phase and only differ more clearly at the merger and the ringdown parts. Therefore, most of the atoms of the BBH dictionary cover the inspiral part, which can be reconstructed more easily. The larger inaccuracies appear when recovering the merger and the ringdown signals because there are less atoms to cover these parts and the reconstruction is less adaptable. A possible solution to this issue could be to use a couple of dictionaries, one to cover the inspiral part and a second one to cover only the merger and the ringdown parts. This may be worth investigating in the future. In addition, we have shown that using the LASSO algorithm iteratively can improve the results. Once a collection of triggers (i.e. arrival times) is available, it is possible to obtain the signal in a few iterations even for low SNR values. However, in order to determine the number of iterations that produces the optimal results, a more detailed study is necessary.

We have also reported results on the use of the LASSO algorithm as a classification method (i.e. for parameter estimation). The classification depends on how dense is the catalog and on how much noise can be removed from the original signal. The results become more accurate the larger the collection of waveforms available in the catalogs and the larger the physical parameters those catalogs cover, which is a major computational task. In particular, in the case of burst signals from core-collapse, the computational cost involved in calculating the GW waveforms renders unfeasible to obtain a large enough template bank. This classification method has therefore the same limitations

than matched filtering. Even so, using the LASSO algorithm as a classification method deserves attention, particularly if used jointly with matched filtering techniques. Finally, we have also briefly shown the performance of dictionary-learning techniques for actual GW signals under real noise conditions. The results for the discovery signal GW150914 seem promising. The study reported in this chapter has shown that if the data are in good enough agreement with the morphology of the atoms used to produce the dictionary, dictionary-learning algorithms may be used to extract signals from noise and to infer physical parameters. These algorithms could thus be a complementary addition to the GW wave data analysis toolkit.

Part V

Summary and outlook

Chapter 11

Summary and outlook

During the development of the research carried out in this thesis, I have worked in several important aspects of GW astronomy. The results of each specific project have already been discussed in the respective chapters. This final chapter summarises the main conclusions, providing an overview of the findings obtained in this work. It also outlines a possible list of future extensions.

11.1 Accretion onto neutron stars and the hidden magnetic-field scenario

We have studied the process of submergence of the magnetic field in a new-born neutron star during a hypercritical accretion stage in coincidence with a core-collapse supernovae explosion. This is one of the possible scenarios proposed to explain the apparently low external dipolar field of CCOs. Our approach has been based on 1D solutions of the relativistic Riemann problem, which provide the location of the spherical boundary (magnetopause) matching an external non-magnetised accretion solution with an internal magnetic field potential solution. For a given total accreted mass and magnetic-field strength, the magnetopause keeps moving inwards if the total (matter plus ram) pressure of the accreting fluid exceeds the magnetic pressure below the magnetopause. Exploring a wide range of accreted masses and field strengths, we have found the conditions for the magnetopause to reach the equilibrium point below the neutron star surface, which implies the burial of the magnetic field. Our study has considered several models with different specific entropy, composition, and neutron star masses, finding that these parameters do not have an important impact over the results.

Our main conclusion has been that, given the modest accreted mass required to bury the typical magnetic fields found in NSs, the CCO scenario is not unusual. On the contrary, the required accreted mass to bury the magnetar-like magnetic field are so large, that the NSs is most likely to collapse to a BH. However, our 1D approach, although sufficient to obtain a good approximation to the problem, does not take into account other effects, like convection, that can modify the results.

The natural extension of our initial study of the hidden magnetic-field scenario involves performing 2D numerical simulations. However, this is not an easy task due a number of reasons: (a) the large difference between the values of the magnetic pressure and the thermal pressure in some of the extreme regimes of this scenario, may lead to numerical inaccuracies or even to the normal termination of the code's execution; (b) the coupling between the magnetosphere and the hot fluid should be handled with care to avoid possible overheating effects. Even so, 2D simulations are essential to improve our understanding of the burial process of the magnetic field after a supernova explosion.

11.2 Linear-oscillation spectrum of proto-neutron stars

Studying the physics involved in the collapse of of the core of massive stars and the subsequent emission of GWs is not well-known. Moreover, the numerical modelling of such kind of systems is computationally very expensive - 3D simulations may take several months in present-day supercomputers. Therefore, it is not possible to infer the physics from the possible GW signals detected, as it is done in the case of BBH waveforms. The latter are significantly simpler systems than core-collapse supernovae and multiple waveforms from general relativity (using PPN and EOB approaches) and even numerical relativity for the merger phase, can be employed to match-filter with actual GW waveforms. In this section of the thesis our goal has been to devise a simple yet robust model, based on the theory of normal modes of oscillation of a spherical system, which allowed to infer physical parameters of the progenitor given a GW signal from the collapse.

As a first step we have determined if the oscillation modes of the PNS can be related with the GW spectrum of the progenitor. To do that, we have used a 1D model to obtain the eigenfunction of the oscillation modes from the data of an existing, state-of-the-art numerical simulation of this kind of scenario. Once the

origin of the modes has been identified, we have separated the distinct types of modes and have calculated their contribution in terms of GW power. Our results show that there exists a clear correspondence between the mode spectrum and the GW spectrum. Therefore, it seems possible to analyse a GW signal from core collapse in terms of the modes of oscillation of the PNS.

Having fulfilled this necessary, proof-of-concept, first step, we plan to take the next step in the near future. The idea is to perform several one-dimensional numerical simulations changing the parameters of the progenitor in order to study the dependence of the time-frequency distribution and power of the modes with those physical parameters (and possibly relaxing the Cowling approximation employed in the current work). The ultimate goal of this study is to develop a model to relate the parameters of the collapse progenitor with the corresponding PNS modes (and therefore with the GW spectrum) that allows us to infer the parameters directly from the GW spectrum without the need to perform the costly numerical simulations.

11.3 Glitch Classification

The presence of glitches inside the GW channels of the advanced LIGO/Virgo detectors is an effect inevitably linked to the extreme sensitivity required to detect GWs from cosmological distances. There are two possible strategies to eliminate them, or at least limit their impact. One is to try to determine the origin and take the necessary measures to solve the problem that causes the glitch. If this is not possible, or the origin of the glitch cannot be determined, the strategy consists in vetoing the data and not using it for science. The automatic glitch-classification techniques we have presented in this thesis can contribute to improve both strategies. On the one hand, they can relate the information provided by the auxiliary detector channels with the data on the GW channel and help to determine the cause of a given type of glitch. On the other hand, determining the type of glitch automatically helps the veto pipelines.

Our results have shown that our three pipelines can classify the glitches in advanced LIGO ER7 data with a 95% accuracy. In addition, using all three pipelines over the same data is a wining strategy since glitches missed by one pipeline can be classified by the others, increasing the overall efficiency. In addition, the WDF-ML pipeline includes an event-trigger generator different to the one used in the LIGO/Virgo detectors, and can therefore be used to cross-check the triggers generated by the latter. Our three pipelines are examples of how machine-learning methods can be applied to GW astronomy. From their

use in the detector subsystems and hardware design to their use in the realm of data analysis, there exists a large variety of machine-learning approaches to improve the possibilities of GW signal detection.

There are some interesting improvements that we plan to apply to the glitch-classification pipelines. To begin with, the whitening procedure for data-conditioning performed by the WDF-ML pipeline will be improved by using a technique known as adaptive whitening [Cuoco, Cella, and Guidi 2004]. Further improvements can also be made by using a training set of pre-classified waveforms or exploring the use of dictionary-learning algorithms, as the one presented in this thesis, for the specific problem of glitch classification. It is worth mentioning also the NSF-funded Gravity Spy project, which aims at building waveform datasets through a citizen-science (zooniverse) program [Zevin et al. 2016, Simpson, Page, and De Roure 2014]. The potential of glitch classification is maximised if the datasets can be employed on real time. In order to achieve this goal, the computational efficiency of all our pipelines will have to improved. An effort to build a more efficient code, adding parallelisation capabilities and even able to run on GPUs, has started at the time of writing of this thesis and will be continued in the near future in close collaborations with Drs. Elena Cuoco (EGO) and Massimiliano Razzano (University of Pisa).

11.4 TV methods for gravitational-wave denoising

Our study of TV-methods in the context of GW astronomy, introduced in this thesis, has begun with their testing under simplistic, Gaussian noise conditions. This initial step was necessary to understand their performance and to explore the parameter space of the methods to determine the values that produce the best results. This initial study has led to interesting results. The most important one is that the two algorithms we have used to solve the ROF problem are able to successfully remove noise and recover a recognisable GW signal. We have found that the choice of the regularisation parameter λ is crucial to obtain proper results. Therefore, it is important to devise an approach to bound the values of λ within an appropriate interval. As the variance of the noise in a GW detector is unknown, it is not possible to use the discrepancy principle to determine the regularisation parameter. As a result, performing an heuristic search based on some quality estimator (such as the PSNR or the SSIM) seems to be the best solution.

The Gaussian case has allowed us to understand the effect λ has on the different scales of the variational problem. It has become necessary to use different values of λ to recover different parts of the signal described by different (frequency or temporal) scales. In the case of BBH signals, this effect has become more apparent than for core-collapse signals, as they are significantly longer. In particular, we have seen that the low-frequency part of the inspiral signal can be recovered with more accuracy by using a value of λ lower than that required to recover the merger and the ringdown parts.

With the lessons learned with the Gaussian case, we have next considered a realistic scenario, using real data from initial LIGO. We have only employed the rROF method, because it is the method that gives the best results in the time domain, avoiding the problems with the borders that usually appear in the frequency domain. We have also modified the algorithm to use it iteratively. This change yields a smoother dependence with λ, increasing the interval of appropriate values. However, the computational cost also increases as the algorithm has to be executed more times. We have also observed that the low-frequency component of the noise present in GW detectors is difficult to remove using only TV-methods. This is related to the fact that the denoising properties of TV-methods is scale-dependent. As the rROF algorithms reduce the total-variation of the input signal, the higher frequencies will be removed faster then the lower ones. Nevertheless, one can use rROF with a low value of λ to isolate the low-frequency components and then remove them. Moreover, we have combined the rROF algorithm with both, a simple filter preprocessing and a whitening method. Our study has shown that TV-methods can improve the results of both pipelines and leads to identifiable denoised signals. Even with actual detector noise, TV-methods are an interesting tool to be combined with other common approaches of the GW data analysis toolbox.

The algorithms to remove noise from a GW signal presented in this thesis constitute a very promising line of research to further pursue in the near future. So far, we have only explored two algorithms based on the linearisation of the Euler-Lagrange equation to solve the ROF problem. It would be interesting to implement the dual-primal algorithm and compare the three methods under the same conditions, trying to determine which is the best algorithm to use in each case. It is also worth further exploring the relation of the regularisation parameter λ with the different scales, in order to devise an algorithm with an adjustable λ that produces the best results for each noise conditions. In this context, the patch-based algorithm as non-local means [Buades, Coll, and Morel 2011] may be a useful tool. Moreover, to characterise the possible improvements

of using a TV-denoising step in the results of other GW pipelines, it is necessary to perform a combined study. Finally, we also mention the graph cuts algorithms for denosing, which have received great attention in recent years. Applying them to GW data may be a very interesting project to tackle next.

