

Dissertation  
submitted to the  
Combined Faculty of Mathematics, Engineering and Natural Sciences  
of the Ruprecht Karl University of Heidelberg, Germany  
for the degree of  
Doctor of Natural Sciences

Presented by  
Shayan Hemmatyar

born in: Tehran

Oral examination: 25-06-2025

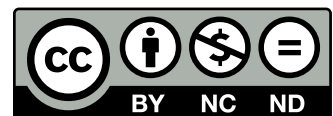


Theoretical and Machine Learning Approaches to Beyond General  
Relativity: Stability of Generalized Proca Theories and Multi-Method  
Classification of Gravitational Wave Observables

Referees: Prof. Dr. Lavinia Heisenberg  
Prof. Dr. Luca Amendola



This work is licensed under a [Creative Commons](#)  
“Attribution-NonCommercial-NoDerivs 3.0 Unported”  
license.





# Abstract

This thesis explores two complementary approaches to testing and understanding gravity beyond General Relativity (GR). The first part focuses on *Generalized Proca theories*—vector-tensor models that extend the Proca action through derivative self-interactions and non-minimal couplings, while maintaining second-order equations of motion and avoiding ghost instabilities. We analyze the quantum consistency of these theories in both flat Minkowski spacetime and weakly curved backgrounds. In flat space, we compute one-loop corrections and observe the emergence of gauge-invariant structures, suggesting a form of radiative stability. In curved spacetime, we develop a scalar-vector-tensor (SVT) decomposition to isolate physical modes and consistently integrate out non-dynamical fields. Our results show that the theories remain well-behaved under quantum corrections, supporting their viability as effective field theories.

The second part leverages gravitational wave (GW) observations as precision probes of strong-field gravity. Using convolutional neural networks (CNNs), we construct a machine learning framework to classify GW signals as either consistent with GR or exhibiting beyond-GR (BGR) deviations. The dataset includes both artificial phase deformations and physically motivated waveforms derived using the parameterized post-Einsteinian (ppE) formalism. A key tool is the *response function*, which captures the sensitivity of the waveform to small deformations. We show that training neural networks on response functions significantly improves classification accuracy and lowers detection thresholds. Applied to massive graviton models, this approach allows us to estimate the smallest graviton mass distinguishable from GR predictions.

Together, these investigations form a coherent program to study modified gravity from both theoretical and observational perspectives, contributing to the broader effort of developing consistent and testable alternatives to Einstein’s theory.





# Zusammenfassung

Diese Dissertation untersucht zwei komplementäre Ansätze zur Überprüfung und zum besseren Verständnis der Gravitation jenseits der Allgemeinen Relativitätstheorie (ART). Der erste Teil konzentriert sich auf *Generalized Proca*-Theorien – Vektor-Tensor-Modelle, die die klassische Proca-Wirkung durch derivative Selbstwechselwirkungen und nicht-minimale Kopplungen erweitern, wobei die Feldgleichungen zweiter Ordnung erhalten bleiben und Geisterfreiheiten vermieden werden. Wir analysieren die quantenkonsistente Struktur dieser Theorien sowohl im flachen Minkowski-Raumzeit-Hintergrund als auch in schwach gekrümmten Raumzeiten. In flachem Raum berechnen wir Ein-Schleifen-Korrekturen und beobachten das Auftreten gauginvarianter Strukturen, was auf eine Form quantenmechanischer Stabilität hindeutet. Für gekrümmte Hintergründe entwickeln wir eine Skalar-Vektor-Tensor (SVT)-Zerlegung, um physikalische Freiheitsgrade zu isolieren und nicht-dynamische Felder systematisch zu eliminieren. Unsere Ergebnisse zeigen, dass diese Theorien unter quantenmechanischen Korrekturen stabil bleiben und als effektive Feldtheorien konsistent sind.

Der zweite Teil nutzt Gravitationswellen (GW) als präzise Testinstrumente für starke Gravitationsfelder. Mithilfe von Convolutional Neural Networks (CNNs) entwickeln wir ein Machine-Learning-Framework, das Gravitationswellen-Signale klassifiziert – entweder als konsistent mit ART oder als Anzeichen für Abweichungen jenseits der ART (BGR). Der Datensatz umfasst sowohl künstlich veränderte Phasen als auch physikalisch motivierte Signale, die mithilfe des parameterisierten post-Einsteinischen (ppE)-Formalismus konstruiert wurden. Ein zentrales Werkzeug ist dabei die *Response-Funktion*, die beschreibt, wie empfindlich beobachtbare Größen auf kleine Phasenverformungen reagieren. Wir zeigen, dass das Training von neuronalen Netzwerken auf Basis der Response-Funktion die Klassifizierungsgenauigkeit verbessert und die Nachweisgrenzen für BGR-Signale deutlich senkt. Am Beispiel massiver Graviton-Theorien schätzen wir die kleinste vom Netzwerk detektierbare gravitative Massenschwelle ab.

Zusammen bilden diese Untersuchungen ein kohärentes Programm zur theoretischen und beobachtungsbasierten Analyse modifizierter Gravitation und leisten einen Beitrag zur Entwicklung konsistenter und überprüfbarer Alternativen zur Einstein'schen Theorie.

## Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisor, Prof. Dr. Lavinia Heisenberg, for providing me with the invaluable opportunity to undertake my PhD under her supervision. Her constant guidance, insightful suggestions, and unwavering support throughout my studies significantly shaped the development of my research and greatly enhanced my academic growth.

I am profoundly grateful to my colleague and friend, Nadine Nussbaumer, whose companionship has been essential throughout this journey. Our numerous engaging and productive discussions, particularly during our collaboration on the quantum stability project, led to crucial insights and meaningful breakthroughs.

Special thanks go to Dr. Héctor Villarrubia-Rojó for his invaluable insights and expertise, particularly regarding gravitational waves and response function analysis. His thoughtful advice greatly enriched my understanding and significantly contributed to my work.

I would also like to extend my appreciation to Dr. Stefan Zentarra, whom I had the pleasure of working with during my Master's studies and at the outset of my PhD. His mentorship during those formative stages provided a solid foundation and motivated me to pursue my research further.

A special acknowledgment goes to my wife, Elnaz, whose unwavering love, patience, and encouragement have been the pillars of strength throughout my academic journey. Her support has been indispensable, helping me overcome challenges and persist through difficult times.

Finally, my heartfelt appreciation goes to my family, whose constant belief in my capabilities and unconditional support have been my greatest motivation. Their sacrifices, encouragement, and love have made this accomplishment possible, and I am forever grateful for their presence in my life.



# Contents

<b>Abstract</b>	<b>vii</b>
<b>Zusammenfassung</b>	<b>ix</b>
<b>Acknowledgements</b>	<b>xi</b>
<b>Contents</b>	<b>xiii</b>
<b>List of Figures</b>	<b>xix</b>
<b>List of Tables</b>	<b>xxiii</b>
<b>1 Introduction and Motivations</b>	<b>1</b>
1.1 Probing Modified Gravity: A Unified Perspective . . . . .	1
1.2 Generalized Proca Theories and Quantum Stability . . . . .	1
1.3 Gravitational Waves as a Probe of New Physics . . . . .	2
1.4 Connecting the Threads . . . . .	3
1.5 Thesis Outline . . . . .	4
<b>I Quantum Stability of Generalized Proca Theories</b>	<b>7</b>
<b>2 Spin-One Fields</b>	<b>9</b>
2.1 Irreducible Representations of the Poincaré Group . . . . .	9
2.2 Tension Between Unitarity and Lorentz Invariance . . . . .	12
2.3 From Abstract States to Fields: Realizing Spin in Spacetime . . . . .	14
2.3.1 Spin-0 Field . . . . .	15
2.4 Local Symmetries and the Origin of Conserved Currents . . . . .	15
2.5 Quantum Operators and the Structure of Vector Propagators . . . . .	17

2.5.1	Massive Spin-1 Field . . . . .	18
2.5.2	Massless Spin-1 Field . . . . .	21
2.6	Identifying Ghosts: Stability Constraints in Vector Theories . . . . .	24
<b>3</b>	<b>Construction of Self-Derivative Interactions</b>	<b>27</b>
3.1	Ostrogradsky Instability . . . . .	28
3.2	Galileons . . . . .	29
3.2.1	Scalar Galileons . . . . .	30
3.2.2	Generalized Proca Theories . . . . .	32
3.3	Generalized Proca Theories in Curved Spacetime . . . . .	38
<b>4</b>	<b>Quantum Stability in Flat Spacetime</b>	<b>41</b>
4.1	Definition of Stability . . . . .	41
4.2	Power Counting . . . . .	45
4.2.1	Diagrammatic Power Counting . . . . .	46
4.2.2	Dimensional Analysis and Rescaling . . . . .	48
4.2.3	Lorentz Invariance . . . . .	49
4.3	Quantum Stability of Generalized Proca Theories in Flat Spacetime . . . . .	55
4.4	Stückelberg Expansion and the Decoupling Limit . . . . .	56
4.4.1	Two-Point Functions . . . . .	58
4.4.2	Three-Point Functions . . . . .	63
4.5	Quantum Stability from the Decoupling Limit . . . . .	65
<b>5</b>	<b>Extending Stability Analysis to Curved Spacetime</b>	<b>69</b>
5.1	Framework and Background Setup . . . . .	69
5.1.1	Quadratic Order . . . . .	70
5.1.2	Cubic interactions . . . . .	73
5.1.3	Quartic interactions . . . . .	73
5.2	Sub-Models and Explicit Lagrangians . . . . .	74
5.2.1	Decoupling Limit . . . . .	77
5.3	Quantum Stability . . . . .	78
5.3.1	Two-point Functions . . . . .	79
5.3.2	Three-point Functions . . . . .	81
5.3.3	Four-point Functions . . . . .	81
5.4	Cutoff Estimation . . . . .	82

## **6 Propagator Structure and SVT Decomposition in the Presence of Background**

<b>Fields</b>	<b>87</b>
6.1 Background Equations of Motion . . . . .	87
6.2 Mixing Terms from a Non-Vanishing Background . . . . .	88
6.3 Scalar-Vector-Tensor Decomposition . . . . .	90
6.4 Gauge Choice: The Uniform Vector Gauge . . . . .	93
6.4.1 Perturbative Elimination of Non-Dynamical Fields . . . . .	95

## **II Multi-Method Classification of Gravitational Wave Observables**

**99**

### **7 Supervised Learning and Neural Networks in Physics** **101**

7.1 Mathematical Foundations of Supervised Learning . . . . .	102
7.2 Optimization and Training . . . . .	104
7.3 Building Neural Networks . . . . .	109
7.4 Convolutional Neural Networks (CNNs) . . . . .	112
7.5 Softmax Activation and Multi-Class Classification . . . . .	114
7.6 Comparison with Other Supervised Learning Methods . . . . .	115

### **8 Problem Setup and Neural Network Architecture** **117**

8.1 Frequency-Domain Gravitational Waveform . . . . .	117
8.2 Waveform Extensions . . . . .	118
8.3 Extending Binary Black Hole Waveforms Beyond GR . . . . .	120
8.3.1 Extracting Binary Black Hole Event Data . . . . .	120
8.3.2 Data Augmentation for Neural Network Training . . . . .	121
8.4 Constructing Beyond-GR Waveforms from GR . . . . .	123
8.4.1 Frequency-Domain Representation and Phase Modification . . . . .	123
8.5 Noise Injection . . . . .	125
8.5.1 LIGO Power Spectral Density (PSD) . . . . .	125
8.5.2 Statistical Properties of the Noise . . . . .	126
8.5.3 Generating and Adding Realistic Detector Noise . . . . .	127
8.6 Neural Network Implementation . . . . .	128
8.6.1 Dataset Splitting and Its Purpose . . . . .	128
8.6.2 Neural Network Architecture . . . . .	129

<b>9</b>	<b>Training and Evaluation of the Classification Neural Network</b>	<b>133</b>
9.1	Toy Model: Gaussian Phase Shift . . . . .	133
9.2	Mismatch as a Preliminary Metric . . . . .	134
9.3	Classification Performance at Strong Deformation ( $\beta = 1$ ) . . . . .	136
9.3.1	Evaluation Setup . . . . .	136
9.3.2	Results for $\beta = 1$ . . . . .	136
9.3.3	Evaluation Metrics and Threshold Criterion . . . . .	137
9.4	Sensitivity to Smaller Phase Shifts . . . . .	140
9.5	Deviation in the Post-Newtonian Phase Coefficients . . . . .	145
<b>10</b>	<b>Response Function Formalism</b>	<b>151</b>
10.1	Theoretical Framework and Response Function Formalism . . . . .	151
10.1.1	Boundary Conditions . . . . .	153
10.2	Mismatch and its Response Function . . . . .	154
10.3	Gaussian Toy Model Revisited: Response Function Performance . . . . .	156
10.4	Limits of Classification: Accuracy, Bayes Optimal Error, and Interpretability	161
<b>11</b>	<b>Analysis of Modified Theories of Gravity</b>	<b>169</b>
11.1	The Parameterized Post-Einsteinian (ppE) Framework . . . . .	170
11.2	Modified Theories of Gravity and Their ppE Signatures . . . . .	172
11.2.1	Brans–Dicke Theory (Scalar–Tensor Gravity) . . . . .	172
11.2.2	Massive Gravity Theories . . . . .	173
11.2.3	Dynamical Chern–Simons Gravity (Parity-Violating Gravity) . . . . .	174
11.3	Neural Network Classification of Massive Gravity Waveforms . . . . .	175
11.3.1	Detectability Threshold for the Graviton Mass . . . . .	176
11.3.2	Comparison with Existing Bounds . . . . .	176
<b>12</b>	<b>Conclusion</b>	<b>181</b>
	<b>Appendix</b>	<b>185</b>
<b>A</b>	<b>Lagrangians</b>	<b>187</b>
A.1	Cubic Interaction Terms . . . . .	187
A.2	Feynman Rules for Flat-Space Generalized Proca Theory . . . . .	199
A.3	Feynman Rules from Curved-Space Lagrangian (Expanded on Flat Back-ground) . . . . .	201



Contents	xvii
<b>B 3.5PN Phasing Coefficients in the Stationary Phase Approximation</b>	<b>205</b>
<b>Bibliography</b>	<b>207</b>



# List of Figures

4.1	One-loop, two-point diagrams. The diagram on the left represents a standard bubble diagram with two internal propagators. The diagram on the right corresponds to a tadpole diagram, where a single internal propagator forms a loop attached to an external leg. . . . .	59
4.2	One-loop, three-point diagrams contributing to the amplitude. The diagram on the left represents the triangle topology, where all three external legs are attached to different vertices of the loop. The diagram on the right corresponds to a bubble-type diagram, where two external legs couple to the same vertex. This second diagram appears in three different channel permutations, depending on which pair of external legs share a vertex. . .	64
7.1	Comparison of Linear Regression and Logistic Regression. The left panel illustrates linear regression, where a continuous line is fitted to the data points, allowing predictions over a continuous range. In contrast, the right panel depicts logistic regression, where the sigmoid function maps input values to probabilities between 0 and 1. This enables binary classification by establishing a decision boundary at a probability threshold (typically 0.5), distinguishing between two classes. . . . .	103
7.2	Visualization of gradient descent on $f(x, y) = x^2 + y^2$ , showing the iterative path toward the minimum. . . . .	105
7.3	Comparison of gradient descent trajectories for the function $f(x) = x^2$ with different learning rates. The left plot demonstrates slow convergence with a small learning rate, while the right plot shows overshooting and divergence with a large learning rate. . . . .	106

7.4	Comparison of polynomial fits (degree 9) via ridge regression for different values of the regularization parameter $\lambda$ . The left panel ( $\lambda = 0$ ) overfits the noisy training data, the center panel ( $\lambda = 0.001$ ) achieves a good balance by closely following the true function and data, while the right panel ( $\lambda = 100$ ) underfits due to excessive regularization. . . . .	108
7.5	Schematic representations of three common activation functions in machine learning: sigmoid, tanh, and ReLU. The ReLU function introduces nonlinearity by outputting zero for negative input values (i.e., “turning off” the neuron), which is crucial for enabling networks to learn complex patterns. 110	
7.6	A Neural Network with 3 Layers: two hidden and one output Layer. [55] .	111
7.7	A 2D convolutional network Filtering. Based on the dimension of the filter (kernel), and the padding, the dimension of the output changes [55]. . . .	113
8.1	BBH events taken from LIGO’s database. . . . .	121
8.2	BBH events and their augmented counterparts. Blue dots represent the original data, while red dots denote the augmented events generated via Gaussian perturbations. . . . .	123
8.3	The red dashed line represents a small shift in the waveform during the inspiral (low-frequency) regime, illustrating the effect of the applied phase modification in the BGR waveform. . . . .	125
8.4	The power spectral density (PSD) used for noise injection in our analysis, corresponding to the aLIGOZeroDetHighPower configuration [72]. This PSD represents the expected sensitivity of Advanced LIGO in its zero-detuned, high-power mode, providing a realistic noise model for simulating gravitational wave detections from binary black hole mergers. . . . .	126
9.1	Average mismatch between BGR and GR waveforms as a function of the deformation parameter $\beta$ . The shaded region shows the standard deviation across the waveform dataset. As $\beta$ increases, the deviation becomes more pronounced, leading to a higher mismatch. . . . .	135
9.2	Training and validation loss for $\beta = 1$ . The loss curves converge smoothly, showing stable and well-regularized learning. . . . .	137
9.3	Test predictions for $\beta = 1$ . The classifier cleanly separates GR (blue) and BGR (red) events. The background shading denotes the classification threshold at 0.5. . . . .	138

9.4	Confusion matrix for the test set at $\beta = 1$ . The model achieves perfect classification. . . . .	139
9.5	Test predictions for $\beta = 0.5$ . The classifier maintains perfect separation. .	141
9.6	Test predictions for $\beta = 0.45$ : miss-classification starts to kick in at this scale. . . . .	141
9.7	Test predictions for $\beta = 0.4$ . . . . .	142
9.8	Test predictions for $\beta = 0.35$ . . . . .	142
9.9	Test predictions for $\beta = 0.3$ . . . . .	143
9.10	Test predictions for $\beta = 0.28$ . This marks the boundary of detectability as defined by the 95% accuracy threshold. . . . .	143
9.11	Confusion matrix for $\beta = 0.28$ . This is the lowest deformation for which the classifier still exceeds 95% test accuracy. . . . .	144
9.12	Softmax classification results for PN-deformed waveforms. The network was trained to distinguish between GR waveforms and BGR waveforms constructed by modifying the $n = 1$ , $n = 4$ , or $n = 6$ PN coefficients individually. The y-axis corresponds to the predicted class probabilities for each sample in the test set. Colored regions and markers indicate the classifier's output. The model successfully learns to differentiate the origin of each waveform, identifying the PN order of the deviation. . . . .	148
10.1	Average mismatch between BGR and GR waveforms using response functions as input, plotted against the deformation parameter $\beta$ . The shaded region indicates the standard deviation across the dataset. Compared to waveform-based classification, response functions provide a considerably lower detection threshold, allowing deviations at significantly smaller $\beta$ values to be accurately identified. . . . .	157
10.2	Classification comparison at $\beta = 0.28$ . The top plot shows waveform-based classification, which fails to reliably distinguish between GR and BGR waveforms. The bottom plot demonstrates perfect classification performance using response functions, highlighting the superiority of response functions in detecting subtle deviations from GR. . . . .	158
10.3	Classification comparison at $\beta = 0.0003$ . The top plot illustrates complete failure of waveform-based classification. The bottom plot shows response-function-based classification maintaining accuracy above 95%, underscoring the vastly improved sensitivity provided by response functions. . . . .	159

10.4	Confusion matrix for response-function classification at $\beta = 0.0003$ . High values along the diagonal reflect the classifier's strong capability to distinguish GR from BGR waveforms even at minimal deviations. . . . .	160
10.5	Illustrative comparison of machine learning performance with human-level performance and the Bayes optimal error. The x-axis denotes schematic training time or model improvement, while the y-axis represents classification accuracy. Human-level accuracy typically approaches the Bayes limit, while model performance initially improves rapidly and then gradually saturates. . . . .	162
10.6	Comparison of normalized response functions at $\beta = 0.03$ . The GR and BGR responses show clear differences, easily discernible by eye. At this level of deformation, classification is straightforward. . . . .	164
10.7	Comparison of normalized response functions at $\beta = 0.003$ . Though less pronounced than in the large $\beta$ case, the BGR response retains visible structure that sets it apart from GR. The classifier continues to achieve near-perfect accuracy. . . . .	165
10.8	Comparison of normalized response functions at $\beta = 0.0003$ . The GR and BGR responses exhibit highly similar noise-dominated behavior, making visual discrimination infeasible. However, the classifier still performs above the 95% accuracy threshold. . . . .	165
10.9	Normalized response functions for $\beta = 0.0003$ across multiple gravitational wave events. <b>Left column:</b> GR responses. <b>Right column:</b> Corresponding BGR responses. Their indistinguishability at first glance is a key reason neural networks outperform human inspection in this regime. . . .	167
10.10	Normalized averaged response functions at $\beta = 0.0003$ . The GR average (blue, dashed) converges to zero due to noise cancellation. The BGR average (green, solid), however, reveals a residual pattern—a coherent imprint of the small deformation. . . . .	168
11.1	Average waveform mismatch as a function of the graviton mass $m_g$ . The classifier becomes insensitive to MG-induced phase shifts for masses smaller than $\sim 2.1 \times 10^{-61}$ kg. . . . .	177
11.2	Phase shift $\delta\psi_{\text{MG}}(f)$ for various values of $m_g$ . The magnitude of the shift increases with graviton mass, affecting the waveform primarily in the detector-sensitive frequency band. . . . .	178

# List of Tables

5.1	Examples of cutoff scales $\Lambda$ and $\Lambda_3$ for various values of $p$ . . . . .	85
9.1	Classification performance metrics and average mismatch for various values of $\beta$ . Boldface indicates the threshold value where test accuracy drops to 95%. . . . .	144





# Chapter 1

## Introduction and Motivations

### 1.1 Probing Modified Gravity: A Unified Perspective

Over the past two decades, gravitational physics has undergone a remarkable transformation, driven by both theoretical insights and experimental breakthroughs. On the observational side, the direct detection of gravitational waves (GWs) by the LIGO and Virgo collaborations has opened an entirely new window into the strong-field regime of gravity. On the theoretical side, a wide array of models have emerged to address persistent tensions in cosmology and quantum gravity. Despite its enormous success, General Relativity (GR) is widely regarded as an effective theory, likely to break down at some energy scale or under extreme conditions.

This thesis investigates two complementary directions in the study of gravity beyond Einstein: (i) the quantum consistency of vector-tensor theories known as Generalized Proca models, and (ii) the detection and classification of deviations from GR in gravitational wave signals using machine learning. While these topics may seem disparate at first glance, they are united by a common goal: to understand the structure and limits of gravitational dynamics in the broader context of modified gravity theories. In what follows, we first discuss the motivation and scope of each direction individually, then highlight how they fit together into a coherent program.

### 1.2 Generalized Proca Theories and Quantum Stability

General Relativity describes gravity as a massless spin-2 field propagating on a curved spacetime, but it lacks a consistent quantum field theoretic completion. One promising

avenue for extending GR involves the introduction of massive spin-1 fields with derivative self-interactions, leading to so-called *Generalized Proca theories*. These vector-tensor models extend the standard Proca action by including higher-order derivative and non-minimal couplings, constructed such that the equations of motion remain second order—thus avoiding Ostrogradsky instabilities.

Generalized Proca theories are natural vector-field analogues of Horndeski and Galileon scalar-tensor models. Their structure ensures that only three propagating degrees of freedom remain: two transverse and one longitudinal polarization. These theories have been widely studied in cosmology, where a background vector field may spontaneously break Lorentz symmetry while maintaining spatial isotropy, potentially offering new insights into dark energy or early universe dynamics.

A key requirement for any effective field theory is quantum stability: the theory should remain consistent under loop corrections, with no dangerous operators reintroduced by renormalization. In flat Minkowski spacetime, previous studies have shown that Generalized Proca theories exhibit a surprising degree of radiative stability, akin to Galileon models. This thesis extends those results by explicitly computing one-loop divergences in simplified Proca models, including 2-point, 3-point, and 4-point functions. We find that certain gauge-invariant structures, such as  $\square\eta_{\mu\nu} - \partial_\mu\partial_\nu$ , emerge from loop corrections, hinting at an effective restoration of gauge symmetry at the quantum level.

We then turn to curved backgrounds, where a non-zero vacuum expectation value of the vector field introduces mixing between gravitational and vector perturbations. This modifies the propagator structure and complicates loop calculations. We develop tools to handle these challenges using both covariant methods and a scalar-vector-tensor (SVT) decomposition, which isolates the physical degrees of freedom and facilitates perturbative computations. Our results show how to consistently eliminate non-dynamical fields and evaluate loop integrals even in the presence of anisotropic dispersion relations.

Together, these investigations provide a comprehensive analysis of Generalized Proca theories as candidate low-energy effective descriptions of modified gravity, with careful attention paid to quantum consistency in both flat and weakly curved spacetimes.

### 1.3 Gravitational Waves as a Probe of New Physics

The first direct detection of gravitational waves in 2015 marked a turning point in experimental gravity. These signals, produced by the inspiral and merger of compact binaries,

carry detailed information about the dynamics of spacetime under extreme conditions. As such, they offer an unparalleled opportunity to test GR in regimes far beyond what was previously accessible.

While the observed waveforms are so far consistent with GR, many extensions to Einstein’s theory predict subtle modifications, particularly in the inspiral phase of binary coalescences. These deviations could arise from additional fields, modified propagation speeds, parity-violating terms, or extra polarizations. A widely used framework for capturing such effects is the *parameterized post-Einsteinian* (ppE) formalism, which encodes beyond-GR (BGR) corrections as deformations of standard waveforms.

Traditional matched filtering techniques are highly effective for detecting known signals, but they rely on explicit waveform templates. This poses a limitation when searching for small or unmodeled deviations. Recent advances in machine learning (ML), particularly convolutional neural networks (CNNs), offer a promising alternative. These data-driven models excel at pattern recognition and can identify features in noisy time-series data that may evade traditional methods.

In this thesis, we construct a machine learning pipeline to classify gravitational wave signals according to their underlying gravitational theory. We generate datasets containing both GR and BGR waveforms, including artificial phase shifts, modified post-Newtonian coefficients, and physically motivated ppE corrections. We train CNNs to distinguish between these classes and evaluate their performance under realistic noise conditions.

A key innovation in our approach is the use of the *response function*, which maps phase deformations to changes in observable overlap. This representation improves classification accuracy and interpretability, enabling the network to detect deviations smaller than those visible to the human eye. We demonstrate that the classifier approaches the Bayes optimal limit in idealized settings, and apply it to estimate the minimal graviton mass detectable via waveform distortions in massive gravity scenarios.

## 1.4 Connecting the Threads

At first glance, the quantum consistency of vector-tensor theories and the machine learning classification of gravitational waveforms may seem to belong to different domains. However, both address fundamental questions about the validity and structure of gravity beyond GR. The former investigates whether certain modifications are theoretically viable, while the latter probes whether such modifications can be observationally detected.

Together, these efforts contribute to a broader program: testing the limits of Einstein's theory through both theoretical consistency and empirical observability. On one hand, Generalized Proca theories provide well-defined, ghost-free modifications of gravity that may serve as realistic alternatives to GR. On the other hand, gravitational wave observations offer a direct means to test these modifications, especially when combined with modern machine learning techniques capable of detecting faint imprints of new physics.

This thesis thus represents a two-pronged approach to exploring modified gravity: theoretical validation through quantum field theoretic methods, and empirical classification through data-driven analysis of gravitational waveforms. Both avenues are crucial for a complete understanding of gravity in the 21st century.

## 1.5 Thesis Outline

The structure of the thesis is as follows:

- **Chapter 2** provides a review of spin-1 field theories, focusing on their classical and quantum consistency. We discuss the representation theory of the Poincaré group, the role of gauge invariance, and the embedding of spin-1 particles into field theory frameworks.
- **Chapter 3** develops the construction of self-derivative interaction theories, such as scalar Galileons and Generalized Proca theories. We discuss the symmetry principles, second-order dynamics, and effective field theory structure that underlie these models.
- **Chapter 4** analyzes the quantum stability of Generalized Proca theories in flat Minkowski spacetime. We compute the divergent parts of loop amplitudes and explore the emergence of gauge-invariant structures in the quantum corrections.
- **Chapter 5** extends the quantum analysis to weakly curved spacetimes. We derive interaction terms in the presence of background curvature and investigate whether loop corrections introduce instabilities or modify the effective action.
- **Chapter 6** investigates the propagator structure and scalar-vector-tensor decomposition in the presence of a non-zero background vector field. We study the mixing between metric and vector perturbations and construct a gauge-invariant basis for perturbative analysis.

- **Chapter 7** introduces the basics of machine learning, including definitions, training paradigms, and a discussion of neural network architectures. We describe how these tools can be applied to gravitational waveform classification.
- **Chapter 8** presents the full architecture of the machine learning framework used to classify gravitational waveforms. This includes the construction of the waveform dataset, convolutional neural network design, and preprocessing techniques.
- **Chapter 9** evaluates the classification performance using a controlled toy model. We assess the detectability threshold for Gaussian phase deformations and perform multiclass classification using post-Newtonian phase coefficient modifications.
- **Chapter 10** introduces the response function formalism, showing how this representation improves classifier sensitivity and interpretability. We demonstrate that neural networks trained on response functions outperform those trained on raw waveforms.
- **Chapter 11** applies the classification framework to waveforms derived from physically motivated beyond-GR theories. Using the parameterized post-Einsteinian (ppE) formalism, we estimate the minimal graviton mass distinguishable from GR.
- **Chapter 12** summarizes the main findings of the thesis and outlines future directions. These include applications to real detector data, more general waveform features, and further exploration of modified gravity theories using machine learning techniques.



**Part I**

**Quantum Stability of Generalized Proca  
Theories**





# Chapter 2

## Spin-One Fields

A natural starting point for exploring the internal consistency requirements of quantum field theories is an investigation of spin and its interplay with Lorentz invariance. While in non-relativistic quantum mechanics spin appears as an abstract internal degree of freedom, its true geometric significance becomes apparent only in the context of relativistic theories. Historical insights from atomic spectroscopy had already revealed that electrons possess two discrete spin states, while photons are characterized by two polarization states. Although the latter can be intuitively understood within classical electrodynamics, the notion of a single massless and structureless quantum—such as the photon—exhibiting polarization poses a deeper conceptual challenge. Unlike electrons, for which the analogy of a spinning top offers some classical intuition, photons defy such interpretation. These subtleties are resolved by recognizing the role of Lorentz symmetry and its implications for constructing consistent quantum field theories.

This chapter provides an overview of the key aspects of spin-1 field theories, focusing on how unitarity, Lorentz invariance, and gauge symmetry constrain the dynamics of massive and massless vector fields. The presentation is structured to lay the groundwork for the construction of generalized Proca theories discussed in the next chapter. For a more detailed and comprehensive treatment, one can refer to [\[1, 2, 3, 4, 5, 6, 7, 8\]](#).

### 2.1 Irreducible Representations of the Poincaré Group

Quantum field theory must be constructed in accordance with the fundamental symmetries of spacetime. In particular, it should exhibit invariance under spacetime translations and Lorentz transformations. The combined group of these transformations—the isometry

group of Minkowski spacetime—is known as the Poincaré group, denoted by  $\text{ISO}(1, 3)$ . Translational invariance asserts that physics is indifferent to the absolute position in spacetime, meaning that observables must remain unchanged under  $x^\mu \rightarrow x^\mu + a^\mu$  for any constant four-vector  $a^\mu$ . Similarly, Lorentz invariance requires that physical predictions remain invariant under changes of inertial frame.

In nature, particles are classified not only by their mass and spin, but also by other conserved quantum numbers. Under Poincaré transformations, certain quantities such as momentum and spin projection may change, while internal quantum numbers remain invariant. Accordingly, we may define a particle species as a collection of quantum states that transform into one another under the action of the Poincaré group.

Formally, a quantum state  $|\psi\rangle$  transforms under a Poincaré operation  $\mathcal{P}$  according to

$$|\psi\rangle \rightarrow \mathcal{P}|\psi\rangle. \quad (2.1)$$

A collection of states that mix among themselves under the group action forms a representation of the group. For instance, scalar fields  $\phi(x)$  define a representation of the translation group via  $\phi(x) \rightarrow \phi(x + a)$ . More generally, if  $\{|\psi_i\rangle\}$  denotes a basis for a given representation, then the transformation law takes the form

$$|\psi_i\rangle \rightarrow \mathcal{P}_{ij}|\psi_j\rangle, \quad (2.2)$$

where the summation over repeated indices is implied. If the representation cannot be decomposed into smaller invariant subsets, it is deemed irreducible.

Beyond irreducibility, physical relevance demands that representations be unitary. This requirement arises because physical observables are constructed from inner products of states—quantities such as matrix elements

$$\mathcal{M} = \langle\psi_1|\psi_2\rangle \quad (2.3)$$

must remain invariant under symmetry transformations. Applying a Poincaré transformation, we obtain

$$\mathcal{M} = \langle\psi_1|\mathcal{P}^\dagger\mathcal{P}|\psi_2\rangle, \quad (2.4)$$

which remains unchanged only if  $\mathcal{P}^\dagger\mathcal{P} = \mathbf{1}$ —the defining property of unitarity.

It is important to note that unitary representations form only a narrow subset of all representations of the Poincaré group. For example, the commonly encountered four-vector representation  $A_\mu$  is not itself unitary. Nevertheless, only unitary representations can yield

well-defined, Lorentz-invariant matrix elements, and therefore they serve as the foundational building blocks for the particle spectrum of any consistent field theory.

**Hence, one may regard particles as quantum states that transform under irreducible unitary representations of the Poincaré group.**

This viewpoint effectively serves as a precise definition of what a particle is in the language of quantum field theory.

Moreover, there exists a more stringent constraint at the level of the full interacting theory: the  $S$ -matrix, which encodes scattering amplitudes, must itself be unitary. While the unitarity of Poincaré representations pertains to the kinematics of free particles,  $S$ -matrix unitarity governs the dynamics and ensures conservation of probability in physical processes.

The decomposition into irreducible representations reflects an observational reality: different types of particles exist, each with their own distinct properties and transformation laws. When external fields such as magnetic fields or optical polarizers are applied to particles like electrons or photons, they interact differently depending on their spin and momentum. These manipulations reveal the underlying group structure that governs how states mix under transformations, with irreducible representations serving as the minimal invariant subspaces.

The classification of all unitary irreducible representations of the Poincaré group was achieved by Eugene Wigner [9]. His framework remains one of the cornerstones of relativistic quantum theory. Wigner demonstrated that these representations are labeled by the particle's mass  $m \geq 0$  and spin  $J$ , where  $J$  is a non-negative half-integer:  $J = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots$ . For  $m > 0$  and  $J > 0$ , the representation contains  $2J + 1$  distinct polarization states. For massless particles with  $J > 0$ , only two physical polarization states remain, while for  $J = 0$ , a single state exists regardless of the mass. A detailed derivation can be found in [10].

While Wigner's classification offers a systematic way to label the kinematic content of particle states, it does not in itself provide a recipe for constructing interacting quantum field theories. The next step involves embedding these representations into local field operators such as scalars  $\phi(x)$ , vectors  $V_\mu(x)$ , spinors  $\psi(x)$ , and higher-rank tensors  $T_{\mu\nu}(x)$ . This embedding enables us to write local Lagrangians and compute physical observables.

A technical challenge arises here: the number of components in such tensorial objects grows rapidly, e.g., 1, 4, 16, 64,  $\dots$  for tensors of increasing rank, whereas the number of physical degrees of freedom in a spin- $j$  representation is  $2j + 1$ . Thus, one must carefully construct the theory to isolate only the physical degrees of freedom. This difficulty leads

naturally to the introduction of additional structures—such as gauge symmetries—that serve to project out the unphysical components. We will explore these aspects in detail as we develop the theory of spin-one fields.

## 2.2 Tension Between Unitarity and Lorentz Invariance

The interplay between unitarity and Lorentz invariance presents a subtle but fundamental challenge in relativistic quantum field theory. To illustrate this issue, we begin with a familiar context from non-relativistic quantum mechanics. Consider an electron whose spin states are represented by the orthonormal basis vectors  $|\uparrow\rangle$  and  $|\downarrow\rangle$ . A general state in this Hilbert space can be expressed as a superposition:

$$|\psi\rangle = c_1|\uparrow\rangle + c_2|\downarrow\rangle. \quad (2.5)$$

The norm of such a state is given by

$$\langle\psi|\psi\rangle = |c_1|^2 + |c_2|^2 > 0, \quad (2.6)$$

and is invariant under rotations. In fact, the rotational symmetry is part of a larger  $SU(2)$  symmetry group, though that detail will not be essential here.

Now consider a different example: a set of four orthonormal states  $\{|V_0\rangle, |V_1\rangle, |V_2\rangle, |V_3\rangle\}$ , transforming as components of a four-vector under Lorentz transformations. A general state in this basis takes the form

$$|\psi\rangle = c_0|V_0\rangle + c_1|V_1\rangle + c_2|V_2\rangle + c_3|V_3\rangle, \quad (2.7)$$

with a corresponding norm

$$\langle\psi|\psi\rangle = |c_0|^2 + |c_1|^2 + |c_2|^2 + |c_3|^2. \quad (2.8)$$

This norm is positive definite, as required by the probabilistic interpretation of quantum mechanics. However, it fails to be invariant under Lorentz transformations. For example, consider the state  $|\psi\rangle = |V_0\rangle$ , which clearly has unit norm. Under a boost along the  $x^1$ -direction with rapidity  $\beta$ , the state transforms to

$$|\psi'\rangle = \cosh\beta |V_0\rangle + \sinh\beta |V_1\rangle, \quad (2.9)$$

and the new norm becomes

$$\langle\psi'|\psi'\rangle = \cosh^2\beta + \sinh^2\beta \neq 1. \quad (2.10)$$

Hence, the norm is not preserved by the Lorentz transformation, implying a breakdown of unitarity.

This discrepancy can also be understood from a matrix perspective. The Lorentz boost in this two-dimensional subspace is represented by the matrix

$$\Lambda = \begin{pmatrix} \cosh \beta & \sinh \beta \\ \sinh \beta & \cosh \beta \end{pmatrix}, \quad (2.11)$$

which is clearly non-unitary, since  $\Lambda^\dagger \neq \Lambda^{-1}$ .

One might attempt to resolve this issue by redefining the inner product to be Lorentz invariant. A natural candidate is

$$\langle \psi | \psi \rangle = |c_0|^2 - |c_1|^2 - |c_2|^2 - |c_3|^2, \quad (2.12)$$

which mirrors the Minkowski metric structure. Although this form is Lorentz invariant, it is not positive definite. This alone is not necessarily problematic—inner products in quantum theory can be complex—but it severely complicates any probabilistic interpretation.

To see this explicitly, consider again the boosted state  $|\psi'\rangle$ . It satisfies  $\langle \psi' | \psi' \rangle = 1$  under the modified norm, preserving Lorentz invariance. However, the overlap

$$|\langle V_0 | \psi' \rangle|^2 = \cosh^2 \beta \quad (2.13)$$

is strictly greater than 1 for any nonzero  $\beta$ . Thus, projection amplitudes can exceed unity, violating the probabilistic interpretation of quantum theory. The issue stems from the mixing of positive- and negative-norm components under Lorentz transformations, which destroys the bound  $0 \leq P \leq 1$  for probabilities.

In summary, we encounter a fundamental tension: on the one hand, unitarity demands a positive-definite inner product, typically preserved by unitary transformations and associated with the Euclidean metric  $\delta^{\mu\nu}$ . On the other hand, Lorentz invariance requires that the theory respect the indefinite Minkowski metric  $g^{\mu\nu} = \text{diag}(+1, -1, -1, -1)$ , which leads to non-unitary representations when naively applied to finite-dimensional Hilbert spaces.

To resolve this conflict, two key observations guide us forward. First, the squared norm of a vector in Minkowski space,  $V_\mu V^\mu = V_0^2 - V_1^2 - V_2^2 - V_3^2$ , suggests that the four-vector representation  $V_\mu$  is reducible. Specifically, it decomposes into a direct sum of a spin-0 component (a Lorentz scalar) and a spin-1 component (a three-dimensional vector under rotations). Therefore, by appropriately projecting out one of these sectors—typically the scalar—we can isolate a subspace in which the remaining states possess a consistent physical interpretation.

Second, although finite-dimensional irreducible unitary representations of the Poincaré group do not exist beyond scalars, there do exist infinite-dimensional unitary representations. The key is to replace constant basis vectors (e.g.,  $(1, 0, 0, 0)$ ,  $(0, 1, 0, 0)$ , etc.) with momentum-dependent polarization vectors  $\epsilon_\mu(p)$  that satisfy appropriate transversality and normalization conditions. These basis vectors are not fixed globally but depend on the on-shell momentum  $p^\mu$ , satisfying  $p^2 = m^2$  for a massive particle.

Thus, our strategy will be as follows: we first identify how to consistently embed the correct number of degrees of freedom for a given mass and spin into tensor fields such as  $A_\mu(x)$ . Then, we will demonstrate how the requirement of Lorentz invariance leads naturally to an infinite-dimensional structure of polarization states, forming a unitary representation of the Poincaré group consistent with both the field-theoretic and probabilistic interpretation of quantum mechanics.

## 2.3 From Abstract States to Fields: Realizing Spin in Space-time

We now turn to the question of how to realize specific spin states within a field-theoretic framework. The goal is to construct Lagrangians that describe fields corresponding to particles of definite spin. Since we begin in the classical regime, we do not yet impose unitarity — as the notion of a Hilbert space inner product is absent — but we do require that the classical energy of the system be bounded from below. The existence of negative-energy solutions would signal instabilities upon quantization, as they would allow decay processes such as the vacuum producing particle-antiparticle pairs with vanishing total momentum,  $p_1^\mu + p_2^\mu = 0$ . This decay would conserve both energy and momentum unless additional constraints prevent negative-energy solutions. A broader discussion of classical pathologies and their quantum implications appears in Section 8.7.

The classical energy density is encoded in the 00 component of the energy-momentum tensor,

$$\mathcal{E} = \mathcal{T}_{00} = \sum_n \frac{\partial \mathcal{L}}{\partial \dot{\phi}_n} \dot{\phi}_n - \mathcal{L}, \quad (2.14)$$

and the total energy of the field configuration is given by the spatial integral

$$E = \int d^3x \mathcal{E}. \quad (2.15)$$

### 2.3.1 Spin-0 Field

The realization of a spin-0 (scalar) particle is straightforward. Since such a particle possesses only one degree of freedom, it can be embedded directly into a real scalar field  $\phi(x)$ . A simple Lorentz-invariant Lagrangian for a free scalar field of mass  $m$  is

$$\mathcal{L}(x) = \frac{1}{2} \partial_\mu \phi(x) \partial^\mu \phi(x) - \frac{1}{2} m^2 \phi^2(x). \quad (2.16)$$

This Lagrangian is invariant under spacetime translations and yields the Klein-Gordon equation of motion,

$$(\square + m^2) \phi(x) = 0, \quad (2.17)$$

which admits plane wave solutions  $\phi(x) \sim e^{\pm i p \cdot x}$  with on-shell condition  $p^2 = m^2$ .

The corresponding energy density derived from this Lagrangian is

$$\mathcal{E} = \frac{1}{2} \left[ (\partial_t \phi)^2 + (\vec{\nabla} \phi)^2 + m^2 \phi^2 \right], \quad (2.18)$$

which is manifestly positive definite and bounded from below. The sign of the kinetic and mass terms has thus been chosen to ensure the stability of the classical theory.

## 2.4 Local Symmetries and the Origin of Conserved Currents

Symmetries in field theory can be broadly classified as either global or local, depending on whether their transformation parameters are constant or space-time dependent. A symmetry characterized by a constant parameter  $\alpha$  is referred to as a global symmetry. When  $\alpha$  is allowed to vary as a function of spacetime,  $\alpha(x)$ , the symmetry becomes local, or gauge. Gauge symmetries therefore represent redundancies in our description that are specified independently at each point in space-time. Importantly, any gauge symmetry automatically contains a global symmetry as a special case, and global symmetries, by Noether's theorem, correspond to conserved currents.

To illustrate this connection concretely, consider a free complex scalar field with the Lagrangian

$$\mathcal{L} = -\phi^\star \square \phi. \quad (2.19)$$

This theory is invariant under global phase rotations  $\phi \rightarrow e^{-i\alpha} \phi$ , with constant  $\alpha$ , and hence possesses an associated conserved current. However, this Lagrangian is not invariant under local phase transformations where  $\alpha = \alpha(x)$ .

Let us now examine what happens when we promote the symmetry to a local one by introducing a gauge field. This leads us to scalar quantum electrodynamics (QED), whose Lagrangian reads:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 + (D_\mu\phi)^\star(D^\mu\phi) - m^2\phi^\star\phi, \quad (2.20)$$

where the covariant derivative is defined as  $D_\mu\phi \equiv (\partial_\mu + ieA_\mu)\phi$ .

Expanding the covariant derivatives, we obtain:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 + \partial_\mu\phi^\star\partial^\mu\phi + ieA_\mu(\phi\partial^\mu\phi^\star - \phi^\star\partial^\mu\phi) + e^2A_\mu A^\mu\phi^\star\phi - m^2\phi^\star\phi. \quad (2.21)$$

The field equations for  $\phi$  and  $\phi^\star$  derived from this Lagrangian are:

$$(\square + m^2)\phi = -2ieA_\mu\partial^\mu\phi + e^2A_\mu^2\phi, \quad (2.22)$$

$$(\square + m^2)\phi^\star = +2ieA_\mu\partial^\mu\phi^\star + e^2A_\mu^2\phi^\star. \quad (2.23)$$

Let us now identify the Noether current associated with the global U(1) symmetry, where the infinitesimal variations are  $\delta\phi = -i\alpha\phi$  and  $\delta\phi^\star = i\alpha\phi^\star$ . Applying the standard Noether procedure yields:

$$J_\mu = \sum_n \frac{\partial\mathcal{L}}{\partial(\partial^\mu\phi_n)} \frac{\delta\phi_n}{\delta\alpha} = -i(\phi\partial_\mu\phi^\star - \phi^\star\partial_\mu\phi) - 2eA_\mu\phi^\star\phi. \quad (2.24)$$

Here, the first term reproduces the Noether current of the free theory (i.e., when  $e = 0$ ). The full current remains conserved when evaluated on the equations of motion. One can explicitly verify that  $\partial^\mu J_\mu = 0$  holds, confirming the consistency of the theory.

It is also illuminating to understand how the coupling between the gauge field  $A_\mu$  and the current  $J_\mu$  generically emerges. Suppose  $\mathcal{L}_0$  denotes the globally symmetric part of the Lagrangian (i.e., the Lagrangian with  $A_\mu = 0$ ). While  $\mathcal{L}_0$  is invariant under constant phase rotations, allowing  $\alpha$  to vary with space-time introduces a variation of the form:

$$\delta\mathcal{L}_0 = (\partial_\mu\alpha)J^\mu + O(\alpha^2). \quad (2.25)$$

This form follows from the general structure of the variation, where only derivatives of  $\alpha$  can appear for a local transformation.

As an explicit example, consider scalar QED with  $A_\mu = 0$ , so that

$$\mathcal{L}_0 = (\partial_\mu\phi^\star)(\partial^\mu\phi) - m^2\phi^\star\phi, \quad (2.26)$$

and the corresponding variation is

$$\delta\mathcal{L}_0 = (\partial_\mu\alpha)J^\mu + (\partial_\mu\alpha)^2\phi^\star\phi. \quad (2.27)$$



Focusing only on the linear term, and integrating by parts, we find:

$$\delta \mathcal{L}_0 = -\alpha \partial^\mu J_\mu. \quad (2.28)$$

Thus, requiring that the action be invariant under local transformations implies that  $\partial^\mu J_\mu = 0$ , which is precisely the condition of current conservation. This argument provides an alternative derivation of Noether's theorem.

To render the full theory invariant under local U(1) transformations without relying on the equations of motion, one introduces a gauge field  $A_\mu$  that transforms as

$$\delta A_\mu = \partial_\mu \alpha. \quad (2.29)$$

One then modifies the Lagrangian to

$$\mathcal{L} = \mathcal{L}_0 - A_\mu J^\mu. \quad (2.30)$$

Under a gauge transformation, this becomes:

$$\delta \mathcal{L} = (\partial_\mu \alpha) J^\mu - (\partial_\mu \alpha) J^\mu = 0, \quad (2.31)$$

showing that the combination is gauge invariant. Therefore, the interaction term  $A_\mu J^\mu$  is not only natural, but also universal: it emerges as a necessary consequence of promoting a global symmetry to a local one, ensuring gauge invariance of the full theory.

## 2.5 Quantum Operators and the Structure of Vector Propagators

To quantize a massive spin-1 field, we promote the classical field  $A_\mu(x)$  to an operator-valued distribution. The quantum field operator is then expressed as

$$A_\mu(x) = \int \frac{d^3 p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_p}} \sum_{j=1}^3 \left( \epsilon_\mu^j(p) a_{p,j} e^{-ip \cdot x} + \epsilon_\mu^{j*}(p) a_{p,j}^\dagger e^{ip \cdot x} \right), \quad (2.32)$$

where  $\omega_p = \sqrt{\vec{p}^2 + m^2}$ , and  $\epsilon_\mu^j(p)$  denote a canonical basis of polarization vectors satisfying  $p^\mu \epsilon_\mu^j(p) = 0$  and  $\epsilon_\mu^j(p) \epsilon^{k\mu*}(p) = -\delta^{jk}$ . The operators  $a_{p,j}^\dagger$  and  $a_{p,j}$  create and annihilate particles with momentum  $p$  and polarization  $j$ , respectively.

The one-particle state is defined by acting with a creation operator on the vacuum:

$$a_{p,j}^\dagger |0\rangle = \frac{1}{\sqrt{2\omega_p}} |p, \epsilon^j\rangle, \quad (2.33)$$

up to an overall normalization. The field operator creates the corresponding one-particle state as:

$$\langle 0|A_\mu(x)|p, \epsilon^j\rangle = \epsilon_\mu^j(p)e^{-ip\cdot x}, \quad (2.34)$$

demonstrating that the field creates a localized excitation with a definite polarization, which can be projected out by contracting with the polarization vector.

The two-point correlation function or propagator of the theory is defined as the vacuum expectation value of the time-ordered product:

$$\langle 0|T\{A^\mu(x)A^\nu(y)\}|0\rangle = D^{\mu\nu}(x-y) = \int \frac{d^4p}{(2\pi)^4} e^{ip\cdot(x-y)} D^{\mu\nu}(p). \quad (2.35)$$

To compute the momentum-space propagator  $D^{\mu\nu}(p)$ , we begin by writing down the quadratic part of the Proca Lagrangian,

$$\mathcal{L} = -\frac{1}{2}(\partial_\mu A_\nu)(\partial^\mu A^\nu) + \frac{1}{2}(\partial_\mu A_\nu)(\partial^\nu A^\mu) + \frac{1}{2}m^2 A_\mu A^\mu. \quad (2.36)$$

In momentum space, this becomes

$$\mathcal{L} = \frac{1}{2}\tilde{A}_\mu(-p) \left[ -p^2\eta^{\mu\nu} + p^\mu p^\nu + m^2\eta^{\mu\nu} \right] \tilde{A}_\nu(p). \quad (2.37)$$

We define the kinetic operator in momentum space as

$$O^{\mu\nu}(p) = -(p^2 - m^2)\eta^{\mu\nu} + p^\mu p^\nu. \quad (2.38)$$

The Feynman propagator is the inverse of this operator, satisfying the identity

$$O^{\mu\lambda}(p)D_{\lambda\nu}(p) = i\delta_\nu^\mu. \quad (2.39)$$

Inverting the operator yields the explicit expression for the Proca propagator:

$$D_{\mu\nu}(p) = \frac{-i}{p^2 - m^2 + i\epsilon} \left( \eta_{\mu\nu} - \frac{p_\mu p_\nu}{m^2} \right). \quad (2.40)$$

This form ensures that the propagator respects Lorentz invariance and correctly encodes the three physical degrees of freedom associated with a massive spin-1 particle. The term proportional to  $p_\mu p_\nu/m^2$  ensures the decoupling of unphysical modes, consistent with the transversality condition  $p^\mu \epsilon_\mu = 0$ .

### 2.5.1 Massive Spin-1 Field

Embedding a massive spin-1 particle into a relativistic field theory requires accommodating exactly three propagating degrees of freedom. Since a four-vector field  $A_\mu(x)$  has four

components, it presents the minimal tensor structure into which such a particle can be embedded. However, not all four components correspond to physical degrees of freedom: a spin-1 representation of the Lorentz group contains three degrees of freedom, while the remaining component typically corresponds to a scalar (spin-0) mode. In group-theoretic terms, this decomposition is reflected in  $4 = 3 \oplus 1$  under  $SO(3)$  rotations. Our objective is to construct a Lagrangian such that only the spin-1 component propagates, and the unphysical mode is dynamically removed.

A naive attempt might be to start with the Lagrangian

$$\mathcal{L} = -\frac{1}{2}(\partial_\nu A_\mu)(\partial^\nu A^\mu) + \frac{1}{2}m^2 A_\mu A^\mu, \quad (2.41)$$

which leads to the equations of motion

$$(\square + m^2) A_\mu = 0. \quad (2.42)$$

At first glance, this appears to describe four independent massive degrees of freedom. However, a closer inspection reveals that each component  $A_\mu$  behaves as a free scalar field with mass  $m$ . This implies that the theory actually describes four massive spin-0 particles, and thus fails to reproduce the desired spin-1 spectrum.

The energy density corresponding to this Lagrangian is

$$\begin{aligned} \mathcal{E} &= \frac{\partial \mathcal{L}}{\partial(\partial_t A_\mu)} \partial_t A_\mu - \mathcal{L} \\ &= -\frac{1}{2} \left[ (\partial_t A_0)^2 + (\vec{\nabla} A_0)^2 + m^2 A_0^2 \right] + \frac{1}{2} \left[ (\partial_t \vec{A})^2 + (\nabla_i A_j)^2 + m^2 \vec{A}^2 \right], \end{aligned} \quad (2.43)$$

where  $\vec{A}$  denotes the spatial components of  $A_\mu$ . The presence of a negative contribution from  $A_0$  indicates an unbounded energy density, a clear sign of instability. Even reversing the overall sign of the Lagrangian would not resolve the issue, as it would simply shift the instability to the spatial components.

This raises a subtle but important question: how do we know whether  $A_\mu$  truly transforms as a Lorentz four-vector, rather than simply a set of four scalar fields? The Lagrangian above is formally invariant under both interpretations. However, in field theory, one does not impose transformation properties by hand — rather, one defines a Lagrangian and then checks whether the resulting dynamics respect a given symmetry. If the Lagrangian is properly constructed, the symmetry will manifest itself in physical observables such as matrix elements, even if it is not manifest at the level of the fields themselves.

For example, the classical Maxwell equations are Lorentz invariant regardless of whether one formulates the theory in terms of electric and magnetic fields  $(\vec{E}, \vec{B})$  or the four-potential  $A_\mu$ . Although the symmetry is more transparent in the latter case, it remains

valid in either description. Historically, many foundational advances in physics emerged by recognizing hidden symmetries in observed phenomena and retroactively identifying their underlying field-theoretic origins. Such was the case in electromagnetism, special and general relativity, weak interactions, and the quark model, among others.

Returning to the spin-1 case, one can write down a more general Lorentz-invariant Lagrangian involving two-derivative terms:

$$\mathcal{L} = \frac{a}{2} A_\mu \square A^\mu + \frac{b}{2} A_\mu \partial^\mu \partial^\nu A_\nu + \frac{1}{2} m^2 A_\mu A^\mu, \quad (2.44)$$

with arbitrary coefficients  $a$  and  $b$ . The term involving  $\partial^\mu A_\mu$  ensures that  $A_\mu$  must transform as a four-vector, since this divergence is not Lorentz invariant for generic field configurations unless  $A_\mu$  carries vectorial transformation properties. Consequently, this structure allows for a decomposition  $4 = 3 \oplus 1$  under  $\text{SO}(3)$ , raising the possibility of isolating the physical spin-1 component.

The equations of motion derived from this Lagrangian are

$$a \square A_\mu + b \partial_\mu (\partial \cdot A) + m^2 A_\mu = 0. \quad (2.45)$$

Taking the divergence yields a secondary equation:

$$[(a + b) \square + m^2] (\partial_\mu A^\mu) = 0. \quad (2.46)$$

Choosing  $a = -b$  and  $m \neq 0$  enforces the Lorentz-invariant constraint  $\partial_\mu A^\mu = 0$ , which eliminates a single degree of freedom. Since the spin-0 mode corresponds to one real scalar component, this condition projects out precisely that unwanted piece.

With the choice  $a = 1$ ,  $b = -1$ , the Lagrangian simplifies to the form

$$\begin{aligned} \mathcal{L} &= \frac{1}{2} A_\mu \square A^\mu - \frac{1}{2} A_\mu \partial^\mu \partial^\nu A_\nu + \frac{1}{2} m^2 A_\mu A^\mu \\ &= -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{1}{2} m^2 A_\mu A^\mu, \end{aligned} \quad (2.47)$$

where we have introduced the field strength tensor  $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ . This Lagrangian is known as the *Proca Lagrangian*. Notably, the appearance of  $F_{\mu\nu}$  arises here not from a requirement of gauge invariance or electromagnetic analogy, but purely from the demand to eliminate the unphysical degree of freedom in a Lorentz-invariant way.

The equations of motion following from the Proca Lagrangian are

$$(\square + m^2) A_\mu = 0, \quad \partial_\mu A^\mu = 0. \quad (2.48)$$

These describe precisely three propagating modes, as required for a massive spin-1 particle.

### 2.5.2 Massless Spin-1 Field

A natural strategy to construct a theory for massless spin-1 particles is to consider the massless limit of the Proca theory. Setting  $m \rightarrow 0$  in the Lagrangian leads to

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \quad (2.49)$$

which coincides with the standard Lagrangian of classical electrodynamics. This correspondence strongly suggests that we are on the correct theoretical track. However, taking the massless limit introduces certain subtleties that must be addressed.

First, the constraint equation that eliminated the unphysical degree of freedom in the Proca theory,

$$m^2 \partial_\mu A^\mu = 0, \quad (2.50)$$

becomes trivial when  $m = 0$ , implying that  $\partial_\mu A^\mu = 0$  no longer follows automatically. This opens the possibility for an additional, potentially unphysical, mode to propagate — namely, the spin-0 component we had previously projected out.

A second issue emerges when analyzing the behavior of the longitudinal polarization in the massless limit. Recall that the longitudinal polarization vector in the Proca theory is given by

$$\epsilon_\mu^L = \left( \frac{p_z}{m}, 0, 0, \frac{E}{m} \right), \quad (2.51)$$

which diverges as  $m \rightarrow 0$ . This divergence is partly due to normalization, and in the  $m \rightarrow 0$  limit, one finds that  $\epsilon_\mu^L$  becomes proportional to  $p_\mu$ . Since massless particles satisfy  $p^2 = 0$ , this corresponds to a light-like polarization direction. Thus, we expect that  $\epsilon_\mu^L \sim p_\mu$  in the limit  $m \rightarrow 0$ , modulo normalization.

Nevertheless, the irreducible representation of the Poincaré group corresponding to a massless spin-1 particle is known to contain only two physical polarizations. Therefore, both the longitudinal mode and the scalar degree of freedom must somehow decouple from the physical spectrum.

Rather than continuing from the massive case, it is more illuminating to start from the massless Lagrangian directly and perform a careful analysis of the gauge structure and degrees of freedom. We begin with:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \quad (2.52)$$

This Lagrangian enjoys a local gauge symmetry:

$$A_\mu(x) \rightarrow A_\mu(x) + \partial_\mu \alpha(x), \quad (2.53)$$

where  $\alpha(x)$  is an arbitrary scalar function. Consequently, two vector potentials differing by the gradient of a scalar are physically indistinguishable — they represent the same field configuration.

The equations of motion obtained from the Lagrangian are

$$\square A_\mu - \partial_\mu(\partial_\nu A^\nu) = 0, \quad (2.54)$$

which, when decomposed into temporal and spatial components, become

$$-\nabla^2 A_0 + \partial_t \nabla \cdot \vec{A} = 0, \quad (2.55)$$

$$\square \vec{A} - \nabla \left( \partial_t A_0 - \nabla \cdot \vec{A} \right) = 0. \quad (2.56)$$

To determine the physical content of the theory, we proceed by fixing a gauge — that is, by exploiting the redundancy introduced by the gauge symmetry to eliminate unphysical degrees of freedom. One convenient choice is the Coulomb gauge,

$$\nabla \cdot \vec{A} = 0. \quad (2.57)$$

This can always be imposed (under mild regularity assumptions) by choosing an appropriate gauge function  $\alpha(x)$ . Under this condition, the equation for  $A_0$  becomes

$$\nabla^2 A_0 = 0, \quad (2.58)$$

which contains no time derivatives. Consequently,  $A_0$  is not a dynamical degree of freedom. Moreover, residual gauge transformations preserving the Coulomb condition must satisfy  $\nabla^2 \alpha = 0$ , and since  $A_0 \rightarrow A_0 + \partial_t \alpha$ , we can use this residual freedom to set  $A_0 = 0$  entirely. We have thus eliminated one unphysical degree of freedom.

With  $A_0 = 0$  and  $\nabla \cdot \vec{A} = 0$ , the spatial components satisfy

$$\square A_i = 0, \quad (2.59)$$

subject to the divergence-free constraint. These equations describe three fields obeying wave equations, but the constraint reduces the number of independent solutions to two. Therefore, only two transverse modes propagate, in accordance with the known structure of massless spin-1 representations.

This can also be made explicit in Fourier space. Expanding the field as

$$A_\mu(x) = \int \frac{d^4 p}{(2\pi)^4} \epsilon_\mu(p) e^{ip \cdot x}, \quad (2.60)$$

we impose the on-shell condition  $p^2 = 0$ , the gauge condition  $p^i \epsilon_i = 0$ , and  $\epsilon_0 = 0$ . Choosing a reference frame where  $p_\mu = (E, 0, 0, E)$ , one finds two linearly independent solutions:

$$\epsilon_\mu^1 = (0, 1, 0, 0), \quad \epsilon_\mu^2 = (0, 0, 1, 0), \quad (2.61)$$

which correspond to linear polarization states of the photon.

Alternatively, one can employ a helicity basis using circular polarizations:

$$\epsilon_\mu^R = \frac{1}{\sqrt{2}}(0, 1, i, 0), \quad \epsilon_\mu^L = \frac{1}{\sqrt{2}}(0, 1, -i, 0). \quad (2.62)$$

These correspond to eigenstates of the spin projection along the direction of motion and are often used in the analysis of scattering amplitudes.

Had we chosen Lorenz gauge,  $\partial_\mu A^\mu = 0$ , we would have found three solutions to  $p^\mu \epsilon_\mu = 0$ . Besides the two physical polarizations above, one finds a third vector,

$$\epsilon_\mu^f = (1, 0, 0, 1), \quad (2.63)$$

referred to as the forward or gauge mode. This mode is not physical: it is proportional to  $p^\mu$  and therefore corresponds to a pure gauge configuration  $A_\mu = \partial_\mu \phi$ , which can be removed by an appropriate choice of  $\alpha(x)$ . Moreover, it fails to satisfy the usual normalization condition  $\epsilon_\mu^* \epsilon^\mu = -1$ .

Similarly, another unphysical mode can appear if one does not impose  $\epsilon_0 = 0$ , namely

$$\epsilon_\mu^0 = (1, 0, 0, 0), \quad (2.64)$$

which again cannot be normalized consistently and represents a timelike polarization.

To summarize: for the massive spin-1 case, the Proca Lagrangian imposes the constraint  $\partial_\mu A^\mu = 0$ , removing one unphysical mode and retaining the three physical degrees of freedom. The theory lacks gauge symmetry, but the  $F_{\mu\nu} F^{\mu\nu}$  kinetic term remains essential for positive energy. In the massless case, gauge invariance becomes a fundamental feature, and through appropriate gauge fixing we can eliminate both scalar and longitudinal polarizations, leaving two transverse physical modes — the correct number for a massless spin-1 field. These polarization vectors depend on momentum  $p^\mu$ , so the full representation space is infinite-dimensional.

Finally, the massless and massive representations of the Poincaré group differ in their little groups. For a massive particle with four-momentum  $p^\mu = (m, 0, 0, 0)$ , the little group is  $\text{SO}(3)$ , with spin- $J$  representations of dimension  $2J + 1$ . For massless particles with  $p^\mu = (E, 0, 0, E)$ , the little group becomes  $\text{ISO}(2)$  — the isometry group of the Euclidean plane — which admits representations with only two helicity states for each  $J$ .

