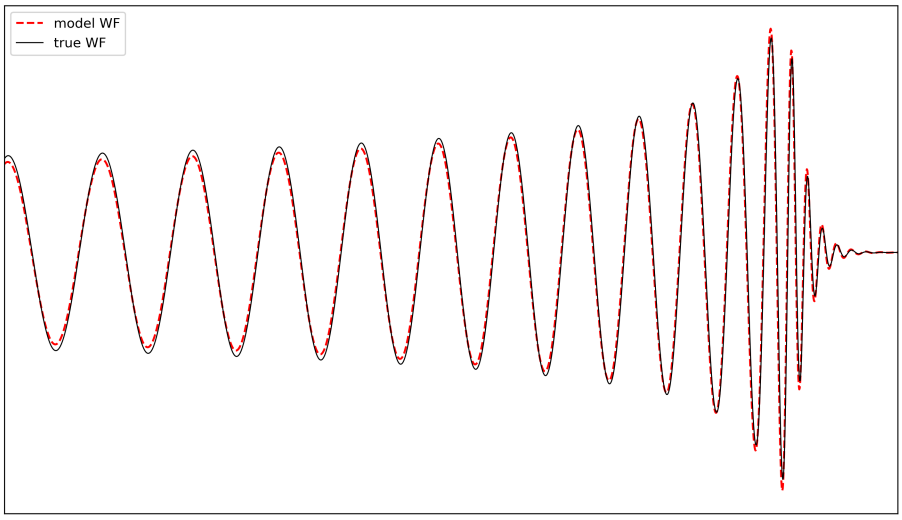
Deriving the Einstein Field Equations and predicting its Gravitational Wave Solutions using Neural Networks

Tim Grimbergen

a thesis submitted to the Department of Mathematics and Physics at Utrecht University in partial fulfillment of the requirements for the degrees of

Bachelor in Mathematics and Physics



supervisor mathematics: Fabian Ziltener
supervisors physics: Chris van den Broeck, Stefano Schmidt, Chinmay Kalaghatgi

14/06/2023

Universiteit Utrecht

Abstract

Nowadays, the most successful theory for describing gravity is Albert Einstein's general relativity. One of its predictions is the phenomenon of gravitational waves. In this work, we first discuss essential concepts from differential geometry to describe this theory. Furthermore, we derive the Einstein Field Equations (EFE) in vacuum from the Einstein-Hilbert action and we prove that small deviations from the flat spacetime metric that are solutions of the wave equation *almost* solve the EFE in vacuum.

Moreover, we develop `mlgw-NN`, which is a machine learning based method that utilizes neural networks (NN) that quickly and accurately predict the form of gravitational waves for the full inspiral of a non-precessing binary black hole system. We found `mlgw-NN` to produce state-of-the-art results for accuracy with a median mismatch of $\mathcal{F}_{\text{med}} = 7.00 \times 10^{-5}$ compared to the `IMRPhenomTPHM` method while having a speed-up factor of about 1.66 (on CPU). With further optimization in mind, this shows `mlgw-NN` to be a promising method for the fast and accurate generation of gravitational waveforms.

Acknowledgments

I want to thank my (4!) supervisors for their guidance. In particular, I want to thank Fabian Ziltener and Stefano Schmidt for dedicating their time to weekly discussions with me and on top of that for always being available for correspondence by e-mail.

Contents

1 Introduction

General relativity is one of the best tested theories in physics. The theory explains gravitational phenomena by means of a *Lorentzian metric*. The metric essentially provides the notion of distances and angles for every point in space, or more specifically *spacetime*, which combines the spacial dimensions with the temporal one. In general, the metric is not static and is influenced by matter. How the metric is altered by matter is governed by the Einstein Field Equations (EFE).

Under certain valid assumptions, the EFE produce gravitational wave solutions. A gravitational wave is a generally small perturbation of flat spacetime propagating through spacetime as a wave. Over the last decade there have been various detections of gravitational waves that are created by the inspiral of very heavy objects such as black holes and neutron stars millions of light years away from the earth.

This thesis is a combination of more mathematically styled chapters 2 and 3 and more physics-styled chapters 4,5,6 and 7. For a more complete overview of each chapter, see the "Chapter Overview" below.

In the mathematical chapters, we first introduce important concepts used in the general relativity in the setting of differential geometry. In chapter 2 for instance, we define pseudo-Riemannian manifolds and thereby a more concrete definition of the (pseudo-)metric g. In chapter 3, we provide a derivation of the EFE from the Einstein-Hilbert action. We also prove proposition 3.2, which essentially states that for small deviations from the flat spacetime metric the wave-equation *almost* solves the EFE in vacuum. This justifies the approximation schemes in physics for describing gravitational wave solutions of the metric.

In the physics chapters, we predict gravitational wave solutions for the Einstein field equations in the important case of binary inspiraling black holes by using neural networks (NN) and other machine learning methods. We restrict the spins of the black holes to be aligned with each other and the orbital plane in which the inspiral happens.

Our approach expands upon `mlgw` [12], which has the same goal but does not employ NNs. Hence, we call our approach `mlgw-NN`. The approach provides a way for creating *surrogate* models for already existing models that produce gravitational waves. Surrogate models approximate the behavior of more complex models. This way, we can possibly generate gravitational waves with a speed-up while still preserving accuracy. The `mlgw-NN` approach is explained in more detail in chapter 5.

In chapter 6, we found the `mlgw-NN` approach to be successful. We were able to make predictions for gravitational waves with a median mismatch of $\mathcal{F}_{\text{med}} = 7.00 \times 10^{-5}$ with `IMRPhenomTPHM` waveforms (WF). Mismatch is the metric used for determining the accuracy of the model. These results seem to be in line with state-of-the-art surrogate gravitational wave models (see Fig. 2, [8]). We were able to obtain a mild speed-up of a factor of 1.66, though for reasons discussed in chapter 7 this might be better than it seems.

The plane in which the inspiral of the two black holes happens could itself be changing

its orientation with respect to the detectors on earth. Furthermore, the spins of the black holes are not necessarily aligned with the orbital plane. The changes in orientations of the rotating bodies is called precession. It is crucial that future developments of `mlgw-NN` focus on implementing this phenomenon as well.

Chapter Overview

Chapter 2: Differential Geometry

We first cover some essential theory from differential geometry. Most of the theory is based on Loring Tu's "Differential Geometry: Connections Curvature and Characteristic Classes" [14], in particular chapters 1, 6, 7, 10, 18 and 21.

The central object of general relativity, for instance, is spacetime, which combines the spacial dimensions together with the temporal dimension. Mathematically, spacetime is modeled by a 4-dimensional pseudo-Riemannian manifold. In chapter 2, we explain that a pseudo-Riemannian manifold is a manifold M with a pseudo-metric g. A pseudo-metric essentially allows for the notion of distance at every point $x \in M$.

More formally, we will see that the metric is described by a $(0, 2)$ tensor field on the tangent bundle TM. The tangent bundle is a special case of the more general concept of vector bundles. Essentially, one can think of a vector bundle as a manifold with a vector space associated to every point. In 2.1, a formal definition is given. A $(0, 2)$ tensor field on the tangent bundle associated to every point $x \in M$ an element (tensor) of the tensor product $T_xM \otimes T_xM$. The tensor product allows multilinear maps to be mapped to linear maps. We discuss tensors and tensor fields more in depth in 2.2.

Other concepts that we discuss in chapter 2 are connections on vector bundles and the curvature of the connection. The ordinary derivative fails for functions on vector bundles. To this extent, a new notion of the derivative should be defined that "connects" the various vector spaces corresponding to different points, therefore it is called the connection. In proposition 2.9, we show there exists a canonical connection corresponding to a pseudo-Riemannian metric g called the Levi-Civita connection. The two defining properties of the Levi-Civita connection is that its torsion-free and that the covariant derivative of g vanishes. In general relativity, the tensors describing curvature are constructed from the Levi-Civita connection.

Chapter 3: Einstein Field Equations

In chapter 3, we provide a derivation of the Einstein Field Equations (EFE) in vacuum

$$\mathrm{Ric} - \tfrac{1}{2} g S = 0, \tag{1.1}$$

by varying the Einstein-Hilbert action

$$F : \mathcal{M} \to \mathbb{R}, \quad F(g) = -\int_M S\mu,$$

where $\mathcal{M}$ is the space of Lorentzian metrics. In equation (1.1), g is the metric, Ric is the Ricci tensor and S is the Ricci scalar, which is the contraction of the Ricci tensor. The Ricci tensor contains (partial) derivatives of the metric, so the EFE is a partial differential equation for the metric g. Our derivation is inspired by the one in chapter 8.4 of Shlomo Sternberg's "Semi-Riemannian Geometry and General Relativity" [13].

We then prove the proposition below

Proposition 3.2 [Small metric perturbations] *Let $\epsilon > 0$ and $K \subset \mathbb{R}^n$ a compact subset of $\mathbb{R}^n$. Then there exists an open (C^2, K)-neighborhood U of the Minkowski metric η in the space $\mathcal{M}$ of pseudo-metrics with $(1, n-1)$ signature, such that for all $g \in U$ we have*

$$||\Box(g-\eta)||_{C^0,K} \leq \epsilon ||g-\eta||_{C^2,K}$$

This proposition justifies the approximation methods for solving the EFE in the following chapters. In 3.2, the notation and other necessary concepts are explained. We also reflect on the assumptions on which the proposition is based and its applications to physics.

Chapter 4: Gravitational Wave Solutions of Einstein Field Equations

Exact solutions of the EFE equation are difficult to find. This is especially the case when the energy momentum tensor, which describes matter, enters the right hand side of the EFE. This leads to approximate theories for gravity, where for instance we assume that the metric is a weakly perturbed flat metric and/or that the matter is slowly moving. These methods, and the phenomenon of gravitational waves that follow from them, are reviewed in chapter 4. In linearized gravity only terms linear in (derivatives of) the metric perturbation $h_{\mu\nu} := g_{\mu\nu} - \eta_{\mu\nu}$ are considered where η is the flat Minkowski metric. We will show that the EFE reduce to

$$\Box \bar{h}_{\mu\nu} = -\frac{16\pi G}{c^4} T_{\mu\nu},$$

with $\Box$ the d'Alembert (or wave) operator, $\bar{h}_{\mu\nu}$ the trace reversed metric and $T_{\mu\nu}$ the energy-momentum tensor.

Chapter 5: Machine Learning Methods

In chapter 5, we provide a fast method for generating accurate waveforms (WF) by using machine learning. We first cover the appropriate data-format of a gravitational wave $h(t)$, namely $h(t) = A(t)e^{i\phi(t)}$ with amplitude $A(t)$ and phase $\phi(t)$ calculated over a discrete time grid t with D points. We refer to the *vectors* $h(t)$ as waveform. The goal is to predict the D-dimensional vectors $A_{\boldsymbol{\theta}}(\mathbf{t})$ and $\phi_{\boldsymbol{\theta}}(\mathbf{t})$ corresponding to parameter inputs $\boldsymbol{\theta}$ such as masses and spins of the black holes.

Since D is large, we represent the vectors $A_{\boldsymbol{\theta}}(\mathbf{t})$ and $\phi_{\boldsymbol{\theta}}(\mathbf{t})$ in a lower dimensional space of dimension K using Principal Component Analysis (PCA).

After the dimensionality reduction, we train Neural Networks (NN) to predict the functions $\boldsymbol{\theta} \mapsto A_{\boldsymbol{\theta}}(\mathbf{t})$ and $\boldsymbol{\theta} \mapsto \phi_{\boldsymbol{\theta}}(\mathbf{t})$.

Chapter 6: Model Results

The accuracy of the method given in chapter 5 is discussed in chapter 6. Furthermore, we evaluate the speed with which our waveforms are generated by comparing this to the speed with which `IMRPhenomTPHM` WFs are generated, which is the method used for generating the dataset on which our models are trained.

In chapter 6, we evaluate the performance of a model created using the `mlgw-NN` method. The models were trained on a dataset consisting of WFs generated by the `IMRPhenomTPHM` method [3],[4], which is based on Post-Newtonian approximation theory.

In particular, we will evaluate the accuracy of the model compared to `IMRPhenomTPHM` WFs and the speed-up of our WF generation time compared to that of `IMRPhenomTPHM`.

2 Differential Geometry

The goal of this chapter is to provide an overview of some important concepts in general relativity. We will do so in a way that emphasizes the mathematically geometric structures of the theory. The base manifold of spacetime, for instance, is a pseudo-Riemmanian manifold, or even more specific a Lorentzian manifold. These objects consists of a smooth manifold M together with a map g, called the pseudo-metric, that associates to all points $x \in M$ a symmetric, bilinear, non-degenerate and smoothly varying map $g_x : T_xM \times T_xM \to \mathbb{R}$. This map almost yields an inner product on the tangent space at the point x, as g_x is not necessarily positive definite. This means that we have a notion of distance and angles on M, which is crucial for doing physics. Definitions and notation in this chapter are mainly based on *Differential Geometry Connections, Curvature, and Characteristic Classes* by Loring Tu [14].

2.1 Vector bundles

The idea of vector bundles is that a (topological) space X has a vector space associated to it at every point $x \in X$ in such a way that for an open subset $U \subset X$ the vector bundle "looks like" $U \times \mathbb{R}^n$. Intuitively the vector spaces have to vary in a continuous manner over the space X. In general relativity, the topological space X is a smooth manifold M and the vector spaces are real, so we will introduce vector bundles in this setting. The dimensions of the manifold M and the vector spaces are n and k respectively. The construction of vector bundles is important as for instance vector fields are described as functions on the tangent bundle, a special case of a vector bundle. Furthermore, the curvature tensor fields in general relativity are formed from the connection, which is a type of object that can be defined for every vector bundle. A definition is given below.

Definition 2.1 (Smooth real vector bundle)**.** Let E (total space) and M (base space) be smooth manifolds and let $\pi : E \to M$ be a smooth and surjective map. A tuple (E, M, π) is called a smooth real vector bundle if:

1. for all $x \in M$ we have that $E_x = \pi^{-1}(\{x\})$, called the fiber of x, is a finite dimensional real vector space,

2. for all $x \in M$ there exists an open neighborhood $U \subset M$ of x and a diffeomorphism

$$\phi : \pi^{-1}(U) \to U \times \mathbb{R}^k,$$

so that for all $x \in U$ we have that ϕ restricts to a linear isomorphism $E_x \to \{x\} \times \mathbb{R}^k$.

We assume that M consists of one connected component, so that the dimension of the vector spaces E_x for all $x \in M$ have the dimension $k \in \mathbb{N}$. This is also called the **rank** of the vector bundle. Whenever M and π are clear from the context, we also denote the vector bundle by E.

From now on we will always assume vector bundles to be smooth real vector bundles. We will now look at a special type of functions on the vector bundle E, namely sections.

Definition 2.2 (sections, space of sections)**.** A smooth map $s : M \to E$ such that $\pi \circ s = \mathrm{id}_M$ is called a section. We denote by $\Gamma(E)$ the space of all sections on the vector bundle E. The space $\Gamma(E)$ is a module over the smooth functions on M, because for $f \in C^\infty(M, \mathbb{R})$ we have that $(fs)(x) = f(x)s(x) \in E_x$, so fs is also a section.

A natural example of the above definitions is given by the tangent bundle $TM := (TM, M, \pi)$.

Example 2.3 (tangent bundle and vector fields)**.** We denote by TM the space of all points $x \in M$ together with all tangent vectors of M so

$$TM := \{(x, v) \,|\, x \in M, v \in T_xM\}.$$

There exists a natural projection $\pi : TM \to M$, $(x, v) \mapsto x$. For a vector bundle, it is necessary that TM itself is a smooth manifold, but what is its smooth structure? Because M is a smooth manifold, we have an atlas $\mathcal{A}$ which is a collection of pairs (U, ϕ), also called charts, where U is a (topologically) open subset of M and ϕ is a diffeomorphism $U \to \mathbb{R}^n$. We briefly sketch how the smooth structure on TM arises from this atlas $\mathcal{A}$. For a complete construction with proofs, see Chapter 5 from Curtis's "Differentiable Manifolds and Theoretical Physics" [2]. Take a chart (U, ϕ). For $x \in U$, we have the linear map

$$(d\phi)_x : T_xM \to \mathbb{R}^n.$$

Then define

$$\phi_{TU} : \pi^{-1}(U) \to \phi(U) \times \mathbb{R}^n, \quad \text{such that} \quad \phi_{TU}((x, v)) = (\phi(x), (d\phi)_x(v)).$$

Then ϕ_{TU} is bijective. We then define the atlas

$$\mathcal{A}_{TM} := \{(\pi^{-1}(U), \phi_{TU}) : (U, \phi) \in \mathcal{A}\},$$

which yields the smooth structure on TM. The triplet (TM, M, π) is the tangent bundle over M, which we denote by TM. A section of TM is also called a vector field. We denote the space of vector fields as $\mathfrak{X}(M)$ and its elements by $X, Y, \ldots$.

Sections essentially attribute to every point $x \in M$ an element of the fiber E_x. One can then wonder if there exists a natural notion for the derivative of such sections. The natural notion of a derivative for a map $f : \mathbb{R}^n \to \mathbb{R}^m$ in the point x and direction v is

$$(df)_x(v) = \lim_{h \to 0} \frac{f(x + hv) - f(x)}{h}.$$

When $f : M \to E$ is instead a section, there are two problems that arise. The first one is that $x + hv$ will not necessarily be an element of M. An easy fix for this problem is

to instead take a path $\gamma : (-1,1) \to M$ so that $\gamma(0) = x$ and $\gamma'(0) = v$ so that the derivative can be defined as

$$(df)_x(v) = \lim_{t\to 0} \frac{f(\gamma(t)) - f(x)}{t}.$$

This then leads to the more subtle problem that the the two vectors $f(\gamma(t))$ and $f(x)$ are elements of the *different* vector space $E_{\gamma(t)}$ and E_x, so the operation of subtraction is not defined. To take derivatives of sections, we need to "connect" the different vector spaces with each other. This leads to the notion of the connection on a vector bundle.

Definition 2.4 (connections)**.** Let (E, M, π) be a vector bundle. A connection is a map

$$\nabla : \mathfrak{X}(M) \times \Gamma(E) \to \Gamma(E),$$

such that for all $X \in \mathfrak{X}(M)$, $s \in \Gamma(E)$ and smooth functions $f \in C^\infty(M, \mathbb{R})$ we have that

1. $\nabla_{fX} s = f \nabla_X s$;
2. $\nabla_X(fs) = (Xf)s + f\nabla_X s$ (Leibniz rule).

One can prove that such a connection always exists on any vector bundle by means of a partition of unity argument. However, it is interesting to note that there does not exist a unique, or even canonical, connection on any given vector bundle. In section 2.3 we will discuss the important case in which there does exist a canonical choice for the connection, namely the Levi-Civita connection.