11.5 Gravitational-wave denoising via dictionary learning

We have also studied the capabilities of learned dictionaries to recover GW signals from a noise-dominated background. Our LASSO algorithm has been tested using signals from two main sources, bursts from rotational core collapse and chirps from BBH coalescence. To obtain the respective dictionaries, we have used 80% of the waveforms for the training, 15% for the validation, i.e. to obtain the best set of parameters that produces the best results, and the last 5% waveforms to assess the method. An interesting feature of LASSO is that, for most Gaussian noise realisations considered, it returns zero if the input signal cannot be reconstructed by the atoms on the dictionary. As a result, the method may provide a fairly clear signal reconstruction. On the other hand, an intrinsic limitation of the method is that the results strongly depend on the selection of the regularisation parameter λ, whose optimal value cannot be set a priori, and must be obtained with validation studies.

There exist a large variety of learning techniques in the literature. In this thesis we have only considered one specific method but in the near future we plan to implement additional methods to perform the learning and to compute the LASSO algorithm more efficiently. Obtaining the denoised solution for one patch of 256 samples takes typically a few tens of ms on an Apple iMac computer with Intel Core i7 processor and 16 Gb of Ram. The most expensive computational cost is associated with the learning task. Reducing the time involved in this part of the method is a key issue in order to eventually apply the method in real time to the actual data generated by the detectors. In the next few months, advanced LIGO and advanced Virgo will (re)start observing runs with improved sensitivity, increasing the number of detections. The development of sophisticated data analysis techniques to improve the opportunities of detection, especially for low SNR events, is therefore a most crucial effort.

We defer for a future study with dictionary-learning methods the analysis of the false-alarm rate using simulated (or real) glitches. We believe that such an analysis, together with validation studies using real noise, are mandatory before

using dictionaries in a detector's pipeline. To avoid false positives caused by glitches, it is possible to employ dictionaries built from a collection of known glitches, which would allow to use LASSO as a glitch classifier. Moreover, as all methods discussed in this work have been originally developed for image reconstruction (2D data), we plan to apply them on GW spectrograms.

Part VI

Appendices

Appendix A

Noise Generation

We generate non-white Gaussian noise whose shape corresponds to Advanced LIGO in the proposed broadband configuration. For this purpose, we employ the algorithm libraries (LAL) ["LAL software documentation (2007)"] provided by the LIGO Scientific Collaboration (LSC) that include one-sided detector noise power spectral density $\sqrt{S(f)}$. If necessary, random noise series can be obtained weighting samples from Normal distribution by noise power spectral density,

$$\mathbb{R}\mathrm{e}(\tilde{n}(f)) = \frac{\sqrt{S(f)}}{2}\mathrm{N}(\mu, \sigma_f^2) \,, \tag{A.1}$$

$$\mathbb{I}\mathrm{m}(\tilde{n}(f)) = \frac{\sqrt{S(f}}{2}\mathrm{N}(\mu, \sigma_f^2) \,, \tag{A.2}$$

$$\{\tilde{n}(f) \in \mathbb{C} \ : \ f = 0, \Delta_f, 2\Delta_f, ..., (N-1)\Delta_f\}, \tag{A.3}$$

where $\Delta_f = \frac{F_s}{N}$ being F_s the sampling rate and N the number of samples. Mean (μ) and variance (σ_f^2) are set to 0 and 1 respectively as corresponds to the standard normal distribution.

As noise time-series are real-valued, in the frequency domain noise must satisfy

$$\tilde{n}(-f) = \tilde{n}^*(f) \,, \tag{A.4}$$

and

$$\mathbb{I}\mathrm{m}\{\tilde{n}(0)\} = \mathbb{I}\mathrm{m}\{\tilde{n}(f_{Ny})\} = \mathbb{I}\mathrm{m}\{\tilde{n}(-f_{Ny})\} = 0 \,, \tag{A.5}$$

where $f_{Ny} = \frac{F_s}{2}$ corresponds to the Nyquist frequency.

Properties of the Fourier transform assure that the Gaussian character of the noise signal is preserved when changing domain. Therefore, we obtain noise

time-series applying the discrete Fourier Transform (DFT)

$$\tilde{n}(f_k) = \sum_{j=0}^{N-1} n(t_j) e^{\frac{-2\pi i}{N} kj} , \tag{A.6}$$

$$n(t_j) = \frac{1}{N} \sum_{k=0}^{N-1} N(f_k) e^{\frac{2\pi i}{N} kj} . \tag{A.7}$$

Appendix B

Gravitational-wave catalogs

B.1 Rotating core-collapse catalog

At the final phase of their evolution, massive stars in the 9-40 $M_\odot$ range develop iron-group element cores which are dynamically unstable against gravitational collapse. According to the standard model of type II/Ib/Ic supernovae, the collapse is initiated by electron captures and photo-disintegration of atomic nuclei when the iron core exceeds the effective Chandrasekhar mass. As the inner core reaches nuclear density and the EoS stiffens, the collapse stops and is followed by the bounce of the inner core. A strong shock wave appears in the boundary between the inner core and the supersonically infalling outer core. Numerical simulations try to elucidate if this shock wave is powerful enough to propagate from the outer core and across the external layers of the star. In the most accepted scenario, energy deposition by neutrinos, convective motions, and instabilities in the standing shock wave, together with general-relativistic effects, are necessary elements leading to successful explosions.

In the core collapse scenario, conservation of angular momentum makes rotating cores with a period of one second to produce millisecond period PNSs, with a rotational energy of about 10^{52} erg. The bulk of gravitational radiation is emitted during bounce, when the quadrupole moment changes rapidly, which produces a burst of GWs with a duration of about 10 ms and a maximum (dimensionless) amplitude of about 10^{-21} at a distance of 10 kpc. Broadly speaking, GW signals from this mechanism exhibit a distinctive morphology characterized by a steep rise in amplitude to positive values before bounce followed by a negative peak at bounce and a series of damped oscillations

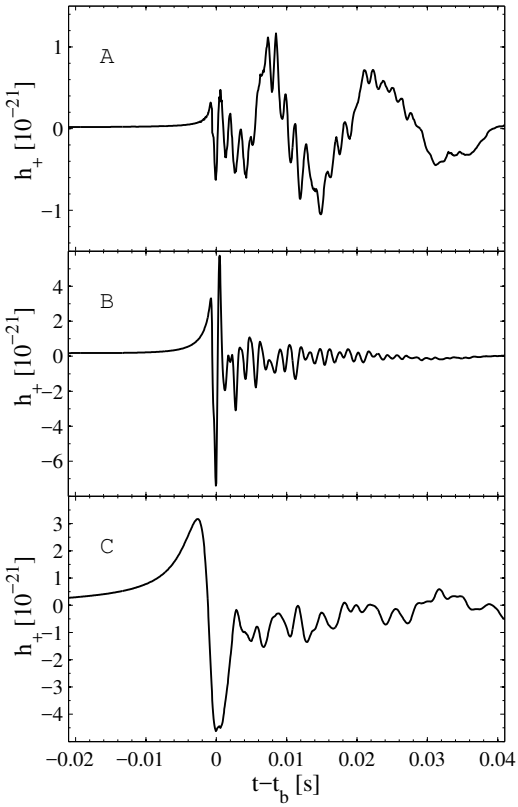


Figure B.1 Gravitational waveforms of three representative signals from the core collapse catalog of Dimmelmeier et al. [2008] with different values of the degree of differential rotation. The EoS and the progenitor mass are fixed for all represented signals. Signal "s20a1o05_shen" (A) is show in the upper panel, signal "s20a2o09_shen" (B) is shown in the middle panel, and signal "s20a3o15_shen" (C) in the bottom panel. $t_{\rm b}$ indicates the time of bounce.

associated with the vibrations of the newly formed PNS around its equilibrium solution.

Catalogs of gravitational waveforms from core collapse supernovae have been obtained through numerical simulations with increasing realism in the input physics (see Fryer and New [2011] and references therein). In our study, and for illustrative purposes, we employ only the catalog developed by Dimmelmeier et al. [2008], who obtained 128 waveforms from general relativistic simulations of rotating stellar core collapse to a NS. The simulations were performed with the CoCoNuT code and include a microphysical treatment of the nuclear EoS, electron capture on heavy nuclei and free protons, and an approximate deleptonization scheme based on spherically symmetric calculations with Boltzmann

neutrino transport. The simulations considered two tabulated EoS, those of Shen et al. [1998] and Lattimer and Swesty [1991] and a wide variety of rotation rates and profiles and progenitor masses.

The morphology and temporal evolution of the waveform signals of the catalog of Dimmelmeier et al. [2008] are determined by the various parameters of the simulations. The initial models include solar metallicity, non-rotating progenitors with masses at zero age main sequence of 11.2 $M_\odot$, 15.0 $M_\odot$, 20.0 $M_\odot$ and 40.0 $M_\odot$. Rotation has a strong effect on the resulting waveforms and it is fixed by the precollapse central angular velocity which is set to values from $\Omega_{\rm c,i} = 0.45$ to 13.31 rad s^{-1}. The angular velocity of the models is given by

$$\Omega = \Omega_{\rm c,i} \frac{A^2}{A^2 + r^2 \sin^2\theta}\,, \tag{B.1}$$

$r \sin\theta$ being the distance to the rotation axis, and A being a length parameterizing the degree of differential rotation. The specific values used in the simulations are $A = 50,000$ km (almost uniform rotation), $A = 1,000$ km, and $A = 500$ km (strong differential rotation). The GW amplitude maximum, $|h_{\rm max}|$, is proportional to the ratio of rotational energy to gravitational energy at bounce, $\frac{T_b}{|W|_b}$. The gravitational waveforms of the catalog are computed employing the Newtonian quadrupole formula where the maximum dimensionless GW strain is related to the wave amplitude $A_{20}^{\rm E2}$ by

$$h = \frac{1}{8}\sqrt{\frac{15}{\pi}}\frac{A_{20}^{\rm E2}}{D} = 8.8524 \times 10^{-21}\frac{A_{20}^{\rm E2}}{10^3\ {\rm cm}}\frac{10\ {\rm kpc}}{D}, \tag{B.2}$$

where D is the distance to the source.

We focus on three representative signals from the catalog to assess our algorithms. These three waveforms, shown in Fig. B.1, cover the signal morphology of the catalog, as explained in Logue et al. [2012]. These signals are labelled "s20a1o05_shen", "s20a2o09_shen", and "s20a3o15_shen" in the original catalog of Dimmelmeier et al. [2008]. We rename them, respectively, as signals "A", "B", and "C", in this work, to simplify the notation. The three values of the degree of differential rotation produce, in particular, the most salient variations in the waveform morphology.

In addition, for the results reported in Chapter 10, we also consider two extra signals, one from a core collapse catalog developed by Abdikamalov et al. [2014b] and a BBH signal from Baker et al. [2007]. These last two signals allow us to investigate the ability of our approach to extract independent waveforms using dictionaries built from atoms that do not contain explicit information on the signals to be denoised. The core-collapse template bank

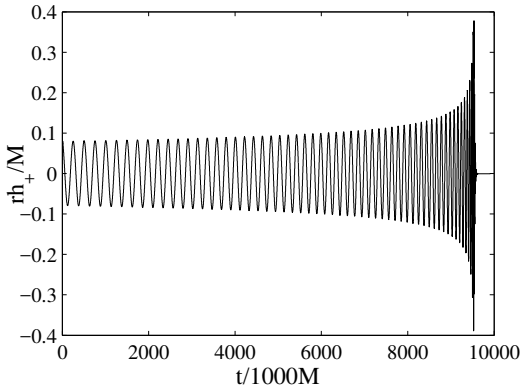


Figure B.2 Gravitational waveform of a representative signal from the BBH catalog of Mroué et al. [2013], signal "0001", used for our test purposes.

computed by Abdikamalov et al. [2014b] has also been built through axisymmetric simulations with the `CoCoNuT` code. The progenitors investigated have different initial angular momentum distributions in the core and the simulations include a microphysical finite-temperature EoS, an approximate electron-capture treatment during collapse, and a neutrino-leakage scheme for the postbounce evolution.