## 2.6 Identifying Ghosts: Stability Constraints in Vector Theories

In quantum field theory, a *ghost* refers to a state with either a negative norm or a wrong-sign kinetic term. An example is the  $A_0$  component in a vector field when described by the Lagrangian  $\mathcal{L} = \frac{1}{2}A_\mu(\square + m^2)A^\mu$ . To identify such pathological degrees of freedom, we analyze the field by separating it into transverse and longitudinal parts. This technique, originally developed for Abelian gauge theories by Stueckelberg [11, 12], was extended to non-Abelian cases by Coleman and collaborators [13, 14], and to gravity by Arkani-Hamed et al. [15].

Consider a general decomposition of the vector field:

$$A_\mu(x) = A_\mu^T(x) + \partial_\mu\pi(x), \quad (2.65)$$

subject to the constraint

$$\partial^\mu A_\mu^T = 0. \quad (2.66)$$

This decomposition is not unique—shifts  $A_\mu^T \rightarrow A_\mu^T + \partial_\mu\alpha$  and  $\pi \rightarrow \pi - \alpha$  leave  $A_\mu$  invariant. However, by choosing an appropriate gauge, one can always impose the condition (2.66).

This parametrization makes it easy to track the behavior of longitudinal modes. Consider the most general Lorentz-invariant quadratic Lagrangian for a vector field:

$$\mathcal{L} = aA_\mu\square A^\mu + bA_\mu\partial^\mu\partial^\nu A_\nu + m^2A_\mu A^\mu. \quad (2.67)$$

Substituting the decomposition into the Lagrangian and using the transversality condition yields

$$\mathcal{L} = aA_\mu^T\square A^{T\mu} + m^2A_\mu^T A^{T\mu} - (a+b)\pi\square^2\pi - m^2\pi\square\pi. \quad (2.68)$$

We see that the scalar mode  $\pi$  acquires a higher-derivative kinetic term unless  $a+b=0$ . The presence of four-derivative terms in the kinetic part signals the existence of ghost degrees of freedom and non-unitarity.

To make this explicit, consider the propagator of  $\pi$ . In momentum space, the two-point function reads

$$\Pi_\pi(k) = \frac{-1}{2(a+b)k^4 - 2m^2k^2} = \frac{1}{2m^2} \left[ \frac{1}{k^2} - \frac{a+b}{(a+b)k^2 - m^2} \right]. \quad (2.69)$$

This expression reveals two propagating modes, one of which necessarily carries a negative residue for generic  $a+b \neq 0$ , indicating a ghost. Only when  $a+b=0$  does the propagator reduce to  $1/(m^2k^2)$ , avoiding ghosts and ensuring unitarity.



Therefore, demanding ghost-free propagation uniquely selects the Proca Lagrangian. Setting  $a = -b = 1/2$  and rescaling  $m^2 \rightarrow \frac{1}{2}m^2$ , we obtain

$$\mathcal{L} = \frac{1}{2}A_\mu \Box A^\mu - \frac{1}{2}A_\mu \partial^\mu \partial^\nu A_\nu + \frac{1}{2}m^2 A_\mu A^\mu = -\frac{1}{4}F_{\mu\nu}^2 + \frac{1}{2}m^2 A_\mu A^\mu. \quad (2.70)$$

In this formulation, the longitudinal mode acquires a kinetic term proportional to the mass, ensuring its proper dynamics.

However, in the massless limit, the longitudinal mode loses its kinetic term and becomes non-dynamical. A mode with interaction but no kinetic energy leads to an ill-defined quantum theory. For instance, consider

$$\mathcal{L} = Z\pi\Box\pi + \lambda\pi^3, \quad (2.71)$$

and rescale  $\pi$  to a canonically normalized field  $\pi_c = \sqrt{Z}\pi$ , yielding

$$\mathcal{L} = \pi_c \Box \pi_c + \frac{\lambda}{Z^{3/2}} \pi_c^3. \quad (2.72)$$

As  $Z \rightarrow 0$ , the interaction becomes non-perturbative, signaling a breakdown of the theory.

To avoid this,  $\pi$  must never appear explicitly, which is ensured by demanding gauge invariance under  $A_\mu \rightarrow A_\mu + \partial_\mu \pi$ . For interacting theories, this is satisfied if  $A_\mu$  couples to a conserved current:

$$\mathcal{L} \supset A_\mu J^\mu, \quad \text{with } \partial_\mu J^\mu = 0. \quad (2.73)$$

To examine interactions, consider a real scalar field  $\phi$ . The simplest Lorentz-invariant interaction term,

$$\mathcal{L}_{\text{int}} = A_\mu \phi \partial^\mu \phi, \quad (2.74)$$

is not gauge invariant. Moreover, no field redefinition of  $\phi$  can restore invariance. For a complex field, the only allowed gauge-invariant coupling is

$$\mathcal{L}_{\text{int}} = -iA_\mu (\phi^\star \partial^\mu \phi - \phi \partial^\mu \phi^\star). \quad (2.75)$$

Upon substituting  $A_\mu \rightarrow A_\mu + \partial_\mu \pi$ , the Lagrangian acquires an extra term

$$i\pi (\phi^\star \Box \phi - \phi \Box \phi^\star). \quad (2.76)$$

This term can be canceled by postulating that  $\phi$  transforms as  $\phi \rightarrow \phi - i\pi\phi$ , so that the kinetic term transforms as

$$(\partial_\mu \phi^\star) (\partial^\mu \phi) \rightarrow (\partial_\mu \phi^\star) (\partial^\mu \phi) - i\pi (\phi^\star \Box \phi - \phi \Box \phi^\star) - (\pi \phi^\star) \Box (\pi \phi). \quad (2.77)$$

The unwanted term is canceled, though higher-order terms in  $\pi$  appear. These can be absorbed by completing the transformation rule for  $\phi$  to second order:

$$\phi \rightarrow \phi - i\pi\phi - \frac{1}{2}\pi^2\phi. \quad (2.78)$$

Including all such terms, we arrive at the fully gauge-invariant scalar QED Lagrangian:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 + (\partial_\mu\phi^\star)(\partial^\mu\phi) - iA_\mu(\phi^\star\partial^\mu\phi - \phi\partial^\mu\phi^\star) + A_\mu^2\phi^\star\phi. \quad (2.79)$$

This construction ensures gauge invariance under  $A_\mu \rightarrow A_\mu + \partial_\mu\pi$  and  $\phi \rightarrow e^{-i\pi}\phi$ , demonstrating that scalar QED can be derived from first principles by eliminating ghost degrees of freedom and demanding gauge invariance.

In this chapter, we have developed a systematic understanding of spin-1 fields by examining the interplay between Lorentz invariance and unitarity, embedding irreducible representations into field-theoretic frameworks, and identifying the conditions under which consistent, ghost-free Lagrangians can be constructed. We explored both massive and massless vector fields, highlighted the role of gauge symmetry, and derived the Proca Lagrangian as the unique healthy theory for a massive spin-1 particle. We also examined how these principles manifest in interactions, quantization, and propagators, providing the foundation for understanding how spin-1 fields behave in quantum field theory.

In the next chapter, we extend this analysis by constructing more general and non-linear interactions for vector fields. This includes theories with derivative self-interactions such as the Galileon and its covariant extensions, as well as the broader class of Generalized Proca theories. These models will allow us to probe the dynamics of spin-1 fields beyond the standard Proca framework, particularly in curved space-times and cosmological contexts.

## Chapter 3

# Construction of Self-Derivative Interactions

In the previous chapter, we focused on the formulation and properties of Proca theory, emphasizing the propagation of three degrees of freedom for a massive spin-1 field. As we now aim to generalize this framework, a natural next step is to explore the possibility of adding derivative self-interactions. These arise in many contexts, including effective field theories and modified gravity models, where one seeks to go beyond standard kinetic terms while maintaining consistency.

Higher-derivative terms naturally emerge in various extensions of classical field theories. Their appearance is often motivated by quantum corrections, renormalization group flows, or attempts to encode richer infrared dynamics in gravity. Prominent examples include the Horndeski [16] and generalized Proca frameworks, both of which introduce derivative interactions in a carefully controlled way. However, incorporating such terms must be done with great care. If not appropriately constrained, higher derivatives can lead to pathological instabilities and spoil the physical viability of the theory.

The goal of this chapter is to systematically construct self-derivative interactions for a massive vector field that avoid these pathologies. In particular, we require that the equations of motion remain of second order in time derivatives and that no extra ghost degrees of freedom are introduced. As we will see, these criteria are deeply connected to a classical result known as the Ostrogradsky instability, which we begin by reviewing.

### 3.1 Ostrogradsky Instability

Before constructing consistent higher-derivative interactions, it is instructive to revisit a fundamental obstruction that arises in their naive implementation. This obstruction is known as the *Ostrogradsky instability* [17, 18, 19], a result established in the 19th century by Mikhail Ostrogradsky. It applies to non-degenerate systems whose Lagrangians contain higher-order time derivatives.

Ostrogradsky's theorem places a strong constraint on the formulation of any theory that involves such terms: unless the theory is degenerate, the associated Hamiltonian is unbounded from below. This implies the existence of arbitrarily negative energy states, which leads to catastrophic instabilities. These instabilities manifest as violations of unitarity and vacuum decay, rendering the theory physically inconsistent.

In what follows, we will explore the mathematical structure of this instability in detail, laying the groundwork for constructing theories that circumvent it by enforcing specific degeneracy conditions.

To clarify the problem, let us consider a mechanical system described by a Lagrangian that depends on second-order time derivatives:

$$L = L(q, \dot{q}, \ddot{q}). \quad (3.1)$$

The Euler-Lagrange equations for this system are

$$\frac{\partial L}{\partial q} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}} \right) + \frac{d^2}{dt^2} \left( \frac{\partial L}{\partial \ddot{q}} \right) = 0. \quad (3.2)$$

The critical assumption in Ostrogradsky's theorem is *non-degeneracy*, i.e., that the highest-derivative term enters the Lagrangian in a nontrivial way:

$$\frac{\partial^2 L}{\partial \ddot{q}^2} \neq 0. \quad (3.3)$$

Under this assumption, one finds that the theory propagates additional degrees of freedom. The system's phase space must then be extended to accommodate initial data not just for  $(q, \dot{q})$ , but also for  $(\ddot{q}, \dddot{q})$ . Ostrogradsky introduced a canonical formalism in which the coordinates and conjugate momenta are defined as

$$X_1 = q, \quad P_1 = \frac{\partial L}{\partial \dot{q}} - \frac{d}{dt} \left( \frac{\partial L}{\partial \ddot{q}} \right), \quad (3.4)$$

$$X_2 = \dot{q}, \quad P_2 = \frac{\partial L}{\partial \ddot{q}}. \quad (3.5)$$

The resulting Hamiltonian is given by the Legendre transform:

$$H(X_1, X_2, P_1, P_2) = P_1 X_2 + P_2 A(X_1, X_2, P_2) - L(X_1, X_2, A), \quad (3.6)$$

where  $A$  is a function implicitly defined by the relation between  $P_2$  and  $\ddot{q}$ .

A key feature of this Hamiltonian is its linear dependence on  $P_1$ , which implies that it is unbounded from below. In physical terms, this allows the system to lower its energy without bound by exciting the ghost mode associated with  $P_1$ , leading to catastrophic vacuum instabilities.

This issue becomes manifest in a simple toy model: a harmonic oscillator modified by a higher-derivative term:

$$L = \frac{m}{2} \dot{q}^2 - \frac{m\omega^2}{2} q^2 - \frac{\epsilon m}{2\omega^2} \ddot{q}^2, \quad (3.7)$$

where  $\epsilon$  is a small parameter. Upon performing the canonical analysis, one finds that the Hamiltonian includes ghost-like excitations with negative energy, confirming the presence of an instability.

To avoid this problem, one must ensure that the theory is *degenerate*, meaning that the equations of motion remain second order despite the presence of higher-derivative terms in the Lagrangian [8]. This can be achieved through special structures or cancellations. For example, in scalar-tensor theories, Horndeski's construction systematically builds the most general Lagrangian that leads to second-order field equations. A similar logic underlies the generalized Proca theories, where derivative self-interactions are introduced for a massive vector field without violating the second-order condition or introducing additional propagating modes.

In conclusion, the Ostrogradsky instability sets a strong theoretical constraint on viable field theories. Any extension involving higher derivatives must be crafted carefully to avoid ghost degrees of freedom. This motivates the study of specially constructed interactions—such as those found in Galileon and generalized Proca frameworks—which preserve stability by keeping the equations of motion second order. The rest of this chapter is devoted to systematically building such interactions.

## 3.2 Galileons

Having established that higher-derivative theories can remain stable if they are constructed to yield second-order equations of motion, we now turn to an explicit class of such interac-

tions: the Galileons [20, 21, 22]. These theories serve as a key example of how non-trivial derivative self-interactions can be incorporated without introducing Ostrogradsky ghosts.

Galileon theories were originally developed in the context of scalar fields, where a specific symmetry—known as the Galilean symmetry—plays a central role in constraining the form of the interactions. We begin by reviewing the scalar case, which not only provides a valuable prototype but also lays the foundation for later generalizations to vector fields.

### 3.2.1 Scalar Galileons

Scalar field theories with non-canonical kinetic terms—i.e., those constructed from functions of  $(\partial\pi)^2$ —do not alter the number of propagating degrees of freedom, as only first derivatives appear and the equations of motion remain second order. However, one may wonder whether more general self-interactions involving *two* derivatives per field can also avoid introducing Ostrogradsky instabilities.

Remarkably, this is possible. A concrete example arises in the decoupling limit of the five-dimensional Dvali-Gabadadze-Porrati (DGP) brane-world model [23], where the helicity-0 component of a massive graviton acquires a self-interaction of the form

$$\mathcal{L}_3 = (\partial\pi)^2 \square\pi. \quad (3.8)$$

At first glance, this term appears problematic due to the presence of higher derivatives. Yet, the equation of motion derived from it,

$$\mathcal{E}_3 = (\square\pi)^2 - (\partial_\mu\partial_\nu\pi)^2, \quad (3.9)$$

contains no more than two time derivatives, thus avoiding the Ostrogradsky instability.

In addition to the favorable derivative structure, the interaction is invariant under a symmetry known as the *Galilean shift*:

$$\pi(x) \rightarrow \pi(x) + c + b_\mu x^\mu, \quad (3.10)$$

where  $c$  and  $b_\mu$  are constant parameters. This symmetry is a remnant of five-dimensional Poincaré invariance in the DGP setup and constrains the form of allowed interactions.

Motivated by this example, one can ask whether more such derivative self-interactions can be constructed [20]. Proceeding order-by-order, we can build interactions with increasing powers of  $\pi$  and derivatives, ensuring that the equations of motion remain second order. For instance, consider quartic self-interactions:

$$\mathcal{L}_4 = \alpha_1 (\partial\pi)^2 (\square\pi)^2 + \alpha_2 (\partial\pi)^2 (\partial_\mu\partial_\nu\pi)^2. \quad (3.11)$$

Although this expression involves two derivatives per field, it generically gives rise to higher-than-second-order equations of motion unless the coefficients are tuned to cancel such terms. Specifically, third time derivatives cancel if we impose  $\alpha_2 = -\alpha_1$ , yielding

$$\mathcal{L}_4 = (\partial\pi)^2 [(\Box\pi)^2 - (\partial_\mu\partial_\nu\pi)^2]. \quad (3.12)$$

Continuing this systematic construction leads to a finite set of Galilean-invariant interactions in four dimensions. These five terms, known as the scalar Galileons, are given by:

$$\mathcal{L}_1 = \pi, \quad (3.13)$$

$$\mathcal{L}_2 = (\partial\pi)^2, \quad (3.14)$$

$$\mathcal{L}_3 = (\partial\pi)^2\Box\pi, \quad (3.15)$$

$$\mathcal{L}_4 = (\partial\pi)^2 [(\Box\pi)^2 - (\partial_\mu\partial_\nu\pi)^2], \quad (3.16)$$

$$\mathcal{L}_5 = (\partial\pi)^2 [(\Box\pi)^3 - 3\Box\pi(\partial_\mu\partial_\nu\pi)^2 + 2(\partial_\mu\partial_\nu\pi)^3]. \quad (3.17)$$

Each of these terms satisfies the Galilean symmetry and gives rise to second-order equations of motion, thus preserving stability.

These interactions can also be defined recursively [8] using the equations of motion  $E_n = \delta\mathcal{L}_n/\delta\pi$ , through the relation

$$\mathcal{L}_{n+1} = -(\partial\pi)^2 E_n, \quad (3.18)$$

with

$$E_1 = 1, \quad (3.19)$$

$$E_2 = \Box\pi, \quad (3.20)$$

$$E_3 = (\Box\pi)^2 - (\partial_\mu\partial_\nu\pi)^2, \quad (3.21)$$

$$E_4 = (\Box\pi)^3 - 3\Box\pi(\partial_\mu\partial_\nu\pi)^2 + 2(\partial_\mu\partial_\nu\pi)^3, \quad (3.22)$$

$$E_5 = (\Box\pi)^4 - 6(\Box\pi)^2(\partial_\mu\partial_\nu\pi)^2 + 8\Box\pi(\partial_\mu\partial_\nu\pi)^3 + 3[(\partial_\mu\partial_\nu\pi)^2]^2 - 6(\partial_\mu\partial_\nu\pi)^4. \quad (3.23)$$

The construction of these Lagrangians is greatly simplified by employing antisymmet-

ric Levi-Civita tensors. The Galileon terms can be expressed compactly as

$$\begin{aligned}
\mathcal{L}_1 &= \pi \epsilon^{\mu\alpha\rho\sigma} \epsilon_{\mu\alpha\rho\sigma}, \\
\mathcal{L}_2 &= \pi \epsilon^{\mu\alpha\rho\sigma} \epsilon^\nu_{\alpha\rho\sigma} \Pi_{\mu\nu}, \\
\mathcal{L}_3 &= \pi \epsilon^{\mu\alpha\rho\sigma} \epsilon^{\nu\beta}_{\rho\sigma} \Pi_{\mu\nu} \Pi_{\alpha\beta}, \\
\mathcal{L}_4 &= \pi \epsilon^{\mu\alpha\rho\sigma} \epsilon^{\nu\beta\gamma}_{\sigma} \Pi_{\mu\nu} \Pi_{\alpha\beta} \Pi_{\rho\gamma}, \\
\mathcal{L}_5 &= \pi \epsilon^{\mu\alpha\rho\sigma} \epsilon^{\nu\beta\gamma\delta} \Pi_{\mu\nu} \Pi_{\alpha\beta} \Pi_{\rho\gamma} \Pi_{\sigma\delta},
\end{aligned} \tag{3.24}$$

where  $\Pi_{\mu\nu} = \partial_\mu \partial_\nu \pi$ . Since  $\Pi_{\mu\nu}$  is symmetric, its contraction with two totally antisymmetric tensors  $\epsilon^{\dots}$  is unique at each order. In four dimensions, the maximum number of distinct contractions with Levi-Civita tensors is reached at  $\mathcal{L}_5$ , which is why the Galileon series terminates at that order.

To summarize, scalar Galileon theories provide a consistent framework for including derivative self-interactions of a scalar field without introducing ghost-like degrees of freedom. Their defining features—second-order equations of motion and internal Galilean invariance—will serve as guiding principles in constructing analogous interactions for massive vector fields in the generalized Proca theories.

### 3.2.2 Generalized Proca Theories

Having reviewed the construction of consistent scalar Galileon interactions, we now extend the same logic to massive spin-1 fields [8, 24, 25]. The goal is to build self-interactions for the vector field that preserve the second-order nature of the equations of motion and avoid introducing an additional propagating degree of freedom, which would signal an Ostrogradsky ghost. Specifically, the consistent theory should propagate exactly three degrees of freedom: two transverse and one longitudinal mode, as dictated by the representation theory of the Lorentz group for a massive vector.

To achieve this, we impose the following two consistency conditions:

- The equations of motion must remain second order.
- The temporal component  $A_0$  must remain non-dynamical.

The second condition is essential. Even if the equations of motion are second order, a dynamical  $A_0$  would indicate the presence of a ghostly degree of freedom, since the correct number of physical degrees of freedom for a massive vector is fixed by symmetry.

We begin by noting that the mass term in the Proca action can be generalized to a potential function  $V(A^2)$ , which involves no derivatives. Since  $A_0$  appears algebraically,



such terms do not alter the number of propagating degrees of freedom and are thus trivially safe.

Similarly, gauge-invariant terms constructed purely from the field strength tensor  $F_{\mu\nu}$  and its dual  $\tilde{F}_{\mu\nu}$ ,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \quad \tilde{F}_{\mu\nu} = \frac{1}{2} \epsilon_{\mu\nu\rho\sigma} F^{\rho\sigma}, \quad (3.25)$$

do not contribute to the dynamics of  $A_0$ . For instance, terms like  $F_{\mu\nu} F^{\mu\nu}$  or  $F_{\mu\nu} \tilde{F}^{\mu\nu}$  respect gauge invariance and remain harmless. Interactions involving contractions like  $A^\mu A^\nu F_\mu^\alpha F_{\nu\alpha}$  are also safe, as long as they do not involve derivatives acting on  $A_0$ .

All such contributions can be encapsulated in a generalized function,

$$\mathcal{L}_2 = f_2(A_\mu, F_{\mu\nu}, \tilde{F}_{\mu\nu}), \quad (3.26)$$

which can be further simplified using scalar combinations of its arguments. Defining

$$X = -\frac{1}{2} A_\mu A^\mu, \quad F = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}, \quad Y = A^\mu A^\nu F_\mu^\alpha F_{\nu\alpha}, \quad (3.27)$$

we can rewrite the Lagrangian as

$$\mathcal{L}_2 = f_2(X, F, Y), \quad (3.28)$$

where we neglect parity-violating terms involving  $\tilde{F}^{\mu\nu}$  for simplicity. This constitutes the most general class of interactions that are safe up to zeroth order in derivatives of  $A_\mu$ .

At the next order in derivatives, the only admissible interaction that preserves the constraint structure is

$$\mathcal{L}_3 = f_3(X) \partial_\mu A^\mu, \quad (3.29)$$

where  $f_3(X)$  is an arbitrary function of the norm of the vector field. This interaction does not contribute to the dynamical evolution of  $A_0$  and the corresponding Hessian matrix with respect to time derivatives [8, 10],

$$H_{\mathcal{L}_3}^{\mu\nu} = \frac{\partial^2 \mathcal{L}_3}{\partial \dot{A}_\mu \partial \dot{A}_\nu}, \quad (3.30)$$

vanishes identically, confirming the absence of higher-order time derivatives.

The uniqueness of this interaction becomes more transparent when expressed in terms of Levi-Civita tensors:

$$\mathcal{L}_3 = -\frac{f_3(X)}{6} \epsilon^{\mu\nu\rho\sigma} \epsilon_{\nu\rho\sigma}^\alpha \partial_\mu A_\alpha = f_3(X) \partial_\mu A^\mu. \quad (3.31)$$

This contraction is unique in four dimensions, and in the decoupling limit where  $A_\mu \rightarrow \partial_\mu \pi$ , it generates the cubic Galileon for the scalar mode.

One could consider extending this structure by coupling the Levi-Civita tensors to additional powers of  $A_\mu$ . For instance,

$$\tilde{\mathcal{L}}_3 = \tilde{f}_3(X) \epsilon^{\mu\nu\rho\sigma} \epsilon^{\alpha\beta}_{\rho\sigma} \partial_\mu A_\alpha A_\nu A_\beta, \quad (3.32)$$

leads to terms of the schematic form

$$\tilde{\mathcal{L}}_3 = \tilde{f}_3(X) A^\mu A^\nu \partial_\mu A_\nu, \quad (3.33)$$

which is equivalent to  $\mathcal{L}_3$  up to a disformal transformation of the metric,  $\eta_{\mu\nu} \rightarrow \eta_{\mu\nu} + \tilde{f}_3(X) A_\mu A_\nu$ . Since the physical content remains unchanged, we will not distinguish between such equivalent forms.

Attempts to construct more complex terms of the type  $\epsilon \epsilon A^4 \partial A$  fail due to the anti-symmetric nature of the Levi-Civita tensor: any term involving two symmetric  $A^\mu$  fields contracted into antisymmetric indices vanishes. Even if such terms are written, they can either be absorbed into a redefinition of the metric (as in a disformal transformation) or vanish identically upon integration by parts.

The interactions  $\mathcal{L}_2$  and  $\mathcal{L}_3$  represent the first steps in building a consistent theory of derivative self-interactions for a massive vector field that propagates only three degrees of freedom and avoids the pitfalls of Ostrogradsky instabilities.

In the next order of derivatives, we encounter the first nontrivial structure involving second-order derivative combinations of the vector field. The most general Lorentz-invariant interaction that is quadratic in derivatives of  $A_\mu$  and preserves our consistency conditions can be written as

$$\mathcal{L}_4 = f_4(X) \left[ c_1 (\partial \cdot A)^2 + c_2 \partial_\rho A_\sigma \partial^\rho A^\sigma + c_3 \partial_\rho A_\sigma \partial^\sigma A^\rho \right], \quad (3.34)$$

where  $f_4(X)$  is an arbitrary function of the vector norm  $X = -A_\mu A^\mu/2$ , and  $c_1$ ,  $c_2$ , and  $c_3$  are constant parameters. Although  $c_1$  and  $c_3$  are not independent under integration by parts, we keep all three coefficients to analyze the general structure.

To ensure that this interaction preserves the desired number of propagating degrees of freedom, we must verify that the temporal component  $A_0$  remains non-dynamical. This requirement is met by imposing the vanishing of the determinant of the Hessian matrix

with respect to time derivatives, which reflects the presence of a second-class constraint:

$$H_{\mathcal{L}_4}^{\mu\nu} = \frac{\partial^2 \mathcal{L}_4}{\partial \dot{A}_\mu \partial \dot{A}_\nu} = f_4(X) \begin{pmatrix} 2(c_1 + c_2 + c_3) & 0 & 0 & 0 \\ 0 & -2c_2 & 0 & 0 \\ 0 & 0 & -2c_2 & 0 \\ 0 & 0 & 0 & -2c_2 \end{pmatrix}. \quad (3.35)$$

The Hessian has a vanishing determinant in two special cases [8]:

- $c_2 = 0$ , in which case three eigenvalues vanish and only  $A_0$  propagates.
- $c_1 + c_2 + c_3 = 0$ , in which case only one eigenvalue vanishes and the correct three physical degrees of freedom propagate.

We are interested in the latter scenario. Setting  $c_1 = 1$  for convenience, the constraint becomes  $c_3 = -(1 + c_2)$ , and the Lagrangian simplifies to

$$\mathcal{L}_4 = f_4(X) \left[ (\partial \cdot A)^2 + c_2 \partial_\rho A_\sigma \partial^\rho A^\sigma - (1 + c_2) \partial_\rho A_\sigma \partial^\sigma A^\rho \right]. \quad (3.36)$$

The structure of this interaction becomes more transparent when rewritten using contractions of Levi-Civita tensors. Unlike the scalar Galileon case, where  $\partial_\mu \partial_\nu \pi$  is symmetric, the derivatives  $\partial_\mu A_\nu$  of the vector field possess both symmetric and antisymmetric components. This allows for two independent contractions with Levi-Civita tensors:

$$\begin{aligned} \mathcal{L}_4 &= -\frac{1}{2} \epsilon^{\mu\nu\rho\sigma} \epsilon_{\rho\sigma}^{\alpha\beta} \left[ f_4(X) \partial_\mu A_\alpha \partial_\nu A_\beta + c_2 \tilde{f}_4(X) \partial_\mu A_\nu \partial_\alpha A_\beta \right] \\ &= f_4(X) \left[ (\partial \cdot A)^2 - \partial_\rho A_\sigma \partial^\sigma A^\rho \right] + c_2 \tilde{f}_4(X) (\partial_\rho A_\sigma \partial^\rho A^\sigma - \partial_\rho A_\sigma \partial^\sigma A^\rho). \end{aligned} \quad (3.37)$$

The terms involving  $c_2$  are proportional to the square of the field strength tensor:

$$F_{\mu\nu} F^{\mu\nu} = 2 (\partial_\rho A_\sigma \partial^\rho A^\sigma - \partial_\rho A_\sigma \partial^\sigma A^\rho). \quad (3.38)$$

Thus, the Lagrangian can be rewritten as [8, 24]:

$$\mathcal{L}_4 = f_4(X) \left[ (\partial \cdot A)^2 - \partial_\rho A_\sigma \partial^\sigma A^\rho \right] + c_2 \tilde{f}_4(X) F_{\rho\sigma} F^{\rho\sigma}. \quad (3.39)$$

Since gauge-invariant  $F^2$  terms can be absorbed into the function  $f_2$  introduced earlier, we discard the last term and retain only the unique non-redundant structure in  $\mathcal{L}_4$ .

Another possible structure involves contracting Levi-Civita tensors with the vector field itself, for example,

$$\epsilon^{\mu\nu\rho\delta} \epsilon^{\alpha\beta\sigma}_\delta \partial_\mu A_\alpha \partial_\nu A_\beta A_\rho A_\sigma, \quad (3.40)$$

which gives rise to interactions of the form

$$\mathcal{L}_4 = \hat{f}_4(X) A^\mu A^\nu \left[ \partial_\nu A_\mu (\partial \cdot A) - \partial_\nu A_\rho \partial^\rho A_\mu \right]. \quad (3.41)$$

These terms are equivalent to those we already constructed, up to disformal transformations of the metric and integrations by parts. Similarly, contractions like

$$\epsilon^{\mu\nu\rho\delta} \epsilon^{\alpha\beta\sigma}{}_\delta \partial_\mu A_\nu \partial_\alpha A_\beta A_\rho A_\sigma \quad (3.42)$$

lead to structures of the type  $F_{\mu\nu} F_{\alpha\beta} A^\mu A^\alpha$ , which are already included in  $\mathcal{L}_2$ . Finally, higher-order contractions such as  $\epsilon \epsilon A^4 \partial A \partial A$  vanish due to the antisymmetric properties of the Levi-Civita tensors. The Levi-Civita formulation once again reveals the underlying simplicity and redundancy in possible contractions, guiding us to a minimal and consistent form for this quartic-order interaction.

As we previously mentioned, ensuring that the Hessian matrix vanishes is essential to eliminate the unwanted temporal degree of freedom in a consistent theory of massive spin-1 fields. This constraint manifests explicitly when calculating the conjugate momentum associated with the Lagrangian  $\mathcal{L}_4$ , given by

$$\Pi^\mu_{\mathcal{L}_4} = \frac{\partial \mathcal{L}_4}{\partial \dot{A}_\mu}. \quad (3.43)$$

In particular, the time component reads

$$\Pi^0_{\mathcal{L}_4} = -2f_4 \vec{\nabla} \cdot \vec{A}, \quad (3.44)$$

which contains no time derivatives and thus imposes a primary constraint:

$$C_1 = \Pi^0_{\mathcal{L}_4} + 2f_4 \vec{\nabla} \cdot \vec{A} \approx 0. \quad (3.45)$$

From this, a secondary constraint emerges by evaluating its Poisson bracket with the Hamiltonian:

$$\{H, C_1\} = \frac{\partial H}{\partial A_\mu} \frac{\partial C_1}{\partial \Pi^\mu} - \frac{\partial H}{\partial \Pi^\mu} \frac{\partial C_1}{\partial A_\mu}. \quad (3.46)$$

These constraints ensure that the temporal component  $A_0$  is nondynamical and the theory propagates only the three physical degrees of freedom.

Extending this method to higher-order interactions, we proceed to cubic order in derivatives. The general form of the quintic interaction is

$$\mathcal{L}_5 = f_5(X) \left[ d_1 (\partial \cdot A)^3 - 3d_2 (\partial \cdot A) \partial_\rho A_\sigma \partial^\rho A^\sigma - 3d_3 (\partial \cdot A) \partial_\rho A_\sigma \partial^\sigma A^\rho \right] \quad (3.47)$$

$$+ 2d_4 \partial_\rho A_\sigma \partial^\gamma A^\rho \partial^\sigma A_\gamma + 2d_5 \partial_\rho A_\sigma \partial^\gamma A^\rho \partial_\gamma A^\sigma \Big], \quad (3.48)$$

where the free parameters  $d_i$  must be chosen to ensure a single vanishing eigenvalue of the Hessian matrix. The required conditions are

$$d_3 = 1 - d_2, \quad d_4 = 1 - \frac{3d_2}{2}, \quad d_5 = \frac{3d_2}{2}, \quad (3.49)$$

with  $d_1 = 1$ . Substituting these into the Lagrangian, we obtain

$$\mathcal{L}_5 = f_5(X) \left[ (\partial \cdot A)^3 - 3d_2 (\partial \cdot A) \partial_\rho A_\sigma \partial^\rho A^\sigma - 3(1 - d_2) (\partial \cdot A) \partial_\rho A_\sigma \partial^\sigma A^\rho \right] \quad (3.50)$$

$$+ 2 \left( 1 - \frac{3d_2}{2} \right) \partial_\rho A_\sigma \partial^\gamma A^\rho \partial^\sigma A_\gamma + 2 \left( \frac{3d_2}{2} \right) \partial_\rho A_\sigma \partial^\gamma A^\rho \partial_\gamma A^\sigma \Big]. \quad (3.51)$$

Just like in the quartic case, some of the  $d_i$  terms can be associated with distinct functions. For example, the term proportional to  $d_2$  may appear with an independent function  $\tilde{f}_5(X)$ :

$$\mathcal{L}_5 = f_5(X) \left[ (\partial \cdot A)^3 - 3(\partial \cdot A) \partial_\rho A_\sigma \partial^\sigma A^\rho + 2\partial_\rho A_\sigma \partial^\gamma A^\rho \partial^\sigma A_\gamma \right] \quad (3.52)$$

$$+ d_2 \tilde{f}_5(X) \left[ \frac{1}{2} (\partial \cdot A) F_{\rho\sigma}^2 - \partial_\sigma A_\gamma F_\rho{}^\sigma F^{\rho\gamma} \right]. \quad (3.53)$$

This decomposition also arises from the Levi-Civita tensor contractions:

$$\mathcal{L}_5 = -\epsilon^{\mu\nu\rho\sigma} \epsilon^{\alpha\beta\delta} \left[ f_5(X) \partial_\mu A_\alpha \partial_\nu A_\beta \partial_\rho A_\delta + d_2 \tilde{f}_5(X) \partial_\mu A_\nu \partial_\rho A_\alpha \partial_\beta A_\delta \right]. \quad (3.54)$$

In fact, the term with  $\tilde{f}_5$  can be compactly rewritten using the dual field strength tensor as

$$\mathcal{L}_5 = f_5(X) \left[ (\partial \cdot A)^3 - 3(\partial \cdot A) \partial_\rho A_\sigma \partial^\sigma A^\rho + 2\partial_\rho A_\sigma \partial^\gamma A^\rho \partial^\sigma A_\gamma \right] + \tilde{f}_5(X) \tilde{F}^{\alpha\mu} \tilde{F}_\mu{}^\beta \partial_\alpha A_\beta. \quad (3.55)$$

If one demanded that the longitudinal mode possess only non-trivial interactions, the series would terminate at  $\mathcal{L}_5$ . However, relaxing this assumption allows us to continue to quartic derivative order. Once again, Levi-Civita contractions guide the consistent construction:

$$\mathcal{L}_6 = -\epsilon^{\mu\nu\rho\sigma} \epsilon^{\alpha\beta\delta\kappa} \left[ f_6(X) \partial_\mu A_\alpha \partial_\nu A_\beta \partial_\rho A_\delta \partial_\sigma A_\kappa + e_2 \tilde{f}_6(X) \partial_\mu A_\nu \partial_\alpha A_\beta \partial_\rho A_\delta \partial_\sigma A_\kappa \right]. \quad (3.56)$$

Only the second term contributes non-trivially to the equations of motion; the first is a total derivative. Simplifying, we get

$$\mathcal{L}_6 = e_2 f_6(X) \tilde{F}^{\alpha\beta} \tilde{F}^{\mu\nu} \partial_\alpha A_\mu \partial_\beta A_\nu. \quad (3.57)$$

Summarizing, the complete generalized Proca Lagrangian on flat spacetime is [8, 26]:

$$\mathcal{L}_{\text{gen.Proca}} = -\frac{1}{4}F_{\mu\nu}^2 + \sum_{n=2}^6 \alpha_n \mathcal{L}_n, \quad (3.58)$$

with the interactions

$$\mathcal{L}_2 = f_2(X, F, Y), \quad (3.59)$$

$$\mathcal{L}_3 = f_3(X) \partial_\mu A^\mu, \quad (3.60)$$

$$\mathcal{L}_4 = f_4(X) [(\partial \cdot A)^2 - \partial_\rho A_\sigma \partial^\sigma A^\rho], \quad (3.61)$$

$$\mathcal{L}_5 = f_5(X) [(\partial \cdot A)^3 - 3(\partial \cdot A) \partial_\rho A_\sigma \partial^\sigma A^\rho + 2\partial_\rho A_\sigma \partial^\gamma A^\rho \partial^\sigma A_\gamma] \quad (3.62)$$

$$+ \tilde{f}_5(X) \tilde{F}^{\alpha\mu} \tilde{F}_\mu^\beta \partial_\alpha A_\beta, \quad (3.63)$$

$$\mathcal{L}_6 = f_6(X) \tilde{F}^{\alpha\beta} \tilde{F}^{\mu\nu} \partial_\alpha A_\mu \partial_\beta A_\nu. \quad (3.64)$$

These represent the most general ghost-free derivative self-interactions for a massive spin-1 field in four-dimensional Minkowski spacetime, terminating naturally at sixth order due to the contraction limits of Levi-Civita tensors.

### 3.3 Generalized Proca Theories in Curved Spacetime

A natural follow-up question is how to consistently couple a massive vector field to gravity, building on the insights obtained from the flat spacetime construction [8, 26]. While the coupling of massless spin-1 fields to gravity proceeds naturally via gauge invariance, the inclusion of derivative self-interactions for massive vector fields—analogueous to Galileon terms—requires greater care. In particular, it is essential to ensure that the equations of motion remain second order in derivatives to avoid Ostrogradsky instabilities.

A naive covariantization of the flat-space generalized Proca interactions would generally introduce higher-order derivatives due to the non-commutativity of covariant derivatives. However, these problematic terms can be canceled by introducing specific non-minimal couplings to curvature tensors. Such couplings involve divergenceless tensors like the Ricci and Einstein tensors, multiplied by functions of the scalar quantity  $X = -\frac{1}{2}A_\mu A^\mu$ .

For the pure Stueckelberg sector—equivalent to a scalar Galileon—the necessary non-minimal terms are structurally identical to those in scalar Horndeski theories [16]. These counter-terms can therefore be imported directly into the vector field case.

When the full vector field is involved, including both longitudinal and transverse components, additional care is required to identify the allowed non-minimal couplings. The

generalized Proca Lagrangian on a curved spacetime takes the form

$$\mathcal{L}_{\text{gen.Proca}}^{\text{curved}} = -\frac{1}{4}\sqrt{-g}F_{\mu\nu}F^{\mu\nu} + \sqrt{-g}\sum_{n=2}^6\mathcal{L}_n, \quad (3.65)$$

where the individual interaction terms are given by

$$\mathcal{L}_2 = G_2(A_\mu, F_{\mu\nu}), \quad (3.66)$$

$$\mathcal{L}_3 = G_3(X)\nabla_\mu A^\mu, \quad (3.67)$$

$$\mathcal{L}_4 = G_4(X)R + G_{4,X}\left[(\nabla_\mu A^\mu)^2 - \nabla_\rho A_\sigma \nabla^\sigma A^\rho\right], \quad (3.68)$$

$$\begin{aligned} \mathcal{L}_5 = G_5(X)G_{\mu\nu}\nabla^\mu A^\nu - \frac{1}{6}G_{5,X}\left[(\nabla \cdot A)^3 + 2\nabla_\rho A_\sigma \nabla^\sigma A^\rho \nabla^\sigma A_\gamma \right. \\ \left. - 3(\nabla \cdot A)\nabla_\rho A_\sigma \nabla^\sigma A^\rho\right] - g_5(X)\tilde{F}^{\alpha\mu}\tilde{F}_\mu^\beta \nabla_\alpha A_\beta, \end{aligned} \quad (3.69)$$

$$\mathcal{L}_6 = G_6(X)\mathcal{L}^{\mu\nu\alpha\beta}\nabla_\mu A_\nu \nabla_\alpha A_\beta + \frac{G_{6,X}}{2}\tilde{F}^{\alpha\beta}\tilde{F}^{\mu\nu}\nabla_\alpha A_\mu \nabla_\beta A_\nu. \quad (3.70)$$

Here,  $\nabla_\mu$  is the covariant derivative, and  $\mathcal{L}^{\mu\nu\alpha\beta}$  denotes the double dual of the Riemann tensor:

$$\mathcal{L}^{\mu\nu\alpha\beta} = \frac{1}{4}\epsilon^{\mu\nu\rho\sigma}\epsilon^{\alpha\beta\gamma\delta}R_{\rho\sigma\gamma\delta}. \quad (3.71)$$

The function  $G_2$  depends on the scalar quantities  $X = -\frac{1}{2}A_\mu A^\mu$ ,  $F = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$ , and  $Y = A^\mu A^\nu F_\mu^\alpha F_{\nu\alpha}$ . Parity-violating contributions like  $F_{\mu\nu}\tilde{F}^{\mu\nu}$  are typically omitted.

These terms constitute the most general derivative self-interactions for a massive vector field on a curved background that yield second-order field equations, up to disformal transformations. The non-minimal couplings—such as  $G_4(X)R$ ,  $G_5(X)G_{\mu\nu}\nabla^\mu A^\nu$ , and  $G_6(X)\mathcal{L}^{\mu\nu\alpha\beta}\nabla_\mu A_\nu \nabla_\alpha A_\beta$ —are required to cancel higher-derivative contributions arising from the covariantization process. By contrast, terms like  $G_3(X)\nabla_\mu A^\mu$  and  $\tilde{G}_5(Y)\tilde{F}^{\mu\alpha}\tilde{F}_\alpha^\nu \nabla_\mu A_\nu$  are already safe under naive covariantization and do not require additional curvature terms.

To clarify the structure of these interactions, it is useful to introduce the symmetric tensor

$$S_{\mu\nu} = \nabla_\mu A_\nu + \nabla_\nu A_\mu. \quad (3.72)$$

In terms of  $S_{\mu\nu}$ , the Lagrangians can be recast as

$$\mathcal{L}_2 = \hat{G}_2(X, F, Y), \quad (3.73)$$

$$\mathcal{L}_3 = \frac{1}{2}G_3(X)[S], \quad (3.74)$$

$$\mathcal{L}_4 = G_4(X)R + G_{4,X}\frac{[S]^2 - [S^2]}{4}, \quad (3.75)$$

$$\mathcal{L}_5 = \frac{G_5(X)}{2} G^{\mu\nu} S_{\mu\nu} - \frac{G_{5,X}}{6} \frac{[S]^3 - 3[S][S^2] + 2[S^3]}{8} + g_5(X) \tilde{F}^{\mu\alpha} \tilde{F}_\alpha{}^\nu S_{\mu\nu}, \quad (3.76)$$

$$\mathcal{L}_6 = G_6(X) \mathcal{L}^{\mu\nu\alpha\beta} F_{\mu\nu} F_{\alpha\beta} + \frac{G_{6,X}}{2} \tilde{F}^{\alpha\beta} \tilde{F}^{\mu\nu} S_{\alpha\mu} S_{\beta\nu}. \quad (3.77)$$

The scalar Horndeski theory is recovered in the limit  $S_{\mu\nu} \rightarrow 2\nabla_\mu \nabla_\nu \pi$ , where  $\pi$  is the Stueckelberg scalar. The interactions that involve  $\tilde{F}^{\mu\nu}$  are intrinsically vectorial and have no scalar analog.

Finally, to ensure that the equations of motion remain second order, it is essential that derivatives acting on  $S_{\mu\nu}$  obey the identity

$$2\nabla_{[\alpha} S_{\beta]\gamma} = [\nabla_\alpha, \nabla_\beta] A_\gamma + [\nabla_\alpha, \nabla_\gamma] A_\beta - [\nabla_\beta, \nabla_\gamma] A_\alpha + \nabla_\gamma F_{\alpha\beta}. \quad (3.78)$$

This ensures that commutators of covariant derivatives only contribute curvature-dependent terms and do not introduce higher derivatives of the field.

In this chapter, we constructed the most general self-derivative interactions for a massive vector field that propagate only three degrees of freedom and yield second-order equations of motion, thereby avoiding the Ostrogradsky instability. These generalized Proca theories, both in flat and curved backgrounds, represent consistent extensions of the Proca framework that incorporate rich phenomenology while maintaining theoretical consistency. Having established their classical structure, the next natural step is to investigate their behavior under quantum corrections. In the following chapter, we examine the quantum stability of these self-derivative interactions and analyze the potential emergence of new counter-terms at the one-loop level.



# Chapter 4

## Quantum Stability in Flat Spacetime

### 4.1 Definition of Stability

In the previous chapter, we presented scalar Galileon and generalized Proca theories as effective field theories (EFTs) that include higher-derivative interactions while preserving second-order equations of motion. These constructions rely on a delicate balance between terms, carefully tuned to avoid the propagation of Ostrogradsky ghosts. With their classical consistency established, we now investigate their behavior under quantum corrections.

The central question of this chapter is whether these theories maintain their structural consistency at the quantum level. That is, do loop corrections spoil the required relations between terms, or does the theory remain radiatively stable?

We define *quantum stability* as the property that radiative corrections do not generate new operators that violate the symmetries or second-order nature of the classical action, and that the coefficients of existing operators are not dangerously renormalized. Since Galileon and generalized Proca theories are constructed with highly constrained interactions, small quantum corrections could, in principle, destabilize their delicate structure.

In the case of scalar Galileons, the interactions enjoy a symmetry known as Galileon symmetry, given by

$$\pi \rightarrow \pi + c + b_\mu x^\mu, \tag{4.1}$$

where  $c$  and  $b_\mu$  are constant parameters. This internal symmetry, preserved up to total derivatives, plays a crucial role in protecting the theory from dangerous quantum effects. While it forbids the generation of symmetry-breaking local operators, it is less obvious whether the symmetry-preserving Galileon operators themselves receive quantum correc-

tions.

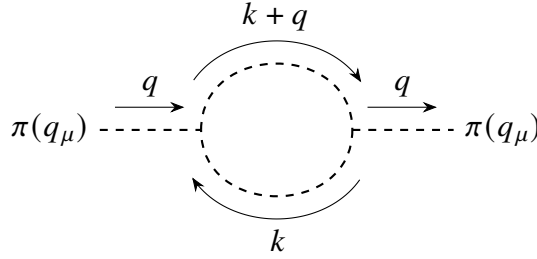
Remarkably, scalar Galileons [27] are protected by a non-renormalization theorem. Loop diagrams do not generate new Galileon-type operators, nor do they renormalize the coefficients of the existing ones. This implies that the classical structure is technically natural: no fine-tuning is required to maintain the theory's consistency under quantum corrections.

To make this concrete, consider the cubic Galileon interaction written in Levi-Civita tensor form:

$$\mathcal{L}_3 = \pi \epsilon^{\mu\alpha\rho\sigma} \epsilon^{\nu\beta}_{\rho\sigma} \Pi_{\mu\nu} \Pi_{\alpha\beta}, \quad (4.2)$$

where  $\Pi_{\mu\nu} = \partial_\mu \partial_\nu \pi$  and  $\epsilon^{\mu\alpha\rho\sigma}$  is the antisymmetric Levi-Civita symbol.

We examine a one-loop correction in which an external  $\pi$  leg carrying momentum  $q_\mu$  is attached to a vertex, while the other two  $\pi$  fields run inside the loop with internal momenta  $k_\mu$  and  $(q + k)_\mu$ . The corresponding diagram takes the form:



The loop integral corresponding to the cubic Galileon vertex takes the form

$$\mathcal{A} \propto \int \frac{d^4 k}{(2\pi)^4} D_k D_{k+q} \epsilon^{\mu\alpha\rho\sigma} \epsilon^{\nu\beta}_{\rho\sigma} k_\mu k_\nu (q+k)_\alpha (q+k)_\beta \cdots, \quad (4.3)$$

where  $D_k = 1/k^2$  is the massless scalar propagator. Owing to the antisymmetry properties of the Levi-Civita tensors, any term in the integrand that is independent of the external momentum  $q_\mu$ , or linear in  $q_\mu$ , vanishes upon integration [8]. This means the leading non-zero contributions involve at least two powers of  $q_\mu$ , which produce higher-derivative operators.

Such operators lie outside the Galileon class and are suppressed within the low-energy effective theory. Consequently, loop corrections do not renormalize the Galileon operators. Instead, the quantum-generated terms preserve Galileon symmetry exactly, whereas the classical terms do so only up to total derivatives. This structure protects the original Galileon interactions and ensures their technical naturalness.

Nonetheless, this radiative stability can be compromised when the scalar field is explicitly coupled to matter. Typical couplings include conformal interactions like  $\pi T$ , disformal terms such as  $\partial_\mu \pi \partial_\nu \pi T^{\mu\nu}$ , or longitudinal couplings of the form  $\partial_\mu \partial_\nu \pi T^{\mu\nu}$ . These types of interactions naturally appear in modified gravity models, including extensions of massive gravity.

Once such couplings are introduced, one-loop quantum corrections may generate new operators that both renormalize Galileon interactions and introduce higher-derivative ghostly terms. Fortunately, these contributions are suppressed by the scale of the matter coupling and remain under control within the validity regime of the effective theory.

In summary, the quantum stability of derivative EFTs such as Galileons hinges on internal symmetries that tightly constrain quantum corrections. As long as these symmetries are preserved and the theory remains decoupled from external matter fields, the Galileon Lagrangian remains radiatively stable and predictive.

Having clarified the quantum robustness of Galileon scalar theories, we now turn to generalized Proca models. These vector field theories can be separated into two classes: (i) interactions that are direct analogues of scalar Galileons, obtained by promoting  $\partial_\mu \pi$  to  $A_\mu$ , and (ii) interactions intrinsic to the vector field, which have no scalar counterpart.

These two classes exhibit markedly different behaviors under quantum corrections. The Galileon-like vector interactions inherit symmetry protections from their scalar counterparts, and are thus expected to obey a similar non-renormalization theorem. In contrast, the intrinsically vectorial terms—those with no analogue in scalar Galileon theories—lack this protection and may receive non-trivial loop corrections.

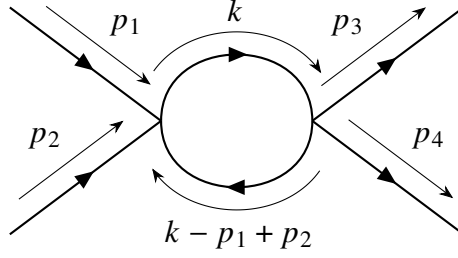
To analyze this more precisely, we consider Feynman diagrams involving only the vector field. The momentum-space propagator for a massive vector field is given by

$$D_{\mu\nu}(k) = \frac{-i}{k^2 - m^2} \left( \eta_{\mu\nu} - \frac{k_\mu k_\nu}{m^2} \right). \quad (4.4)$$

As a representative example, consider a Galileon-type quartic interaction of the schematic form

$$A^\mu A^\nu \epsilon_\mu^{\alpha\rho\gamma} \epsilon_\nu^{\beta\sigma} \gamma \partial_\alpha A_\beta \partial_\rho A_\sigma. \quad (4.5)$$

Focusing on the most relevant diagrams—those involving the smallest number of external momenta—we assign momenta  $p_1$  and  $p_2$  to the external  $A^\mu$  legs, and the loop momenta are  $k$  and  $p_1 + p_2 - k$ .



The corresponding loop integral takes the form

$$\mathcal{A} \sim \int \frac{d^4 k}{(2\pi)^4} D^{\kappa\mu}(p_1) D^{\delta\nu}(p_2) \epsilon_\mu^{\alpha\rho\gamma} \epsilon_\nu^{\beta\sigma} \gamma k_\alpha (p_1 + p_2 - k)_\rho D_{\lambda\beta}(k) D_{\xi\sigma}(p_1 + p_2 - k). \quad (4.6)$$

Due to the antisymmetric structure of the Levi-Civita contractions, any terms that do not depend on external momenta—or that are linear in them—cancel out [8]. Thus, the loop corrections produce only higher-derivative operators that are irrelevant in the low-energy EFT. The original Galileon-type operators remain unrenormalized, confirming their quantum stability.

The loop integral for such a diagram is schematically given by

$$\mathcal{A} \sim \int \frac{d^4 k}{(2\pi)^4} D^{\kappa\mu}(p_1) D^{\delta\nu}(p_2) \epsilon_\mu^{\alpha\rho\gamma} \epsilon_\nu^{\beta\sigma} \gamma k_\alpha (p_1 + p_2 - k)_\rho D_{\lambda\beta}(k) D_{\xi\sigma}(p_1 + p_2 - k). \quad (4.7)$$

Due to the antisymmetric structure of the Levi-Civita tensors, terms that are independent of, or linear in, the external momenta vanish after integration. The leading non-zero contributions must include at least two powers of external momenta, which implies that the generated operators involve more derivatives than the original interaction. These higher-derivative operators are suppressed in the effective field theory, meaning that the original Galileon-type terms are not renormalized. This mirrors the scalar Galileon case and confirms the quantum stability of these vector interactions.

In contrast, vector interactions that have no scalar analogue behave differently. Consider, for example, the following quartic term:

$$A^\mu A^\nu \epsilon_\mu^{\alpha\rho\gamma} \epsilon_\nu^{\beta\sigma} \gamma \partial_\alpha A_\rho \partial_\beta A_\sigma. \quad (4.8)$$

Following the same procedure as above and assigning external momenta to the undifferentiated fields, the corresponding loop amplitude becomes

$$\mathcal{A} \sim \int \frac{d^4 k}{(2\pi)^4} D^{\kappa\mu}(p_1) D^{\delta\nu}(p_2) \epsilon_\mu^{\alpha\rho\gamma} \epsilon_\nu^{\beta\sigma} \gamma k_\alpha (p_1 + p_2 - k)_\beta D_{\lambda\rho}(k) D_{\xi\sigma}(p_1 + p_2 - k). \quad (4.9)$$

Unlike the Galileon-type case, this interaction admits contributions proportional to products of loop momenta such as  $k_\alpha k_\beta$ , even in the absence of external momentum. These contributions generate operators with the same structure as the original interaction, thereby renormalizing it. As a result, genuinely vectorial terms do not enjoy the same quantum protection and are subject to radiative corrections.

It is therefore crucial to quantify the magnitude of these corrections and assess their impact within the validity regime of the effective field theory. This motivates a more detailed analysis of loop effects and their scaling behavior, which we will explore in the following chapters.

## 4.2 Power Counting

Before performing explicit loop computations, it is useful to develop a general framework for estimating the structure of ultraviolet (UV) divergences in a quantum field theory. Such estimates can help us anticipate the form of counter-terms required for renormalization, without having to compute every diagram in detail. This method, known as power counting, allows us to determine whether a given Feynman diagram is divergent, and if so, what kind of divergence we should expect.

The basic idea is that, for diagrams with a non-negative superficial degree of divergence  $D \geq 0$ , the part of the momentum integral where all internal momenta become large simultaneously will contribute a divergence. As a heuristic example, such behavior resembles an integral of the form  $\int_0^\infty k^{D-1} dk$ , which diverges for  $D \geq 0$ .

To gain some intuition, suppose we differentiate the diagram's amplitude with respect to one of the external momenta. Each such derivative effectively reduces the power of loop momentum in the integrand. After taking  $D+1$  derivatives, the divergent part of the integral becomes finite, indicating that the divergent contribution must be a polynomial of degree at most  $D$  in the external momenta. The coefficients of this polynomial are divergent, while any non-polynomial remainder remains finite.

To illustrate this idea concretely, consider the one-dimensional example

$$\mathcal{J}(q) = \int_0^\infty \frac{dk}{k+q}, \quad (4.10)$$

which has a logarithmic divergence corresponding to  $D = 0$ . Differentiating once gives

$$\mathcal{J}'(q) = - \int_0^\infty \frac{dk}{(k+q)^2} = -\frac{1}{q}, \quad (4.11)$$

so the original integral can be written as

$$\mathcal{J}(q) = -\ln q + c, \quad (4.12)$$

where  $c$  is a divergent constant. The divergent part is independent of the external momentum  $q$ , while the finite part is analytic in  $q$ .

A similar structure appears in integrals with higher degrees of divergence. For instance, a linearly divergent integral like

$$\int_0^\infty \frac{k dk}{k + q} \quad (4.13)$$

produces an expression of the form

$$a + bq + q \ln q, \quad (4.14)$$

where  $a$  and  $b$  are divergent constants, and the remaining  $q \ln q$  term is finite.

From a Lagrangian perspective, this means that any divergence arising from a diagram with  $E_f$  external legs of field type  $f$  and superficial degree  $D$  can be absorbed into counter-terms constructed from local operators involving  $E_f$  fields and at most  $D$  derivatives. These are precisely the types of terms already present or allowed in the original Lagrangian if the theory is to remain renormalizable or effectively renormalizable. Consequently, UV divergences manifest themselves as corrections to existing coupling constants or the introduction of new, higher-dimensional operators in the effective action.

In the following sections, we will make these arguments more precise by analyzing specific examples involving two- and three-point functions, including tensor structure and index contractions. This will allow us to move beyond the heuristic and establish a systematic approach to identifying and controlling divergences in effective field theories.

### 4.2.1 Diagrammatic Power Counting

To move beyond the heuristic argument and make the discussion more precise, we now consider the diagrammatic structure of loop integrals in theories with derivative self-interactions. Let us denote the amputated one-loop amplitude as

$$I^{\mu\nu\alpha\cdots}(p_1, p_2, p_3, \dots), \quad (4.15)$$

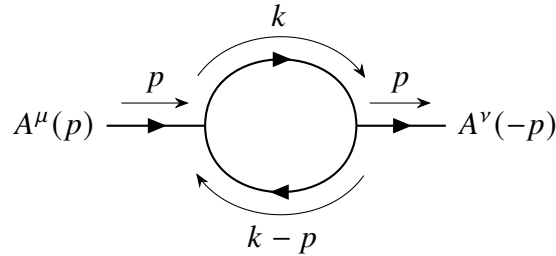
where the indices correspond to Lorentz structures from external fields or vertices, and  $\{p_i\}$  represent the external momenta. In general, due to the complexity of the interaction terms, each Feynman diagram will include multiple contributions: propagators may carry

different tensor structures, and interaction vertices—especially in self-interacting theories like generalized Proca—typically include several terms with different momentum dependencies.

To isolate the ultraviolet-divergent part of the diagram, we restrict our attention to the most divergent pieces of the integrand. These correspond to the terms with the highest number of loop momentum factors in the numerator. Lower-order terms contribute either finite pieces or divergences of lower degree, which are subleading in the UV limit and do not affect the leading structure of the divergence.

The superficial degree of divergence  $D$  of a diagram dictates the maximum power of external momenta appearing in the divergent part of the amplitude. Once  $D$  is identified, we can follow a systematic procedure: we take  $D + 1$  derivatives of the amplitude with respect to external momenta to reduce the divergence to a finite integral. Then, we reintegrate, as in the simple scalar example discussed previously, to reconstruct the divergent part as a polynomial of degree  $D$  in the external momenta.

As a concrete example, let us consider the one-loop two-point function generated by the interaction vertex from  $\mathcal{L}_3$ :



The full tensor structure of this diagram will be specified later. For the purpose of power counting, however, it suffices to observe that the most divergent contribution arises from the part of the propagator containing additional powers of loop momentum. Specifically, the propagator takes the form

$$D_{\mu\nu}(k) = \frac{1}{k^2 - m^2 + i\epsilon} \left( -\eta_{\mu\nu} + \frac{k_\mu k_\nu}{m^2} \right), \quad (4.16)$$

and the second term in the numerator, proportional to  $k_\mu k_\nu$ , leads to enhanced UV divergence due to the presence of extra powers of the loop momentum  $k$  in the numerator.

When the propagator is combined with the momentum-dependent structure of the  $\mathcal{L}_3$  vertex, the resulting loop integrand contains several terms with different momentum dependence. As we will show explicitly in a later section, the term with the highest superficial

degree of divergence behaves as  $D = 6$ . This dominant contribution arises from the interplay between the derivative interactions in the vertex and the  $k_\mu k_\nu$  term in the propagator numerator, which introduces additional powers of loop momentum.

Taking a derivative with respect to an external momentum lowers the overall degree of divergence by one. For example, the first derivative of the amplitude reduces the divergence from  $D = 6$  to  $D = 5$ :

$$\frac{\partial I^{\mu\nu}(p)}{\partial p^\alpha} \sim \text{divergence of degree } D = 5. \quad (4.17)$$

Continuing this process, taking a total of  $D + 1 = 7$  derivatives renders the expression finite. This ensures that the original divergent part of the amplitude is a local polynomial of degree six in the external momentum. After taking 7 such derivatives, the resulting expression becomes UV finite:

$$\frac{\partial^7 I^{\mu\nu}(p)}{\partial p^{\alpha_1} \dots \partial p^{\alpha_7}} = f_{\alpha_1 \dots \alpha_7}^{\mu\nu}(p), \quad (4.18)$$

where  $f_{\alpha_1 \dots \alpha_7}^{\mu\nu}(p)$  is a finite, momentum-dependent tensor function. Since this expression is finite, it can be discarded in the context of isolating the divergent part. Reintegrating this equation seven times with respect to the external momentum reconstructs the divergent polynomial part of the original amplitude:

$$I^{\mu\nu}(p) \Big|_{\text{div}} = C_{(0)}^{\mu\nu} + C_{(1)\alpha_1}^{\mu\nu} p^{\alpha_1} + C_{(2)\alpha_1\alpha_2}^{\mu\nu} p^{\alpha_1} p^{\alpha_2} + \dots + C_{(6)\alpha_1 \dots \alpha_6}^{\mu\nu} p^{\alpha_1} \dots p^{\alpha_6}, \quad (4.19)$$

where the coefficients  $C_{(n)}^{\mu\nu}$  are divergent tensors that depend on the details of the loop integration, but are independent of external momenta. This expansion confirms that the UV divergence takes the form of a local polynomial in the external momenta, consistent with the structure expected from the effective action.

### 4.2.2 Dimensional Analysis and Rescaling

It is often advantageous to express divergent amplitudes in terms of dimensionless coefficients. This makes the scaling behavior of divergences more transparent, especially in effective field theories where all quantities are organized in a derivative expansion normalized by a characteristic energy scale.

To proceed, we recall a general result from dimensional analysis: in  $d = 4$  spacetime dimensions, the mass dimension of an  $n$ -point Green's function is given by

$$[I_n] = 4 - \sum_i [\phi_i], \quad (4.20)$$



where  $[\phi_i]$  is the mass dimension of the  $i$ -th external field. In theories involving vector fields with canonical dimension  $[A_\mu] = 1$ , this simplifies for an  $n$ -point function involving only vector fields to

$$[I_n] = 4 - n. \quad (4.21)$$

For instance, a two-point function has dimension 2, a three-point function has dimension 1, and so on.

Therefore, the divergent part of a two-point function such as  $I^{\mu\nu}(p)$  must have overall mass dimension 2. Since the momentum  $p^\alpha$  carries dimension 1, each term in the polynomial expansion of  $I^{\mu\nu}(p)$  must be accompanied by a coefficient with dimension such that the total dimension remains 2. To express these coefficients in a dimensionless form, we rescale them by appropriate powers of a mass scale  $m$  and write:

$$\begin{aligned} I^{\mu\nu}(p) \Big|_{\text{div}} = & m^2 \tilde{C}_{(0)}^{\mu\nu} + m \tilde{C}_{(1)\alpha_1}^{\mu\nu} p^{\alpha_1} + \tilde{C}_{(2)\alpha_1\alpha_2}^{\mu\nu} p^{\alpha_1} p^{\alpha_2} + \frac{1}{m} \tilde{C}_{(3)\alpha_1\alpha_2\alpha_3}^{\mu\nu} p^{\alpha_1} p^{\alpha_2} p^{\alpha_3} \\ & + \frac{1}{m^2} \tilde{C}_{(4)\alpha_1\cdots\alpha_4}^{\mu\nu} p^{\alpha_1} \cdots p^{\alpha_4} + \frac{1}{m^3} \tilde{C}_{(5)\alpha_1\cdots\alpha_5}^{\mu\nu} p^{\alpha_1} \cdots p^{\alpha_5} + \frac{1}{m^4} \tilde{C}_{(6)\alpha_1\cdots\alpha_6}^{\mu\nu} p^{\alpha_1} \cdots p^{\alpha_6}. \end{aligned} \quad (4.22)$$

Each  $\tilde{C}_{(n)}^{\mu\nu}$  is a dimensionless tensor, constructed from contractions of the metric and Levi-Civita symbols with loop integrals, and encodes the tensorial structure of the UV divergent part. This form makes manifest how the UV behavior is governed by local operators with increasing numbers of derivatives, suppressed by increasing powers of the cutoff scale  $m$ , in line with the principles of effective field theory.