We finish this paragraph by introducing two important objects that are defined in terms of the connection and that are essential for describing curvature. The *curvature of the connection* is given by the map

$$\begin{aligned} R : \mathfrak{X}(M) \times \mathfrak{X}(M) \times \Gamma(E) &\to \Gamma(E) \\ R(X,Y)s &= \nabla_X \nabla_Y s - \nabla_Y \nabla_X s - \nabla_{[X,Y]} s, \end{aligned} \tag{2.1}$$

where $[X, Y]$ denotes the Lie-bracket of the vector fields X and Y. Finally, the *torsion of the connection* is given by the map

$$\mathrm{Tor} : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M), \quad \mathrm{Tor}(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y]. \tag{2.2}$$

2.2 Tensor fields

In this paragraph, we start by introducing tensors from a mathematical viewpoint. The construction of the tensor product on vector spaces allows us to convert *multilinear* maps (such as the inner product or the determinant) to *linear* maps. We then give a characterization of tensors as F-linear maps, and we use this to show that the curvature and torsion of the connection (equations 2.2 and 2.1) are in fact tensor fields. We start with the definition of a tensor field.

Definition 2.5 (tensor field)**.** Given a tangent bundle TM, a (p, q) tensor field is a section of the vector bundle

$$T^{(p,q)}M := \underbrace{TM \otimes \cdots \otimes TM}_{p} \otimes \underbrace{T^*M \otimes \cdots \otimes T^*M}_{q},$$

where the symbol "$\otimes$" denotes the tensor product. We denote by $\Gamma(T^{(p,q)}M)$ the space of all (p, q) tensor fields on M.

Of course, now we need to define what a tensor product is. One way to do this, is by the following universal property. The tensor product of two vector spaces V and W, denoted by $V \otimes W$, is a vector space together with a bilinear map

$$\otimes : V \times W \to V \otimes W, \quad (v, w) \mapsto v \otimes w,$$

such that for every bilinear map $b : V \times W \to U$, there is a unique linear map $\tilde{b} : V \otimes W \to U$ such that $b = \tilde{b} \circ \otimes$.

There does not exists a *unique* tensor product of two vector spaces, in the sense that there are many possible explicit constructions of $V \otimes W$ together with a bilinear map satisfying the property above. It is instructive to give an explicit construction for the tensor product $V \otimes W$ of two vector spaces over a field $\mathbb{F}$. We denote by L the space which has as a basis pairs of functions $\delta_{(v,w)}$, where $\delta_{(v_1,w_1)}(v_2, w_2)$ is 1 if $(v_1, w_1) = (v_2, w_2)$ and 0 if $(v_1, w_1) \neq (v_2, w_2)$. Elements of L are then finite linear combinations of $\delta_{(v,w)}$ for $(v, w) \in V \times W$. Note that this space is a huge (infinite dimensional) space. For convenience, we identify $\delta_{(v,w)}$ with (v, w). We then define the set S spanned by all elements of the form

$$\begin{gathered}(v_1 + v_2, w) - (v_1, w) - (v_2, w), \quad (v, w_1 + w_2) - (v, w_1) - (v, w_2) \\ (av, w) - a(v, w), \quad (v, aw) - a(v, w),\end{gathered}$$

with $v, v_1, v_2 \in V$, $w, w_1, w_2 \in W$ and $a \in \mathbb{F}$. Now the tensor product $V \otimes W$ is defined as the quotient L/R. Furthermore, we denote by $\otimes$ the projection from L to L/S mapping $L \ni (v, w) \mapsto v \otimes w \in V \otimes W$. Notice that by construction the map $\otimes$ is bilinear, because

$$\begin{aligned}(v_1 + v_2) \otimes w &= v_1 \otimes w + v_2 \otimes w, \\ v \otimes (w_1 + w_2) &= v \otimes w_1 + v \otimes w_2, \\ (av) \otimes w &= a(v \otimes w) = v \otimes (aw).\end{aligned}$$

It can be shown that this construction indeed satisfies the universal property for the tensor product. Both the universal property and the construction above naturally extend to tensor products of any finite number of vector spaces, as is used in definition 2.5. Strictly speaking, in definition 2.5 we are taking the tensor product of vector bundles, which is defined point-wise for $x \in M$ by taking the tensor product of all fibers at the point x.

Denote by $\mathcal{F}$ the space of all $C^\infty(M, \mathbb{R})$ functions and by $\Omega^1(M)$ the space of covector fields (sections of the cotangent bundle T^*M). A tensor field $T \in \Gamma(T^{(p,q)}M)$ then defines an $\mathcal{F}$-multilinear function

$$\begin{aligned} \Omega^1(M)^p \times \mathfrak{X}(M)^q &\to \mathcal{F} \\ (\omega_1, \ldots, \omega_p, X_1, \ldots, X_q) &\mapsto T(\omega_1, \ldots, \omega_p, X_1, \ldots, X_q). \end{aligned}$$

We can use this fact together with the proposition below to identify $\mathcal{F}$-multilinear maps with tensors.

Proposition 2.6 (Tensors and $\mathcal{F}$-multilinear maps)**.** *There is a one-to-one correspondence between*

$$\{F\text{-multilinear functions } T : \Omega^1(M)^p \times \mathfrak{X}(M)^q \to F\}$$

and

$$\{\text{tensor fields } \tilde{T} \in \Gamma(T^{(p,q)}M)\}$$

such that for $x \in M$ we have

$$T(\omega_1, \ldots, \omega_p, X_1, \ldots, X_q)(x) = \tilde{T}_x(\omega_1(x), \ldots, \omega_p(x), X_1(x), \ldots, X_q(x)).$$

Proof. See Tu, Proposition 21.11 [14]. □

Using proposition 2.6, we will now argue that the torsion Tor and the curvature R for sections of the tangent bundle TM can be identified with tensor fields.

Example 2.7 (Torsion and curvature as tensor fields)**.** We start by showing that Tor is $\mathcal{F}$-multilinear. The summation part of multilinearity is trivial, so assume $f, g \in \mathcal{F}$, then we wish to show that

$$\mathrm{Tor}(fX, gY) = \nabla_{fX}(gY) - \nabla_{gY}(fX) - [fX, gY] = fg\mathrm{Tor}(X, Y).$$

First of all, for the Lie Bracket we have

$$[X, gY] = X(g)Y + g[X, Y],$$

and the Lie bracket is skew-symmetric so that

$$[fX, gY] = fX(g)Y - gY(f)X + fg[X, Y].$$

For the connection, we have by definition 2.4

$$\nabla_{fX}(gY) = f\nabla_X(gY) = f(X(g)Y + g\nabla_X Y) = fX(g)Y + fg\nabla_X Y,$$

and similarly

$$\nabla_{gY}(fX) = gY(f)X + fg\nabla_Y X,$$

so putting everything together we find

$$\begin{aligned}\mathrm{Tor}(fX, gY) &= fX(g)Y + fg\nabla_X Y - gY(f)X - fg\nabla_Y X - fX(g)Y + gY(f)X - fg[X,Y] \\ &= fg\nabla_X Y - fg\nabla_Y X - fg[X,Y] = fg\mathrm{Tor}(X,Y),\end{aligned}$$

so that Tor is indeed $\mathcal{F}$-multilinear and by proposition 2.6, there exists a tensor field $\widetilde{\mathrm{Tor}}$.

To show the Riemann curvature tensor defined as in 2.1 for sections on the tangent bundle TM, we need to show that

$$R(fX, gY, hZ) = \nabla_{fX}\nabla_{gY}hZ - \nabla_{gY}\nabla_{fX}hZ - \nabla_{[fX,gY]}hZ = fghR(X,Y,Z),$$

for all $X, Y, Z \in \mathfrak{X}(M)$ and $f, g, h \in \mathcal{F}$. The derivation is similar to that of the torsion because we make use of exactly the same properties of the Lie bracket and the connection. Since the derivation is quite lengthy we omit it here.

Finally, for a (p, q) tensor field T we can define the covariant derivative of T with respect to a vector field Y as

$$\begin{aligned}\nabla_Y(T(\omega_1, \ldots, \omega_p, X_1, \ldots, X_q)) = &(\nabla_Y T)(\omega_1, \ldots, \omega_p, X_1, \ldots, X_q) \\ &+ \sum_{i=1}^{p} T(\omega_1, \ldots, \nabla_Z\omega_i, \ldots, \omega_p, X_1, \ldots, X_q) \\ &+ \sum_{i=1}^{q} T(\omega_1, \ldots, \omega_p, X_1, \ldots, \nabla_Z X_i, \ldots, X_q).\end{aligned} \tag{2.3}$$

The covariant derivative of a smooth function is defined as the usual derivative

$$\nabla_Y f = Y(f).$$

2.3 Pseudo-Riemannian manifolds

A smooth manifold M can be given an additional structure called the pseudo-metric g. For convenience, we will abbreviate pseudo-metric to just metric. We start by giving the definition.

Definition 2.8 (Pseudo-Riemannian manifold)**.** A pseudo-Riemannian manifold M is a manifold with a $(0, 2)$-tensor field g on M, called the (pseudo-)metric that assigns a non-degenerate, smooth, symmetric and bilinear map g_x to all points $x \in M$ as follows,

$$\begin{aligned}g_x : T_xM \times T_xM \to \mathbb{R}, \quad (X, Y) \mapsto g_x(X, Y) \\ g_x(X, Y) = g_x(Y, X) \\ g_x(a(X + Z), Y) = ag_x(X, Y) + ag_x(Z, Y).\end{aligned}$$

Notice that g_x has almost all of the properties of an inner-product on the tangent space T_xM, except for positive definiteness. For this reason, g is commonly denoted by $\langle\ ,\ \rangle$ and g_x by $\langle\ ,\ \rangle_x$. The smoothness property is such that for smooth vector fields X and Y, the map $M \to \mathbb{R}$, $x \mapsto \langle X_x, Y_x\rangle_x$ is a smooth map in the usual sense.

Locally, the metric can be expressed by a symmetric $n \times n$ matrix $A := A_{g_x}$, so that for tangent vectors $v, w \in T_xM$ we have the matrix multiplication

$$g_x(v, w) = v^T A w = \sum A_{ij} v^i w^j, \tag{2.4}$$

where a basis $\{\partial_i\}_{1 \leq i \leq n}$ for T_xM is chosen and the coefficients of A are denoted by $A_{ij} = g_x(\partial_i, \partial_j)$. The signature of the metric can then be defined as (p, n, r), where p are the number of positive eigenvalues of A, n are the number of negative eigenvalues and r the multiplicity of 0 as an eigenvalue. The non-degeneracy of g implies that $r = 0$ for all metrics, and therefore we specify metrics only by (p, n). Physicists have found that so-called Lorentzian manifolds are useful for relativity theory. A Lorentzian manifold is a pseudo-Riemannian manifold of which the metric has signature $(n-1, 1)$ or $(1, n-1)$.

It turns out that the metric yields a canonical choice for the connection, namely the Levi-Civita connection, on the tangent bundle TM.

Proposition 2.9 (Levi-Civita connection)**.** *Let (M, g) be a pseudo-Riemannian manifold. There exists a unique symmetric connection on the tangent bundle such that the connection*

1. *conserves the metric (so $(\nabla_Z \langle , \rangle)(X, Y) = 0$ for all $X, Y, Z \in \mathfrak{X}(M)$),*
2. *is torsion free (so $\mathrm{Tor}(X, Y) = 0$ for all $X, Y \in \mathfrak{X}(M)$).*

This connection is called the Levi-Civita connection.

Proof. Let X, Y and Z be vector fields. We will show that the Koszul formula given by

$$2\langle \nabla_X Y, Z \rangle = X\langle Y, Z \rangle + Y\langle Z, X \rangle - Z\langle X, Y \rangle - \langle X, [Y, Z] \rangle + \langle Y, [Z, X] \rangle + \langle Z, [X, Y] \rangle, \tag{2.5}$$

uniquely determines the connection $\nabla_X Y$.

This is due to the fact that a vector field V is uniquely defined by $\langle V, W \rangle$ for all vector fields $W \in \mathfrak{X}(M)$. Assume that $\langle V, W \rangle = \langle V', W \rangle$, or $\langle V - V', W \rangle = 0$ for all W, then by the non-degeneracy of the metric at each point $x \in M$ we have with $W = V - V'$ that $\langle V_x - V'_x, V_x - V'_x \rangle_x = 0$, so $V_x - V'_x = 0$ for all x, and therefore $V = V'$.

Assuming that ∇ as defined by this formula indeed satisfies the two conditions, then the uniqueness and existence of the Levi-Civita connection is proved. It remains to show that the two conditions are indeed satisfied. We start by showing that the metric is conserved. By our formula for the covariant derivative of a tensor field (2.3), we have that

$$\nabla_Z(\langle X, Y \rangle) := (\nabla_Z \langle , \rangle)(X, Y) + \langle \nabla_z X, Y \rangle + \langle X, \nabla_Z Y \rangle,$$

showing that ∇ conserves the metric is equivalent to showing

$$\langle \nabla_Z X, Y \rangle + \langle X, \nabla_Z Y \rangle = \nabla_Z(\langle X, Y \rangle) = Z\langle X, Y \rangle, \tag{2.6}$$

where the last equality is due to the fact that the covariant derivative of a smooth function coincides with the usual derivative. By the Koszul formula (2.5) and the fact that $\langle X, \nabla_Z Y\rangle = \langle \nabla_Z Y, X\rangle$ we have the expressions

$$2\langle \nabla_Z X, Y\rangle = Z\langle X, Y\rangle + X\langle Y, Z\rangle - Y\langle Z, X\rangle - \langle Z, [X, Y]\rangle + \langle X, [Y, Z]\rangle + \langle Y, [Z, X]\rangle,$$
$$2\langle X, \nabla_Z Y\rangle = Z\langle Y, X\rangle + Y\langle X, Z\rangle - X\langle Z, Y\rangle - \langle Z, [Y, X]\rangle + \langle Y, [X, Z]\rangle + \langle X, [Z, Y]\rangle.$$

By using the symmetry property of $\langle , \rangle$ and the skew-symmetry of the Lie-bracket $[X, Y] = -[Y, X]$, equation (2.6) immediately follows, so ∇ conserves the metric.
We then show that ∇ is torsion free, or

$$\nabla_X Y - \nabla_Y X = [X, Y].$$

This is equivalent with the condition that for all Z

$$\langle \nabla_X Y - \nabla_Y X, Z\rangle = \langle \nabla_X Y, Z\rangle - \langle \nabla_Y X, Z\rangle = \langle [X, Y], Z\rangle.$$

The Koszul formula (2.5) yields

$$2\langle \nabla_X Y, Z\rangle = X\langle Y, Z\rangle + Y\langle Z, X\rangle - Z\langle X, Y\rangle - \langle X, [Y, Z]\rangle + \langle Y, [Z, X]\rangle + \langle Z, [X, Y]\rangle,$$
$$2\langle \nabla_Y X, Z\rangle = Y\langle X, Z\rangle + X\langle Z, Y\rangle - Z\langle Y, X\rangle - \langle Y, [X, Z]\rangle + \langle X, [Z, Y]\rangle + \langle Z, [Y, X]\rangle.$$

Because of the symmetry property of $\langle , \rangle$ we immediately see that the first three terms in both expressions cancel each other when subtracting, so we are left with

$$\begin{aligned} 2\langle \nabla_X Y - \nabla_Y X, Z\rangle =& -\langle X, [Y, Z]\rangle + \langle Y, [Z, X]\rangle + \langle Z, [X, Y]\rangle \\ &+ \langle Y, [X, Z]\rangle - \langle X, [Z, Y]\rangle - \langle Z, [Y, X]\rangle \\ =& 2\langle Z, [X, Y]\rangle = 2\langle [X, Y], Z\rangle, \end{aligned}$$

where we used the skew-symmetry of the Lie-bracket again to cancel terms. This shows that indeed

$$\nabla_X Y - \nabla_Y X = [X, Y],$$

so the metric is torsion free. This concludes the proof of proposition 2.9. □

In physics, the terminology of "raising and lowering" indices of a tensor is frequently used. We will cover these operations in the setting of pseudo-Riemannian geometry. The idea is that there exists a canonical isomorphism between a vector space V and its dual space V^*. We define these maps below for $V = T_x M$.

Definition 2.10 (flat, sharp)**.** We define

$$\flat_g : T_x M \to T_x^* M, \quad v \mapsto g_x(v, \cdot), \tag{2.7}$$

where the dot $\cdot$ indicates that the object is a function of a variable, which in this case is an element of the tangent space, that should take the place of the dot.
Then we define

$$\sharp_g : T_x^*M \to T_xM, \quad \omega \mapsto \omega^\sharp, \text{ such that } \omega(w) = g_x(\omega^\sharp, w). \tag{2.8}$$

We call $\flat_g$ *flat* and $\sharp_g$ *sharp*.[1]

Note that these maps are indeed each others inverse as

$$\begin{aligned} \sharp_g \circ \flat_g(v) = \sharp_g(g_x(v,\cdot)) = v \quad &\text{because} \quad g_x(v,w) = g_x(\omega^\sharp, w) \quad \text{so } \omega^\sharp = v \\ \flat_g \circ \sharp_g(\omega) = \flat_g(\omega^\#) = g_x(\omega^\sharp, \cdot) = \omega \quad &\text{because} \quad \omega(w) = g_x(\omega^\sharp, w). \end{aligned}$$

By analyzing the behavior of the flat and sharp maps in charts, the notation in physics for raising and lowering indices will become clear. From now on we employ the *Einstein summation convention*, where a sum over repeated indices is implied. The metric g can then be written as $g = A_{ij}dx^i dx^j$, where $\{dx^i\}_{1\le i\le n}$ are the basis vectors for the cotangent space of dimension n. For smooth vector fields X, Y we then have

$$\flat_g(X)(Y) = A_{ij}X^kY^l\partial_k\partial_l dx^i dx^j = A_{ij}X^kY^l\delta^i_k\delta^j_l = A_{ij}X^iY^j,$$

where we used $dx^j\partial_i = \delta^j_i$ which is one if $i = j$ and zero otherwise. Essentially, we have obtained the covector field $\flat_g(X)$ denoted by

$$\flat_g(X) = X_j dx^j,$$

where $X_j = A_{ij}X^i$ so that the index is lowered. Analogously, the sharp map $\sharp_g$ (2.8) can be applied to *co*vector fields to create vector fields, $\omega^i = (A^*)^{ij}\omega_j$. In example 2.11 we will find that the numbers $(A^*)^{ij}$, expressed in the dual basis, are the ij^{th} components of the matrix inverse A^{-1}.

The notion of raising and lowering indices can be extended to sections of the space

$$\underbrace{TM \otimes \cdots \otimes TM}_{p} \otimes \underbrace{T^*M \otimes \cdots \otimes T^*M}_{q}.$$

The following example shows how to apply the flat and sharp maps on the metric to define the contravariant metric.

Example 2.11 (contravariant metric)**.** By definition 2.8, the metric is described as an element of $\Gamma(T^{(0,2)}M)$. At each point $x \in M$ it takes two vectors and produces a number. The contravariant metric g^* is an element of $\Gamma(T^{(2,0)}M)$ which locally operates as

$$g^*(\omega, \eta) = (A^*)^{ij}\omega_i\eta_j.$$

[1]The names and symbols for the flat and sharp map are inspired by musical notation, where these symbols are used to respectively lower or raise a tone. Analogously, vectors (which are thought of as pointy) are flattened/lowered by the flat map, while one forms are sharpened/raised by the sharp map.