B.2 Binary-black-hole catalog

The second type of GW signals we use for our study is that from the inspiral and merger of BBH. Along with the merger of binary NS stars, such events are considered the most promising sources for the first direct detection of GWs. This kind of systems evolve through three distinctive phases, inspiral, merger, and ringdown, each one of them giving rise to a different signal morphology.

During the inspiral phase, the orbital separation between the two compact objects decays due to GW emission. Analytic approximations of GR such as post-Newtonian expansions or the Effective One Body method describe to high accuracy the motion of the system and the radiated GW content during such inspiral phase. These were, until recently, the only approaches available to model the GW signal before the merger of the compact binary. After the merger, on the other hand, the resulting single BH asymptotically reaches a stationary state characterized by the ringdown of its quasi-normal modes of oscillation, whose waveform signal can be computed using standard techniques from BH perturbation theory. Either of these techniques, however, cannot be used to model the signal during the merger phase when the peak of the gravitational

radiation is produced. At this phase, the signal waveform has to be computed solving the full Einstein equations with the techniques of numerical relativity, a long-lasting challenging problem that was only recently finally solved [Pretorius 2005, Campanelli et al. 2006, Baker et al. 2006]. Since those breakthroughs, many numerical-relativity simulations of BBH mergers followed, which led to the first BBH waveform catalogs of the entire signal (late inspiral, merger, and quasi-normal mode ringdown) based solely on fully numerical-relativity approaches.

Nowadays, numerical-relativity waveforms for BBH mergers have become increasingly more accurate and span the entire seven-dimensional parameter space, namely initial spin magnitudes, angles between the initial spin vectors and the initial orbital angular momentum vector, angles between the line segment connecting the centers of the BHs and the initial spin vectors projected onto the initial orbital plane, mass ratio, number of orbits before merger (late inspiral), initial eccentricity, and final spin. State-of-the-art waveforms are in particular reported by Mroué et al. [2013], and these are the ones we use for our tests. This catalog includes 174 numerical simulations of which 167 cover more than 12 orbits and 91 represent precessing binaries. It also extends previous simulations to large mass ratios (from 1 to 8) and includes simulations with the first systematic sampling of eccentric BBH waveforms. In addition, the catalog incorporates new simulations with the highest BH spin studied to date (0.98). As for the case of the core collapse burst catalog, it is sufficient to focus on a representative signal from the BBH catalog in order to illustrate the performance of our denoising algorithms. To such purpose we select signal labelled “0001” from the Mroué et al. [2013] catalog, which is shown in Fig. B.2.

B.3 Correspondence with signal catalogs

For the sake of completeness and to facilitate the identification of the NR waveforms used in this study, Table B.1 reports the correspondence of the waveforms used in Chapter 10 with the original naming of the burst catalog [Dimmelmeier et al. 2008] and of the BBH catalog [Mroué et al. 2013].

Table B.1 Relation between the number of the GW signals employed in Chapter 10 for the validation set (first two columns) and for the test set (last two columns) with the corresponding signal in the core-collapse catalog [Dimmelmeier et al. 2008] and in the BBH catalog [Mroué et al. 2013].

	Validation		Test	
#	Core Collapse	BBH (SXS)	Core Collapse	BBH (SXS)
#1	S15A2O13_shen	SXS:BBH:0022	S11A3O09_shen	SXS:BBH:0030
#2	S15A3O07_ls	SXS:BBH:0003	S15A1O09_ls	SXS:BBH:0047
#3	S15A3O15_shen	SXS:BBH:0028	S15A1O09_shen	SXS:BBH:0068
#4	S15A2O13_ls	SXS:BBH:0076	S15A3O15_ls	SXS:BBH:0087
#5	S15A1O01_ls	SXS:BBH:0001	S20A2O07_shen	
#6	S20A1O09_shen	SXS:BBH:0077	S40A1O01_shen	
#7	S11A1O01_shen	SXS:BBH:0090	S40A3O05_ls	
#8	S40A2O15_ls	SXS:BBH:0053	S40A3O12_ls	
#9	S20A2O05_ls	SXS:BBH:0019		
#10	S20A1O09_ls	SXS:BBH:0041		
#11	S15A2O07_ls	SXS:BBH:0091		
#12	S20A3O07_ls	SXS:BBH:0050		
#13	S11A2O07_ls	SXS:BBH:0080		
#14	S11A1O13_ls	SXS:BBH:0062		
#15	S15A2O05_ls	SXS:BBH:0059		
#16	S11A3O13_shen			
#17	S11A3O09_ls			
#18	S40A1O01_ls			
#19	S11A1O05_shen			
#20	S20A3O13_shen			

Bibliography

[1] J. Aasi et al. "Characterization of the LIGO detectors during their sixth science run." In: *Classical and Quantum Gravity* 32.11 (2015), p. 115012.

[2] J. Aasi et al. "Advanced ligo." In: *Classical and Quantum Gravity* 32.7 (2015), p. 074001.

[3] J. Abadie et al. "All-sky search for periodic gravitational waves in the full S5 LIGO data." In: *Phys. Rev. D* 85.2, 022001 (Jan. 2012), p. 022001. DOI: 10.1103/PhysRevD.85.022001. arXiv: 1110.0208 [gr-qc].

[4] B Abbott et al. "Implications for the Origin of GRB 070201 from LIGO Observations." In: *The Astrophysical Journal* 681.2 (2008), p. 1419.

[5] B. P. Abbott et al. "LIGO: the Laser Interferometer Gravitational-Wave Observatory." In: *Reports on Progress in Physics* 72.7, 076901 (July 2009), p. 076901. DOI: 10.1088/0034-4885/72/7/076901. arXiv: 0711.3041 [gr-qc].

[6] B. P. Abbott et al. "Binary Black Hole Mergers in the First Advanced LIGO Observing Run." In: *Physical Review X* 6.4, 041015 (Oct. 2016), p. 041015. DOI: 10.1103/PhysRevX.6.041015. arXiv: 1606.04856 [gr-qc].

[7] B. P. Abbott et al. "Observation of Gravitational Waves from a Binary Black Hole Merger." In: *Physical Review Letters* 116.6, 061102 (Feb. 2016), p. 061102. DOI: 10.1103/PhysRevLett.116.061102. arXiv: 1602.03837 [gr-qc].

[8] B. P. Abbott et al. "Properties of the Binary Black Hole Merger GW150914." In: *Physical Review Letters* 116.24, 241102 (June 2016), p. 241102. DOI: 10.1103/PhysRevLett.116.241102. arXiv: 1602.03840 [gr-qc].

[9] B. P. Abbott et al. "GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence." In: *Physical Review Letters* 116.24, 241103 (June 2016), p. 241103. DOI: 10.1103/PhysRevLett.116.241103. arXiv: 1606.04855 [gr-qc].

[10] B. P. Abbott et al. "Characterization of transient noise in Advanced LIGO relevant to gravitational wave signal GW150914." In: *Classical and Quantum Gravity* 33.13 (2016), p. 134001.

[11] B. P. Abbott et al. "Prospects for observing and localizing gravitational-wave transients with Advanced LIGO and Advanced Virgo." In: *Living Reviews in Relativity* 19.1 (2016), p. 1.

[12] B. Abbott et al. "All-sky search for long-duration gravitational wave transients with initial LIGO." In: *Physical Review D* 93.4 (2016), p. 042005.

[13] B. Abbott et al. "Directly comparing GW150914 with numerical solutions of Einstein's equations for binary black hole coalescence." In: *Physical Review D* 94.6 (2016), p. 064035.

[14] B. Abbott et al. "Observing gravitational-wave transient GW150914 with minimal assumptions." In: *Physical Review D* 93.12 (2016), p. 122004.

[15] E. Abdikamalov et al. "Measuring the angular momentum distribution in core-collapse supernova progenitors with gravitational waves." In: *Phys. Rev. D* 90.4, 044001 (Aug. 2014), p. 044001. DOI: 10.1103/PhysRevD.90.044001. arXiv: 1311.3678 [astro-ph.SR].

[16] E. Abdikamalov et al. "Measuring the angular momentum distribution in core-collapse supernova progenitors with gravitational waves." In: *Phys. Rev. D* 90.4, 044001 (Aug. 2014), p. 044001. DOI: 10.1103/PhysRevD.90.044001. arXiv: 1311.3678 [astro-ph.SR].

[17] T. Accadia et al. "Plans for the Upgrade of the Gravitational Wave Detector Virgo: Advanced Virgo." In: *Twelfth Marcel Grossmann Meeting on General Relativity*. Ed. by A. H. Chamseddine. 2012, pp. 1738–1742. DOI: 10.1142/9789814374552_0313.

[18] F. Acernese et al. "NAP: a tool for noise data analysis. Application to Virgo engineering runs." In: *Classical and Quantum Gravity* 22.18 (2005), S1041.

[19] F. Acernese et al. "Gravitational waves by gamma-ray bursts and the Virgo detector: the case of GRB 050915a." In: *Classical and Quantum Gravity* 24.19 (2007), S671.

[20] F Acernese et al. "Status of VIRGO." In: *Classical and Quantum Gravity* 25.11 (2008), p. 114045.

[21] F Acernese et al. "Advanced Virgo: a second-generation interferometric gravitational wave detector." In: *Classical and Quantum Gravity* 32.2 (2014), p. 024001.

[22] D. N. Aguilera, J. A. Pons, and J. A. Miralles. "2D Cooling of magnetized neutron stars." In: *Astron. Astrophys.* 486 (July 2008), pp. 255–271. DOI: 10.1051/0004-6361:20078786. arXiv: 0710.0854.

[23] M. Aharon, M. Elad, and A. Bruckstein. "*rmk*-SVD: An algorithm for designing overcomplete dictionaries for sparse representation." In: *IEEE Transactions on signal processing* 54.11 (2006), pp. 4311–4322.

[24] P. Ajith et al. "Instrumental vetoes for transient gravitational-wave triggers using noise-coupling models: The bilinear-coupling veto." In: *Phys. Rev. D* 89.12, 122001 (June 2014), p. 122001. DOI: 10.1103/PhysRevD.89.122001. arXiv: 1403.1431 [gr-qc].

[25] P. Ajith et al. "Instrumental vetoes for transient gravitational-wave triggers using noise-coupling models: The bilinear-coupling veto." In: *Physical Review D* 89.12 (2014), p. 122001.

[26] A Akmal, V. Pandharipande, and D. Ravenhall. "Equation of state of nucleon matter and neutron star structure." In: *Physical Review C* 58.3 (1998), p. 1804.

[27] M. Alcubierre. *Introduction to 3+ 1 numerical relativity*. Vol. 140. Oxford University Press, 2008.

[28] W. G. Anderson and R. Balasubramanian. "Time-frequency detection of gravitational waves." In: *Phys. Rev. D* 60.10, 102001 (Nov. 1999), p. 102001. DOI: 10.1103/PhysRevD.60.102001. eprint: gr-qc/9905023.