### 4.2.3 Lorentz Invariance

A crucial constraint on the structure of the divergent amplitude comes from Lorentz invariance. Since the amplitude  $I^{\mu\nu}(p)$  must transform covariantly under Lorentz transformations, the coefficients  $\tilde{C}_{(n)\alpha_1\cdots\alpha_n}^{\mu\nu}$  must be built entirely from Lorentz-invariant tensors. In flat spacetime, the only available building blocks for constructing such coefficients are the Minkowski metric  $\eta^{\mu\nu}$  and the totally antisymmetric Levi-Civita tensor  $\varepsilon^{\mu\nu\rho\sigma}$ <sup>1</sup>.

Let us now examine the tensorial structure of the sixth-order term. The coefficient  $\tilde{C}_{(6)\alpha_1\cdots\alpha_6}^{\mu\nu}$  must be constructed using six symmetric Lorentz indices. Possible contractions

<sup>1</sup>The Levi-Civita tensor can, in principle, appear in parity-violating theories. However, in many cases including parity-preserving interactions, such terms are excluded. Even when parity is not assumed, index symmetries may prevent the construction of appropriate tensor structures using  $\varepsilon^{\mu\nu\rho\sigma}$ .

using the metric include:

$$\eta^{\mu\nu}\eta^{\alpha_1\alpha_2}\eta^{\alpha_3\alpha_4}\eta^{\alpha_5\alpha_6}, \quad \eta^\mu_{\alpha_1}\eta^\nu_{\alpha_2}\eta^{\alpha_3\alpha_4}\eta^{\alpha_5\alpha_6}, \quad \text{etc.} \quad (4.23)$$

These are consistent with Lorentz invariance and yield scalar expressions upon contraction with polarization vectors and external momenta. When the full amplitude is contracted with polarization vectors  $\epsilon_\mu, \epsilon_\nu$ , the sixth-order term contributes combinations such as

$$(\epsilon \cdot \epsilon) \frac{p^6}{m^4}, \quad (\epsilon \cdot p)^2 \frac{p^4}{m^4}, \quad (4.24)$$

which are fully Lorentz invariant and of the correct dimension.

Now consider the fifth-order term. The coefficient  $\tilde{C}_{(5)\alpha_1\cdots\alpha_5}^{\mu\nu}$  carries a total of seven Lorentz indices. In constructing such a tensor purely from the metric, one quickly sees that no combination of  $\eta^{\mu\nu}$  can yield an object with \*seven\* free indices, due to the symmetric nature of  $\eta^{\mu\nu}$  and the fact that each metric contracts two indices. Therefore, it is not possible to construct a rank-7 tensor with the required index symmetries using only  $\eta^{\mu\nu}$ .

One might ask whether the antisymmetric Levi-Civita tensor could rescue the situation. However,  $\epsilon^{\mu\nu\rho\sigma}$  has only four indices, and even using it in combination with the metric, one cannot construct a fully symmetric rank-7 tensor compatible with the Lorentz structure and the index symmetries required by the amplitude and polarization contractions. Moreover, unless the theory contains parity-violating terms, contributions involving  $\epsilon^{\mu\nu\rho\sigma}$  are absent altogether.

Therefore, we conclude that odd powers of momenta — such as the fifth-order term — cannot appear in the Lorentz-invariant UV-divergent part of the amplitude. The absence of such terms is a direct consequence of the impossibility of constructing the necessary tensor structures with the available Lorentz-invariant objects. This reasoning applies equally to all odd powers, leading to the simplified structure:

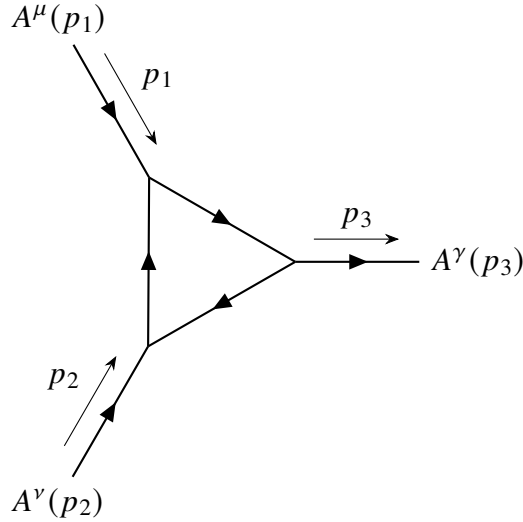
$$I^{\mu\nu}(p)\Big|_{\text{div}} = m^2 \tilde{C}_{(0)}^{\mu\nu} + \tilde{C}_{(2)\alpha_1\alpha_2}^{\mu\nu} p^{\alpha_1} p^{\alpha_2} + \frac{1}{m^2} \tilde{C}_{(4)\alpha_1\cdots\alpha_4}^{\mu\nu} p^{\alpha_1} \cdots p^{\alpha_4} + \frac{1}{m^4} \tilde{C}_{(6)\alpha_1\cdots\alpha_6}^{\mu\nu} p^{\alpha_1} \cdots p^{\alpha_6}. \quad (4.25)$$

Thus, Lorentz invariance not only constrains the form of divergent terms but also eliminates entire classes of them, significantly simplifying the structure of UV divergences in effective theories.

While Lorentz invariance often rules out certain terms in the momentum expansion—particularly those involving odd powers—these symmetry constraints depend crucially on the total number of free Lorentz indices in the diagram. This number is determined by the

structure of the amplitude and the number of external legs. As a result, odd powers of momenta are not universally forbidden, and in some cases even powers may be excluded due to the absence of compatible Lorentz structures.

To illustrate this point, consider the one-loop three-point function arising from a triangle diagram:



We denote the amputated amplitude by

$$I^{\mu\nu\gamma}(p_1, p_2, p_3), \quad (4.26)$$

where  $p_1$ ,  $p_2$ , and  $p_3$  are the external momenta, and  $\mu$ ,  $\nu$ , and  $\gamma$  are the indices associated with the external vector fields.

Suppose the expansion includes a term of seventh order in momenta,

$$\frac{1}{m^6} \tilde{C}_{(7)\alpha_1 \dots \alpha_7}^{\mu\nu\gamma} p^{\alpha_1} \dots p^{\alpha_7}, \quad (4.27)$$

which contains a total of ten Lorentz indices. Since this is an even number, it is possible to construct such a tensor using contractions of the Minkowski metric  $\eta^{\mu\nu}$ . For instance, after contraction with polarization vectors  $\epsilon^\mu$ ,  $\epsilon^\nu$ ,  $\epsilon^\gamma$  and external momenta  $p_i$ , one can obtain Lorentz-invariant structures such as

$$(\epsilon \cdot \epsilon)(\epsilon \cdot p) \frac{p^6}{m^6}, \quad (\epsilon \cdot p)^3 \frac{p^4}{m^6}, \quad (4.28)$$

where each  $\epsilon$  and  $p$  can correspond to any of the external legs. These expressions are consistent with the required symmetries. Therefore, in this case, a seventh-order term in momenta is allowed.

On the other hand, consider a sixth-order term of the form

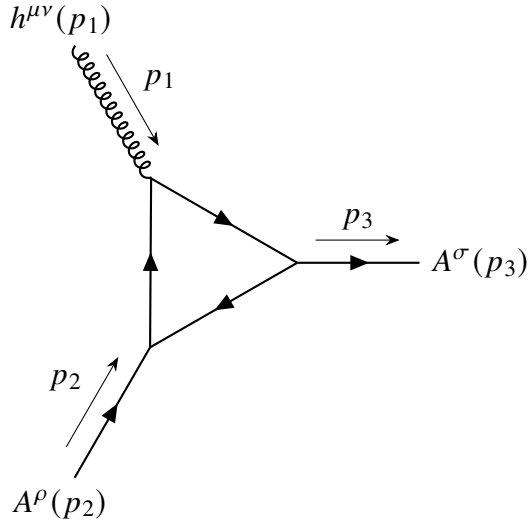
$$\frac{1}{m^4} \tilde{C}_{(6)\alpha_1 \dots \alpha_6}^{\mu\nu\gamma} p^{\alpha_1} \dots p^{\alpha_6}, \quad (4.29)$$

which brings the total number of Lorentz indices to nine. In this case, constructing a fully Lorentz-invariant tensor from the Minkowski metric alone is not possible, as any such construction yields tensors with an even number of indices. Furthermore, since the theory under consideration is assumed to preserve parity, terms involving a single Levi-Civita tensor are not allowed, and double Levi-Civita contractions always result in tensors with even rank. As a result, no admissible tensor structure exists for this term, and it must vanish.

This example underscores an important subtlety: the appearance of a term in the UV-divergent part of an amplitude depends not only on the power of momentum but also on the specific index structure of the diagram. Lorentz invariance imposes strict constraints on which tensor structures are permitted, and these constraints must be evaluated on a case-by-case basis depending on the number of external fields and the symmetry properties of the integrand.

As we have seen, the presence or absence of certain momentum powers in the divergent part of an amplitude is governed by the total number of free Lorentz indices available for constructing invariant tensor structures. These indices come not only from the tensor coefficients  $\tilde{C}$  in the loop expansion but also from the external field polarizations. Therefore, it is essential to keep track of the number and type of external legs, as they directly affect the allowed terms in the polynomial expansion.

As a final example, consider again the one-loop three-point function from a triangle diagram, but now suppose that one of the external legs corresponds to a graviton rather than a vector field:



In this case, the amplitude takes the form

$$I^{\mu\nu\rho\sigma}(p_1, p_2, p_3), \quad (4.30)$$

where  $\mu$  and  $\nu$  are associated with the polarization vectors of the two vector fields, and  $\rho\sigma$  with the polarization tensor of the external graviton.

The corresponding polarization structure now involves three external legs, but the total number of Lorentz indices has increased to four due to the symmetric rank-2 polarization tensor  $\epsilon^{\rho\sigma}$  for the graviton. Upon contraction with the tensor coefficient from the loop integral, the total number of indices becomes

$$n_{\text{total}} = 4 + n,$$

where  $n$  is the number of momentum factors  $p^{\alpha_1}, \dots, p^{\alpha_n}$  in the expansion term.

For the full amplitude to be Lorentz-invariant,  $n_{\text{total}}$  must be even so that all indices can be contracted using the metric tensor. This immediately implies that only even powers of momenta can appear in the expansion. Odd powers would lead to an odd total number of indices, for which no Lorentz-invariant tensor structure exists using only symmetric combinations of  $\eta^{\mu\nu}$  (and assuming again that parity is preserved, so single  $\epsilon^{\mu\nu\rho\sigma}$  tensors are excluded). This example reinforces the broader conclusion: the structure of UV divergences is not determined solely by the loop momentum power counting, but also by the full Lorentz index structure, which depends on the spins of the external particles. The interplay between power counting and Lorentz invariance is highly sensitive to the particle content of the diagram and must be treated with care.

To conclude this section, we summarize the key steps and guiding principles for determining the form of the UV-divergent part of loop amplitudes using power counting and Lorentz invariance:

1. **Determine the superficial degree of divergence.** For a given diagram, identify the number of loops, propagators, and derivatives in the vertices. The superficial degree of divergence  $D$  provides an upper bound on the degree of momentum dependence in the divergent part of the amplitude.
2. **Focus on the most divergent terms.** Isolate the leading behavior in loop momentum by retaining only those terms in the numerator that maximize powers of loop momenta (e.g., from momentum-dependent propagators or derivative interactions).
3. **Differentiate and reintegrate.** Apply  $D + 1$  derivatives with respect to external momenta to render the integral finite. This procedure guarantees that the divergent part is a polynomial in external momenta of degree at most  $D$ . Reintegrating yields the structure of the divergent terms.
4. **Express coefficients in dimensionless form.** Introduce a mass scale  $m$  and factor out appropriate powers so that the polynomial coefficients become dimensionless tensors:

$$I^{\mu\nu}(p)\Big|_{\text{div}} = m^2 \tilde{C}_{(0)}^{\mu\nu} + \tilde{C}_{(2)\alpha_1\alpha_2}^{\mu\nu} p^{\alpha_1} p^{\alpha_2} + \frac{1}{m^2} \tilde{C}_{(4)\alpha_1\cdots\alpha_4}^{\mu\nu} p^{\alpha_1} \cdots p^{\alpha_4} + \cdots \quad (4.31)$$

5. **Ensure even total number of Lorentz indices.** The sum of Lorentz indices from external fields (polarizations) and from momentum powers must be even to allow contraction into Lorentz-invariant scalars using only the Minkowski metric. Terms with an odd total number of indices are excluded in parity-preserving theories, as no such invariant tensor structure can be constructed.

This systematic procedure provides a practical tool to identify the structure of divergences without performing the full loop integration. It also clarifies the interplay between UV behavior and symmetry constraints in effective field theories involving higher-spin fields and derivative interactions.

### 4.3 Quantum Stability of Generalized Proca Theories in Flat Spacetime

In the preceding chapters, we presented the structure of generalized Proca theories formulated on a flat Minkowski background, as encapsulated in Eq. (3.64). These theories are characterized by a set of functions  $f_i$  that depend on Lorentz-invariant combinations of the vector field  $A_\mu$  and its derivatives. In particular, the relevant scalar building blocks entering the Lagrangian are given by:

$$X = -\frac{1}{2}A_\mu A^\mu, \quad F = F_{\mu\nu}F^{\mu\nu}, \quad Y = A^\mu A^\nu F_\mu{}^\alpha F_{\nu\alpha}, \quad (4.32)$$

where the field strength tensor is defined as  $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ .

In this chapter, we focus on a simplified subclass of these theories by choosing the functions  $f_i$  to be either constant or linear in their arguments. This restriction significantly reduces the complexity of the interaction terms, allowing for a more tractable quantum analysis. Our treatment closely follows the methodology developed in [28], where the quantum behavior of such simplified models was studied in detail. In this simplified subclass, the functions  $f_i(X)$  and  $\tilde{f}_i(X)$  introduced in Eq. (3.64) take on linear or constant forms. Explicitly, we define these functions as follows:

$$f_3(X) = J_3 X, \quad f_4(X) = J_4 X, \quad \tilde{f}_4(X) = \tilde{J}_4 X, \quad f_{5,6}(X) = J_{5,6}, \quad \tilde{f}_{5,6}(X) = \tilde{J}_{5,6}. \quad (4.33)$$

With these choices, the theory is simplified significantly, facilitating the analysis of quantum stability in subsequent sections. Concretely, we choose:

$$\begin{aligned} \mathcal{L}_2 &= -\frac{1}{4}F^2 + \frac{1}{2}m^2 A^2, \\ \mathcal{L}_3 &= \frac{m^2}{\Lambda_2^2} J_3 A^2 \partial_\mu A^\mu, \\ \mathcal{L}_4 &= \frac{m^2}{\Lambda_2^4} A^2 \left[ J_4 \left( (\partial_\mu A^\mu)^2 - \partial_\mu A_\nu \partial^\nu A^\mu \right) + \tilde{J}_4 F^2 \right], \\ \mathcal{L}_5 &= -\frac{1}{\Lambda_2^2} \tilde{J}_5 \epsilon^{\mu\nu\rho\sigma} \epsilon^{\alpha\beta\gamma}{}_\sigma \partial_\mu A_\nu \partial_\alpha A_\beta \partial_\rho A_\gamma, \\ \mathcal{L}_6 &= -\frac{1}{\Lambda_2^4} \tilde{J}_6 \epsilon^{\mu\nu\rho\sigma} \epsilon^{\alpha\beta\gamma\delta} \partial_\mu A_\nu \partial_\alpha A_\beta \partial_\rho A_\gamma \partial_\sigma A_\delta. \end{aligned} \quad (4.34)$$

With this specific choice, the terms proportional to  $J_5$  and  $J_6$  become total derivatives and thus can be neglected. The theory presented here is characterized by two fundamental

classical scales: the mass of the vector field,  $m$ , and the interaction scale,  $\Lambda_2$ . The scale  $\Lambda_2$  determines the strength of interactions and organizes the terms in the Lagrangian according to their number of fields  $n$ . Explicitly, each interaction term is suppressed by factors of  $\frac{1}{\Lambda_2^{2n-4}}$ , where  $n$  counts the total number of fields involved.

Furthermore, the ratio of these two scales,  $m/\Lambda_2$ , forms a dimensionless parameter that can be interpreted as an effective coupling constant. Typically, this parameter is assumed to be small, ensuring the validity of perturbative expansions. The full description of the Feynman rules for this theory is given in appendix A.2.

## 4.4 Stückelberg Expansion and the Decoupling Limit

To investigate the quantum stability of generalized Proca theories, it is convenient to rewrite the theory by explicitly introducing an auxiliary scalar field. This reformulation introduces redundancy by replacing the vector field  $A_\mu$  according to the Stückelberg rescription [12, 28]:

$$A_\mu \rightarrow A_\mu + \frac{1}{m} \partial_\mu \phi, \quad (4.35)$$

where the scalar field  $\phi$  plays the role of the Goldstone boson, which is "eaten" by the vector field. The mass scale  $m$  is fixed by requiring canonical normalization of the scalar kinetic term. This approach explicitly reintroduces the gauge redundancy into the massive vector theory, enabling a gauge-invariant formulation. Importantly, under this replacement, gauge-invariant quantities such as  $F_{\mu\nu}$  and its dual  $\tilde{F}_{\mu\nu}$  remain unaffected.

It is natural then to introduce a gauge-covariant derivative defined as

$$D_\mu \phi = \partial_\mu \phi + m A_\mu, \quad (4.36)$$

with the theory becoming invariant under the simultaneous gauge transformations

$$\phi \rightarrow \phi + m\alpha, \quad A_\mu \rightarrow A_\mu - \partial_\mu \alpha. \quad (4.37)$$

Choosing the unitary gauge  $\alpha = -\frac{\phi}{m}$  restores the original formulation by eliminating  $\phi$ , explicitly showing that only three physical degrees of freedom are propagated.

Another useful gauge choice, implemented via the Fadeev-Popov procedure, is given by

$$\partial_\mu A^\mu + m\phi = 0. \quad (4.38)$$

In this gauge, the propagators for the fields  $A_\mu$  and  $\phi$  become simple:

$$\frac{-i\eta_{\mu\nu}}{p^2 + m^2} \quad \text{and} \quad \frac{-i}{p^2 + m^2}, \quad (4.39)$$



respectively. Crucially, at high energies, these propagators behave as  $\sim 1/p^2$ , in contrast to the original massive vector propagator, which behaved as  $\sim 1/m^2$ . Thus, this gauge choice is particularly advantageous when studying the high-energy behavior and quantum stability of the theory.

To better understand the effective field theory (EFT), we identify the energy scale at which the theory becomes strongly coupled. For instance, considering the scalar sector, a representative  $2 \rightarrow 2$  scattering amplitude arising from the schematic operator  $\sim \frac{m^2}{\Lambda_2^4} \frac{1}{m^4} (\partial\phi)^2 (\partial^2\phi)^2$  from  $\mathcal{L}_4$  scales as

$$\mathcal{M}_{2 \rightarrow 2} \sim \frac{E^6}{\Lambda_2^4 m^2}. \quad (4.40)$$

Therefore, the scale at which strong coupling occurs is given by

$$\Lambda_3 \equiv (\Lambda_2^2 m)^{1/3}. \quad (4.41)$$

Provided that the parameter  $m/\Lambda_2$  remains small (which corresponds to a weak classical coupling regime), there exists a parametrically large separation between this strong coupling scale  $\Lambda_3$  and the vector mass scale  $m$ . This scale hierarchy ensures the consistency and validity of the EFT, and it is crucial for the applicability of the so-called decoupling limit. In this limit, the theory is analyzed by zooming in at energies close to  $\Lambda_3$ , while simultaneously sending the other scales away, i.e.,

$$m \rightarrow 0, \quad \Lambda_2 \rightarrow \infty, \quad \text{with} \quad \Lambda_3 = (\Lambda_2^2 m)^{1/3} = \text{const.} \quad (4.42)$$

This decoupling limit significantly simplifies the theory, effectively decoupling the vector modes from the scalar mode, except in gauge-invariant structures involving  $F_{\mu\nu}$  and  $\tilde{F}_{\mu\nu}$ . In particular, the gauge symmetry (introduced in Eq. (3.3)) splits into two separate symmetries: a pure gauge symmetry for  $A_\mu$  and a global shift symmetry for  $\phi$ ,

$$\phi \rightarrow \phi + c, \quad A_\mu \rightarrow A_\mu - \partial_\mu \alpha. \quad (4.43)$$

In this simplified regime, the Lagrangian is obtained by substituting

$$A_\mu \rightarrow \frac{1}{m} \partial_\mu \phi, \quad F \rightarrow F, \quad \tilde{F} \rightarrow \tilde{F}, \quad (4.44)$$

and subsequently taking the limit (3.8). This procedure yields the decoupling limit La-

grangian:

$$\begin{aligned}
\mathcal{L}_2 &= -\frac{1}{4}F^2 + \frac{1}{2}(\partial\phi)^2, \\
\mathcal{L}_3 &= \frac{J_3}{\Lambda_3^3}(\partial\phi)^2\Box\phi, \\
\mathcal{L}_4 &= \frac{J_4}{\Lambda_3^6}(\partial\phi)^2\left[(\Box\phi)^2 - (\partial_\mu\partial_\nu\phi)^2\right], \\
\mathcal{L}_5 &= -\frac{\tilde{J}_5}{\Lambda_3^3}\tilde{F}^{\mu\alpha}\tilde{F}_\alpha{}^\nu\partial_\mu\partial_\nu\phi, \\
\mathcal{L}_6 &= -\frac{\tilde{J}_6}{\Lambda_3^6}\tilde{F}^{\mu\alpha}\tilde{F}^{\nu\beta}\partial_\mu\partial_\nu\phi\partial_\alpha\partial_\beta\phi.
\end{aligned} \tag{4.45}$$

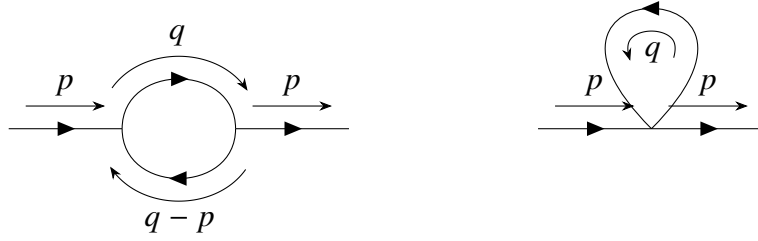
Notice that, in this limit, the mass term for the vector field and the terms proportional to  $\tilde{J}_4$  vanish, whereas the remaining scalar-dependent terms reduce precisely to scalar Galileon interactions [27].

From these considerations, we conclude that the decoupling limit can equivalently be viewed as a high-energy limit of the original massive vector theory. Specifically, lowering the mass  $m$  is analogous to probing higher energies, where the longitudinal polarization mode (represented by the scalar field  $\phi$ ) dominates. This aligns naturally with the Goldstone boson equivalence theorem: at low energies all three polarizations appear equivalent, but at energies approaching  $\Lambda_3$ , the longitudinal mode is clearly distinguished from the transverse modes.

Thus, analyzing quantum stability in the decoupling limit provides clarity by isolating the longitudinal mode, simplifying calculations of radiative corrections and potential instabilities. In this limit, the radiative stability of the theory is predominantly governed by the scalar sector, reducing the complexity of the quantum corrections significantly. Before systematically addressing the hierarchy and stability of interactions, we proceed in the following section with explicit computations of one-loop divergences and the associated counterterms.

#### 4.4.1 Two-Point Functions

In the following analysis, we employ dimensional regularization combined with the  $\overline{\text{MS}}$  subtraction scheme to systematically isolate and extract the one-loop divergences arising from these Feynman diagrams using the tools designed for the loop computations [29, 30, 31, 32, 33].



**Figure 4.1.** One-loop, two-point diagrams. The diagram on the left represents a standard bubble diagram with two internal propagators. The diagram on the right corresponds to a tadpole diagram, where a single internal propagator forms a loop attached to an external leg.

The diagrams for the two-point functions are represented in figure 4.1.

For the diagram on the left in Figure 4.1, commonly referred to as the *bubble diagram*, one-loop contributions can arise from interactions involving two insertions of  $\mathcal{L}_3$  or  $\mathcal{L}_5$ . As a representative example, we consider here the case of two insertions of  $\mathcal{L}_3$ , leading to the following amplitude:

$$\begin{aligned} \Gamma^{\mu\nu}(p) = \frac{-2i}{\Lambda_2^4} \int \frac{d^4 q}{(2\pi)^4} \frac{1}{(q^2 - m^2)((q-p)^2 - m^2)} \cdot \left( -q^\alpha q^\beta + m^2 \eta^{\alpha\beta} \right) \\ \cdot \left( -(q-p)^\gamma (q-p)^\theta + m^2 \eta^{\gamma\theta} \right) \cdot \left( -p^\mu \eta_{\alpha\gamma} - (q-p)^\gamma \eta_\alpha^\mu + q^\alpha \eta_\gamma^\mu \right) \\ \cdot \left( -p^\nu \eta_{\beta\theta} + (q-p)^\theta \eta_\beta^\nu + q^\beta \eta_\theta^\nu \right) \end{aligned} \quad (4.46)$$

Using the power counting techniques developed in the previous section, we can isolate the UV-divergent part of the amplitude by retaining only the highest powers of the loop momentum  $q$ . One such contribution is:

$$\Gamma_{\text{div}}^{\mu\nu}(p) \supset \frac{-2i}{\Lambda_2^4} \int \frac{d^4 q}{(2\pi)^4} \frac{q^\mu q^\nu (q \cdot q)^2}{(q^2 - m^2)((q-p)^2 - m^2)}. \quad (4.47)$$

The UV-divergent part of this integral, which has superficial degree of divergence  $D = 6$ , takes the general polynomial form:

$$\Gamma_{\text{div}}^{\mu\nu}(p) = m^2 C_{(0)}^{\mu\nu} + C_{(2)\alpha\beta}^{\mu\nu} p^\alpha p^\beta + \frac{1}{m^2} C_{(4)\alpha\beta\gamma\delta}^{\mu\nu} p^\alpha p^\beta p^\gamma p^\delta + \frac{1}{m^4} C_{(6)\alpha_1 \dots \alpha_6}^{\mu\nu} p^{\alpha_1} \dots p^{\alpha_6}, \quad (4.48)$$

where each  $C_{(n)}^{\mu\nu}$  is a dimensionless Lorentz-covariant tensor built from contractions of the Minkowski metric, encoding the tensor structure of the divergent contribution.

Following the power counting machinery developed earlier and omitting Lorentz indices for clarity, the divergent part of the amplitude scales as:

$$\Gamma_{\text{div}}(p) \sim \frac{m^4}{\Lambda_2^4} \left( m^2 + p^2 + \frac{p^4}{m^2} + \frac{p^6}{m^4} \right), \quad (4.49)$$

where each term corresponds to a local operator in the effective action with an increasing number of derivatives. The presence of higher-derivative contributions is a direct consequence of the derivative self-interaction in  $\mathcal{L}_3$ , while the overall scaling is governed by the suppression scale  $\Lambda_2$ .

Repeating the same procedure for the remaining loop diagrams—including mixed contributions such as  $\mathcal{L}_3\mathcal{L}_5$  and contact terms from  $\mathcal{L}_4$ , as well as higher-order combinations like  $\mathcal{L}_5\mathcal{L}_5$  and  $\mathcal{L}_6$ —and also including the tadpole diagram shown on the right in Figure 4.1, we obtain the following schematic expressions for the one-loop divergent part of the two-point function:

$$\begin{aligned} \Gamma_{\mathcal{L}_3\mathcal{L}_3}^{\text{div}}(p) &\sim \frac{m^4}{\Lambda_2^4} \left( m^2 + p^2 + \frac{p^4}{m^2} + \frac{p^6}{m^4} \right), \\ \Gamma_{\mathcal{L}_3\mathcal{L}_5, \mathcal{L}_4}^{\text{div}}(p) &\sim \frac{m^2}{\Lambda_2^4} \left( m^4 + m^2 p^2 + p^4 + \frac{p^6}{m^2} + \frac{p^8}{m^4} \right), \\ \Gamma_{\mathcal{L}_5\mathcal{L}_5, \mathcal{L}_6}^{\text{div}}(p) &\sim \frac{1}{\Lambda_2^4} \left( m^6 + m^4 p^2 + m^2 p^4 + p^6 + \frac{p^8}{m^2} + \frac{p^{10}}{m^4} \right), \end{aligned} \quad (4.50)$$

where  $p$  denotes the external momentum. Each term in the expansions corresponds to a local operator consistent with Lorentz invariance and the symmetries of the theory, organized in increasing powers of derivatives acting on the fields. Based on the structure of the one-loop divergent terms presented above, we observe that the classical operators appearing in the tree-level Lagrangian do receive quantum corrections. That is, the theory is not radiatively stable in the strict sense, and the classical structure is renormalized by loop-induced counterterms.

One-loop terms proportional to  $p^2$  unavoidably induce corrections to the kinetic structure of the theory. In particular, they generate contributions of the form

$$\sim \frac{m^4}{\Lambda_2^4} (\partial_\mu A^\mu)^2, \quad (4.51)$$

which detune the gauge-invariant structure appearing in the Proca kinetic operator. While the classical kinetic term is given by

$$\mathcal{L}_2 = -\frac{1}{4}F^2 + \frac{1}{2}m^2 A^2,$$

the operator  $(\partial_\mu A^\mu)^2$  explicitly breaks the accidental gauge symmetry of the massless limit and must be introduced as a counterterm to cancel the associated divergence.

Other examples of radiatively generated operators include higher-derivative terms such as

$$\sim \frac{m^2}{\Lambda_2^4} A^2 \partial_\mu A_\nu \partial^\mu A^\nu, \quad \sim \frac{1}{\Lambda_2^4} A^2 \square^2 A^2, \quad (4.52)$$

as well as detunings of the specific combinations present in  $\mathcal{L}_4$ , for example:

$$\sim \frac{m^2}{\Lambda_2^4} A^2 \left( (\partial_\mu A^\mu)^2 + \partial_\mu A_\nu \partial^\nu A^\mu \right). \quad (4.53)$$

These loop-induced operators illustrate that, in the absence of symmetry protection, quantum corrections generically regenerate all local structures allowed by dimensional analysis and Lorentz invariance. As such, any classical tuning is generically destabilized under renormalization.

To evaluate the physical relevance of these corrections—especially their impact on the propagating degrees of freedom—it is important to examine the theory in the decoupling limit introduced in Eq. (4.42). This limit isolates the longitudinal mode, which controls the high-energy dynamics and is most sensitive to radiative corrections.

In this regime, many of the loop-induced operators become suppressed. For instance, the operator

$$(\partial_\mu A^\mu)^2 \sim \frac{1}{m^2} (\square \phi)^2$$

generates a ghost with mass

$$m_t^2 \sim \frac{\Lambda_3^6}{m^4},$$

which remains well above the cutoff  $\Lambda_3$ , making it irrelevant for the low-energy theory.

However, the most dangerous contributions are those involving the highest powers of external momentum, as they dominate in the ultraviolet and can jeopardize the consistency of the effective field theory. For instance, the  $p^8$  term in the loop expansion induces a counterterm of the form

$$\sim \frac{\partial^8}{\Lambda_2^4 m^2} A^2 \xrightarrow{\text{DL}} \frac{\partial^6}{\Lambda_3^6} \cdot \frac{\partial^2}{m^2} (\partial \phi)^2, \quad (4.54)$$

where we have used the longitudinal mode identification  $A_\mu \sim \partial_\mu \phi / m$  valid in the decoupling limit.

In this limit, the mass parameter  $m$  acts as a suppression scale for the scalar mode, and as  $m \rightarrow 0$ , the induced term becomes increasingly divergent. The presence of the explicit

$1/m^2$  factor indicates that the quantum correction grows without bound in the decoupling limit and eventually overwhelms the classical kinetic term. This signals a breakdown of perturbative control and suggests that the effective theory may not be radiatively stable.

At this stage, everything appears problematic: from the perspective of naive power counting, the theory seems to be destabilized by quantum corrections, and the proliferation of higher-derivative operators threatens the very structure of the classical action. However, we go beyond this superficial analysis and perform the loop integrals explicitly.

To go beyond the naive power counting arguments, we now compute the one-loop two-point amplitude explicitly using dimensional regularization in  $d = 4 + 2\epsilon$ . The UV-divergent part of the amplitude takes the form:

$$\begin{aligned} \Gamma_2^{\text{div}} = \frac{\epsilon_p^\alpha \epsilon_{-p}^\beta}{16\pi^2 \epsilon \Lambda_2^4} & \left[ p^2 \eta_{\alpha\beta} m^4 \left( -3J_3^2 + 6\tilde{J}_4 - 4J_3\tilde{J}_5 + 2\tilde{J}_5^2 \right) + \eta_{\alpha\beta} m^6 \left( -3J_3^2 + 6\tilde{J}_4 \right) \right. \\ & + p_\alpha p_\beta m^4 \left( 12J_3^2 - 6\tilde{J}_4 + 16J_3\tilde{J}_5 + \frac{11}{2}\tilde{J}_5^2 \right) \\ & + p^2 p_\alpha p_\beta m^2 \left( -3J_3^2 - 2J_3\tilde{J}_5 + \frac{2}{3}\tilde{J}_5^2 \right) - p^4 \eta_{\alpha\beta} \frac{m^2}{\Lambda_2^4} \cdot \frac{19}{6}\tilde{J}_5^2 \\ & + p^4 p_\alpha p_\beta \left( \frac{1}{2}J_3^2 - \frac{13}{12}\tilde{J}_5^2 \right) + p^6 \eta_{\alpha\beta} \cdot \frac{4}{3}\tilde{J}_5^2 \\ & \left. + \frac{p^6}{m^2} \cdot \frac{1}{6}\tilde{J}_5^2 \left( p_\alpha p_\beta - p^2 \eta_{\alpha\beta} \right) \right] \end{aligned} \quad (4.55)$$

This result explicitly displays the UV divergence structure of the two-point function, including both gauge-invariant and non-gauge-invariant tensor structures. The polarization tensors  $\epsilon_p^\alpha$  and  $\epsilon_{-p}^\beta$  project the amplitude onto the physical states of the external vector fields. We now analyze the structure of the divergent part given in Eq. (4.55). First, we observe that certain operators do not contribute to the two-point function at one-loop order. In particular, the quartic self-interaction in  $\mathcal{L}_4$  proportional to  $J_4$ , as well as the quartic interaction  $\mathcal{L}_6$ , leave no imprint on the divergence structure at this level. This is consistent with expectations, as these terms do not contain the tensor structures that could contribute to the relevant diagrams.

Among the operators that do contribute, the term proportional to  $\tilde{J}_4$  leads to a divergence that preserves the gauge-compatible structure of the quadratic operator,

$$\left( \square + m^2 \right) \eta_{\alpha\beta} - \partial_\alpha \partial_\beta,$$

which is consistent with the nature of  $\tilde{J}_4 F^2$  as a gauge-invariant interaction. In contrast, the contributions involving  $J_3$  and  $\tilde{J}_5$ —originating from  $\mathcal{L}_3$  and  $\mathcal{L}_5$ , respectively—generate

deviations from this structure. These result in the expected detuning, introducing terms that could break gauge invariance and potentially reintroduce higher-derivative ghosts, as discussed earlier.

However, and crucially, only terms with sufficiently high powers of external momentum pose a real threat to the consistency of the effective field theory. Referring to the final line of Eq. (4.55), we note that while terms of order  $p^8$  are present, those scaling as  $p^{10}$  are entirely absent—even though nothing forbids their appearance from a purely dimensional or symmetry standpoint. This non-trivial cancellation indicates that the structure of the generalized Proca theory is specially arranged to eliminate the most dangerous contributions at one loop.

Furthermore, the  $p^8$  terms that do survive are remarkably well-behaved: they appear only in the specific gauge-preserving combination

$$(\Box\eta_{\alpha\beta} - \partial_\alpha\partial_\beta),$$

which ensures that the induced counterterm does not spoil the structure of the theory. In the decoupling limit, this term leads to a harmless quantum correction of the form

$$\sim \frac{\partial^6}{\Lambda_2^4 m^2} F^2 \xrightarrow{\text{DL}} \frac{\partial^6}{\Lambda_3^6} F^2, \quad (4.56)$$

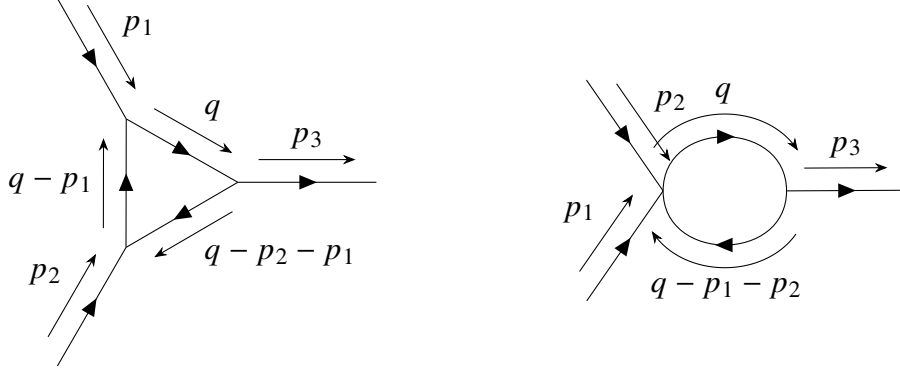
which fits naturally within the hierarchy of terms expected in the effective theory. The correction remains under control relative to the classical action, confirming that the theory is protected against the most dangerous loop effects despite the absence of an explicit symmetry enforcing this cancellation.

#### 4.4.2 Three-Point Functions

We now extend our analysis to the three-point function, following the same strategy as in the two-point case. That is, we compute the one-loop diagrams contributing to the three-point amplitude, isolate the divergent parts using dimensional regularization, and study their tensor structure and dependence on external momenta. Our goal is to determine whether dangerous, radiatively generated terms arise and whether the effective theory remains under control.

The relevant diagrams are shown in Figure 4.2. The diagram on the left is the familiar triangle topology, where each external momentum enters at a different vertex of the loop. The diagram on the right represents a bubble-type contribution in which two of the external

legs couple to the same vertex. This second topology appears in three different channel permutations, corresponding to the different possible pairings of external legs.



**Figure 4.2.** One-loop, three-point diagrams contributing to the amplitude. The diagram on the left represents the triangle topology, where all three external legs are attached to different vertices of the loop. The diagram on the right corresponds to a bubble-type diagram, where two external legs couple to the same vertex. This second diagram appears in three different channel permutations, depending on which pair of external legs share a vertex.

Before diving into explicit computations, we can again anticipate the possible structure of UV divergences using dimensional analysis and Lorentz invariance. Unlike the two-point function, the three-point amplitude carries three Lorentz indices associated with the external polarization vectors, which must be contracted with momenta or metric tensors to form invariant combinations. This increases the number of available tensor structures and also the number of potentially dangerous terms.

Based on operator insertions and loop power counting, the divergent part of the amplitude is expected to scale schematically as follows:

$$\begin{aligned}
 \Gamma_{\mathcal{L}_3^3}^{\text{div}} &\sim \frac{m^6}{\Lambda_2^6} \left( p + \frac{p^3}{m^2} + \frac{p^5}{m^4} + \frac{p^7}{m^6} \right), \\
 \Gamma_{\mathcal{L}_3^2 \mathcal{L}_5, \mathcal{L}_3 \mathcal{L}_4}^{\text{div}} &\sim \frac{m^4}{\Lambda_2^6} \left( m^2 p + p^3 + \frac{p^5}{m^2} + \frac{p^7}{m^4} + \frac{p^9}{m^6} \right), \\
 \Gamma_{\mathcal{L}_3 \mathcal{L}_5^2, \mathcal{L}_3 \mathcal{L}_6, \mathcal{L}_5 \mathcal{L}_4}^{\text{div}} &\sim \frac{m^2}{\Lambda_2^6} \left( m^4 p + m^3 p^3 + p^5 + \frac{p^7}{m^2} + \frac{p^9}{m^4} + \frac{p^{11}}{m^6} \right), \\
 \Gamma_{\mathcal{L}_5^3, \mathcal{L}_5 \mathcal{L}_6}^{\text{div}} &\sim \frac{1}{\Lambda_2^6} \left( m^6 p + m^4 p^3 + m^3 p^5 + p^7 + \frac{p^9}{m^2} + \frac{p^{11}}{m^4} + \frac{p^{13}}{m^6} \right),
 \end{aligned} \tag{4.57}$$



where  $p$  schematically denotes external momenta, and the series truncates at  $1/m^4$  or  $1/m^6$  depending on the number of propagators and insertions involved in the diagram.

Once again, we find that the contributions involving high powers of external momenta—especially those scaling as  $p^{11}$  or higher—deserve special attention. In particular, loop diagrams involving combinations such as  $\mathcal{L}_3\mathcal{L}_5^2$ ,  $\mathcal{L}_5^3$ , and  $\mathcal{L}_5\mathcal{L}_6$  potentially generate dangerous counterterms that could destabilize the effective field theory if they are not suppressed or protected by symmetry. We now proceed to compute the divergent part of the amplitude explicitly and assess whether such terms actually appear.

Regardless of the precise form of the individual contributions, the most significant outcome of the expansion in Eq. (4.2) is the absence of any terms scaling as  $\sim p^{13}$  or  $\sim p^{11}$  in the divergent part of the one-loop three-point function. This absence is non-trivial: such terms would have been allowed by dimensional analysis and Lorentz invariance, and their presence could have led to severe radiative instabilities in the effective field theory. Their cancellation is therefore a non-obvious consistency check on the structure of the generalized Proca interactions. These cancellations suggest that the theory is protected against the most dangerous quantum corrections, even in the absence of an explicit symmetry that would forbid such terms. The UV-divergent contributions that do remain are consistent with the EFT expansion and do not threaten its validity below the strong coupling scale.

For a detailed account of the computational techniques and methods employed in deriving these results, the reader is referred to [28]. The structure of the one-loop four-point function is also discussed in the above reference. Although we do not reproduce the four-point analysis here, the outcome is entirely consistent with the findings from the two- and three-point sectors. Specifically, the four-point function does not introduce any instabilities: either the generated terms are sufficiently suppressed—so that any associated ghost degrees of freedom lie above the EFT cutoff—or the potentially dangerous high-derivative operators are absent altogether. In cases where higher-derivative structures do appear, they are organized into gauge-invariant combinations that do not threaten the consistency of the effective theory. This reinforces the overall conclusion that the generalized Proca theory, despite its derivative self-interactions, exhibits remarkable quantum stability at one loop.

## 4.5 Quantum Stability from the Decoupling Limit

In the previous sections, we explicitly computed one-loop amplitudes for the two- and three-point functions in a restricted generalized Proca model on flat spacetime. These

results revealed non-trivial cancellations that eliminate dangerous high-momentum divergences, suggesting that the theory is radiatively stable despite its derivative self-interactions and the absence of manifest symmetry protection. We now summarize how this structure can be understood and extended to all loop orders and all  $n$ -point functions using the decoupling limit.

The decoupling limit (DL), as defined in equation (4.42), isolates the high-energy behavior of the theory by sending the vector mass to zero while keeping the strong coupling scale  $\Lambda_3 = (m^2 \Lambda_2)^{1/3}$  fixed. This procedure introduces a scalar Stückelberg field  $\phi$  through the replacement

$$A_\mu \rightarrow A_\mu + \frac{1}{m} \partial_\mu \phi, \quad (4.58)$$

which makes the longitudinal mode of the vector field explicit. In this limit, the transverse and longitudinal degrees of freedom decouple at the level of symmetries, and the interactions reduce to a set of Galileon-like scalar terms and gauge-invariant vector structures.

Quantum corrections in the DL are organized in terms of three expansion parameters:

$$\alpha_{\text{cl}} = \frac{\partial^2 \phi}{\Lambda_3^3}, \quad \alpha_{\text{q}} = \frac{\partial^2}{\Lambda_3^2}, \quad \alpha_{\bar{\text{q}}} = \frac{F^2}{\Lambda_3^4}, \quad (4.59)$$

where  $\alpha_{\text{cl}}$  characterizes the size of classical non-linearities, while  $\alpha_{\text{q}}$  and  $\alpha_{\bar{\text{q}}}$  quantify quantum loop contributions. Since every loop-induced counterterm must include a non-zero power of  $\alpha_{\text{q}}$  or  $\alpha_{\bar{\text{q}}}$ , quantum corrections are always suppressed relative to the classical operators provided the energy remains below the cutoff scale  $\Lambda_3$ . In particular, the classical Galileon-like terms are non-renormalized, and all potentially dangerous operators are absent or appear only in a gauge-invariant form, such as  $F^2$ .

This structure guarantees that the effective field theory remains under control: classical non-linearities can become important (e.g. for Vainshtein [34] screening) without inducing large quantum corrections that destabilize the EFT. As emphasized in [28], the absence of dangerous operators like  $F^3$  or  $\partial^{10} A^2$  at all loop orders and for arbitrary  $n$ -point functions can be understood purely from dimensional and symmetry considerations in the decoupling limit. Furthermore, this argument confirms that taking the decoupling limit and computing loop corrections commute, thereby validating the consistency of the loop-level results obtained in the unitary gauge.

We refer the reader to [28] for a complete and rigorous derivation of these results and their implications for the full generalized Proca effective field theory.

In this chapter, we analyzed the quantum stability of generalized Proca theories in the flat Minkowski background. By computing explicit one-loop amplitudes for two- and

three-point functions, we identified the structure of UV divergences and demonstrated that potentially dangerous operators are either absent or appear only in safe, gauge-invariant combinations. These results suggest that the effective field theory remains under control despite its non-renormalizable interactions. In the next chapter, we extend this analysis to the case of curved spacetime, where gravitational interactions are included via metric perturbations.



# Chapter 5

## Extending Stability Analysis to Curved Spacetime

In this section, we study the dynamics of Generalized Proca theories in curved spacetime, with a particular focus on their behavior in the presence of weak gravitational fields. More precisely, we aim to understand how the Proca field interacts with small perturbations of the metric around flat Minkowski spacetime. This setup corresponds to the regime where gravity is weak, and the spacetime metric can be treated as a perturbation around the flat background [35, 36, 37, 38].

### 5.1 Framework and Background Setup

We begin our analysis with the full Lagrangian for the Generalized Proca theory (see Eq. (3.70)). We then apply a perturbative expansions:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \quad (5.1)$$

$$A^\mu = \bar{A}^\mu + \delta A^\mu, \quad (5.2)$$

where  $h_{\mu\nu}$  represents a small perturbation describing weak gravitational effects [39].

For a zero background configuration of the Proca field,  $\bar{A}_\mu = 0$ , the equations of motion at the background level are determined entirely by the gravitational sector. In this case, the background field equations reduce to the following condition:

$$G_2 \left( \frac{1}{2} A^\mu A_\mu, -\frac{1}{4} F^2 \right) \Big|_{A_\mu=0} = 0. \quad (5.3)$$

From here on, we use a bar to indicate that a function is evaluated on the background, e.g.,  $\bar{G}_2 \equiv G_2|_{A_\mu=0}$ . With this notation, the background equation simplifies to:

$$\bar{G}_2 = 0. \quad (5.4)$$

### 5.1.1 Quadratic Order

Expanding the action up to quadratic order in the perturbations  $\delta A_\mu$  and  $h_{\mu\nu}$ , we obtain:

$$\begin{aligned} \mathcal{L}_2^{(2)} &= \frac{1}{2} (\partial_\mu \delta A_\nu \partial^\nu \delta A^\mu - \partial_\mu \delta A_\nu \partial^\mu \delta A^\nu) \bar{G}_2^{(0,1)} + \frac{1}{2} \delta A_\mu \delta A^\mu \bar{G}_2^{(1,0)}, \\ \mathcal{L}_3^{(2)} &= \frac{1}{2} \bar{G}_3 \delta A^\mu \partial_\mu h + \frac{1}{2} \bar{G}_3 h \partial_\mu \delta A^\mu, \\ \mathcal{L}_4^{(2)} &= \bar{G}_4 h^{\mu\nu} \partial_\nu \partial_\mu h - \frac{1}{4} \bar{G}_4 \partial_\nu h \partial^\nu h - \bar{G}_4 \partial_\mu h^{\mu\nu} \partial^\lambda h_{\nu\lambda} + \bar{G}_4 \partial^\nu h \partial^\lambda h_{\nu\lambda} \\ &\quad - 2 \bar{G}_4 h^{\mu\nu} \partial^\lambda \partial_\nu h_{\mu\lambda} + \frac{1}{2} \bar{G}_4 h \partial^\lambda \partial^\nu h_{\nu\lambda} + \bar{G}_4 h^{\mu\nu} \partial^\lambda \partial_\lambda h_{\mu\nu} - \frac{1}{2} \bar{G}_4 h \partial^\lambda \partial_\lambda h \\ &\quad - \frac{1}{2} \bar{G}_4 \partial_\nu h_{\mu\lambda} \partial^\lambda h^{\mu\nu} + \frac{3}{4} \bar{G}_4 \partial^\lambda h_{\mu\nu} \partial_\lambda h^{\mu\nu} + (\partial_\mu \delta A^\mu \partial_\nu \delta A^\nu - \partial_\mu \delta A_\nu \partial^\nu \delta A^\mu) \bar{G}_4^{(1)}, \\ \mathcal{L}_5^{(2)} &= \frac{1}{2} \bar{G}_5 \partial_\mu \partial^\mu h \partial_\nu \delta A^\nu - \frac{1}{2} \bar{G}_5 \partial_\nu \delta A^\nu \partial_\rho \partial_\mu h^{\rho\mu} - \frac{1}{2} \bar{G}_5 \partial_\mu \partial^\mu h^{\nu\rho} \partial_\rho \delta A_\nu \\ &\quad + \frac{1}{2} \bar{G}_5 \partial_\nu \partial_\mu h^{\rho\mu} \partial_\rho \delta A^\nu + \frac{1}{2} \bar{G}_5 \partial_\rho \partial_\mu h^{\nu\mu} \partial^\rho \delta A_\nu - \frac{1}{2} \bar{G}_5 \partial_\rho \partial_\nu h \partial^\rho \delta A^\nu, \\ \mathcal{L}_6^{(2)} &= 0. \end{aligned} \quad (5.5)$$

where,  $h \equiv h^\mu{}_\mu$  denotes the trace of the metric perturbation [40]. From the quadratic Lagrangian, we observe the presence of mixing terms between the Proca field perturbation  $\delta A_\mu$  and the metric perturbation  $h_{\mu\nu}$ . These terms indicate that, at this order, the kinetic operators for the two fields are not diagonal. In other words, the fluctuation Lagrangian contains off-diagonal terms that couple the vector and tensor sectors, implying that the propagator structure is not initially block-diagonal.

Explicitly, the mixed part of the quadratic Lagrangian is given by:

$$\begin{aligned} \mathcal{L}_{\text{mix}}^{(2)} &= \frac{1}{2} \bar{G}_3 \delta A^\nu \partial_\nu h + \frac{1}{2} \bar{G}_3 h \partial_\mu \delta A^\mu \\ &\quad + \frac{1}{2} \bar{G}_5 \partial_\mu \partial^\mu h \partial_\nu \delta A^\nu - \frac{1}{2} \bar{G}_5 \partial_\nu \delta A^\nu \partial_\rho \partial_\mu h^{\rho\mu} - \frac{1}{2} \bar{G}_5 \partial_\mu \partial^\mu h^{\nu\rho} \partial_\rho \delta A_\nu \\ &\quad + \frac{1}{2} \bar{G}_5 \partial_\nu \partial_\mu h^{\rho\mu} \partial_\rho \delta A^\nu + \frac{1}{2} \bar{G}_5 \partial_\rho \partial_\mu h^{\nu\mu} \partial^\rho \delta A_\nu - \frac{1}{2} \bar{G}_5 \partial_\rho \partial_\nu h \partial^\rho \delta A^\nu. \end{aligned} \quad (5.6)$$

However, these mixing terms involve total derivatives and can be eliminated through integration by parts. After appropriate rearrangement, all such terms cancel out, and the quadratic Lagrangian effectively becomes block-diagonal in the tensor and vector sectors. As a result, the graviton and Proca field decouple at the quadratic level, making it straightforward to define the propagators independently for each field.

After performing integration by parts, the mixing terms vanish and the vector and tensor sectors decouple at the quadratic level. That is,

$$\mathcal{L}_{\text{mix}}^{(2)} = 0. \quad (5.7)$$

The simplified quadratic Lagrangians then become:

$$\begin{aligned} \mathcal{L}_2^{(2)} &= -\frac{1}{2} \delta A^\mu \partial^\nu \partial_\mu \delta A_\nu \bar{G}_2^{(0,1)} + \frac{1}{2} \delta A^\mu \partial^\nu \partial_\nu \delta A_\mu \bar{G}_2^{(0,1)} + \frac{1}{2} \delta A_\mu \delta A^\mu \bar{G}_2^{(1,0)}, \\ \mathcal{L}_3^{(2)} &= 0, \\ \mathcal{L}_4^{(2)} &= -\frac{1}{2} \bar{G}_4 h^{\mu\nu} \partial^\rho \partial_\nu h_{\mu\rho} + \frac{1}{2} \bar{G}_4 h \partial^\rho \partial^\nu h_{\nu\rho} + \frac{1}{4} \bar{G}_4 h^{\mu\nu} \partial^\rho \partial_\rho h_{\mu\nu} - \frac{1}{4} \bar{G}_4 h \partial^\rho \partial_\rho h, \\ \mathcal{L}_5^{(2)} &= 0, \\ \mathcal{L}_6^{(2)} &= 0. \end{aligned} \quad (5.8)$$

Before proceeding to compute the propagators, we must perform two essential steps: canonical normalization and gauge fixing.

First, we canonically normalize the fields to bring their kinetic terms into a standard form. This is necessary to correctly identify the propagators. To bring the quadratic Lagrangian into canonical form, we perform the following field redefinitions:

$$\delta A_\mu \rightarrow \frac{\delta A_\mu}{\sqrt{\bar{G}_2^{(0,1)}}}, \quad h_{\mu\nu} \rightarrow -\frac{\sqrt{-2\bar{G}_4}}{\bar{G}_4} h_{\mu\nu}. \quad (5.9)$$

In terms of these canonically normalized fields, the quadratic Lagrangian becomes:

$$\begin{aligned} \mathcal{L}_{\text{canon}}^{(2)} &= -\frac{1}{2} \delta A^\mu \partial^\nu \partial_\mu \delta A_\nu + \frac{1}{2} \delta A^\mu \partial^\nu \partial_\nu \delta A_\mu + \frac{1}{2} \frac{\bar{G}_2^{(1,0)}}{\bar{G}_2^{(0,1)}} \delta A^\mu \delta A_\mu \\ &\quad + h^{\mu\nu} \partial^\rho \partial_\nu h_{\mu\rho} - h \partial^\rho \partial^\nu h_{\nu\rho} - \frac{1}{2} h^{\mu\nu} \partial^\rho \partial_\rho h_{\mu\nu} + \frac{1}{2} h \partial^\rho \partial_\rho h. \end{aligned} \quad (5.10)$$

Second, we need to fix the gauge for the gravitational perturbations. The Lagrangian  $\mathcal{L}_4$ , which contains the kinetic terms for  $h_{\mu\nu}$ , is invariant under linearized diffeomorphisms

of the form

$$h_{\mu\nu} \rightarrow h_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\mu, \quad (5.11)$$

where  $\xi_\mu$  is an arbitrary vector field parametrizing infinitesimal coordinate transformations. This gauge symmetry introduces redundancies in the description of the graviton degrees of freedom, making the kinetic operator non-invertible. To correctly define the graviton propagator, we must add a gauge fixing term that breaks this symmetry and renders the kinetic operator invertible.

We work in the harmonic gauge (also known as de Donder gauge), which imposes the condition

$$\partial^\mu h_{\mu\nu} - \frac{1}{2} \partial_\nu h = 0. \quad (5.12)$$

The corresponding gauge fixing Lagrangian is

$$\mathcal{L}_{\text{GF}} = \frac{1}{\zeta} \left( \partial^\mu h_{\mu\nu} - \frac{1}{2} \partial_\nu h \right) \left( \partial^\rho h_\rho{}^\nu - \frac{1}{2} \partial^\nu h \right), \quad (5.13)$$

where  $\zeta$  is the gauge parameter [39]. In what follows, we choose  $\zeta = 1$ . This term removes the degeneracy in the graviton kinetic operator and allows us to consistently define the free graviton propagator.

After including the gauge fixing term, the full quadratic Lagrangian becomes

$$\mathcal{L}_{\text{quad}} = \mathcal{L}_2^{(2)} + \mathcal{L}_4^{(2)} + \mathcal{L}_{\text{GF}}, \quad (5.14)$$

where  $\mathcal{L}_2^{(2)}$  is the canonically normalized Proca sector,  $\mathcal{L}_4^{(2)}$  is the canonically normalized graviton sector, and  $\mathcal{L}_{\text{GF}}$  is the gauge fixing term derived in harmonic gauge.

Performing integration by parts and simplifying, the total quadratic Lagrangian becomes:

$$\begin{aligned} \mathcal{L}_{\text{quad}} = & -\frac{1}{2} \delta A^\mu \partial^\nu \partial_\mu \delta A_\nu + \frac{1}{2} \delta A^\mu \partial^\nu \partial_\nu \delta A_\mu + \frac{1}{2} \frac{\bar{G}_2^{(1,0)}}{\bar{G}_2^{(0,1)}} \delta A^\mu \delta A_\mu \\ & - \frac{1}{2} h^{\mu\nu} \partial^\rho \partial_\rho h_{\mu\nu} + \frac{1}{4} h \partial^\rho \partial_\rho h. \end{aligned} \quad (5.15)$$

Here, the first two terms represent the kinetic part of the Proca field, while the third term acts as an effective mass term. The last two terms correspond to the graviton kinetic structure after including the gauge fixing contribution.

From the quadratic Lagrangian, we can now read off the momentum-space propagators for the canonically normalized fields:



$$\mu \quad \xrightarrow{p} \quad \nu = D_{\mu\nu}(p) = \frac{1}{p^2 - m_A^2 + i\epsilon} \left( -\eta_{\mu\nu} + \frac{p_\mu p_\nu}{m_A^2} \right), \quad (5.16)$$

for the vector field, where the mass is given by  $m_A^2 = \bar{G}_2^{(1,0)} / \bar{G}_2^{(0,1)}$ . and

$$\mu\nu \quad \xrightarrow{p} \quad \rho\sigma = D_{\mu\nu\rho\sigma}(p) = \frac{1}{2(p^2 + i\epsilon)} (\eta_{\mu\rho}\eta_{\nu\sigma} + \eta_{\mu\sigma}\eta_{\nu\rho} - \eta_{\mu\nu}\eta_{\rho\sigma}) \quad (5.17)$$

For the graviton.

### 5.1.2 Cubic interactions

The cubic Lagrangian contains interactions involving combinations of up to three perturbations from the metric  $h_{\mu\nu}$  and the vector field  $\delta A_\mu$ . Due to the complexity of these terms, here we briefly summarize the structure of each Lagrangian  $\mathcal{L}_n^{(3)}$ , providing full explicit expressions in Appendix A.1.

The primary cubic interaction terms are as follows:

- $\mathcal{L}_2^{(3)}$ : Terms with two vector perturbations and one graviton ( $h\delta A\delta A$ ).
- $\mathcal{L}_3^{(3)}$ : Includes purely vector terms ( $\delta A\delta A\delta A$ ) and mixed graviton-vector terms with two gravitons ( $hh\delta A$ ).
- $\mathcal{L}_4^{(3)}$ : Purely gravitational interactions ( $hhh$ ) and mixed terms involving one graviton and two vector perturbations ( $h\delta A\delta A$ ).
- $\mathcal{L}_5^{(3)}$ : Purely vector cubic interactions ( $\delta A\delta A\delta A$ ), mixed interactions involving one graviton and two vectors ( $h\delta A\delta A$ ), and terms with two gravitons and one vector ( $hh\delta A$ ).
- $\mathcal{L}_6^{(3)}$ : Contains terms involving two vectors and one graviton ( $h\delta A\delta A$ ), typically with higher-order derivatives.

### 5.1.3 Quartic interactions

We now summarize the structure of the quartic interaction terms arising in generalized Proca theories at fourth order in perturbations. At this order, interactions involve up to

four perturbation fields: the vector perturbation  $\delta A_\mu$  and the metric perturbation  $h_{\mu\nu}$ . The possible combinations and the interaction structures are summarized below:

- $\mathcal{L}_2^{(4)}$ : Contains purely vector interactions ( $\delta A^4$ ) as well as mixed vector-graviton interactions involving two gravitons and two vectors ( $h^2 \delta A^2$ ).
- $\mathcal{L}_3^{(4)}$ : Involves terms with three gravitons and one vector ( $h^3 \delta A$ ), and terms with three vectors and one graviton ( $h \delta A^3$ ).
- $\mathcal{L}_4^{(4)}$ : Includes purely gravitational terms ( $h^4$ ), purely vector terms ( $\delta A^4$ ), and mixed terms with two gravitons and two vectors ( $h^2 \delta A^2$ ).
- $\mathcal{L}_5^{(4)}$ : Contains terms with three vectors and one graviton ( $h \delta A^3$ ), typically involving antisymmetric Levi-Civita contractions.
- $\mathcal{L}_6^{(4)}$ : Includes purely vector quartic interactions ( $\delta A^4$ ) and mixed interactions with two gravitons and two vectors ( $h^2 \delta A^2$ ). These terms have the richest derivative and curvature-dependent structures.

## 5.2 Sub-Models and Explicit Lagrangians

We now consider a specific subset of the generalized Proca theories by choosing the following forms for the functions  $G_i$  and  $g_5$ :

$$\begin{aligned}
 G_2(X, F) &= -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{1}{2} m^2 A^2, \\
 G_3(X) &= \frac{m^2}{\Lambda^2} J_3 X, \\
 G_4(X) &= -\frac{2}{\kappa^2}, \\
 G_5(X) &= 0, \\
 g_5(X) &= 0, \\
 G_6(X) &= 0,
 \end{aligned} \tag{5.18}$$

where  $\kappa = \frac{\sqrt{32\pi}}{M_{\text{pl}}}$ , and  $M_p l$  is the Planck mass.

This particular choice corresponds to a theory minimally coupled to gravity due to the simple Einstein-Hilbert form of the gravitational sector represented by the choice  $G_4(X) = -2/\kappa^2$  with no explicit vector-curvature interactions.