Locally we may write $g_x(X,Y) = A_{ij}dx^i dx^j$ where $A_{ij} = g_x(\partial_i, \partial_j)$ for a basis $\{\partial_i\}_{1\leq i\leq n}$. In coordinates we have seen how the flat and sharp maps act on vectors and covectors respectively. By applying the flat and sharp map on the vector ∂_i and corresponding covector dx^i we obtain

$$\begin{aligned}\flat_g(\sharp_g(dx^i)) &= A_{ij}(A^*)^{jk}\partial_k = dx^i \\ \sharp_g(\flat_g(\partial_i)) &= (A^*)^{ij}A_{jk}dx^k = \partial_i,\end{aligned}$$

where the last equalities are true because we have shown that $\flat_g$ and $\sharp_g$ are inverse maps. It follows that, when interpreting $A_{ij}(A^*)^{jk}$ as a matrix multiplication, we obtain

$$A_{ij}(A^*)^{jk} = \delta_i^k.$$

This means that the components of A^* are the components of the inverse matrix A^{-1}, so $(A^*)^{ij} = (A^{-1})_{ij}$. For this reason, the contravariant metric is also called the inverse metric.

Next, consider the tangent bundle TM with (M,g) a pseudo-Riemannian manifold. Sections of this tangent bundle are then precisely the vector fields $\mathfrak{X}(M)$ as seen in example 2.3. The curvature of the connection then turns into the *Riemann curvature tensor*

$$\begin{aligned}R &: \mathfrak{X}(M)\times\mathfrak{X}(M)\times\mathfrak{X}(M)\to\mathfrak{X}(M) \\ R(X,Y)Z &= \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z.\end{aligned} \tag{2.9}$$

This tensor "eats" three vector fields and "spits out" another one, so according to definition 2.5 R is an element of $\Gamma(T^{(1,3)}M)$. We can also see this as an object that eats two vector fields X and Y and spits out an object that sends vector fields to vector fields, or an endomorphism $Z\mapsto R(X,Y)Z$. By taking the trace of this endomorphism, we obtain the Ricci tensor

$$\mathrm{Ric} = \mathrm{Tr}(Z\mapsto R(X,Y)Z), \tag{2.10}$$

which is then a symmetric element of $\Gamma(T^{(0,2)}M)$. We can "contract" the Ricci tensor to obtain the Ricci scalar or scalar curvature S, formally defined by

$$S = \mathrm{Tr}_g\mathrm{Ric}, \tag{2.11}$$

which is a $C^\infty(M,\mathbb{R})$ function. In physics, S is often denoted by $R = R_i^i = g^{ij}R_{ij} = g^{ij}R^k_{ikj}$. Concretely, for each point $x\in M$ we can apply the sharp map on Ric_x to create the unique linear map $\rho_x : T_xM\to T_xM$ so that

$$\mathrm{Ric}_x(v,w) = g_x(\rho_x(v),w),$$

for all $v,w\in T_xM$ or in indices

$$(\rho_x)_i^j = g^{jk}(\mathrm{Ric}_x)_{ki}.$$

The scalar curvature S_x is defined pointwise as the trace of ρ_x, so for an orthonormal basis $\{\partial_i\}$ we have

$$S_x = \mathrm{Tr}(v \mapsto \rho_x(v)) = \sum g_x(\rho_x(\partial_i), \partial_i) = \sum \mathrm{Ric}(\partial_i, \partial_i),$$

which also explains the physicists notation $g^{ij}R_{ij}$ (here R_{ij} denotes the Ricci tensor). To emphasize the dependence of S on the metric g we define the contraction map

$$C : \Gamma(T^{(2,0)}M) \times \Gamma(T^{(0,2)}M) \to C^\infty(M, \mathbb{R}), \quad (A, B) \mapsto C(A, B), \tag{2.12}$$

which is locally equivalent to the above construction, but for arbitrary tensors fields A and B, so that it easily follows that C is bilinear. In chapter 3, we will see the role that these curvature objects play in the Einstein equation.

2.4 Calculus on Pseudo-Riemannian manifolds

In chapter 3, we will derive the Einstein Field Equations by applying the stationary-action principle on the Einstein-Hilbert action. The goal of this chapter is to make sense of this procedure. The Einstein-Hilbert action is given by

$$F : \mathcal{M} \to \mathbb{R}, \quad g \mapsto -\int_M S(g)\mu, \tag{2.13}$$

where $\mathcal{M}$ is the space of Lorentzian metrics on M, $S(g) \in C^\infty(M, \mathbb{R})$ is the Ricci scalar of g and μ is the *density* induced by g. To make sense of (2.13), we need to define integration of densities on a manifold M.

Definition 2.12 (densities)**.** Let V be an n-dimensional vector space and denote by $\mathcal{D}(V)$ the space of functions σ of n-tuples of vectors so that for linear maps $A : V \to V$ they satisfy

$$\sigma(Av_1, \ldots, Av_n) = |\det A|\sigma(v_1, \ldots, v_n).$$

A density μ on M is then defined as a smooth map that assigns to each point $x \in M$ an element of $\mathcal{D}(T_xM)$.

For pseudo-Riemannian manifolds, the metric g naturally induces a density μ_g, by

$$\mu_g : T_xM \ni v_1, \ldots, v_n \mapsto \sqrt{|\det(g_x(v_i, v_j))|}$$

where $\det(g_x(v_i, v_j))$ is the determinant of the $n \times n$ matrix obtained by taking the inner products $g_x(v_i, v_j)$ for all $1 \leq i, j \leq n$. Note that for a linear map A we have

$$\begin{aligned}\sqrt{|\det(\langle Av_i, Av_j\rangle)|} &= \sqrt{|\det(A\langle v_i, v_j\rangle A^T)|} = \sqrt{\det(A)^2}\mu_g(v_1, \ldots, v_n) \\ &= |\det(A)|\mu_g(v_1, \ldots, v_n),\end{aligned}$$

so that the definition of a density is satisfied. Note that this density μ_g is nowhere vanishing because the metric is non-degenerate. Let f be a scalar function. In the case of $M = \mathbb{R}^n$, we write

$$\int_M f(x)dx,$$

for the integral of f. For coordinate transformations $\phi : M \to M$ such that $\phi(y) = x$, the integrand would change to $f(y)|\det(D\phi(y))|dy$ so the value of the integral is invariant. In our case, where M is pseudo-Riemannian, the objects that we can integrate are not functions but densities μ. By taking the pullback of a density μ through a smooth function f, so $f^*\mu$, is again a density. This is what we do when we want to calculate the integral of the Einstein-Hilbert action (2.13).

Once a chart is chosen, any density can be identified with a top degree n-form ω, and in the case of μ_g we have

$$\mu_g = \sqrt{g}dx^1 \wedge \cdots \wedge dx^n =: \omega.$$

If we assume for simplicity that M is described by a single chart, integration of a function multiplied by the density μ_g is then defined as

$$\int_M f\mu_g := \int_M f\sqrt{g}dx^1 \wedge \cdots \wedge dx^n,$$

where the right hand side is defined in the usual way for integration on a smooth manifold. It turns out that by defining the integrand this way, it transforms under diffeomorphisms in such a way that the integral remains invariant. It can be shown that the Stokes and Divergence theorems from ordinary calculus are still true in this more general setting.

Finally, we explain what we mean by the "stationary action principle". Let the action F be defined as in (2.13). The derivative of this map at some "point" $g \in \mathcal{M}$ is given by

$$dF_g : T_g\mathcal{M} \to \mathbb{R}, \quad \hat{g} \mapsto dF_g(\hat{g}).$$

We then define a *stationary point* of F as an element $g \in \mathcal{M}$ such that

$$dF_g(\hat{g}) = 0 \quad \text{for all } \hat{g} \in T_g\mathcal{M}.$$

3 Einstein Field Equations

Albert Einstein's general relativity states that gravitational phenomena are described by a (pseudo-)metric g on a manifold M. From the metric g, the important curvature objects, such as the Riemann curvature tensor and the Ricci tensor, can be constructed. The relation between matter (energy, momentum, stress and shear) and the curvature objects is given by the Einstein Field Equations (EFE)

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}.$$

This form of the EFE is how one would generally find it in physics textbooks (see [1], Eq. (4.44)). Here $R_{\mu\nu}$ is the Ricci tensor and R is the Ricci scalar. Furthermore, G is Newton's constant of gravity, c the speed of light and $T_{\mu\nu}$ is the *energy momentum* tensor. In this section, we will derive the EFE (for arbitrary dimension) in the form

$$\mathrm{Ric}^* - \frac{1}{2}Sg^* = 0,$$

from the Einstein-Hilbert action. There are a few differences between the two forms. Firstly, we denote the Ricci tensor by Ric and the Ricci scalar by S, instead of $R_{\mu\nu}$ and R respectively. Secondly, we derive the equations for the vacuum situation, so we have no energy momentum term on the right hand side. Lastly, the $*$'s on the tensors indicate that we take the *contravariant* of the tensor, but it turns out that in the end the covariant and contravariant form of the EFE are equivalent. We have already defined what $*$ does on g in example 2.11. The object Ric^* is defined differently as we will see in this paragraph.

3.1 Derivation of the EFE

We first briefly outline the derivation. Then we will compute the different components separately, and put everything together at the end.

Derivation outline

We start with the Einstein-Hilbert action, on which we apply the stationary-action principle. This is in essence a physical law which states that upon variation of the action, the critical points determine the trajectories that occur in nature. The Einstein-Hilbert action is described by the map

$$F : \mathcal{M} \to \mathbb{R}$$
$$F(g) = -\int_M S(g)\mu,$$

with $\mathcal{M}$ the space of Lorentzian metrics, M the manifold over which we are integrating, $S(g)$ the Ricci scalar corresponding to g (as in definition 2.11) and μ the density induced

by g as introduced in 2.4. Applying the stationary action principle (see section 2.4) yields for a variation $\hat{g}$

$$-(dF)_g(\hat{g}) = \int_M dS_g(\hat{g})\mu + \int_M S d\mu_g(\hat{g}).$$

By defining the maps

$$\begin{aligned} * \,:\, \mathcal{M} &\to \Gamma(T^{(2,0)}M), \quad g \mapsto *(g) = g^* \\ \mathrm{Ric} \,:\, \mathcal{M} &\to \Gamma(T^{(0,2)}M), \quad g \mapsto \mathrm{Ric}(g), \end{aligned}$$

the scalar curvature can then be described as a function of the metric by

$$S \,:\, \mathcal{M} \to C^\infty(M, \mathbb{R}), \quad S(g) = C(g^*, \mathrm{Ric}(g)).$$

The contraction map C (see also (2.12)) is bilinear. The derivative of a bilinear map C at the point (A, B) acting on a variation $(\hat{A}, \hat{B})$ is given by

$$(dC)_{(A,B)}(\hat{A}, \hat{B}) = C(A, \hat{B}) + C(\hat{A}, B).$$

From the chain rule it follows that

$$\begin{aligned} dS \,:\, T_g\mathcal{M} &\to T_{S(g)}C^\infty(M, \mathbb{R}) \cong C^\infty(M, \mathbb{R}), \\ (dS)_g(\hat{g}) &= (dC)_{S(g)}((d*)_g(\hat{g}), (d\mathrm{Ric})_g(\hat{g})) \\ &= C((d*)_g(\hat{g}), \mathrm{Ric}(g)) + C(g^*, (d\mathrm{Ric})_g(\hat{g})). \end{aligned}$$

The variation of the Einstein-Hilbert action is then given by

$$-(dF)_g(\hat{g}) = \int_M C((d*)_g(\hat{g}), \mathrm{Ric}(g))\mu + \int_M C(g^*, (d\mathrm{Ric})_g(\hat{g}))\mu + \int_M S(d\mu)_g(\hat{g}). \quad (3.1)$$

Most of the work resides in rewriting these integrals in such a way that the $\hat{g}$ is "isolated", so that we can read off the equations of motion. After doing these computations in the following 3 paragraphs, we obtain equations (3.4),(3.5) and (3.14), so that

$$(dF)_g(\hat{g}) = \int_M C(\mathrm{Ric}^* - \tfrac{1}{2}Sg^*, \hat{g})\mu_g. \quad (3.2)$$

By applying the stationary action principle, i.e., requiring $dF_g(\hat{g}) = 0$ for all variations $\hat{g}$, the EFE in vacuum

$$\mathrm{Ric}^* - \tfrac{1}{2}Sg^* = 0 \quad (3.3)$$

follow.

$\int_M C((d*)_g(\hat{g}), \mathrm{Ric}(g))\mu$

In order to calculate $(d*)_g(\hat{g})$, we use that in charts we have the identity

$$(g^*)^{ik}g_{kj} = g_{jk}(g^*)^{ki} = \delta^i_j,$$

which we have seen in example 2.11. Differentiating this identity with respect to the metric g, we obtain by Leibniz

$$((d*)_g(\hat{g}))^{ik}g_{kj} + (g^*)^{ik}(((di)_g)(\hat{g}))_{kj} = 0,$$

where i is the identity map, which is linear, so that its derivative equals the identity map at all points, i.e., $(di)_g(\hat{g}) = \hat{g}$. We then obtain by rearranging

$$\begin{aligned}
((d*)_g(\hat{g}))^{ik}g_{kj} &= -(g^*)^{ik}(\hat{g})_{kj} \\
((d*)_g(\hat{g}))^{ik}g_{kj}(g^*)^{jl} &= -(g^*)^{ik}(\hat{g})_{kj}(g^*)^{jl} \\
((d*)_g(\hat{g}))^{ik}\delta^l_k &= -(g^*)^{ik}(\hat{g})_{kj}(g^*)^{jl} \\
((d*)_g(\hat{g}))^{lk} &= -(g^*)^{ik}(\hat{g})_{kj}(g^*)^{jl},
\end{aligned}$$

which can be written in the index-free notation as

$$(d*)_g(\hat{g}) = -\sharp_g\hat{g}\sharp_g,$$

where the two sharp maps indicate that we are raising the indices of the $(0,2)$ tensor field $\hat{g}$. The integrand can then be written as

$$C((d*)_g(\hat{g}), \mathrm{Ric}(g)) = C(-\sharp_g\hat{g}\sharp_g, \mathrm{Ric}(g)) = C(\hat{g}, -\sharp_g\mathrm{Ric}(g)\sharp_g),$$

which follows from the local definition of the contraction. The object $\sharp_g\mathrm{Ric}(g)\sharp_g$ is then the contravariant Ricci tensor with both of its indices raised, which we define as Ric^*. In the end we obtain

$$\int_M C((d*)_g(\hat{g}), \mathrm{Ric}(g))\mu = \int_M C(-\mathrm{Ric}^*(g), \hat{g})\mu. \tag{3.4}$$

$\int_M C(g^*, (d\mathrm{Ric})_g(\hat{g}))\mu$

We will argue that

$$\int_M C(g^*, (d\mathrm{Ric})_g(\hat{g}))\mu = 0. \tag{3.5}$$

We will need to compute the derivative of the Ricci tensor. For this purpose, we first introduce some auxiliary maps. First, remember that from proposition 2.9, it followed that for any metric $g \in \mathcal{M}$ on a pseudo-Riemannian manifold, there exists a unique connection called the Levi-Civita connection. This allows us to introduce the well-defined map

$$\mathrm{LC} : \mathcal{M} \to \mathcal{C}, \quad g \mapsto LC(g),$$

where $\mathcal{C}$ is the space of torsion-free connections on the tangent bundle TM. Next we define the map curv, that is closely related to the Riemann curvature tensor R as

$$\text{curv} : \mathcal{C} \to \mathcal{R}, \quad \text{curv}(\nabla)(X,Y) \mapsto \nabla_{[X,Y]} - [\nabla_X, \nabla_Y],$$

where we denote by $\mathcal{R}$ the space of "curvature-like" tensors. Notice that we can obtain the Riemann curvature tensor R, corresponding to a certain metric, by composing the previous two maps

$$R = (\text{curv} \circ \text{LC})(g).$$

We are ultimately interested in the derivative of *the Ricci tensor* dRic. The Ricci tensor is defined as the trace of the Riemann curvature tensor, so by the chain rule we have

$$(d\text{Ric})_g = d(\text{Tr}(\cdot) \circ R)_g = d(\text{Tr}\cdot)_{R(g)} \circ (dR)_g = \text{Tr}((dR)_g), \tag{3.6}$$

because the trace is a linear map and the derivative of a linear map is itself at every point. We will calculate the derivative dR by the chain rule,

$$(dR)_g = (d\text{curv})_{\text{LC(g)}} \circ (d\text{LC})_g,$$

or, if we let $(dR)_g$ act on a variation $\hat{g} \in T_g\mathcal{M}$

$$(dR)_g(\hat{g}) = (d\text{curv})_{\text{LC(g)}}((d\text{LC})_g(\hat{g})).$$

First we look at the derivative of the LC map at a point $g \in \mathcal{M}$

$$(d\text{LC})_g : T_g\mathcal{M} \to T_{LC(g)}\mathcal{C}, \quad \hat{g} \mapsto A := (d\text{LC})_g(\hat{g}). \tag{3.7}$$

The tangent space $T_{LC(g)}\mathcal{C}$ can be identified with the space $\mathcal{A}$ defined by

$$\mathcal{A} = \{\nabla - \nabla' : \nabla, \nabla' \in \mathcal{C}\},$$

which is a space consisting of symmetric tensors. The elements of $\mathcal{A}$ are symmetric because ∇ and ∇' are torsion-free connections, so that

$$A_X Y = \nabla_X Y - \nabla'_X Y = \nabla_Y X + [X,Y] - (\nabla'_Y X + [X,Y]) = \nabla_Y X - \nabla'_Y X = A_Y X.$$

To see that the elements of $\mathcal{A}$ are tensors we introduce a smooth function f and use that for connections we have

$$\nabla_{fY} X = f \nabla_Y X,$$

so we have

$$(\nabla_X - \nabla'_X)(fY) = \nabla_{fY} X - \nabla'_{fY} X = f(\nabla_Y - \nabla'_Y)X = f(\nabla_X - \nabla'_X)Y,$$

where we used the symmetry of $\nabla - \nabla'$ twice. It follows that elements of $\mathcal{A}$ are $C^\infty(M)$-linear. By proposition 2.6 it follows they are tensors. We say that $\mathcal{A}$ is the associated vector space to the affine space of torsion-free connections $\mathcal{C}$. It is then a general fact for affine spaces that their tangent spaces at any point can be identified by the associated vector space, so

$$T_{\mathrm{LC}(g)}\mathcal{C} \cong \mathcal{A}.$$

Secondly, we look at the derivative of the curv map

$$\begin{aligned}(d\mathrm{curv})_\nabla &: \mathcal{A} \to T_{\mathrm{curv}(A)}\mathcal{R} \\ \big((d\mathrm{curv})_\nabla(A)\big)(X,Y) &= A_{[X,Y]} - A_X\nabla_Y - \nabla_X A_Y + A_Y\nabla_X + \nabla_Y A_X.\end{aligned}$$

Since the Levi-Civita connection is torsion free, it follows that

$$[X,Y] = \nabla_X Y - \nabla_Y X,$$

so

$$A_{[X,Y]} = A_{\nabla_X Y} - A_{\nabla_Y X},$$

and fully written out the derivative of the curv map then becomes

$$\big((d\mathrm{curv})_\nabla(A)\big)(X,Y) = A_{\nabla_X Y} - A_{\nabla_Y X} - A_X\nabla_Y - \nabla_X A_Y + A_Y\nabla_X + \nabla_Y A_X. \tag{3.8}$$

In addition by formula (2.3) for the covariant derivative of a $(1,2)$ tensor field, we have

$$\begin{aligned}\nabla_X(A(Y,Z)) &= (\nabla_X A)(Y,Z) + A(\nabla_X Y, Z) + A(Y, \nabla_X Z) \\ &= (\nabla_X A)(Y,Z) + A_{\nabla_X Y} Z + A_Y \nabla_X Z.\end{aligned}$$

By rearranging the terms in 3.8, we notice that

$$\begin{aligned}\big((d\mathrm{curv})_\nabla(A)\big)(X,Y)Z &= \Big(A_{\nabla_X Y} - A_{\nabla_Y X} - A_X\nabla_Y - \nabla_X A_Y + A_Y\nabla_X + \nabla_Y A_X\Big)Z \\ &= \Big(\nabla_Y A_X Z - A_{\nabla_Y X} Z - A_X \nabla_Y Z\Big) + \Big(- \nabla_X A_Y Z + A_{\nabla_X Y} Z + A_Y \nabla_X Z\Big) \\ &= (\nabla_Y A)(X,Z) - (\nabla_X A)(Y,Z).\end{aligned}$$

By equation (3.6), the derivative of the Ricci tensor in terms of a variation $\hat{g}$ is then given by

$$\begin{aligned}\big((d\mathrm{Ric})_g(\hat{g})\big)(Y,Z) &= \mathrm{Tr}\left(X \mapsto \big((d\mathrm{curv})_\nabla(A)\big)(X,Y)Z\right) \\ &= \mathrm{Tr}\big(X \mapsto (\nabla_Y A)(X,Z) - (\nabla_X A)(Y,Z)\big) \\ &= \mathrm{Tr}\big(X \mapsto (\nabla_Y A)(X,Z)\big) - \mathrm{Tr}\big(X \mapsto (\nabla_X A)(Y,Z)\big),\end{aligned} \tag{3.9}$$

where the metric dependence is inside A. When we plug in a vector field X into the $(1,2)$-tensor field A, we obtain for all points $x \in M$ a map

$$A(\cdot, X)(x) : T_xM \to T_xM, \quad w \mapsto A(w, X(x)),$$

We then define the trace of A as the map

$$\mathrm{Tr}(A) : \Gamma(T^{(1,2)}M) \to \Gamma(T^{(0,1)}M), \text{ such that } \mathrm{Tr}(A)(Y) = \mathrm{Tr}(X \mapsto A(X,Y)), \ Y \in \mathfrak{X}(M),$$

which is a covector field. We introduce the vector field

$$V = \#_g \mathrm{Tr}(A) - C(g^*, A), \tag{3.10}$$

where now the contraction map C only works on the two lower indices of A.