[29] H. Andresen et al. "Gravitational wave signals from 3D neutrino hydrodynamics simulations of core-collapse supernovae." In: *Mon. Not. R. Astron. Soc.* 468 (June 2017), pp. 2032–2051. DOI: 10.1093/mnras/stx618. arXiv: 1607.05199 [astro-ph.HE].

[30] A. M. Anile. *Relativistic fluids and magneto-fluids: With applications in astrophysics and plasma physics*. 1989.

[31] L. Antón et al. "Numerical 3+1 General Relativistic Magnetohydrodynamics: A Local Characteristic Approach." In: *Astrophys. J.* 637 (Jan. 2006), pp. 296–312. DOI: 10.1086/498238. eprint: astro-ph/0506063.

[32] J. Antoniadis et al. "A Massive Pulsar in a Compact Relativistic Binary." In: *Science* 340 (Apr. 2013), p. 448. DOI: 10.1126/science.1233232. arXiv: 1304.6875 [astro-ph.HE].

[33] J Areeda and J Smith. In: *URL https://wiki.ligo.org/DetChar/DataQuality/PreO1DQShifts#ER7* (2016).

[34] J. Arons and S. M. Lea. "Accretion onto magnetized neutron stars - Structure and interchange instability of a model magnetosphere." In: *Astrophys. J.* 207 (Aug. 1976), pp. 914–936. DOI: 10.1086/154562.

[35] Y. Aso et al. "Interferometer design of the KAGRA gravitational wave detector." In: *Phys. Rev. D* 88.4, 043007 (Aug. 2013), p. 043007. DOI: 10.1103/PhysRevD.88.043007. arXiv: 1306.6747 [gr-qc].

[36] J. G. Baker et al. "Gravitational-Wave Extraction from an Inspiraling Configuration of Merging Black Holes." In: *Physical Review Letters* 96.11, 111102 (Mar. 2006), p. 111102. DOI: 10.1103/PhysRevLett.96.111102. eprint: gr-qc/0511103.

[37] J. G. Baker et al. "Comparisons of binary black hole merger waveforms." In: *Classical and Quantum Gravity* 24 (June 2007), S25–S31. DOI: 10.1088/0264-9381/24/12/S03. eprint: gr-qc/0701016.

[38] R. Balasubramanian, B. S. Sathyaprakash, and S. Dhurandhar. "Gravitational waves from coalescing binaries: Detection strategies and Monte Carlo estimation of parameters." In: *Physical Review D* 53.6 (1996), p. 3033.

[39] F. Banyuls et al. "Numerical {3+ 1} general relativistic hydrodynamics: a local characteristic approach." In: *The Astrophysical Journal* 476.1 (1997), p. 221.

[40] T. W. Baumgarte and S. L. Shapiro. *Numerical relativity: solving Einstein's equations on the computer*. Cambridge University Press, 2010.

[41] M. Belkin and P. Niyogi. "Laplacian eigenmaps for dimensionality reduction and data representation." In: *Neural computation* 15.6 (2003), pp. 1373–1396.

[42] C. G. Bernal, D. Page, and W. H. Lee. "Hypercritical Accretion onto a Newborn Neutron Star and Magnetic Field Submergence." In: *Astrophys. J.* 770, 106 (June 2013), p. 106. DOI: 10.1088/0004-637X/770/2/106. arXiv: 1212.0464 [astro-ph.HE].

[43] C. G. Bernal, W. H. Lee, and D. Page. "Hypercritical accretion onto a magnetized neutron star surface: a numerical approach." In: *arXiv preprint arXiv:1006.3003* (2010).

[44] G. F. Bignami et al. "The magnetic field of an isolated neutron star from X-ray cyclotron absorption lines." In: *Nature* 423 (June 2003), pp. 725–727. DOI: 10.1038/nature01703.

[45] L. Blanchet, T. Damour, and B. R. Iyer. "Gravitational waves from inspiralling compact binaries: Energy loss and waveform to second-post-Newtonian order." In: *Phys. Rev. D* 51 (May 1995), pp. 5360–5386. DOI: 10.1103/PhysRevD.51.5360. eprint: gr-qc/9501029.

[46] L. Blanchet et al. "Gravitational Radiation from Inspiralling Compact Binaries Completed at the Third Post-Newtonian Order." In: *Physical Review Letters* 93.9, 091101 (Aug. 2004), p. 091101. DOI: 10.1103/PhysRevLett.93.091101. eprint: gr-qc/0406012.

[47] J. M. Blondin. "Hypercritical spherical accretion onto compact objects." In: *Astrophys. J.* 308 (Sept. 1986), pp. 755–764. DOI: 10.1086/164548.

[48] J. M. Blondin, A. Mezzacappa, and C. DeMarino. "Stability of Standing Accretion Shocks, with an Eye toward Core-Collapse Supernovae." In: *Astrophys. J.* 584 (Feb. 2003), pp. 971–980. DOI: 10.1086/345812. eprint: astro-ph/0210634.

[49] S. Bogdanov, C.-Y. Ng, and V. M. Kaspi. "Constraining the Evolutionary Fate of Central Compact Objects: "Old" Radio Pulsars in Supernova Remnants." In: *Astrophys. J. Lett.* 792, L36 (Sept. 2014), p. L36. DOI: 10.1088/2041-8205/792/2/L36. arXiv: 1408.2740 [astro-ph.HE].

[50] A. Bonanno, V. Urpin, and G. Belvedere. "Protoneutron star dynamos and pulsar magnetism." In: *Astron. Astrophys.* 440 (Sept. 2005), pp. 199–205. DOI: 10.1051/0004-6361:20042098. eprint: astro-ph/0504328.

[51] S. Bose, A. Pai, and S. Dhurandhar. "Detection of Gravitational Waves from Inspiraling, Compact Binaries Using a Network of Interferometric Detectors." In: *International Journal of Modern Physics D* 9 (2000), pp. 325–329. DOI: 10.1142/S0218271800000360. eprint: gr-qc/0002010.

[52] O. Bousquet and L. Bottou. "The tradeoffs of large scale learning." In: *Advances in neural information processing systems*. 2008, pp. 161–168.

[53] L. M. Bregman. "The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming." In: *USSR computational mathematics and mathematical physics* 7.3 (1967), pp. 200–217.

[54] A. Buades, B. Coll, and J.-M. Morel. "Non-local means denoising." In: *Image Processing On Line* 1 (2011), pp. 208–212.

[55] A. Buonanno and T. Damour. "Transition from inspiral to plunge in binary black hole coalescences." In: *Phys. Rev. D* 62.6, 064015 (Sept. 2000), p. 064015. DOI: 10.1103/PhysRevD.62.064015. eprint: gr-qc/0001013.

[56] A. Buonanno and Y. Chen. "Quantum noise in second generation, signal-recycled laser interferometric gravitational-wave detectors." In: *Physical Review D* 64.4 (2001), p. 042006.

[57] A. Burrows and J. M. Lattimer. "The birth of neutron stars." In: *Astrophys. J.* 307 (Aug. 1986), pp. 178–196. DOI: 10.1086/164405.

[58] G. Camelio et al. "Evolution of a proto-neutron star with a nuclear many-body equation of state: neutrino luminosity and gravitational wave frequencies." In: *ArXiv e-prints* (Apr. 2017). arXiv: 1704.01923 [astro-ph.HE].

[59] M. Campanelli et al. "Accurate Evolutions of Orbiting Black-Hole Binaries without Excision." In: *Physical Review Letters* 96.11, 111101 (Mar. 2006), p. 111101. DOI: 10.1103/PhysRevLett.96.111101. eprint: gr-qc/0511048.

[60] E. J. Candès, J. Romberg, and T. Tao. "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information." In: *IEEE Transactions on information theory* 52.2 (2006), pp. 489–509.

[61] J. L. Carter. "Dual methods for total variation-based image restoration." PhD thesis. University of California Los Angeles, 2001.

[62] P. Cerdá-Durán et al. "Gravitational Wave Signatures in Black Hole Forming Core Collapse." In: *Astrophys. J. Lett.* 779, L18 (Dec. 2013), p. L18. DOI: 10.1088/2041-8205/779/2/L18. arXiv: 1310.8290 [astro-ph.SR].

[63] A. Chambolle. "An algorithm for total variation minimization and applications." In: *Journal of Mathematical imaging and vision* 20.1 (2004), pp. 89–97.

[64] A. Chambolle. "Total variation minimization and a class of binary MRF models." In: *International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*. Springer. 2005, pp. 136–152.

[65] A. Chambolle and P.-L. Lions. "Image recovery via total variation minimization and related problems." In: *Numerische Mathematik* 76.2 (1997), pp. 167–188.

[66] A. Chambolle et al. "An introduction to total variation for image analysis." In: *Theoretical foundations and numerical methods for sparse recovery* 9.263-340 (2010), p. 227.

[67] T. Chan, A. Marquina, and P. Mulet. "High-order total variation-based image restoration." In: *SIAM Journal on Scientific Computing* 22.2 (2000), pp. 503–516.

[68] T. F. Chan and S. Esedoglu. "Aspects of total variation regularized L 1 function approximation." In: *SIAM Journal on Applied Mathematics* 65.5 (2005), pp. 1817–1837.

[69] T. F. Chan, G. H. Golub, and P. Mulet. "A nonlinear primal-dual method for total variation-based image restoration." In: *SIAM journal on scientific computing* 20.6 (1999), pp. 1964–1977.

[70] S. Chatterji et al. "Multiresolution techniques for the detection of gravitational-wave bursts." In: *Classical and Quantum Gravity* 21.20 (2004), S1809.

[71] S. S. Chen, D. L. Donoho, and M. A. Saunders. "Atomic Decomposition by Basis Pursuit." In: *SIAM Review* 43 (Jan. 2001), pp. 129–159. DOI: 10.1137/S003614450037906X.

[72] R. A. Chevalier. "Neutron star accretion in a supernova." In: *Astrophys. J.* 346 (Nov. 1989), pp. 847–859. DOI: 10.1086/168066.

[73] N. Christensen et al. “LIGO S6 detector characterization studies.” In: *Classical and Quantum Gravity* 27.19 (2010), p. 194010.

[74] R. Ciolfi et al. “Instability-driven evolution of poloidal magnetic fields in relativistic stars.” In: *The Astrophysical Journal Letters* 736.1 (2011), p. L6.

[75] S. A. Colgate. “Neutron-Star Formation, Thermonuclear Supernovae, and Heavy-Element Reimplosion.” In: *Astrophys. J.* 163 (Jan. 1971), p. 221. DOI: 10.1086/150760.

[76] I. Cordero-Carrión et al. “Improved constrained scheme for the Einstein equations: An approach to the uniqueness issue.” In: *Phys. Rev. D* 79.2, 024017 (Jan. 2009), p. 024017. DOI: 10.1103/PhysRevD.79.024017. arXiv: 0809.2325 [gr-qc].

[77] N. J. Cornish and T. B. Littenberg. “Bayeswave: Bayesian inference for gravitational wave bursts and instrument glitches.” In: *Classical and Quantum Gravity* 32.13 (2015), p. 135012.

[78] J Creighton et al. “LAL software documentation (2007).” In: *URL http://www.lsc-group. phys. uwm. edu/daswg/projects/lal. html* ().

[79] A. Cumming et al. “Design and development of the advanced LIGO monolithic fused silica suspension.” In: *Classical and Quantum Gravity* 29.3 (2012), p. 035003.