With this setup, the quadratic-order action in perturbations reads:

$$\begin{aligned}\mathcal{L}^{(2)} = & \frac{1}{2}m^2 \delta A_\mu \delta A^\mu - \frac{1}{2}\delta A_\mu \partial_\nu \partial^\mu \delta A^\nu + \frac{1}{2}\delta A_\mu \partial_\nu \partial^\nu \delta A^\mu \\ & - \frac{1}{2}h_{\mu\nu} \partial_\alpha \partial^\alpha h^{\mu\nu} + \frac{1}{4}h \partial_\alpha \partial^\alpha h.\end{aligned}\quad (5.19)$$

For the cubic interactions, we have:

$$\begin{aligned}\mathcal{L}_2^{(3)} = & \frac{m^2 \sqrt{2\pi}}{M_{\text{pl}}} h \delta A_\alpha \delta A^\alpha + \frac{2m^2 \sqrt{2\pi}}{M_{\text{pl}}} h_{\alpha\beta} \delta A^\alpha \delta A^\beta + \frac{\sqrt{2\pi}}{M_{\text{pl}}} h \partial_\alpha \delta A_\beta \partial^\beta \delta A^\alpha \\ & - \frac{\sqrt{2\pi}}{M_{\text{pl}}} h \partial_\beta \delta A_\alpha \partial^\beta \delta A^\alpha - \frac{2\sqrt{2\pi}}{M_{\text{pl}}} h_{\alpha\gamma} \partial_\beta \delta A^\gamma \partial^\beta \delta A^\alpha + \frac{2\sqrt{2\pi}}{M_{\text{pl}}} h_{\beta\gamma} \partial^\beta \delta A^\alpha \partial^\gamma \delta A_\alpha \\ & + \frac{4\sqrt{2\pi}}{M_{\text{pl}}} \delta A^\alpha \partial_\beta h_{\alpha\gamma} \partial^\gamma \delta A^\beta - \frac{4\sqrt{2\pi}}{M_{\text{pl}}} \delta A^\alpha \partial_\gamma h_{\alpha\beta} \partial^\gamma \delta A^\beta,\end{aligned}\quad (5.20)$$

$$\mathcal{L}_3^{(3)} = \frac{J_3 m^2}{2\Lambda^2} \delta A_\beta \delta A^\beta \partial_\alpha \delta A^\alpha, \quad (5.21)$$

$$\begin{aligned}\mathcal{L}_4^{(3)} = & \frac{6\sqrt{2\pi}}{M_{\text{pl}}} h_{\alpha\beta} \partial^\alpha h_{\gamma\delta} \partial^\beta h^{\gamma\delta} - \frac{2\sqrt{2\pi}}{M_{\text{pl}}} h_{\alpha\beta} \partial^\alpha h \partial^\beta h + \frac{\sqrt{2\pi}}{M_{\text{pl}}} h^2 \partial^\beta \partial^\alpha h_{\alpha\beta} - \frac{\sqrt{2\pi}}{M_{\text{pl}}} h^2 \partial^\beta \partial_\beta h \\ & + \frac{8\sqrt{2\pi}}{M_{\text{pl}}} h_{\alpha\beta} \partial^\beta h \partial^\gamma h^\alpha{}_\gamma + \frac{8\sqrt{2\pi}}{M_{\text{pl}}} h_{\alpha\beta} \partial^\beta h^\alpha{}_\gamma \partial^\gamma h + \frac{8\sqrt{2\pi}}{M_{\text{pl}}} h^\alpha{}_\gamma h_{\alpha\beta} \partial^\gamma \partial^\beta h \\ & - \frac{4\sqrt{2\pi}}{M_{\text{pl}}} h h_{\beta\gamma} \partial^\gamma \partial^\beta h - \frac{4\sqrt{2\pi}}{M_{\text{pl}}} h_{\alpha\beta} \partial^\gamma h \partial_\gamma h^{\alpha\beta} + \frac{\sqrt{2\pi}}{M_{\text{pl}}} h \partial^\gamma h \partial_\gamma h \\ & - \frac{8\sqrt{2\pi}}{M_{\text{pl}}} h_{\alpha\beta} \partial^\gamma h^\alpha{}_\gamma \partial^\delta h^\beta{}_\delta - \frac{16\sqrt{2\pi}}{M_{\text{pl}}} h_{\alpha\beta} \partial^\beta h^\alpha{}_\gamma \partial^\delta h^\gamma{}_\delta + \frac{4\sqrt{2\pi}}{M_{\text{pl}}} h \partial^\beta h_{\beta\gamma} \partial^\delta h^\gamma{}_\delta \\ & + \frac{8\sqrt{2\pi}}{M_{\text{pl}}} h_{\alpha\beta} \partial_\gamma h^{\alpha\beta} \partial^\delta h^\gamma{}_\delta - \frac{4\sqrt{2\pi}}{M_{\text{pl}}} h \partial_\gamma h \partial^\delta h^\gamma{}_\delta - \frac{8\sqrt{2\pi}}{M_{\text{pl}}} h_{\alpha\beta} h_{\gamma\delta} \partial^\delta \partial^\beta h^{\alpha\gamma} \\ & + \frac{8\sqrt{2\pi}}{M_{\text{pl}}} h_{\alpha\beta} h_{\gamma\delta} \partial^\delta \partial^\gamma h^{\alpha\beta} - \frac{16\sqrt{2\pi}}{M_{\text{pl}}} h^\alpha{}_\gamma h_{\alpha\beta} \partial^\delta \partial^\gamma h^\beta{}_\delta + \frac{8\sqrt{2\pi}}{M_{\text{pl}}} h h_{\beta\gamma} \partial^\delta \partial^\gamma h^\beta{}_\delta \\ & + \frac{2\sqrt{2\pi}}{M_{\text{pl}}} h_{\alpha\beta} h^{\alpha\beta} \partial^\delta \partial^\gamma h_{\gamma\delta} - \frac{2\sqrt{2\pi}}{M_{\text{pl}}} h^2 \partial^\delta \partial^\gamma h_{\gamma\delta} + \frac{8\sqrt{2\pi}}{M_{\text{pl}}} h^\alpha{}_\gamma h_{\alpha\beta} \partial^\delta \partial_\delta h^{\beta\gamma} \\ & - \frac{4\sqrt{2\pi}}{M_{\text{pl}}} h h_{\beta\gamma} \partial^\delta \partial_\delta h^{\beta\gamma} - \frac{2\sqrt{2\pi}}{M_{\text{pl}}} h_{\alpha\beta} h^{\alpha\beta} \partial^\delta \partial_\delta h + \frac{2\sqrt{2\pi}}{M_{\text{pl}}} h^2 \partial^\delta \partial_\delta h \\ & - \frac{8\sqrt{2\pi}}{M_{\text{pl}}} h_{\alpha\beta} \partial^\beta h_{\gamma\delta} \partial^\delta h^{\alpha\gamma} - \frac{4\sqrt{2\pi}}{M_{\text{pl}}} h_{\alpha\beta} \partial^\gamma h_{\beta\delta} \partial^\delta h^{\alpha\gamma} + \frac{12\sqrt{2\pi}}{M_{\text{pl}}} h_{\alpha\beta} \partial^\delta h_{\beta\gamma} \partial^\delta h^{\alpha\gamma} \\ & + \frac{2\sqrt{2\pi}}{M_{\text{pl}}} h \partial^\gamma h_{\beta\delta} \partial^\delta h^{\beta\gamma} - \frac{3\sqrt{2\pi}}{M_{\text{pl}}} h \partial^\delta h_{\beta\gamma} \partial^\delta h^{\beta\gamma},\end{aligned}\quad (5.22)$$

$$\mathcal{L}_5^{(3)} = 0, \quad (5.23)$$

$$\mathcal{L}_6^{(3)} = 0. \quad (5.24)$$

And for the quartic interactions, we have:

$$\begin{aligned} \mathcal{L}_2^{(4)} = & -\frac{4m^2\pi}{M_{\text{pl}}^2} h_{\beta\gamma} h^{\beta\gamma} \delta A_\alpha \delta A^\alpha + \frac{2m^2\pi}{M_{\text{pl}}^2} h^2 \delta A_\alpha \delta A^\alpha + \frac{8m^2\pi}{M_{\text{pl}}^2} h_{\alpha\beta} h \delta A^\alpha \delta A^\beta - \frac{4\pi}{M_{\text{pl}}^2} h_{\gamma\delta} h^{\gamma\delta} \partial_\alpha \delta A_\beta \partial^\beta \delta A^\alpha \\ & + \frac{2\pi}{M_{\text{pl}}^2} h^2 \partial_\alpha \delta A_\beta \partial^\beta \delta A^\alpha + \frac{4\pi}{M_{\text{pl}}^2} h_{\gamma\delta} h^{\gamma\delta} \partial_\beta \delta A_\alpha \partial^\beta \delta A^\alpha - \frac{2\pi}{M_{\text{pl}}^2} h^2 \partial_\beta \delta A_\alpha \partial^\beta \delta A^\alpha \\ & - \frac{8\pi}{M_{\text{pl}}^2} h_{\alpha\gamma} h \partial_\beta \delta A^\gamma \partial^\beta \delta A^\alpha - \frac{16\pi}{M_{\text{pl}}^2} h^\delta_\beta h_{\gamma\delta} \partial^\beta \delta A^\alpha \partial^\gamma \delta A_\alpha + \frac{8\pi}{M_{\text{pl}}^2} h_{\beta\gamma} h \partial^\beta \delta A^\alpha \partial^\gamma \delta A_\alpha \\ & + \frac{16\pi}{M_{\text{pl}}^2} h \delta A^\alpha \partial_\beta h_{\alpha\gamma} \partial^\gamma \delta A^\beta - \frac{32\pi}{M_{\text{pl}}^2} h^\delta_\gamma \delta A^\alpha \partial_\beta h_{\alpha\delta} \partial^\gamma \delta A^\beta - \frac{16\pi}{M_{\text{pl}}^2} h \delta A^\alpha \partial_\gamma h_{\alpha\beta} \partial^\gamma \delta A^\beta \\ & + \frac{32\pi}{M_{\text{pl}}^2} h^\delta_\gamma \delta A^\alpha \partial^\gamma \delta A^\beta \partial_\delta h_{\alpha\beta} + \frac{16\pi}{M_{\text{pl}}^2} \delta A^\alpha \delta A^\beta \partial_\gamma h_{\beta\delta} \partial^\delta h_{\alpha}{}^\gamma - \frac{16\pi}{M_{\text{pl}}^2} \delta A^\alpha \delta A^\beta \partial_\delta h_{\beta\gamma} \partial^\delta h_{\alpha}{}^\gamma \\ & + \frac{16\pi}{M_{\text{pl}}^2} h_{\alpha\gamma} h_{\beta\delta} \partial^\beta \delta A^\alpha \partial^\delta \delta A^\gamma, \end{aligned} \quad (5.25)$$

$$\mathcal{L}_3^{(4)} = \frac{J_3 m^2 \sqrt{2\pi}}{M_{\text{pl}} \Lambda^2} \delta A^\alpha \delta A_\beta \delta A^\beta \partial_\alpha h + \frac{J_3 m^2 \sqrt{2\pi}}{M_{\text{pl}} \Lambda^2} h \delta A_\beta \delta A^\beta \partial_\alpha \delta A^\alpha + \frac{2J_3 m^2 \sqrt{2\pi}}{M_{\text{pl}} \Lambda^2} h_{\beta\alpha} \delta A^\alpha \delta A^\beta \partial_\gamma \delta A^\gamma, \quad (5.26)$$

$$\begin{aligned} \mathcal{L}_4^{(4)} = & -\frac{4\pi}{M_{\text{pl}}^2} h h h \partial_\beta \partial_\alpha h^{\alpha\beta} + \frac{8\pi}{M_{\text{pl}}^2} h_{\alpha\beta} h h \partial^\alpha \partial^\beta h + \frac{4\pi}{M_{\text{pl}}^2} h h h \partial^\beta \partial_\beta h - \frac{2\pi}{M_{\text{pl}}^2} h h h \partial_\beta h \partial^\beta h \\ & + \frac{32\pi}{M_{\text{pl}}^2} h_{\alpha\beta} h_{\gamma\delta} \partial^\beta h^{\delta\epsilon} \partial^\gamma h_{\epsilon}{}^\alpha - \frac{8\pi}{M_{\text{pl}}^2} h h \partial_\alpha h^{\alpha\beta} \partial_\gamma h_{\beta}{}^\gamma + \frac{8\pi}{M_{\text{pl}}^2} h h \partial^\beta h \partial_\gamma h_{\beta}{}^\gamma \\ & - \frac{48\pi}{M_{\text{pl}}^2} h_{\alpha}{}^\gamma h_{\alpha\beta} \partial^\beta h^{\delta\epsilon} \partial^\gamma h_{\delta\epsilon} + \frac{24\pi}{M_{\text{pl}}^2} h h_{\beta\gamma} \partial^\beta h^{\delta\epsilon} \partial^\gamma h_{\delta\epsilon} + \frac{16\pi}{M_{\text{pl}}^2} h_{\alpha}{}^\gamma h_{\alpha\beta} \partial^\beta h \partial^\gamma h \\ & - \frac{8\pi}{M_{\text{pl}}^2} h h_{\beta\gamma} \partial^\beta h \partial^\gamma h - \frac{16\pi}{M_{\text{pl}}^2} h_{\alpha\beta} h h \partial^\gamma \partial^\alpha h_{\beta}{}^\gamma - \frac{64\pi}{M_{\text{pl}}^2} h_{\alpha}{}^\gamma h_{\alpha\beta} h_{\delta\epsilon} \partial^\gamma \partial^\beta h^{\delta\epsilon} \\ & + \frac{8\pi}{M_{\text{pl}}^2} h_{\alpha\beta} h h \partial^\gamma \partial_\gamma h^{\alpha\beta} - \frac{4\pi}{M_{\text{pl}}^2} h h \partial^\beta h^{\alpha\gamma} \partial_\gamma h_{\alpha\beta} + \frac{6\pi}{M_{\text{pl}}^2} h h \partial^\gamma h^{\alpha\beta} \partial_\gamma h_{\alpha\beta} \\ & - \frac{64\pi}{M_{\text{pl}}^2} h_{\alpha}{}^\gamma h_{\alpha\beta} \partial^\gamma h \partial^\delta h_{\beta}{}^\delta + \frac{32\pi}{M_{\text{pl}}^2} h h_{\beta\gamma} \partial^\gamma h \partial^\delta h_{\beta}{}^\delta - \frac{96\pi}{M_{\text{pl}}^2} h_{\alpha\beta} h_{\gamma\delta} \partial^\gamma h^{\alpha\epsilon} \partial^\delta h_{\epsilon}{}^\beta \\ & + \frac{64\pi}{M_{\text{pl}}^2} h_{\alpha\beta} h_{\gamma\delta} \partial^\beta h^{\alpha\epsilon} \partial^\delta h_{\epsilon}{}^\gamma - \frac{64\pi}{M_{\text{pl}}^2} h_{\alpha\beta} h_{\gamma\delta} \partial^\beta h^{\alpha\gamma} \partial^\delta h + \frac{32\pi}{M_{\text{pl}}^2} h_{\alpha\beta} h_{\gamma\delta} \partial^\gamma h^{\alpha\beta} \partial^\delta h \\ & - \frac{64\pi}{M_{\text{pl}}^2} h_{\alpha}{}^\gamma h_{\alpha\beta} \partial^\gamma h^{\beta\delta} \partial^\delta h + \frac{32\pi}{M_{\text{pl}}^2} h h_{\beta\gamma} \partial^\gamma h^{\beta\delta} \partial^\delta h - \frac{64\pi}{M_{\text{pl}}^2} h_{\alpha}{}^\gamma h_{\alpha\beta} h_{\delta}{}^\beta \partial^\delta \partial^\gamma h \end{aligned}$$

$$\begin{aligned}
& + \frac{32\pi}{M_{\text{pl}}^2} h h^\beta_\gamma h_{\beta\delta} \partial^\delta \partial^\gamma h + \frac{16\pi}{M_{\text{pl}}^2} h_{\alpha\beta} h^{\alpha\beta} h_{\gamma\delta} \partial^\delta \partial^\gamma h - \frac{16\pi}{M_{\text{pl}}^2} h^2 h_{\gamma\delta} \partial^\delta \partial^\gamma h \\
& + \frac{32\pi}{M_{\text{pl}}^2} h^\alpha_\gamma h_{\alpha\beta} \partial^\delta h \partial_\delta h^{\beta\gamma} - \frac{16\pi}{M_{\text{pl}}^2} h h_{\beta\gamma} \partial^\delta h \partial_\delta h^{\beta\gamma} - \frac{4\pi}{M_{\text{pl}}^2} h_{\alpha\beta} h^{\alpha\beta} \partial^\delta h \partial_\delta h \\
& + \frac{4\pi}{M_{\text{pl}}^2} h^2 \partial^\delta h \partial_\delta h + \frac{128\pi}{M_{\text{pl}}^2} h^\alpha_\gamma h_{\alpha\beta} h_{\delta\epsilon} \partial^\epsilon \partial^\gamma h^{\beta\delta} - \frac{32\pi}{M_{\text{pl}}^2} h h_{\beta\gamma} h_{\delta\epsilon} \partial^\epsilon \partial^\gamma h^{\beta\delta} \\
& - \frac{64\pi}{M_{\text{pl}}^2} h^\alpha_\gamma h_{\alpha\beta} h_{\delta\epsilon} \partial^\epsilon \partial^\delta h^{\beta\gamma} + \frac{32\pi}{M_{\text{pl}}^2} h h_{\beta\gamma} h_{\delta\epsilon} \partial^\epsilon \partial^\delta h^{\beta\gamma} + \frac{16\pi}{M_{\text{pl}}^2} h_{\alpha\beta} h^{\alpha\beta} h_{\gamma\delta} \partial^\delta \partial^\gamma h \\
& - \frac{16\pi}{M_{\text{pl}}^2} h^2 h_{\gamma\delta} \partial^\delta \partial^\gamma h - \frac{32\pi}{M_{\text{pl}}^2} h^\alpha_\gamma h_{\alpha\beta} h_{\delta\epsilon} \partial^\epsilon \partial^\gamma h^{\beta\delta} + \frac{16\pi}{M_{\text{pl}}^2} h h_{\beta\gamma} h_{\delta\epsilon} \partial^\epsilon \partial^\gamma h^{\beta\delta}, \quad (5.27)
\end{aligned}$$

$$\mathcal{L}_5^{(4)} = 0, \quad (5.28)$$

$$\mathcal{L}_6^{(4)} = 0. \quad (5.29)$$

### 5.2.1 Decoupling Limit

After performing the Stückelberg transformation, and upon examining the cubic Lagrangians

$\mathcal{L}_{2,\text{Sb}}^{(3)}$  and  $\mathcal{L}_{3,\text{Sb}}^{(3)}$ ,

$$\begin{aligned}
\mathcal{L}_{2,\text{Sb}}^{(3)} = & -\frac{4\sqrt{2\pi}}{mM_{\text{pl}}} h^{\rho\mu} \partial_\mu \partial_\nu \phi \partial^\nu \delta A_\rho + \frac{4\sqrt{2\pi}}{mM_{\text{pl}}} h^\nu_\rho \partial^\rho \partial_\mu \phi \partial_\nu \delta A^\mu + \frac{4\sqrt{2\pi}}{mM_{\text{pl}}} \partial^\rho \phi \partial^\nu \delta A^\mu \partial_\mu h_{\nu\rho} \\
& - \frac{4\sqrt{2\pi}}{mM_{\text{pl}}} \partial^\rho \phi \partial^\nu \delta A^\mu \partial_\nu h_{\mu\rho} + \frac{\sqrt{2\pi}}{M_{\text{pl}}} h \partial^\mu \delta A^\nu \partial_\nu \delta A_\mu - \frac{2\sqrt{2\pi}}{M_{\text{pl}}} h^{\mu\rho} \partial^\nu \delta A_\mu \partial_\nu \delta A_\rho \\
& + \frac{2\sqrt{2\pi}}{M_{\text{pl}}} h_{\nu\mu} \partial^\mu \delta A^\rho \partial^\nu \delta A_\rho - \frac{\sqrt{2\pi}}{M_{\text{pl}}} h \partial_\nu \delta A^\rho \partial^\nu \delta A_\rho + \frac{4\sqrt{2\pi}}{M_{\text{pl}}} \delta A^\mu \partial^\rho \delta A^\nu \partial_\nu h_{\mu\rho} \\
& - \frac{4\sqrt{2\pi}}{M_{\text{pl}}} \delta A^\mu \partial^\rho \delta A^\nu \partial_\rho h_{\mu\nu} + \frac{4\sqrt{2\pi}m}{M_{\text{pl}}} \delta A^\mu h_{\mu\nu} \partial^\nu \phi + \frac{2\sqrt{2\pi}}{M_{\text{pl}}} h_{\mu\nu} \partial^\mu \phi \partial^\nu \phi \\
& + \frac{2\sqrt{2\pi}m}{M_{\text{pl}}} \delta A^\mu h \partial_\mu \phi + \frac{\sqrt{2\pi}}{M_{\text{pl}}} h \partial_\mu \phi \partial^\mu \phi + \frac{2\sqrt{2\pi}m^2}{M_{\text{pl}}} \delta A^\mu \delta A^\nu h_{\mu\nu} + \frac{\sqrt{2\pi}m^2}{M_{\text{pl}}} \delta A_\mu \delta A^\mu h \\
\mathcal{L}_{3,\text{Sb}}^{(3)} = & \frac{J_3}{\Lambda^2} \delta A^\mu \partial_\mu \phi \square \phi + \frac{J_3}{2\Lambda^2} \partial_\mu \phi \partial^\mu \phi \partial_\nu \delta A^\nu + \frac{J_3 m^2}{2\Lambda^2} \delta A_\mu \delta A^\mu \partial_\nu \delta A^\nu \\
& + \frac{J_3 m}{2\Lambda^2} \delta A_\mu \delta A^\mu \square \phi + \frac{J_3 m}{\Lambda^2} \delta A^\mu \partial_\mu \phi \partial_\nu \delta A^\nu + \frac{J_3}{2\Lambda^2 m} \partial_\mu \phi \partial^\mu \phi \square \phi
\end{aligned} \quad (5.30)$$

one observes that the theory introduces interaction terms involving scalar, vector, and tensor modes that are controlled by distinct energy scales. These scales can be extracted by identifying the leading operators in the high-energy regime. The two relevant suppression

scales that emerge from this analysis are:

$$\tilde{\Lambda}_2 = (mM_{\text{pl}})^{1/2}, \quad \Lambda_3 = (\Lambda^2 m)^{1/3}. \quad (5.31)$$

To consistently isolate the physically relevant interactions at high energies, and in particular to suppress the mixing between scalar, vector, and tensor modes, we define a decoupling limit in which all redundant scales are sent to their appropriate asymptotic values while keeping the physically meaningful scales fixed. Specifically, we take:

$$m \rightarrow 0, \quad M_{\text{pl}} \rightarrow \infty, \quad \Lambda \rightarrow \infty, \quad \text{with} \quad \tilde{\Lambda}_2 = (mM_{\text{pl}})^{1/2}, \quad \Lambda_3 = (\Lambda^2 m)^{1/3} \quad \text{fixed.} \quad (5.32)$$

This definition ensures that the kinetic terms for all fields remain properly normalized and that the interactions among the different helicity components of the Proca field and the graviton persist in a controlled and finite way. While  $m \rightarrow 0$  removes the explicit mass scale and  $M_{\text{pl}} \rightarrow \infty$  decouples gravitational backreaction, the two derived scales  $\tilde{\Lambda}_2$  and  $\Lambda_3$  remain finite and define the regime where the effective field theory is still predictive.

Such scaling limits are common in multi-scale gravitational theories, such as bi-gravity, where similar techniques are employed to retain finite, non-trivial dynamics for certain modes while sending other scales to their asymptotic limits. In this context,  $\tilde{\Lambda}_2$  governs the mixing between the helicity-0 mode and gravity, while  $\Lambda_3$  controls the self-interactions of the Proca field, including the helicity-0 sector.

This decoupling limit serves as a useful theoretical tool to test the consistency and quantum stability of the theory, by isolating the dominant interaction terms that survive in the high-energy regime.

### 5.3 Quantum Stability

In this section, we summarize our findings regarding the quantum stability of the defined model. Rather than providing a comprehensive treatment of all possible loop contributions, we focus on the most relevant results and examples that illustrate the underlying mechanisms at play.

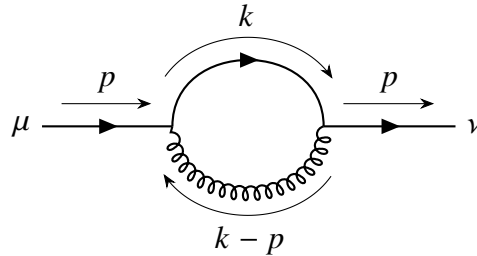
After performing a systematic power-counting analysis, we were able to filter out the majority of Feynman diagrams. These diagrams do not exhibit any signs of instability at high energies—specifically, they do not generate operators with dangerously large momentum dependence, nor do they induce unsuppressed interactions that could violate the validity of the effective theory.

For the remaining subset of diagrams—those which could not be excluded purely by power counting—we carried out explicit one-loop computations. These diagrams typically involve lower-dimensional operators where subtle cancellations or nontrivial momentum structures might arise. In particular, we analyzed selected two-, three-, and four-point functions to examine whether any divergent or non-decoupling terms survive in the ultraviolet limit.

In the following, we present a few representative examples of these computations, including their diagrammatic structure and the key steps of the analysis. These examples serve to illustrate the general strategy and confirm that all potentially dangerous contributions are either absent or under control in the decoupling limit.

### 5.3.1 Two-point Functions

As a concrete example, consider the following two-point diagram, which in the power-counting analysis appeared to have a dangerous momentum scaling of  $p^8$ ; however, the explicit computation shows that all such high-momentum contributions cancel, and the resulting divergence is much softer, consistent with the expected decoupling structure.



1. **Original UV-divergent part of the two-point function (highest momentum power):**

$$\frac{J_3^2 m^4}{48 \epsilon_{\text{UV}} \pi^2 \Lambda^4} \left[ -p^\mu p^\nu (23 m^2 + 3 p^2) + \eta^{\mu\nu} (-9 m^4 + 26 m^2 p^2 + 3 p^4) \right].$$

The term with the largest power of  $p$  is  $3 p^4 \eta^{\mu\nu}$ .

2. **Field content: two fields.** Being a two-point function, this divergence corresponds to an operator with two powers of the field  $A$ .
3. **Convert momenta to derivatives on the field:** In position space, each  $p^\mu$  becomes  $\partial^\mu$ . Consequently,  $p^4 \rightarrow \partial^4$ .

4. **Replace  $A^\mu$  by its Stueckelberg form for high-energy analysis:**

$$A^\mu = \frac{1}{m} \partial^\mu \phi.$$

Hence,

$$A^2 \longrightarrow \left(\frac{1}{m} \partial \phi\right)^2 = \frac{1}{m^2} (\partial \phi)^2.$$

5. **Combine steps:** The highest-momentum piece  $p^4$  in momentum space translates to

$$\partial^4(A^2) \longrightarrow \partial^4\left(\frac{1}{m^2}(\partial \phi)^2\right) = \frac{1}{m^2} \partial^4((\partial \phi)^2).$$

Putting back the overall prefactor  $\frac{m^4}{\Lambda^4}$ , we get

$$\sim \frac{m^4}{\Lambda^4} \frac{1}{m^2} \partial^4((\partial \phi)^2) = \frac{m^2}{\Lambda^4} \partial^4((\partial \phi)^2).$$

6. **Take the decoupling limit:**

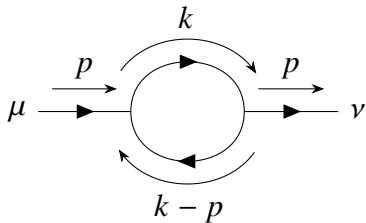
$$m \rightarrow 0, \quad M_{\text{pl}} \rightarrow \infty, \quad \Lambda \rightarrow \infty,$$

while  $\tilde{\Lambda}_2$  and  $\Lambda_3$  remain fixed (as previously defined). The factor

$$\frac{m^2}{\Lambda^4}$$

then controls whether this operator is suppressed or enhanced in that limit, depending on how  $\Lambda$  scales with  $m$ . In typical scenarios where  $\Lambda_3 \equiv (\Lambda^2 m)^{1/3}$  is held constant,  $m^2/\Lambda^4$  can vanish, thus suppressing the induced two-field divergence at high energies.

One can repeat the process for the other diagrams, for example:

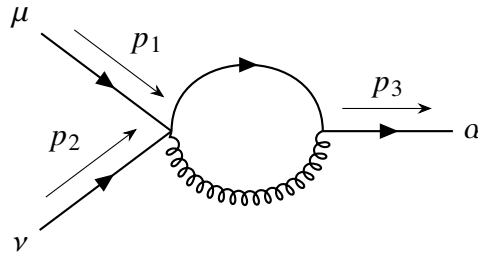


$$= \frac{1}{16 \epsilon_{\text{UV}} m^4 M_{\text{pl}}^2 \pi} \left[ -6m^4 \eta^{\mu\nu} (m^2 + p^2) + p^\mu p^\nu (24m^4 - 6m^2 p^2 + p^4) \right] \quad (5.33)$$



### 5.3.2 Three-point Functions

For the three-point functions, all diagrams turned out to be safe based on the power-counting analysis—none of them displayed dangerously large momentum scaling that could jeopardize the stability of the theory. However, to illustrate the procedure and confirm the expectations, we include a simple three-point diagram along with its explicit one-loop result as a representative example.

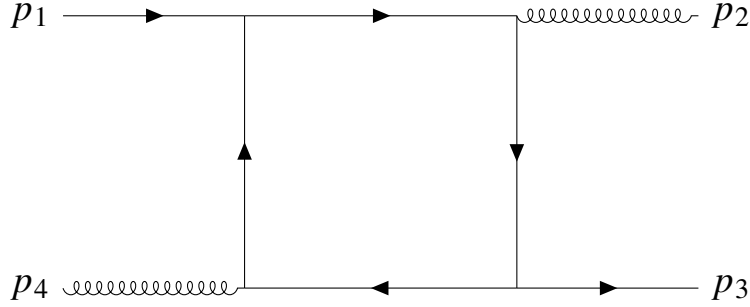


$$\begin{aligned}
 &= \frac{i J_3}{12 \epsilon_{\text{UV}} M_{\text{pl}}^2 \pi \Lambda^2} \left[ \eta^{\alpha\nu} \left( 5m^2 p_3^\mu (-2m^2 + p_3^2) + p_1^\mu (27m^4 + 12m^2 p_3^2 - p_3^4) \right) \right. \\
 &\quad + \eta^{\alpha\mu} \left( 5m^2 p_3^\nu (-2m^2 + p_3^2) + p_2^\nu (27m^4 + 12m^2 p_3^2 - p_3^4) \right) \\
 &\quad + p_3^\alpha \left( p_1^\mu p_3^\nu (-21m^2 + p_3^2) + \eta^{\mu\nu} (-22m^4 + 8m^2 p_3^2) \right. \\
 &\quad \left. \left. + p_3^\mu \left( -6m^2 p_3^\nu + p_2^\nu (-21m^2 + p_3^2) \right) \right) \right]. \tag{5.34}
 \end{aligned}$$

After taking the decoupling limit as defined in (5.32), the above result is further suppressed. This confirms that even in the loop-corrected theory, the relevant interaction remains under control at high energies, and no dangerous non-decoupling behavior emerges from the three-point sector.

### 5.3.3 Four-point Functions

For the four-point functions, almost all diagrams pass the power-counting filter and exhibit safe high-energy behavior. However, there remains one class of interactions that requires special attention—namely, the box-type diagrams shown below. These diagrams are not excluded by power counting and may potentially lead to unsuppressed operators in the UV.



The explicit loop computation and extraction of the divergent part for this class of diagrams is technically involved and remains a work in progress [40]. Nonetheless, preliminary indications suggest that the same suppression mechanisms observed in the lower-point functions may also play a role here. Moreover, we are very confident that the result of such computation will also place these diagrams on the safe side, consistent with the overall quantum stability of the theory.

## 5.4 Cutoff Estimation

In the Lagrangian setup presented in equation (5.18), the theory involves a general cutoff scale  $\Lambda$ , which parametrizes the suppression of higher-dimensional operators. A natural question arises regarding the validity and physical meaning of such a cutoff [41]: what is the energy range over which the effective field theory remains predictive and free of inconsistencies?

In this section, we address this question by analyzing the conditions under which the theory remains healthy in the decoupling limit. Our goal is to identify the constraints on the cutoff that ensure the stability of the Lagrangian, particularly in the presence of graviton interactions. We begin with the Lagrangians introduced in (5.18), but instead of using the dimensionless ratio  $(m/\Lambda)^2$ , we introduce a small parameter  $\lambda$  defined purely in terms of the mass parameter  $m$  and the Planck mass  $M_{\text{pl}}$ . This redefinition allows us to write

$$\lambda = \left( \frac{m}{M_{\text{pl}}} \right)^p, \quad (5.35)$$

and reduces the number of independent scales in the theory. Consequently, the Lagrangian depends only on the two physically meaningful mass parameters,  $m$  and  $M_{\text{pl}}$ , simplifying the structure of the effective field theory and clarifying the scaling behavior of different operators.

We will now show that consistency of the theory at the decoupling limit imposes constraints on the parameter  $p$ , and thereby on the original cutoff  $\Lambda$ . In particular, we will find that requiring healthy interactions and suppression of non-renormalizable terms restricts the allowed range of  $p$ .

$$\begin{aligned}\mathcal{L}_3^{(3)} = & \frac{J_3 m^{p-2}}{M_{\text{pl}}^p} \delta A^\mu \partial_\mu \phi \partial_\nu \partial^\nu \phi + \frac{1}{2} \frac{J_3 m^{p-2}}{M_{\text{pl}}^p} \partial_\mu \phi \partial^\mu \phi \partial^\nu \delta A_\nu \\ & + \frac{1}{2} \frac{J_3 m^{p-1}}{M_{\text{pl}}^p} \delta A_\mu \delta A^\mu \partial_\nu \partial^\nu \phi + \frac{J_3 m^{p-1}}{M_{\text{pl}}^p} \delta A^\mu \partial_\mu \phi \partial^\nu \delta A_\nu \\ & + \frac{1}{2} J_3 \delta A_\mu \delta A^\mu \left( \frac{m}{M_{\text{pl}}} \right)^p \partial^\nu \delta A_\nu + \frac{1}{2} \frac{J_3 m^{p-3}}{M_{\text{pl}}^p} \partial_\mu \phi \partial^\mu \phi \partial_\nu \partial^\nu \phi\end{aligned}\quad (5.36)$$

Looking at the structure of  $\mathcal{L}_3$  at the cubic order, we have:

$$\begin{aligned}\mathcal{L}_3^{(3)} = & \frac{J_3 m^{p-2}}{M_{\text{pl}}^p} \delta A^\mu \partial_\mu \phi \partial_\nu \partial^\nu \phi + \frac{1}{2} \frac{J_3 m^{p-2}}{M_{\text{pl}}^p} \partial_\mu \phi \partial^\mu \phi \partial^\nu \delta A_\nu \\ & + \frac{1}{2} \frac{J_3 m^{p-1}}{M_{\text{pl}}^p} \delta A_\mu \delta A^\mu \partial_\nu \partial^\nu \phi + \frac{J_3 m^{p-1}}{M_{\text{pl}}^p} \delta A^\mu \partial_\mu \phi \partial^\nu \delta A_\nu \\ & + \frac{1}{2} J_3 \delta A_\mu \delta A^\mu \left( \frac{m}{M_{\text{pl}}} \right)^p \partial^\nu \delta A_\nu + \frac{1}{2} \frac{J_3 m^{p-3}}{M_{\text{pl}}^p} \partial_\mu \phi \partial^\mu \phi \partial_\nu \partial^\nu \phi.\end{aligned}\quad (5.37)$$

Among all the interaction terms, the most dangerous one—i.e., the one that becomes problematic most rapidly as we approach the decoupling limit  $m \rightarrow 0$ —is the last term. This term contains the most negative power of  $m$ , and hence dominates in the low-mass limit. Specifically, it reads

$$\frac{1}{2} J_3 \frac{m^{p-3}}{M_{\text{pl}}^p} \partial_\mu \phi \partial^\mu \phi \partial_\nu \partial^\nu \phi. \quad (5.38)$$

Since we require  $p > 0$  for the small parameter  $\lambda$  to remain well-defined, we see that the power of  $m$  in the numerator becomes increasingly negative as  $p$  decreases.

Based on the scaling behavior of this term, we define the corresponding decoupling scale  $\Lambda_3$ , which characterizes the energy at which interactions become strongly coupled. This scale is given by

$$\Lambda_3 = \left( m^{3-p} M_{\text{pl}}^p \right)^{1/3}. \quad (5.39)$$

$$m \rightarrow 0, \quad M_{\text{pl}} \rightarrow \infty, \quad \text{with} \quad \Lambda_3 = \left( m^{3-p} M_{\text{pl}}^p \right)^{1/3} = \text{const.} \quad (5.40)$$

Taking the decoupling limit, the cubic Lagrangian  $\mathcal{L}_3^{(3)}$  reduces to a single term. In this limit, it becomes

$$\mathcal{L}_3^{(3)}|_{\text{DL}} = \frac{J_3}{2\Lambda_3^3} \partial_\mu \phi \partial^\mu \phi \partial_\nu \partial^\nu \phi. \quad (5.41)$$

As expected, the only remaining contribution at this scale comes from the pure scalar mode  $\phi$ . All other interactions involving the vector fluctuation  $\delta A_\mu$  are suppressed by additional powers of the mass parameter  $m$ , and hence vanish in the decoupling limit.

We now return to the Lagrangian  $\mathcal{L}_2^{(3)}$ . Using the cutoff scale  $\Lambda_3$  with its expression in terms of  $m$  and  $M_{\text{pl}}$ ,

$$\begin{aligned} \mathcal{L}_2^{(3)} = & -\frac{4\sqrt{2\pi} m^{3/p-2}}{\Lambda_3^{3/p}} h^{\rho\mu} \partial_\mu \partial_\nu \phi \partial^\nu \delta A_\rho + \frac{4\sqrt{2\pi} m^{3/p-2}}{\Lambda_3^{3/p}} h^\nu{}_\rho \partial^\rho \partial_\mu \phi \partial_\nu \delta A^\mu \\ & + \frac{4\sqrt{2\pi} m^{3/p-2}}{\Lambda_3^{3/p}} \partial^\rho \phi \partial^\nu \delta A^\mu \partial_\mu h_{\nu\rho} - \frac{4\sqrt{2\pi} m^{3/p-2}}{\Lambda_3^{3/p}} \partial^\rho \phi \partial^\nu \delta A^\mu \partial_\nu h_{\mu\rho} \\ & + \frac{\sqrt{2\pi} m^{3/p-1}}{\Lambda_3^{3/p}} h \partial^\mu \delta A^\nu \partial_\nu \delta A_\mu - \frac{2\sqrt{2\pi} m^{3/p-1}}{\Lambda_3^{3/p}} h^{\mu\rho} \partial^\nu \delta A_\mu \partial_\nu \delta A_\rho \\ & + \frac{2\sqrt{2\pi} m^{3/p-1}}{\Lambda_3^{3/p}} h_{\nu\mu} \partial^\mu \delta A^\rho \partial^\nu \delta A_\rho - \frac{\sqrt{2\pi} m^{3/p-1}}{\Lambda_3^{3/p}} h \partial_\nu \delta A^\rho \partial^\nu \delta A_\rho \\ & + \frac{4\sqrt{2\pi} m^{3/p-1}}{\Lambda_3^{3/p}} \delta A^\mu \partial^\rho \delta A^\nu \partial_\nu h_{\mu\rho} - \frac{4\sqrt{2\pi} m^{3/p-1}}{\Lambda_3^{3/p}} \delta A^\mu \partial^\rho \delta A^\nu \partial_\rho h_{\mu\nu} \\ & + \frac{4\sqrt{2\pi} m^{3/p}}{\Lambda_3^{3/p}} \delta A^\mu h_{\mu\nu} \partial^\nu \phi + \frac{2\sqrt{2\pi} m^{3/p-1}}{\Lambda_3^{3/p}} h_{\mu\nu} \partial^\mu \phi \partial^\nu \phi \\ & + \frac{2\sqrt{2\pi} m^{3/p}}{\Lambda_3^{3/p}} \delta A^\mu h \partial_\mu \phi + \frac{\sqrt{2\pi} m^{3/p-1}}{\Lambda_3^{3/p}} h \partial_\mu \phi \partial^\mu \phi \\ & + \frac{2\sqrt{2\pi} m^{3/p+1}}{\Lambda_3^{3/p}} \delta A^\mu \delta A^\nu h_{\mu\nu} + \frac{\sqrt{2\pi} m^{3/p+1}}{\Lambda_3^{3/p}} \delta A_\mu \delta A^\mu h, \end{aligned} \quad (5.42)$$

we observe that some non-trivial factors of  $m$  could appear in the denominators of the interaction terms. These modify the scaling of the graviton couplings and, in particular, can introduce inverse powers of  $m$  that potentially grow large in the decoupling limit  $m \rightarrow 0$ . To avoid a destructive strong coupling behavior and ensure that all graviton-mediated interactions remain under control, we must also bound the dangerous terms in  $\mathcal{L}_2^{(3)}$ . This leads to the constraint

$$0 < p < \frac{3}{2}. \quad (5.43)$$

This range ensures that no graviton interaction becomes strongly coupled before the cutoff scale is reached, preserving the consistency of the effective field theory in the low-energy limit.

It is important to emphasize that the boundary value  $p = \frac{3}{2}$  must be excluded from the allowed range. In this case, the decoupling of modes in  $\mathcal{L}_2^{(3)}$  fails to occur, and we are left with residual interactions between the scalar, vector, and tensor degrees of freedom. Specifically, at  $p = \frac{3}{2}$ , the Lagrangian becomes

$$\mathcal{L}_2^{(3)}\big|_{p=\frac{3}{2}} = \frac{4\sqrt{2}\pi}{\Lambda_3^2} \left( \partial^\rho \phi \partial^\nu \delta A^\mu \partial_\mu h_{\nu\rho} - \partial^\rho \phi \partial^\nu \delta A^\mu \partial_\nu h_{\mu\rho} + h_\nu{}^\sigma \partial_\sigma \partial_\mu \phi \partial^\nu \delta A^\mu - h^{\rho\sigma} \partial_\sigma \partial_\nu \phi \partial^\nu \delta A_\rho \right). \quad (5.44)$$

This residual interaction implies that, contrary to naive expectations, the vector and scalar modes do not fully decouple at this critical value of  $p$ . This observation corrects the conclusion drawn in [42], where the upper bound  $p = \frac{3}{2}$  was included in the allowed range and it was incorrectly claimed that all interactions vanish in the decoupling limit. Our analysis shows that true decoupling only occurs for

$$0 < p < \frac{3}{2}, \quad (5.45)$$

with strict inequality at the upper boundary.

One can see the connection between different values of  $p$  and the corresponding cutoff scale  $\Lambda$ . Below is a summary for several representative choices:

$p$	$\Lambda$	$\Lambda_3$
$\frac{1}{2}$	$m^{3/4} M_{\text{pl}}^{1/4}$	$(m^5 M_{\text{pl}})^{1/6}$
1	$m^{1/2} M_{\text{pl}}^{1/2}$	$(m^2 M_{\text{pl}})^{1/3}$
$\frac{4}{3}$	$m^{1/2} M_{\text{pl}}^{2/3}$	$(m^5 M_{\text{pl}}^4)^{1/9}$
$\frac{22}{15}$	$m^{4/15} M_{\text{pl}}^{11/15}$	$(m^{23} M_{\text{pl}}^{22})^{1/45}$
$\frac{3}{2}$	$m^{1/4} M_{\text{pl}}^{3/4}$	$(m M_{\text{pl}})^{1/2}$

**Table 5.1.** Examples of cutoff scales  $\Lambda$  and  $\Lambda_3$  for various values of  $p$ .

For values of  $p$  within the bound established above, the quantum stability of the Lagrangians (5.18) has been explicitly verified at the level of two-, three-, and four-point functions at one loop. Fortunately, no detailed loop computations were required. As we showed earlier using a power-counting analysis, all loop-induced terms that are allowed by symmetry are either sufficiently suppressed by the cutoff scale, or involve structures whose potentially dangerous high-energy behavior is absent due to the specific tensor contractions and derivative structure of the theory.

This result confirms that the effective theory remains consistent and perturbatively stable in the decoupling limit, provided the parameter  $p$  lies within the range

$$0 < p < \frac{3}{2}. \quad (5.46)$$

## Chapter 6

# Propagator Structure and SVT Decomposition in the Presence of Background Fields

In this chapter, we revisit our initial framework through an alternative approach. Our primary objective is to eliminate all mixing terms at the quadratic level, which we achieve through the scalar-vector-tensor (SVT) decomposition. This decomposition systematically separates the perturbative modes of our system, ensuring their complete decoupling at the quadratic level. The power of this approach lies in its ability to isolate and analyze different dynamical components independently. However, this mathematical transformation introduces two significant considerations. First, we must work with an expanded set of degrees of freedom, some of which are purely gauge artifacts rather than physical modes. Second, the formulation necessarily breaks manifest covariance, requiring an explicit 3+1 decomposition where spatial and temporal derivatives are treated distinctly. While this sacrifices some geometric elegance, it provides a powerful computational framework for analyzing the dynamics of our system.

### 6.1 Background Equations of Motion

We begin by setting up the field configurations for our analysis:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \tag{6.1}$$

$$A^\mu = \bar{A}^\mu + \delta A^\mu, \tag{6.2}$$

where  $\eta_{\mu\nu}$  is the Minkowski spacetime metric,  $\bar{A}^\mu$  represents the background value of the vector field, and  $\delta A^\mu$  denotes its perturbations. The metric perturbations are given by  $h_{\mu\nu}$ . Here and throughout the text, we use the notation  $G_{i,X}$  to denote the derivative of the function  $G_i$  with respect to its first argument  $X = \frac{1}{2}A_\mu A^\mu$ , while  $G_{i,F}$  represents the derivative with respect to the second argument  $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$ , where  $F_{\mu\nu}$  is the field strength tensor. We work in the perturbative regime where these fluctuations are small:

$$|h_{\mu\nu}| \ll 1, \quad |\delta A^\mu| \ll 1, \quad (6.3)$$

For our analysis, we consider a constant background vector field  $\bar{A}^\mu$ . The background equations of motion, derived from variations with respect to the metric, take the form:

$$-\frac{1}{2}\bar{G}_2\eta_{\mu\nu} + \frac{1}{2}\bar{A}_\mu\bar{A}_\nu\bar{G}_{2,X} = 0, \quad (6.4)$$

while variations with respect to the vector field yield:

$$\bar{A}^\mu\bar{G}_{2,X} = 0. \quad (6.5)$$

For a constant Minkowski background with a constant background vector field, these equations admit two possible solutions [43]:

Case A:  $\bar{G}_2 = 0, \bar{G}_{2,X} \neq 0$  with  $\bar{A}^\mu = 0$

Case B:  $\bar{G}_2 = 0, \bar{G}_{2,X} = 0$  with  $\bar{A}^\mu = \text{constant}$

In our subsequent analysis, we focus on Case B, which allows for a non-zero background vector field. This choice is compatible with our setup and can be realized through a Mexican hat type potential within the  $G_2$  term of our theory.

## 6.2 Mixing Terms from a Non-Vanishing Background

In the previous setup, we considered the case where the background vector field vanishes,  $\bar{A}_\mu = 0$ . Under this condition, the quadratic action simplifies considerably, and the propagators for the Proca field and the graviton become diagonal. This allows one to treat the kinetic and mass terms of the spin-1 and spin-2 perturbations independently at leading order.

However, once a non-trivial background  $\bar{A}_\mu \neq 0$  is introduced, the structure of the theory changes significantly. In general, the quadratic action contains mixing terms between the vector and tensor perturbations, and the resulting kinetic matrix is no longer diagonal. The mixing quadratic Lagrangian is given by:



$$\begin{aligned}
\mathcal{L}_{\text{mix}}^{(2)} = & \bar{A}^\mu \bar{G}_2^{(0,1)} \partial^\nu \delta A^\rho \partial_\rho h_{\mu\nu} - \bar{A}^\mu \bar{G}_2^{(0,1)} \partial^\nu \delta A^\rho \partial_\nu h_{\mu\rho} \\
& + \frac{1}{2} \bar{A}^\mu \bar{A}^\nu \delta A_\mu \bar{G}'_3 \partial_\nu h^\rho{}_\rho - \frac{1}{2} \bar{A}^\mu \bar{A}^\nu \delta A^\rho \bar{G}'_3 \partial_\rho h_{\mu\nu} \\
& - \frac{2}{\kappa^2} \bar{A}^\mu \delta A_\mu \bar{G}'_4 \partial^\rho \partial^\nu h_{\nu\rho} + \frac{2}{\kappa^2} \bar{A}^\mu \delta A_\mu \bar{G}'_4 \partial^\rho \partial_\rho h^\nu{}_\nu \\
& + \bar{A}^\mu \bar{G}'_4 \partial^\nu \delta A_\nu \partial_\mu h^\rho{}_\rho - \bar{A}^\mu \bar{G}'_4 \partial^\rho \delta A^\nu \partial_\mu h_{\nu\rho} \\
& - \bar{A}^\mu \bar{G}'_4 \partial^\rho \delta A^\nu \partial^\nu h_{\mu\rho} + \bar{A}^\mu \bar{G}'_4 \partial^\rho \delta A^\nu \partial_\rho h_{\mu\nu} \\
& + \frac{1}{2} \bar{A}^\mu \bar{A}^\nu \bar{A}^\rho \delta A_\mu \bar{G}_2^{(2,0)} h_{\nu\rho}.
\end{aligned} \tag{6.6}$$

For a generic background configuration, this term does not vanish and leads to non-diagonal elements in the inverse propagator. As a result, the propagator structure becomes non-trivial. Schematically, the quadratic fluctuations of the vector and tensor fields can be represented by the following structure:

$$\mathcal{L}^{(2)} \sim \begin{pmatrix} \delta A & h \end{pmatrix} \begin{pmatrix} O_{AA} & O_{Ah} \\ O_{hA} & O_{hh} \end{pmatrix} \begin{pmatrix} \delta A \\ h \end{pmatrix} \tag{6.7}$$

The corresponding propagator structure can be schematically represented as a matrix  $\mathbf{D}(p)$  of the form:

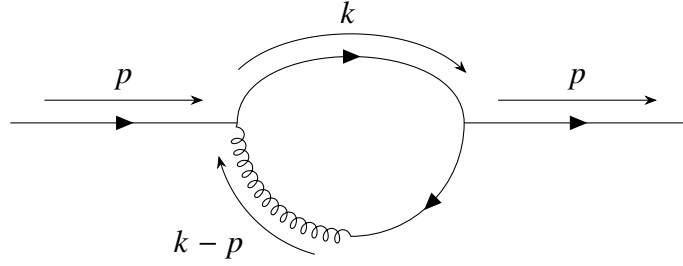
$$\mathbf{D}(p) = \begin{pmatrix} D_{11}^{\mu\nu}(p) & D_{12}^{\mu|\alpha\beta}(p) \\ D_{21}^{\alpha\beta|\mu}(p) & D_{22}^{\rho\sigma|\alpha\beta}(p) \end{pmatrix} \tag{6.8}$$

Diagrammatically, the off-diagonal component  $D_{12}^{\mu|\alpha\beta}(p)$ , representing the mixing between the Proca and graviton fields, can be illustrated as follows:

$$\begin{array}{ccc}
\mu & \xrightarrow{p} & \alpha\beta \\
\longrightarrow & \text{---} & \text{---}
\end{array}$$

Although the presence of mixing terms suggests that the kinetic operator is not diagonal in general, this does not obstruct the perturbative computation of physical quantities. In particular, one can still proceed with standard loop calculations, provided that all relevant diagrammatic contributions — including those involving mixed propagators — are taken into account.

While the mixing increases the number of diagrams and introduces additional tensor structures, the usual Feynman diagram techniques remain valid. For instance, in the case of one-loop bubble diagrams contributing to the two-point function, we now encounter new diagrams involving mixed internal lines that were absent in the diagonal case. One such example is illustrated below:



Notice the different and additional vertex structures that can now appear at each end of the diagram. These arise from the mixing terms and lead to an increased number of possible one-loop contributions that must be considered in the computation.

An alternative — though not very conventional in standard quantum field theory setups — is to deal with the non-diagonal quadratic terms by decomposing the fields into their scalar, vector, and tensor components. This mode decomposition can sometimes simplify the analysis by making the physical degrees of freedom more transparent. We will explore this approach in the next section.

### 6.3 Scalar-Vector-Tensor Decomposition

An alternative approach — though not necessarily simpler — is to work directly with the physical degrees of freedom by decomposing the perturbations into scalar, vector, and tensor components. This scalar-vector-tensor (SVT) decomposition [44, 45] does not eliminate the complexities of the system, but it can offer a different perspective, especially in the presence of symmetries such as isotropy, where distinct mode sectors often decouple.

Indeed, as we will see shortly, the quadratic action written in SVT variables often leads to a decoupling of scalar, vector, and tensor modes. In such cases, the propagators for individual modes become diagonal within each sector, which can make certain aspects of the analysis more transparent.

To set the stage, we begin by decomposing the metric perturbation as follows:

$$\begin{aligned} h_{00} &= -2\phi, \\ h_{0i} &= B_i^T + \partial_i B, \\ h_{ij} &= \frac{1}{3}\delta_{ij}\psi + E_{ij}^{TT} + \partial_{(i}E_{j)}^T + \left(\partial_i\partial_j - \frac{1}{3}\delta_{ij}\partial^2\right)E, \end{aligned} \tag{6.9}$$

Similarly, the perturbation of the vector field can be decomposed as:

$$\begin{aligned} \delta A_0 &= -\varphi, \\ \delta A_i &= v_i^T + \partial_i\omega, \end{aligned} \tag{6.10}$$

These decompositions are subject to the following constraints:

$$\partial^i B_i^T = 0, \quad \partial^i E_{ij}^{TT} = 0, \quad \delta^{ij} E_{ij}^{TT} = 0, \quad \partial^i v_i^T = 0, \quad (6.11)$$

These conditions ensure that vector modes are transverse and tensor modes are both transverse and traceless, leading to a clean separation between different types of perturbations.

Altogether, the decomposition yields ten degrees of freedom from the metric perturbation in Eq. (6.9), and four from the vector field perturbation in Eq. (6.10). Diffeomorphism invariance grants us gauge freedom, which allows for the transformation of any field  $T = \bar{T} + \delta T$  under the Lie derivative along a vector  $\xi^\mu$  as:

$$\delta T \rightarrow \delta T - \mathcal{L}_\xi \bar{T}. \quad (6.12)$$

Under these transformations, the metric components transform as:

$$\phi \rightarrow \phi - \dot{\xi}_0, \quad (6.13a)$$

$$B_i^T \rightarrow B_i^T - \dot{\xi}_i^T, \quad (6.13b)$$

$$B \rightarrow B - \xi_0 - \dot{\xi}, \quad (6.13c)$$

$$\psi \rightarrow \psi - 2\nabla^2 \xi, \quad (6.13d)$$

$$E \rightarrow E - 2\xi, \quad (6.13e)$$

$$E_i^T \rightarrow E_i^T - 2\xi_i^T, \quad (6.13f)$$

$$E_{ij}^{TT} \rightarrow E_{ij}^{TT}, \quad (6.13g)$$

where the gauge parameter  $\xi_\mu$  decomposes as:

$$\xi_\mu = \left( \xi_0, \xi_i^T + \partial_i \xi \right). \quad (6.14)$$

For the vector field perturbations, given a constant background field  $\bar{A}_\mu$ , the transformations are:

$$\delta A_0 \rightarrow \delta A_0 - \bar{A}^\nu \partial_\nu \xi_0, \quad (6.15a)$$

$$\delta A_i \rightarrow \delta A_i - \bar{A}^\nu \partial_\nu \xi_i^T - \bar{A}^\nu \partial_i \xi. \quad (6.15b)$$

For an isotropic background that respects the  $SO(3)$  symmetry, the background configuration is constrained as:

$$A^\mu = (A_0, 0, 0, 0). \quad (6.16)$$

With this consideration, the gauge transformations in (6.15) simplify to:

$$\varphi \rightarrow \varphi - A_0 \dot{\xi}_0, \quad (6.17a)$$

$$v_i^T \rightarrow v_i^T - A_0 \dot{\xi}_i^T, \quad (6.17b)$$

$$\omega \rightarrow \omega - A_0 \dot{\xi}. \quad (6.17c)$$

These transformations enable the reduction of the degrees of freedom from 14 to 10 physical ones. Rather than fixing a specific gauge, we construct gauge-invariant combinations that capture the physical degrees of freedom:

$$\Phi = \phi + \dot{B} - \frac{1}{2} \dot{E}, \quad (6.18a)$$

$$\Theta = \frac{1}{3} (-\nabla^2 E + \psi), \quad (6.18b)$$

$$\beta = \varphi + \frac{1}{2} A_0 \ddot{E} - A_0 \dot{B}, \quad (6.18c)$$

$$\gamma = \omega + \frac{1}{2} A_0 \dot{E}, \quad (6.18d)$$

$$\Gamma_i = v_i + A_0 \dot{E}_i, \quad (6.18e)$$

$$\Xi_i = B_i - \dot{E}_i. \quad (6.18f)$$

With four scalars, two transverse vectors, and one traceless-transverse tensor  $E_{ij}$ , we have successfully identified our ten gauge-invariant degrees of freedom. To analyze which of these variables represent dynamical degrees of freedom, we expand the Lagrangian to quadratic order and derive the equations of motion. These equations reveal the true physical content of our theory and determine which variables propagate dynamically. The equations

of motion at first order are:

$$\begin{aligned}
& 2G_4 (\partial_\nu \partial_\mu h - \partial_\rho \partial_\mu h_{\nu}{}^\rho - \partial_\rho \partial_\nu h_{\mu}{}^\rho + \partial_\rho \partial^\rho h_{\mu\nu} + \eta_{\mu\nu} (\partial^\alpha \partial_\rho h_{\alpha}{}^\rho - \partial^\alpha \partial_\alpha h)) \\
& + 2\bar{A}^\rho \bar{A}^\alpha \eta_{\mu\nu} \partial_\alpha \delta A_\rho \bar{G}_{3,X} + \bar{A}^\rho \bar{A}^\alpha \bar{A}^\beta \eta_{\mu\nu} \partial_\beta h_{\rho\alpha} \bar{G}_{3,X} - 4\bar{A}^\rho \partial_\nu \partial_\mu \delta A_\rho \bar{G}_{4,X} \\
& - 2\bar{A}^\rho \bar{A}^\alpha \partial_\nu \partial_\mu h_{\rho\alpha} \bar{G}_{4,X} + 2\bar{A}^\rho \partial_\rho \partial_\mu \delta A_\nu \bar{G}_{4,X} + 2\bar{A}_\nu \partial_\rho \partial_\mu \delta A^\rho \bar{G}_{4,X} \\
& + 2\bar{A}^\rho \partial_\rho \partial_\nu \delta A_\mu \bar{G}_{4,X} - 2\bar{A}_\nu \partial_\rho \partial^\rho \delta A_\mu \bar{G}_{4,X} + 2\bar{A}_\nu \bar{A}^\rho \partial_\alpha \partial_\mu h_{\rho}{}^\alpha \bar{G}_{4,X} \\
& - 4\bar{A}^\rho \eta_{\mu\nu} \partial_\alpha \partial_\rho \delta A^\alpha \bar{G}_{4,X} + 2\bar{A}^\rho \bar{A}^\alpha \partial_\alpha \partial_\rho h_{\mu\nu} \bar{G}_{4,X} - 2\bar{A}^\rho \bar{A}^\alpha \eta_{\mu\nu} \partial_\alpha \partial_\rho h \bar{G}_{4,X} \\
& + 4\bar{A}^\rho \eta_{\mu\nu} \partial_\alpha \partial^\alpha \delta A_\rho \bar{G}_{4,X} - 2\bar{A}_\nu \bar{A}^\rho \partial_\alpha \partial^\alpha h_{\mu\rho} \bar{G}_{4,X} + 2\bar{A}^\rho \bar{A}^\alpha \eta_{\mu\nu} \partial_\beta \partial^\beta h_{\rho\alpha} \bar{G}_{4,X} \\
& + 2\bar{A}_\nu (-\partial_\rho \partial_\mu \delta A^\rho + \partial_\rho \partial^\rho \delta A_\mu + \bar{A}^\rho (-\partial_\alpha \partial_\mu h_{\rho}{}^\alpha + \partial_\alpha \partial^\alpha h_{\mu\rho})) \bar{G}_{2,FX} \\
& + \bar{A}_\mu \left( 2 \left( \partial_\rho \partial_\nu \delta A^\rho - \partial_\rho \partial^\rho \delta A_\nu + \bar{A}^\rho \left( \partial_\alpha \partial_\nu h_{\rho}{}^\alpha - \partial_\alpha \partial^\alpha h_{\nu\rho} \right) \right) (\bar{G}_{4,X} - \bar{G}_{2,FX}) \right. \\
& \left. + \bar{A}_\nu (- (2\partial_\rho \delta A^\rho + \bar{A}^\rho \partial_\rho h) \bar{G}_{3,X} + 2 (-\partial_\alpha \partial_\rho h^{\rho\alpha} + \partial_\alpha \partial^\alpha h) \bar{G}_{4,X} + \bar{A}^\rho (2\delta A_\rho + \bar{A}^\alpha h_{\rho\alpha}) \bar{G}_{2,XX}) \right) = 0,
\end{aligned} \tag{6.19}$$

For the variation with respect to  $h^{\mu\nu}$ , and

$$\begin{aligned}
& \frac{1}{2} \bar{A}^\nu \left( 2\bar{G}_{2,F} (\partial^\rho \partial_\rho h_{\mu\nu} - \partial^\rho \partial_\mu h_{\nu\rho}) + \bar{G}_{3,X} (\bar{A}^\rho \partial_\mu h_{\nu\rho} - \bar{A}_\mu \partial^\nu h + 2\partial_\mu \delta A_\nu) \right. \\
& \quad \left. + 2\bar{G}_{4,X} (-\partial_\nu \partial_\mu h + \partial^\rho \partial_\mu h_{\nu\rho} + \partial^\rho \partial_\nu h_{\mu\rho} - \partial^\rho \partial_\rho h_{\mu\nu}) + \bar{A}_\mu \bar{G}_{2,XX} (\bar{A}^\rho h_{\nu\rho} + 2\delta A_\nu) \right) \\
& \quad + \bar{A}_\mu (\bar{G}_{4,X} (\partial^\rho \partial_\rho h - \partial^\rho \partial^\nu h_{\nu\rho}) - \bar{G}_{3,X} \partial_\nu \delta A^\nu) + \bar{G}_{2,F} (\partial^\nu \partial_\nu \delta A_\mu - \partial^\nu \partial_\mu \delta A_\nu) = 0,
\end{aligned} \tag{6.20}$$

for the variation with respect to  $\delta A^\mu$ . We analyze these equations by decomposing them into scalar, vector, and tensor sectors, which due to our SVT decomposition remain decoupled at the quadratic level.

## 6.4 Gauge Choice: The Uniform Vector Gauge

Gauge-invariant variables are extremely useful for identifying the true physical degrees of freedom in a system with gauge redundancy. They allow one to isolate the combinations of fields that remain unaffected by coordinate transformations or internal symmetries, thus providing a clear understanding of what propagates and what does not.

However, for computational purposes, especially when dealing with perturbative expansions and higher-order interactions, it is often advantageous to fix a gauge. A suitable gauge choice can significantly simplify the structure of the equations and reduce the number of terms that need to be computed explicitly.

For this reason, we work in the *uniform vector gauge*, which is defined by the following decomposition of the metric and vector field perturbations:

$$\begin{aligned} h_{00} &= -2\phi, \\ h_{0j} &= \partial_j B + B_j, \\ h_{ij} &= E_{ij} + 2\psi \delta_{ij}, \end{aligned} \tag{6.21}$$

$$\begin{aligned} \delta A_0 &= -\varphi, \\ \delta A_j &= v_j - \bar{A}_0 \partial_j B, \end{aligned} \tag{6.22}$$

In this gauge, the scalar, vector, and tensor degrees of freedom are organized in a way that facilitates both canonical analysis and perturbative expansion. The tensor perturbations, which are gauge-invariant by construction, remain unaffected by this choice and their quadratic action reads:

$$\mathcal{L}_{\text{tensor}}^{(2)} = -\frac{M_{\text{Pl}}^2}{64\pi} \dot{E}_{ab} \dot{E}^{ab} + \frac{M_{\text{Pl}}^2}{64\pi} \partial^c E_{ab} \partial_c E^{ab}. \tag{6.23}$$

In this gauge, the quadratic Lagrangians for the different sectors take the following form.

The vector sector reads:

$$\begin{aligned} \mathcal{L}_{\text{vector}}^{(2)} &= \frac{1}{2} \bar{A}_0^2 \dot{B}^a \dot{B}_a + \bar{A}_0 \dot{B}^a \dot{v}_a + \frac{1}{2} \dot{v}^a \dot{v}_a \\ &+ \frac{M_{\text{Pl}}^2}{8\pi} B^a \partial^b \partial_b B_a + \frac{1}{2} \bar{A}_0^2 \partial^a B^b \partial_b B_a - \frac{M_{\text{Pl}}^2}{32\pi} \partial^a B^b \partial_b B_a \\ &+ \bar{A}_0 \partial^a v^b \partial_b B_a - \frac{1}{2} \bar{A}_0^2 \partial^b B^a \partial_b B_a + \frac{3M_{\text{Pl}}^2}{32\pi} \partial^b B^a \partial_b B_a \\ &- \bar{A}_0 \partial^b v^a \partial_b B_a + \frac{1}{2} \partial^a v^b \partial_b v_a - \frac{1}{2} \partial^b v^a \partial_b v_a \end{aligned} \tag{6.24}$$

The scalar sector is given by:

$$\begin{aligned} \mathcal{L}_{\text{scalar}}^{(2)} &= -\frac{3M_{\text{Pl}}^2}{8\pi} \psi \ddot{\psi} + \bar{A}_0^2 m^2 \varphi^2 + 2\bar{A}_0^3 m^2 \varphi \phi + \bar{A}_0^4 m^2 \phi^2 \\ &- \frac{M_{\text{Pl}}^2}{4\pi} \dot{\psi} \partial^a \partial_a B + \frac{M_{\text{Pl}}^2}{8\pi} \psi \partial^a \partial_a \psi + \frac{M_{\text{Pl}}^2}{4\pi} \phi \partial^a \partial_a \psi \\ &+ \frac{1}{2} \partial^a \varphi \partial_a \varphi + 2\bar{A}_0 \partial^a \phi \partial_a \varphi + 2\bar{A}_0^2 \partial^a \phi \partial_a \phi \end{aligned} \tag{6.25}$$

At first sight, the quadratic Lagrangians in this gauge still exhibit mixing between various fields, particularly in the scalar and vector sectors. However, not all fields appearing in

the Lagrangians are dynamical — several of them act as auxiliary variables whose equations of motion are purely constraint equations. By solving these constraint equations and substituting back into the action, we can eliminate the non-dynamical fields.