Lemma 3.1. *Let V be as in (3.10). Denote by ∇V the covariant derivative of the vector field V, g^* the contravariant metric and $(dRic)_g(\hat{g})$ the variation of the Ricci tensor as in 3.9. Then we have*

$$\mathrm{Tr}(\nabla V) = C(g^*, (d\mathrm{Ric})_g(\hat{g})). \tag{3.11}$$

Proof. In a chart, we can write for the LHS of (3.11)

$$\begin{aligned} \mathrm{Tr}(\nabla V) = \nabla_i V^i &= \nabla_i (g^{ij} A^k_{kj} - g^{kj} A^i_{kj}) \\ &= \nabla_i g^{ij} A^k_{kj} - \nabla_i g^{kj} A^i_{kj} \end{aligned}$$

and for the RHS of (3.11) we get by employing index notation on (3.9)

$$\begin{aligned} C(\hat{g}, (d\mathrm{Ric})_g(\hat{g})) &= g^{ij} \big(d\mathrm{Ric}_g(\hat{g})\big)_{ij} \\ &= g^{ij} (\nabla_i A^k_{kj} - \nabla_k A^k_{ij}) \\ &= \nabla_i g^{ij} A^k_{kj} - \nabla_k g^{ij} A^k_{ij}, \end{aligned}$$

where in the last line we used that g commutes with ∇. Now we notice that the LHS and RHS are indeed equal (after relabeling $k \leftrightarrow i$ for the second terms). □

Because $\hat{g}$ is a variation of compact support, we have that $A := (d\mathrm{LC})_g(\hat{g})$ is also of compact support, and so V is a vector field of compact support. Furthermore, the trace of the covariant derivative is equal to the divergence, in formulas: $\mathrm{Tr}(\nabla V) = \mathrm{div} V$. The divergence here is defined (in local coordinates) as

$$\mathrm{div} V = \sum_{i=1}^{n} \frac{1}{\sqrt{|\det(g_{jk})|}} \frac{\partial(\sqrt{|\det(g_{jk})|} V^i)}{\partial x^i},$$

and one can prove that the divergence theorem on pseudo-Riemannian manifolds still holds. By lemma 3.1, we then find

$$\int_M C(g^*, (d\mathrm{Ric})_g(\hat{g}))\, \mu = \int_M \mathrm{Tr}(\nabla V)\, \mu = \int_M \mathrm{div} V\, \mu = \int_{\partial M} V \cdot n\, dS = 0,$$

where the last line equals zero, because V vanishes outside some compact subset of M.

$\int_M S(d\mu)_g(\hat{g})$

We will show that

$$(d\mu)_g(\hat{g}) = \tfrac{1}{2}C(g^*, \hat{g})\mu.$$

Let $x \in M$, $\{\partial_i\}_{1\leq i\leq n}$ an orthonormal basis for T_xM, and denote by A be the matrix with components $A_{ij} = g_x(\partial_i, \partial_j)$, so that

$$\mu = \sqrt{\sigma \det(A)}, \tag{3.12}$$

where σ is the sign of the determinant. Notice that g is non-degenerate and symmetric, so the matrix A will have a non-zero determinant. Moreover, all eigenvalues $\lambda_1, \ldots, \lambda_n$ are non-zero real numbers. This means that the logarithms of σA and $\sigma \det(A)$ are defined. For convenience, assume that $\sigma = +1$. We have that

$$\log(\det(A)) = \mathrm{Tr}(\log(A)),$$

which follows from $\det(A) = \lambda_1 \cdots \lambda_n$ and $\mathrm{Tr}(\log(A)) = \log(\lambda_1) + \cdots + \log(\lambda_n)$. Then we use the chain rule for differentiation on the left side to obtain

$$d(\log \circ \det)_A(\hat{g}) = d(\log \cdot)_{\det(A)} \circ d(\det \cdot)_A(\hat{g}) = \frac{1}{\det(A)} d(\det \cdot)_A(\hat{g})$$

but if we apply the chain rule on the right side we obtain

$$d(\mathrm{Tr} \circ \log)_A(\hat{g}) = d(\mathrm{Tr}\cdot)_{\log(A)} \circ d(\log \cdot)_A(\hat{g}) = \mathrm{Tr}(A^{-1}\hat{g})$$

where we used that the derivative of the logarithm map in A for matrices associates the inverse matrix A^{-1} to the point A, so combining the two above equations we have

$$d(\det \cdot)_A(\hat{g}) = \mathrm{Tr}(A^{-1}\hat{g}) \det(A). \tag{3.13}$$

In the case $\sigma = -1$ we would also obtain this expression because the two minus signs cancel. Now we use the chain rule one final time on 3.12, which yields

$$(d\mu)_g(\hat{g}) = d(\sqrt{\cdot} \circ (\sigma \det(\cdot))_A(\hat{g}) = d(\sqrt{\cdot})_{\sigma \det(A)} \circ d(\sigma \det(\cdot))_A(\hat{g}),$$

and because σ is just a constant this gives

$$(d\mu)_g(\hat{g}) = \frac{\sigma}{2\sqrt{\sigma \det(A)}} d(\det(\cdot))_A(\hat{g}),$$

and by substituting (3.13) we obtain

$$(d\mu)_g(\hat{g}) = \frac{\sigma}{2}\mathrm{Tr}(A^{-1}\hat{g}) \frac{\det(A)}{\sqrt{\sigma \det(A)}} = \frac{1}{2}\mathrm{Tr}(A^{-1}\hat{g})\mu.$$

Finally, by example 2.11 we can conclude that locally multiplying $\hat{g}$ by the inverse matrix A^{-1} and then taking the trace, is equivalent to our coordinate-free expression for $C(g^*, \hat{g})$, so that indeed

$$(d\mu)_g(G) = \tfrac{1}{2}C(g^*, \hat{g})\mu.$$

This then allows us to rewrite

$$\int_M S(d\mu)_g(\hat{g}) = \int_M \tfrac{1}{2}C(Sg^*, \hat{g})\mu \tag{3.14}$$

Remarks on EFE

So, we derived from the Einstein-Hilbert action

$$F(g) = -\int_M S\mu$$

the EFE in vacuum

$$\mathrm{Ric}^* - \tfrac{1}{2}Sg^* = 0.$$

If one instead would view the objects in the integrand of the action as functions of the *contravariant* metric, and took derivatives with respect to the contravariant metric, we would get the equations

$$\mathrm{Ric} - \tfrac{1}{2}Sg = 0,$$

without stars. Because the metric and contravariant metric uniquely define each other, the equations must be equivalent.

Also take note of the interesting fact that we have not put any restrictions on the dimension of the manifold M. So, even though general relativity is generally regarded as a four-dimensional theory, the above derivation is applicable to higher and lower dimensions.

The term corresponding to matter, T, can be added to the EFE by slightly altering the Einstein-Hilbert action so that it contains a term corresponding to the matter fields. This would yield as a final result

$$\mathrm{Ric} - \tfrac{1}{2}Sg = T.$$

3.2 Approximate theories for EFE

It is notoriously difficult to find analytic solutions to the EFE. One of the ways of dealing with this problem is by numerically solving the system of differential equations corresponding to the EFE. Another approach is using approximation methods. In this section, we will focus on *why* these approximation methods could be fruitful instead of how to actually apply them. In this section we will employ the standard physics index notation for the more lengthy equations.

As a starting point we adopt the Landau-Lifshitz formulation of general relativity from Chapter 6 of "Gravity: Newtonian, Post-Newtonian, Relativistic" by Poisson and Will [11]. We adopt the convention that for the contravariant form g^* of the metric g we denote in indices $g^{ij} := (g^*)^{ij}$. First we introduce

$$\tilde{g}^{ij} = \sqrt{-\det(g)}\, g^{ij} \quad \text{(6.1) in [11]}, \tag{3.15}$$

where the metric determinant $\det(g) := \det(g_{ij})$ is locally defined by taking the determinant of the matrix g_{ij} representing the metric g in some basis (see equation (2.4)). The minus sign in (3.15) is necessary because g is a Lorentzian metric with signature $(1, n-1)$ so that the $\det(g) < 0$. We then define the *metric perturbation* h^{ij} as

$$h^{ij} = \eta^{ij} - \tilde{g}^{ij} \quad \text{(6.48) in [11]}, \tag{3.16}$$

where η is the constant flat/Minkowski metric[2]. Another way of introducing the metric perturbation is by defining $h_{ij} = g_{ij} - \eta_{ij}$ (we will do this in chapter 4). The formulation in terms of $\tilde{g}$ is an artifact of the Landau-Lifshitz formalism. The metric g_{ij} is uniquely defined by (3.15) for dimensions $n > 2$ because

$$\begin{aligned}\det(\tilde{g}^{ij}) &= \det\left(\sqrt{-\det(g)}\, g^{ij}\right) = \left(\sqrt{-\det(g)}\right)^n \det(g^{ij}) \\ &= \left(\sqrt{-\det(g)}\right)^n \frac{1}{\det(g)} = -\left(\sqrt{-\det(g)}\right)^{n-2},\end{aligned}$$

so that g^{ij} can be calculated and inverted to obtain g_{ij} and then $h_{ij} = g_{ij} - \eta_{ij}$ can be calculated. It is important to stress that the way we defined h^{ij} in (3.16) is in general *not equal* to the contravariant form of h_{ij} by the other definition, but they do *define each other*.

In equation (6.51) of [11] it is stated that the EFE can be rewritten as

$$\Box h^{ij} = -\frac{16\pi G}{c^4}\tau^{ij} \quad \text{(6.51) in [11]}, \tag{3.17}$$

$$\partial_i h^{ij} = 0 \quad \text{(6.49) in [11]}, \tag{3.18}$$

where (3.18) indicates the *harmonic gauge condition* that h has to satisfy, the $\Box$ denotes the flat spacetime d'Alembert operator, and the τ, called the "effective energy-momentum pseudotensor", on the right hand side of (3.17). The expression for τ is

$$\tau^{ij} := (-\det(g))(T^{ij} + t^{ij}_{LL} + t^{ij}_H) \quad \text{(6.52) in [11]}$$

$$\frac{16\pi G}{c^4}(-\det(g))t^{ij}_H := \partial_k h^{il}\partial_l h^{jk} - h^{kl}\partial_{kl}h^{ij} \quad \text{(6.53) in [11]} \tag{3.19}$$

$$\begin{aligned}\frac{16\pi G}{c^4}(-\det(g))t^{ij}_{LL} :=\ & \partial_m\tilde{g}^{ij}\partial_k\tilde{g}^{mk} - \partial_m\tilde{g}^{im}\partial_k\tilde{g}^{jk} + \tfrac{1}{2}g^{ij}g_{mk}\partial_n\tilde{g}^{ml}\partial_l\tilde{g}^{kn} \\ & - g^{im}g_{kl}\partial_n\tilde{g}^{jl}\partial_m\tilde{g}^{kn} - g^{jm}g_{kl}\partial_n\tilde{g}^{il}\partial_m\tilde{g}^{kn} + g_{mk}g^{ln}\partial_l\tilde{g}^{im}\partial_n\tilde{g}^{jk} \\ & + \tfrac{1}{8}(2g^{im}g^{jk} - g^{ij}g^{mk})(2g_{ln}g_{sr} - g_{ns}g_{lr})\partial_m\tilde{g}^{lr}\partial_k\tilde{g}^{ns} \quad \text{(6.5) in [11]},\end{aligned} \tag{3.20}$$

where T^{ij} is the usual energy-momentum tensor. The objects t_H and t_{LL} (see (6.53) and (6.5) in [11]) arise from rewriting the EFE in terms of h and $\tilde{g}$ instead of only g. They are quite complicated, in particular t_{LL}, though luckily for us, their specific form will not matter too much. Our goal is to proof proposition 3.2, which essentially states that small deviations from the flat spacetime metric that are solutions of the wave-equation *almost* solve the EFE in vacuum ($T^{ij} = 0$).

The statement could be generalized by perturbing around some general metric, say g_0, instead of η. Physically this might be more interesting, because the signal h that we want to measure is actually a perturbation of the metric already present due to earth.

[2]The Minkowski spacetime is represented by the matrix $\mathrm{diag}(-1, 1, \ldots, 1)$ in local orthonormal coordinates $(ct, x_1, \ldots, x_{n-1})$.

However, we will assume for simplicity that $g_0 = \eta$ so that we do not have to take into account any derivatives of g_0.

Another generalization would be to set $T^{ij} \neq 0$. However, this requires making assumptions about the energy and momentum characteristics of the source, which at the moment is not of our interest.

A prerequisite for stating proposition 3.2, is to give meaning to the notion of "closeness of metrics". We can view the EFE (3.17) as a differential operator which takes as input a two times continuously differentiable metric and returns a continuous metric. We introduce a norm on the space of metrics. For simplicity, we will consider the case in which $M = \mathbb{R}^n$. The metrics are defined on the entire manifold $\mathbb{R}^n$. We introduce for compact subsets $K \subset \mathbb{R}^n$ and k times continuously differentiable metrics

$$g : \mathbb{R}^n \to \mathrm{Sym}(n) \leq \mathbb{R}^{n\times n}$$

the norm

$$||g||_{C^k,K} = \sum_{i=0}^{k} ||D^i g||_K,$$

where $D^i g$ denotes the i^{th} derivative of g. For every $x \in K$ we have that $(D^i g)_x$ is a linear map, so by $||D^i g||$ we denote the operator norm of the differential, which takes on a maximum and minimum on K. The proposition is then the following statement.

Proposition 3.2 (Small metric perturbations)**.** *Let $\epsilon > 0$ and $K \subset \mathbb{R}^n$ a compact subset of $\mathbb{R}^n$. Then there exists an open (C^2, K)-neighborhood U of the Minkowski metric η in the space $\mathcal{M}$ of pseudo-metrics with $(1, n-1)$ signature, such that for all $g \in U$ we have*

$$||\Box(g-\eta)||_{C^0,K} \leq \epsilon ||g-\eta||_{C^2,K}$$

Proof. We define the set

$$\mathcal{S}(n) := \{n \times n \text{ real symmetric and invertible matrices of signature } (1, n-1)\},$$

and the map

$$f : \mathcal{S}(n) \to \mathcal{S}(n), \quad A \mapsto \sqrt{-\det(A)}A^{-1}, \tag{3.21}$$

so that $h = \eta - f \circ g$, with $g : \mathbb{R}^n \to \mathcal{S}(n)$. The extension of f to all invertible matrices is smooth, so the restriction to $\mathcal{S}(n)$ is also smooth. The EFE in (3.17) is expressed in terms of h (see (3.15) and (3.16)), but we want to make a statement about $g - \eta$. The derivatives of $f \circ g$ are given by

$$\begin{aligned} D(f\circ g) &= ((Df)\circ g)Dg \\ D^2(f\circ g) &= ((D^2 f)\circ g)(Dg, Dg) + ((Df)\circ g)D^2 g \end{aligned}$$

So for the d'Alembertian we have

$$\Box(f\circ g) = -((D^2 f)\circ g)(D_1 g, D_1 g) + \sum_{i=2}^{n} \left[((D^2 f)\circ g)(D_i g, D_i g)\right] + ((Df)\circ g)\Box g,$$

so since $\Box\eta = 0$, we have

$$\Box(g-\eta) = ((Df)\circ g)^{-1}\left(\Box(f\circ g) + ((D^2f)\circ g)(D_1g, D_1g) - \sum_{i=2}^{n}((D^2f)\circ g)(D_ig, D_ig)\right), \tag{3.22}$$

provided $(Df)\circ g$ is invertible. We will now first focus on bounding the d'Alembertian $\Box(f\circ g)$, and at the end of the proof use the relation above to make the conclusion we want.

In the Einstein Field Equations (3.17), $\Box(h) = \Box(f\circ g)$ is given as an expression in terms of other terms. We need to show that the $||.||_{C^0,K}$ norm of every term on the right hand side of (3.17) can be bounded by $C\epsilon||g-\eta||_{C^2,K}$ where $C > 0$ is a constant that only depends on K, n (dimension) and ϵ. By η we denote the constant function $\mathbb{R}^n \to \mathcal{S}(n)$, $x \mapsto \eta_M$, where η_M is just the diagonal matrix $\text{diag}(-1, 1, \ldots, 1)$.