[80] E. Cuoco, G. Cella, and G. M. Guidi. “Whitening of non-stationary noise from gravitational wave detectors.” In: *Classical and Quantum Gravity* 21.5 (2004), S801.

[81] T. Damour and A. Vilenkin. “Gravitational radiation from cosmic (super)strings: Bursts, stochastic background, and observational windows.” In: *Phys. Rev. D* 71.6, 063510 (Mar. 2005), p. 063510. DOI: 10.1103/PhysRevD.71.063510. eprint: hep-th/0410222.

[82] K. Davidson and J. P. Ostriker. “Neutron-Star Accretion in a Stellar Wind: Model for a Pulsed X-Ray Source.” In: *Astrophys. J.* 179 (Jan. 1973), pp. 585–598. DOI: 10.1086/151897.

[83] A. De Luca et al. “XMM-Newton and VLT observations of the isolated neutron star 1E 1207.4-5209.” In: *Astron. Astrophys.* 418 (May 2004), pp. 625–637. DOI: 10.1051/0004-6361:20031781. eprint: astro-ph/0312646.

[84] P. B. Demorest et al. “A two-solar-mass neutron star measured using Shapiro delay.” In: *Nature* 467 (Oct. 2010), pp. 1081–1083. DOI: 10.1038/nature09466. arXiv: 1010.5788 [astro-ph.HE].

[85] A. Di Credico and LIGO Scientific Collaboration. “Gravitational wave burst vetoes in the LIGO S2 and S3 data analyses.” In: *Classical and Quantum Gravity* 22 (Sept. 2005), S1051–S1058. DOI: 10.1088/0264-9381/22/18/S19. eprint: gr-qc/0504106.

[86] H. Dimmelmeier, J. A. Font, and E. Müller. “Relativistic simulations of rotational core collapse II. Collapse dynamics and gravitational radiation.” In: *Astron. Astrophys.* 393 (Oct. 2002), pp. 523–542. DOI: 10.1051/0004-6361:20021053. eprint: astro-ph/0204289.

[87] H. Dimmelmeier et al. "Gravitational wave burst signal from core collapse of rotating stars." In: *Phys. Rev. D* 78.6, 064056 (Sept. 2008), p. 064056. DOI: 10.1103/PhysRevD.78.064056. arXiv: 0806.4953.

[88] H. Dimmelmeier, J. A. Font, and E. Müller. "Relativistic simulations of rotational core collapse I. Methods, initial models, and code tests." In: *Astronomy & Astrophysics* 388.3 (2002), pp. 917–935.

[89] H. Dimmelmeier et al. "Combining spectral and shock-capturing methods: A new numerical approach for 3D relativistic core collapse simulations." In: *Physical Review D* 71.6 (2005), p. 064023.

[90] D. L. Donoho. "De-noising by soft-thresholding." In: *IEEE transactions on information theory* 41.3 (1995), pp. 613–627.

[91] D. L. Donoho. "Compressed sensing." In: *IEEE Transactions on information theory* 52.4 (2006), pp. 1289–1306.

[92] D. L. Donoho and I. M. Johnstone. "Ideal spatial adaptation by wavelet shrinkage." In: *biometrika* (1994), pp. 425–455.

[93] F. Douchin and P. Haensel. "A unified equation of state of dense matter and neutron star structure." In: *Astron. Astrophys.* 380 (Dec. 2001), pp. 151–167. DOI: 10.1051/0004-6361:20011402. eprint: astro-ph/0111092.

[94] B. Efron et al. "Least angle regression." In: *The Annals of statistics* 32.2 (2004), pp. 407–499.

[95] M. Elad and M. Aharon. "Image denoising via sparse and redundant representations over learned dictionaries." In: *IEEE Transactions on Image processing* 15.12 (2006), pp. 3736–3745.

[96] E. Endeve et al. "Turbulent Magnetic Field Amplification from Spiral SASI Modes: Implications for Core-collapse Supernovae and Proto-neutron Star Magnetization." In: *Astrophys. J.* 751, 26 (May 2012), p. 26. DOI: 10.1088/0004-637X/751/1/26. arXiv: 1203.3108 [astro-ph.SR].

[97] W. J. Engels, R. Frey, and C. D. Ott. "Multivariate regression analysis of gravitational waves from rotating core collapse." In: *Phys. Rev. D* 90.12, 124026 (Dec. 2014), p. 124026. DOI: 10.1103/PhysRevD.90.124026. arXiv: 1406.1164 [gr-qc].

[98] C. M. Espinoza et al. "The Braking Index of PSR J1734-3333 and the Magnetar Population." In: *Astrophys. J. Lett.* 741, L13 (Nov. 2011), p. L13. DOI: 10.1088/2041-8205/741/1/L13. arXiv: 1109.2740 [astro-ph.HE].

[99] R. Essick et al. "Localization of short duration gravitational-wave transients with the early advanced LIGO and Virgo detectors." In: *The Astrophysical Journal* 800.2 (2015), p. 81.

[100] L. C. Evans. "Partial differential equations and Monge-Kantorovich mass transfer." In: *Current developments in mathematics* 1997.1 (1997), pp. 65–126.

[101] C.-A. Faucher-Giguère and V. M. Kaspi. "Birth and Evolution of Isolated Radio Pulsars." In: *Astrophys. J.* 643 (May 2006), pp. 332–355. DOI: 10.1086/501516. eprint: astro-ph/0512585.

[102] V. Ferrari et al. "Gravitational waves from rotating proto-neutron stars." In: *Classical and Quantum Gravity* 21 (Mar. 2004), S515–S519. DOI: 10.1088/0264-9381/21/5/019. eprint: astro-ph/0409578.

[103] E. E. Flanagan and S. A. Hughes. "Measuring gravitational waves from binary black hole coalescences. I. Signal to noise for inspiral, merger, and ringdown." In: *Physical Review D* 57.8 (1998), p. 4535.

[104] T. Foglizzo and M. Tagger. "Entropic-acoustic instability in shocked accretion flows." In: *Astron. Astrophys.* 363 (Nov. 2000), pp. 174–183. eprint: astro-ph/0009193.

[105] J. A. Font. "Numerical hydrodynamics and magnetohydrodynamics in general relativity." In: *Living Reviews in Relativity* 11.1 (2008), p. 7.

[106] J. Friedman, T. Hastie, and R. Tibshirani. *The elements of statistical learning*. Vol. 1. Springer series in statistics Springer, Berlin, 2001.

[107] J. Friedman, T. Hastie, and R. Tibshirani. *The elements of statistical learning*. Vol. 1. Springer series in statistics Springer, Berlin, 2001.

[108] J. L. Friedman and N. Stergioulas. *Rotating relativistic stars*. Cambridge University Press, 2013.

[109] C. L. Fryer and K. C. B. New. "Gravitational Waves from Gravitational Collapse." In: *Living Reviews in Relativity* 14, 1 (Dec. 2011), p. 1. DOI: 10.12942/lrr-2011-1.

[110] C. L. Fryer. "Fallback in stellar collapse." In: *New Astronomy Reviews* 50.7 (2006), pp. 492–495.

[111] J. Fuller et al. "Supernova seismology: gravitational wave signatures of rapidly rotating core collapse." In: *Mon. Not. R. Astron. Soc.* 450 (June 2015), pp. 414–427. DOI: 10.1093/mnras/stv698. arXiv: 1501.06951 [astro-ph.HE].

[112] M. Gabler et al. "Magneto-elastic oscillations of neutron stars: exploring different magnetic field configurations." In: *Mon. Not. R. Astron. Soc.* 430 (Apr. 2013), pp. 1811–1831. DOI: 10.1093/mnras/sts721. arXiv: 1208.6443 [astro-ph.SR].

[113] N. P. Galatsanos and A. K. Katsaggelos. "Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation." In: *IEEE Transactions on image processing* 1.3 (1992), pp. 322–336.

[114] U. Geppert, D. Page, and T. Zannias. "Submergence and re-diffusion of the neutron star magnetic field after the supernova." In: *Astron. Astrophys.* 345 (May 1999), pp. 847–854.

[115] P. Getreuer. "Rudin-Osher-Fatemi Total Variation Denoising using Split Bregman." In: *Image Processing On Line* 2 (2012), pp. 74–95. DOI: 10.5201/ipol.2012.g-tvd.

[116] B. Giacomazzo and L. Rezzolla. "The exact solution of the Riemann problem in relativistic magnetohydrodynamics." In: *Journal of Fluid Mechanics* 562 (2006), pp. 223–259.

[117] S. K. Godunov. "A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics." In: *Matematicheskii Sbornik* 89.3 (1959), pp. 271–306.

[118] T. Goldstein and S. Osher. "The split Bregman method for L1-regularized problems." In: *SIAM journal on imaging sciences* 2.2 (2009), pp. 323–343.

[119] E. V. Gotthelf and J. P. Halpern. "Precise Timing of the X-ray Pulsar 1E 1207.4-5209: A Steady Neutron Star Weakly Magnetized at Birth." In: *Astrophys. J. Lett.* 664 (July 2007), pp. L35–L38. DOI: 10.1086/520637. arXiv: 0704.2255.

[120] E. V. Gotthelf, J. P. Halpern, and J. Alford. "The Spin-down of PSR J0821-4300 and PSR J1210-5226: Confirmation of Central Compact Objects as Anti-magnetars." In: *Astrophys. J.* 765, 58 (Mar. 2013), p. 58. DOI: 10.1088/0004-637X/765/1/58. arXiv: 1301.2717 [astro-ph.HE].

[121] E. V. Gotthelf, J. P. Halpern, and F. D. Seward. "Discovery of a 105 ms X-Ray Pulsar in Kesteven 79: On the Nature of Compact Central Objects in Supernova Remnants." In: *Astrophys. J.* 627 (July 2005), pp. 390–396. DOI: 10.1086/430300. eprint: astro-ph/0503424.

[122] J.-O. Goussard, P. Haensel, and J. L. Zdunik. "Rapid differential rotation of protoneutron stars and constraints on radio pulsars periods." In: *Astron. Astrophys.* 330 (Feb. 1998), pp. 1005–1016. eprint: astro-ph/9711347.

[123] H. Grote, L. S. Collaboration, et al. "The GEO 600 status." In: *Classical and Quantum Gravity* 27.8 (2010), p. 084003.

[124] M. Gullón et al. "Population synthesis of isolated neutron stars with magneto-rotational evolution." In: *Mon. Not. R. Astron. Soc.* 443 (Sept. 2014), pp. 1891–1899. DOI: 10.1093/mnras/stu1253. arXiv: 1406.6794 [astro-ph.HE].

[125] J. P. Halpern and E. V. Gotthelf. "Spin-Down Measurement of PSR J1852+0040 in Kesteven 79: Central Compact Objects as Anti-Magnetars." In: *Astrophys. J.* 709 (Jan. 2010), pp. 436–446. DOI: 10.1088/0004-637X/709/1/436. arXiv: 0911.0093 [astro-ph.HE].

[126] J. P. Halpern et al. "X-Ray Timing of PSR J1852+0040 in Kesteven 79: Evidence of Neutron Stars Weakly Magnetized at Birth." In: *Astrophys. J.* 665 (Aug. 2007), pp. 1304–1310. DOI: 10.1086/519557. arXiv: 0705.0978.

[127] N. J. Hammer, H.-T. Janka, and E. Müller. "Three-dimensional Simulations of Mixing Instabilities in Supernova Explosions." In: *Astrophys. J.* 714 (May 2010), pp. 1371–1385. DOI: 10.1088/0004-637X/714/2/1371. arXiv: 0908.3474 [astro-ph.SR].