The resulting quadratic Lagrangian, expressed entirely in terms of the physical degrees of freedom, takes the form [44]:

$$\begin{aligned} \mathcal{L}_{\text{quadratic}} = & -\frac{M_{\text{Pl}}^4}{32\pi^2 \bar{A}_0^2} \partial^a \psi \partial_a \psi + \frac{M_{\text{Pl}}^2}{64\pi} \partial^c E_{ab} \partial_c E^{ab} - \frac{1}{2} \partial^b \Gamma^a \partial_b \Gamma_a \\ & - \frac{M_{\text{Pl}}^2}{8\pi} \partial^a \psi \partial_a \psi - \frac{M_{\text{Pl}}^2}{64\pi} \dot{E}_{ab} \dot{E}^{ab} + \frac{1}{2} \dot{\Gamma}^a \dot{\Gamma}_a \\ & + \frac{\Lambda^4 M_{\text{Pl}}^4}{16\pi^2 \bar{A}_0^2 J_3^2 m^2} \dot{\psi}^2 - \frac{3M_{\text{Pl}}^2}{8\pi} \dot{\psi}^2, \end{aligned} \quad (6.26)$$

where we have defined the gauge-invariant combination

$$\Gamma_a = v_a + \bar{A}_0 B_a. \quad (6.27)$$

This Lagrangian clearly shows that the dynamical degrees of freedom at quadratic order consist of:

- One transverse-traceless tensor mode  $E_{ab}$ ,
- One transverse vector mode  $\Gamma_a$ ,
- One scalar mode  $\psi$ .

Each of these propagates independently, with canonically normalized kinetic terms up to overall prefactors, confirming the consistency of the decomposition and the gauge choice.

### 6.4.1 Perturbative Elimination of Non-Dynamical Fields

To eliminate the non-dynamical fields such as  $\phi$ ,  $\varphi$ ,  $B$ , and  $B_a$ , we must proceed perturbatively. This is because the equations of motion for these fields become non-linear when higher-order interactions are included — particularly due to their appearance in cubic and quartic terms in the action. As a result, we expand these fields in powers of a small perturbation parameter  $\alpha$ , which organizes the computation order by order:

$$\phi = \alpha \phi^{(1)} + \alpha^2 \phi^{(2)} + \alpha^3 \phi^{(3)} + \dots \quad (6.28)$$

$$\varphi = \alpha \varphi^{(1)} + \alpha^2 \varphi^{(2)} + \alpha^3 \varphi^{(3)} + \dots \quad (6.29)$$

$$B = \alpha B^{(1)} + \alpha^2 B^{(2)} + \alpha^3 B^{(3)} + \dots \quad (6.30)$$

$$B_a = \alpha B_a^{(1)} + \alpha^2 B_a^{(2)} + \alpha^3 B_a^{(3)} + \dots \quad (6.31)$$

Here,  $\alpha$  is a bookkeeping parameter that keeps track of the perturbative order. The crucial point is that to compute the Lagrangian at a given order  $O(\alpha^n)$ , it is sufficient to solve the constraint equations for the non-dynamical fields up to order  $\alpha^{n-1}$ . For instance, at quadratic order, only the first-order components  $\phi^{(1)}$ ,  $\varphi^{(1)}$ ,  $B^{(1)}$ , and  $B_a^{(1)}$  are needed. However, when analyzing the cubic Lagrangian, the second-order solutions  $\phi^{(2)}$ ,  $\varphi^{(2)}$ , etc., also contribute and must be included.

This perturbative strategy ensures consistency of the expansion and allows us to iteratively eliminate non-dynamical fields while maintaining control over the order of the approximation. It is particularly useful in theories with derivative interactions, where constraints become increasingly intricate at higher orders.

To illustrate this procedure explicitly, let us consider the second-order equation for the metric vector perturbation  $B_x^{(2)}$ . At this order, the equation of motion includes both linear and non-linear terms in the dynamical fields, such as  $\psi$ ,  $\Gamma_a$ , and  $E_{ab}$ , as well as their derivatives. Solving this equation allows us to express  $B_x^{(2)}$  in terms of the dynamical variables alone.

The full equation of motion at  $O(\alpha^2)$  reads

$$\begin{aligned} 0 = & -\frac{\Lambda^2 M_{\text{Pl}}^6}{128 \bar{A}_0^5 J_3 m^2 \pi^3} \partial^a \partial_a \psi \partial_x \psi - \frac{\Lambda^4 M_{\text{Pl}}^6}{128 \bar{A}_0^6 J_3^2 m^4 \pi^3} \partial^a \partial_a \psi \partial_x \dot{\psi} + \frac{\Lambda^2 M_{\text{Pl}}^6}{128 \bar{A}_0^5 J_3 m^2 \pi^3} \partial^a \psi \partial_x \partial_a \psi \\ & + \frac{\Lambda^4 M_{\text{Pl}}^6}{128 \bar{A}_0^6 J_3^2 m^4 \pi^3} \partial^a \dot{\psi} \partial_x \partial_a \psi + \frac{M_{\text{Pl}}^4}{64 \bar{A}_0^2 \pi^2} E_x^a \partial_a \psi + \frac{\Lambda^2 M_{\text{Pl}}^4}{64 \bar{A}_0^3 J_3 m^2 \pi^2} E_x^a \partial_a \dot{\psi} \\ & + \frac{\Lambda^2 M_{\text{Pl}}^4}{64 \bar{A}_0^3 J_3 m^2 \pi^2} \partial^a \psi \partial^b \partial_b E_{xa} + \frac{\Lambda^2 M_{\text{Pl}}^4}{32 \bar{A}_0^3 J_3 m^2 \pi^2} E_{xa} \partial^b \partial_b \partial^a \psi - \frac{\Lambda^2 M_{\text{Pl}}^4}{32 \bar{A}_0^3 J_3 m^2 \pi^2} E_{xb} \partial^b \partial^a \partial_a \psi \\ & + \frac{\Lambda^2 M_{\text{Pl}}^4}{16 \bar{A}_0^3 J_3 m^2 \pi^2} \partial^a E_{xb} \partial^b \partial_a \psi - \frac{\Lambda^2 M_{\text{Pl}}^4}{16 \bar{A}_0^3 J_3 m^2 \pi^2} \partial^b E_{xa} \partial^b \partial^a \psi + \frac{J_3 m^2 M_{\text{Pl}}^4}{32 \bar{A}_0 \pi^2 \Lambda^2} \psi \partial_x \psi \end{aligned} \quad (6.32)$$

where we are solving for the second-order metric vector perturbation component  $B_x^{(2)}$  in this equation.

Importantly,  $B_x^{(2)}$  appears under a Laplacian operator, specifically as  $\partial^a \partial_a B_x^{(2)}$ , which implies that in order to isolate it, one must apply the inverse Laplacian. This leads to a formally non-local solution of the form

$$B_x^{(2)} = \nabla^{-2} (\dots), \quad (6.33)$$



This non-locality [46] reflects the constrained nature of the system, where certain metric components are not independent but are determined by elliptic-type constraint equations. Such structures are common in gravitational theories when working in a fixed gauge, and they do not hinder consistency as long as the inversion is well-defined within the perturbative regime.

## Implications for Loop Computations

From the structure of the reduced Lagrangian, we can extract the propagators for the physical degrees of freedom. These propagators exhibit modified dispersion relations, where the propagation speed may differ from unity due to interactions with the background field. For example [47], a scalar field propagator in this setup takes the form

$$\text{-----} \xrightarrow{q} \text{-----} = \frac{-i}{-q_0^2 + c_s^2 \mathbf{q}^2 - i\epsilon} = \frac{-i}{\bar{q}^2 - i\epsilon}, \quad (6.34)$$

where we have defined the rescaled four-momentum

$$\bar{q}^\mu \equiv (q_0, c_s \mathbf{q}),$$

and  $c_s$  denotes the effective propagation speed of the mode.

As a result, loop integrals of the type

$$\int \frac{d^d q}{(2\pi)^d}$$

must be treated with care. In particular, it is often necessary to separate time and spatial components in the integrals, performing the  $q_0$  and  $\mathbf{q}$  integrations independently. This becomes essential when dealing with loop diagrams involving non-relativistic dispersion relations or anisotropic propagation speeds. These features are a hallmark of theories with background structure, such as the one considered here.

## Summary and Outlook

The goal of this chapter was to develop tools for analyzing loop-level quantum corrections in theories with a non-vanishing vector background. In such setups, the presence of background structure generically prevents the diagonalization of the propagator at quadratic order, leading to mixing between different field components. To tackle this, we proposed and compared two complementary approaches.

The first method was to work directly with the non-diagonal propagator structure. This approach preserves the covariant form of the theory, allowing us to utilize the standard machinery of covariant loop computations. However, the price to pay is an increase in the number of Feynman diagrams, due to the presence of off-diagonal propagators. All possible contractions and mixed propagator insertions must be carefully included in the loop expansion. Despite the added combinatorial complexity, the calculations remain well-defined within the covariant formalism.

The second method involved performing a scalar-vector-tensor (SVT) decomposition, separating the perturbations into their irreducible components [48]. In this formulation, the quadratic Lagrangian becomes block-diagonal, with no mixing between different sectors. As a result, the propagators are diagonal, greatly simplifying the identification of the propagating degrees of freedom and the analysis of stability. However, this comes at the cost of breaking manifest covariance. The decomposition explicitly separates time and spatial components, which complicates the structure of interaction terms and integrals. Moreover, going to higher orders in perturbation theory introduces non-localities, particularly when solving constraint equations. These non-localities can be handled in momentum space using Fourier transforms, but the resulting loop integrals are non-covariant and, in general, more difficult to evaluate. While such integrals can be computed in simple settings, a general and systematic method for evaluating them is still lacking.

In summary, both approaches have distinct advantages and limitations. The covariant method maintains formal structure but requires careful bookkeeping of diagrammatic contributions. The SVT-based method simplifies propagators but sacrifices covariance and introduces new technical challenges at higher orders.

## **Part II**

# **Multi-Method Classification of Gravitational Wave Observables**



## Chapter 7

# Supervised Learning and Neural Networks in Physics

Machine learning (ML) has rapidly become an important tool across many areas of physics, enabling data-driven discovery in fields from high-energy particle collisions to astrophysical observations [49]. In particular, supervised learning – where models are trained on labeled data – has seen widespread success in tackling classification and regression problems in physics. In a supervised learning task, we have a dataset of input examples  $X$  (such as experimental measurements or detector signals) each paired with an output label  $Y$  (such as a particle identity or signal/noise tag). The goal is to learn a function  $f : X \rightarrow Y$  that can predict the correct output for new, unseen inputs. Supervised ML techniques have been used, for instance, to distinguish rare physics events from background noise, identify phases of matter, and recognize complex patterns in detector data. Neural networks have become especially prominent supervised learning models in physics due to their flexibility and power. A neural network (NN) can approximate highly complex, nonlinear relationships, making it ideal for analyzing the enormous volumes of high-dimensional data common in modern experiments [49, 50]. High-energy physics (HEP) offers a prime example: proton collisions at the Large Hadron Collider produce terabytes of data per second, and physicists must sift out a few interesting events (like Higgs boson decays) from overwhelming backgrounds. Early applications of neural networks in HEP date back to the 1990s, improving particle identification and event selection in collider experiments [51]. More recently, deep neural networks (with many layers) have outperformed previous methods on challenging HEP tasks. For example, a deep network achieved about 8% better classification accuracy than the best earlier approaches in distinguishing exotic particle signals from background,

all without the need for manual feature engineering [49]. This boost in performance can significantly increase the discovery potential for new particles. Another exciting domain is gravitational-wave astronomy. The detection of gravitational waves – ripples in spacetime from violent astrophysical events – traditionally relies on matched filtering, which involves correlating detector data with dozens of thousands of theoretical waveform templates. This is computationally intensive and can become a bottleneck for real-time detection as the network of detectors and the event rate grow. Deep learning offers a powerful alternative: convolutional neural networks (CNNs) have been developed to scan noisy time-series data for wave signatures, learning directly from simulated waveforms. Neural networks have been applied to real LIGO/Virgo data, showing they can identify true signals and adapt to nonstationary detector noise [52, 53, 54]. These successes underscore the relevance of neural networks in both high-energy physics and gravitational wave analysis. In this chapter, we explore the foundations of supervised learning and neural networks, discuss how they are trained and optimized, and delve into the special case of convolutional neural networks – including why they are well-suited to structured data like images and waveforms.

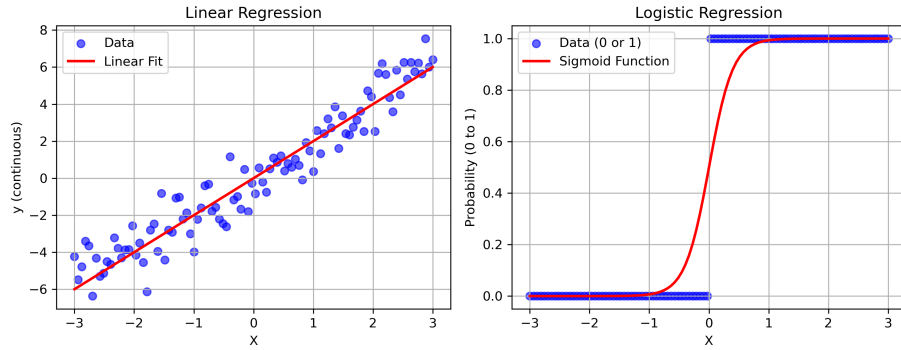
## 7.1 Mathematical Foundations of Supervised Learning

We begin by formalizing supervised learning problems. Regression and classification are two primary categories. In regression, the goal is to predict a continuous output  $y$  from an input  $\mathbf{x}$ . For example, given a detector signal, we might predict a physical parameter (like the energy of a particle or the mass of an astrophysical object). In classification, the aim is to assign inputs to one of several discrete categories. An example is identifying whether an event is "signal" or "background", or classifying particle tracks as electron vs. muon vs. pion. Formally, we have a training dataset  $(\mathbf{x}^{(i)}, y^{(i)})_{i=1}^N$  with inputs  $\mathbf{x}^{(i)} \in \mathbb{R}^d$  and labels  $y^{(i)}$  (either continuous values for regression or class indices for classification). We seek a function (model)  $f(\mathbf{x}; \theta)$ , parameterized by  $\theta$ , that maps an input to a prediction  $\hat{y}$ , such that  $\hat{y} \approx y$  for future examples. The parameters  $\theta$  are adjusted during training to minimize a measure of error on the training set, called the cost function or loss function.

For regression tasks, a common choice of loss is the mean squared error (MSE):

$$J(\theta) = \frac{1}{N} \sum_{i=1}^N \left( \hat{y}^{(i)} - y^{(i)} \right)^2, \quad (7.1)$$

where  $\hat{y}^{(i)} = f(\mathbf{x}^{(i)}; \theta)$  is the model's prediction.  $J(\theta)$  is essentially the average squared deviation between predictions and true targets [55]. It is a convenient, differentiable proxy



**Figure 7.1.** Comparison of Linear Regression and Logistic Regression. The left panel illustrates linear regression, where a continuous line is fitted to the data points, allowing predictions over a continuous range. In contrast, the right panel depicts logistic regression, where the sigmoid function maps input values to probabilities between 0 and 1. This enables binary classification by establishing a decision boundary at a probability threshold (typically 0.5), distinguishing between two classes.

for measuring accuracy in predicting continuous quantities, and its smooth quadratic form often leads to simple analytical properties (e.g. the gradient is proportional to the error). For classification, especially binary classification (two classes), a prevalent loss is the cross-entropy associated with logistic regression. If we denote the model's predicted probability for class 1 as  $\hat{p}^{(i)} = f(\mathbf{x}^{(i)}; \theta)$  (and thus  $1 - \hat{p}^{(i)}$  for class 0), and the true label  $y^{(i)} \in \{0, 1\}$ , the binary cross-entropy loss for a single example is:

$$\ell(\theta; \mathbf{x}^{(i)}, y^{(i)}) = - \left[ y^{(i)} \log \hat{p}^{(i)} + (1 - y^{(i)}) \log (1 - \hat{p}^{(i)}) \right]. \quad (7.2)$$

The total cost is the average cross-entropy over all training examples:

$$J(\theta) = \frac{1}{N} \sum_i \ell(\theta; \mathbf{x}^{(i)}, y^{(i)}). \quad (7.3)$$

Cross-entropy originates from information theory and measures the dissimilarity between the true distribution and the predicted distribution [56]. Minimizing cross-entropy is equivalent to maximizing the likelihood of the data under a Bernoulli model in this case. This loss heavily penalizes confident but wrong predictions – if the true label  $y^{(i)} = 1$  but the predicted  $\hat{p}^{(i)}$  is close to 0, the  $-\log(\hat{p}^{(i)})$  term is huge. Thus, the model is encouraged not only to be correct, but confidently correct. Cross-entropy is the preferred loss for classification problems in neural networks because it works seamlessly with probabilistic outputs (such as those produced by a softmax, discussed in Section 6) and tends to yield faster

convergence than alternatives like MSE in classification contexts. In multi-class classification with  $K$  classes, this generalizes to the multinomial cross-entropy (or categorical cross-entropy): if the network outputs a vector of predicted probabilities  $(\hat{p}_1, \hat{p}_2, \dots, \hat{p}_K)$  for the  $K$  classes, and  $y^{(i)}$  is represented as a one-hot vector (all zeros except a 1 at the true class index), then:

$$J(\theta) = -\frac{1}{N} \sum_{i=1}^N \sum_{k=1}^K y_k^{(i)} \log \hat{p}_k^{(i)} \quad (7.4)$$

which again penalizes any deviation of the predicted probability for the true class away from 1. These cost functions provide a quantitative target for the learning algorithm: finding parameters  $\theta$  that minimize  $J(\theta)$ . In most realistic ML problems, especially with neural networks, this minimization cannot be done analytically; we must resort to iterative numerical optimization. Before turning to optimization, it's worth noting that these cost choices are not arbitrary. MSE corresponds to maximum likelihood under a Gaussian noise model, and cross-entropy corresponds to maximum likelihood for a Bernoulli or multinomial model. Thus, the choice of cost function often reflects assumptions about the data (Gaussian-distributed regression noise, or classification viewed as probabilistic inference). When those assumptions hold, these losses lead to efficient learning of the underlying patterns.

## 7.2 Optimization and Training

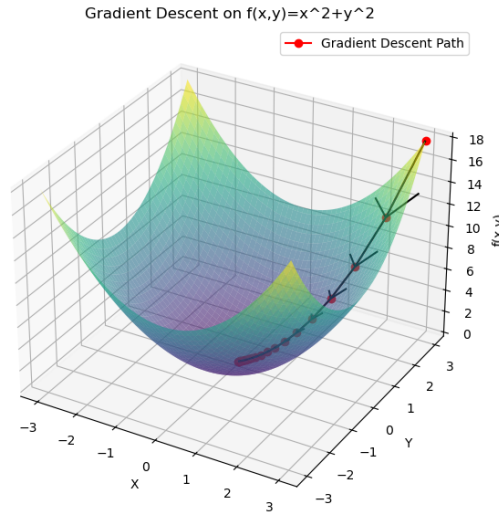
Training a supervised model means adjusting its parameters to minimize the cost function. The main algorithm for this is gradient descent [57], due to its simplicity and effectiveness in high-dimensional parameter spaces. In batch gradient descent, we compute the gradient of  $J(\theta)$  with respect to all parameters  $\theta = (\theta_1, \theta_2, \dots)$ , and update the parameters in the opposite direction of the gradient (since the gradient points toward increasing cost). The update rule is:

$$\theta \leftarrow \theta - \alpha \nabla_{\theta} J(\theta) \quad (7.5)$$

where  $\alpha$  is the learning rate, a small positive scalar that controls the step size. Conceptually, one can imagine  $J(\theta)$  as defining a high-dimensional surface; gradient descent moves the parameters “downhill” on this surface in small steps until hopefully reaching a minimum (ideally the global minimum, though in practice, we often settle for a good local minimum).



For example, if  $J$  is the MSE,  $\nabla_{\theta}J$  will be proportional to the residual errors between predictions and targets, and the update reduces those errors.



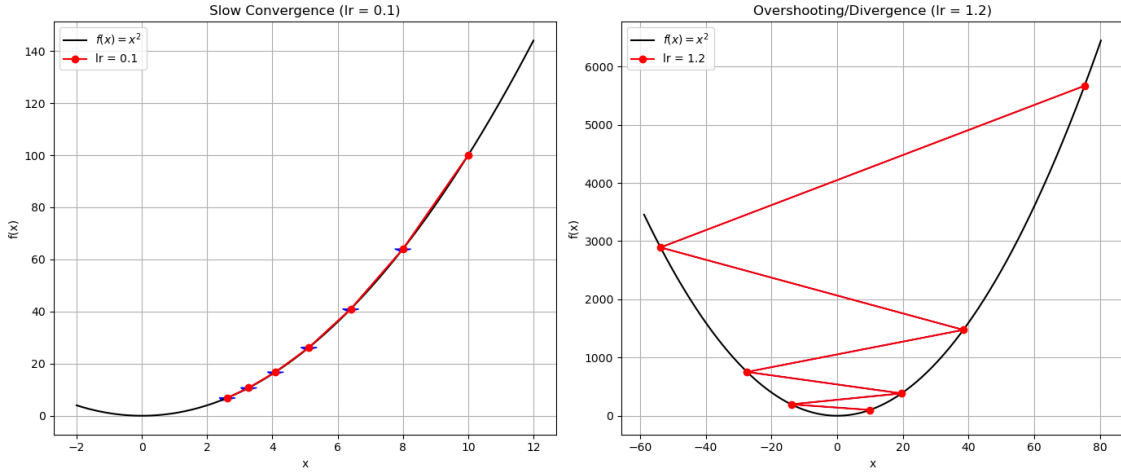
**Figure 7.2.** Visualization of gradient descent on  $f(x, y) = x^2 + y^2$ , showing the iterative path toward the minimum.

Choosing an appropriate learning rate  $\alpha$  is crucial: too large  $\alpha$  can overshoot minima and cause divergence, while too small  $\alpha$  leads to painfully slow convergence. In practice, especially for large datasets, stochastic gradient descent (SGD) and its variants are used [55, 56]. In SGD, rather than computing the gradient on the entire dataset (which can be very costly for millions of examples), we approximate the gradient using a mini-batch of training examples. For instance, we might use 32 or 128 randomly selected examples at each iteration to compute an approximate gradient. The update then becomes  $\theta \leftarrow \theta - \alpha \nabla_{\theta} J_{batch}(\theta)$ , where  $J_{batch}$  is the cost computed on that mini-batch. Using mini-batches introduces noise into the updates but drastically reduces computation per update and often helps escape shallow local minima, acting as a form of regularization. We iterate over many mini-batches (often cycling through the dataset multiple times, which are called epochs) until the cost stops decreasing significantly. Gradient descent variants improve basic SGD by adapting the learning rate during training or by incorporating momentum. One simple improvement is to add a momentum term [56]. Here we maintain a velocity vector  $v$  (of same dimension

as  $\theta$ ) that accumulates a decaying history of past gradients:

$$\begin{aligned} v &\leftarrow \beta v + (1 - \beta) \nabla_{\theta} J(\theta), \\ \theta &\leftarrow \theta - \alpha v, \end{aligned} \quad (7.6)$$

where  $0 < \beta < 1$  (commonly  $\beta \approx 0.9$ ). Momentum can thereby lead to faster and more stable convergence.



**Figure 7.3.** Comparison of gradient descent trajectories for the function  $f(x) = x^2$  with different learning rates. The left plot demonstrates slow convergence with a small learning rate, while the right plot shows overshooting and divergence with a large learning rate.

Perhaps the most widely used optimizer in deep learning is Adam (Adaptive Moment Estimation) [58]. Adam combines ideas from momentum and adaptive learning rates. It keeps an exponentially decaying average of past gradients  $m_t$  (first moment) and of past squared gradients  $v_t$  (second moment), and updates them at each time step  $t$ :

$$\begin{aligned} m_t &= \beta_1 m_{t-1} + (1 - \beta_1) \nabla_{\theta} J(\theta)_t \\ v_t &= \beta_2 v_{t-1} + (1 - \beta_2) [\nabla_{\theta} J(\theta)_t]^2, \end{aligned} \quad (7.7)$$

with decay rates  $\beta_1, \beta_2$  typically around 0.9 and 0.999. Adam then uses bias-corrected estimates of these moments (to account for their initialization at zero) and updates the parameters as:

$$\theta \leftarrow \theta - \alpha \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon}, \quad (7.8)$$

where  $\hat{m}_t = m_t / (1 - \beta_1^t)$  and  $\hat{v}_t = v_t / (1 - \beta_2^t)$ , and  $\epsilon$  is a small constant (like  $10^{-8}$ ) to prevent division by zero. Intuitively, Adam scales each parameter's learning rate by an adaptive

factor: if a parameter's gradient has been large and volatile,  $v_t$  will be large and thus the effective learning rate for that parameter is reduced; conversely, parameters with consistently small gradients get a relatively larger step. This per-parameter adaptation, along with momentum on the first moment, makes Adam quite robust and fast-converging for a variety of problems. Indeed, Adam has been successfully used in training deep networks for physics applications ranging from jet tagging to gravitational wave signal recognition. Other popular optimizers include RMSprop and Adagrad, which also adapt learning rates, but Adam often gives the best of both worlds and is the default choice in many frameworks. We will use Adam in our later examples unless otherwise stated. Even with an effective optimization algorithm, one must be cautious of **overfitting**, a phenomenon where a model learns not only the general trends in the training data but also its noise and specific features. Overfitting severely limits a model's ability to generalize to new, unseen data. In physics, this is particularly problematic since the goal is not merely to memorize a dataset but to extract underlying physical laws or signals.

To mitigate overfitting, **regularization techniques** [55, 56] are essential in supervised learning. One of the simplest and most widely used regularization methods is **weight decay**, which discourages excessively large model parameters by adding a penalty term to the loss function.

In the case of **L2 regularization** (also known as ridge regression), a quadratic penalty is added to the cost function:

$$J(\theta) = J_0(\theta) + \frac{\lambda}{2} \sum_j \theta_j^2, \quad (7.9)$$

where  $J_0(\theta)$  is the original cost function,  $\lambda$  is a hyperparameter that controls the strength of the penalty, and  $\theta_j$  are the model parameters. This regularization term effectively **pulls the weights toward zero**, preventing them from growing too large. In deep networks, this has the effect of smoothing the function learned by the model, as large weights tend to correspond to highly oscillatory functions that fit noise rather than true patterns.

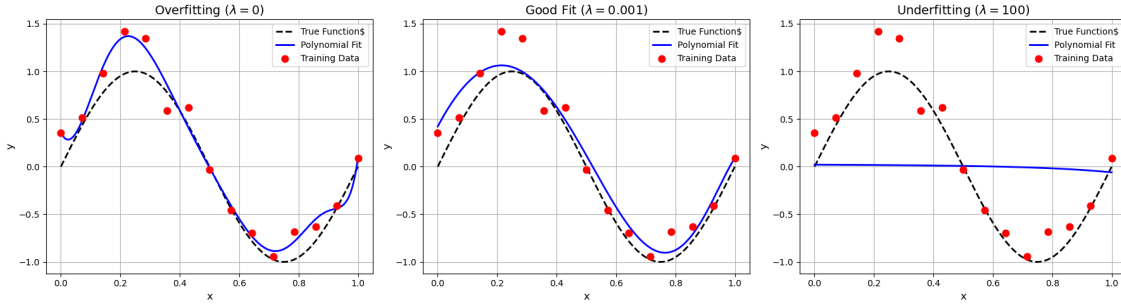
Alternatively, **L1 regularization** (lasso) imposes an absolute-value penalty:

$$J(\theta) = J_0(\theta) + \lambda \sum_j |\theta_j|. \quad (7.10)$$

Unlike L2 regularization, which shrinks weights continuously, L1 regularization promotes sparsity, meaning that many weights are driven exactly to zero. This can be useful in settings where one expects only a subset of features to be relevant. However, in deep neural

networks, L2 regularization is generally preferred since it allows for a smoother reduction of weights during optimization.

For convolutional neural networks (CNNs), L2 regularization plays an important role in ensuring that learned filters focus on meaningful features rather than noise. Another



**Figure 7.4.** Comparison of polynomial fits (degree 9) via ridge regression for different values of the regularization parameter  $\lambda$ . The left panel ( $\lambda = 0$ ) overfits the noisy training data, the center panel ( $\lambda = 0.001$ ) achieves a good balance by closely following the true function and data, while the right panel ( $\lambda = 100$ ) underfits due to excessive regularization.

useful technique is **dropout** [59], which randomly disables a fraction of neurons during training to prevent over-reliance on specific paths in the network. This enforces redundancy in representations and reduces overfitting, ensuring that the model learns robust and generalizable features.

Beyond weight decay and dropout, several other regularization strategies can help improve generalization:

- **Early stopping:** By monitoring validation performance, training can be halted once the model starts overfitting, preventing unnecessary complexity.
- **Data augmentation:** By introducing physics-informed transformations or noise to the training data, the model can learn to be invariant to irrelevant variations, improving robustness.

By carefully incorporating appropriate regularization techniques, we ensure that our neural networks generalize well, which is critical for making reliable physics predictions. The combination of effective optimization algorithms and well-chosen regularization methods allows us to extract meaningful physical insights from data while minimizing the risk of overfitting.

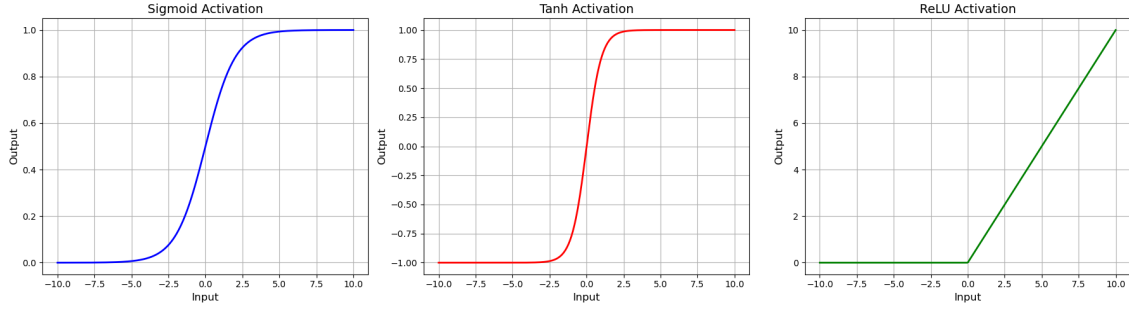
## 7.3 Building Neural Networks

A feedforward neural network (also called a multilayer perceptron, MLP) is organized into layers of nodes (neurons) with directed connections from each layer to the next [60]. The first layer (input layer) consists of neurons representing the input features, and the final layer (output layer) produces the predictions. In between are one or more hidden layers which give the network its flexibility to learn complex functions. Each neuron in a hidden or output layer computes a weighted sum of its inputs and then applies a nonlinear activation function. Mathematically, if  $\mathbf{f}^{(l-1)}$  is the vector of activations from layer  $(l - 1)$ , then the activations of layer  $l$  (a vector  $\mathbf{f}^{(l)}$ ) are:

$$\mathbf{f}^{(l)} = g^l \left( W^{(l)} \mathbf{h}^{(l-1)} + \mathbf{b}^{(l)} \right), \quad (7.11)$$

where  $W^{(l)}$  is a weight matrix and  $\mathbf{b}^{(l)}$  a bias vector for layer  $l$ , and  $g(\cdot)$  is the activation function applied elementwise. The activation  $g$  introduces nonlinearity; without it, the whole network would collapse into a single linear transformation no matter how many layers we stack. Common choices of  $g$  include [55]:

- Sigmoid ( $g(x) = 1/(1 + e^{-x})$ ): squashes the input into the range  $[0, 1]$ . Historically used in early networks, but now less common (except perhaps in the output layer for binary probability outputs) because of issues like vanishing gradients.
- Hyperbolic tangent ( $\tanh(x)$ ): ranges from  $-1$  to  $1$ , zero-centered, often yields better training than sigmoid for hidden layers, but still can saturate for large  $|x|$ .
- ReLU (Rectified Linear Unit,  $\text{ReLU}(x) = \max(0, x)$ ): outputs 0 for negative inputs and linear (identity) for positive inputs. ReLU and its variants are now standard for hidden layers in deep networks because they mitigate the vanishing gradient problem and promote sparse activations. Physically, a ReLU neuron can be thought of as a feature detector that either is inactive (0) or linearly active if a certain threshold is exceeded. This simplicity leads to easier optimization in practice.
- Softmax, used in output layer for multi-class classification, which we detail in Section 6. Other activations like leaky ReLU, ELU, and GELU have been introduced, but the above are sufficient for most of our discussion. We often use ReLUs in hidden layers and an appropriate activation for the output depending on the task (e.g., identity for regression, sigmoid for binary classification, softmax for multi-class).

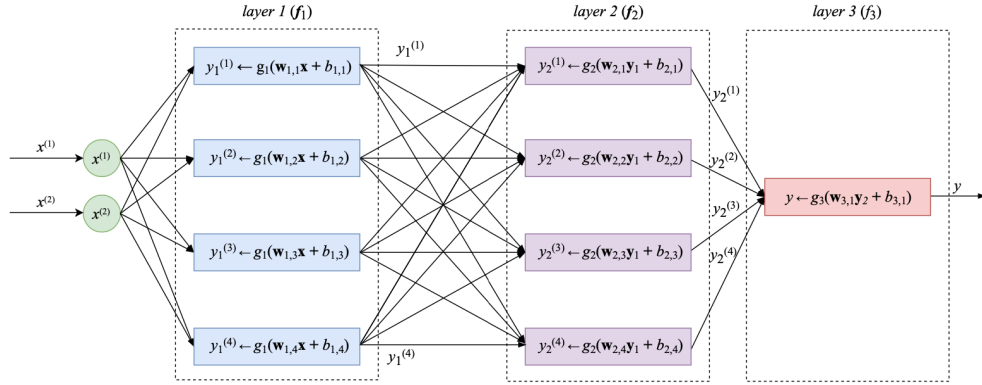


**Figure 7.5.** Schematic representations of three common activation functions in machine learning: sigmoid, tanh, and ReLU. The ReLU function introduces nonlinearity by outputting zero for negative input values (i.e., “turning off” the neuron), which is crucial for enabling networks to learn complex patterns.

Let’s illustrate a simple feedforward network architecture. Suppose we are classifying events as signal vs. background based on some features (detector measurements, etc.). A basic network might have an input layer with  $d$  neurons (one per feature), one or two hidden layers (say with  $m$  and  $n$  neurons, respectively), and an output layer with a single neuron outputting  $\hat{y}$  (the predicted probability of “signal”). The computations would be:

- Layer 1:  $\mathbf{f}^{(1)} = g_1(W^{(1)}\mathbf{x} + \mathbf{b}^{(1)})$ , yielding an  $m$ -dimensional vector.
- Layer 2:  $\mathbf{f}^{(2)} = g_2(W^{(2)}\mathbf{f}^{(1)} + \mathbf{b}^{(2)})$ , yielding an  $n$ -dimensional vector.
- Output:  $\hat{y} = g_{\text{out}}(W^{(3)}\mathbf{f}^{(2)} + b^{(3)})$ , a scalar.

Here  $g_1, g_2$  might be ReLU, and  $g_{\text{out}}$  could be sigmoid for a binary classification. The network’s overall function is a nested composition of linear maps and nonlinearities:  $\hat{y} = g_{\text{out}}(W^{(3)}g_2(W^{(2)}g_1(W^{(1)}\mathbf{x} + b^{(1)}) + b^{(2)}) + b^{(3)})$ . Although this looks complex, it is just a series of matrix multiplications and element-wise operations - easily implemented on a computer. The expressive power of neural networks comes from stacking many such layers: even a single hidden layer network can approximate any continuous function on  $\mathbb{R}^d$  to arbitrary accuracy (the universal approximation theorem) [56], given enough neurons. Deeper networks (multiple hidden layers) can do so more efficiently, reusing and combining low-level features to build high-level ones. In HEP, for example, one can imagine the first hidden layer learning to pick out local energy deposit patterns in a calorimeter, the second layer combining those to identify entire particle showers, and so on, culminating in an output that distinguishes electrons from jets. Training a neural network means finding the weights  $W^{(l)}$  and biases  $b^{(l)}$  that minimize the cost on the training set. We use



**Figure 7.6.** A Neural Network with 3 Layers: two hidden and one output Layer. [55]

the gradient-based optimization methods discussed in the previous section, but computing the gradient  $\nabla_{\theta} J(\theta)$  for a network with thousands or millions of parameters is non-trivial. This is where the backpropagation algorithm is crucial [61]. Backpropagation efficiently computes the gradient of the cost with respect to every weight in the network by propagating the error backwards through the network. It is essentially an application of the chain rule of calculus to the nested function  $f(\mathbf{x}; \theta)$ . In practice, we perform a forward pass to compute the predictions and the cost  $J$ , then a backward pass to compute gradients. In the backward pass, one introduces an error term  $\delta^{(l)}$  for each layer, which represents how a small change in that layer's activation would affect the cost. For the output layer, if we have a scalar output,  $\delta^{(\text{out})} = \frac{\partial J}{\partial \hat{y}} g'_{\text{out}}(z^{(\text{out})})$  where  $z^{(\text{out})} = W^{(3)}\mathbf{f}^{(2)} + b^{(3)}$  is the pre-activation. For layers  $l = L - 1$  down to 1,

$$\delta^{(l)} = \left( \left( W^{(l+1)} \right)^T \delta^{(l+1)} \right) \odot g'_l(z^{(l)}), \quad (7.12)$$

where  $\odot$  denotes elementwise multiplication and  $z^{(l)} = W^{(l)}\mathbf{f}^{(l-1)} + b^{(l)}$ . This formula comes from the chain rule, effectively distributing the error from layer  $l + 1$  back to layer  $l$  by the weights  $W^{(l+1)}$ . Once we have the  $\delta$  for a layer, the gradients for its parameters are:

$$\frac{\partial J}{\partial W^{(l)}} = \delta^{(l)} \left( \mathbf{f}^{(l-1)} \right)^T, \quad \frac{\partial J}{\partial b^{(l)}} = \delta^{(l)}, \quad (7.13)$$

i.e. the error term times the inputs to that layer (for weights) or 1 (for biases). With backpropagation, we can compute all these partial derivatives in time proportional to doing two forward passes. This algorithm, discovered multiple times and popularized in the 1980s [61], enabled the training of deep networks by drastically reducing the computational

burden of gradient computation. Modern automatic differentiation frameworks implement backprop for us, but understanding it is useful: for instance, it helps in diagnosing why vanishing or exploding gradients occur in very deep networks (gradients  $\delta^{(l)}$  either diminish or blow up as they propagate back through many layers, depending on the Jacobians  $W^{(k)}$  and activation slopes). Techniques like careful weight initialization and normalization are used to mitigate these issues so that even 100-layer networks can be trained. In summary, a feedforward neural network is a powerful function approximator defined by a composition of linear transformations and simple nonlinearities. By choosing an architecture (number of layers and neurons) and an appropriate loss function, we can train the network on example data using gradient-based optimization. The network then “learns” internal representations of the data that are useful for the desired prediction task. We emphasize that the design of the network (often called the architecture) can be guided by physical insight. For example, if we know certain symmetries or locality properties of the problem, we can build them into the architecture. This brings us to convolutional networks [55, 56], which incorporate translational symmetry and local receptive fields – a perfect fit for many physics problems with spatial or temporal structure.

## 7.4 Convolutional Neural Networks (CNNs)

Feedforward networks process inputs as fixed-length vectors without assuming any structure in the data. However, many physics datasets, such as frequency-series waveforms, have inherent structure. Convolutional neural networks (CNNs) leverage this structure through local connectivity and parameter sharing, making them well-suited for analyzing gravitational waveforms [56].

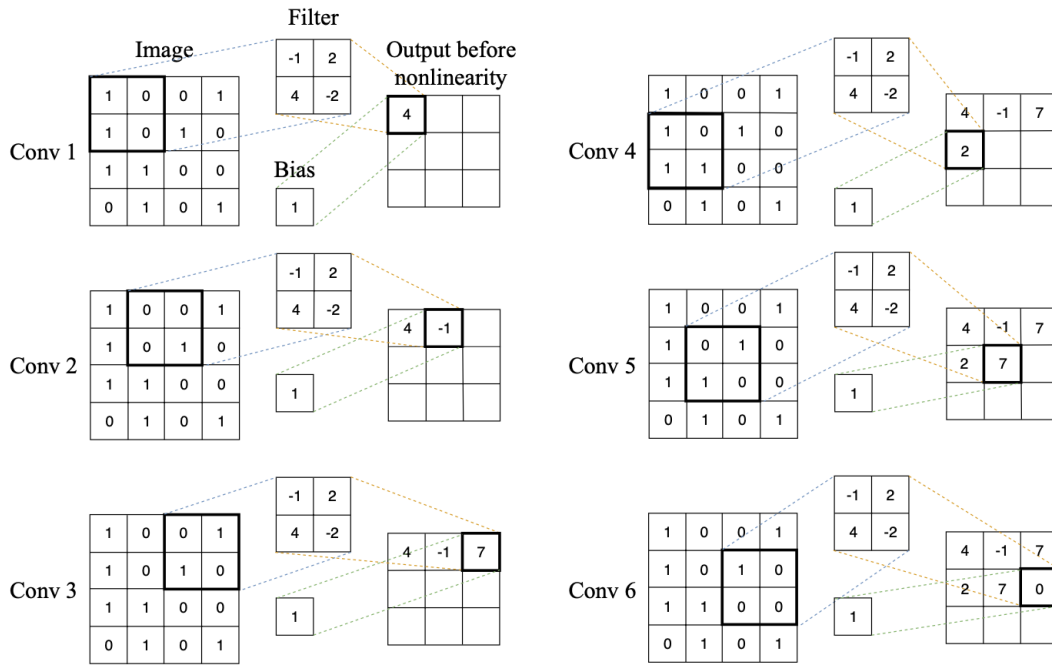
The core operation in a CNN is the convolution. For a 1D signal  $x[i]$ , convolving with a filter  $w[k]$  produces an output feature map  $y[i]$ :

$$y[i] = \sum_k w[k]x[i - k]. \quad (7.14)$$

This sliding operation detects localized patterns, such as oscillatory features in gravitational wave signals. Unlike traditional signal processing where filters are pre-defined, CNNs learn optimal filters from data. Two key advantages arise: *locality*, where each filter captures relevant short-time structures, and *weight sharing*, which enforces translational invariance—critical for detecting signals occurring at unknown times.



A typical CNN consists of multiple convolutional layers interleaved with nonlinear activations (e.g., ReLU) and pooling layers. Pooling reduces feature map resolution, improving robustness to small shifts and reducing computation. In waveform analysis, max-pooling over a few time steps helps summarize key patterns while retaining essential signal characteristics. After several convolutional and pooling layers, the output is processed by fully connected layers for classification or regression tasks.



**Figure 7.7.** A 2D convolutional network Filtering. Based on the dimension of the filter (kernel), and the padding, the dimension of the output changes [55].

CNNs drastically reduce the number of parameters compared to fully connected networks. For instance, detecting waveform features with small  $k$ -length filters requires far fewer weights than directly learning from raw time-series data. This efficiency allows CNNs to generalize well while retaining sensitivity to key structures.

In gravitational wave analysis, CNNs have been used to classify signals against noise by learning template-like filters optimized for signal detection [52]. Unlike matched filtering, which relies on predefined waveform templates, CNNs can automatically extract signal features and adapt to variations in waveform morphology. Additionally, spectrogram-based CNNs enable the classification of transient noise artifacts in detector data, improving data quality for further analysis.

Despite their success, interpretability remains an important challenge. Visualizing

learned filters and feature maps can provide insights into what patterns the network detects, enhancing trust in CNN-based analyses. Used appropriately, CNNs offer a powerful tool for gravitational wave searches, combining efficiency with the ability to learn data-driven features.

In the next section, we discuss the softmax activation function, commonly used in CNNs for multi-class classification.

## 7.5 Softmax Activation and Multi-Class Classification

While we have focused on binary classification, many problems involve multiple categories. In waveform analysis, for example, one may classify signals into different astrophysical sources. Neural networks handle multi-class outputs using the softmax activation [56, 55] in the output layer. Given scores  $z_1, z_2, \dots, z_K$  for  $K$  classes, softmax converts them into probabilities:

$$\hat{y}_k = \frac{\exp(z_k)}{\sum_{j=1}^K \exp(z_j)}, \quad k = 1, \dots, K. \quad (7.15)$$

This ensures  $\hat{y}_k \in (0, 1)$  and  $\sum_k \hat{y}_k = 1$ , making softmax useful for categorical predictions. The largest  $z_k$  leads to the highest  $\hat{y}_k$ , treating  $z_k$  as an unnormalized log-probability.

Training uses the cross-entropy loss, which encourages the network to increase the probability of the correct class while reducing others. The gradient of the loss with respect to  $z_k$  is simply  $\hat{y}_k - 1_{k=y}$ , reinforcing the correct classification. This formulation effectively trains the network to assign high confidence to the right class while suppressing incorrect ones.

Softmax is translation-invariant: adding a constant  $c$  to all  $z_k$  does not affect the output, meaning only relative values matter. The exponential nature of softmax amplifies differences, often leading to confident predictions. However, to mitigate overconfidence, techniques like temperature scaling (dividing logits by a factor) can adjust probability distributions for better uncertainty estimation [56].

In physics applications, softmax outputs provide a likelihood-based classification akin to hypothesis testing. For example, in gravitational wave analysis, a neural network can assign probabilities to different signal classes based on waveform characteristics [52]. Proper calibration ensures these probabilities reflect actual classification confidence.

In summary, softmax enables neural networks to perform multi-class classification by generating probabilistic outputs that integrate naturally with cross-entropy loss. It is widely

used in physics and other domains where distinguishing between multiple categories is essential. In the next section, we compare neural networks with other supervised learning methods.

## 7.6 Comparison with Other Supervised Learning Methods

Neural networks are a powerful approach to supervised learning, but they are not the only option. Methods such as support vector machines (SVMs) [62] and decision trees [63] also play roles in physics applications, each with their strengths and limitations.

**Support Vector Machines (SVMs)** classify data by maximizing the margin between classes, relying on support vectors near decision boundaries. They can efficiently handle non-linear patterns via kernel methods and were state-of-the-art before deep learning. SVMs perform well on smaller datasets and offer a convex optimization solution, but they scale poorly to large datasets and struggle with high-dimensional raw inputs like waveforms. For example, classifying gravitational wave signals with an SVM would require manual feature extraction (e.g., Fourier transforms), whereas a CNN learns features directly from raw time-series data.

**Decision Trees** recursively split feature space into regions based on threshold tests. While interpretable, single trees tend to overfit. Ensemble methods, such as Random Forests and Boosted Decision Trees (BDTs), improve performance and have been widely used in physics, including Higgs boson searches. Trees are particularly strong for tabular data with predefined features but do not naturally handle raw signals like waveforms. Compared to neural networks, tree ensembles require less parameter tuning and handle missing data well but lack the flexibility of deep learning models.

**Other Methods** include Bayesian classifiers, which rely on probabilistic models but often make oversimplified assumptions, and linear models like logistic regression, which work well for low-dimensional data but fail to capture complex structures without feature engineering.

In modern physics applications with large datasets—such as collider physics, cosmology, and gravitational wave analysis—deep learning has become dominant due to its ability to learn directly from raw data and leverage high-performance computing. However, classical methods remain useful, especially for small datasets or when interpretability is crucial. A practical approach may involve starting with simpler models like Random Forests to as-

sess data complexity before moving to neural networks if greater performance is required.

Hybrid approaches also exist, such as using a CNN to extract features from waveforms and feeding them into a tree-based classifier. While end-to-end deep learning is increasingly favored, traditional methods remain valuable, particularly in theoretical physics contexts where model interpretability and analytical tractability are priorities.

In this chapter, we explored supervised learning with a focus on neural networks and convolutional architectures, highlighting their advantages for physics applications. We covered key concepts, including optimization techniques, regularization, and activation functions, leading to a discussion on CNNs and their effectiveness in structured data analysis. Compared to traditional methods like SVMs and decision trees, deep learning excels in handling large datasets and learning complex patterns [\[64\]](#) directly from raw signals. Neural networks, particularly CNNs, provide a powerful approach to detecting structured features in data, making them well-suited for gravitational wave analysis. In the next chapter, we apply these concepts to waveform classification, constructing and training a neural network to distinguish gravitational wave signals.

# Chapter 8

## Problem Setup and Neural Network Architecture

After the pedagogical introduction to neural networks in the previous chapter, we now turn our focus to the specific problem of classifying gravitational waveforms using supervised learning. This chapter outlines the problem setup, the derivation of the gravitational waveform in the frequency domain, and the neural network architecture that we employ for classification.

### 8.1 Frequency-Domain Gravitational Waveform

To begin with, we derive the frequency-domain representation of gravitational waves sourced from compact binary mergers [65, 66, 67, 68]. These waveforms provide the basis for our machine learning models, which classify signals as either General Relativity (GR) or beyond-GR (BGR). The gravitational wave strain,  $h(t)$ , is typically decomposed into two independent polarization modes:

$$h(t) \equiv h_+(t) - ih_\times(t) \quad (8.1)$$

where  $h_+$  and  $h_\times$  are the plus and cross polarizations, respectively. These polarizations are conveniently expressed in terms of spin-weighted spherical harmonics:

$$h(t; \lambda, \iota, \varphi_c) = \sum_{\ell=2}^{+\infty} \sum_{m=-\ell}^{\ell} -2Y_{\ell m}(\iota, \varphi_c) h_{\ell m}(t, \lambda). \quad (8.2)$$

The intrinsic parameters of the binary, such as component masses  $m_1, m_2$  and spins  $\chi_1, \chi_2$ , determine the time evolution of  $h_{\ell m}(t)$ .

Using the general time-domain representation,

$$h_{\ell m}(t) = A_{\ell m}(t) e^{i\Phi_{\ell m}(t)}, \quad (8.3)$$

we now compute the Fourier transform of the gravitational wave signal. The Fourier transforms of the plus and cross polarizations are given by [66]:

$$\begin{aligned} \tilde{h}_+(f) &= \sum_{\ell=2}^{\infty} \sum_{m=1}^{\ell} \sqrt{\frac{2\ell+1}{4\pi}} [(-1)^{\ell} d_{\ell,-m}^2(\theta) + d_{\ell m}^2(\theta)] \tilde{h}_{\ell m}(f), \\ \tilde{h}_{\times}(f) &= -i \sum_{\ell=2}^{\infty} \sum_{m=1}^{\ell} \sqrt{\frac{2\ell+1}{4\pi}} [(-1)^{\ell} d_{\ell,-m}^2(\theta) - d_{\ell m}^2(\theta)] \tilde{h}_{\ell m}(f), \end{aligned} \quad (8.4)$$

where  $d_{\ell m}^2(\theta)$  are the Wigner  $d$ -functions,  $\theta$  and  $\phi_0$  describe the line of sight in the detector frame, and the Fourier transform of the spherical harmonic mode is

$$\tilde{h}_{\ell m}(f) = \int_{-\infty}^{\infty} e^{2\pi i f t} A_{\ell m}(t) e^{i\Phi_{\ell m}(t)} dt. \quad (8.5)$$

The real and imaginary parts of the Fourier transform can be computed separately as:

$$\begin{aligned} \tilde{h}_{\ell m}^R(f) &= \int_{-\infty}^{\infty} e^{2\pi i f t} A_{\ell m}(t) \cos(\Phi_{\ell m}(t)) dt \\ \tilde{h}_{\ell m}^I(f) &= \int_{-\infty}^{\infty} e^{2\pi i f t} A_{\ell m}(t) \sin(\Phi_{\ell m}(t)) dt \end{aligned} \quad (8.6)$$

For non-spinning and non-precessing binaries, the imaginary part satisfies  $\tilde{h}_{\ell m}^I(f) = -i\tilde{h}_{\ell m}^R(f)$ , allowing us to express the Fourier transforms as:

$$\begin{aligned} \tilde{h}_+(f) &= \sum_{\ell=2}^{\infty} \sum_{m=1}^{\ell} \left[ (-1)^{\ell} \frac{d_{\ell,-m}^2(\theta)}{d_{\ell m}^2(\theta)} + 1 \right] {}_{-2}Y_{\ell m}(\theta, \phi_0) \tilde{h}_{\ell m}^R(f), \\ \tilde{h}_{\times}(f) &= -i \sum_{\ell=2}^{\infty} \sum_{m=1}^{\ell} \left[ (-1)^{\ell} \frac{d_{\ell,-m}^2(\theta)}{d_{\ell m}^2(\theta)} - 1 \right] {}_{-2}Y_{\ell m}(\theta, \phi_0) \tilde{h}_{\ell m}^R(f). \end{aligned} \quad (8.7)$$

In terms of the Fourier domain, phase and amplitude  $\tilde{h}_{\ell m}^R(f)$  can be written as follows:

$$\tilde{h}_{\ell m}^R(f) = A_{\ell m}(f) e^{i\psi_{\ell m}(f)} \quad (8.8)$$

The phase  $\psi_{\ell m}(f)$  in the above equation will be the main object of our study.

## 8.2 Waveform Extensions

The primary objective of our study is to investigate the modifications in the phase and amplitude of the waveform [69]. The quantities  $\psi_{\ell m}(f)$  and  $A_{\ell m}(f)$  represent the phase and

amplitude, respectively, in the Fourier domain [70, 71]. We introduce potential deviations, denoted as  $\delta\psi_{\ell m}$  and  $\delta A_{\ell m}$ , which modify these quantities as follows:

$$\psi_{\ell m}(\eta, M, \chi_S, \chi_A) = \psi_{\ell m}^{(GR)}(\eta, M, \chi_S, \chi_A) + \delta\psi_{\ell m}, \quad (8.9)$$

for the phase, and

$$A_{\ell m}(\eta, M, \chi_S, \chi_A) = A_{\ell m}^{(GR)}(\eta, M, \chi_S, \chi_A) + \delta A_{\ell m}, \quad (8.10)$$

for the amplitude.

Here,  $\eta$  represents the symmetric mass ratio:

$$\eta = \frac{q}{(1+q)^2}, \quad \text{with } q = \frac{m_1}{m_2}, \quad (8.11)$$

where  $M$  is the total mass of the binary:

$$M = m_1 + m_2. \quad (8.12)$$

The parameters  $\chi_S$  and  $\chi_A$  are the symmetric and anti-symmetric spin components, defined as:

$$\chi_S = \frac{\chi_1 + \chi_2}{2}, \quad \chi_A = \frac{\chi_1 - \chi_2}{2}. \quad (8.13)$$

The deviations  $\delta\psi_{\ell m}$  and  $\delta A_{\ell m}$  introduce small shifts in  $\eta$ ,  $M$ , and the spin components. These shifts are denoted by  $\delta\eta$ ,  $\delta M$ ,  $\delta\chi_S$ , and  $\delta\chi_A$ . Accounting for these shifts, the total observational shift in phase  $\Delta\psi_{\ell m}$  is given by:

$$\psi_{\ell m}(\eta + \delta\eta, M + \delta M, \chi_S + \delta\chi_S, \chi_A + \delta\chi_A) = \psi_{\ell m}^{(GR)}(\eta, M, \chi_S, \chi_A) + \Delta\psi_{\ell m}. \quad (8.14)$$

Expanding to first order in the shifts, we obtain:

$$\Delta\psi_{\ell m} = \delta\psi_{\ell m} + \frac{\partial\psi_{\ell m}^{(GR)}}{\partial\eta}\delta\eta + \frac{\partial\psi_{\ell m}^{(GR)}}{\partial M}\delta M + \frac{\partial\psi_{\ell m}^{(GR)}}{\partial\chi_S}\delta\chi_S + \frac{\partial\psi_{\ell m}^{(GR)}}{\partial\chi_A}\delta\chi_A, \quad (8.15)$$

for the phase, and

$$\Delta A_{\ell m} = \delta A_{\ell m} + \frac{\partial A_{\ell m}^{(GR)}}{\partial\eta}\delta\eta + \frac{\partial A_{\ell m}^{(GR)}}{\partial M}\delta M + \frac{\partial A_{\ell m}^{(GR)}}{\partial\chi_S}\delta\chi_S + \frac{\partial A_{\ell m}^{(GR)}}{\partial\chi_A}\delta\chi_A, \quad (8.16)$$

for the amplitude.

To properly assess the significance of these shifts, it is essential to compare them relative to the corresponding parameters. The absolute magnitude of the shifts alone does not provide meaningful insight; rather, their relative effect on the waveform is crucial. Hence, we normalize the shifts by the parameter values:

$$\begin{aligned}\Delta\psi_{\ell m} &= \delta\psi_{\ell m} + \left(\eta \frac{\partial\psi_{\ell m}^{(GR)}}{\partial\eta}\right) \frac{\delta\eta}{\eta} + \left(M \frac{\partial\psi_{\ell m}^{(GR)}}{\partial M}\right) \frac{\delta M}{M} + \left(\chi_S \frac{\partial\psi_{\ell m}^{(GR)}}{\partial\chi_S}\right) \frac{\delta\chi_S}{\chi_S} + \left(\chi_A \frac{\partial\psi_{\ell m}^{(GR)}}{\partial\chi_A}\right) \frac{\delta\chi_A}{\chi_A}, \\ \Delta A_{\ell m} &= \delta A_{\ell m} + \left(\eta \frac{\partial A_{\ell m}^{(GR)}}{\partial\eta}\right) \frac{\delta\eta}{\eta} + \left(M \frac{\partial A_{\ell m}^{(GR)}}{\partial M}\right) \frac{\delta M}{M} + \left(\chi_S \frac{\partial A_{\ell m}^{(GR)}}{\partial\chi_S}\right) \frac{\delta\chi_S}{\chi_S} + \left(\chi_A \frac{\partial A_{\ell m}^{(GR)}}{\partial\chi_A}\right) \frac{\delta\chi_A}{\chi_A}.\end{aligned}\tag{8.17}$$

Thus, the computation of observational shifts requires knowledge of both the intrinsic extensions  $\delta\psi_{\ell m}$  and  $\delta A_{\ell m}$ , as well as the variations in the mass and spin parameters:  $\delta\eta$ ,  $\delta M$ ,  $\delta\chi_S$ , and  $\delta\chi_A$ .

### 8.3 Extending Binary Black Hole Waveforms Beyond GR

In this section, we investigate binary black hole (BBH) events and extend their phase to construct beyond General Relativity (BGR) waveforms [66, 69]. The key idea is to introduce artificial modifications directly to the phase rather than deriving them from a specific alternative theory of gravity. This approach allows us to train a model that can classify any deviations from GR without being constrained to a particular theoretical framework.

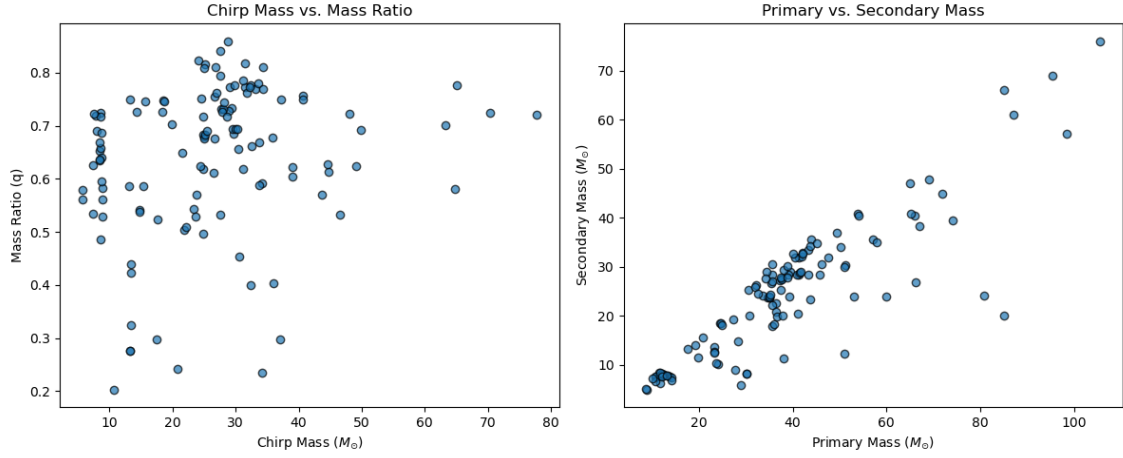
Instead of working from a specific theory and deriving its effect on the phase—for example, through modifications to the post-Newtonian parameters—we introduce direct modifications to the phase itself. This method ensures a more generalized classification model that can identify arbitrary deviations from GR. However, if one is interested in testing a particular theory, they can derive the explicit phase modification from that theory and still utilize our neural network setup for classification.

#### 8.3.1 Extracting Binary Black Hole Event Data

We begin by considering 121 BBH events with their data extracted from the LIGO database [72]. These events serve as the basis for generating the corresponding gravitational waveforms.

To generate the waveforms, we utilize the PyCBC package [73] in Python, specifically employing the IMRPhenomXHM approximant. This waveform model incorporates higher-order modes, making it well-suited for accurately modeling BBH signals observed





**Figure 8.1.** BBH events taken from LIGO's database.

by LIGO. Moreover, since our primary focus is on modifications in the inspiral regime, this approximant is particularly appropriate for our analysis.

### 8.3.2 Data Augmentation for Neural Network Training

For effective neural network training, a sufficiently large dataset is required to ensure generalization and robustness. However, our initial dataset contains only 121 BBH events, which is relatively small for deep learning applications. To address this, we employ data augmentation, a common strategy in machine learning, where new data points are generated by introducing controlled variations to existing ones. This allows us to expand the dataset without acquiring additional observational data.

In gravitational wave astrophysics, BBH parameters such as component masses and spins are continuous variables rather than discrete categories. Consequently, realistic variations in these parameters should not fundamentally alter the physics of the waveform but should still introduce enough diversity for training a neural network. By perturbing these parameters slightly, we ensure that the model learns robust features rather than memorizing a limited set of waveforms.

The key principle behind our augmentation approach is to apply Gaussian perturbations to certain physical parameters while ensuring that the resulting values remain physically meaningful. Specifically, we perturb:

- Component masses  $m_1, m_2$
- Effective spin parameter  $\chi_{\text{eff}}$

To generate augmented samples, we apply small Gaussian-distributed perturbations to the parameters. The perturbations are defined as follows:

### Component Mass Perturbations

For a given binary system with primary mass  $m_1$  and secondary mass  $m_2$ , we introduce perturbations using a normal distribution:

$$m'_1 = m_1 + \mathcal{N}(0, \sigma_{m_1}), \quad m'_2 = m_2 + \mathcal{N}(0, \sigma_{m_2}), \quad (8.18)$$

where  $\mathcal{N}(0, \sigma)$  represents a Gaussian distribution with zero mean and standard deviation  $\sigma$ . The standard deviation is chosen as a fraction  $p$  (randomly selected between 3% and 5%) of the original mass:

$$\sigma_{m_1} = p \cdot m_1, \quad \sigma_{m_2} = p \cdot m_2. \quad (8.19)$$

Since the primary mass must always be larger than the secondary mass, we impose the constraint:

$$m'_1 \geq m'_2. \quad (8.20)$$

### Effective Spin Perturbations

The effective spin parameter  $\chi_{\text{eff}}$  is defined as:

$$\chi_{\text{eff}} = \frac{m_1 \chi_1 + m_2 \chi_2}{m_1 + m_2}. \quad (8.21)$$

To introduce variations, we apply a Gaussian perturbation:

$$\chi'_{\text{eff}} = \chi_{\text{eff}} + \mathcal{N}(0, \sigma_{\chi_{\text{eff}}}), \quad (8.22)$$

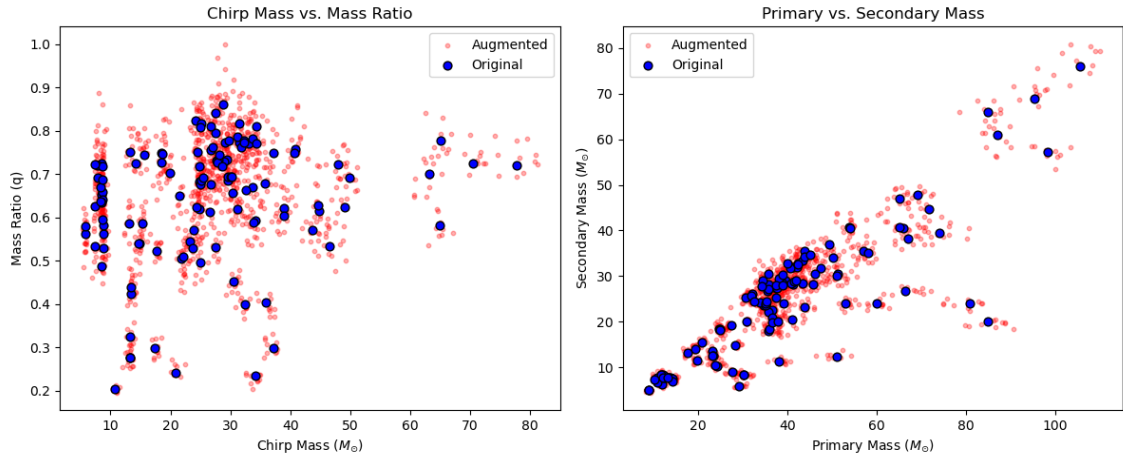
where the standard deviation is defined as:

$$\sigma_{\chi_{\text{eff}}} = p \cdot |\chi_{\text{eff}}| + 0.01. \quad (8.23)$$

To ensure physical validity, the perturbed effective spin must remain within the allowed range:

$$-1 \leq \chi'_{\text{eff}} \leq 1. \quad (8.24)$$

By iterating this augmentation process over all events in the dataset, we effectively expand the number of training examples, ensuring a diverse and well-distributed dataset for training our neural network. The augmented dataset retains the essential physical characteristics of the original BBH signals while introducing variability that improves the model's ability to generalize beyond the limited set of observed waveforms. With the augmentation, we expand our original dataset tenfold.