For the proof we need the following. By continuity of f there exists $\tilde{\delta}_1 > 0$ such that

$$||f(A) - \eta_M|| = ||f(A) - f(\eta_M)| \leq \epsilon, \tag{3.23}$$

for all $A \in \overline{B_{\tilde{\delta}_1}(\eta_M)} \subset \mathcal{S}(n)$. By continuity of g, this $\tilde{\delta}_1$ correspond to a δ_1 such that all metrics $g \in B_{\delta_1}(\eta) \subset \mathcal{M}$ (measured with the $||.||_{C^2,K}$ norm) have the property that $g(x) \in \overline{B_{\tilde{\delta}_1}(\eta_M)}$.

We also need $(Df)_{g(x)}$ to be invertible. By the fact that Df is continuous and $(Df)_{\eta_M}$ is invertible, it follows that there exists $\tilde{\delta}_2$ such that $(Df)_A$ is invertible for all $A \in \overline{B_{\tilde{\delta}_2}(\eta_M)} \subset \mathcal{S}(n)$. By continuity of g, this $\tilde{\delta}_2$ corresponds to a δ_2 such that for all metrics $g \in B_{\delta_2}(\eta) \subset \mathcal{M}$ we have that $g(x) \in \overline{B_{\tilde{\delta}_2}(\eta_M)}$.

We choose the neighborhood U as the δ-ball

$$U := B_\delta(\eta) = \{g \in \mathcal{M} \,|\, ||g-\eta||_{C^2,K} < \delta\},$$

with

$$0 < \delta < \min\{1, \delta_1, \delta_2, \epsilon\}.$$

Also define $\tilde{\delta} = \min\{\tilde{\delta}_1, \tilde{\delta}_2\}$ and

$$Q := \overline{B_{\tilde{\delta}}(\eta_M)},$$

which is the compact subset of $\mathcal{S}(n)$. Since $\delta < \min\{\delta_1, \delta_2\}$, we know that $g(x) \in Q$ for all $g \in U$ and $x \in K$.

First of all, notice that all $\tilde{g}$'s in (3.20) can be replaced with $h = \eta - \tilde{g}$, because $\partial\tilde{g} = -\partial h$, and all $\partial\tilde{g}$ come in pairs. With this in mind, we plug in (3.19) and (3.20) into equation (3.17) to obtain

$$\begin{aligned}\Box h^{ij} = {} & \partial_m h^{ij}\partial_k h^{mk} - \partial_m h^{im}\partial_k h^{jk} + \partial_k h^{il}\partial_l h^{jk} - h^{kl}\partial_{kl}h^{ij} + \tfrac{1}{2}g^{ij}g_{mk}\partial_n h^{ml}\partial_l h^{kn} \\ & - g^{im}g_{kl}\partial_n h^{jl}\partial_m h^{kn} - g^{jm}g_{kl}\partial_n h^{il}\partial_m h^{kn} + g_{mk}g^{ln}\partial_l h^{im}\partial_n h^{jk} \\ & + \tfrac{1}{8}(2g^{im}g^{jk} - g^{ij}g^{mk})(2g_{ln}g_{st} - g_{ns}g_{lt})\partial_m h^{lt}\partial_k h^{ns}.\end{aligned} \tag{3.24}$$

One can now observe that all terms on the RHS of (3.24) are quadratic in h, ∂h and $\partial^2 h$, meaning that the right hand side of (3.24) "shrinks" as δ^2 while the neighborhood U shrinks as δ. Another remark for these terms, is that the index notation implies that each term is a *finite sum* of terms of the specified form. Since the sums are finite, it is sufficient to provide an upper bound for the norm of the *summand* of these terms and multiply this upper bound by the number of terms in the sum. For this reason, we will only bound the summands.

Then define the constants

$$
\begin{aligned}
C_1 &:= \max_{A\in Q}\{||(Df)_A||\},\\
C_2 &:= \max_{A\in Q}\{||(D^2f)_A||\},\\
C_3 &:= \max_{A\in Q}\{||A||\},\\
C_4 &:= \max_{A\in Q}\{||A^{-1}||\},\\
C_5 &:= \text{a positive real number such that } ||\partial_i g^{jk}||_{C^0,K} \leq C_3||Dg|| \text{ for all } 1\leq i,j,k\leq n,\\
C_6 &:= \text{a positive real number such that } ||\partial_{ij} g^{kl}||_{C^0,K} \leq C_4||D^2 g|| \text{ for all } 1\leq i,j,k,l\leq n,\\
C_7 &:= \max_{A\in Q}\{||((Df)_A)^{-1}||\}.
\end{aligned}
$$

where in the last line we used that $(Df)_A$ is invertible for all $A \in Q$ since $\delta < \delta_2$. We then have

$$||\partial_k h^{il}||_{C^0,K} =\leq C_5||D(f\circ g)||_{C^0,K} \leq C_1C_5\delta.$$

For the first term in (3.24), we then get

$$||\partial_m h^{ij}\partial_k h^{mk}||_{C^0,K} \leq C_1^2C_5^2\delta^2 \leq C\delta\epsilon,$$

where we defined $C = C_1^2C_5^2$ and used $\delta < \epsilon$, so it has the desired bound. Some terms contain contractions with the metric tensor, for instance the first term on the second line of (3.24). This term is bounded by

$$||g^{im}g_{kl}\partial_n h^{jl}\partial_m h^{kn}||_{C^0,K} \leq C_3C_4C_1^2C_5^2\delta^2 \leq C\delta\epsilon,$$

where we defined $C = C_3C_4C_1^2C_5^2$ and used $\delta < \epsilon$, so it has the desired bound. All other terms in (3.24) are treated similarly to the above two terms, except $h^{kl}\partial_{kl}h^{ij}$. For this term we have

$$||h^{kl}\partial_{kl}h^{ij}||_{C^0,K} \leq C_6||\eta - f\circ g||_{C^0,K}||D^2(\eta - f\circ g)||_{C^0,K}. \tag{3.25}$$

For the left term in (3.25) we have, since $\delta < \delta_1$, that we satisfy the continuity-condition (3.23), so

$$||\eta - f\circ g||_{C^0,K} \leq ||f\circ g - \eta||_{C^0,K} \leq \epsilon$$

For the right term in (3.25) we have

$$\begin{aligned}||D^2(\eta - f\circ g)||_{C^0,K} &= ||D(((Df)\circ g)Dg)||_{C^0,K}\\ &= ||(((D^2f)\circ g)Dg)Dg + ((Df)\circ g)D^2g||_{C^0,K}\\ &\leq C_2||Dg||^2_{C^0,K} + C_1||D^2g||_{C^0,K}\\ &\leq C_2\delta^2 + C_1\delta,\end{aligned}$$

So inequality (3.25) turns into

$$||h^{kl}\partial_{kl}h^{ij}||_{C^0,K} \leq C_6(C_2\delta + C_1)\epsilon\delta \leq C\epsilon\delta,$$

where we defined $C = C_6(C_2 + C_1)$ and used $\delta < \min\{1,\epsilon\}$, so it has the desired bound. We have now described the way to bound all terms on the right hand side of (3.24) in the form $C||g-\eta||_{C^2,K}$. Redefine C then as the constant such that $||\Box(\eta - f\circ g)||_{C^0,K} \leq C\epsilon||g-\eta||_{C^2,K}$. What we want now is to bound $\Box(g-\eta)$. Using (3.22), we have

$$\begin{aligned}&||\Box(g-\eta)||_{C^0,K}\\ &= \left|\left|((Df)\circ g)^{-1}\left(\Box(f\circ g) + ((D^2f)\circ g)(D_1g, D_1g) - \sum_{i=2}^{n}((D^2f)\circ g)(D_ig, D_ig)\right)\right|\right|_{C^0,K}\\ &\leq C_7(C\epsilon||g-\eta||_{C_2,K} + nC_2\delta^2)\\ &\leq C_7(C + nC_2)\epsilon||g-\eta||_{C_2,K}\end{aligned}$$

where we used that $(Df)\circ g)$ is invertible and bounded by C_7 and we used $\delta < \min\{1,\epsilon\}$ to bound $((D^2f)\circ g)(D_1g, D_1g)$.

Finally, we conclude that $\Box(g-\eta)$ indeed has the appropriate bound so that the proposition holds. □

There are a few comments to be made about proposition 3.2 besides the ones stated above the proposition statement. We have proved that the wave equation $\Box(g-\eta)$ is an arbitrarily good approximation to the flat spacetime solution of the Einstein Field Equations as long as the metric perturbation is small enough. In the Landau-Lifshitz form of the Einstein Field Equations (3.17), it is also required for a metric to satisfy the Lorenz gauge $\partial_i h^{ij}$ (3.18) in order for it to be a solution to the Einstein Field Equations. However, for proposition 3.2 it was not necessary to assume that the metrics satisfy this gauge. If one wants to argue that all perturbed metrics in the neighborhood U are indeed close to an actual solution of the EFE, this restriction should most likely be required.

A physicist would want to know what the *biggest* neighborhood U is, but this question is a lot more difficult. Furthermore, one should make sense of the numbers ϵ and δ physically, i.e., associate units to it, before any results about reality could be deduced. This is all beyond the scope of this thesis, and therefore we leave it at this (physically less interesting) result.

4 Gravitational Wave Solutions of Einstein Field Equations

As already stated, it is difficult to find analytical results to the EFE. Under the approximation that the metric is a weakly perturbed Minkowski metric

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \quad |h_{\mu\nu}| \ll 1, \tag{4.1}$$

it turns out that the EFE can be reduced to a wave equation for $h_{\mu\nu}$[3]. In 4.1, the units of 1 are taken to be the units such that $\eta_{\mu\nu} = (-1, 1, 1, 1)$. How to arrive at this wave equation will be covered in this chapter.

4.1 EFE for Linearized gravity

In this section we will linearize the EFE, which means that we neglect all terms that are of order $O(h^2)$ or higher, including derivatives of h. The derivation is mostly based on chapters 1.1 and 1.2 of Maggiore's "Gravitational Waves: Volume 1: Theory and Experiment" [9].

We make use of "small" coordinate transformations

$$x^\mu \mapsto x'^\mu := x^\mu + \xi^\mu(x^\nu) \tag{4.2}$$

that at most make alterations of order $O(h)$. So ξ is of the same order as h, but we will soon see that we also require that $\partial_\mu \xi^\nu$ is of order $O(h)$. The metric is changed under coordinate transformations by the equation

$$g_{\mu\nu}(x) \mapsto g'_{\mu\nu}(x') = \frac{\partial x^\rho}{\partial x'^\mu} \frac{\partial x^\sigma}{\partial x'^\nu} g_{\rho\sigma}(x). \tag{4.3}$$

We have

$$\begin{aligned} x^\mu &= x'^\mu - \xi^\mu(x^\nu) \\ x^\mu &= x'^\mu - \xi^\mu(x'^\nu - \xi^\nu(x^\rho)) \\ x^\mu &= x'^\mu - \xi^\mu(x'^\nu) + O(\xi^2) \quad \text{(Taylor expansion)}, \end{aligned}$$

so that

$$\frac{\partial x^\rho}{\partial x'^\mu} = \delta^\rho_\mu - \frac{\partial \xi^\rho}{\partial x'^\mu}, \tag{4.4}$$

where it does not matter whether we view ξ^ρ as a function of x^ν or x'^ν because

$$\frac{\partial \xi^\rho}{\partial x^\mu} = \frac{\partial \xi^\rho}{\partial x'^\nu} \frac{\partial x'^\nu}{\partial x^\mu} = \left(\delta^\nu_\mu + \frac{\partial \xi^\nu}{\partial x'^\mu}\right) \frac{\partial \xi^\rho}{\partial x'^\nu} = \frac{\partial \xi^\rho}{\partial x'^\mu} + O(h^2).$$

[3]From now on, we assume spacetime to be 4-dimensional and we use Greek letters for index notation. This distinguishes the more physics-styled approach that we will employ for the rest of the chapters from the more mathematical approach of earlier chapters.

Substituting (4.4) into (4.3), we find

$$\begin{aligned} g'_{\mu\nu}(x') &= \frac{\partial x^\rho}{\partial x'^\mu}\frac{\partial x^\sigma}{\partial x'^\nu} g_{\rho\sigma}(x) \\ &= (\delta^\rho_\mu - \frac{\partial \xi^\rho}{\partial x'^\mu})(\delta^\sigma_\nu - \frac{\partial \xi^\sigma}{\partial x'^\nu})(\eta_{\sigma\rho} + h_{\sigma\rho}) \\ &= \delta^\rho_\mu \delta^\sigma_\nu (\eta_{\sigma\rho} + h_{\sigma\rho}) - \delta^\sigma_\nu \frac{\partial \xi^\rho}{\partial x'^\mu}\eta_{\rho\sigma}(x) - \delta^\rho_\mu \frac{\partial \xi^\sigma}{\partial x'^\nu}\eta_{\rho\sigma}(x) + O(h^2) \\ &= g_{\mu\nu}(x) - \frac{\partial \xi_\nu}{\partial x'^\mu} - \frac{\partial \xi_\mu}{\partial x'^\nu} + O(h^2), \end{aligned}$$

so we find for the metric perturbation under small transformations

$$h'_{\mu\nu} = h_{\mu\nu} - \frac{\partial \xi_\nu}{\partial x'^\mu} - \frac{\partial \xi_\mu}{\partial x'^\nu} = h_{\mu\nu} - \partial_\mu \xi_\nu - \partial_\nu \xi_\mu. \tag{4.5}$$

Note that we have been raising and lowering indices of ξ using the flat spacetime metric η, which is allowed because

$$g_{\mu\nu}\xi^\nu - \eta_{\mu\nu}\xi^n = h_{\mu\nu}\xi^\nu = O(h^2).$$

In earlier chapters we already saw the Riemann curvature tensor (2.9) expressed in terms of the connection (2.4). Locally, the connection is given by Christoffel symbols $\Gamma^\rho_{\mu\nu}$ which can be expressed in terms of the metric as

$$\Gamma^\rho_{\mu\nu} = \tfrac{1}{2} g^{\rho\lambda}(\partial_\mu g_{\lambda\nu} + \partial_\nu g_{\mu\lambda} - \partial_\lambda g_{\mu\nu}),$$

and, one can check, the curvature tensor is then locally given by

$$R^\rho_{\sigma\mu\nu} = \partial_\mu \Gamma^\rho_{\nu\sigma} - \partial_\nu \Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda}\Gamma^\lambda_{\nu\sigma} - \Gamma^\rho_{\nu\lambda}\Gamma^\lambda_{\mu\sigma}.$$

By using the identity $g^{\mu\nu}g_{\nu\rho} = \delta^\mu_\rho$, it is easy to see that to first order in h we have for the inverse metric

$$g^{\mu\nu} = \eta^{\mu\nu} - h^{\mu\nu},$$

so that the Christoffel symbols become

$$\Gamma^\rho_{\mu\nu} = \tfrac{1}{2}(\partial_\mu h^\rho_\nu + \partial_\nu h^\rho_\mu - \partial^\rho h_{\mu\nu}),$$

and so the linearized curvature tensor becomes

$$R^\rho_{\sigma\mu\nu} = \tfrac{1}{2}(\partial_\mu \partial_\sigma h^\rho_\nu - \partial_\mu \partial^\rho h_{\mu\nu} - \partial_\nu \partial_\sigma h^\rho_\mu + \partial_\nu \partial^\rho h_{\mu\sigma}).$$

From the curvature tensor, we obtain the Ricci tensor $R_{\mu\nu}$ and Ricci scalar R

$$\begin{aligned} R_{\mu\nu} &= \tfrac{1}{2}(-\Box h_{\mu\nu} + \partial_\mu \partial_\rho h^\rho_\nu + \partial_\nu \partial^\rho h_{\rho\mu} - \partial_\mu \partial_\nu h) \\ R &= -\Box h + \partial_\mu \partial_\nu h^{\mu\nu}, \end{aligned}$$

where we defined $\Box := \eta^{\mu\nu}\partial_\mu\partial_\nu$ as the flat-spacetime d'Alembertian and $h = h^\mu_\mu$ as the trace of the metric. We now have all the ingredients to plug into the left-hand side of the EFE

$$R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R = \frac{8\pi G}{c^4}T_{\mu\nu}.$$

We first define the trace-reversed metric perturbation $\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h$, so that

$$\begin{aligned} R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R &= R_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}R \\ &= \tfrac{1}{2}(-\Box h_{\mu\nu} + \partial_\mu\partial_\rho h^\rho_\nu + \partial_\nu\partial^\rho h_{\rho\mu} - \partial_\mu\partial_\nu h) - \tfrac{1}{2}\eta_{\mu\nu}(-\Box h + \partial_\rho\partial_\lambda h^{\rho\lambda}) \\ &= -\tfrac{1}{2}(\Box\bar{h}_{\mu\nu} - \partial_\mu\partial_\rho\bar{h}^\rho_\nu - \partial_\nu\partial_\rho\bar{h}^\rho_\mu + \eta_{\mu\nu}\partial_\rho\partial_\lambda\bar{h}^{\rho\lambda}). \end{aligned} \tag{4.6}$$

The amount of terms can be reduced even further when imposing the the Lorenz gauge on $\bar{h}$, meaning

$$\partial^\mu\bar{h}_{\nu\mu} = 0. \tag{4.7}$$

Of course, we have to show that it is possible that this condition can be imposed, i.e., we need to show that there exists a small ξ^μ such that we obtain (4.7) for the transformed perturbation $\bar{h}'$. From (4.5) and the definition of the trace-reversed metric, we find that under a gauge transformation

$$\bar{h}'_{\mu\nu} = \bar{h}_{\mu\nu} - \partial_\mu\xi_\nu - \partial_\nu\xi_\mu + \eta_{\mu\nu}\partial_\rho\xi^\rho,$$

from which it follows that

$$\partial^\nu\bar{h}'_{\mu\nu} = \partial^\nu\bar{h}_{\mu\nu} - \Box\xi_\mu,$$

so that the Lorenz gauge is satisfied when

$$\Box\xi_\mu = \partial^\nu\bar{h}_{\mu\nu}.$$

Such a ξ_μ can always be found. Consider for instance a Green's function $G(x)$ for the d'Alembertian operator, so that $\Box_x G(x-y) = \delta^4(x-y)$, then we obtain

$$\xi_\mu(x) = \int G(x-y)\partial^\nu\bar{h}_{\mu\nu}d^4x.$$

In this gauge, three terms vanish from the left-hand side of the Einstein field equations for the trace-reversed metric (4.6), so that the EFE take the simple form[4]

$$\Box\bar{h}_{\mu\nu} = -\frac{16\pi G}{c^4}T_{\mu\nu}. \tag{4.8}$$

[4]Note that in the third chapter, equation (3.17), we already saw the EFE written a form very similar to this, except that we kept all of the higher order terms in h as well. In proposition 3.2, we proved mathematically that small enough metric perturbations solve the EFE to arbitrary accuracy.