[128] G. Hammond et al. "Reducing the suspension thermal noise of advanced gravitational wave detectors." In: *Classical and Quantum Gravity* 29.12 (2012), p. 124009.

[129] P. C. Hansen. "Analysis of discrete ill-posed problems by means of the L-curve." In: *SIAM review* 34.4 (1992), pp. 561–580.

[130] P. C. Hansen, T. K. Jensen, and G. Rodriguez. "An adaptive pruning algorithm for the discrete L-curve criterion." In: *Journal of computational and applied mathematics* 198.2 (2007), pp. 483–492.

[131] G. M. Harry, L. S. Collaboration, et al. "Advanced LIGO: the next generation of gravitational wave detectors." In: *Classical and Quantum Gravity* 27.8 (2010), p. 084006.

[132] G. M. Harry et al. "Thermal noise in interferometric gravitational wave detectors due to dielectric optical coatings." In: *Classical and Quantum Gravity* 19.5 (2002), p. 897.

[133] I. S. Heng. "Rotating stellar core-collapse waveform decomposition: a principal component analysis approach." In: *Classical and Quantum Gravity* 26.10 (2009), p. 105005.

[134] W. C. G. Ho. "Superfluid effects on gauging core temperatures of neutron stars in low-mass X-ray binaries." In: *Mon. Not. R. Astron. Soc.* 418 (Nov. 2011), pp. L99–L103. DOI: 10.1111/j.1745-3933.2011.01152.x. arXiv: 1109.0018 [astro-ph.HE].

[135] J. C. Houck and R. A. Chevalier. "Steady spherical hypercritical accretion onto neutron stars." In: *Astrophys. J.* 376 (July 1991), pp. 234–244. DOI: 10.1086/170272.

[136] C. Y. Hui and W. Becker. "X-ray observations of RX J0822-4300 and Puppis-A." In: *Astron. Astrophys.* 454 (Aug. 2006), pp. 543–552. DOI: 10.1051/0004-6361:20053554. eprint: astro-ph/0508655.

[137] M. Irani and S. Peleg. "Motion analysis for image enhancement: Resolution, occlusion, and transparency." In: *Journal of Visual Communication and Image Representation* 4.4 (1993), pp. 324–335.

[138] J. A. Isenberg. "Waveless approximation theories of gravity." In: *International Journal of Modern Physics D* 17.02 (2008), pp. 265–273.

[139] J. D. Jackson. *Classical Electrodynamics*. 1962.

[140] H.-T. Janka et al. "Theory of core-collapse supernovae." In: *Phys. Rep.* 442 (Apr. 2007), pp. 38–74. DOI: 10.1016/j.physrep.2007.02.002. eprint: astro-ph/0612072.

[141] P. Jaranowski and A. Kròlak. "Gravitational-Wave Data Analysis. Formalism and Sample Applications: The Gaussian Case." In: *Living Reviews in Relativity* 8, 3 (Dec. 2005), p. 3. DOI: 10.12942/lrr-2005-3.

[142] P. Jaranowski and A. Królak. *Analysis of gravitational-wave data*. Vol. 29. Cambridge University Press, 2009.

[143] C. C. Joggerst, A. Almgren, and S. E. Woosley. "Three-dimensional Simulations of Rayleigh-Taylor Mixing in Core-collapse Supernovae with Castro." In: *Astrophys. J.* 723 (Nov. 2010), pp. 353–363. DOI: 10.1088/0004-637X/723/1/353. arXiv: 1009.3336 [astro-ph.HE].

[144] J. B. Kanner et al. "Leveraging waveform complexity for confident detection of gravitational waves." In: *Physical Review D* 93.2 (2016), p. 022002.

[145] K. Kifonidis et al. "Non-spherical core collapse supernovae. I. Neutrino-driven convection, Rayleigh-Taylor instabilities, and the formation and propagation of metal clumps." In: *Astron. Astrophys.* 408 (Sept. 2003), pp. 621–649. DOI: 10.1051/0004-6361:20030863. eprint: astro-ph/0302239.

[146] K. Kifonidis et al. "Non-spherical core collapse supernovae. II. The late-time evolution of globally anisotropic neutrino-driven explosions and their implications for SN 1987 A." In: *Astron. Astrophys.* 453 (July 2006), pp. 661–678. DOI: 10.1051/0004-6361:20054512. eprint: astro-ph/0511369.

[147] S. Klimenko et al. "Method for detection and reconstruction of gravitational wave transients with networks of advanced detectors." In: *Phys. Rev. D* 93.4, 042004 (Feb. 2016), p. 042004. DOI: 10.1103/PhysRevD.93.042004. arXiv: 1511.05999 [gr-qc].

[148] K. Kokkotas and B Schmidt. *Living Reviews in Relativity*. 1999.

[149] M. Kruskal and M. Schwarzschild. "Some Instabilities of a Completely Ionized Plasma." In: *Proceedings of the Royal Society of London Series A* 223 (May 1954), pp. 348–360. DOI: 10.1098/rspa.1954.0120.

[150] A. K. Kulkarni and M. M. Romanova. "Accretion to magnetized stars through the Rayleigh-Taylor instability: global 3D simulations." In: *Mon. Not. R. Astron. Soc.* 386 (May 2008), pp. 673–687. DOI: 10.1111/j.1365-2966.2008.13094.x. arXiv: 0802.1759.

[151] T. Kuroda, K. Kotake, and T. Takiwaki. "A New Gravitational-wave Signature from Standing Accretion Shock Instability in Supernovae." In: *Astrophys. J. Lett.* 829, L14 (Sept. 2016), p. L14. DOI: 10.3847/2041-8205/829/1/L14. arXiv: 1605.09215 [astro-ph.HE].

[152] F. K. Lamb, C. J. Pethick, and D. Pines. "A Model for Compact X-Ray Sources: Accretion by Rotating Magnetic Stars." In: *Astrophys. J.* 184 (Aug. 1973), pp. 271–290. DOI: 10.1086/152325.

[153] J. M. Lattimer. "The Nuclear Equation of State and Neutron Star Masses." In: *Annual Review of Nuclear and Particle Science* 62 (Nov. 2012), pp. 485–515. DOI: 10.1146/annurev-nucl-102711-095018. arXiv: 1305.3510 [nucl-th].

[154] J. M. Lattimer and M. Prakash. "Ultimate Energy Density of Observable Cold Baryonic Matter." In: *Physical Review Letters* 94.11, 111101 (Mar. 2005), p. 111101. DOI: 10.1103/PhysRevLett.94.111101. eprint: astro-ph/0411280.

[155] J. M. Lattimer and F. D. Swesty. "A generalized equation of state for hot, dense matter." In: *Nuclear Physics A* 535 (Dec. 1991), pp. 331–376. DOI: 10.1016/0375-9474(91)90452-C.

[156] R. J. LeVeque. *Numerical methods for conservation laws*. Springer Science & Business Media, 1992.

[157] A. Lichnerowicz. *Relativistic Hydrodynamics and Magnetohydrodynamics*. 1967.

[158] A. Lichnerowicz. "L'intégration des équations de la gravitation relativiste et le probleme des n corps." In: *J. Math. Pures Appl* 23.37 (1944), p. 4.

[159] *LIGO Open Science Center*. http://losc.ligo.org/.

[160] T. B. Littenberg et al. "Enabling high confidence detections of gravitational-wave bursts." In: *Phys. Rev. D* 94.4, 044050 (Aug. 2016), p. 044050. DOI: 10.1103/PhysRevD.94.044050. arXiv: 1511.08752 [gr-qc].

[161] J. Logue et al. "Inferring core-collapse supernova physics with gravitational waves." In: *Phys. Rev. D* 86.4, 044023 (Aug. 2012), p. 044023. DOI: 10.1103/PhysRevD.86.044023. arXiv: 1202.3256 [gr-qc].

[162] C. Louchet. "Modèles variationnels et bayésiens pour le débruitage d'images: de la variation totale vers les moyennes non-locales." PhD thesis. Université René Descartes-Paris V, 2008.

[163] R. Lynch et al. "An information-theoretic approach to the gravitational-wave burst detection problem." In: *ArXiv e-prints* (Nov. 2015). arXiv: 1511.05955 [gr-qc].

[164] R. Lynch et al. “An information-theoretic approach to the gravitational-wave burst detection problem.” In: *arXiv preprint arXiv:1511.05955* (2015).

[165] J. Mairal et al. “Online dictionary learning for sparse coding.” In: *Proceedings of the 26th annual international conference on machine learning*. ACM. 2009, pp. 689–696.

[166] S. G. Mallat and Z. Zhang. “Matching pursuits with time-frequency dictionaries.” In: *IEEE Transactions on signal processing* 41.12 (1993), pp. 3397–3415.

[167] M. W. Marcellin et al. “An overview of JPEG-2000.” In: *Data Compression Conference, 2000. Proceedings. DCC 2000*. IEEE. 2000, pp. 523–541.

[168] A. Marquina and S. Osher. “Explicit algorithms for a new time dependent model based on level set motion for nonlinear deblurring and noise removal.” In: *SIAM Journal on Scientific Computing* 22.2 (2000), pp. 387–405.

[169] A. Marquina and S. J. Osher. “Image super-resolution by TV-regularization and Bregman iteration.” In: *Journal of Scientific Computing* 37.3 (2008), pp. 367–382.

[170] A. Marquina and S. J. Osher. “Image super-resolution by TV-regularization and Bregman iteration.” In: *Journal of Scientific Computing* 37.3 (2008), pp. 367–382.

[171] J. M. Martí, J. M. Ibáñez, and J. A. Miralles. “Numerical relativistic hydrodynamics: local characteristic approach.” In: *Physical Review D* 43.12 (1991), p. 3794.

[172] J. M. Martí and E. Müller. “The analytical solution of the Riemann problem in relativistic hydrodynamics.” In: *Journal of Fluid Mechanics* 258 (1994), pp. 317–333.

[173] D. V. Martynov et al. “Sensitivity of the Advanced LIGO detectors at the beginning of gravitational wave astronomy.” In: *Phys. Rev. D* 93.11, 112004 (June 2016), p. 112004. DOI: 10.1103/PhysRevD.93.112004. arXiv: 1604.00439 [astro-ph.IM].

[174] S. Mereghetti et al. “Pulse Phase Variations of the X-Ray Spectral Features in the Radio-quiet Neutron Star 1E 1207-5209.” In: *Astrophys. J.* 581 (Dec. 2002), pp. 1280–1285. DOI: 10.1086/344435. eprint: astro-ph/0207296.

[175] V. Mewes et al. “Numerical relativity simulations of thick accretion disks around tilted Kerr black holes.” In: *Physical Review D* 93.6 (2016), p. 064055.

[176] Y. Meyer. *Oscillating patterns in image processing and nonlinear evolution equations: the fifteenth Dean Jacqueline B. Lewis memorial lectures*. Vol. 22. American Mathematical Soc., 2001.

[177] C. A. Micchelli, L. Shen, and Y. Xu. “Proximity algorithms for image models: denoising.” In: *Inverse Problems* 27.4, 045009 (Apr. 2011), p. 045009. DOI: 10.1088/0266-5611/27/4/045009.

[178] F. C. Michel. “Accretion of matter by condensed objects.” In: *Astrophysics and Space Science* 15.1 (1972), pp. 153–160.