**Figure 8.2.** BBH events and their augmented counterparts. Blue dots represent the original data, while red dots denote the augmented events generated via Gaussian perturbations.

## 8.4 Constructing Beyond-GR Waveforms from GR

To systematically construct beyond-GR (BGR) waveforms, we start from the original GR waveform and introduce a phase modification. For each GR event, we generate a corresponding BGR waveform by applying a controlled phase shift to the  $(\ell, m) = (2, 2)$  mode [69]. The motivation behind modifying only this mode is its dominant contribution to the total gravitational waveform. While higher-order modes ( $\ell > 2$ ) can be significant in asymmetric mass-ratio systems or high-inclination binaries, the  $(2, 2)$  mode primarily dictates the overall structure of the signal. By restricting modifications to this mode, we ensure computational efficiency while still capturing leading-order deviations from GR.

### 8.4.1 Frequency-Domain Representation and Phase Modification

A gravitational waveform in the frequency domain can be expressed as a sum over all modes:

$$h_{\text{GR}}(f) = \sum_{\ell,m} h_{\ell m, \text{GR}}(f) e^{i\psi_{\ell m, \text{GR}}(f)}, \quad (8.25)$$

where:

- $h_{\ell m, \text{GR}}(f)$  is the amplitude of the  $(\ell, m)$  mode in the frequency domain,
- $\psi_{\ell m, \text{GR}}(f)$  is the corresponding phase function.

To construct a BGR waveform, we introduce a small modification to the phase function, parameterized by a deviation  $\delta\psi(f)$ , such that:

$$h_{\text{BGR}}(f) = \sum_{\ell,m} h_{\ell m, \text{GR}}(f) e^{i(\psi_{\ell m, \text{GR}}(f) + \beta \delta\psi_{\ell m}(f))}, \quad (8.26)$$

where  $\beta$  is a free parameter that controls the strength of the modification.

Since the (2,2) mode dominates the waveform, we approximate the total signal by considering only this mode:

$$h_{\text{BGR}}(f) \approx h_{\text{GR}}(f) e^{i(\psi_{22, \text{GR}}(f) + \beta \delta\psi_{22}(f))}. \quad (8.27)$$

For small  $\delta\psi_{22}(f)$ , we expand the exponential function using a first-order Taylor expansion:

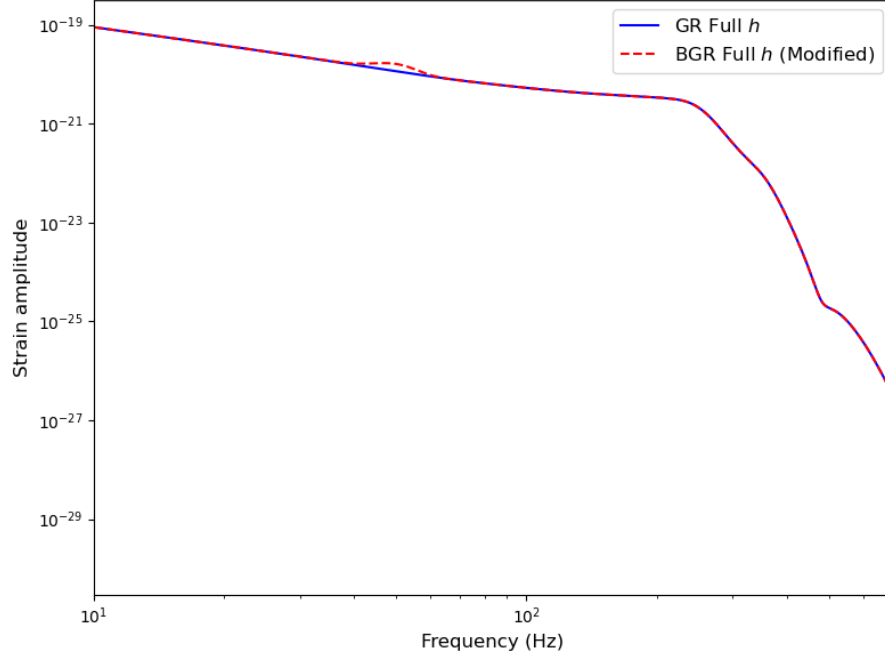
$$e^{i\beta \delta\psi_{22}(f)} \approx 1 + i\beta \delta\psi_{22}(f) + O(\delta\psi^2). \quad (8.28)$$

Substituting this into the expression for  $h_{\text{BGR}}(f)$ , we obtain:

$$h_{\text{BGR}}(f) \approx h_{\text{GR}}(f) + i\beta h_{22, \text{GR}}(f) \delta\psi_{22}(f). \quad (8.29)$$

This construction provides a model-independent way to extend GR waveforms to BGR waveforms. Instead of deriving  $\delta\psi_{22}(f)$  from a specific beyond-GR theory, we allow it to be a free function, parameterized for general deviations. This enables our neural network to be trained on a broad class of deviations rather than being restricted to a single theoretical framework. However, if a specific beyond-GR model predicts an explicit phase shift, it can still be incorporated into this framework by using its corresponding  $\delta\psi_{22}(f)$ .

By systematically modifying only the (2,2) mode, we retain the essential physical characteristics of the original GR waveforms while allowing for deviations that can be detected and classified by a machine learning model.



**Figure 8.3.** The red dashed line represents a small shift in the waveform during the inspiral (low-frequency) regime, illustrating the effect of the applied phase modification in the BGR waveform.

## 8.5 Noise Injection

To make the dataset more representative of real gravitational wave detections, we inject detector noise into the waveforms. In actual observations, gravitational wave signals are buried in instrumental noise, and training a neural network on noiseless waveforms would not generalize well to real data. To simulate a more realistic scenario, we add noise to both the GR and BGR waveforms, following the characteristics of LIGO’s detector noise.

### 8.5.1 LIGO Power Spectral Density (PSD)

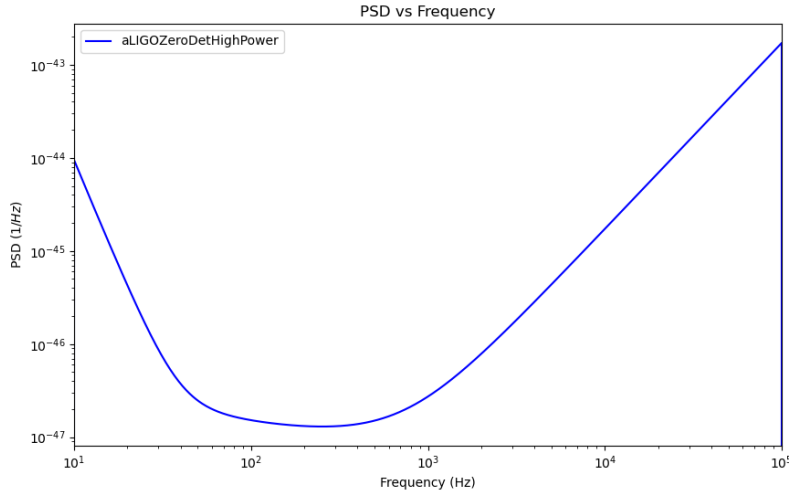
The noise in gravitational wave detectors is characterized by the power spectral density (PSD), which quantifies the frequency-dependent noise amplitude. Given a detector’s strain noise  $n(f)$ , the PSD is defined as [74]:

$$S_n(f) = \lim_{T \rightarrow \infty} \frac{1}{T} \langle |\tilde{n}(f)|^2 \rangle, \quad (8.30)$$

where:

- $\tilde{n}(f)$  is the Fourier transform of the time-domain detector noise  $n(t)$ ,
- $T$  is the observation time window,
- $\langle \cdot \rangle$  denotes an ensemble average over realizations of the noise.

The PSD provides a measure of how noise varies with frequency. In the case of LIGO, the noise is colored, meaning that it is not uniform across all frequencies but instead follows a characteristic shape dictated by instrumental and environmental effects.



**Figure 8.4.** The power spectral density (PSD) used for noise injection in our analysis, corresponding to the aLIGOZeroDetHighPower configuration [72]. This PSD represents the expected sensitivity of Advanced LIGO in its zero-detuned, high-power mode, providing a realistic noise model for simulating gravitational wave detections from binary black hole mergers.

### 8.5.2 Statistical Properties of the Noise

The detector noise is assumed to be stationary and Gaussian, meaning that different frequency components are uncorrelated. The two-sided PSD  $S_n(f)$  satisfies the following expectation value:

$$\langle \tilde{n}(f) \tilde{n}^*(f') \rangle = \delta(f - f') S_n(f), \quad (8.31)$$

where:

- $\tilde{n}(f)$  is the frequency-domain representation of the noise,
- $\delta(f - f')$  ensures that noise at different frequencies is uncorrelated,
- $S_n(f)$  determines the variance of the noise at frequency  $f$ .

This expression shows that the noise power at each frequency component is dictated by  $S_n(f)$ , while different frequencies remain statistically independent.

### 8.5.3 Generating and Adding Realistic Detector Noise

To generate realistic noise, we sample from a zero-mean Gaussian distribution with variance given by the PSD:

$$\tilde{n}(f) \sim \mathcal{N}(0, S_n(f)). \quad (8.32)$$

However, to ensure proper scaling of the noise, we normalize it as:

$$\tilde{n}(f) = \frac{z(f)}{\sqrt{S_n(f)}}, \quad (8.33)$$

where  $z(f)$  is a complex Gaussian random variable with zero mean and unit variance. The final noisy waveform is then obtained by adding this noise to the clean signal:

$$\tilde{h}_{\text{noisy}}(f) = \tilde{h}_{\text{clean}}(f) + \tilde{n}(f). \quad (8.34)$$

We apply this noise injection procedure separately to both GR and BGR waveforms. For each event, we:

- Generate a unique noise realization using LIGO's PSD.
- Apply the same noise realization to both the full waveform and the dominant (2,2) mode.
- Use a different seed for GR and BGR waveforms to ensure independent noise realizations.

This approach ensures that the dataset not only reflects theoretical waveform modifications but also accounts for realistic observational conditions, improving the robustness of the neural network's classification capability.

## 8.6 Neural Network Implementation

With our dataset fully prepared, we now focus on the design and implementation of a neural network to classify gravitational waveforms as either GR or beyond GR(BGR). The goal is to construct a model that effectively learns patterns from the noisy waveform data and generalizes well to unseen examples [69]. In this section, we describe the dataset split, the structure of the neural network, and its implementation.

### 8.6.1 Dataset Splitting and Its Purpose

To ensure that the model learns effectively while also being evaluated on unseen data, we split the dataset into three distinct subsets: training, validation, and testing [55]. Each subset serves a specific purpose in the learning process.

#### Training Set

The training set contains the majority of the data and is used for optimizing the neural network's parameters. During training, the model learns by minimizing a loss function, which measures the discrepancy between predicted and true labels. The dataset for training, denoted as  $\mathcal{D}_{\text{train}}$ , consists of waveform samples  $X_{\text{train}}$  and corresponding labels  $y_{\text{train}}$ :

$$\mathcal{D}_{\text{train}} = \{(X_i, y_i)\}_{i=1}^{N_{\text{train}}}. \quad (8.35)$$

The neural network aims to approximate a function  $f_{\theta}(X)$  parameterized by  $\theta$  that maps input waveforms to output labels:

$$\hat{y} = f_{\theta}(X). \quad (8.36)$$

#### Validation Set and Cross-Validation

The validation set is used to fine-tune hyperparameters and monitor model performance. Unlike the training set, it is not used for weight updates. Instead, after each training epoch, the model evaluates its performance on  $\mathcal{D}_{\text{val}}$ , providing an estimate of its generalization ability.

To assess performance, the validation loss is computed as:



$$\mathcal{L}_{\text{val}} = -\frac{1}{N_{\text{val}}} \sum_{i=1}^{N_{\text{val}}} [y_i \log \sigma(f_{\theta}(X_i)) + (1 - y_i) \log(1 - \sigma(f_{\theta}(X_i)))] , \quad (8.37)$$

where  $\sigma(y) = 1/(1 + e^{-y})$  is the sigmoid activation function.

A key use of the validation set is detecting overfitting. If the validation loss starts increasing while the training loss continues decreasing, this indicates that the model is memorizing the training data rather than learning general features.

### Testing Set

After training is complete, the model is evaluated on the test set to assess its real-world performance. The test accuracy is computed as:

$$A_{\text{test}} = \frac{1}{N_{\text{test}}} \sum_{i=1}^{N_{\text{test}}} \mathbf{1}\{\hat{y}_i = y_i\}, \quad (8.38)$$

where  $\mathbf{1}$  is an indicator function that equals 1 if the prediction matches the true label and 0 otherwise. A high test accuracy suggests strong generalization ability.

## 8.6.2 Neural Network Architecture

We implement a one-dimensional convolutional neural network (CNN) to classify gravitational waveforms. CNNs are particularly effective for structured data such as time-series or frequency-domain signals because they apply local feature extraction.

The architecture consists of the following components:

- **Input layer:** Takes in the frequency-domain waveform as a one-dimensional input array.
- **Convolutional layers:** Apply a set of trainable filters to capture local frequency patterns. Each convolution operation is followed by a non-linear activation function  $g$  that introduces non-linearity into the model.
- **Fully connected layers:** Flatten the extracted features and pass them through dense layers that progressively learn high-level representations.
- **Output layer:** A single neuron that outputs a real number (logit), which is then passed through a sigmoid function to interpret the result as a probability of being a BGR waveform

Mathematically, each convolutional layer performs:

$$y_j = g \left( \sum_i w_{ij} x_i + b_j \right), \quad (8.39)$$

where: -  $x_i$  are input features (frequency bins), -  $w_{ij}$  are learnable filter weights, -  $b_j$  is a bias term, -  $g(\cdot)$  is the activation function.

The final dense layer combines extracted features into a weighted sum:

$$y_{\text{out}} = W y_{\text{conv}} + b. \quad (8.40)$$

In code, this corresponds to the following implementation using the TensorFlow package [75, 76]:

```
model = models.Sequential([
    layers.Input(shape=input_shape),
    layers.Conv1D(16, kernel_size=5, activation='relu'),
    layers.Conv1D(32, kernel_size=5, activation='relu'),
    layers.Conv1D(64, kernel_size=5, activation='relu'),
    layers.Flatten(),
    layers.Dense(128, activation='relu'),
    layers.Dense(1, activation='linear')
])
```

The model is trained using the Adam optimizer [58] with a learning rate of  $10^{-4}$ . Training is performed using mini-batches of size 64 over 50 epochs. Throughout training, the model learns to extract distinguishing features from the frequency-domain waveforms.

To monitor performance, we track both training and validation loss at each epoch. If validation loss increases while training loss decreases, this suggests overfitting.

Once training is complete, the model is tested on unseen data, and its accuracy is computed. A strong test accuracy indicates that the model has learned generalizable patterns for classifying gravitational waveforms.

In this chapter, we have developed all the necessary tools for training our classification neural network. We constructed a dataset of gravitational waveforms, introduced noise injection to simulate realistic observational conditions, and designed a convolutional neural network capable of distinguishing GR from BGR waveforms.

With the dataset and model architecture in place, we are now prepared to apply these methods in training. In the next chapter, we train multiple neural networks, each based on

---

different modifications to the phase  $\delta\psi(f)$ , and analyze their performance in classifying beyond-GR signals.



## Chapter 9

# Training and Evaluation of the Classification Neural Network

In this chapter, we apply the neural network architecture developed earlier to a controlled toy model, where deviations from General Relativity (GR) are introduced in a simplified and tunable way. The goal is to explore how sensitive the classifier is to small departures from GR, and to determine the minimum strength of modification that remains detectable.

To this end, we start from a set of GR waveforms and construct corresponding beyond-GR (BGR) waveforms by modifying only their phase in the frequency domain. Specifically, we consider a Gaussian deformation localized in the inspiral regime, and study its impact on the classification task.

### 9.1 Toy Model: Gaussian Phase Shift

We define the BGR waveform by applying a phase deformation to the GR waveform in the frequency domain. In this toy model, the modification is localized to the dominant  $(\ell, m) = (2, 2)$  mode and takes the form

$$\delta\psi_{22}(f) = \exp\left[-\frac{(f - 50)^2}{100}\right], \quad (9.1)$$

where  $f$  is the frequency. This shift is a Gaussian centered at 50 Hz with a width  $\sqrt{100} = 10$  Hz, which lies well within the inspiral regime of the waveform. The specific choice of 50 and 100 is not physically significant—it serves only to localize the modification in the low-frequency part of the signal, consistent with the illustrative purpose of the toy model.

This Gaussian deformation is then plugged into the general modification formula defined previously in Eq. (8.29), which in our case becomes

$$h_{\text{BGR}}(f) = h_{\text{GR}}(f) + i\beta h_{22,\text{GR}}(f) \delta\psi_{22}(f), \quad (9.2)$$

where  $h_{\text{GR}}(f)$  is the original GR waveform,  $h_{22,\text{GR}}(f)$  is the GR waveform restricted to the (2, 2) mode, and  $\beta$  is a free parameter controlling the strength of the deviation from GR.

For each value of  $\beta$ , we generate a set of BGR waveforms using Eq. (9.2), and compare them to their GR counterparts. In the next section, we will evaluate how distinguishable these waveforms are by computing average matches between GR and BGR signals. This will serve as a preliminary measure of the magnitude of the modification before proceeding to neural network classification.

## 9.2 Mismatch as a Preliminary Metric

Before attempting classification with a neural network, it is useful to first assess the magnitude of the deviation introduced by the Gaussian phase shift. This gives us a sense of how distinguishable the modified BGR waveforms are from their original GR counterparts. A natural and widely used similarity measure in gravitational-wave data analysis is the *match* between two waveforms, defined with respect to the detector noise properties [74].

Let  $A(f)$  and  $B(f)$  be two waveforms expressed in the frequency domain. The inner product between them is given by

$$(A | B) = 4 \Re \int_0^\infty \frac{A^*(f)B(f)}{S_n(f)} df, \quad (9.3)$$

where  $S_n(f)$  denotes the one-sided power spectral density (PSD) of the detector noise, and  $\Re\{\cdot\}$  indicates the real part. The *match* between  $A$  and  $B$  is then defined as the normalized inner product, maximized over relative time and phase shifts:

$$\text{Match}(A, B) = \max_{t_0, \phi_0} \frac{(A | B)}{\sqrt{(A | A)(B | B)}}. \quad (9.4)$$

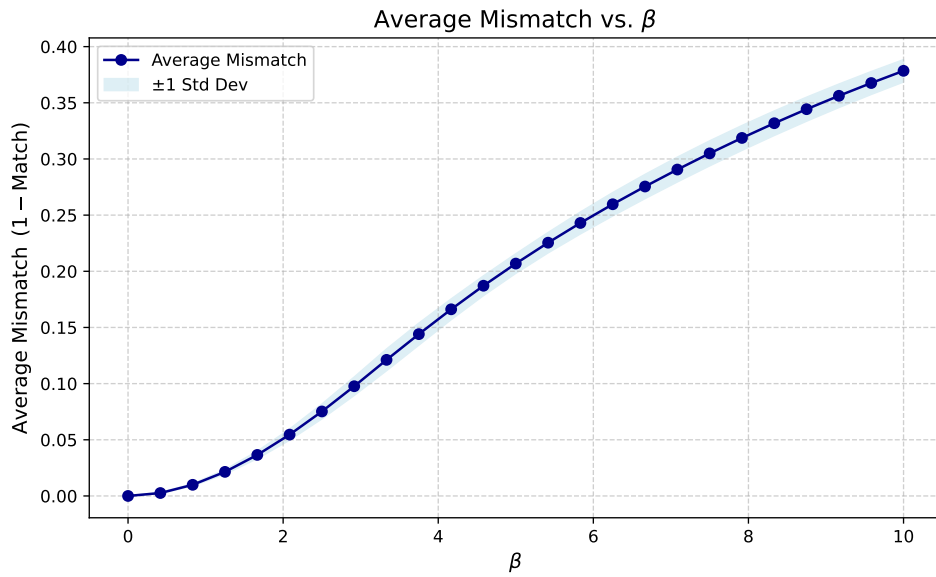
To quantify the difference between BGR and GR waveforms, we compute the average *mismatch*, defined as

$$\text{Mismatch} = 1 - \text{Match}. \quad (9.5)$$

A mismatch value close to zero indicates that the waveforms are nearly indistinguishable, while a value closer to one signals strong deviation.

In our analysis, we compute the mismatch between the original GR waveform and its BGR counterpart for each event in the dataset, at several values of the phase deformation strength  $\beta$ . We then average over all events to obtain the average mismatch as a function of  $\beta$ .

Figure 9.1 shows the average mismatch as a function of the deformation strength  $\beta$ . As expected, the mismatch increases with  $\beta$ , reflecting the growing deviation from GR. We also display a shaded band corresponding to one standard deviation across the waveform population to indicate the spread of values.



**Figure 9.1.** Average mismatch between BGR and GR waveforms as a function of the deformation parameter  $\beta$ . The shaded region shows the standard deviation across the waveform dataset. As  $\beta$  increases, the deviation becomes more pronounced, leading to a higher mismatch.

While the mismatch provides a useful first indication of waveform similarity, it is not always a reliable measure of physical deviation. In this analysis, we set the inclination angle to zero and consider only the dominant (2, 2) mode. Under these simplifying assumptions, the phase modification applied to the waveform appears clearly in the strain signal and is captured by the mismatch.

However, this is not generally the case. For instance, if the inclination angle were set to  $\pi/2$ , the observed strain from the (2, 2) mode would be significantly suppressed, and any deformation applied to that mode—such as the Gaussian shift considered here—would become practically invisible. In such scenarios, the mismatch could remain close to zero

even in the presence of a meaningful physical deviation. Similarly, if the deviation manifests in higher-order modes or in a way that interferes constructively or destructively with other components, the mismatch may fail to reflect it accurately.

Therefore, while mismatch serves as a useful baseline diagnostic in this toy setup, more robust and comprehensive metrics are needed for general analyses that include multiple modes, varying inclination angles, or realistic astrophysical signals.

### 9.3 Classification Performance at Strong Deformation ( $\beta = 1$ )

In this section, we apply the convolutional neural network (CNN) introduced in the previous chapter to classify gravitational waveforms into GR and BGR categories. The BGR waveforms are constructed using the Gaussian phase shift toy model described earlier, with the deformation strength fixed to  $\beta = 1$ . This case corresponds to the largest deviation from GR in our setup and serves as a reference point for evaluating the classifier's effectiveness.

#### 9.3.1 Evaluation Setup

Each waveform is represented by its amplitude spectrum, interpolated onto a common frequency grid ranging from 10 to 400 Hz. The dataset is split into training, validation, and test sets in a 70%-15%-15% ratio, with stratification to preserve class balance. Each sample is normalized individually using z-score normalization prior to being passed to the neural network.

The CNN architecture, consisting of multiple convolutional layers followed by max-pooling, fully connected layers, and L2 regularization, was described in chapter 8. We use the same network structure for this experiment.

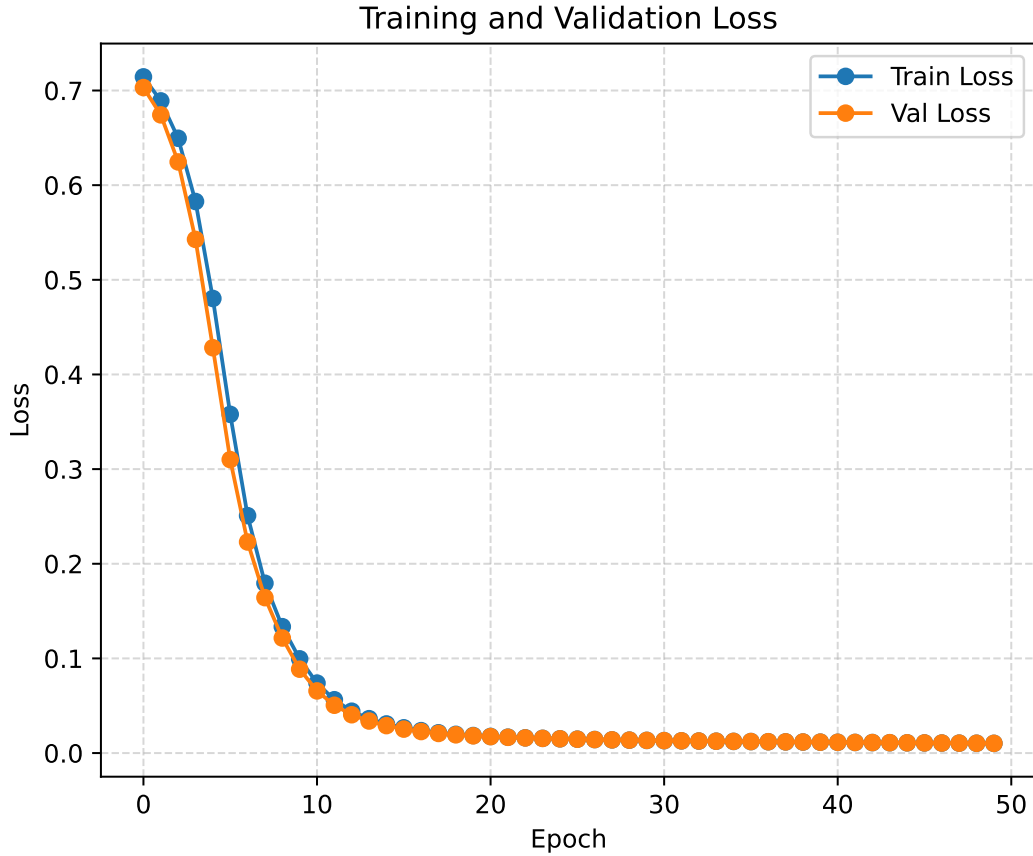
#### 9.3.2 Results for $\beta = 1$

- **Training loss:** The training and validation loss over 50 epochs is shown in Figure 9.2. The loss decreases steadily and stabilizes, indicating that the model converges without overfitting.
- **Classification predictions:** Figure 9.3 displays the predicted probabilities for test samples. The GR and BGR signals are clearly separated, and the classifier assigns



confident scores to both classes.

- **Confusion matrix:** As seen in Figure 9.4, the classifier correctly labels all test samples. No misclassifications occur at  $\beta = 1$ .



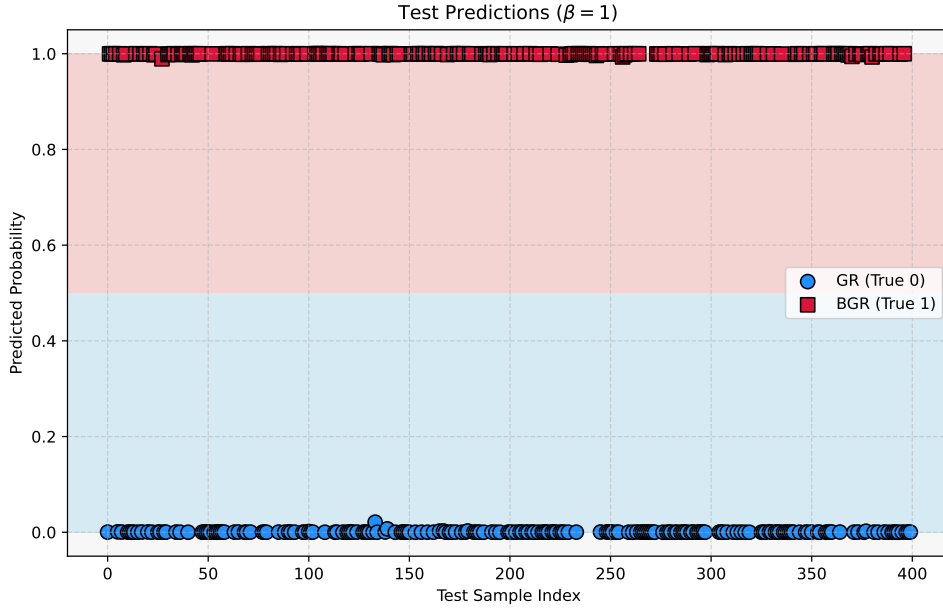
**Figure 9.2.** Training and validation loss for  $\beta = 1$ . The loss curves converge smoothly, showing stable and well-regularized learning.

### 9.3.3 Evaluation Metrics and Threshold Criterion

To quantify the classifier's performance beyond visual inspection, we compute several standard metrics used in binary classification tasks: accuracy, precision, recall, and the F1 score [77]. These are defined as follows:

- **Accuracy:**

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}, \quad (9.6)$$



**Figure 9.3.** Test predictions for  $\beta = 1$ . The classifier cleanly separates GR (blue) and BGR (red) events. The background shading denotes the classification threshold at 0.5.

where  $TP$  and  $TN$  are the number of true positives and true negatives, and  $FP$  and  $FN$  are false positives and false negatives, respectively. Accuracy measures the overall fraction of correctly classified examples.

- **Precision:**

$$\text{Precision} = \frac{TP}{TP + FP}, \quad (9.7)$$

which quantifies the proportion of events classified as BGR that are actually BGR.

- **Recall (Sensitivity):**

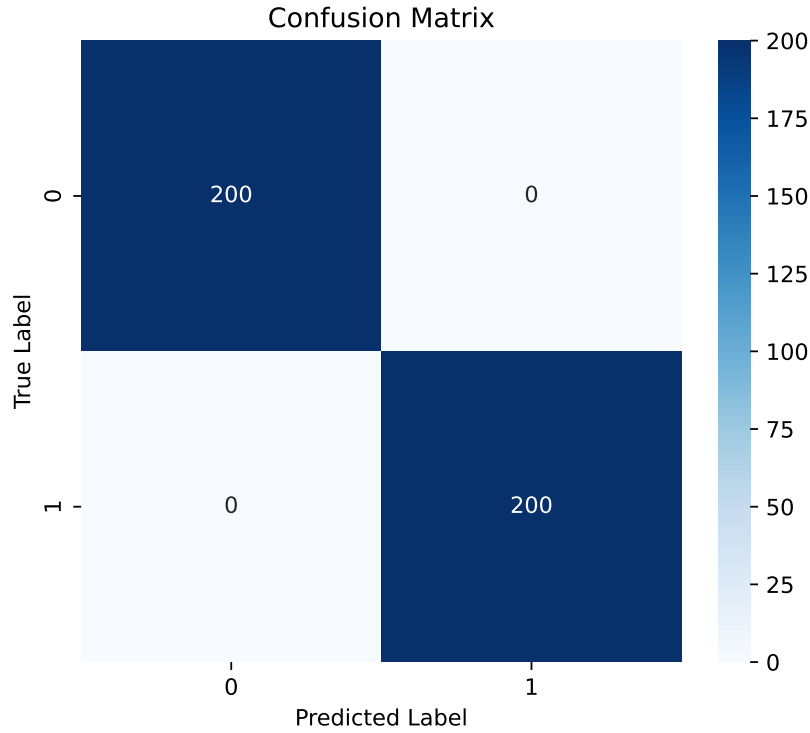
$$\text{Recall} = \frac{TP}{TP + FN}, \quad (9.8)$$

representing the proportion of actual BGR signals that were correctly identified by the model.

- **F1 Score:**

$$\text{F1} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}. \quad (9.9)$$

The F1 score is the harmonic mean of precision and recall. It is particularly useful when dealing with imbalanced datasets where one class is significantly underrepresented.



**Figure 9.4.** Confusion matrix for the test set at  $\beta = 1$ . The model achieves perfect classification.

In our setup, the dataset is perfectly balanced: we include an equal number of GR and BGR waveforms. Therefore, accuracy provides a direct and interpretable measure of the model’s effectiveness, and we adopt it as our primary performance metric. Specifically, we define a successful classification as one where the test accuracy exceeds 95%.

Although the F1 score is often used in situations where class imbalance skews the interpretation of accuracy, we compute it here for completeness. Since our model makes no misclassifications in this case, all performance metrics reach their maximum values:

$$\text{Accuracy} = 1.0000,$$

$$\text{Precision} = 1.0000,$$

$$\text{Recall} = 1.0000,$$

$$\text{F1 Score} = 1.0000.$$

These results confirm that the neural network perfectly distinguishes GR and BGR

waveforms for strong deformations ( $\beta = 1$ ). This establishes a baseline for the smallest detectable deviation that the model can reliably classify in this toy model.

## 9.4 Sensitivity to Smaller Phase Shifts

Having established that the neural network successfully classifies strongly modified waveforms at  $\beta = 1$ , we now investigate the model’s sensitivity to progressively smaller deviations from GR. In this section, we fix the network architecture and training procedure, and vary only the deformation strength  $\beta$ . Our goal is to identify the smallest detectable deviation that satisfies our performance threshold.

**Detection criterion:** As introduced earlier, we define a successful classification as one in which the test accuracy exceeds 95%. Since our dataset is balanced, accuracy provides a direct and interpretable performance measure. The F1 score is also computed for completeness, although it coincides closely with accuracy in this case.

### Classification Results for Selected $\beta$ Values

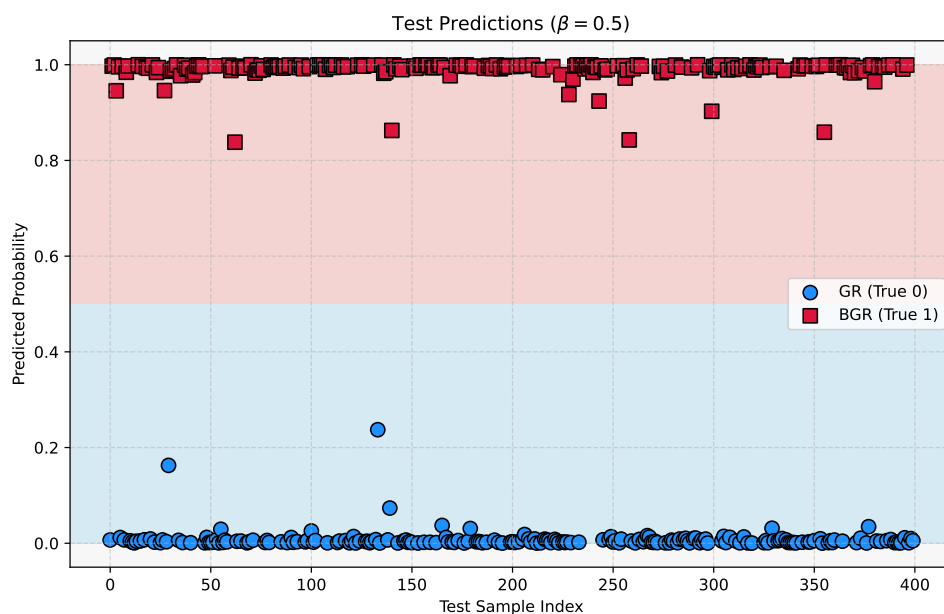
To illustrate the performance trend, we evaluate the model at several values of  $\beta$  between 1.0 and 0.28. The classification plots for each case are shown in Figures 9.5–9.10, and the final confusion matrix at the 95% threshold is shown in Figure 9.11.

Each figure shows the predicted probabilities for the test samples, with background shading indicating the classification threshold at 0.5. As  $\beta$  decreases, the predicted classes become increasingly mixed near the decision boundary, reflecting the reduced distinguishability of the BGR waveforms.

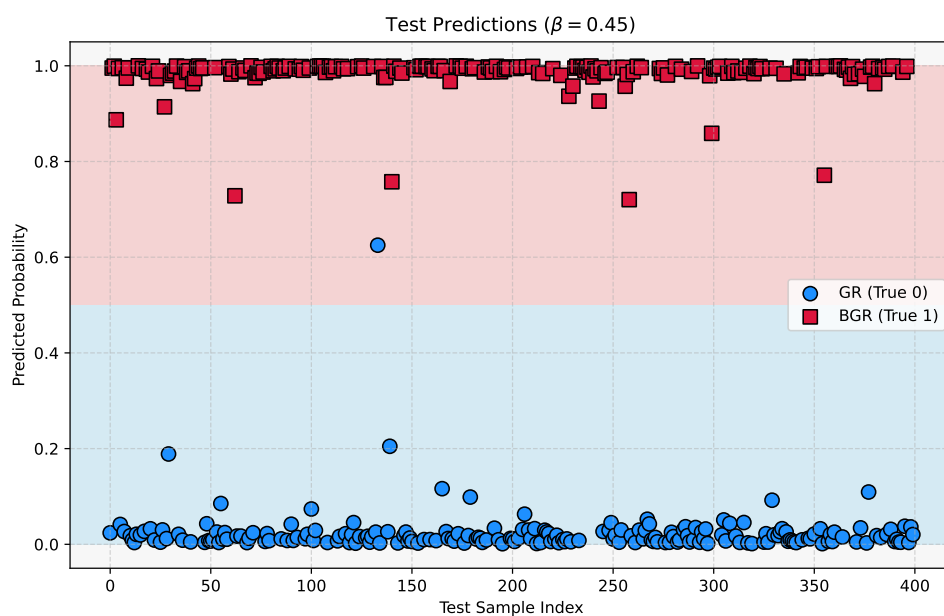
Table 9.1 summarizes the classification performance for each value of  $\beta$ , along with the average mismatch between GR and BGR waveforms. The results show a smooth degradation in performance as the deformation strength decreases. Based on our 95% accuracy criterion, we identify  $\beta = 0.28$  as the smallest reliably detectable phase deformation in this toy model.

## Discussion and Outlook

The results presented in Table 9.1 demonstrate that the neural network can reliably detect phase deformations down to  $\beta = 0.28$ , corresponding to an average mismatch of approx-

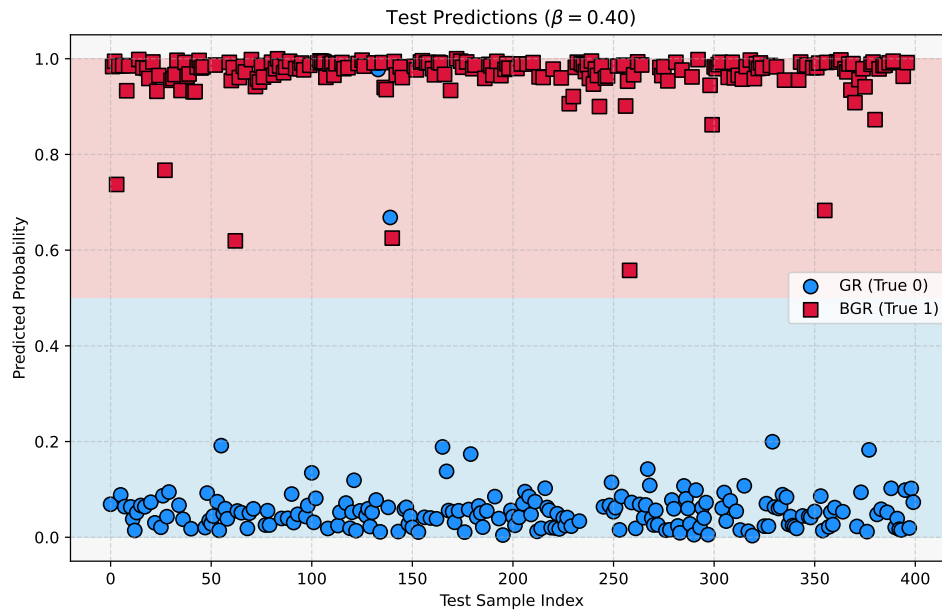


**Figure 9.5.** Test predictions for  $\beta = 0.5$ . The classifier maintains perfect separation.

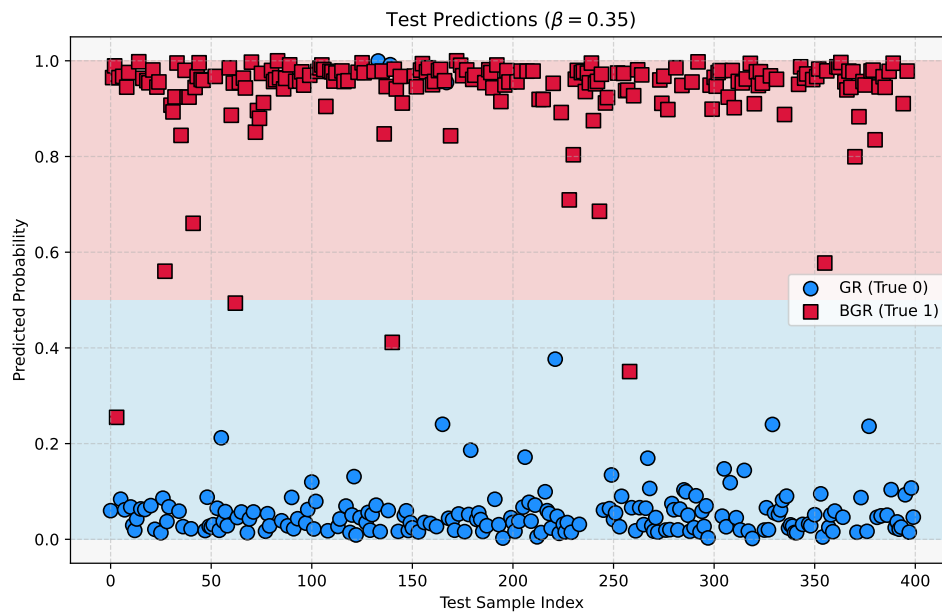


**Figure 9.6.** Test predictions for  $\beta = 0.45$ : miss-classification starts to kick in at this scale.

imately 0.0012. Below this threshold, the classifier's performance drops below our 95% accuracy criterion, making the BGR waveforms statistically indistinguishable from their

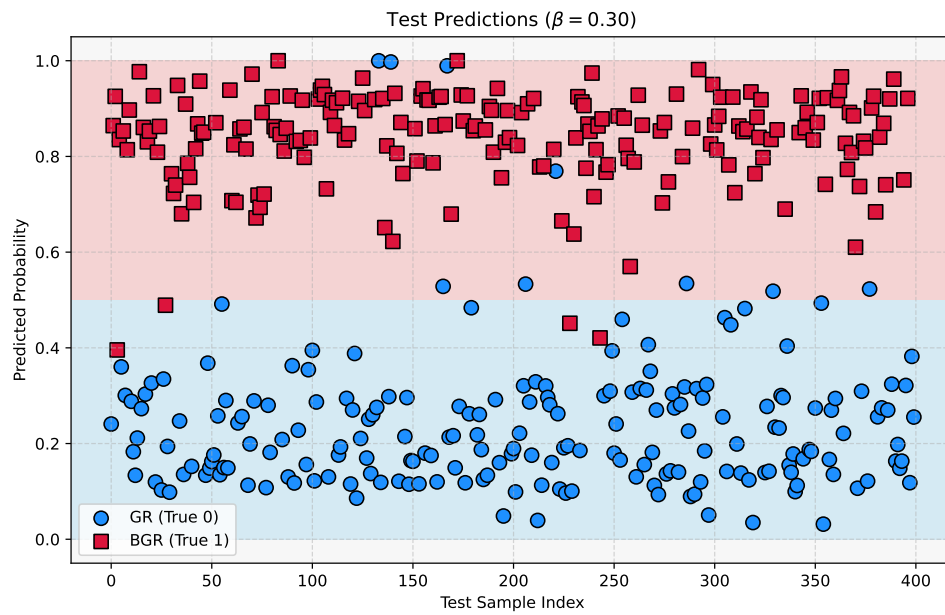


**Figure 9.7.** Test predictions for  $\beta = 0.4$ .

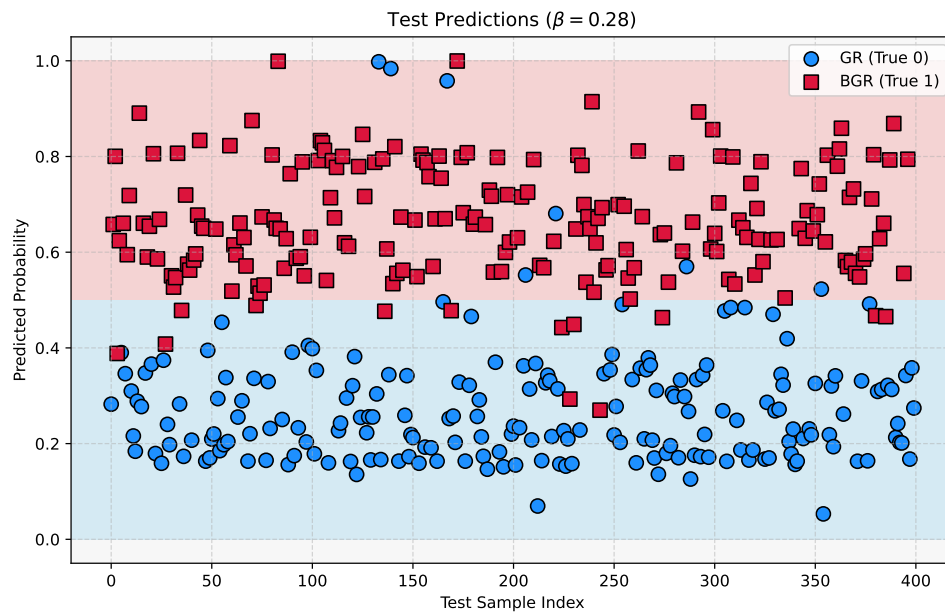


**Figure 9.8.** Test predictions for  $\beta = 0.35$ .

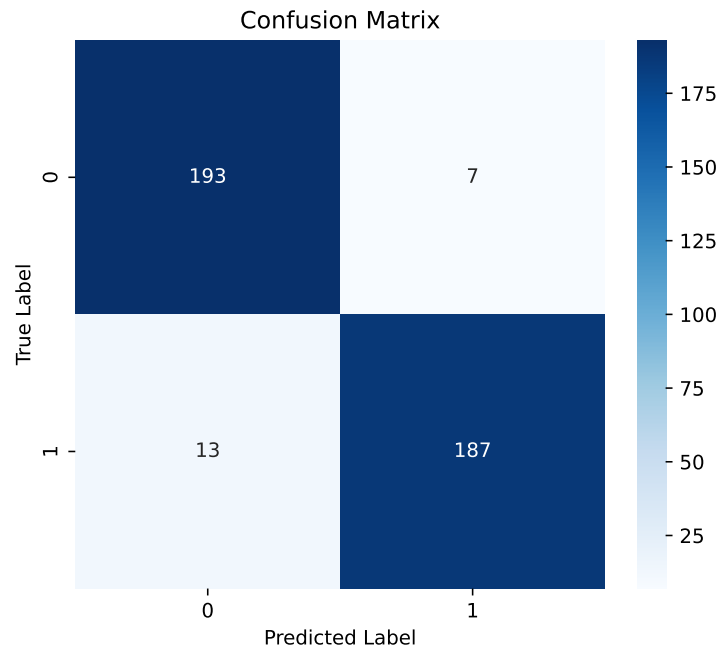
GR counterparts under the current setup. This boundary provides a practical measure of the smallest detectable deviation in our toy model.



**Figure 9.9.** Test predictions for  $\beta = 0.3$ .



**Figure 9.10.** Test predictions for  $\beta = 0.28$ . This marks the boundary of detectability as defined by the 95% accuracy threshold.



**Figure 9.11.** Confusion matrix for  $\beta = 0.28$ . This is the lowest deformation for which the classifier still exceeds 95% test accuracy.

**Table 9.1.** Classification performance metrics and average mismatch for various values of  $\beta$ . Boldface indicates the threshold value where test accuracy drops to 95%.

$\beta$	Accuracy	Precision	Recall	F1 Score	Avg. Mismatch
1.00	1.0000	1.0000	1.0000	1.0000	0.0140
0.50	1.0000	1.0000	1.0000	1.0000	0.0037
0.45	0.9975	0.9950	1.0000	0.9975	0.0030
0.40	0.9950	0.9901	1.0000	0.9950	0.0024
0.35	0.9825	0.9849	0.9800	0.9825	0.0018
0.30	0.9675	0.9561	0.9800	0.9679	0.0014
<b>0.28</b>	<b>0.9500</b>	<b>0.9639</b>	<b>0.9350</b>	<b>0.9492</b>	<b>0.0012</b>

It is important to note that these results are model-specific: they depend on the waveform content (limited here to the (2, 2) mode), the fixed inclination angle, and the specific form of the deformation applied. In more realistic scenarios, the sensitivity could shift depending on these factors.

In the next section, we move beyond toy models and explore a more systematic param-



eterization of deviations from GR using the post-Newtonian (PN) formalism [78, 71]. We investigate how small changes in PN phase coefficients affect waveform classification and analyze the detectability of such deviations using the same neural network pipeline.

## 9.5 Deviation in the Post-Newtonian Phase Coefficients

Having explored toy-model deformations using localized phase shifts, we now turn to a more structured and physically motivated approach. In this section, we examine how deviations in the post-Newtonian (PN) phase parameters affect the gravitational waveform, and study whether such changes can be detected by our classification pipeline [66, 69].

We consider the frequency-domain representation of a gravitational waveform [66, 71, 79] mode  $(\ell, m)$ , given by:

$$\tilde{h}_{\ell m}^R(f) = A_{\ell m}(f) e^{i\psi_{\ell m}(f)}, \quad (9.10)$$

where  $A_{\ell m}(f)$  is the amplitude and  $\psi_{\ell m}(f)$  is the phase of the  $(\ell, m)$ -mode.

During the quasi-circular, adiabatic inspiral [71], the phase  $\psi_{\ell m}(f)$  can be computed using PN theory under the stationary phase approximation (SPA). In General Relativity (GR), the phase takes the form:

$$\psi_{\ell m}^{(\text{GR})}(f, \lambda) = \frac{3}{128 \eta v^5} \cdot \frac{m}{2} \left[ \sum_{n=0}^7 \psi_n^{(\text{PN})}(\lambda) v^n + \sum_{n=5}^6 \psi_{n(l)}^{(\text{PN})}(\lambda) v^n \log v \right]. \quad (9.11)$$

where  $\lambda$  denotes the binary parameters, and  $v$  is the PN expansion parameter (related to the orbital velocity), defined as

$$v \equiv (2\pi F M)^{1/3} = \left( \frac{2\pi f M}{m} \right)^{1/3}. \quad (9.12)$$

Here,  $F$  is the orbital frequency, which relates to the gravitational wave frequency  $f$  for a given  $(\ell, m)$ -mode. The coefficients  $\psi_n^{(\text{PN})}$  and  $\psi_{n(l)}^{(\text{PN})}$  are known PN contributions up to 3.5PN order, which depend explicitly on the binary's physical parameters (e.g., masses, spins).

### Introducing Deviations in PN Coefficients

To introduce a deviation from GR in a controlled way, we apply a small shift to a single PN coefficient in the GR phase expression (9.11). Specifically, we consider [69]:

$$\psi_n^{(\text{PN})} \longrightarrow \psi_n^{(\text{PN})} + \delta\psi_n^{(\text{PN})}, \quad (9.13)$$

or, for the logarithmic terms,

$$\psi_{n(l)}^{(\text{PN})} \longrightarrow \psi_{n(l)}^{(\text{PN})} + \delta\psi_{n(l)}^{(\text{PN})}. \quad (9.14)$$

Plugging these shifts into Eq. (9.11), the modified BGR phase becomes

$$\psi_{\ell m}^{(\text{BGR})}(f) = \psi_{\ell m}^{(\text{GR})}(f) + \frac{3}{128\eta v^5} \cdot \frac{m}{2} \left[ \sum_n v^n \delta\psi_n^{(\text{PN})} + \sum_n v^n \log v \delta\psi_{n(l)}^{(\text{PN})} \right]. \quad (9.15)$$

We identify the second term as the deviation from the GR phase, and define

$$\delta\psi_{\ell m}^{(n)}(f) = \frac{3}{128\eta v^5} \frac{m}{2} v^n \delta\psi_n^{(\text{PN})}, \quad (9.16)$$

for a deviation in the  $n$ -th PN term, and

$$\delta\psi_{\ell m}^{(n,\log)}(f) = \frac{3}{128\eta v^5} \frac{m}{2} v^n \log v \delta\psi_{n(l)}^{(\text{PN})}, \quad (9.17)$$

for a deviation in a logarithmic PN term.

These expressions define  $\delta\psi_{\ell m}(f)$  in the physically motivated PN-based BGR model and can be used directly in Eq. (8.29) to construct modified waveforms. In the following section, we implement such deformations and investigate whether the classifier can detect them.

**Focusing on the dominant (2, 2) mode.** In the remainder of this analysis, we restrict our attention to the  $(\ell, m) = (2, 2)$  mode, which dominates the signal during the inspiral phase of quasi-circular, non-spinning binaries. Setting  $m = 2$  in the expressions above, the phase deformation becomes:

$$\delta\psi_{22}^{(n)}(f) = \frac{3}{128\eta v^5} \cdot v^n \delta\psi_n^{(\text{PN})}, \quad (9.18)$$

and

$$\delta\psi_{22}^{(n,\log)}(f) = \frac{3}{128\eta v^5} \cdot v^n \log v \delta\psi_{n(l)}^{(\text{PN})}. \quad (9.19)$$

In our implementation, we treat each  $\delta\psi_n^{(\text{PN})}$  as a free numerical parameter, analogous to the deformation strength  $\beta$  used in the toy model. Each such parameter governs the size of a deviation in a particular PN order. This allows us to construct beyond-GR (BGR) waveforms by introducing controlled, interpretable shifts in the phase structure.

One of the strengths of this formalism is its flexibility: rather than modifying a single PN term, one can simultaneously introduce deviations in multiple PN orders. In principle,

this opens the door to a more general framework, where a combination of  $\delta\psi_n^{(\text{PN})}$  and  $\delta\psi_{n(l)}^{(\text{PN})}$  values can be explored. A systematic grid search over these parameters could, for example, identify combinations that lead to detectable waveform deviations — i.e., those that cross the mismatch threshold or exceed a given classification confidence.

This structured approach provides a direct bridge between the data-driven analysis and the underlying physical theory, enabling us to probe not just the existence of deviations from GR, but also where in the PN expansion such deviations might lie.

## Multiclass Classification via Deviations in PN Coefficients

Having established the formalism for introducing deviations in individual PN coefficients, we now move beyond binary classification and consider a multiclass setup. Specifically, instead of simply distinguishing between GR and BGR waveforms, we aim to determine which particular PN coefficient has been modified.

We construct four waveform classes:

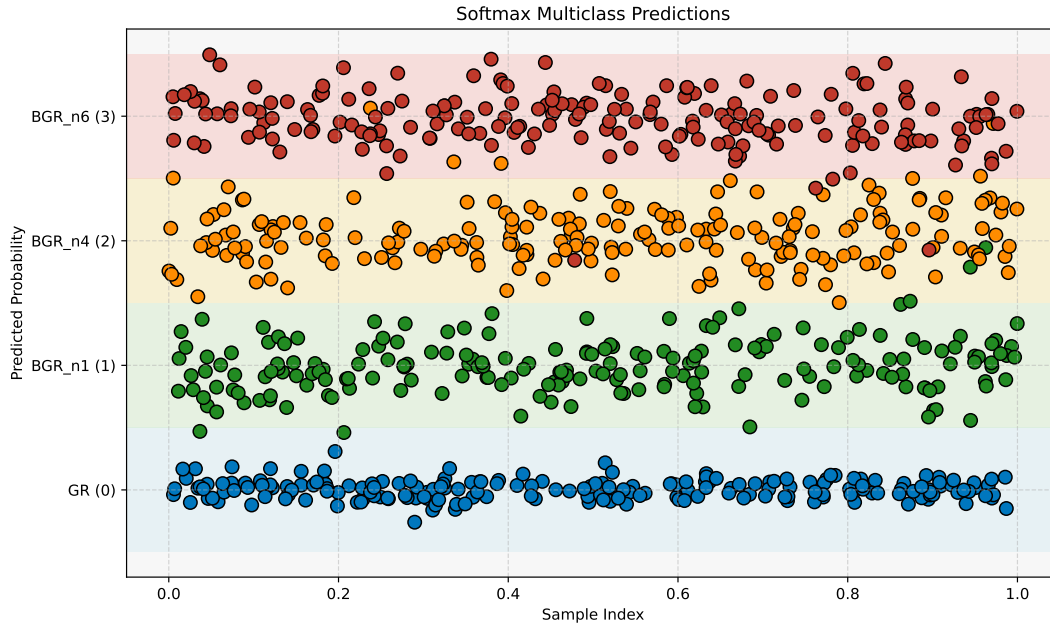
- **Class 0:** GR waveforms with no deformation.
- **Class 1:** BGR waveforms with a deviation in the  $n = 1$  PN coefficient (i.e., the 0.5PN term).
- **Class 2:** BGR waveforms with a deviation in the  $n = 4$  PN coefficient (i.e., the 2PN term).
- **Class 3:** BGR waveforms with a deviation in the  $n = 6$  PN coefficient (i.e., the 3PN term).

Each of these deviations corresponds to a phase shift of the form:

$$\delta\psi_{22}^{(n)}(f) = \frac{3}{128\eta v^5} \cdot v^n \delta\psi_n^{(\text{PN})}, \quad (9.20)$$

where  $v = (2\pi f M/m)^{1/3}$  and  $m = 2$  for the  $(2, 2)$  mode. Only one coefficient is perturbed at a time in each case, and the others are left unchanged. This allows us to interpret the classifier's decision as identifying the specific PN structure being violated.

To address this task, we adopt a softmax-based classification approach, as introduced in Section 7.5. The neural network is trained to assign a probability distribution over the four possible classes, providing not only a prediction but also a confidence level for each



**Figure 9.12.** Softmax classification results for PN-deformed waveforms. The network was trained to distinguish between GR waveforms and BGR waveforms constructed by modifying the  $n = 1$ ,  $n = 4$ , or  $n = 6$  PN coefficients individually. The y-axis corresponds to the predicted class probabilities for each sample in the test set. Colored regions and markers indicate the classifier’s output. The model successfully learns to differentiate the origin of each waveform, identifying the PN order of the deviation.

classification. This probabilistic output enables us to construct a colorful visualization of the classifier’s behavior on the test set, which is shown in Figure 9.12.

This result highlights the potential of softmax classifiers not only for detection of deviations from GR, but also for pinpointing the nature of the deviation in a physically interpretable way. Such capabilities may be particularly valuable for future gravitational wave data analysis pipelines, where explainability and physical grounding of decisions will be crucial.

The PN-based formalism described in this section is especially powerful because it connects the data-driven classification approach directly to the structure of the underlying theory. By modifying specific PN coefficients, we mimic the leading-order effects that arise in modified gravity theories such as Horndeski, beyond-Horndeski, and Generalized Proca theories. Many of these models have well-established PN expansions derived from their field equations. Therefore, a classifier trained in this framework could, in principle, be used not only to detect a deviation from GR but also to constrain or identify viable

beyond-GR theories based on their PN imprint.

This flexibility and theoretical grounding make the PN-based classification framework a promising tool for gravitational wave phenomenology in the era of precision GW astronomy.

**Summary and Outlook.** In this chapter, we demonstrated the effectiveness of a neural network classifier in distinguishing between GR and beyond-GR (BGR) waveforms using controlled phase deformations. By analyzing both a Gaussian toy model and structured deviations in post-Newtonian (PN) coefficients, we established the minimum strength of modification that remains reliably detectable under our current framework. These results highlight the sensitivity of the model and provide a baseline for classification performance when using the waveform itself as input.

However, classifying directly from the waveform has certain limitations, especially when the distinguishing features between GR and BGR signals are subtle, buried in noise, or affected by observational degeneracies. In such cases, the waveform itself may not provide the most informative structure for the classification task. To address this, the next chapter introduces a more general and theoretically motivated framework: the *response function*. While in our specific implementation the response function is related to the match and mismatch between signals, the formalism itself is much broader and can be adapted to a variety of contexts. We will see that this alternative representation captures more discriminative features and enables a significant improvement in performance.



# Chapter 10

## Response Function Formalism

The goal of this chapter is to go beyond waveform-based classification by introducing the response function [69] as a refined input to the neural network. Although in this work we focus on a version of the response function related to the match between waveforms, the formalism is general and can be adapted to different scenarios. We will show that using the response function leads to improved accuracy and a lower detection threshold for deviations from GR. In particular, our results indicate that the classification performance approaches the Bayes error limit—the theoretical boundary for any statistical decision process—and in some regimes even outperforms human-level discrimination. This shift marks a key step toward more sensitive and interpretable gravitational wave classification methods. The response function approach was originally introduced in a cosmological context in [80, 81], where it was used to analyze deformations of the Hubble parameter and their impact on other observables.

### 10.1 Theoretical Framework and Response Function Formalism

To improve the performance and interpretability of waveform classification, we now turn to a more structured representation based on the concept of the *response function*. This formalism allows us to study how physical parameters and waveform deformations influence observable quantities in a systematic way.

Let us begin by revisiting Eq. (8.17), which expresses the relative shift in the phase and amplitude of a given  $(\ell, m)$  mode in terms of intrinsic waveform deformations  $\delta\psi_{\ell m}$  and  $\delta A_{\ell m}$ , as well as shifts in the physical parameters of the binary system:

$$\begin{aligned}\Delta\psi_{\ell m} &= \delta\psi_{\ell m} + \left(\eta \frac{\partial\psi_{\ell m}^{(GR)}}{\partial\eta}\right) \frac{\delta\eta}{\eta} + \left(M \frac{\partial\psi_{\ell m}^{(GR)}}{\partial M}\right) \frac{\delta M}{M} + \left(\chi_S \frac{\partial\psi_{\ell m}^{(GR)}}{\partial\chi_S}\right) \frac{\delta\chi_S}{\chi_S} + \left(\chi_A \frac{\partial\psi_{\ell m}^{(GR)}}{\partial\chi_A}\right) \frac{\delta\chi_A}{\chi_A}, \\ \Delta A_{\ell m} &= \delta A_{\ell m} + \left(\eta \frac{\partial A_{\ell m}^{(GR)}}{\partial\eta}\right) \frac{\delta\eta}{\eta} + \left(M \frac{\partial A_{\ell m}^{(GR)}}{\partial M}\right) \frac{\delta M}{M} + \left(\chi_S \frac{\partial A_{\ell m}^{(GR)}}{\partial\chi_S}\right) \frac{\delta\chi_S}{\chi_S} + \left(\chi_A \frac{\partial A_{\ell m}^{(GR)}}{\partial\chi_A}\right) \frac{\delta\chi_A}{\chi_A}.\end{aligned}$$

This structure naturally motivates the introduction of frequency-dependent response functions that characterize how changes in the waveform phase and amplitude propagate into changes in the binary parameters. We define the following integral relations:

$$\begin{aligned}\frac{\delta\eta}{\eta} &= \int df R_\eta(f) \delta\psi_{\ell m}(f) + \int df \tilde{R}_\eta(f) \delta A_{\ell m}(f), \\ \frac{\delta M}{M} &= \int df R_M(f) \delta\psi_{\ell m}(f) + \int df \tilde{R}_M(f) \delta A_{\ell m}(f), \\ \frac{\delta\chi_S}{\chi_S} &= \int df R_{\chi_S}(f) \delta\psi_{\ell m}(f) + \int df \tilde{R}_{\chi_S}(f) \delta A_{\ell m}(f), \\ \frac{\delta\chi_A}{\chi_A} &= \int df R_{\chi_A}(f) \delta\psi_{\ell m}(f) + \int df \tilde{R}_{\chi_A}(f) \delta A_{\ell m}(f),\end{aligned}\tag{10.1}$$

where

$$R_\eta(f), \quad R_M(f), \quad R_{\chi_S}(f), \quad R_{\chi_A}(f)$$

are the phase response functions, and

$$\tilde{R}_\eta(f), \quad \tilde{R}_M(f), \quad \tilde{R}_{\chi_S}(f), \quad \tilde{R}_{\chi_A}(f)$$

are the amplitude response functions. These functions encode how perturbations in the waveform structure map onto shifts in the system's physical parameters.