The metric perturbation can be recovered by

$$h_{\mu\nu} = \bar{h}_{\mu\nu} - \tfrac{1}{2}\eta_{\mu\nu}\bar{h}.$$

Equation (4.8) makes it clear that wave-like solutions exists for the metric perturbation, which is the phenomenon that we call gravitational waves. There are some comments to be made about the number of degrees of freedom that is implied by equation (4.8). First of all, the metric is a symmetric $(0, 2)$ tensor for a four-dimensional space, so that on first glance it has ten independent components, i.e., degrees of freedom. Fixing a coordinate system, like we do in (4.7), reduces this to six degrees of freedom. It turns out, however, that far away from any source terms ($T_{\mu\nu} = 0$), the number of degrees of freedom reduces even further. This is accomplished by making yet another gauge transformation ξ^μ that satisfies

$$\Box\xi^\mu = 0, \quad |\xi^\mu| \ll 1,$$

so that the conditions $\bar{h} = 0$ and $\bar{h}_{0\nu} = 0$ are satisfied. This is called the transverse-traceless (TT) gauge. For a detailed description, see chapter 1.2 of Maggiore [9]. The main result that we want to emphasize here is that the metric perturbation takes the form

$$h_{ij}^{TT}(t, z) = \begin{pmatrix} h_+ & h_\times & 0 \\ h_\times & -h_+ & 0 \\ 0 & 0 & 0 \end{pmatrix} \cos(\omega(t - \tfrac{z}{c})), \tag{4.9}$$

where i, j are spacial indices, h_+ and $h_\times$ are called the "plus" and "cross" polarizations respectively, and ω is the wave frequency.

4.2 Further approximation schemes

In 4.1, we discussed linear gravity in which the EFE assume the simple form of a wave equation. Though providing insight in the behavior of gravitational waves, the applicability in the field of gravitational wave detection is very limited due to its poor accuracy. The accuracy can be improved by considering higher and higher orders of h and its derivatives. The idea is that all terms on the left hand side of the EFE can be (formally) expanded in Newton's gravitational constant G. This method is called the *post-Minkowskian expansion.*

Further approximations can be made to the right hand side of the EFE by formally expanding $T_{\mu\nu}$ in powers of v/c, where v is the characteristic velocity of the matter. This method is called the *post-Newtonian expansion.*

How to apply these methods to obtain more accurate formulas for the metric perturbation $h_{\mu\nu}$ is beyond the scope of this work, but they are described in some detail in ([11], Chapters 6-9) and ([9], Chapter 5).

5 A Machine Learning Method for Waveform Generation

Fast waveform (WF) generation is crucial for both the detection of gravitational waves and for inferring the physical properties of the source. Currently, waveforms are generated by relatively slow methods. In this chapter we provide a method for fast and accurate waveform generation describing the full inspiral of a non-precessing binary black hole system. We will use several neural networks (NN) that take as input the relevant variables of such a system and that together produce as output the corresponding waveform. We build upon the `mlgw` package published in 2021 by Stefano Schmidt [12]. Naturally, we call our approach `mlgw-NN`. The type of models that we build are called *surrogates*, which are models that try to approximate the behavior of some other (more complex) model.

The quantity that we want to predict is the metric perturbation, or *strain*, which can be decomposed into modes H_{lm} as

$$h(t; d, \iota, \varphi, \boldsymbol{\theta}) = h_+ + ih_\times = \frac{G}{c^2}\frac{M}{d}\sum_{l=2}^{\infty}\sum_{m=-l}^{l} {}_{-2}Y_{lm}(\iota, \varphi)H_{lm}(t/M, \bar{\boldsymbol{\theta}}), \tag{5.1}$$

as given in equation (1) from [12]. The variables are explained further below. The NNs are trained on waveforms produced by the `IMRPhenomTPHM` model [3]. Separate models are made for the $(2,2), (2,1), (3,3), (4,4)$ and $(5,5)$ modes, where the tuples are formatted as (l, m). In order to make predictions for $h(t; ...)$ using machine learning methods, we must first represent WFs on a discrete time grid. This is covered in 5.1.

Before moving to NNs, we recapitulate the method given in the `mlgw`-paper [12] and we make adjustments where necessary. The steps for training an `mlgw-NN` model can be summarized as follows

1. Choosing a time grid,
2. Creating a dataset, covered in 5.2
3. Reducing the dimensionality,
4. Training the neural networks, covered in 5.3

5.1 Waveform representation

As a function of time, WFs can be represented by the formula

$$h(t, \boldsymbol{\theta}) = A(t, \boldsymbol{\theta})e^{i\phi(t,\boldsymbol{\theta})}, \tag{5.2}$$

where $A(t, \boldsymbol{\theta})$ and $\phi(t, \boldsymbol{\theta})$, respectively, describe the amplitude and phase of the waveform as a function of time and the system parameters. The decomposition in amplitude and phase (5.2) applies to the individual modes $H_{lm}(t, \bar{\boldsymbol{\theta}})$ as well. The parameters contained in the $\boldsymbol{\theta}$ vector are $\boldsymbol{\theta} = (m_1, m_2, \boldsymbol{s}_1, \boldsymbol{s}_2)$, with m_i the masses of the black holes and $\boldsymbol{s}_i$ the

spin vectors of the black holes. The full WF in the form of (5.1) is also dependent on d the distance from the source, and ι (polar angle) and φ (azimuthal angle) the polar angles of the orbital plane, but these are irrelevant for training. We will assume that $m_1 \geq m_2$, and introduce the mass ratio $q := m_1/m_2$ and total mass $M := m_1 + m_2$. In this work, we restrict the spins to be aligned with the orbital plane, i.e., $\boldsymbol{s}_i = (0, 0, s_{i,z}) =: (0, 0, s_i)$.

Furthermore, we need to specify a discrete *time grid* in order to generate WFs of the form (5.2). This means that we specify a set of D points $\mathbf{t} = \{t_1, \ldots, t_D\}$ for which A and ϕ will be calculated in a way that a sufficiently large part of the inspiral is captured. The amplitude and phase can then be represented by D-dimensional vectors

$$A_{\boldsymbol{\theta}}(\mathbf{t}) = \big(A_{\boldsymbol{\theta}}(t_1), \ldots, A_{\boldsymbol{\theta}}(t_D)\big) \quad \text{and} \quad \phi_{\boldsymbol{\theta}}(\mathbf{t}) = \big(\phi_{\boldsymbol{\theta}}(t_1), \ldots, \phi_{\boldsymbol{\theta}}(t_D)\big)\,. \tag{5.3}$$

In order to obtain a "smooth" function for amplitude and phase, the vectors should be (non-linearly) interpolated on a more dense time grid.

5.2 Preparing the data

The goal is to "predict" the amplitude and phase functions

$$\boldsymbol{\theta} \mapsto A_{\boldsymbol{\theta}}(\mathbf{t}), \tag{5.4a}$$

$$\boldsymbol{\theta} \mapsto \phi_{\boldsymbol{\theta}}(\mathbf{t}), \tag{5.4b}$$

corresponding to the different modes H_{lm}. For training purposes we restrict the parameters to $\bar{\boldsymbol{\theta}} = (q, s_1, s_2)$. This is justified because the dependence of the WF on ι and φ is only through the spherical harmonics ${}_{-2}Y_{lm}(\iota, \varphi)$. Also, the distance d will be set to 1 Mpc as it is only a scaling factor of the amplitude. We fix the total mass $M = 20\,\mathrm{M}_\odot$ as total mass only scales the time grid. The functions (5.4a,5.4b), for each mode separately, are predicted by several neural networks trained on a *training (data)set*.

Choosing a time grid

Because the length of the time grid scales linearly with total mass, it is convenient to introduce the dimensionless reduced time $\tau = t/M$. After rescaling the time grid (so $\mathbf{t} \to \boldsymbol{\tau}$), the amplitude and phase vectors are described by $A(\boldsymbol{\tau}, \bar{\boldsymbol{\theta}})$ and $\phi(\boldsymbol{\tau}, \bar{\boldsymbol{\theta}})$, where $\bar{\boldsymbol{\theta}} = (q, s_1, s_2)$, because the time grid is linearly scaled by total mass M. The actual predicted functions are then given by

$$\bar{\boldsymbol{\theta}} \mapsto A_{\bar{\boldsymbol{\theta}}}(\boldsymbol{\tau}), \tag{5.5a}$$

$$\bar{\boldsymbol{\theta}} \mapsto \phi_{\bar{\boldsymbol{\theta}}}(\boldsymbol{\tau}). \tag{5.5b}$$

Time grids are shifted so that the peak of the amplitude function $A_{\bar{\boldsymbol{\theta}}}(\boldsymbol{\tau})$ corresponding to the $(2, 2)$ mode is located at $\tau = 0$. The amplitude of a gravitational wave varies quickly around its peak but slowly at the early stages of the inspiral. Hence, a time grid

with equally spaced points is not suitable. This problem is solved by transforming the equally-spaced time grid by

$$\tau_i \mapsto \text{sign}(\tau_i) \cdot |\tau_i|^{\frac{1}{\alpha}} \quad i \in \{1, \dots, D\}, \tag{5.6}$$

where $0 < \alpha < 1$ is the *distortion parameter*. We also define the minimum of the transformed reduced time grid $\boldsymbol{\tau}$ as $\tau_{\min} = -\tau_1 > 0$. By this definition the actual WF starts at $\tau_{\min} M$ before the merger. The parameters α and $\tau_{\min}$ are called *hyperparameters*. Hyperparameters have nothing to do with the physics of the problem (like mass), but rather influence the training of our model.

Creating a dataset

The neural networks are trained on N WFs generated by the `IMRPhenomTPHM` method. The training set is defined by the $(N, 3 + 2D)$ dimensional matrix X where the rows are formatted as

$$X_i = (q, s_1, s_2, A^T_{\bar{\boldsymbol{\theta}}}(\boldsymbol{\tau}), \phi^T_{\bar{\boldsymbol{\theta}}}(\boldsymbol{\tau})).$$

The parameters (q, s_1, s_2) are sampled from a, not necessarily uniform, distribution. The parameter range

$$\mathcal{P} = (q_{\min}, q_{\max}) \times (s_{1,\min}, s_{1,\max}) \times (s_{2,\min}, s_{2,\max}) \tag{5.7}$$

should be specified so that the method used for creating the training set of WFs, in our case `IMRPhenomTPHM` WFs, is also valid on this region.

To predict a smooth function for the phase as a function of $\bar{\boldsymbol{\theta}}$, the phase vectors $\phi_{\bar{\boldsymbol{\theta}}}(\boldsymbol{\tau})$ are shifted so that at $\bar{\tau}$ we have $\phi_{\bar{\boldsymbol{\theta}}}(\bar{\tau}) = 0$. Since $\tau = 0$ is always in the time grid and model performance is not dependent on the value of $\bar{\tau}$, we conveniently set $\bar{\tau} = 0$.

Reducing the dimensionality

For computational reasons it is not feasible to predict all of the D components of the amplitude and phase vectors. Fortunately, it turns out that it is not required to predict all of the D values. The idea is that we first apply a linear transformation on the amplitude and phase vectors to a lower dimensional space of dimension K. Then, we predict the K-dimensional vectors and finally we apply the (approximately) inverse projection to reconstruct the D-dimensional vectors.
Just like `mlgw`, we use Principal Component Analysis (PCA) for the dimensionality reduction (see chapter 12.2 of [10] for details). For readability and generality, we introduce the vector $\mathbf{f} \in \mathbb{R}^D$ which in our case represents either $A_{\bar{\boldsymbol{\theta}}}(\boldsymbol{\tau})$ or $\phi_{\bar{\boldsymbol{\theta}}}(\boldsymbol{\tau})$. We also define the empirical mean

$$\boldsymbol{\mu} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{f}_i$$

and the empirical covariance matrix

$$\Sigma = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{f}_i - \boldsymbol{\mu})(\mathbf{f}_i - \boldsymbol{\mu})^T,$$

which is a $D \times D$ matrix. The objective of the PCA method is to find an orthonormal basis $V = \{\mathbf{v}_1, \ldots, \mathbf{v}_K \,|\, \mathbf{v}_i \in \mathbb{R}^D\}$ together with coefficients $Z = \{z_i, \ldots, z_K \,|\, z_i \in \mathbb{R}\}$, so that the (averaged) reconstruction error

$$E(V, Z) = \frac{1}{N}\sum_{i=1}^{N} ||\mathbf{f}_i - (H^T(H(\mathbf{f}_i - \boldsymbol{\mu})) + \boldsymbol{\mu})||, \tag{5.8}$$

is minimized, where H is the $K \times D$ matrix containing the vectors $z_i\mathbf{v}_i$ as rows. In theorem 12.2.1 of [10], it is proved that this objective is accomplished by choosing z_i to be the K largest eigenvalues[5] of Σ and choosing $\mathbf{v}_i$ to be the eigenvectors of Σ corresponding to z_i. We call the vectors $z_i\mathbf{v}_i$ principal components (PCs)[6]. So, the D-dimensional vectors $\mathbf{f}_i$ are projected to the K-dimensional space by

$$\mathbf{f}_i \mapsto \mathbf{g}_i = H(\mathbf{f}_i - \boldsymbol{\mu}), \tag{5.9}$$

and the reduced vectors $\mathbf{g}_i$ are transformed back to the D-dimensional space by

$$\mathbf{g}_i \mapsto \hat{\mathbf{f}}_i = H^T\mathbf{g}_i + \boldsymbol{\mu}. \tag{5.10}$$

The dimension K is an important hyperparameter that has an impact on both the speed and the accuracy of the predicted WFs compared to the true WFs. The average reconstruction error (5.8) decreases monotonically as K increases. If K is too small, the predictions $\hat{A}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\tau})$ and $\hat{\phi}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\tau})$ are quick but the WFs are inaccurate due to the large reconstruction error. If K is too large, the WFs are (possibly) predicted accurately but the predictions might be slow due to the increase in computations.

5.3 Training neural networks

After the first three steps, we are in the position to create neural networks to predict the functions (5.5a) and (5.5b). For our purpose, we will be using feedforward neural networks, or multi-layer perceptrons.

[5]The matrix Σ, which consists of real numbers, is symmetric and semi-positive definite, therefore its eigenvalues are real and non-negative.

[6]Usually, only the vectors $\mathbf{v}_i$ are called principal components. In our case it is more convenient to give this name to the scaled version of the vectors with the corresponding eigenvalues.

Structure of a NN

A NN is a machine learning tool, originally inspired by the workings of neurons in a brain, that is able to predict non-linear functions. NNs consist of *layers*, *nodes* and *edges*. The layers can be classified as the input layer, hidden layers and the output layer. Each layer consists of nodes and the nodes of neighboring layers are connected by edges that have weights. The values that the nodes of a layer take on are computed by applying an *activation function* on the sum of the nodes, weighted by the edges, of the previous layer. The values of the nodes are "fed forward". By providing the NN with inputs, i.e., setting the values of the nodes of the input layer, the information is fed forward through the network until it reaches the output layer. In our case, the outputs correspond to the coefficients of the PCs.

In 1991 it was proved by Hornik that NNs are "universal approximators" for continuous functions [6]. This means that continuous functions can be approximated by a NN to arbitrary accuracy, given enough hidden nodes. In theory this makes NNs a powerful tool for our problem, as we expect the amplitude and phase to be (at least) continuous functions of time.

Training of a NN

A NN is trained on a labeled *training set* $T = \{(\mathbf{x}_1, \mathbf{y}_1), \ldots, (\mathbf{x}_N, \mathbf{y}_N)\}$ consisting of N pairs of the form $(\mathbf{x}_i, \mathbf{y}_i)$ corresponding to the input and output of some sample. Explaining the full training scheme is beyond the scope of this work, but it is important to have the conceptual idea in mind. The NN can be thought of as a function

$$f : X \to Y,\ (\mathbf{x}; w_1, w_2, \ldots, w_L) \mapsto \hat{\mathbf{y}},$$

where $\mathbf{x}$ is a vector in the input space X, $\hat{\mathbf{y}}$ is the *predicted* vector in the output space Y and w_i are the weights of the edges connecting nodes of consecutive layers. The performance of the NN can be measured on the entire training set by the mean squared error

$$\mathrm{MSE}(f, T) = \frac{1}{N} \sum_{i=1}^{N} ||f(\mathbf{x}_i; w_1, \ldots, w_L) - \mathbf{y}_i||^2 = \frac{1}{N} \sum_{i=1}^{N} ||\hat{\mathbf{y}}_i - \mathbf{y}_i||^2,$$

where $||.||$ is the Euclidean norm. Essentially, we then calculate the *gradient* $\boldsymbol{\nabla}_{\mathbf{w}}\mathrm{MSE}(f, T)$ of $\mathrm{MSE}(f, T)$ with respect to the vector containing all of the weights and then replace the weights according to $\mathbf{w} \to \mathbf{w} - c\boldsymbol{\nabla}_{\mathbf{w}}\mathrm{MSE}(f, T)$ where c is a small positive constant. This algorithm is called *gradient descent*. After updating the weights, the MSE will (generally) be smaller for the next iteration through all of the training data. This process is then repeated a number of times, called *epochs*, until the performance stagnates.

Again, the above description is an oversimplification. The key take-away is that we repeatedly update the weights in such a way to reduce the error of our predicted values $\hat{\mathbf{y}}$ with the acutal values $\mathbf{y}$.

Concrete set-up

The `mlgw-NN` approach uses *Sequential* models from the *Keras API*, which is an interface for the *TensorFlow* library [7]. The functions that our NNs will predict are

$$\bar{\boldsymbol{\theta}} \mapsto \tilde{A}_{\bar{\boldsymbol{\theta}}}(\tau), \tag{5.11a}$$

$$\bar{\boldsymbol{\theta}} \mapsto \tilde{\phi}_{\bar{\boldsymbol{\theta}}}(\tau). \tag{5.11b}$$

The vector $\tilde{A}_{\bar{\boldsymbol{\theta}}}(\tau)$ is the K-dimensional representation of the full D-dimensional amplitude vector $\hat{A}_{\bar{\boldsymbol{\theta}}}(\tau)$, which can be obtained by applying the inverse PCA transformation (5.10). Of course, the same applies for $\tilde{\phi}_{\bar{\boldsymbol{\theta}}}(\tau)$.

The input layer has three nodes for the parameters $\bar{\theta}$. Additionally, we can add *features*, which are non-linear combinations of q, s_1 and s_2, to the dataset so that the input layer also has more modes. Examples of features commonly used are

$$q^2, \quad qs_1, \quad \text{(reduced) chirp mass } \mathcal{M}_c = \left(\frac{q}{(1+q^2)}\right)^{\frac{3}{5}}, \quad \text{or effective spin } \chi = \frac{s_1 + qs_2}{1+q}. \tag{5.12}$$

In theory, these features are not needed to predict the functions (5.11a,5.11b) because NNs can approximate any continuous function to arbitrary accuracy, as stated earlier. In practice, however, features appear to increase the performance of the NNs and to decrease the time it takes to train the NNs.

Each node of the output layer predicts one of the components of the K-dimensional vectors $\tilde{A}_{\bar{\boldsymbol{\theta}}}(\tau))$ or $\tilde{\phi}_{\bar{\boldsymbol{\theta}}}(\tau)$. The output layer of a NN consists of K_j nodes. We use a subscript j, because it can be useful to create several NNs predicting different PCs. This method comes at the expense of speed, but the increase in accuracy might outweigh this consideration. The structure of the NN is summarized in Fig. 1.

The number of hidden layers, the number of nodes per hidden layer and the activation functions between layers are all hyperparameters of the NN. Additional hyperparamters for the training phase of the NN are the number of epochs, the batch size and the optimizer. The batch size determines how many training samples are put through the NN during training at once. The optimizer is the algorithm that handles the updates of the weights. The most important consideration to take into account when setting these hyperparameters is the performance of the NN, i.e., the MSE of each PC should be as small as possible. Other considerations are training speed and computational time for predictions.