[179] F. C. Michel. "Accretion magnetosphere stability. I - Adiabatic gas model." In: *Astrophys. J.* 214 (May 1977), pp. 261–267. DOI: 10.1086/155250.

[180] A. H. Mroué et al. "Catalog of 174 Binary Black Hole Simulations for Gravitational Wave Astronomy." In: *Physical Review Letters* 111.24, 241104 (Dec. 2013), p. 241104. DOI: 10.1103/PhysRevLett.111.241104. arXiv: 1304.6077 [gr-qc].

[181] D. Mukherjee, D. Bhattacharya, and A. Mignone. "MHD instabilities in accretion mounds - I. 2D axisymmetric simulations." In: *Mon. Not. R. Astron. Soc.* 430 (Apr. 2013), pp. 1976–1987. DOI: 10.1093/mnras/stt020. arXiv: 1212.3897 [astro-ph.HE].

[182] D. Mukherjee, D. Bhattacharya, and A. Mignone. "MHD instabilities in accretion mounds - II. 3D simulations." In: *Mon. Not. R. Astron. Soc.* 435 (Oct. 2013), pp. 718–727. DOI: 10.1093/mnras/stt1344. arXiv: 1307.5052 [astro-ph.HE].

[183] S Mukherjee, R Obaid, and B Matkarimov. "Classification of glitch waveforms in gravitational wave detector characterization." In: *Journal of Physics: Conference Series*. Vol. 243. 1. IOP Publishing. 2010, p. 012006.

[184] B. Müller, H.-T. Janka, and A. Marek. "A New Multi-dimensional General Relativistic Neutrino Hydrodynamics Code of Core-collapse Supernovae. III. Gravitational Wave Signals from Supernova Explosion Models." In: *Astrophys. J.* 766, 43 (Mar. 2013), p. 43. DOI: 10.1088/0004-637X/766/1/43. arXiv: 1210.6984 [astro-ph.SR].

[185] J. W. Murphy, C. D. Ott, and A. Burrows. "A Model for Gravitational Wave Emission from Neutrino-Driven Core-Collapse Supernovae." In: *Astrophys. J.* 707 (Dec. 2009), pp. 1173–1190. DOI: 10.1088/0004-637X/707/2/1173. arXiv: 0907.4762 [astro-ph.SR].

[186] A. Muslimov and D. Page. "Delayed switch-on of pulsars." In: *Astrophys. J. Lett.* 440 (Feb. 1995), pp. L77–L80. DOI: 10.1086/187765.

[187] A. Nagar and L. Rezzolla. "TOPICAL REVIEW: Gauge-invariant non-spherical metric perturbations of Schwarzschild black-hole spacetimes." In: *Classical and Quantum Gravity* 22 (Aug. 2005), R167–R192. DOI: 10.1088/0264-9381/22/16/R01. eprint: gr-qc/0502064.

[188] J. W. Negele and D. Vautherin. "Neutron star matter at sub-nuclear densities." In: *Nuclear Physics A* 207 (June 1973), pp. 298–320. DOI: 10.1016/0375-9474(73)90349-7.

[189] A. Y. Ng, M. I. Jordan, Y. Weiss, et al. "On spectral clustering: Analysis and an algorithm." In: *NIPS*. Vol. 14. 2. 2001, pp. 849–856.

[190] M. Obergaulinger and M. Aloy. "Protomagnetar and black hole formation in high-mass stars." In: *ArXiv e-prints* (Mar. 2017). arXiv: 1703.09893 [astro-ph.SR].

[191] M. Obergaulinger, H.-T. Janka, and M. A. Aloy. "Magnetic field amplification and magnetically supported explosions of collapsing, non-rotating stellar cores." In: *Mon. Not. R. Astron. Soc.* 445 (Dec. 2014), pp. 3169–3199. DOI: 10.1093/mnras/stu1969. arXiv: 1405.7466 [astro-ph.SR].

[192] E. O'Connor and C. D. Ott. "The Progenitor Dependence of the Pre-explosion Neutrino Emission in Core-collapse Supernovae." In: *Astrophys. J.* 762, 126 (Jan. 2013), p. 126. DOI: 10.1088/0004-637X/762/2/126. arXiv: 1207.1100 [astro-ph.HE].

[193] B. A. Olshausen and D. J. Field. "Sparse coding with an overcomplete basis set: A strategy employed by V1?" In: *Vision research* 37.23 (1997), pp. 3311–3325.

[194] S. Osher et al. "An iterative regularization method for total variation-based image restoration." In: *Multiscale Modeling & Simulation* 4.2 (2005), pp. 460–489.

[195] B. J. Owen and B. S. Sathyaprakash. "Matched filtering of gravitational waves from inspiraling compact binaries: Computational cost and template placement." In: *Phys. Rev. D* 60.2, 022002 (July 1999), p. 022002. DOI: 10.1103/PhysRevD.60.022002. eprint: gr-qc/9808076.

[196] D. Page et al. "Minimal Cooling of Neutron Stars: A New Paradigm." In: *Astrophys. J. Suppl. Ser.* 155 (Dec. 2004), pp. 623–650. DOI: 10.1086/424844. eprint: astro-ph/0403657.

[197] C Palomba et al. "Searches for continuous gravitational wave signals and stochastic backgrounds in LIGO and Virgo data." In: *arXiv preprint arXiv:1201.3176* (2012).

[198] V. R. Pandharipande and R. A. Smith. "Nuclear matter calculations with mean scalar fields." In: *Physics Letters B* 59 (Oct. 1975), pp. 15–18. DOI: 10.1016/0370-2693(75)90143-4.

[199] A. Passamonti et al. "Coupling of radial and nonradial oscillations of relativistic stars: Gauge-invariant formalism." In: *Phys. Rev. D* 71.2, 024022 (Jan. 2005), p. 024022. DOI: 10.1103/PhysRevD.71.024022. eprint: gr-qc/0407108.

[200] Y. C. Pati, R. Rezaiifar, and P. Krishnaprasad. "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition." In: *Signals, Systems and Computers, 1993. 1993 Conference Record of The Twenty-Seventh Asilomar Conference on*. IEEE. 1993, pp. 40–44.

[201] D. J. B. Payne and A. Melatos. "Burial of the polar magnetic field of an accreting neutron star - I. Self-consistent analytic and numerical equilibria." In: *Mon. Not. R. Astron. Soc.* 351 (June 2004), pp. 569–584. DOI: 10.1111/j.1365-2966.2004.07798.x. eprint: astro-ph/0403173.

[202] D. J. B. Payne and A. Melatos. "Burial of the polar magnetic field of an accreting neutron star - II. Hydromagnetic stability of axisymmetric equilibria." In: *Mon. Not. R. Astron. Soc.* 376 (Apr. 2007), pp. 609–624. DOI: 10.1111/j.1365-2966.2007.11451.x. eprint: astro-ph/0703203.

[203] F. Pedregosa et al. "Scikit-learn: Machine learning in Python." In: *Journal of Machine Learning Research* 12.Oct (2011), pp. 2825–2830.

[204] G. Peyré. "The numerical tours of signal processing." In: *Computing in Science & Engineering* 13.4 (2011), pp. 94–97.

[205] M. Pitkin et al. "Gravitational wave detection by interferometry (ground and space)." In: *Living Reviews in Relativity* 14.1 (2011), p. 5.

[206] T. Pock et al. “Fast and exact solution of total variation models on the GPU.” In: *Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. IEEE Computer Society Conference on*. IEEE. 2008, pp. 1–8.

[207] J. A. Pons and U. Geppert. “Magnetic field dissipation in neutron star crusts: from magnetars to isolated neutron stars.” In: *Astron. Astrophys.* 470 (July 2007), pp. 303–315. DOI: 10.1051/0004-6361:20077456. eprint: astro-ph/0703267.

[208] J. A. Pons, J. A. Miralles, and U. Geppert. “Magneto-thermal evolution of neutron stars.” In: *Astron. Astrophys.* 496 (Mar. 2009), pp. 207–216. DOI: 10.1051/0004-6361:200811229. arXiv: 0812.3018.

[209] J. A. Pons et al. “Evolution of Proto-Neutron Stars.” In: *Astrophys. J.* 513 (Mar. 1999), pp. 780–804. DOI: 10.1086/306889. eprint: astro-ph/9807040.

[210] J. A. Pons, J. M. Marti, and E. Müller. “The exact solution of the Riemann problem with non-zero tangential velocities in relativistic hydrodynamics.” In: *Journal of Fluid Mechanics* 422 (2000), pp. 125–139.

[211] J. Powell et al. “Classification methods for noise transients in advanced gravitational-wave detectors.” In: *Classical and Quantum Gravity* 32.21, 215012 (Nov. 2015), p. 215012. DOI: 10.1088/0264-9381/32/21/215012. arXiv: 1505.01299 [astro-ph.IM].

[212] J. Powell et al. “Inferring the core-collapse supernova explosion mechanism with gravitational waves.” In: *Physical Review D* 94.12 (2016), p. 123012.

[213] J. Powell et al. “Classification methods for noise transients in advanced gravitational-wave detectors II: performance tests on Advanced LIGO data.” In: *Class. Quant. Grav.* 34.3 (2017), p. 034002. DOI: 10.1088/1361-6382/34/3/034002. arXiv: 1609.06262 [astro-ph.IM].

[214] F. Pretorius. “Evolution of Binary Black-Hole Spacetimes.” In: *Physical Review Letters* 95.12, 121101 (Sept. 2005), p. 121101. DOI: 10.1103/PhysRevLett.95.121101. eprint: gr-qc/0507014.

[215] M. H. Protter and H. F. Weinberger. *Maximum principles in differential equations*. Springer Science & Business Media, 2012.

[216] A. Reisenegger and P. Goldreich. “A new class of g-modes in neutron stars.” In: *Astrophys. J.* 395 (Aug. 1992), pp. 240–249. DOI: 10.1086/171645.

[217] L. Rezzolla and O. Zanotti. *Relativistic hydrodynamics*. Oxford University Press, 2013.

[218] L. Rezzolla, O. Zanotti, and J. A. Pons. “An improved exact Riemann solver for multi-dimensional relativistic flows.” In: *Journal of Fluid Mechanics* 479 (2003), pp. 199–219.

[219] S. Richers et al. “Equation of State Effects on Gravitational Waves from Rotating Core Collapse.” In: *ArXiv e-prints* (Jan. 2017). arXiv: 1701.02752 [astro-ph.HE].

[220] F Robinet. *Omicron: an algorithm to detect and characterize transient noise in gravitational-wave detectors*. 2015.

[221] M. M. Romanova, A. K. Kulkarni, and R. V. E. Lovelace. "Unstable Disk Accretion onto Magnetized Stars: First Global Three-dimensional Magnetohydrodynamic Simulations." In: *Astrophys. J. Lett.* 673, L171 (Feb. 2008), p. L171. DOI: 10.1086/527298.

[222] R. Romero et al. "The exact solution of the Riemann problem in relativistic magnetohydrodynamics with tangential magnetic fields." In: *Journal of Fluid Mechanics* 544 (Dec. 2005), pp. 323–338. DOI: 10.1017/S0022112005006701. eprint: astro-ph/0506527.

[223] C. Röver et al. "Bayesian reconstruction of gravitational wave burst signals from simulations of rotating stellar core collapse and bounce." In: *Phys. Rev. D* 80.10, 102004 (Nov. 2009), p. 102004. DOI: 10.1103/PhysRevD.80.102004. arXiv: 0909.1093 [gr-qc].