Once the response functions are known, we can generalize the formalism to study how any observable quantity  $O(f)$  responds to deformations in the waveform. Assuming that  $O$  depends on the underlying parameters and waveform features, its total shift can be expressed as:

$$\begin{aligned}\Delta O(f) &= I_\eta(f) \frac{\delta\eta}{\eta} + I_M(f) \frac{\delta M}{M} + I_{\chi_S}(f) \frac{\delta\chi_S}{\chi_S} + I_{\chi_A}(f) \frac{\delta\chi_A}{\chi_A} \\ &\quad + \int_{f_{\ell m}^{\text{ref}}}^{f_{22}^{\text{peak}}} dx R_O(x, f) \delta\psi_{\ell m}(x) + \int_{f_{\ell m}^{\text{ref}}}^{f_{22}^{\text{peak}}} dx \tilde{R}_O(x, f) \delta A_{\ell m}(x),\end{aligned}\tag{10.2}$$

where  $I_i(f)$  are weight functions dependent on the system parameters and the frequency  $f$ . Combining the parameter shifts with their respective response functions, we can write the total shift in a compact convolutional form:



$$\Delta O(f) = \int_{f_{\ell m}^{\text{ref}}}^{f_{22}^{\text{peak}}} dx \mathcal{R}_O(x, f) \delta\psi_{\ell m}(x) + \int_{f_{\ell m}^{\text{ref}}}^{f_{22}^{\text{peak}}} dx \tilde{\mathcal{R}}_O(x, f) \delta A_{\ell m}(x), \quad (10.3)$$

with the effective response kernels defined as:

$$\begin{aligned} \mathcal{R}_O(x, f) &= I_\eta(f) R_\eta(x) + I_M(f) R_M(x) + I_{\chi_S}(f) R_{\chi_S}(x) + I_{\chi_A}(f) R_{\chi_A}(x) + R_O(x, f), \\ \tilde{\mathcal{R}}_O(x, f) &= I_\eta(f) \tilde{R}_\eta(x) + I_M(f) \tilde{R}_M(x) + I_{\chi_S}(f) \tilde{R}_{\chi_S}(x) + I_{\chi_A}(f) \tilde{R}_{\chi_A}(x) + \tilde{R}_O(x, f). \end{aligned} \quad (10.4)$$

Here,  $f_{\ell m}^{\text{ref}}$  and  $f_{22}^{\text{peak}}$  represent the lower and upper frequency bounds over which the deformation takes effect. Although this general expression includes contributions from changes in system parameters, in the following sections we will focus solely on the contributions arising directly from the waveform deformations themselves, namely  $\delta\psi_{\ell m}$ .

### 10.1.1 Boundary Conditions

So far, we have introduced a general framework that allows for both phase  $\delta\psi_{\ell m}(f)$  and amplitude  $\delta A_{\ell m}(f)$  extensions, as well as possible shifts in physical parameters like  $\eta$ . However, in many practical scenarios of early-inspiral waveform analysis, the *phase deformation* tends to dominate the observable signature. This is because phase accumulates over many cycles, often revealing small deviations more clearly than amplitude shifts, which are subject to larger observational uncertainties [82].

Moreover, a common simplifying assumption is that beyond-GR (BGR) features vanish by the time the frequency reaches  $f_{22}^{\text{peak}}$ , effectively resetting the waveform to its GR form in the high-frequency regime [82, 70]. Here,  $f_{22}^{\text{peak}}$  denotes the frequency at which the amplitude of the dominant  $(\ell, m) = (2, 2)$  mode reaches its maximum—typically marking the transition from the inspiral phase to the merger. This frequency serves as a natural boundary beyond which the inspiral-based waveform modeling is no longer valid, and any deviations from GR are assumed to be negligible. Mathematically, this boundary condition can be written as

$$\Delta\psi_{\ell m}(f_{22}^{\text{peak}}) = 0,$$

where

$$\Delta\psi_{\ell m}(f) = \delta\psi_{\ell m}(f) + g_{\ell m}(f) \frac{\delta\eta}{\eta}$$

captures both the *direct* phase shift  $\delta\psi_{\ell m}(f)$  and any *parameter-induced* shift proportional to  $\delta\eta/\eta$ . Here,  $g_{\ell m}(f)$  is determined by the derivative of the GR phase with respect to  $\eta$ .

Hence, at  $f_{22}^{\text{peak}}$ ,

$$0 = \delta\psi_{\ell m}(f_{22}^{\text{peak}}) + g_{\ell m}(f_{22}^{\text{peak}}) \frac{\delta\eta}{\eta},$$

which implies

$$\frac{\delta\eta}{\eta} = -\frac{\delta\psi_{\ell m}(f_{22}^{\text{peak}})}{g_{\ell m}(f_{22}^{\text{peak}})}. \quad (10.5)$$

For the types of waveform deformations we consider—such as a localized Gaussian shift in the low-frequency regime (see Sec. 9.1)—the function  $\delta\psi_{\ell m}(f)$  is constructed to approach zero (often exponentially) as  $f \rightarrow f_{22}^{\text{peak}}$ . Consequently, Eq. (10.5) forces  $\delta\eta/\eta \approx 0$ . Put differently, if the phase deviation itself vanishes at high frequencies, there is no leftover constant shift in  $\eta$  to compensate for a residual mismatch.

Under this boundary condition, the overall BGR imprint becomes purely *phase-driven* in the early inspiral, and we need not consider additional parameter or amplitude shifts. In summary:

- **Phase-Dominance:** Phase accumulates over many cycles, making  $\delta\psi_{\ell m}(f)$  a more sensitive probe of small deviations than amplitude-based modifications.
- **High-Frequency Reset:** By design, the deformation is engineered to vanish by  $f_{22}^{\text{peak}}$ , leaving no net shift in physical parameters.
- **Simplicity:** Restricting to a phase-only extension reduces the dimensionality of the problem, making both analytic and numerical analyses more tractable.

Though this simplification omits potentially interesting amplitude effects or multi-parameter shifts, it is well motivated in cases where phase plays the dominant role in detecting and quantifying beyond-GR deviations, especially in the low-frequency (early-inspiral) regime.

## 10.2 Mismatch and its Response Function

A key observable in gravitational wave data analysis is the *mismatch*, which quantifies the dissimilarity between a gravitational wave signal and a theoretical model or template. Given a model waveform  $h(f)$  and a signal waveform  $s(f)$ , the *match* between them is defined as [74]:

$$\mathcal{M}(h, s) = \frac{(h | s)}{\sqrt{(h | h)(s | s)}},$$

with values ranging from 0 to 1, where 1 indicates perfect agreement. The corresponding *mismatch* is defined as

$$\overline{\mathcal{M}}(h, s) = 1 - \mathcal{M}(h, s) = 1 - \frac{(h | s)}{\sqrt{(h | h)(s | s)}}.$$

We are interested in how this mismatch changes under small deformations of the waveform model. Let  $\Delta h(f)$  denote such a variation. Then the variation in the mismatch is given by

$$\Delta \overline{\mathcal{M}}(h, s) = \frac{1}{\sqrt{(h | h)}\sqrt{(s | s)}} \left( -(\Delta h | s) + \frac{(h | s)(\Delta h | h)}{(h | h)} \right). \quad (10.6)$$

To extract more structure from this expression, we define the normalized waveforms:

$$\hat{h}(f) = \frac{h(f)}{\|h\|}, \quad \hat{s}(f) = \frac{s(f)}{\|s\|},$$

where  $\|h\| = \sqrt{(h | h)}$  and  $\|s\| = \sqrt{(s | s)}$  are the normalization factors.

The match now becomes simply

$$\mathcal{M}(h, s) = (\hat{h} | \hat{s}),$$

and the expression for the mismatch variation becomes:

$$\Delta \overline{\mathcal{M}} = -(\Delta \hat{h} | \hat{s}) + (\hat{h} | \hat{s}) \cdot (\Delta \hat{h} | \hat{h}).$$

To identify how the mismatch responds pointwise in frequency, we consider a deformation generated by the dominant (2, 2) mode of the model waveform,  $h_{22}(f)$ . We assume that the deformation takes the form

$$h_{\text{BGR}}(f) \approx h_{\text{GR}}(f) + i\beta h_{22, \text{GR}}(f) \delta\psi_{22}(f), \quad (10.7)$$

where  $\delta\psi_{22}(f)$  encodes the frequency-dependent phase deformation, and  $\beta$  is a parameter controlling the strength of the deviation from GR. This expression describes a leading-order modification to the waveform in the frequency domain and serves as the basis for constructing beyond-GR waveforms.

Using this ansatz, we define the *response function*  $R(f)$  associated with the mismatch:

$$R(f) = \frac{4i h_{22}(f)}{\|h\| S_n(f)} \left[ -\hat{s}^*(f) + \mathcal{M}(h, s) \hat{h}^*(f) \right],$$

where  $h(f)$  and  $s(f)$  are the model and signal waveforms respectively,  $S_n(f)$  is the power spectral density of the detector noise, and the hat denotes normalization.

This response function measures how sensitively the mismatch reacts to phase deformations at each frequency  $f$ , assuming that such deformations enter via Eq. (8.29). The factor of  $ih_{22}(f)$  reflects the fact that the deformation enters through a complex phase rotation generated by the  $(2, 2)$  mode.

Although in GR we expect no intrinsic deviation—i.e.,  $\delta\psi = 0$ —we can still compute the response function  $R(f)$  without assuming the underlying theory. In practice, when the data is consistent with GR, the response function exhibits no coherent structure and appears as purely stochastic noise. However, when a genuine beyond-GR deviation is present, the response function reveals clear, frequency-localized features, making it a powerful and interpretable observable for detecting and characterizing departures from General Relativity.

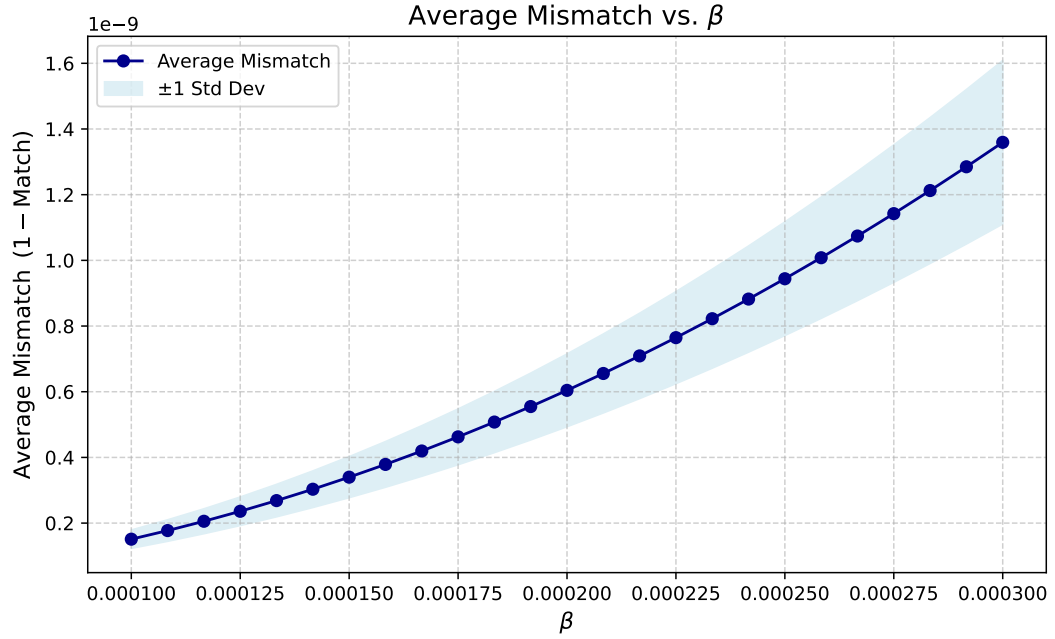
### 10.3 Gaussian Toy Model Revisited: Response Function Performance

Having established the theoretical advantages of the response function formalism, we now revisit the Gaussian toy model previously discussed. Our goal is to quantify the improved classification capability when using response functions instead of waveforms as input to the neural network. Specifically, we examine how the neural network’s accuracy changes as we lower the deformation parameter  $\beta$ .

In Fig. 10.1, we illustrate the average mismatch as a function of  $\beta$  for the response-function-based dataset. Comparing this plot to the waveform-based results presented in earlier chapters, we observe a significant enhancement in sensitivity. This improved sensitivity allows the classifier to reliably detect much smaller deviations from General Relativity, thus lowering the detection threshold considerably.

To illustrate the dramatic improvement in classification accuracy, we revisit the previously challenging scenario at  $\beta = 0.28$ . In the earlier waveform-based analysis, the neural network failed to achieve reliable classification accuracy at this level. However, as demonstrated in Fig. 10.2, the response function classifier achieves perfect classification with accuracy, precision, recall, and F1 score all equal to 1.0. This demonstrates the exceptional sensitivity and discriminative power of response functions in waveform analysis.

To further explore the limits of the response-function approach, we decrease  $\beta$  substantially to 0.0003, a regime previously unattainable for waveform-based classifiers. At this minimal deformation, Fig. 10.3 illustrates that the waveform-based classification completely fails, while the response-function-based classifier maintains high performance met-



**Figure 10.1.** Average mismatch between BGR and GR waveforms using response functions as input, plotted against the deformation parameter  $\beta$ . The shaded region indicates the standard deviation across the dataset. Compared to waveform-based classification, response functions provide a considerably lower detection threshold, allowing deviations at significantly smaller  $\beta$  values to be accurately identified.

rics:

$$\text{Accuracy} = 0.9525,$$

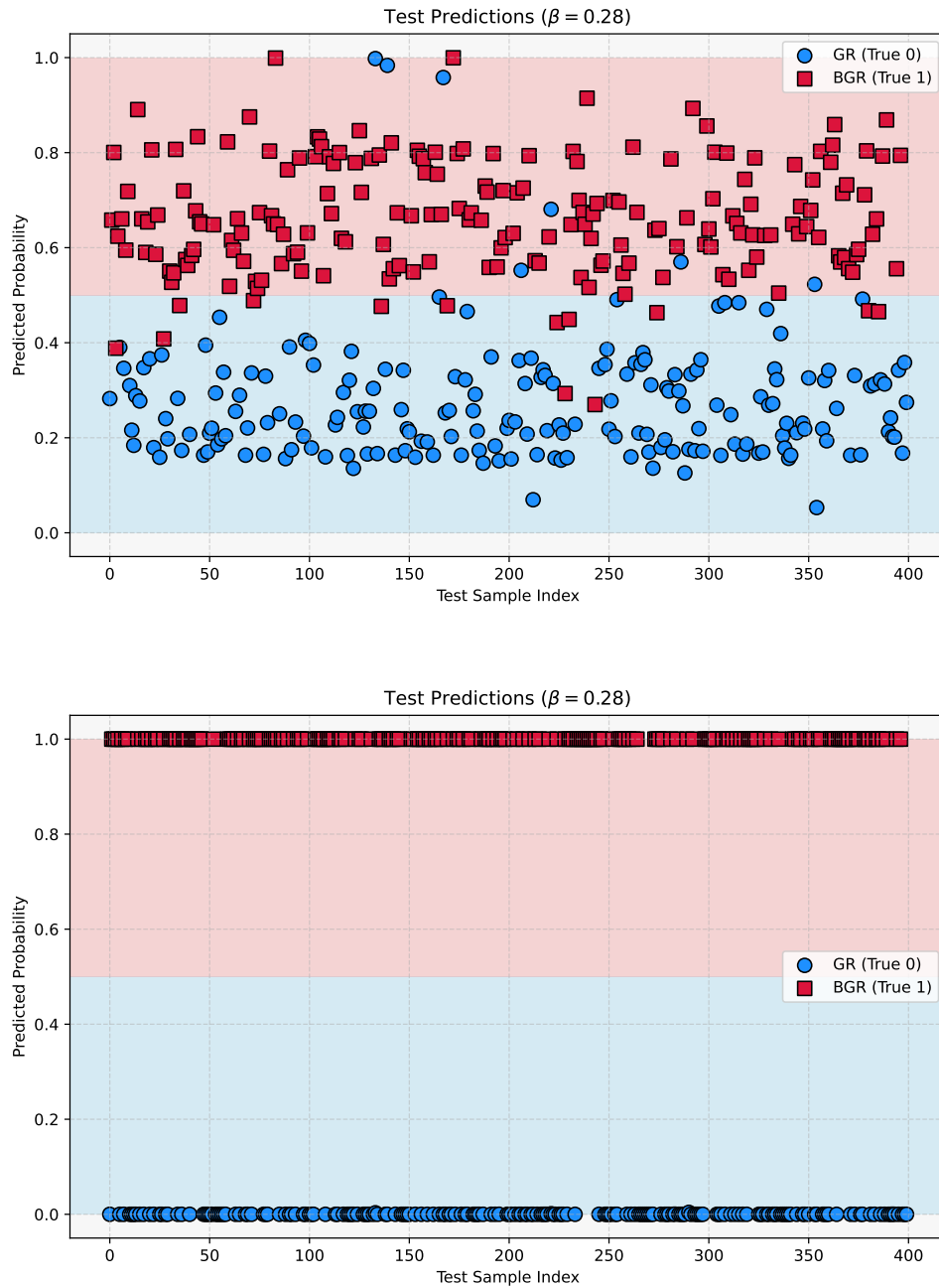
$$\text{Precision} = 0.9594,$$

$$\text{Recall} = 0.9450,$$

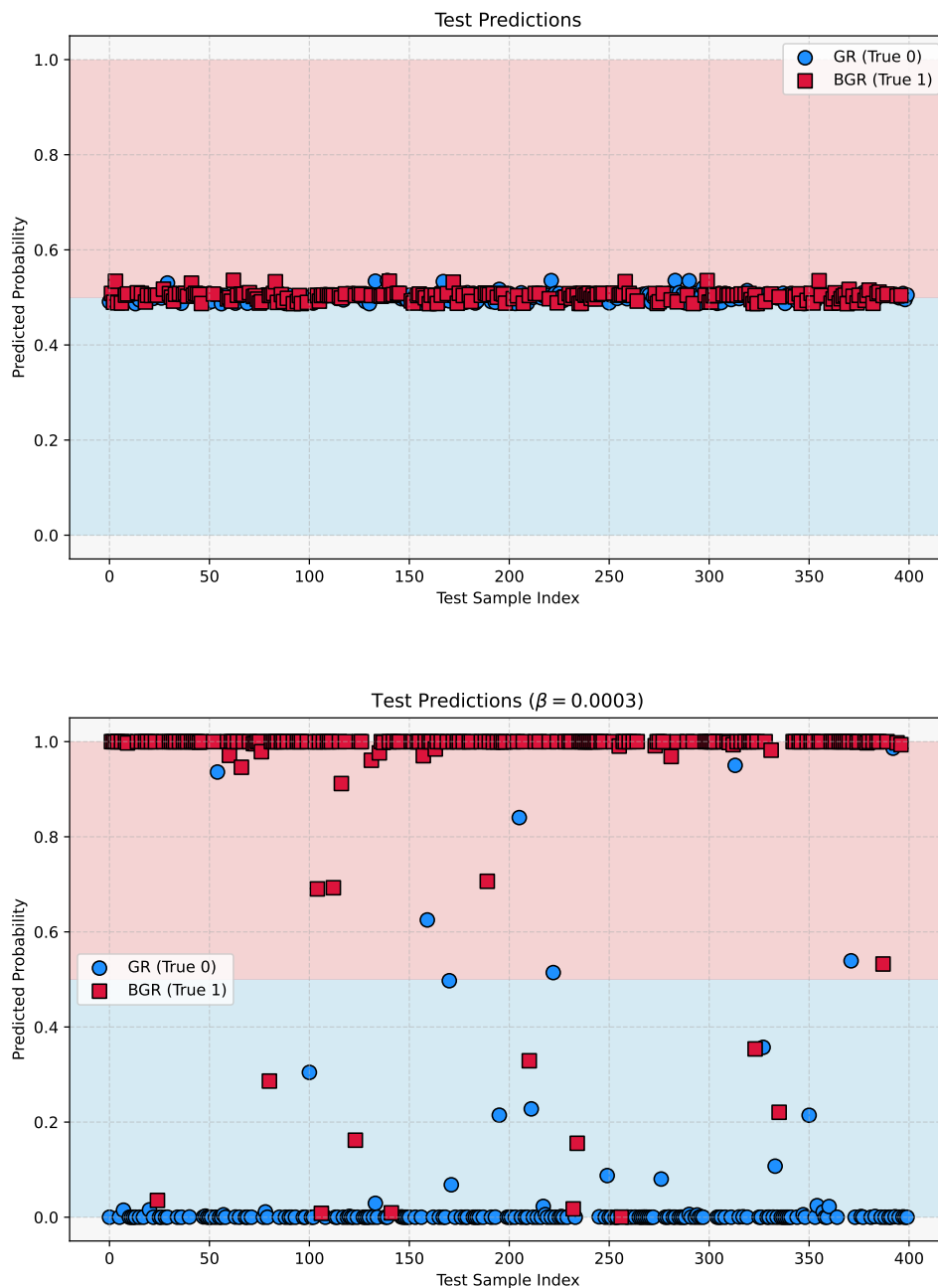
$$\text{F1 Score} = 0.9521.$$

The confusion matrix depicted in Fig. 10.4 further confirms the robustness of the classification, clearly separating GR and BGR waveforms at a deformation scale far below previous thresholds.

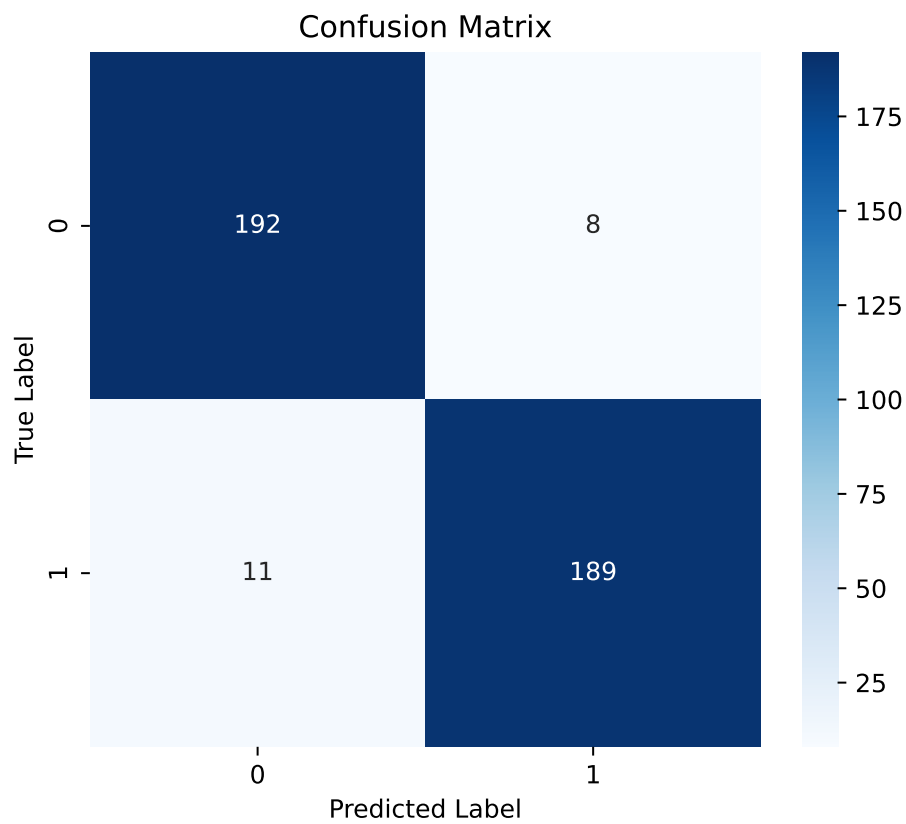
In conclusion, the introduction of the response function significantly enhances the sensitivity and reliability of gravitational waveform classification, providing a powerful and robust tool for detecting subtle beyond-GR deviations.



**Figure 10.2.** Classification comparison at  $\beta = 0.28$ . The top plot shows waveform-based classification, which fails to reliably distinguish between GR and BGR waveforms. The bottom plot demonstrates perfect classification performance using response functions, highlighting the superiority of response functions in detecting subtle deviations from GR.



**Figure 10.3.** Classification comparison at  $\beta = 0.0003$ . The top plot illustrates complete failure of waveform-based classification. The bottom plot shows response-function-based classification maintaining accuracy above 95%, underscoring the vastly improved sensitivity provided by response functions.



**Figure 10.4.** Confusion matrix for response-function classification at  $\beta = 0.0003$ . High values along the diagonal reflect the classifier's strong capability to distinguish GR from BGR waveforms even at minimal deviations.



## 10.4 Limits of Classification: Accuracy, Bayes Optimal Error, and Interpretability

In this final section, we discuss the theoretical and practical limitations inherent to any classification framework, focusing on the concepts of Bayes optimal error, human-level performance, and interpretability [83, 84]. These ideas help contextualize the effectiveness of the response function formalism, not only as a practical tool but also as a representation that approaches the theoretical limit of performance.

### Bayes Optimal Error

In any supervised classification task, the goal is to learn a function  $f : \mathcal{X} \rightarrow \mathcal{Y}$  that minimizes the expected risk under the true data distribution  $P(x, y)$ . The Bayes optimal classifier is defined as

$$f_{\text{Bayes}}(x) = \arg \max_{y \in \mathcal{Y}} P(y | x), \quad (10.8)$$

and achieves the lowest possible classification error:

$$\epsilon_{\text{Bayes}} = \mathbb{E}_x \left[ 1 - \max_{y \in \mathcal{Y}} P(y | x) \right]. \quad (10.9)$$

This quantity represents the irreducible error due to inherent noise or ambiguity in the input features. No classifier, regardless of its capacity or training data, can achieve an error lower than  $\epsilon_{\text{Bayes}}$ .

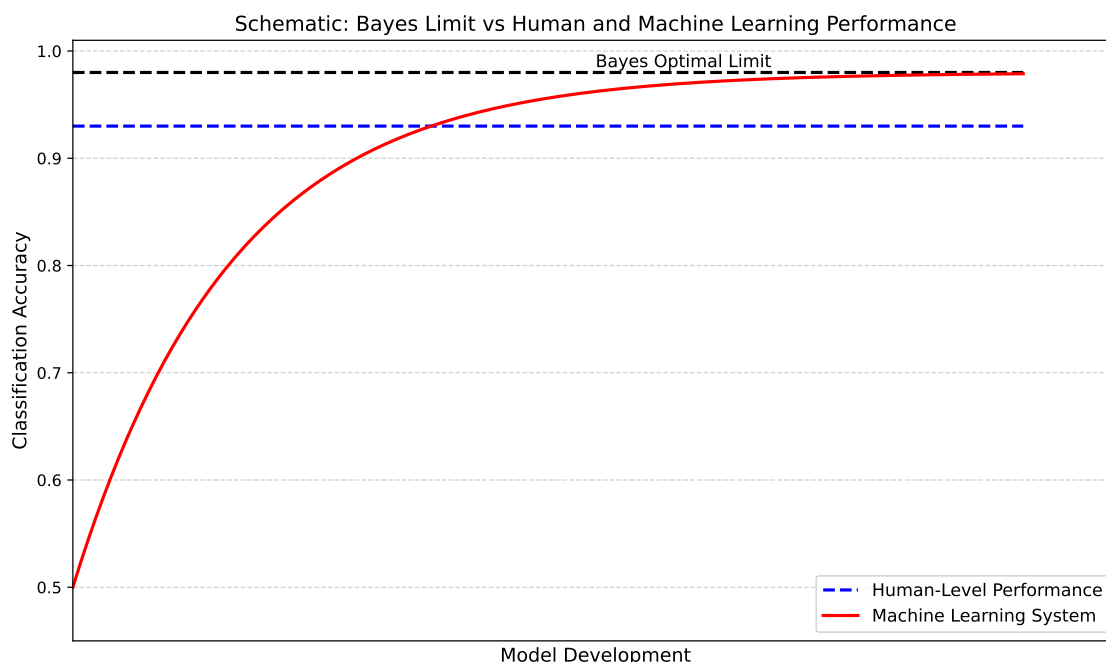
In gravitational wave analysis, such ambiguity can arise from intrinsic degeneracies between GR and BGR waveforms at low signal-to-noise ratios or due to extremely small deviations. The Bayes error thus sets a fundamental bound on what any classifier can hope to achieve, including deep learning models trained on ideal datasets.

### Human-Level Performance and the Role of Feature Representations

For tasks that humans perform well, such as image recognition or waveform discrimination, human-level accuracy provides a natural benchmark. In such cases, one typically finds

$$\epsilon_{\text{Bayes}} \lesssim \epsilon_{\text{human}} \ll 1, \quad (10.10)$$

indicating that while humans can be impressively accurate, they are not strictly optimal. So long as a machine learning system performs worse than humans, one can leverage human



**Figure 10.5.** Illustrative comparison of machine learning performance with human-level performance and the Bayes optimal error. The x-axis denotes schematic training time or model improvement, while the y-axis represents classification accuracy. Human-level accuracy typically approaches the Bayes limit, while model performance initially improves rapidly and then gradually saturates.

intuition and labeled data to guide model improvements. Once a model surpasses this level, however, such tools become less effective, and progress often slows.

Feature selection plays a central role in bridging this performance gap. In classical regression problems, the choice of features greatly affects generalization. For example, predicting house prices based on area, number of rooms, and location is meaningful, whereas using the color of the front door is likely irrelevant or even detrimental. The same principle applies in our setting: waveforms and response functions are alternative feature representations for the same underlying physical event, but their discriminative power differs significantly.

## Interpretability and the Strength of Response Functions

The response function formalism excels not only in accuracy but also in interpretability. Even for small deformation parameters such as  $\beta = 0.0003$ , the response functions exhibit clear and structured deviations between GR and BGR signals—at least in a statistical sense.

By contrast, the raw waveforms in that same regime appear nearly identical, making reliable classification difficult or impossible by eye.

We can quantify the performance of different input representations by their empirical classification error. Let

$$\epsilon_{\text{RF}} \quad \text{and} \quad \epsilon_{\text{WF}} \quad (10.11)$$

denote the classification errors using response functions and waveforms, respectively. Our empirical results indicate the following hierarchy:

$$\epsilon_{\text{RF}} < \epsilon_{\text{WF}} \lesssim \epsilon_{\text{human}}. \quad (10.12)$$

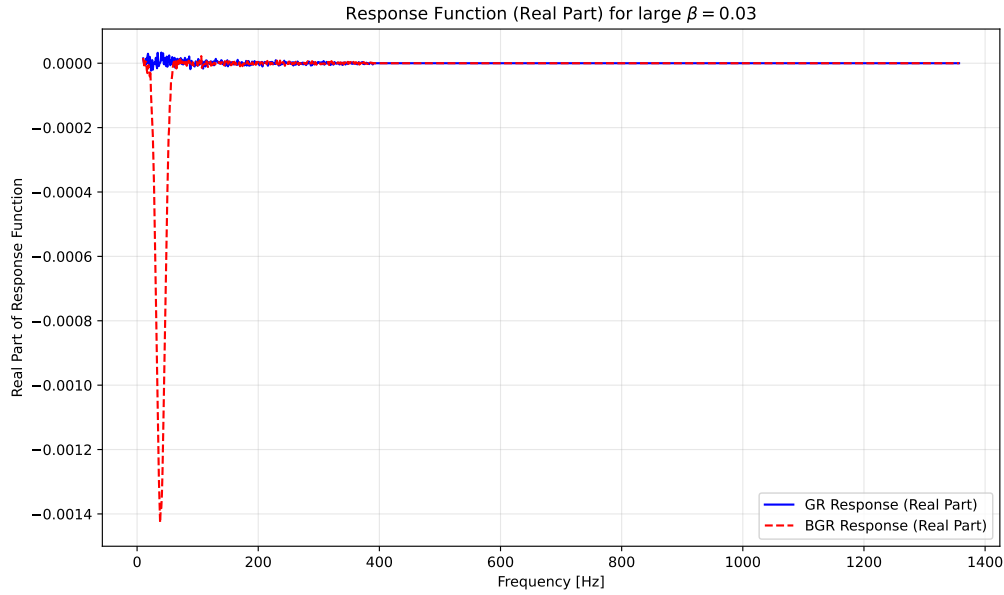
That is, the classifier trained on response functions outperforms both the waveform-based model *and* human-level classification, especially in the regime of small deformations. In summary, the response function formalism demonstrates how physically motivated feature engineering can bring classification systems closer to their theoretical performance ceiling. By isolating and amplifying the parts of the signal most sensitive to deviations from GR, response functions enhance both interpretability and classification accuracy—pushing machine learning models closer to the Bayes optimal limit.

## Visual Diagnostics: Response Function Structures Across Deformation Scales

To further illustrate the effectiveness of the response function formalism, we now visualize and compare the GR and BGR response functions at various deformation strengths. This highlights how, at large or moderate  $\beta$ , the difference is sometimes visually apparent, but becomes invisible at smaller  $\beta$ —yet remains detectable by a neural network.

**Large Deformation** ( $\beta = 0.03$ ). In Fig. 10.6, we show the normalized real part of the GR and BGR response functions for a deformation parameter value of  $\beta = 0.03$ . The difference between the two responses is immediately visible, with the BGR function exhibiting a strong and coherent deviation from the GR baseline. At this level, classification is trivial for both the neural network and human inspection. Nevertheless, the neural network remains valuable for systematically analyzing large datasets, where manual inspection is impractical.

**Intermediate Deformation** ( $\beta = 0.003$ ). Reducing the deformation strength to an intermediate level,  $\beta = 0.003$ , we obtain the response functions shown in Fig. 10.7. While the



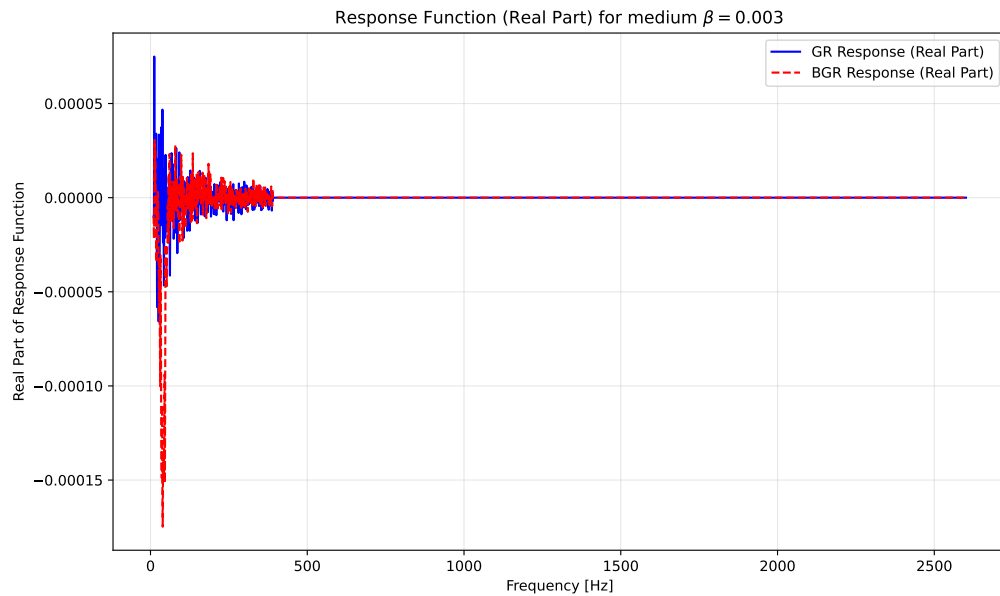
**Figure 10.6.** Comparison of normalized response functions at  $\beta = 0.03$ . The GR and BGR responses show clear differences, easily discernible by eye. At this level of deformation, classification is straightforward.

two signals are more similar than in the large- $\beta$  regime, there remains a distinguishable structure in the BGR response. This subtle, but still recognizable, deviation underscores why the neural network continues to achieve high accuracy—even though the waveform-based approach is less reliable here.

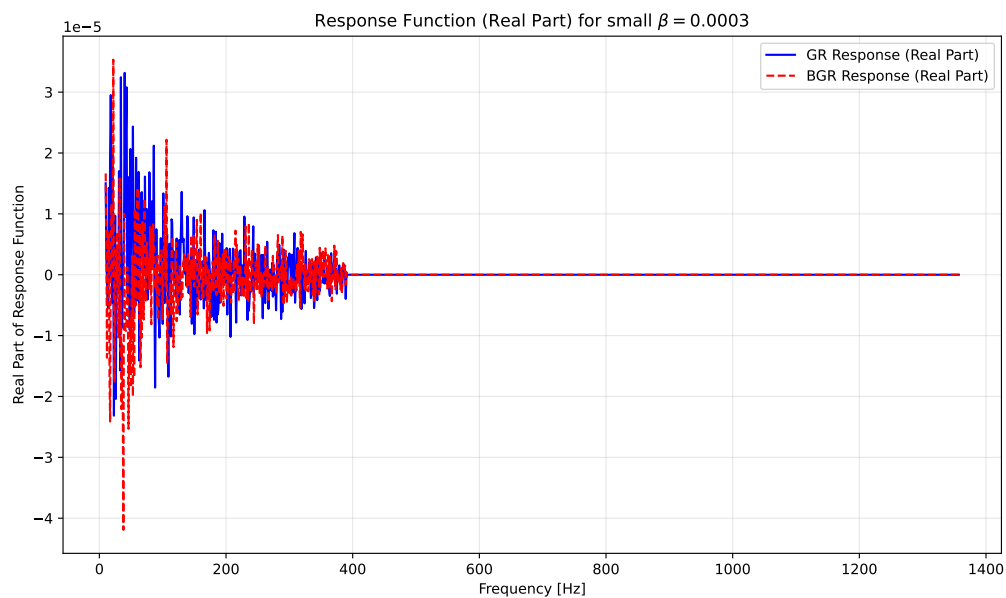
**Small Deformation** ( $\beta = 0.0003$ ). Fig. 10.8 presents the response functions for  $\beta = 0.0003$ . Here, both GR and BGR signals exhibit noisy, overlapping patterns that appear virtually indistinguishable to the human eye. Despite this, the trained neural network still classifies with over 95% accuracy, demonstrating its ability to extract minute, distributed features that remain hidden at the individual-event level.

## Averaging Methods vs. Neural Networks

One might wonder whether the faint BGR pattern could be recovered simply by averaging multiple examples. Indeed, in Fig. 10.9 we see nine pairs of GR/BGR response functions at  $\beta = 0.0003$ . Individually, these pairs look virtually identical, but averaging them yields a *systematic* difference, as illustrated in Fig. 10.10. The GR average converges toward zero from random fluctuation, while the BGR average retains a small but coherent feature.



**Figure 10.7.** Comparison of normalized response functions at  $\beta = 0.003$ . Though less pronounced than in the large  $\beta$  case, the BGR response retains visible structure that sets it apart from GR. The classifier continues to achieve near-perfect accuracy.



**Figure 10.8.** Comparison of normalized response functions at  $\beta = 0.0003$ . The GR and BGR responses exhibit highly similar noise-dominated behavior, making visual discrimination infeasible. However, the classifier still performs above the 95% accuracy threshold.

However, this averaging approach presupposes knowledge of which signals are GR and which are BGR. Thus, it cannot serve as a classification strategy in practice; one must already know the label to group signals correctly. Neural networks overcome this limitation by inferring such subtle regularities without prior labeling, thereby uncovering the same hidden structure at the level of individual events.

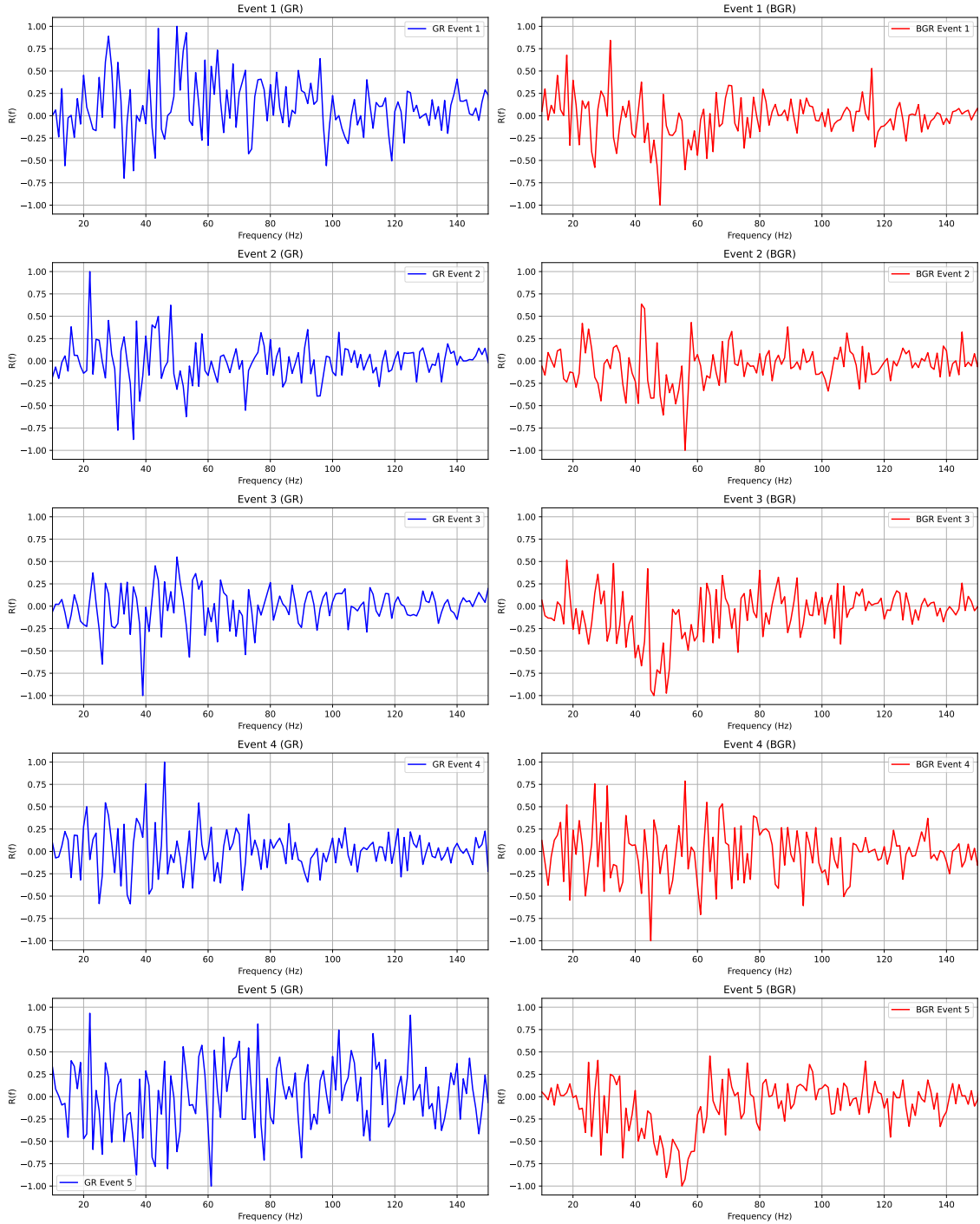
## Conclusion

Taken together, these results demonstrate the power and interpretability of the response function formalism. For large or intermediate deformations, the difference between GR and BGR can be visually distinguished, but manual classification quickly becomes impractical on large datasets. At small deformations, the signals become visually indistinguishable, surpassing human-level performance even though a consistent statistical imprint remains. A neural network trained on response functions can exploit this imprint, achieving high accuracy and pushing classification closer to the Bayes optimal limit.

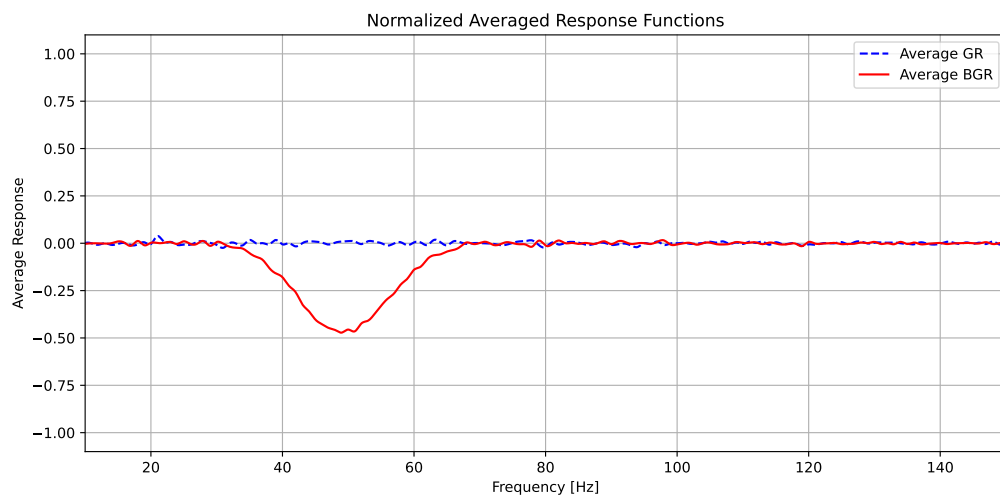
Overall, this study highlights that:

- **Choice of Feature Space Matters:** The response function isolates physically relevant aspects of the waveform, significantly improving classification performance compared to raw waveforms.
- **Neural Networks Surpass Visual Methods:** While human inspection fails at low  $\beta$ , the network maintains  $> 95\%$  accuracy by capturing subtle patterns across the frequency domain.
- **Statistical Methods and Prior Knowledge:** Averaging reveals hidden structure but requires prior labeling, underscoring the greater utility of machine learning for real-time, label-free detection tasks.
- **Approaching the Bayes Limit:** By focusing on the most discriminative features, the response function formalism allows classifiers to operate near the theoretical ceiling imposed by noise and data overlap.

These observations collectively reinforce the conclusion that response functions, combined with deep learning, form a robust, interpretable, and theoretically well-grounded approach to classifying gravitational wave signals—ultimately moving us closer to the elusive Bayes optimal error.



**Figure 10.9.** Normalized response functions for  $\beta = 0.0003$  across multiple gravitational wave events. **Left column:** GR responses. **Right column:** Corresponding BGR responses. Their indistinguishability at first glance is a key reason neural networks outperform human inspection in this regime.



**Figure 10.10.** Normalized averaged response functions at  $\beta = 0.0003$ . The GR average (blue, dashed) converges to zero due to noise cancellation. The BGR average (green, solid), however, reveals a residual pattern—a coherent imprint of the small deformation.



# Chapter 11

## Analysis of Modified Theories of Gravity

In the previous chapters, we demonstrated the effectiveness of neural network classifiers in detecting deviations from general relativity (GR) using controlled modifications of the gravitational wave (GW) phase. Starting from a Gaussian toy model and moving to more structured deviations in post-Newtonian (PN) phase coefficients, we showed how subtle changes in the waveform’s inspiral phase can be systematically identified and attributed to specific PN orders. However, these deformations—while informative—were constructed in a phenomenological manner and lacked direct connection to concrete theoretical models.

In this chapter, we take a decisive step toward physical realism by studying waveform deformations that arise explicitly from modified theories of gravity. In particular, we focus on the parameterized post-Einsteinian (ppE) framework [85, 86, 87], which provides a systematic way to incorporate deviations from GR into the frequency-domain gravitational waveform. This formalism allows us to encode leading-order phase corrections predicted by several well-motivated theories, including scalar–tensor models such as Brans–Dicke theory, massive graviton scenarios, and parity-violating Chern–Simons gravity. By embedding these analytically derived phase corrections into the waveform model, we construct beyond-GR (BGR) signals that are not only detectable but also interpretable within the context of fundamental physics. The goal of this chapter is to assess whether our classification pipeline can reliably distinguish these physically motivated deviations from standard GR predictions, and to identify the theoretical signatures that are most readily captured by data-driven approaches.

## 11.1 The Parameterized Post-Einsteinian (ppE) Framework

In order to systematically explore deviations from General Relativity (GR) in a theory-agnostic way, it is essential to construct a waveform model that can accommodate a wide variety of possible modifications. The *parameterized post-Einsteinian* (ppE) framework provides such a tool, allowing one to encapsulate deviations from GR in a unified and flexible formalism that is suitable for gravitational wave (GW) data analysis. Analogous in spirit to the parameterized post-Newtonian (ppN) approach used in weak-field tests of gravity, the ppE framework aims to model potential strong-field and dynamical deviations from GR by introducing free parameters into the GW waveform, particularly during the inspiral phase.

The central idea behind the ppE framework is to construct a deformation of the GR waveform that remains general enough to interpolate between GR and a broad class of modified theories of gravity, while still respecting basic physical principles such as energy conservation and dimensional consistency. These deformations are introduced directly in the frequency domain, which is well-suited for matched-filtering techniques used in GW detection pipelines.

Let us consider the frequency-domain gravitational waveform for a binary inspiral in GR, in the stationary phase approximation (SPA). It takes the general form:

$$\tilde{h}_{\text{GR}}(f) = \mathcal{A}_{\text{GR}}(f) e^{i\Psi_{\text{GR}}(f)}, \quad (11.1)$$

where  $\mathcal{A}_{\text{GR}}(f)$  is the amplitude and  $\Psi_{\text{GR}}(f)$  is the GW phase as predicted by GR. In the restricted post-Newtonian (PN) approximation, the phase can be expanded as a series in the dimensionless velocity parameter  $v = (\pi M f)^{1/3}$ , where  $M$  is the total mass of the binary system:

$$\Psi_{\text{GR}}(f) = 2\pi f t_c - \phi_c + \sum_{k=0}^N \psi_k v^k. \quad (11.2)$$

The ppE framework generalizes this by introducing parametric deviations in both the amplitude and phase. The general ppE-modified waveform takes the form [85]:

$$\tilde{h}_{\text{ppE}}(f) = \tilde{h}_{\text{GR}}(f) (1 + \alpha u^a) e^{i\beta u^b}, \quad (11.3)$$

where  $u = \pi M f$  is the reduced frequency, and the parameters  $(\alpha, a)$  and  $(\beta, b)$  characterize the deviation in amplitude and phase, respectively. The GR limit is recovered when  $\alpha = \beta = 0$ .

This ansatz allows for a controlled departure from GR in a phenomenological way. The parameters  $\beta$  and  $b$  govern the magnitude and PN order of the phase deviation. Since  $u \sim v^3$ , different values of  $b$  correspond to different PN orders. For example,  $b = -7/3$  corresponds to a  $-1$ PN correction relative to GR's leading order  $v^{-5}$  phase dependence. This flexibility allows the ppE framework to encapsulate a wide array of deviations stemming from different physical effects, such as additional radiation channels, modified dispersion relations, or non-GR conservative dynamics.

The strength of the ppE formalism lies in its ability to:

- Systematically interpolate between GR and a broad class of modified theories.
- Allow for parametrized, waveform-level tests of GR in the strong-field, highly dynamical regime.
- Remain analytically tractable and directly usable in GW data analysis via matched filtering.

In this chapter, we again focus exclusively on **phase modifications**. That is, we set  $\alpha = 0$  and restrict our attention to modifications encoded purely in the waveform's phase. This choice is motivated both by theoretical considerations—since many modifications to GR appear dominantly in the phase—and by practical considerations, as the phase carries the most prominent features of the inspiral dynamics.

Under this restriction, the ppE waveform simplifies to:

$$\tilde{h}_{\text{ppE}}(f) = \tilde{h}_{\text{GR}}(f) e^{i\beta u^b}. \quad (11.4)$$

In the regime where the deviation is small, i.e.,  $\beta u^b \ll 1$ , we can expand the exponential to first order:

$$\tilde{h}_{\text{ppE}}(f) \approx \tilde{h}_{\text{GR}}(f) (1 + i\beta u^b). \quad (11.5)$$

This reveals that the deviation from GR enters as a leading-order phase shift:

$$\delta\psi(f) = \beta u^b = \beta(\pi M f)^b. \quad (11.6)$$

The total phase of the waveform thus becomes:

$$\Psi(f) = \Psi_{\text{GR}}(f) + \delta\psi(f) = \Psi_{\text{GR}}(f) + \beta u^b. \quad (11.7)$$

The term  $\delta\psi(f)$  encapsulates the leading-order correction to the GR phase introduced by beyond-GR physics in the ppE framework. In the remainder of this chapter, we will

explore specific choices of the parameters  $(\beta, b)$  that arise from concrete modified theories of gravity, and analyze whether such deviations can be robustly detected by neural network classifiers trained on waveform data.

## 11.2 Modified Theories of Gravity and Their ppE Signatures

In this section, we introduce several representative modified theories of gravity [88] that provide explicit and physically motivated deviations from General Relativity (GR). Each theory modifies the gravitational dynamics in a different way—through scalar degrees of freedom, modified propagation, or parity-violating interactions—and leads to characteristic corrections in the phase of gravitational wave (GW) signals. We describe the theoretical motivation for each model, provide its action or Lagrangian, identify relevant coupling constants, and extract the leading-order phase correction  $\delta\psi(f)$  as it appears in the ppE framework.

### 11.2.1 Brans–Dicke Theory (Scalar–Tensor Gravity)

Brans–Dicke theory [89] is one of the earliest and simplest scalar–tensor theories of gravity. It introduces a massless scalar field that couples to the Ricci scalar, effectively promoting Newton’s constant  $G$  to a dynamical quantity. This modification is motivated by attempts to incorporate Mach’s principle and to generalize GR within a broader class of metric theories.

**Action:**

$$S = \frac{1}{16\pi} \int d^4x \sqrt{-g} \left[ \phi R - \frac{\omega_{\text{BD}}}{\phi} (\nabla_\mu \phi)(\nabla^\mu \phi) \right] + S_{\text{matter}}[g_{\mu\nu}, \Psi], \quad (11.8)$$

where  $\phi$  is the scalar field, and  $\omega_{\text{BD}}$  is the Brans–Dicke coupling constant.

**Key parameters:**

- $\omega_{\text{BD}}$ : Brans–Dicke coupling constant (dimensionless).
- $s_1, s_2$ : sensitivities of the two compact objects (related to their scalar charges).

**ppE phase parameters:**

$$\beta_{\text{BD}} = -\frac{5}{3584 \omega_{\text{BD}}} (s_1 - s_2)^2 \eta^{2/5}, \quad b = -\frac{7}{3}, \quad (11.9)$$

where  $\eta = \frac{m_1 m_2}{(m_1 + m_2)^2}$  is the symmetric mass ratio.

**Phase correction:**

$$\delta\psi_{\text{BD}}(f) = \beta_{\text{BD}}(\pi M f)^{-7/3}. \quad (11.10)$$

The  $-1\text{PN}$  correction arises from dipole radiation in the scalar field. However, in standard Brans–Dicke theory, stationary black holes do not carry scalar charge due to no-hair theorems. As a result, the scalar charges  $s_1$  and  $s_2$  vanish, leading to  $\beta_{\text{BD}} = 0$  for binary black hole systems. Even in dynamical binaries, Brans–Dicke theory does not generically endow black holes with scalar hair. For this reason, although we include Brans–Dicke theory here for completeness, we do not study it in our numerical analysis, as our setup is specifically designed for non-precessing **black hole–black hole binaries**. This theory is, however, relevant in neutron star binaries or mixed BH–NS systems, where scalar charges can be nonzero.

### 11.2.2 Massive Gravity Theories

Massive graviton theories [90, 91] explore the possibility that the graviton has a small but nonzero mass, leading to a modified dispersion relation and a frequency-dependent propagation speed for gravitational waves. This modifies the gravitational wave phase through a cumulative dephasing effect over cosmological distances, as higher-frequency waves travel faster than lower-frequency ones.

**Action:** There are various formulations, but generically the Einstein–Hilbert term is supplemented with a mass term of Fierz–Pauli type:

$$S = \frac{1}{16\pi G} \int d^4x \sqrt{-g} \left[ R - \frac{1}{2} m_g^2 (h_{\mu\nu} h^{\mu\nu} - h^2) \right], \quad (11.11)$$

where  $m_g$  is the graviton mass and  $h_{\mu\nu}$  is the metric perturbation.

**Key parameters:**

- $\lambda_g = \frac{h}{m_g c}$ : graviton Compton wavelength.
- $D$ : effective luminosity distance to the source.
- $z$ : redshift of the source.
- $\mathcal{M} = \eta^{3/5} M$ : chirp mass.

**ppE phase parameters:**

$$\beta_{\text{MG}} = \frac{\pi^2 D \mathcal{M}}{\lambda_g^2 (1+z)}, \quad b = -1, \quad (11.12)$$

where  $u = \pi M f$  is the dimensionless reduced frequency variable.

**Phase correction:**

$$\delta\psi_{\text{MG}}(f) = -\beta_{\text{MG}} u^{-1} = -\beta_{\text{MG}} (\pi M f)^{-1}. \quad (11.13)$$

This is a +1PN effect arising from the modified dispersion relation for a massive graviton. The phase delay accumulates over the travel time from the source to the detector. The correction vanishes in the limit  $\lambda_g \rightarrow \infty$ , corresponding to a massless graviton. Choosing  $\lambda_g \gg D$  ensures that  $\beta_{\text{MG}} \ll 1$ , consistent with current experimental bounds from LIGO, Virgo, and pulsar timing observations.

### 11.2.3 Dynamical Chern–Simons Gravity (Parity-Violating Gravity)

Dynamical Chern–Simons (dCS) gravity [92, 93] arises in several high-energy physics frameworks, including string theory and loop quantum gravity. It breaks parity symmetry in the gravitational sector by coupling a scalar field to the gravitational Pontryagin density. Unlike its non-dynamical counterpart, dCS gravity treats the scalar field as an independent dynamical degree of freedom, which allows black holes to acquire scalar charge if they are spinning. This leads to detectable modifications to the gravitational waveform during the inspiral.

**Action:**

$$S = \int d^4x \sqrt{-g} \left[ \frac{R}{16\pi G} + \frac{\alpha_{\text{CS}}}{4} \theta R^* R - \frac{1}{2} (\nabla_\mu \theta) (\nabla^\mu \theta) \right] + S_{\text{matter}}, \quad (11.14)$$

where  $\theta$  is the dynamical scalar field and  $R^* R$  is the Pontryagin density.

**Key parameters:**

- $\xi = \alpha_{\text{CS}}/\kappa$ : dimensionful dCS coupling constant (with dimensions of length squared).
- $\chi$ : dimensionless spin of the black hole.

**ppE phase parameters (in dCS gravity):**

$$\beta_{\text{dCS}} \propto \zeta_{\text{CS}} \cdot \left( \frac{m_1^2}{M^2} \chi_1^2 + \frac{m_2^2}{M^2} \chi_2^2 \right), \quad b = -\frac{1}{3}, \quad (11.15)$$

where  $\zeta_{\text{CS}} \propto \xi/M^4$  is a dimensionless small-coupling parameter.

**Phase correction:**

$$\delta\psi_{\text{dCS}}(f) = \beta_{\text{dCS}} (\pi M f)^{-1/3}. \quad (11.16)$$

In dCS gravity, spinning black holes acquire scalar charge and emit parity-violating gravitational radiation. This leads to an inspiral-phase correction at fractional PN order  $b = -1/3$ , making it a clean and distinct deviation from GR. The effect vanishes in the non-spinning limit, but is present in non-precessing binaries with aligned spins — the regime we focus on in this work. Therefore, dCS gravity is fully compatible with our waveform modeling and classification framework, and provides a natural candidate for the first beyond-GR theory to include in our neural network analysis.

Each of the theories above contributes a distinct frequency-dependent phase correction of the form

$$\delta\psi(f) = \beta (\pi M f)^b, \quad (11.17)$$

where the values of  $\beta$  and  $b$  depend on the theory-specific coupling constants and physical mechanisms. These expressions are valid in the leading-order small-coupling approximation and can be directly inserted into the ppE waveform model to simulate realistic beyond-GR (BGR) waveforms. In the next sections, we implement these deformations and study whether such subtle deviations can be reliably detected and classified using our neural network framework.

### 11.3 Neural Network Classification of Massive Gravity Waveforms

Having developed a general framework for modeling post-Einsteinian gravitational waveforms and training neural networks to classify deviations from General Relativity (GR), we now apply this methodology to a specific and physically motivated case: **massive graviton (MG)** theories. In these models, the propagation of gravitational waves is modified due to the presence of a small but nonzero graviton mass, leading to a frequency-dependent dispersion relation and an accumulated phase shift in the waveform.

Now that we have analytic control over the waveform modification through the ppE framework, we are in a position to use the neural network classifier developed in previous chapters to test the detectability of this deviation. In particular, the MG correction introduces a leading-order frequency-domain phase shift of the form:

$$\delta\psi_{\text{MG}}(f) = -\beta_{\text{MG}}(\pi M f)^{-1}, \quad (11.18)$$

where the deformation strength  $\beta_{\text{MG}}$  depends on the graviton Compton wavelength  $\lambda_g$ , the source distance  $D$ , and the redshift  $z$ , as discussed in Section 11.2.

In this section, we focus exclusively on this MG-induced deformation and explore whether it can be robustly identified using a neural network trained to distinguish between GR and beyond-GR (BGR) waveforms. Our classifier receives inspiral-only waveforms as input and is trained to predict whether a given waveform contains the MG-induced phase shift described above.

While we restrict attention here to massive gravity for concreteness and clarity, the same procedure can be applied to other modified gravity theories discussed earlier, such as dynamical Chern–Simons gravity or scalar–tensor theories with dipole radiation. Each of these can be modeled via an analytic ppE phase term  $\delta\psi(f)$ , which can be incorporated into the waveform and passed to the classifier. Thus, this study serves as a template for broader applications of data-driven classification in gravitational wave tests of GR.

### 11.3.1 Detectability Threshold for the Graviton Mass

The primary goal of this section is to determine the smallest graviton mass  $m_g$  — as it appears in the Lagrangian — for which our neural network can reliably distinguish MG-modified waveforms from GR predictions. To quantify detectability, we use the response function formalism described in Chapter 10 and evaluate the classifier’s accuracy over a range of graviton masses.

Our analysis shows that the trained network achieves classification accuracy greater than 95% for all graviton masses satisfying

$$m_g \gtrsim 2.1 \times 10^{-61} \text{ kg} \approx 1.2 \times 10^{-25} \text{ eV}/c^2, \quad (11.19)$$

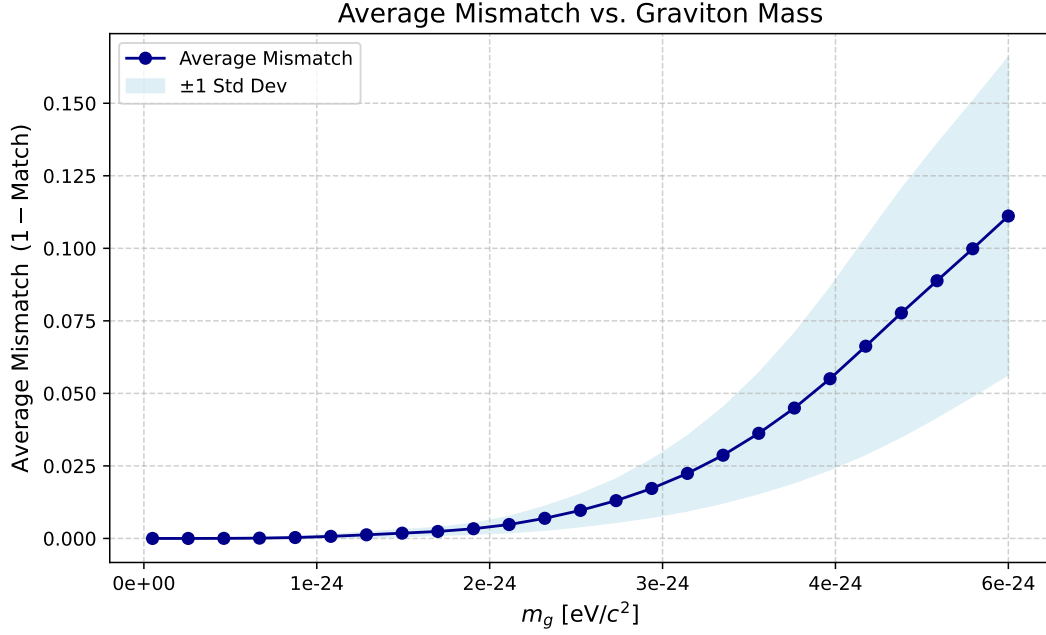
indicating a robust sensitivity to very small deviations from GR. Below this threshold, the accumulated MG phase shift becomes too small to be reliably detected in the inspiral waveform, and the classifier’s performance degrades accordingly.

To illustrate the structure of the phase correction, we also plot the phase shift  $\delta\psi_{\text{MG}}(f)$  as a function of frequency for several representative values of  $m_g$ . As expected, higher graviton masses produce larger dephasing, with noticeable impact in the 10–800 Hz range.

### 11.3.2 Comparison with Existing Bounds

Our machine learning approach enables a complementary strategy to constrain modified gravity theories from waveform structure alone. For context, current observational bounds





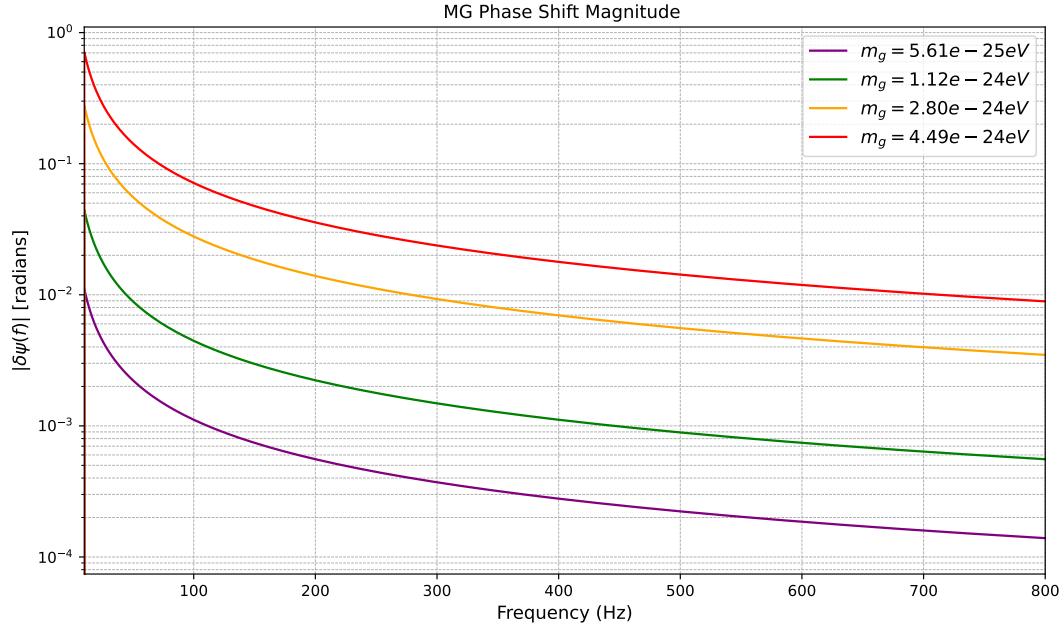
**Figure 11.1.** Average waveform mismatch as a function of the graviton mass  $m_g$ . The classifier becomes insensitive to MG-induced phase shifts for masses smaller than  $\sim 2.1 \times 10^{-61}$  kg.

on the graviton mass derived from binary black hole coalescences constrain [94, 95]:

$$m_g < 7.7 \times 10^{-23} \text{ eV}/c^2 \quad (11.20)$$

at 90% confidence level [94, 95]. While our analysis does not aim to set competitive observational limits, it highlights the potential of data-driven methods to *detect* or *exclude* deviations from GR based on waveform morphology — even when the deviations are much smaller than current observational bounds.