Tricks: There are three important "tricks" that we employ to improve the performance of the NNs.
We call the first of these tricks the *weighted MSE loss*. The idea is that higher order PCs are more difficult to learn than the lower order PCs. This means that the MSE of the first PC *should* be the smallest, then the second PC, and so on. However, when during training the loss function of the NN weighs the loss of all PCs equally, the MSEs of the

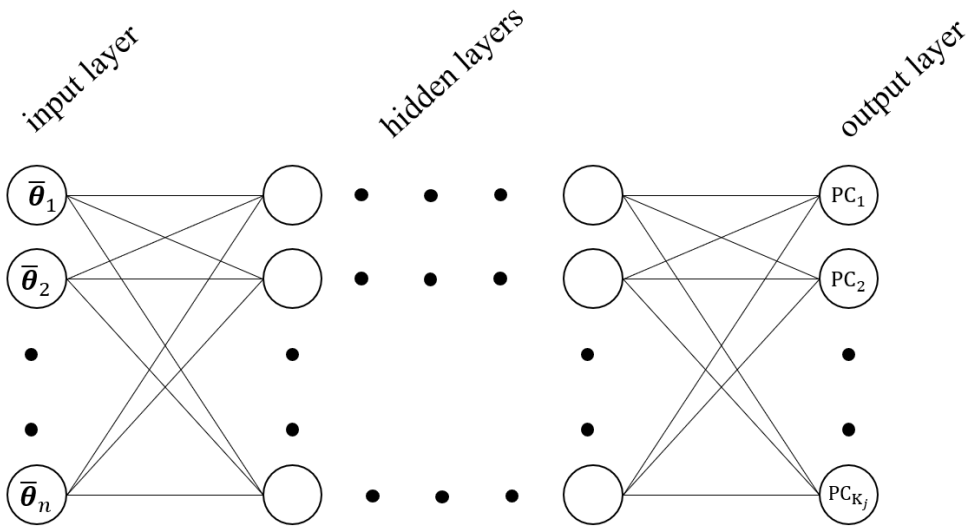


Figure 1: Neural network predicting K_j PCs of the vector $\tilde{A}_{\bar{\boldsymbol{\theta}}}(\tau)$ or $\tilde{\phi}_{\bar{\boldsymbol{\theta}}}(\tau)$. The model takes in n parameters. These parameters include q, s_1 and s_2, but also added features such as described in (5.12).

first few PCs are often not as low as they could be. This issue is resolved by *weighing* the MSE of the lower order PCs more heavily than higher order PCs. A weighted mean squared error MSE_ω can be defined in various ways, we will use

$$\text{MSE}_\omega(\mathbf{y}, \hat{\mathbf{y}}) = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{K}(\omega_j(y_i^j - \hat{y}_i^j))^2,$$

where ω_j are the weights and y_i^j is the j^{th} component of the i^{th} training sample.

The second trick we call *residual models*, which are also based on the principle that lower order PCs can be predicted more accurately than high order PCs. The idea is that we have made a NN which predicts the function $\bar{\boldsymbol{\theta}} \mapsto f(\bar{\boldsymbol{\theta}})$. Then we calculate the *residual of the dataset corresponding to the NN* which consists of the data points $\{f_1 - \hat{f}_1, \ldots, f_N - \hat{f}_N\}$. Then, we train a new NN on the residual dataset. The final predictions are given by

$$\hat{f} = \text{NN}(\bar{\boldsymbol{\theta}}) + \text{NN}_{\text{res}}(\bar{\boldsymbol{\theta}}). \tag{5.13}$$

For our purpose, residual models have shown to significantly improve the accuracy of the lower order PCs, though we do not precisely understand why the performance of the *non*-residual model is not able to capture the full behavior.

The third trick is to exponentially decrease the learning rate during training. This is to make it more likely that the NNs converge to a (local) minimum.

6 Model Results

We now exhibit the performance of an `mlgw-NN` model created by the method of chapter 5. We will assess the accuracy of the `mlgw-NN` model for both the individual modes H_{lm} for $(l,m) \in \{(2,2),(2,1),(3,3),(4,4),(5,5)\}$ and the full WFs. Then we determine the speed-up of `mlgw-NN` compared to `IMRPhenomTPHM` for full WF generation.

For each mode, we use the the following structure of NNs. We predict the first 4 PCs of the amplitude with one NN. For the phase we use three NNs: one for the first 2 PCs, one for PCs 3, 4, 5 and 6, and a residual model for the first 2 PCs. The number of hidden layers and nodes per hidden layer of the NNs vary from 3-6 and 5-30 respectively. In Appendix A we provide the complete description of all of the hyperparameters of the NNs. This structure is mostly based on (sophisticated) trial and error rather than a rigorous analysis of the hyperparameters of the NNs. In Table 1, we state the hyperparameters used for the amplitude model of the $(2,2)$ mode.

Table 1: Hyperparameters for amplitude NN of the $(2,2)$ mode. See Appendix A for the complete description of all hyperparameters of all NNs.

Hyperparameter	value
Principal components	4
MSE weights	200,40,3,3
Nodes per hidden layer	30,20,20,10,10
Activation function (except last layer)	Sigmoid
Initial learning rate, decay factor γ	0.001, 0.0005
Batch size	64

The NNs for each mode are trained on a training sets of 16000 WFs generated by the `IMRPhenomTPHM`-method with parameters q, s_1 and s_2 sampled from the domain

$$\mathcal{P} = [1,10] \times [-0.9,0.9] \times [-0.9,0.9]. \tag{6.1}$$

The sampling for s_1 and s_2 is uniform, and q was sampled from a distribution heavily biased towards 1 and slightly towards 10 (see Fig. 2). During our research, a distribution for q biased towards the boundary of its sampling interval showed to significantly improve the performance of the NNs in these regions. For the training datasets we use $D = 2000$ grid points, we set the distortion parameter $\alpha = 0.5$ and we set $\tau_{\min} = 2\,\mathrm{s}/\mathrm{M}_\odot$. The hyperparameters D and α were chosen based on the analysis in the `mlgw`-paper [12].

The model's performance is evaluated against test WFs generated with a starting frequency of 15 Hz and total mass varying in the range $[13.4..., 20.3...]$[7]. For full WF

[7]These values are based on an approximate relation between $\tau_{\min}$ and $f_{\min}$, namely $f_{\min} = 151\,\mathrm{Hz}\left(\frac{(1+q)^2}{q}\right)^{3/8}\left(\frac{\mathrm{M}_\odot}{\mathrm{M}}\right)\left(\frac{1\,\frac{\mathrm{s}}{\mathrm{M}_\odot}}{\tau_{\min}}\right)$ (Eq. (9) in [12]). WFs can be made longer by varying the total mass as a function of q. This way, we make use of a larger time-grid for which our model is able to make predictions

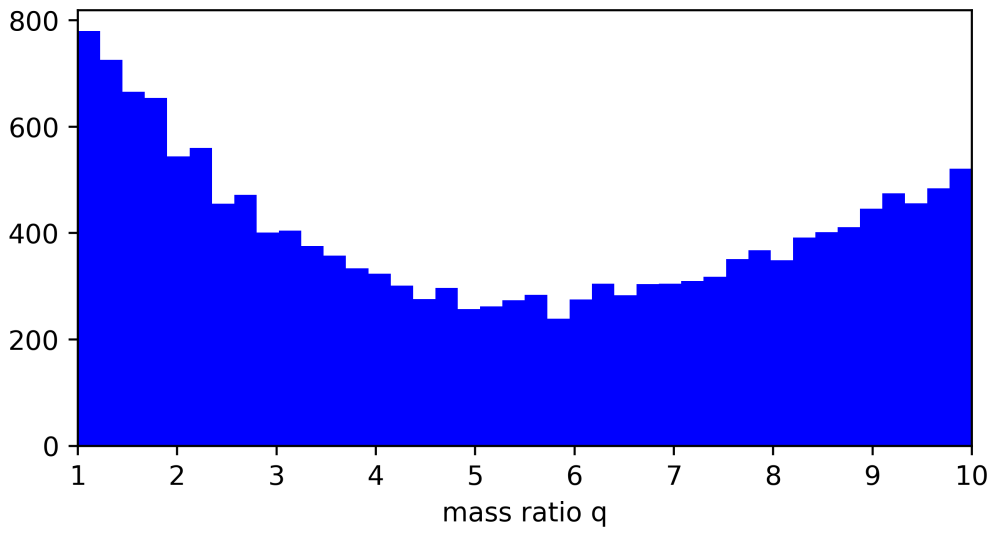


Figure 2: The biased distribution from which the 16000 training values of q are sampled in the interval $[1, 10]$. The training values for spins s_1 and s_2 are both sampled uniformly from the intervals $[-0.9, 0.9]$.

testing we use a set of 10000 WFs and for individual mode testing we use a set of 4000 WFs. For both datasets of test WFs, the parameters (q, s_1, s_2) are uniformly sampled from the training range $\mathcal{P}$ (6.1). Additionally for the full WFs, the variable φ is sampled uniformly from the range $[0, 2\pi]$ and $\cos(\iota)$ is sampled uniformly from the interval $[-1, 1]$. For a complete overview of the test WFs, see Table 2.

6.1 Accuracy

The accuracy of a the **modes** can be quantified by the *optimal mismatch* $\mathcal{F}$, defined by

$$\mathcal{F}(h_1, h_2) = \min_{\Delta\phi,\, \Delta t} \left(1 - \frac{\mathfrak{Re}(\langle h_1, h_2 e^{i\Delta\phi} \rangle)}{\sqrt{\langle h_1, h_1 \rangle \langle h_2, h_2 \rangle}} \right), \tag{6.2}$$

where $\langle\rangle$ is defined by the Wiener product

$$\langle h_1, h_2 \rangle = 4 \int_0^\infty df \frac{\bar{h}_1^*(f) \bar{h}_2(f)}{S_n(f)}, \tag{6.3}$$

where $\bar{h}$ denotes the Fourier transformed of h, the * denotes complex conjugation and $S_n(f)$ is the detector noise (power spectral density). We minimize over a constant phase shift $\Delta\phi$ and a constant time shift Δt, such that $h_1(t)$ and $h_2(t + \Delta t)e^{i\Delta\phi}$ are optimally aligned (for readability we omit the time argument of h). We assume the noise is "flat",

Table 2: Description of the parameters (first column) for the datasets for testing performance of individual modes (second column) and full WFs (third column). For the modes, datasets are created for $(l, m) \in \{(2,2), (2,1), (3,3), (4,4), (5,5)\}$. The parameters q, s_1 and s_2 for test WFs are sampled uniformly from the range $\mathcal{P} = [1, 10] \times [-0.9, 0.9] \times [-0.9, 0.9]$ for the modes and for full WFs. For full WFs we also uniformly sample the angles corresponding to the orbital plane ϕ and ι, and we sample the angles corresponding to sky-location ϕ_s and θ_s, and polarization of the GW ψ. For both modes and full WFs, the total mass M was varied in the range $[13.4..., 20.3...]$ depending on mass ratio q.

	Modes	Full WF
Number of WFs	4000	10000
Number of grid points D	2000	2000
Distortion parameter α	0.5	0.5
Starting frequency	15 Hz	15 Hz
Sampled parameters	q, s_1, s_2	$q, s_1, s_2, \iota, \varphi, \theta_s, \varphi_s, \psi$
Total mass M	variable	variable

i.e., $S_n(f) = 1$ for all frequencies, so that every frequency is weighted equally for determining accuracy.

For **full WFs** the accuracy is measured a bit differently. The point is that gravitational wave signals can enter the detector from arbitrary sky-locations, parameterized by the angles $(\theta_s, \varphi_s) \in [0, \pi] \times [0, 2\pi]$ and with different polarization, parameterized by $\psi \in [0, 2\pi]$. We take this into account by uniformly sampling these angles from their domains and then model the true WF as the signal it produces in a GW detector. The mismatch in this case is calculated differently, see IV.A of [5] where the corresponding match is defined in equation (13). Since it will be clear from the context whether we talk about modes or full WFs, we also denote the mismatch for full WFs with the symbol $\mathcal{F}$.

Ultimately, the accuracy of the models is evaluated with the mismatch metric. During training, however, we use the MSE. Not surprisingly, the mismatch and MSE seem to be heavily positively correlated, in the sense that a low MSE yields a low mismatch.

In Fig. 3, we show the normalized histograms of the logarithm of the mismatch between individual modes generated by `IMRPhenomTPHM` and `mlgw-NN`. The median mismatch of the dominant $(2, 2)$ mode $\mathcal{F}_{\text{med}} \approx 3.64 \times 10^{-5}$ shows that the `mlgw-NN` is able to predict the `IMRPhenomTPHM` accurately. Mismatches of the subdominant modes are at least an order of magnitude higher than the $(2, 2)$ mode. This shows that the behavior of the $(2, 2)$ mode is relatively "regular" and that the higher order modes are more "noisy".

In Fig. 4, we exhibit the normalized histogram of the logarithm of the mismatch between full `IMRPhenomTPHM` test WFs and the predicted `mlgw-NN` WFs. The median value $\mathcal{F}_{\text{med}} \approx 7.00 \times 10^{-5}$ shows about an order of magnitude improvement compared to

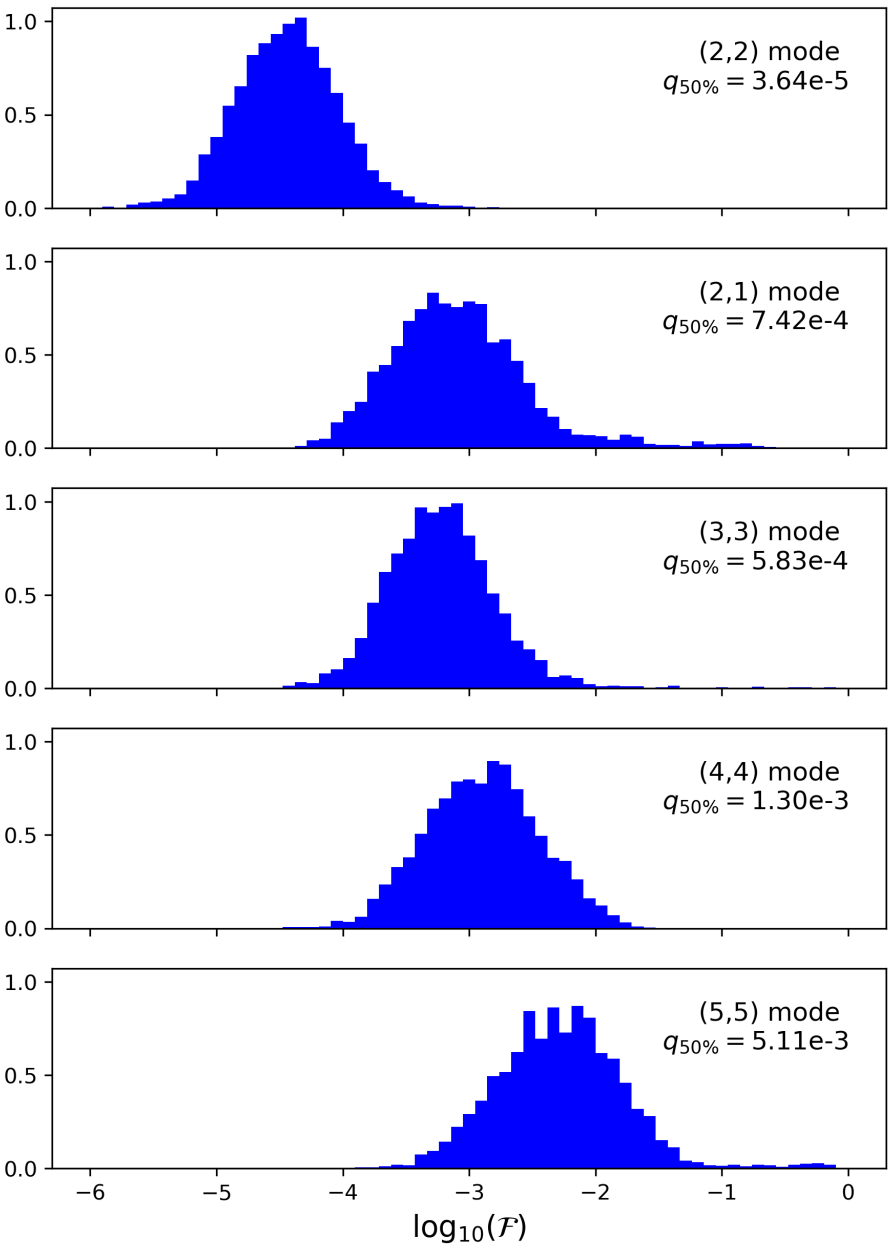


Figure 3: Individual mode comparison between test WFs H_{lm} generated by the `IMRPhenomTPHM` method and by our `mlgw-NN` approach. For each mode, the comparison consists of 4000 WFs with a minimal frequency of 15 Hz. The individual modes are generated with uniformly sampled mass ratio q and spins s_1 and s_2 in the range of $\mathcal{P}$, see also Table 2. In the upper right corners of the panels, the mode and corresponding median $q_{50\%}$ are reported.

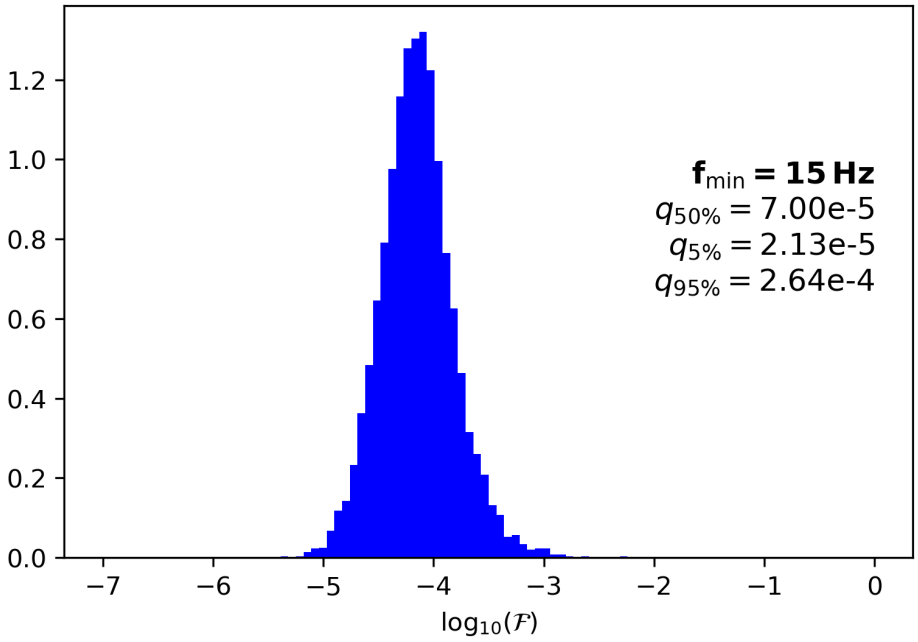


Figure 4: Normalized histogram of the logarithm of the mismatch between 10000 full test WFs generated by the `IMRPhenomTPHM`-method and by our NN approach with a minimal frequency of 15 Hz. Test WFs are generated with uniformly sampled mass ratio q and spins s_1 and s_2 in the range of $\mathcal{P}$, see also Table 2. In the upper right corner, we report the positions of the median $q_{50\%}$, the 5th and 95th percentile.