[224] L. I. Rudin, S. Osher, and E. Fatemi. "Nonlinear total variation based noise removal algorithms." In: *Physica D Nonlinear Phenomena* 60 (Nov. 1992), pp. 259–268. DOI: 10.1016/0167-2789(92)90242-F.

[225] B. S. Sathyaprakash and B. F. Schutz. "Physics, Astrophysics and Cosmology with Gravitational Waves." In: *Living Reviews in Relativity* 12, 2 (Dec. 2009), p. 2. DOI: 10.12942/lrr-2009-2. arXiv: 0903.0338 [gr-qc].

[226] L. Scheck et al. "Multidimensional supernova simulations with approximative neutrino transport. I. Neutron star kicks and the anisotropy of neutrino-driven explosions in two spatial dimensions." In: *Astron. Astrophys.* 457 (Oct. 2006), pp. 963–986. DOI: 10.1051/0004-6361:20064855. eprint: astro-ph/0601302.

[227] I. R. Seitenzahl et al. "Proton-rich Nuclear Statistical Equilibrium." In: *Astrophys. J. Lett.* 685, L129 (Oct. 2008), p. L129. DOI: 10.1086/592501. arXiv: 0808.2033.

[228] F. D. Seward et al. "A Compact Central Object in the Supernova Remnant Kesteven 79." In: *Astrophys. J.* 584 (Feb. 2003), pp. 414–417. DOI: 10.1086/345600. eprint: astro-ph/0210496.

[229] N. Shabaltas and D. Lai. "The Hidden Magnetic Field of the Young Neutron Star in Kesteven 79." In: *Astrophys. J.* 748, 148 (Apr. 2012), p. 148. DOI: 10.1088/0004-637X/748/2/148. arXiv: 1110.3129 [astro-ph.HE].

[230] S. L. Shapiro and S. A. Teukolsky. *Black holes, white dwarfs, and neutron stars: The physics of compact objects*. 1983.

[231] H. Shen et al. "Relativistic Equation of State of Nuclear Matter for Supernova Explosion." In: *Progress of Theoretical Physics* 100 (Nov. 1998), pp. 1013–1031. DOI: 10.1143/PTP.100.1013. eprint: nucl-th/9806095.

[232] M. Shibata. *Numerical Relativity*. Vol. 1. World Scientific, 2015.

[233] T Sidery, A Passamonti, and N Andersson. "The dynamics of pulsar glitches: contrasting phenomenology with numerical evolutions." In: *Monthly Notices of the Royal Astronomical Society* 405.2 (2010), pp. 1061–1074.

[234] R. Simpson, K. R. Page, and D. De Roure. "Zooniverse: observing the world's largest citizen science platform." In: *Proceedings of the 23rd international conference on world wide web*. ACM. 2014, pp. 1049–1054.

[235] D. Sivia. “Data Analysis: A Bayesian Tutorial (Oxford Science Publications).” In: (1996).

[236] J Smith. *Glitch Classes*. https://dcc.ligo.org/LIGO-G1500642/public. 2016.

[237] J. R. Smith, L. S. Collaboration, et al. “The path to the enhanced and advanced LIGO gravitational-wave detectors.” In: *Classical and Quantum Gravity* 26.11 (2009), p. 114013.

[238] J. R. Smith et al. “A hierarchical method for vetoing noise transients in gravitational-wave detectors.” In: *Classical and Quantum Gravity* 28.23 (2011), p. 235005.

[239] H. Sotani and T. Takiwaki. “Gravitational wave asteroseismology with protoneutron stars.” In: *Phys. Rev. D* 94.4, 044043 (Aug. 2016), p. 044043. DOI: 10.1103/PhysRevD.94.044043. arXiv: 1608.01048 [astro-ph.HE].

[240] N. Stergioulas. “Rotating Stars in Relativity.” In: *Living Reviews in Relativity* 6, 3 (June 2003), p. 3. DOI: 10.12942/lrr-2003-3. eprint: gr-qc/0302034.

[241] P. N. Swarztrauber. “A direct Method for the Discrete Solution of Separable Elliptic Equations.” In: *SIAM Journal on Numerical Analysis* 11 (Dec. 1974), pp. 1136–1150. DOI: 10.1137/0711086.

[242] M. E. Taylor. *Partial differential equations. III, volume 117 of Applied Mathematical Sciences*. 1997.

[243] C. Thompson and R. C. Duncan. “Neutron star dynamos and the origins of pulsar magnetism.” In: *Astrophys. J.* 408 (May 1993), pp. 194–217. DOI: 10.1086/172580.

[244] K. S. Thorne and A. Campolattaro. “Non-Radial Pulsation of General-Relativistic Stellar Models. I. Analytic Analysis for L >= 2.” In: *Astrophys. J.* 149 (Sept. 1967), p. 591. DOI: 10.1086/149288.

[245] K. S. Thorne. “Nonradial pulsation of general-relativistic stellar models. IV. The weakfield limit.” In: *The Astrophysical Journal* 158 (1969), p. 997.

[246] K. S. Thorne, C. W. Misner, and J. A. Wheeler. *Gravitation*. Freeman, 2000.

[247] E. Thrane and M. Coughlin. “Detecting Gravitation-Wave Transients at 5 σ : A Hierarchical Approach.” In: *Physical Review Letters* 115.18, 181102 (Oct. 2015), p. 181102. DOI: 10.1103/PhysRevLett.115.181102. arXiv: 1507.00537 [astro-ph.IM].

[248] R. Tibshirani. “Regression shrinkage and selection via the lasso.” In: *Journal of the Royal Statistical Society. Series B (Methodological)* (1996), pp. 267–288.

[249] R. J. Tibshirani et al. *The solution path of the generalized lasso*. Stanford University, 2011.

[250] F. X. Timmes. “Integration of Nuclear Reaction Networks for Stellar Hydrodynamics.” In: *Astrophys. J. Suppl. Ser.* 124 (Sept. 1999), pp. 241–263. DOI: 10.1086/313257.

[251] F. X. Timmes and D. Arnett. “The Accuracy, Consistency, and Speed of Five Equations of State for Stellar Hydrodynamics.” In: *Astrophys. J. Suppl. Ser.* 125 (Nov. 1999), pp. 277–294. DOI: 10.1086/313271.

[252] F. X. Timmes and F. D. Swesty. "The Accuracy, Consistency, and Speed of an Electron-Positron Equation of State Based on Table Interpolation of the Helmholtz Free Energy." In: *Astrophys. J. Suppl. Ser.* 126 (Feb. 2000), pp. 501–516. DOI: 10.1086/313304.

[253] E. F. Toro. *Riemann solvers and numerical methods for fluid dynamics: a practical introduction*. Springer Science & Business Media, 2013.

[254] A. Torres et al. "Total-variation-based methods for gravitational wave denoising." In: *Phys. Rev. D* 90.8, 084029 (Oct. 2014), p. 084029. DOI: 10.1103/PhysRevD.90.084029. arXiv: 1409.7888 [gr-qc].

[255] M. Ugliano et al. "Progenitor-explosion Connection and Remnant Birth Masses for Neutrino-driven Supernovae of Iron-core Progenitors." In: *Astrophys. J.* 757, 69 (Sept. 2012), p. 69. DOI: 10.1088/0004-637X/757/1/69. arXiv: 1205.3657 [astro-ph.SR].

[256] C. Van Eysden and A Melatos. "Gravitational radiation from pulsar glitches." In: *Classical and Quantum Gravity* 25.22 (2008), p. 225020.

[257] J. Veitch et al. "Parameter estimation for compact binaries with ground-based gravitational-wave observations using the LALInference software library." In: *Physical Review D* 91.4 (2015), p. 042003.

[258] D. Viganò and J. A. Pons. "Central compact objects and the hidden magnetic field scenario." In: *Mon. Not. R. Astron. Soc.* 425 (Oct. 2012), pp. 2487–2492. DOI: 10.1111/j.1365-2966.2012.21679.x. arXiv: 1206.2014 [astro-ph.SR].

[259] D. Viganò et al. "Unifying the observational diversity of isolated neutron stars via magneto-thermal evolution models." In: *Mon. Not. R. Astron. Soc.* 434 (Sept. 2013), pp. 123–141. DOI: 10.1093/mnras/stt1008. arXiv: 1306.2156 [astro-ph.SR].

[260] C. R. Vogel and M. E. Oman. "Iterative methods for total variation denoising." In: *SIAM Journal on Scientific Computing* 17.1 (1996), pp. 227–238.

[261] C. R. Vogel and M. E. Oman. "Fast, robust total variation-based reconstruction of noisy, blurred images." In: *IEEE transactions on image processing* 7.6 (1998), pp. 813–824.

[262] Z. Wang et al. "Image quality assessment: from error visibility to structural similarity." In: *IEEE transactions on image processing* 13.4 (2004), pp. 600–612.

[263] J. R. Wilson, G. J. Mathews, and P. Marronetti. "Relativistic numerical model for close neutron-star binaries." In: *Physical Review D* 54.2 (1996), p. 1317.

[264] A. Wongwathanarat, E. Müller, and H.-T. Janka. "Three-dimensional simulations of core-collapse supernovae: from shock revival to shock breakout." In: *Astron. Astrophys.* 577, A48 (May 2015), A48. DOI: 10.1051/0004-6361/201425025. arXiv: 1409.5431 [astro-ph.HE].

[265] S. E. Woosley, A. Heger, and T. A. Weaver. "The evolution and explosion of massive stars." In: *Reviews of Modern Physics* 74 (Nov. 2002), pp. 1015–1071. DOI: 10.1103/RevModPhys.74.1015.

[266] S. E. Woosley and T. A. Weaver. "The Evolution and Explosion of Massive Stars. II. Explosive Hydrodynamics and Nucleosynthesis." In: *Astrophys. J. Suppl. Ser.* 101 (Nov. 1995), p. 181. DOI: 10.1086/192237.

[267] S. Woosley and A. Heger. "The progenitor stars of gamma-ray bursts." In: *The Astrophysical Journal* 637.2 (2006), p. 914.

[268] E. J. Young and G. Chanmugam. "Postaccretion magnetic field evolution of neutron stars." In: *Astrophys. J. Lett.* 442 (Apr. 1995), pp. L53–L56. DOI: 10.1086/187814.

[269] V. E. Zavlin et al. "Discovery of 424 Millisecond Pulsations from the Radio-quiet Neutron Star in the Supernova Remnant PKS 1209-51/52." In: *Astrophys. J. Lett.* 540 (Sept. 2000), pp. L25–L28. DOI: 10.1086/312866. eprint: astro-ph/0005548.

[270] M. Zevin et al. "Gravity Spy: Integrating Advanced LIGO Detector Characterization, Machine Learning, and Citizen Science." In: *arXiv preprint arXiv:1611.04596* (2016).

[271] W. Zhang, S. E. Woosley, and A. Heger. "Fallback and Black Hole Production in Massive Stars." In: *Astrophys. J.* 679, 639-654 (May 2008), pp. 639–654. DOI: 10.1086/526404. eprint: astro-ph/0701083.

[272] B. Zink, P. D. Lasky, and K. D. Kokkotas. "Are gravitational waves from giant magnetar flares observable?" In: *Physical Review D* 85.2 (2012), p. 024030.

[273] H. Zou and T. Hastie. "Regularization and variable selection via the elastic net." In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 67.2 (2005), pp. 301–320.