`mlgw` WFs (Fig. 5, [12])[8]. Another improvement is that the tail of the `mlgw-NN` WFs ends about an order of magnitude earlier than `mlgw` WFs. The results in Fig 4 are similar to state-of-the-art NN approaches by Khan and Green (Fig. 2, [8]).

We finish the discussion of the accuracy by examining the performance of the model as a function of the mass and spin parameters. In Fig. 5, the logarithm of the mismatch is plotted as a function of the parameters q, s_1 and s_2. The middle and right panel show a decrease in performance for high q especially when combined with large spin s_1. In fact, approximately 84% of the WFs with mismatch $\mathcal{F} > 10^{-3}$ are in the region $q > 9$, even though this region only makes up approximately 12% of the parameter space. The left and middle panel of Fig. 5, also show a significant decrease in performance for high spin s_1, the spin of the heavier black hole. Approximately 65% of the WFs with mismatch $\mathcal{F} > 10^{-3}$ are in the region $s_1 > 0.5$ region, which makes up only around 22% of the parameter space. The performance of the model seems reasonably independent from the parameter s_2.

[8]There are two remarks to be made about this comparison. The models in [12] are trained on `TEOBResumS` WFs instead of `IMRPhenomTPHM` WFs. Secondly, the WFs in [12] have a starting frequency of 10 Hz, so they are a bit longer.

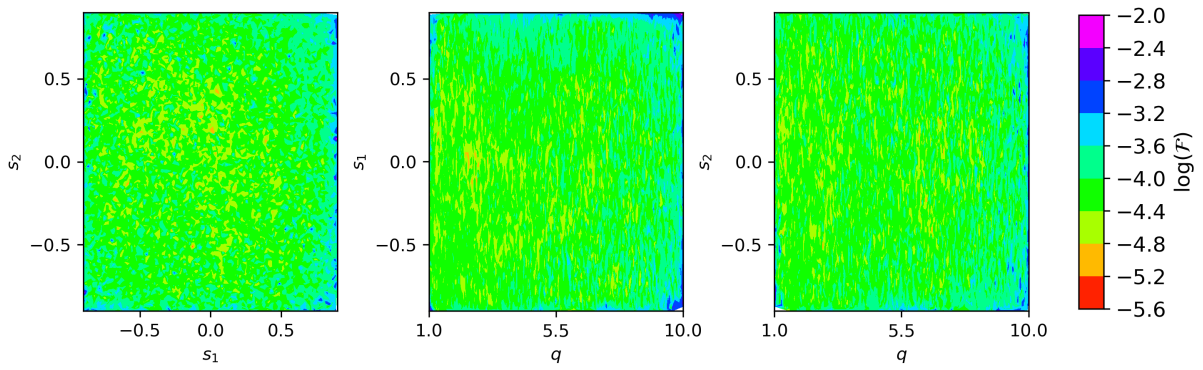


Figure 5: We show the logarithm of the mismatch as a function of the parameters s_1, s_2 and q. On the left panel, we plot the dependence on s_1 and s_2, on the middle panel between q and s_1 and on the right panel between q and s_2. From the panels, we observe that the model performance decreases for large q and for high values of s_1. The model performance does not seem to be correlated with the parameter s_2.

6.2 Speed-up

We then compare the speed with which `mlgw-NN` generates full WFs compared to the speed of the `IMRPhenomTPHM` method. We evaluate the speed-up in two ways:

1. **One-by-one** comparison. First we generate a test WF using `IMRPhenomTPHM` for uniformly sampled parameters. Immediately after this, we generate the `mlgw-NN` WF corresponding to these parameters. These two processes are timed using the `time` library of `Python`, and the process is repeated for, in our case, 10000 test WFs.

2. **Batch** comparison. Unlike the previous approach, we first generate all of the test WFs using `IMRPhenomTPHM` and store the parameters. Then, we set a *batch size* B. The idea behind batches is that we can store large matrices in memory. The execution of the necessary matrix multiplications can be optimized, which makes this approach faster. An important side remark, is that the batch-approach is not applicable to the `IMRPhenomTPHM` method. We then generate the 10000 `mlgw-NN` WFs in $10000/B$ batches of size B. Essentially, the one-by-one comparison has batch size $B = 1$.

   A slight downside to this approach, is that we cannot specify a separate time-grid for the different WFs in one batch. In order to obtain a useful comparison, we will store the shortest and longest time-grids of the `IMRPhenomTPHM` WFs. Then we generate the `mlgw-NN` WFs on both time-grids and time the two approaches separately[9].

[9]Due to different time-grids, the mismatch will vary slightly compared to the one-by-one approach. This is not a concern, because differently sized time-grids only mildly affects mismatch results.

In Fig. 6, the histograms for the generation time of single WFs is shown for the `IMRPhenomTPHM` method (red) and the `mlgw-NN` method (blue). This data shows that even without increasing the batch size, the median speed-up of the `mlgw-NN` method is already a factor of 1.66 compared to `IMRPhenomTPHM`. It should be noted that `IMRPhenomTPHM` is itself already quite a fast WF generation method.

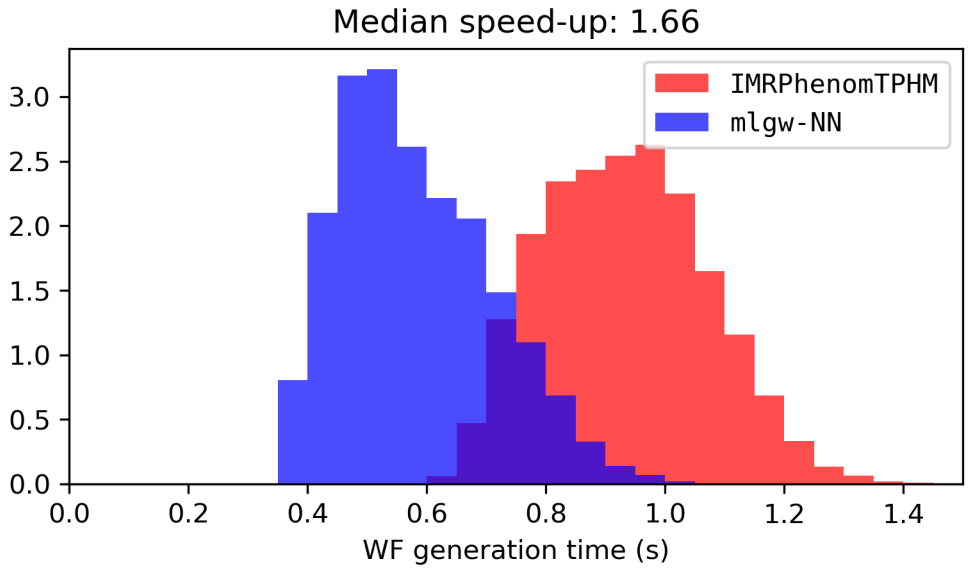


Figure 6: Normalized histograms for generation times of 10000 full test WFs with a starting frequency of 15 Hz. The comparison, using the one-by-one approach, is between the `IMRPhenomTPHM`-method (red) and the `mlgw-NN`-method (blue). The median speed-up with one-by-one generation is approximately a factor of 66%.

To obtain a further increase in speed-up, we use a batch size of $B = 200$. In table 3, we summarize the results. Notably, on the long time-grid the larger batch size of 200 only slightly beats the one-by-one method. This is due to an on average longer time-grid on which the $B = 200$ WFs are generated. By taking the average of the total generation times for the long and short time-grids, we obtain an average speed-up of 2.56 compared to `IMRPhenomTPHM`. Ideally, an even larger batch-size is used to increase the speed-up even more. The problem is that the WFs take up a lot of memory space, causing the computer to run into memory problems if B is too big. Moreover, table 3 shows that the model predictions might not be the bottleneck in speed, but that instead the reconstruction process which involves interpolating the predicted vectors is most time-consuming.

Table 3: Total generation times for 10000 full WFs using different methods. As a benchmark, we state the results for the `IMRPhenomTPHM` method, which can only generate one WF at a time. For the `mlgw-NN` approach, three methods are reported. Firstly, we report the time taken by the one-by-one method where the time-grids vary for every WF. Then, WFs are generated in batches of 200 on both the shortest time-grid and the longest time-grid from the first approach.

Method	Time-grid	Batch size B	Total generation time (s)	Average speed-up
`IMRPhenomTPHM`	variable	1	9335	-
`mlgw-NN`	variable	1	5834	1.60
	long	200	4525	2.06
	short	200	2772	3.37

7 Conclusion

In this work, we have provided `mlgw-NN`, which is a successful method for constructing a system of feedforward neural networks that models the entire process of a non-precessing binary black hole inspiral, merger and ringdown.

A model produced by the `mlgw-NN` method was trained on `IMRPhenomTPHM` WFs with parameters in the domain $(q, s_1, s_2) \in [1, 10] \times [-0.9, 0.9] \times [-0.9, 0.9]$. The model was shown to be highly accurate with a median mismatch $\mathcal{F}_{\text{med}} = 7.00 \times 10^{-5}$ for 10000 test WFs generated with a minimum frequency of $f_{\text{min}} = 15$ Hz by `IMRPhenomTPHM`. A parameter analysis showed that for large mass ratios q and high positive spin s_1 the model performance slightly decreased. To improve results, more training samples in this region could be added. There are a lot of other ways to improve the performance of this NN approach. For instance, a more rigorous assessment of NN hyperparameters could lead to more accurate and efficient NNs.

The speed-up of the `mlgw-NN` approach compared to `IMRPhenomTPHM` is approximately a factor of 1.66 for one-by-one WF generation and 2.56 for a batch sizes $B = 200$. It should be noted that there exist different WF generation methods, for instance the effective one body (EOB) models, that are slower but more faithful to numerical relativity. Numerical relativity is the slowest, but most reliable, method for WF generation which is based on solving the full system of partial differential equations of the system. When applying the `mlgw-NN` method to EOB models, we can expect a much greater speed-up. The accuracy should not decrease, however, it would take a lot longer to obtain a training set of WFs. The strength of `mlgw-NN` resides in its ability to approximate any method of WF generation.

The accuracy of `mlgw-NN` proves to be already at state-of-the-art level, but the speed-up is somewhat lacking. Future ventures of this project should therefore focus mainly on improving the generation time. Other than testing against slower generation methods, some other straightforward speed-up techniques exist. The current implementation of `mlgw-NN` is on a CPU. However, generation speeds seem to increase substantially when employing a GPU instead.

Lastly, future endeavors should work on implementing precession. The plane in which the inspiral of the two black holes happens could itself be changing its orientation with respect to the detectors on earth. This "wobbling" of the plane is called precession. It should be noted that the predicted modes H_{lm} remain useful. [4]

A Detailed neural network structure

We now provide the details of the models used for generating the results of chapter 6. For every of the five modes we used four NNs to predict the PCs of the reduced vectors $\tilde{\phi}_{\bar{\boldsymbol{\theta}}}(\tau)$ and $\tilde{A}_{\bar{\boldsymbol{\theta}}}(\tau)$. This means that in total we use 20 NNs to predict one full WF.

Every model has the sigmoid activation function between input layer and hidden layers and between all hidden layers. For the output layer the linear activation function was used to enable the NN to return negative numbers.

The optimizer that is used for every NN is the Nadam optimizer of `Keras`, which is the Adam optimizer with Nesterov momentum. This optimizer was chosen as it consistently increased the performance of the NNs compared to the more standard Adam optimizer.

The learning rate is constant for the first 500 epochs, and then decays exponentially according to $lr(E) = lr_{in}\, e^{-\gamma(E-500)}$ where lr_{in} is the initial learning rate, γ is the decay factor and E is the number of epochs.

For amplitude models, we use as inputs of the model

$$q,\ s_1,\ s_2,\ q^2,\ s_1^2,\ s_2^2,\ qs_1,\ qs_2,\ s_1s_2,\ \log(q),\ \mathcal{M}_c,\ \chi,\ \chi^2,\ \chi^3,\ \nu,\ \nu^2,\ \nu^3,\ \nu^4,$$

and for phase models we use

$$\begin{gathered} q,\ s_1,\ s_2,\ q^2,\ s_1^2,\ s_2^2, qs_1,\ qs_2,\ s_1s_2,\ \log(q),\ \mathcal{M}_c,\ \chi,\ \chi^2,\ \chi^3,\ \nu,\ \nu^2,\ \nu^3,\ \nu^4, \\ \chi\nu,\ \chi^2\nu,\ \chi\nu^2,\ \chi^3\nu,\ \chi^2\nu^2,\ \chi\nu^3, \end{gathered}$$

where we defined

$$\mathcal{M}_c = \left(\frac{q}{(1+q)^2}\right)^{3/5}, \quad \nu = \frac{1}{(1+q)^2}, \quad \chi = \frac{s_1 + qs_2}{1+q},$$

as the chirp mass, symmetric mass ratio and effective spin respectively.

For the other hyperparameters see Table 4 below.

Table 4: Detailed overview of the hyperparameters of the NNs. In the second column "Amp" refers to the model predicting the 4 PC of amplitude, "Ph1" refers to the model predicting the first PC1 and PC2 of the phase, "Ph2" for PC3, PC4, PC5 and PC6, and "Ph3" refers to the model predicting the residual of the "Ph1" model.

Mode	NN type	Hidden layers	MSE weights	Learning rate, decay (γ)	Batch size
$(2,2)$	Amp	30,20,20,10,10	200,40,3,3	10^{-3}, 5×10^{-4}	64
$(2,2)$	Ph1	20,20,15,10	5,1	10^{-4}, 3×10^{-4}	128
$(2,2)$	Ph2	15,15,15,15,15	7,3,1,1	10^{-3}, 10^{-3}	64
$(2,2)$	Ph3	20,20,15,10	8,1	4×10^{-3}, 7×10^{-4}	128
$(2,1)$	Amp	30,20,20,10,10	500,50,3,3	10^{-3},5×10^{-4}	64
$(2,1)$	Ph1	20,10,5	10,1	10^{-4}, 5×10^{-4}	128
$(2,1)$	Ph2	20,20,15,15	12,3,1,1	5×10^{-3}, 5×10^{-4}	64
$(2,1)$	Ph3	20,15,15,10	12,1	4×10^{-3}, 7×10^{-4}	128
$(3,3)$	Amp	30,20,20,10,10	200,40,3,3	10^{-3}, 5×10^{-4}	64
$(3,3)$	Ph1	30,30,20,18,10	5,1	10^{-4}, 3×10^{-4}	128
$(3,3)$	Ph2	20,20,15,15,15	10,3,1,1	4×10^{-4}, 3×10^{-4}	64
$(3,3)$	Ph3	20,20,15,12	12,1	4×10^{-3}, 2×10^{-4}	128
$(4,4)$	Amp	30,20,20,10,10	200,40,3,3	10^{-3}, 5×10^{-4}	64
$(4,4)$	Ph1	30,30,20,18,10	5,1	10^{-4}, 3×10^{-4}	128
$(4,4)$	Ph2	20,20,15,15,15	10,3,1,1	4×10^{-4}, 3×10^{-4}	64
$(4,4)$	Ph3	20,15,15,10	12,1	4×10^{-4}, 5×10^{-4}	128
$(5,5)$	Amp	30,20,20,10,10	200,40,3,3	10^{-3}, 5×10^{-4}	64
$(5,5)$	Ph1	30,30,20,18,10	5,1	10^{-4}, 3×10^{-4}	128
$(5,5)$	Ph2	20,20,15,15,15	10,3,3,1	4×10^{-4}, 3×10^{-4}	64
$(5,5)$	Ph3	20,15,15,10	12,1	4×10^{-3}, 5×10^{-4}	128

References

[1] S. Caroll. *An introduction to General Relativity Spacetime and Geometry*. 1st. New York: Harmony Books, Crown Publishers, 2019.

[2] W.D. Curtis and F.R. Miller. *Differential Manifolds and Theoretical Physics*. Pure and applied mathematics. Academic Press, 1985. ISBN: 9780122002311. URL: https://books.google.nl/books?id=euzNPgAACAAJ.

[3] Héctor Estellés et al. "New twists in compact binary waveform modeling: A fast time-domain model for precession". In: *Physical Review D* 105.8 (Apr. 2022). DOI: 10.1103/physrevd.105.084040.

[4] Héctor Estellés et al. "Time-domain phenomenological model of gravitational-wave subdominant harmonics for quasicircular nonprecessing binary black hole coalescences". In: *Phys. Rev. D* 105 (8 Apr. 2022), p. 084039. DOI: 10.1103/PhysRevD.105.084039. URL: https://link.aps.org/doi/10.1103/PhysRevD.105.084039.

[5] I. Harry, J.ó n Bustillo Calder, and Alex Nitz. "Searching for the full symphony of black hole binary mergers". In: *Physical Review D* 97.2 (Jan. 2018). DOI: 10.1103/physrevd.97.023004. URL: https://doi.org/10.1103%5C%2Fphysrevd.97.023004.

[6] Kurt Hornik. "Approximation capabilities of multilayer feedforward networks". In: *Neural Networks* 4.2 (1991), pp. 251–257. DOI: https://doi.org/10.1016/0893-6080(91)90009-T.

[7] *Keras: The high-level API for TensorFlow*. URL: https://www.tensorflow.org/guide/keras.

[8] Sebastian Khan and Rhys Green. "Gravitational-wave surrogate models powered by artificial neural networks". In: *Physical Review D* 103.6 (Mar. 2021). DOI: 10.1103/physrevd.103.064015. URL: https://doi.org/10.1103%2Fphysrevd.103.064015.

[9] Michele Maggiore. *Gravitational Waves: Volume 1: Theory and Experiments*. Oxford University Press, Oct. 2007. ISBN: 9780198570745. DOI: 10.1093/acprof:oso/9780198570745.001.0001. URL: https://doi.org/10.1093/acprof:oso/9780198570745.001.0001.

[10] K. Murphy. *Machine Learning: A Probabilistic Perspective*. MIT Press, 2012. URL: https://books.google.nl/books/about/Machine_Learning.html?id=NZP6AQAAQBAJ&redir_esc=y.

[11] E. Poisson and C.M. Will. *Gravity: Newtonian, Post-Newtonian, Relativistic*. Cambridge University Press, 2014. ISBN: 9781107032866. URL: https://books.google.nl/books?id=PZ5cAwAAQBAJ.

[12] Stefano Schmidt et al. “Machine learning gravitational waves from binary black hole mergers”. In: *Physical Review D* 103.4 (Feb. 2021). DOI: 10.1103/physrevd.103.043020.

[13] S. Sternberg. *Semi-Riemann Geometry and General Relativity*. 1st. New York: Harmony Books, Crown Publishers, 2019. URL: https://people.math.harvard.edu/~shlomo/docs/semi_riemannian_geometry.pdf.

[14] L. Tu. *Differential Geometry Connections, Curvature, and Characteristic Classes*. 1st. Medford: Springer, 2017. URL: https://link.springer.com/book/10.1007/978-3-319-55084-8.