Our study therefore complements traditional parameter estimation pipelines by offering a model-independent, classification-based test of gravity, and can serve as a first step in future multi-modal searches for beyond-GR signatures in gravitational wave data.



**Figure 11.2.** Phase shift  $\delta\psi_{\text{MG}}(f)$  for various values of  $m_g$ . The magnitude of the shift increases with graviton mass, affecting the waveform primarily in the detector-sensitive frequency band.

In this chapter, we applied our neural network classification framework to a concrete modified gravity model: massive graviton theories. By leveraging the analytical ppE phase correction induced by the graviton mass, we generated a consistent dataset of GR and BGR waveforms and trained a classifier to detect deviations arising from MG-induced dispersion.

Our results demonstrate that the classifier achieves high sensitivity to the waveform modifications introduced by massive gravity, successfully identifying deviations from GR with over 95% accuracy for graviton masses exceeding  $2.1 \times 10^{-61}$  kg ( $\sim 1.2 \times 10^{-25}$  eV/ $c^2$ ). This detection threshold is well below existing observational upper bounds on the graviton mass, highlighting the discriminative power of machine learning models trained on waveform morphology.

This study provides a clear demonstration of how data-driven classification methods can complement traditional parameter estimation pipelines in gravitational wave astronomy. The techniques developed here are readily applicable to other modified gravity scenarios — including scalar–tensor theories and parity-violating models — which we will explore in future work or subsequent chapters.

Ultimately, our analysis illustrates the promise of neural network classifiers as tools

---

for probing the structure of gravitational waveforms and testing the fundamental nature of gravity itself.



# Chapter 12

## Conclusion

This thesis has pursued a dual investigation into the theoretical foundations and observational consequences of modified gravity, focusing specifically on the quantum stability of Generalized Proca theories and the classification of gravitational waveforms beyond General Relativity (GR). These two lines of research, while distinct in methodology, are unified by the broader goal of understanding gravity as an effective field theory and testing its limits in both high-energy and strong-field regimes.

### Quantum Stability of Generalized Proca Theories

In the first part of this thesis, we conducted a detailed quantum field theoretic analysis of Generalized Proca theories—a class of vector-tensor models that generalize the Proca action through non-trivial derivative self-interactions and non-minimal couplings to gravity. These theories are constructed to propagate exactly three degrees of freedom (two transverse and one longitudinal polarization) while maintaining second-order equations of motion, thereby avoiding the Ostrogradsky instability.

We began with a systematic perturbative study in flat Minkowski spacetime, where the focus was on evaluating the one-loop divergent structures of the 2-point, 3-point, and 4-point correlation functions. Using a simplified yet representative form of the Generalized Proca Lagrangian, we demonstrated that the theory is technically natural: the 2-point and 3-point divergences can be absorbed into redefinitions of existing terms in the Lagrangian, and no higher-derivative or ghost-like operators are generated. Interestingly, certain gauge-invariant combinations, such as  $(\Box\eta_{\mu\nu} - \partial_\mu\partial_\nu)$ , emerge in the loop corrections, hinting at an effective restoration of gauge symmetry at the quantum level despite the absence of

classical gauge invariance.

We extended our analysis to include gravitational interactions by studying Generalized Proca theories on weakly curved backgrounds. In this setting, the coupling to dynamical gravity introduces mixing between the vector field and graviton perturbations, significantly complicating the structure of the kinetic and interaction terms. To address this, we constructed the full quadratic action for metric and vector fluctuations and derived the corresponding Feynman rules.

To manage the complexity introduced by background vector fields, we employed a scalar-vector-tensor (SVT) decomposition that isolates physical degrees of freedom and facilitates the identification of dynamical and auxiliary fields. This decomposition enabled a perturbative elimination of constraint fields and allowed us to express the action in terms of propagating modes.

Our results indicate that the theory remains radiatively stable in curved spacetime, at least at the 2- and 3-point level. Loop-induced corrections preserve the structure of the effective action without introducing unphysical degrees of freedom or violating unitarity. Although a full analysis of the 4-point function in curved space remains for future work [40], our setup and preliminary findings strongly suggest that Generalized Proca theories can be interpreted as well-defined effective field theories up to a natural cutoff scale.

These results not only strengthen the theoretical foundations of vector-tensor modifications of gravity but also open avenues for connecting these models to cosmological phenomena such as inflation, dark energy, or deviations from GR in strong-field astrophysical environments.

## **Classification of Gravitational Waveforms Beyond GR**

The second part of this thesis focused on developing a machine learning framework to classify gravitational waveforms into GR and beyond-GR (BGR) categories. This work is motivated by the growing potential of gravitational wave astronomy to serve as a precision tool for probing the fundamental nature of gravity.

We constructed a diverse dataset of gravitational waveforms that includes both GR signals and deformations representative of modified gravity theories. Initial tests were performed using controlled Gaussian phase deformations, allowing us to quantify the minimum detectable deviation under realistic noise conditions. Convolutional neural networks (CNNs) trained on these waveforms successfully identified deviations with high accuracy,

down to a mismatch level of approximately  $10^{-3}$ .

To improve both performance and interpretability, we introduced the *response function*—a frequency-domain representation of how small phase deformations affect the overlap between waveforms. Neural networks trained on response functions significantly outperformed waveform-based classifiers, achieving reliable detection at deformation levels as small as  $\beta = 3 \times 10^{-4}$ . This response function approach proved both physically transparent and computationally advantageous.

We then transitioned from toy models to structured multiclass classification using perturbed post-Newtonian (PN) coefficients. This allowed us not only to detect BGR deviations but to attribute them to specific PN orders, thus enhancing theoretical interpretability. Finally, we applied our framework to waveform models derived from massive graviton theories via the parameterized post-Einsteinian (ppE) formalism. Our classifiers detected graviton masses as low as  $m_g \sim 1.2 \times 10^{-25} \text{ eV}/c^2$ , demonstrating sensitivity well beyond current observational bounds.

While Generalized Proca theories were not directly included in the waveform classification due to the lack of concrete ppE expressions, our framework is designed to accommodate them once such predictions become available. This integration represents a promising direction for future work.

## Outlook

Several directions naturally follow from this work. On the theoretical side, completing the analysis of 4-point functions in curved backgrounds will further test the robustness of Generalized Proca theories. Exploring their UV completions and embedding them in more fundamental frameworks (e.g., string theory or holography) remains an open challenge.

On the data-driven side, improvements in computational efficiency, uncertainty quantification, and generalization to more realistic waveform features (such as spin, eccentricity, and higher modes) will be critical. Moving beyond softmax-based classification to multi-label architectures could enable the simultaneous detection of multiple deviations, better reflecting the structure of realistic effective field theories.

Finally, the response function formalism introduced here has broader implications. It may serve not only as a feature extraction tool but also as a bridge between machine learning and physical interpretability, enabling future applications in model selection, anomaly detection, and parameter inference.

In summary, this thesis presents a coherent framework for probing modified gravity at both the theoretical and observational level. By combining quantum field theoretic analysis with machine learning techniques, we offer new tools for exploring the landscape of viable gravitational theories and for extracting their imprints from gravitational wave data. These methods will play a crucial role in the next generation of gravitational tests, potentially revealing the first hints of new physics beyond Einstein's theory.



# Appendix



# Appendix A

## Lagrangians

### A.1 Cubic Interaction Terms

This appendix collects the complete cubic-order interaction terms arising from the generalized Proca Lagrangian. The expressions are grouped according to their origin from the Lagrangian pieces  $\mathcal{L}_n$ .

All expressions are given in terms of canonically normalized fields and written using flat spacetime indices. We follow the same conventions used in the main text.

#### Cubic Interactions

$$\begin{aligned} \mathcal{L}_2^{(3)} = & -\frac{\sqrt{-\bar{G}_4} h \partial_a \delta A_b \partial^b \delta A^a}{2\sqrt{2} \bar{G}_4} + \frac{\sqrt{-\bar{G}_4} h \partial_b \delta A_a \partial^b \delta A^a}{2\sqrt{2} \bar{G}_4} + \frac{\sqrt{-\bar{G}_4} h_{ac} \partial_b \delta A^c \partial^b \delta A^a}{\sqrt{2} \bar{G}_4} \\ & - \frac{\sqrt{-\bar{G}_4} h_{bc} \partial^b \delta A^a \partial^c \delta A_a}{\sqrt{2} \bar{G}_4} - \frac{\sqrt{2} \sqrt{-\bar{G}_4} \delta A_a \partial_b h_{ac} \partial^c \delta A^b}{\bar{G}_4} + \frac{\sqrt{2} \sqrt{-\bar{G}_4} \delta A_a \partial_c h_{ab} \partial^c \delta A^b}{\bar{G}_4} \\ & - \frac{\sqrt{-\bar{G}_4} h \delta A_a \delta A^a \bar{G}_2^{(1,0)}}{2\sqrt{2} \bar{G}_4 \bar{G}_2^{(0,1)}} - \frac{\sqrt{-\bar{G}_4} h_{ab} \delta A^a \delta A^b \bar{G}_2^{(1,0)}}{\sqrt{2} \bar{G}_4 \bar{G}_2^{(0,1)}}. \end{aligned} \quad (\text{A.1})$$

$$\begin{aligned} \mathcal{L}_3^{(3)} = & \frac{\delta A_b \delta A^b \square \delta A_a \bar{G}_3'}{2 \left( \bar{G}_2^{(0,1)} \right)^{3/2}} + \frac{\bar{G}_3 h_{ac} \delta A_b \partial^b h^{ac}}{\bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} - \frac{\bar{G}_3 h \delta A_b \partial^b h}{2 \bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} + \frac{\bar{G}_3 h_{ac} h^{ac} \square \delta A_b}{2 \bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} \\ & - \frac{\bar{G}_3 h^2 \square \delta A_b}{2 \bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} + \frac{\bar{G}_3 h^2 \square \delta A_b}{4 \bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}}. \end{aligned} \quad (\text{A.2})$$

$$\begin{aligned}
\mathcal{L}_4^{(3)} = & -\frac{3\sqrt{-\bar{G}_4} h_{ab} \partial^a h_{cd} \partial^b h^{cd}}{\sqrt{2} \bar{G}_4} + \frac{\sqrt{-\bar{G}_4} h_{ab} \partial^a h_c^c \partial^b h_d^d}{\sqrt{2} \bar{G}_4} - \frac{\sqrt{-\bar{G}_4} h^2 \partial^b \partial^a h_{ab}}{2\sqrt{2} \bar{G}_4} \\
& + \frac{\sqrt{-\bar{G}_4} h^2 \partial^b \partial_b h}{2\sqrt{2} \bar{G}_4} - \frac{2\sqrt{2}\sqrt{-\bar{G}_4} h_{ab} \partial^b h_d^d \partial^c h_c^a}{\bar{G}_4} - \frac{2\sqrt{2}\sqrt{-\bar{G}_4} h_{ab} \partial^b h_c^a \partial^c h_d^d}{\bar{G}_4} \\
& - \frac{2\sqrt{2}\sqrt{-\bar{G}_4} h_c^a h_{ab} \partial^c \partial^b h_d^d}{\bar{G}_4} + \frac{\sqrt{2}\sqrt{-\bar{G}_4} h h_{bc} \partial^c \partial^b h_d^d}{\bar{G}_4} + \frac{\sqrt{2}\sqrt{-\bar{G}_4} h_{ab} \partial^c h_d^d \partial_c h^{ab}}{\bar{G}_4} \\
& - \frac{\sqrt{-\bar{G}_4} h \partial^c h_d^d \partial_c h_b^b}{2\sqrt{2} \bar{G}_4} + \frac{2\sqrt{2}\sqrt{-\bar{G}_4} h_{ab} \partial^c h_c^a \partial^d h_b^b}{\bar{G}_4} + \frac{4\sqrt{2}\sqrt{-\bar{G}_4} h_{ab} \partial^b h_c^a \partial^d h_c^d}{\bar{G}_4} \\
& - \frac{\sqrt{2}\sqrt{-\bar{G}_4} h \partial^b h_{bc} \partial^d h_c^d}{\bar{G}_4} - \frac{2\sqrt{2}\sqrt{-\bar{G}_4} h_{ab} \partial_c h^{ab} \partial^d h_c^d}{\bar{G}_4} + \frac{\sqrt{2}\sqrt{-\bar{G}_4} h \partial_c h_b^b \partial^d h_c^d}{\bar{G}_4} \\
& + \frac{2\sqrt{2}\sqrt{-\bar{G}_4} h_{ab} h_{cd} \partial^d \partial^b h^{ac}}{\bar{G}_4} - \frac{2\sqrt{2}\sqrt{-\bar{G}_4} h_{ab} h_{cd} \partial^d \partial^c h^{ab}}{\bar{G}_4} + \frac{4\sqrt{2}\sqrt{-\bar{G}_4} h_c^a h_{ab} \partial^d \partial^c h_b^d}{\bar{G}_4} \\
& - \frac{\sqrt{-\bar{G}_4} h_{ab} h^{ab} \partial^d \partial^c h_{cd}}{\sqrt{2} \bar{G}_4} + \frac{\sqrt{-\bar{G}_4} h^2 \partial^d \partial^c h_{cd}}{\sqrt{2} \bar{G}_4} - \frac{2\sqrt{2}\sqrt{-\bar{G}_4} h_c^a h_{ab} \partial^d \partial_d h^{bc}}{\bar{G}_4} \\
& + \frac{\sqrt{2}\sqrt{-\bar{G}_4} h h_{bc} \partial^d \partial_d h^{bc}}{\bar{G}_4} + \frac{\sqrt{-\bar{G}_4} h_{ab} h^{ab} \partial^d \partial_d h_c^c}{\sqrt{2} \bar{G}_4} - \frac{\sqrt{-\bar{G}_4} h^2 \partial^d \partial_d h_c^c}{\sqrt{2} \bar{G}_4} \\
& + \frac{2\sqrt{2}\sqrt{-\bar{G}_4} h_{ab} \partial^b h_{cd} \partial^d h^{ac}}{\bar{G}_4} - \frac{2\sqrt{2}\sqrt{-\bar{G}_4} h h_{bc} \partial^d \partial^c h_b^d}{\bar{G}_4} - \frac{\sqrt{2}\sqrt{-\bar{G}_4} h_{ab} \partial^c h^{bd} \partial_d h_c^a}{\bar{G}_4} \\
& - \frac{3\sqrt{2}\sqrt{-\bar{G}_4} h_{ab} \partial^d h^{bc} \partial_d h_c^a}{\bar{G}_4} - \frac{\sqrt{-\bar{G}_4} h \partial^c h^{bd} \partial_d h_{bc}}{\sqrt{2} \bar{G}_4} \\
& + \frac{3\sqrt{-\bar{G}_4} h \partial^d h^{bc} \partial_d h_{bc}}{2\sqrt{2} \bar{G}_4} - \frac{\sqrt{2}\sqrt{-\bar{G}_4} \delta A_a \partial^a h_c^c \partial^b \delta A_b \bar{G}'_4}{\bar{G}_4 \bar{G}_2^{(0,1)}} - \frac{\sqrt{-\bar{G}_4} h \partial^a \delta A_a \partial^b \delta A_b \bar{G}'_4}{\sqrt{2} \bar{G}_4 \bar{G}_2^{(0,1)}} \\
& + \frac{\sqrt{-\bar{G}_4} h \partial^a \delta A^b \partial_b \delta A_a \bar{G}'_4}{\sqrt{2} \bar{G}_4 \bar{G}_2^{(0,1)}} - \frac{\sqrt{-\bar{G}_4} \delta A^a \delta A_a \partial^c \partial^b h_{bc} \bar{G}'_4}{\sqrt{2} \bar{G}_4 \bar{G}_2^{(0,1)}} \\
& + \frac{\sqrt{-\bar{G}_4} \delta A^a \delta A_a \partial^c \partial_c h_b^b \bar{G}'_4}{\sqrt{2} \bar{G}_4 \bar{G}_2^{(0,1)}} + \frac{\sqrt{2}\sqrt{-\bar{G}_4} \delta A_a \partial^a h^{bc} \partial_c \delta A_b \bar{G}'_4}{\bar{G}_4 \bar{G}_2^{(0,1)}} \\
& + \frac{\sqrt{2}\sqrt{-\bar{G}_4} \delta A_a \partial^b h^{ac} \partial_c \delta A_b \bar{G}'_4}{\bar{G}_4 \bar{G}_2^{(0,1)}} - \frac{\sqrt{2}\sqrt{-\bar{G}_4} \delta A_a \partial^c h^{ab} \partial_c \delta A_b \bar{G}'_4}{\bar{G}_4 \bar{G}_2^{(0,1)}}.
\end{aligned} \tag{A.3}$$

$$\mathcal{L}_5^{(3)} = -\frac{\bar{g}_5 \partial^b \delta A_a \partial^c \delta A_b \partial_c \delta A^a}{(\bar{G}_2^{(0,1)})^{3/2}} + \frac{\bar{g}_5 \partial^a \delta A^c \partial^b \delta A_a \partial_c \delta A_b}{(\bar{G}_2^{(0,1)})^{3/2}} - \frac{\bar{g}_5 \partial^a \delta A_a \partial^b \delta A^c \partial_c \delta A_b}{(\bar{G}_2^{(0,1)})^{3/2}}$$

$$\begin{aligned}
& + \frac{\bar{g}_5 \partial^a \delta A_a \partial^c \delta A^b \partial_c \delta A_b}{\left(\bar{G}_2^{(0,1)}\right)^{3/2}} - \frac{\partial^a \delta A_a \partial^b \delta A_b \partial^c \delta A_c \bar{G}'_5}{6 \left(\bar{G}_2^{(0,1)}\right)^{3/2}} - \frac{\partial^a \delta A^c \partial^b \delta A_a \partial_c \delta A_b \bar{G}'_5}{3 \left(\bar{G}_2^{(0,1)}\right)^{3/2}} \\
& + \frac{\partial^a \delta A_a \partial^b \delta A^c \partial_c \delta A_b \bar{G}'_5}{2 \left(\bar{G}_2^{(0,1)}\right)^{3/2}} - \frac{\bar{G}_5 \partial^a h_{cd} \partial^b h^{cd} \partial_b \delta A_a}{2 \bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} - \frac{\bar{G}_5 h_{cd} \partial^b \partial^a h^{cd} \partial_b \delta A_a}{\bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} \\
& + \frac{\bar{G}_5 h^c_c \partial^b \partial^a h^d_d \partial_b \delta A_a}{2 \bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} \\
& - \frac{\bar{G}_5 \partial^a h^b_c \partial_b \delta A_a \partial^c h^d_d}{2 \bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} - \frac{\bar{G}_5 \partial^b h^a_c \partial_b \delta A_a \partial^c h^d_d}{2 \bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} - \frac{\bar{G}_5 h^b_c \partial_b \delta A_a \partial^c \partial^a h^d_d}{\bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} \\
& + \frac{\bar{G}_5 \delta A_a \partial^a h_{bc} \partial^c \partial^b h^d_d}{2 \bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} + \frac{\bar{G}_5 h_{bc} \partial^a \delta A_a \partial^c \partial^b h^d_d}{\bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} + \frac{\bar{G}_5 \partial^b \delta A_a \partial^c h^d_d \partial_c h^{ab}}{2 \bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} \\
& - \frac{\bar{G}_5 \partial^a \delta A_a \partial^c h^d_d \partial_c h^b_b}{4 \bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} \\
& - \frac{\bar{G}_5 \partial^a \delta A_a \partial^b h_{bc} \partial^d h^c_d}{\bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} + \frac{\bar{G}_5 \partial^a h^b_c \partial_b \delta A_a \partial^d h^c_d}{\bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} + \frac{\bar{G}_5 \partial^b h^a_c \partial_b \delta A_a \partial^d h^c_d}{\bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} \\
& - \frac{\bar{G}_5 \partial^b \delta A_a \partial^c h^{ab} \partial^d h^c_d}{\bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} + \frac{\bar{G}_5 \partial^a \delta A_a \partial^c h^b_b \partial^d h^c_d}{\bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} + \frac{\bar{G}_5 h_{cd} \partial^b \delta A_a \partial^d \partial^a h^{bc}}{\bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} \\
& - \frac{\bar{G}_5 h^c_c \partial^b \delta A_a \partial^d \partial^a h^b_d}{2 \bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} \\
& + \frac{\bar{G}_5 h^b_c \partial_b \delta A_a \partial^d \partial^a h^c_d}{\bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} + \frac{\bar{G}_5 h_{cd} \partial^b \delta A_a \partial^d \partial_b h^{ac}}{\bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} - \frac{\bar{G}_5 h^c_c \partial^b \delta A_a \partial^d \partial_b h^a_d}{2 \bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} \\
& - \frac{\bar{G}_5 h_{cd} \partial^b \delta A_a \partial^d \partial^c h^{ab}}{\bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} + \frac{\bar{G}_5 h^b_c \partial_b \delta A_a \partial^d \partial^c h^a_d}{\bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} - \frac{\bar{G}_5 \delta A_a \partial^a h_{bc} \partial^d \partial^c h^b_d}{\bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} \\
& - \frac{2 \bar{G}_5 h_{bc} \partial^a \delta A_a \partial^d \partial^c h^b_d}{\bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} + \frac{\bar{G}_5 \delta A_a \partial^a h^b_b \partial^d \partial^c h_{cd}}{2 \bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} + \frac{\bar{G}_5 h^b_b \partial^a \delta A_a \partial^d \partial^c h_{cd}}{2 \bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} \\
& + \frac{\bar{G}_5 h^c_c \partial^b \delta A_a \partial^d \partial_d h^{ab}}{2 \bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} - \frac{\bar{G}_5 h^b_c \partial_b \delta A_a \partial^d \partial_d h^{ac}}{\bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} + \frac{\bar{G}_5 \delta A_a \partial^a h_{bc} \partial^d \partial_d h^{bc}}{2 \bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} \\
& + \frac{\bar{G}_5 h_{bc} \partial^a \delta A_a \partial^d \partial_d h^{bc}}{\bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} - \frac{\bar{G}_5 \delta A_a \partial^a h^b_b \partial^d \partial_d h^c_c}{2 \bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} - \frac{\bar{G}_5 h^b_b \partial^a \delta A_a \partial^d \partial_d h^c_c}{2 \bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}}
\end{aligned}$$

$$\begin{aligned}
& + \frac{\bar{G}_5 \partial^b \delta A_a \partial^c h^{bd} \partial_d h^a_c}{\bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} - \frac{\bar{G}_5 \partial^b \delta A_a \partial^d h^{bc} \partial_d h^a_c}{\bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} - \frac{\bar{G}_5 \partial^a \delta A_a \partial^c h^{bd} \partial_d h_{bc}}{2\bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}} \\
& + \frac{3\bar{G}_5 \partial^a \delta A_a \partial^d h^{bc} \partial_d h_{bc}}{4\bar{G}_4 \sqrt{\bar{G}_2^{(0,1)}}}.
\end{aligned} \tag{A.4}$$

$$\begin{aligned}
\mathcal{L}_6^{(3)} = & \frac{\sqrt{-\bar{G}_4} \bar{G}_6 \partial^c \delta A_d \partial_c \delta A_a \partial^d \partial^a h^e_e}{\sqrt{2} \bar{G}_4 \bar{G}_2^{(0,1)}} + \frac{\sqrt{-\bar{G}_4} \bar{G}_6 \partial^c \delta A_a \partial^d \partial^c h^e_e \partial_d \delta A^a}{\sqrt{2} \bar{G}_4 \bar{G}_2^{(0,1)}} \\
& - \frac{\sqrt{2} \sqrt{-\bar{G}_4} \bar{G}_6 \partial^c \delta A_a \partial^d \partial^a h^e_e \partial_d \delta A^c}{\bar{G}_4 \bar{G}_2^{(0,1)}} + \frac{\sqrt{2} \sqrt{-\bar{G}_4} \bar{G}_6 \partial^c \delta A_a \partial^d \delta A^c \partial^e \partial^a h_{de}}{\bar{G}_4 \bar{G}_2^{(0,1)}} \\
& - \frac{\sqrt{2} \sqrt{-\bar{G}_4} \bar{G}_6 \partial^c \delta A_d \partial_c \delta A_a \partial^e \partial^d h^a_e}{\bar{G}_4 \bar{G}_2^{(0,1)}} + \frac{\sqrt{2} \sqrt{-\bar{G}_4} \bar{G}_6 \partial^c \delta A_a \partial^d \delta A^c \partial^e \partial^d h^a_e}{\bar{G}_4 \bar{G}_2^{(0,1)}} \\
& - \frac{\sqrt{2} \sqrt{-\bar{G}_4} \bar{G}_6 \partial^c \delta A_a \partial^d \delta A^a \partial^e \partial^d h_{ce}}{\bar{G}_4 \bar{G}_2^{(0,1)}} - \frac{\sqrt{-\bar{G}_4} \bar{G}_6 \partial^a \delta A^c \partial_c \delta A_a \partial^e \partial^d h_{de}}{\sqrt{2} \bar{G}_4 \bar{G}_2^{(0,1)}} \\
& + \frac{\sqrt{-\bar{G}_4} \bar{G}_6 \partial^c \delta A^a \partial_c \delta A_a \partial^e \partial^d h_{de}}{\sqrt{2} \bar{G}_4 \bar{G}_2^{(0,1)}} + \frac{\sqrt{-\bar{G}_4} \bar{G}_6 \partial^c \delta A_d \partial_c \delta A_a \partial^e \partial_e h^{ad}}{\sqrt{2} \bar{G}_4 \bar{G}_2^{(0,1)}} \\
& - \frac{\sqrt{2} \sqrt{-\bar{G}_4} \bar{G}_6 \partial^c \delta A_a \partial^d \delta A^c \partial^e \partial_e h^{ad}}{\bar{G}_4 \bar{G}_2^{(0,1)}} + \frac{\sqrt{-\bar{G}_4} \bar{G}_6 \partial^c \delta A_a \partial^d \delta A^a \partial^e \partial_e h^{cd}}{\sqrt{2} \bar{G}_4 \bar{G}_2^{(0,1)}} \\
& + \frac{\sqrt{-\bar{G}_4} \bar{G}_6 \partial^a \delta A^c \partial_c \delta A_a \partial^e \partial_e h^d_d}{\sqrt{2} \bar{G}_4 \bar{G}_2^{(0,1)}} - \frac{\sqrt{-\bar{G}_4} \bar{G}_6 \partial^c \delta A^a \partial_c \delta A_a \partial^e \partial_e h^d_d}{\sqrt{2} \bar{G}_4 \bar{G}_2^{(0,1)}} \\
& - \frac{\sqrt{-\bar{G}_4} \bar{G}_6 \partial^c \delta A_a \partial^d \partial^a h^{ce} \partial_e \delta A_d}{\sqrt{2} \bar{G}_4 \bar{G}_2^{(0,1)}} + \frac{\sqrt{2} \sqrt{-\bar{G}_4} \bar{G}_6 \partial^c \delta A_a \partial^d \partial^c h^{ae} \partial_e \delta A_d}{\bar{G}_4 \bar{G}_2^{(0,1)}} \\
& - \frac{\sqrt{-\bar{G}_4} \bar{G}_6 \partial^c \delta A_a \partial^e \partial^c h^{ad} \partial_e \delta A_d}{\sqrt{2} \bar{G}_4 \bar{G}_2^{(0,1)}}.
\end{aligned} \tag{A.5}$$

### Quartic Interactions

$$\begin{aligned}
\mathcal{L}_2^{(4)} = & \frac{h^{cd} h_{cd} \partial^a \delta A^b \partial_b \delta A_a}{4\bar{G}_4} - \frac{h^c h^d_d \partial^a \delta A^b \partial_b \delta A_a}{4\bar{G}_4} + \frac{h^2 \partial^a \delta A^b \partial_b \delta A_a}{8\bar{G}_4} \\
& - \frac{h^{cd} h_{cd} \partial^b \delta A^a \partial_b \delta A_a}{4\bar{G}_4} + \frac{h^c h^d_d \partial^b \delta A^a \partial_b \delta A_a}{4\bar{G}_4} - \frac{h^2 \partial^b \delta A^a \partial_b \delta A_a}{8\bar{G}_4} \\
& + \frac{h^{ac} h^d_d \partial^b \delta A_c \partial_b \delta A_a}{2\bar{G}_4} + \frac{h^{bd} h^c_c \partial_b \delta A_a \partial_c \delta A^a}{\bar{G}_4} - \frac{h^{bc} h^d_d \partial_b \delta A_a \partial_c \delta A^a}{2\bar{G}_4} \\
& - \frac{h^d_d \delta A_a \partial^b h^{ac} \partial_c \delta A_b}{\bar{G}_4} + \frac{2h^{cd} \delta A_a \partial^b h^a_d \partial_c \delta A_b}{\bar{G}_4} + \frac{h^d_d \delta A_a \partial^c h^{ab} \partial_c \delta A_b}{\bar{G}_4}
\end{aligned}$$

$$\begin{aligned}
& - \frac{2h^{cd} \delta A_a \partial^c \delta A^b \partial_d h^{ab}}{\bar{G}_4} - \frac{\delta A_a \delta A_b \partial^c h^{bd} \partial_d h^a_c}{\bar{G}_4} + \frac{\delta A_a \delta A_b \partial^d h^{bc} \partial_d h^a_c}{\bar{G}_4} \\
& - \frac{h^{ac} h^{bd} \partial_b \delta A_a \partial_d \delta A_c}{\bar{G}_4} + \frac{\partial^a \delta A^b \partial_b \delta A_a \partial^c \delta A^d \partial_d \delta A_c \bar{G}_2^{(0,2)}}{8 \left( \bar{G}_2^{(0,1)} \right)^2} \\
& - \frac{\partial^b \delta A^a \partial_b \delta A_a \partial^c \delta A^d \partial_d \delta A_c \bar{G}_2^{(0,2)}}{4 \left( \bar{G}_2^{(0,1)} \right)^2} + \frac{\partial^b \delta A^a \partial_b \delta A_a \partial^d \delta A^c \partial_d \delta A_c \bar{G}_2^{(0,2)}}{8 \left( \bar{G}_2^{(0,1)} \right)^2} \\
& + \frac{h^{bc} h_{bc} \delta A^a \delta A_a \bar{G}_2^{(1,0)}}{4 \bar{G}_4 \bar{G}_2^{(0,1)}} - \frac{h^b_b h^c_c \delta A^a \delta A_a \bar{G}_2^{(1,0)}}{4 \bar{G}_4 \bar{G}_2^{(0,1)}} + \frac{h^2 \delta A^a \delta A_a \bar{G}_2^{(1,0)}}{8 \bar{G}_4 \bar{G}_2^{(0,1)}} \\
& - \frac{h^{ab} h^c_c \delta A_a \delta A_b \bar{G}_2^{(1,0)}}{2 \bar{G}_4 \bar{G}_2^{(0,1)}} + \frac{\delta A^a \delta A_a \partial^b \delta A^c \partial_c \delta A_b \bar{G}_2^{(1,1)}}{4 \left( \bar{G}_2^{(0,1)} \right)^2} - \frac{\delta A^a \delta A_a \partial^c \delta A^b \partial_c \delta A_b \bar{G}_2^{(1,1)}}{4 \left( \bar{G}_2^{(0,1)} \right)^2} \\
& + \frac{\delta A^a \delta A_a \delta A^b \delta A_b \bar{G}_2^{(2,0)}}{8 \left( \bar{G}_2^{(0,1)} \right)^2}. \tag{A.6}
\end{aligned}$$

$$\begin{aligned}
\mathcal{L}_3^{(4)} = & - \frac{\sqrt{-\bar{G}_4} \delta A_a \delta A^b \delta A_b \partial^a h^c_c \bar{G}'_3}{2\sqrt{2} \bar{G}_4 \left( \bar{G}_2^{(0,1)} \right)^{3/2}} - \frac{\sqrt{-\bar{G}_4} h^c_c \delta A^b \delta A_b \partial^a \delta A_a \bar{G}'_3}{2\sqrt{2} \bar{G}_4 \left( \bar{G}_2^{(0,1)} \right)^{3/2}} \\
& - \frac{\sqrt{-\bar{G}_4} h^{ba} \delta A_a \delta A_b \partial^c \delta A_c \bar{G}'_3}{\sqrt{2} \bar{G}_4 \left( \bar{G}_2^{(0,1)} \right)^{3/2}} - \frac{\bar{G}_3 \sqrt{-\bar{G}_4} h^2 \delta A^b \partial_b h^a_a}{4\sqrt{2} \bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} \\
& + \frac{\sqrt{2} \bar{G}_3 \sqrt{-\bar{G}_4} h^{ad} h^{ac} h_{cd} \delta A^b \partial_b h^{cd}}{\bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} - \frac{\bar{G}_3 \sqrt{-\bar{G}_4} h^a_a h^{cd} \delta A^b \partial_b h_{cd}}{\sqrt{2} \bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} \\
& - \frac{\bar{G}_3 \sqrt{-\bar{G}_4} h^{ac} h_{ac} \delta A^b \partial_b h^d_d}{2\sqrt{2} \bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} + \frac{\bar{G}_3 \sqrt{-\bar{G}_4} h^a_a h^c_c \delta A^b \partial_b h^d_d}{2\sqrt{2} \bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} \\
& + \frac{\sqrt{2} \bar{G}_3 \sqrt{-\bar{G}_4} h^{ad} h^{ac} h_{cd} \partial^b \delta A_b}{3\bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} - \frac{\bar{G}_3 \sqrt{-\bar{G}_4} h^a_a h_{cd} h^{cd} \partial^b \delta A_b}{2\sqrt{2} \bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} \\
& - \frac{\bar{G}_3 \sqrt{-\bar{G}_4} h^a_a h^c_c h^d_d \partial^b \delta A_b}{6\sqrt{2} \bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} + \frac{\bar{G}_3 \sqrt{-\bar{G}_4} h^3 \partial^b \delta A_b}{4\sqrt{2} \bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}}. \tag{A.7}
\end{aligned}$$

$$\begin{aligned}
\mathcal{L}_4^{(4)} = & \frac{h^3 \partial^b \partial^a h_{ab}}{4\bar{G}_4} - \frac{h^{ab} h^2 \partial^b \partial^a h^c_c}{2\bar{G}_4} - \frac{h^3 \partial^b \partial_b h^a_a}{4\bar{G}_4} \\
& + \frac{h^2 \partial^b h^c_c \partial_b h^a_a}{8\bar{G}_4} - \frac{2h^{ab} h^{cd} \partial_b h_{de} \partial_c h^e_a}{\bar{G}_4} + \frac{h^2 \partial^a h_{ab} \partial^c h^b_c}{2\bar{G}_4}
\end{aligned}$$

$$\begin{aligned}
& - \frac{h^2 \partial^b h_a^a \partial^c h_{bc}}{2\bar{G}_4} + \frac{3h^{ac} h^{ab} \partial_b h^{de} \partial_c h_{de}}{\bar{G}_4} - \frac{3h h^{bc} \partial_b h^{de} \partial_c h_{de}}{2\bar{G}_4} \\
& - \frac{h^{ac} h^{ab} \partial_b h_d^d \partial_c h_e^e}{\bar{G}_4} + \frac{h h^{bc} \partial_b h_d^d \partial_c h_e^e}{2\bar{G}_4} + \frac{h^{ab} h^2 \partial^c \partial^b h_{ac}}{\bar{G}_4} \\
& + \frac{4h^{ac} h^{ab} h^{de} \partial_c \partial_b h_{de}}{\bar{G}_4} - \frac{h^{ab} h^2 \partial^c \partial_c h_{ab}}{2\bar{G}_4} + \frac{h^2 \partial^b h_{ac} \partial^c h_a^b}{4\bar{G}_4} \\
& - \frac{3h^2 \partial^c h_{ab} \partial_c h^{ab}}{8\bar{G}_4} + \frac{4h^{ac} h^{ab} \partial_c h_e^e \partial_d h_b^d}{\bar{G}_4} - \frac{2h h^{bc} \partial_c h_e^e \partial_d h_b^d}{\bar{G}_4} \\
& + \frac{6h^{ab} h^{cd} \partial_c h_{ae} \partial_d h_b^e}{\bar{G}_4} - \frac{4h^{ab} h^{cd} \partial_b h_{ae} \partial_d h_c^e}{\bar{G}_4} + \frac{4h^{ab} h^{cd} \partial_b h_{ac} \partial_d h_e^e}{\bar{G}_4} \\
& - \frac{2h^{ab} h^{cd} \partial_c h_{ab} \partial_d h_e^e}{\bar{G}_4} + \frac{4h^{ac} h^{ab} \partial_c h_{bd} \partial_d h_e^e}{\bar{G}_4} - \frac{2h h^{bc} \partial_c h_{bd} \partial_d h_e^e}{\bar{G}_4} \\
& + \frac{4h^{ac} h^{ab} h_{bd} \partial_d \partial_c h_e^e}{\bar{G}_4} - \frac{2h_a^a h^{bd} h_b^c \partial_d \partial_c h_e^e}{\bar{G}_4} - \frac{h^{ab} h_{ab} h^{cd} \partial_d \partial_c h_e^e}{\bar{G}_4} \\
& + \frac{h_a^a h_b^b h^{cd} \partial_d \partial_c h_e^e}{\bar{G}_4} - \frac{2h^{ac} h^{ab} \partial_d h_e^e \partial^d h_{bc}}{\bar{G}_4} + \frac{h_a^a h^{bc} \partial_d h_e^e \partial^d h_{bc}}{\bar{G}_4} \\
& + \frac{h^{ab} h_{ab} \partial_d h_e^e \partial^d h_c^c}{4\bar{G}_4} - \frac{h_a^a h_b^b \partial_d h_e^e \partial^d h_c^c}{4\bar{G}_4} - \frac{4h^{ac} h^{ab} \partial_d h_b^d \partial_e h_c^e}{\bar{G}_4} \\
& + \frac{2h_a^a h^{bc} \partial_d h_b^d \partial_e h_c^e}{\bar{G}_4} - \frac{8h^{ab} h^{cd} \partial_b h_{ac} \partial_e h_d^e}{\bar{G}_4} + \frac{4h^{ab} h^{cd} \partial_c h_{ab} \partial_e h_d^e}{\bar{G}_4} \\
& - \frac{8h^{ac} h^{ab} \partial_c h_{bd} \partial_e h^{de}}{\bar{G}_4} + \frac{4h_a^a h^{bc} \partial_c h_{bd} \partial_e h^{de}}{\bar{G}_4} + \frac{h^{ab} h_{ab} \partial_c h^{cd} \partial_e h_d^e}{\bar{G}_4} \\
& - \frac{h_a^a h_b^b \partial_c h^{cd} \partial_e h_d^e}{\bar{G}_4} + \frac{4h^{ac} h^{ab} \partial^d h_{bc} \partial_e h_d^e}{\bar{G}_4} - \frac{2h_a^a h^{bc} \partial^d h_{bc} \partial_e h_d^e}{\bar{G}_4} \\
& - \frac{h^{ab} h_{ab} \partial^d h_c^c \partial_e h_d^e}{\bar{G}_4} + \frac{h_a^a h_b^b \partial^d h_c^c \partial_e h_d^e}{\bar{G}_4} - \frac{8h^{ac} h^{ab} h^{de} \partial_e \partial_c h_{bd}}{\bar{G}_4} \\
& + \frac{2h_a^a h^{bc} h^{de} \partial_e \partial_c h_{bd}}{\bar{G}_4} + \frac{4h^{ac} h^{ab} h^{de} \partial_e \partial_d h_{bc}}{\bar{G}_4} - \frac{2h_a^a h^{bc} h^{de} \partial_e \partial_d h_{bc}}{\bar{G}_4} \\
& - \frac{8h^{ac} h^{ab} h_b^d \partial_e \partial_d h_c^e}{\bar{G}_4} + \frac{4h_a^a h^{bd} h_b^c \partial_e \partial_d h_c^e}{\bar{G}_4} + \frac{2h^{ab} h_{ab} h^{cd} \partial_e \partial_d h_c^e}{\bar{G}_4} \\
& - \frac{2h_a^a h_b^b h^{cd} \partial_e \partial_d h_c^e}{\bar{G}_4} + \frac{2h^{ac} h^{ab} h_{bc} \partial_e \partial_d h^{de}}{3\bar{G}_4} - \frac{h_a^a h^{bc} h_{bc} \partial_e \partial_d h^{de}}{2\bar{G}_4} \\
& - \frac{h_a^a h_b^b h_c^c \partial_e \partial_d h^{de}}{6\bar{G}_4} + \frac{4h^{ac} h^{ab} h_b^d \partial_e \partial^e h_{cd}}{\bar{G}_4} - \frac{2h_a^a h^{bd} h_b^c \partial_e \partial^e h_{cd}}{\bar{G}_4} \\
& - \frac{h^{ab} h_{ab} h^{cd} \partial_e \partial^e h_{cd}}{\bar{G}_4} + \frac{h_a^a h_b^b h^{cd} \partial_e \partial^e h_{cd}}{\bar{G}_4} - \frac{2h^{ac} h^{ab} h_{bc} \partial_e \partial^e h_d^d}{3\bar{G}_4}
\end{aligned}$$



$$\begin{aligned}
& + \frac{h_a^a h^{bc} h_{bc} \partial_e \partial^e h_d^d}{2\bar{G}_4} + \frac{h_a^a h^b h_c^c \partial_e \partial^e h_d^d}{6\bar{G}_4} + \frac{4 h^{ab} h^{cd} \partial_d h_{ce} \partial^e h_{ab}}{\bar{G}_4} \\
& - \frac{h^{ab} h^{cd} \partial_e h^{cd} \partial_e h_{ab}}{\bar{G}_4} - \frac{4 h^{ab} h^{cd} \partial_d h^{be} \partial_e h_{ac}}{\bar{G}_4} + \frac{3 h^{ab} h^{cd} \partial_e h^{bd} \partial_e h_{ac}}{\bar{G}_4} \\
& - \frac{4 h^{ac} h^{ab} \partial_c h^e \partial_e h_{bc}}{\bar{G}_4} + \frac{2 h_a^a h^{bc} \partial_c h^e \partial_e h_{bd}}{\bar{G}_4} - \frac{2 h^{ac} h^{ab} \partial_d h_{cd} \partial_e h_{bc}}{\bar{G}_4} \\
& + \frac{h^{ab} h^{bc} \partial_d h_{ce} \partial_d h_e}{4\bar{G}_4} + \frac{6 h^{ac} h^{ab} \partial_d h_{ce} \partial_d h_e}{\bar{G}_4} - \frac{3 h^{ab} h^{bc} \partial_d h_{ce} \partial_d h_e}{\bar{G}_4} \\
& + \frac{h^{ab} h^{bc} \partial_e h_{cd} \partial_e h_{dc}}{2\bar{G}_4} - \frac{h^{ab} h^{ac} \partial_e h_{ed} \partial_d h_c}{2\bar{G}_4} - \frac{3 h^{ab} h^{bc} \partial_e h_{cd} \partial_e h_{db}}{4\bar{G}_4} \\
& + \frac{3 h^{ab} h^{bc} \partial_e h_{cd} \partial_e h_{db}}{4\bar{G}_4} + \frac{\delta A_a \delta A^a \partial_b \delta A^b \partial_c \delta A^c \bar{G}_4^{(2)}}{2 \left( \bar{G}_2^{(0,1)} \right)^2} \\
& - \frac{\delta A_a \delta A^a \partial_b \delta A_c \partial^c \delta A^b \bar{G}_4^{(2)}}{2 \left( \bar{G}_2^{(0,1)} \right)^2} + \frac{\delta A^a \delta A^b \partial_a h_{cd} \partial_b h^{cd} \bar{G}_4^{(1)}}{2 \bar{G}_4 \bar{G}_2^{(0,1)}} \\
& - \frac{\delta A^a \delta A^b \partial_a h_c^c \partial_b h_d^d \bar{G}_4^{(1)}}{2 \bar{G}_4 \bar{G}_2^{(0,1)}} + \frac{2 h_{cd} \delta A^a \partial_a h^{cd} \partial_b \delta A^b \bar{G}_4^{(1)}}{\bar{G}_4 \bar{G}_2^{(0,1)}} \\
& - \frac{h_c^c \delta A^a \partial_a h_d^d \partial_b \delta A^b \bar{G}_4^{(1)}}{\bar{G}_4 \bar{G}_2^{(0,1)}} + \frac{h_{cd} h^{cd} \partial_a \delta A^a \partial_b \delta A^b \bar{G}_4^{(1)}}{2 \bar{G}_4 \bar{G}_2^{(0,1)}} \\
& - \frac{h_c^c h_d^d \partial_a \delta A^a \partial_b \delta A^b \bar{G}_4^{(1)}}{2 \bar{G}_4 \bar{G}_2^{(0,1)}} + \frac{(h_a^a)^2 \partial_a \delta A^a \partial_b \delta A^b \bar{G}_4^{(1)}}{4 \bar{G}_4 \bar{G}_2^{(0,1)}} \\
& - \frac{h_{cd} h^{cd} \partial_a \delta A_b \partial^b \delta A^a \bar{G}_4^{(1)}}{2 \bar{G}_4 \bar{G}_2^{(0,1)}} + \frac{h_c^c h_d^d \partial_a \delta A_b \partial^b \delta A^a \bar{G}_4^{(1)}}{2 \bar{G}_4 \bar{G}_2^{(0,1)}} \\
& - \frac{(h_a^a)^2 \partial_a \delta A_b \partial^b \delta A^a \bar{G}_4^{(1)}}{4 \bar{G}_4 \bar{G}_2^{(0,1)}} - \frac{h_{bc} \delta A_a \delta A^a \partial_c \partial_b h_d^d \bar{G}_4^{(1)}}{\bar{G}_4 \bar{G}_2^{(0,1)}} \\
& + \frac{\delta A_a \delta A^a \partial_c h_d^d \partial^c h_b^b \bar{G}_4^{(1)}}{4 \bar{G}_4 \bar{G}_2^{(0,1)}} + \frac{h_d^d \delta A^a \partial_a h_{bc} \partial^c \delta A^b \bar{G}_4^{(1)}}{\bar{G}_4 \bar{G}_2^{(0,1)}} \\
& - \frac{2 h^{cd} \delta A^a \partial_a h_{bd} \partial_c \delta A^b \bar{G}_4^{(1)}}{\bar{G}_4 \bar{G}_2^{(0,1)}} + \frac{h_d^d \delta A^a \partial_b h_{ac} \partial^c \delta A^b \bar{G}_4^{(1)}}{\bar{G}_4 \bar{G}_2^{(0,1)}} \\
& - \frac{2 h^{cd} \delta A^a \partial_b h_{ad} \partial_c \delta A^b \bar{G}_4^{(1)}}{\bar{G}_4 \bar{G}_2^{(0,1)}} - \frac{h_d^d \delta A^a \partial_c h_{ab} \partial^c \delta A^b \bar{G}_4^{(1)}}{\bar{G}_4 \bar{G}_2^{(0,1)}} \\
& + \frac{2 h^{cd} \delta A^a \partial^c \delta A^b \partial_d h_{ab} \bar{G}_4^{(1)}}{\bar{G}_4 \bar{G}_2^{(0,1)}} + \frac{\delta A_a \delta A^a \partial_b h^{bc} \partial_d h_c^d \bar{G}_4^{(1)}}{\bar{G}_4 \bar{G}_2^{(0,1)}}
\end{aligned}$$

$$\begin{aligned}
& - \frac{\delta A_a \delta A^a \partial^c h_b^b \partial_d h^{cd} \bar{G}_4^{(1)}}{\bar{G}_4 \bar{G}_2^{(0,1)}} + \frac{2h^{bc} \delta A_a \delta A^a \partial_d \partial_c h^{bd} \bar{G}_4^{(1)}}{\bar{G}_4 \bar{G}_2^{(0,1)}} \\
& - \frac{h_b^b \delta A_a \delta A^a \partial_d \partial_c h^{cd} \bar{G}_4^{(1)}}{2\bar{G}_4 \bar{G}_2^{(0,1)}} - \frac{h_{ab} \delta A^a \delta A^b \partial_d \partial_c h^{cd} \bar{G}_4^{(1)}}{\bar{G}_4 \bar{G}_2^{(0,1)}} \\
& - \frac{h^{bc} \delta A_a \delta A^a \partial_d \partial^d h_{bc} \bar{G}_4^{(1)}}{\bar{G}_4 \bar{G}_2^{(0,1)}} + \frac{h_b^b \delta A_a \delta A^a \partial_d \partial^d h_c^c \bar{G}_4^{(1)}}{2\bar{G}_4 \bar{G}_2^{(0,1)}} \\
& + \frac{h_{ab} \delta A^a \delta A^b \partial_d \partial^d h_c^c \bar{G}_4^{(1)}}{\bar{G}_4 \bar{G}_2^{(0,1)}} + \frac{\delta A^a \delta A^b \partial_c h_{bd} \partial^d h_a^c \bar{G}_4^{(1)}}{\bar{G}_4 \bar{G}_2^{(0,1)}} \\
& - \frac{\delta A^a \delta A^b \partial_d h_{bc} \partial^d h_a^c \bar{G}_4^{(1)}}{\bar{G}_4 \bar{G}_2^{(0,1)}} + \frac{\delta A_a \delta A^a \partial_c h_{bd} \partial^d h^{bc} \bar{G}_4^{(1)}}{2\bar{G}_4 \bar{G}_2^{(0,1)}} \\
& - \frac{3\delta A_a \delta A^a \partial_d h_{bc} \partial^d h^{bc} \bar{G}_4^{(1)}}{4\bar{G}_4 \bar{G}_2^{(0,1)}}. \tag{A.8}
\end{aligned}$$

$$\begin{aligned}
\mathcal{L}_5^{(4)} = & \frac{\sqrt{-\bar{G}_4} \bar{g}_5 h_d^d \partial^b \delta A^a \partial_c \delta A_b \partial^c \delta A_a}{\sqrt{2} \bar{G}_4 \left( \bar{G}_2^{(0,1)} \right)^{3/2}} - \frac{\sqrt{-\bar{G}_4} \bar{g}_5 h_d^d \partial_a \delta A^c \partial^b \delta A^a \partial^c \delta A_b}{\sqrt{2} \bar{G}_4 \left( \bar{G}_2^{(0,1)} \right)^{3/2}} \\
& + \frac{\sqrt{2} \sqrt{-\bar{G}_4} \bar{g}_5 h^{ad} \partial^b \delta A_a \partial_c \delta A_d \partial^c \delta A_b}{\bar{G}_4 \left( \bar{G}_2^{(0,1)} \right)^{3/2}} + \frac{\sqrt{-\bar{G}_4} \bar{g}_5 \delta A^a \partial_a h_d^d \partial_b \delta A^c \partial^c \delta A^b}{\sqrt{2} \bar{G}_4 \left( \bar{G}_2^{(0,1)} \right)^{3/2}} \\
& + \frac{\sqrt{-\bar{G}_4} \bar{g}_5 h_d^d \partial_a \delta A^a \partial_b \delta A^c \partial^c \delta A^b}{\sqrt{2} \bar{G}_4 \left( \bar{G}_2^{(0,1)} \right)^{3/2}} - \frac{\sqrt{-\bar{G}_4} \bar{g}_5 \delta A^a \partial_a h_d^d \partial_c \delta A^b \partial^c \delta A_b}{\sqrt{2} \bar{G}_4 \left( \bar{G}_2^{(0,1)} \right)^{3/2}} \\
& - \frac{\sqrt{-\bar{G}_4} \bar{g}_5 h_d^d \partial_a \delta A^a \partial_c \delta A^b \partial^c \delta A_b}{\sqrt{2} \bar{G}_4 \left( \bar{G}_2^{(0,1)} \right)^{3/2}} + \frac{\sqrt{-\bar{G}_4} \bar{g}_5 \delta A^a \partial_a h^{bd} \partial_c \delta A_d \partial^c \delta A_b}{\sqrt{2} \bar{G}_4 \left( \bar{G}_2^{(0,1)} \right)^{3/2}} \\
& - \frac{\sqrt{2} \sqrt{-\bar{G}_4} \bar{g}_5 h^{bd} \partial_a \delta A^a \partial_c \delta A_d \partial^c \delta A_b}{\bar{G}_4 \left( \bar{G}_2^{(0,1)} \right)^{3/2}} - \frac{\sqrt{2} \sqrt{-\bar{G}_4} \bar{g}_5 h^{cd} \partial^b \delta A^a \partial^c \delta A_a \partial^d \delta A_b}{\bar{G}_4 \left( \bar{G}_2^{(0,1)} \right)^{3/2}} \\
& + \frac{\sqrt{-\bar{G}_4} \bar{g}_5 \delta A^a \partial_a h^{cd} \partial^c \delta A^b \partial^d \delta A_b}{\sqrt{2} \bar{G}_4 \left( \bar{G}_2^{(0,1)} \right)^{3/2}} + \frac{\sqrt{2} \sqrt{-\bar{G}_4} \bar{g}_5 h^{cd} \partial_a \delta A^a \partial^c \delta A^b \partial^d \delta A_b}{\bar{G}_4 \left( \bar{G}_2^{(0,1)} \right)^{3/2}} \\
& - \frac{\sqrt{2} \sqrt{-\bar{G}_4} \bar{g}_5 \delta A^a \partial_a h^{bd} \partial^c \delta A_b \partial^d \delta A_c}{\bar{G}_4 \left( \bar{G}_2^{(0,1)} \right)^{3/2}} - \frac{2\sqrt{2} \sqrt{-\bar{G}_4} \bar{g}_5 \delta A^a \partial^b \delta A_b \partial^d h_{ac} \partial_d \delta A^c}{\bar{G}_4 \left( \bar{G}_2^{(0,1)} \right)^{3/2}}
\end{aligned}$$

$$\begin{aligned}
& + \frac{\sqrt{-\bar{G}_4} \delta A^a \partial_a h_d^d \partial^b \delta A_b \partial^c \delta A_c \bar{G}'_5}{2\sqrt{2}\bar{G}_4 \left(\bar{G}_2^{(0,1)}\right)^{3/2}} + \frac{\sqrt{-\bar{G}_4} h_d^d \partial^a \delta A_a \partial^b \delta A_b \partial^c \delta A_c \bar{G}'_5}{6\sqrt{2}\bar{G}_4 \left(\bar{G}_2^{(0,1)}\right)^{3/2}} \\
& + \frac{\sqrt{-\bar{G}_4} h_d^d \partial^a \delta A^c \partial^b \delta A_a \partial^c \delta A_b \bar{G}'_5}{3\sqrt{2}\bar{G}_4 \left(\bar{G}_2^{(0,1)}\right)^{3/2}} - \frac{\sqrt{-\bar{G}_4} \delta A^a \partial_a h_d^d \partial^b \delta A^c \partial^c \delta A_b \bar{G}'_5}{2\sqrt{2}\bar{G}_4 \left(\bar{G}_2^{(0,1)}\right)^{3/2}} \\
& - \frac{\sqrt{-\bar{G}_4} h_d^d \partial^a \delta A_a \partial^b \delta A^c \partial^c \delta A_b \bar{G}'_5}{2\sqrt{2}\bar{G}_4 \left(\bar{G}_2^{(0,1)}\right)^{3/2}} + \frac{\sqrt{-\bar{G}_4} \delta A_a \delta A^a \partial^c \partial^b h_d^d \partial_c \delta A_b \bar{G}'_5}{2\sqrt{2}\bar{G}_4 \left(\bar{G}_2^{(0,1)}\right)^{3/2}} \\
& - \frac{\sqrt{-\bar{G}_4} \delta A_a \delta A^a \partial^c \delta A^b \partial^d \partial^b h_{cd} \bar{G}'_5}{2\sqrt{2}\bar{G}_4 \left(\bar{G}_2^{(0,1)}\right)^{3/2}} - \frac{\sqrt{-\bar{G}_4} \delta A_a \delta A^a \partial^c \delta A^b \partial^d \partial^c h_{bd} \bar{G}'_5}{2\sqrt{2}\bar{G}_4 \left(\bar{G}_2^{(0,1)}\right)^{3/2}} \\
& + \frac{\sqrt{-\bar{G}_4} \delta A_a \delta A^a \partial^b \delta A_b \partial^d \partial^c h_{cd} \bar{G}'_5}{2\sqrt{2}\bar{G}_4 \left(\bar{G}_2^{(0,1)}\right)^{3/2}} + \frac{\sqrt{-\bar{G}_4} \delta A_a \delta A^a \partial^c \delta A^b \partial^d \partial_d h_{bc} \bar{G}'_5}{2\sqrt{2}\bar{G}_4 \left(\bar{G}_2^{(0,1)}\right)^{3/2}} \\
& - \frac{\sqrt{-\bar{G}_4} \delta A_a \delta A^a \partial^b \delta A_b \partial^d \partial_d h_c^c \bar{G}'_5}{2\sqrt{2}\bar{G}_4 \left(\bar{G}_2^{(0,1)}\right)^{3/2}} - \frac{\sqrt{2}\sqrt{-\bar{G}_4} \bar{G}_5 h^{cd} \partial^a h_c^e \partial^b h_{de} \partial_b \delta A_a}{\bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} \\
& + \frac{\sqrt{-\bar{G}_4} \bar{G}_5 h_c^c \partial^a h^{de} \partial^b h_{de} \partial_b \delta A_a}{2\sqrt{2}\bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} + \frac{\sqrt{-\bar{G}_4} \bar{G}_5 (h_a^a)^2 \partial^b \partial^a h_c^c \partial_b \delta A_a}{4\sqrt{2}\bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} \\
& - \frac{\sqrt{2}\sqrt{-\bar{G}_4} \bar{G}_5 h_{ce} h^{cd} \partial^b \partial^a h_d^e \partial_b \delta A_a}{\bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} + \frac{\sqrt{-\bar{G}_4} \bar{G}_5 h_c^c h^{de} \partial^b \partial^a h_{de} \partial_b \delta A_a}{\sqrt{2}\bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} \\
& + \frac{\sqrt{-\bar{G}_4} \bar{G}_5 h^{cd} h_{cd} \partial^b \partial^a h_e^e \partial_b \delta A_a}{2\sqrt{2}\bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} - \frac{\sqrt{-\bar{G}_4} \bar{G}_5 (h_a^a)^2 \partial^b \partial^a h_e^e \partial_b \delta A_a}{2\sqrt{2}\bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} \\
& + \frac{\sqrt{-\bar{G}_4} \bar{G}_5 h^{bc} \delta A^a \partial_a h^{de} \partial_b \partial_c h_{de}}{2\sqrt{2}\bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} + \frac{3\sqrt{-\bar{G}_4} \bar{G}_5 h^{bc} \partial^a \delta A_a \partial_b h^{de} \partial_c h_{de}}{2\sqrt{2}\bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} \\
& - \frac{\sqrt{-\bar{G}_4} \bar{G}_5 h_{bc} \partial^a h^{de} \partial^b \delta A_a \partial^c h_{de}}{\sqrt{2}\bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} - \frac{\sqrt{-\bar{G}_4} \bar{G}_5 h^{bc} \partial^a \delta A_a \partial_b h_d^d \partial_c h_e^e}{2\sqrt{2}\bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} \\
& - \frac{\bar{G}_5 \sqrt{-\bar{G}_4} (h_a^a)^2 \partial^b \delta A^a \partial^c \partial_a h_{bc}}{4\sqrt{2}\bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} - \frac{\sqrt{2}\sqrt{-\bar{G}_4} \bar{G}_5 h_{bc} h_{de} \partial^b \delta A^a \partial^c \partial_a h^{de}}{\bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} \\
& - \frac{\sqrt{-\bar{G}_4} \bar{G}_5 \delta A^a \partial_a h^{de} \partial^c h_{de} \partial_c h_b^b}{2\sqrt{2}\bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} + \frac{\sqrt{2}\sqrt{-\bar{G}_4} \bar{G}_5 h^{bc} \partial^a \delta A_a \partial_c h_e^e \partial_d h_b^d}{\bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}}
\end{aligned}$$

$$\begin{aligned}
& - \frac{\sqrt{2}\sqrt{-\bar{G}_4} \bar{G}_5 h^{cd} \partial^b \delta A_a \partial_c h^{ae} \partial_d h^b_e}{\bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} + \frac{\sqrt{2}\sqrt{-\bar{G}_4} \bar{G}_5 h^{cd} \partial^a h^{be} \partial_b \delta A_a \partial_d h_{ce}}{\bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} \\
& + \frac{\sqrt{2}\sqrt{-\bar{G}_4} \bar{G}_5 h^{cd} \partial^b h^{ae} \partial_b \delta A_a \partial_d h_{ce}}{\bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} - \frac{\sqrt{-\bar{G}_4} \bar{G}_5 h^{cd} \partial^a h^b_c \partial_b \delta A_a \partial_d h^e_e}{\sqrt{2} \bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} \\
& + \frac{\sqrt{-\bar{G}_4} \bar{G}_5 h^c_c \partial^a h^{bd} \partial_b \delta A_a \partial_d h^e_e}{2\sqrt{2} \bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} - \frac{\sqrt{-\bar{G}_4} \bar{G}_5 h^c_b \partial^a h^d_c \partial_a \delta A^b \partial_d h^e_e}{\sqrt{2} \bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} \\
& + \frac{\sqrt{-\bar{G}_4} \bar{G}_5 h^c_c \partial^a h^{bd} \partial_a \delta A_b \partial_d h^e_e}{2\sqrt{2} \bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} - \frac{\sqrt{-\bar{G}_4} \bar{G}_5 h^{cd} \delta A^a \partial_a h^{be} \partial_b \partial_c h_{de}}{\sqrt{2} \bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} \\
& + \frac{\sqrt{2}\sqrt{-\bar{G}_4} \bar{G}_5 h^{bc} \delta A^a \partial_a h^d_c \partial_d \partial_b h^e_e}{\bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} - \frac{\sqrt{-\bar{G}_4} \bar{G}_5 h^b_b \delta A^a \partial_a h^{cd} \partial_d \partial_c h^e_e}{2\sqrt{2} \bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} \\
& - \frac{\sqrt{-\bar{G}_4} \bar{G}_5 h_{bc} \partial^a \delta A_a \partial_d h^e_e \partial^d h^{bc}}{\sqrt{2} \bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} + \frac{\delta A^a \partial_a h^b_b \partial_d h^e_e \partial^d h^c_c \bar{G}_5 \sqrt{-\bar{G}_4}}{4\sqrt{2} \bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} \\
& + \frac{h^b_b \partial^a \delta A_a \partial_d h^e_e \partial^d h^c_c \bar{G}_5 \sqrt{-\bar{G}_4}}{4\sqrt{2} \bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} + \frac{\sqrt{2} \sqrt{-\bar{G}_4} \bar{G}_5 h_{cd} \partial_b \delta A_a \partial^c h^a_e \partial^e h^{bd}}{\bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} \\
& - \frac{\sqrt{-\bar{G}_4} \bar{G}_5 h_{cd} \partial^a h_{be} \partial^b \delta A_a \partial^e h^{cd}}{\sqrt{2} \bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} - \frac{\sqrt{2} \sqrt{-\bar{G}_4} \bar{G}_5 h_{bc} \partial^a \delta A_a \partial_d h^b_d \partial_e h^{ce}}{\bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} \\
& + \frac{\sqrt{-\bar{G}_4} \bar{G}_5 \delta A^a \partial_a h_{bc} \partial_d h^{bc} \partial_e h^{de}}{2\sqrt{2} \bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} + \frac{3\sqrt{-\bar{G}_4} \bar{G}_5 h_{bc} \partial^a \delta A_a \partial_d h^{be} \partial^d h^c_e}{2\sqrt{2} \bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} \\
& + \frac{\sqrt{-\bar{G}_4} \bar{G}_5 h^b_b h^c_c \partial^a \delta A_a \partial^e \partial^d h_{de}}{2\sqrt{2} \bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} + \frac{\sqrt{-\bar{G}_4} \bar{G}_5 h^{cd} h_{cd} \partial_b \delta A^a \partial^e \partial_e h_a^b}{2\sqrt{2} \bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} \\
& - \frac{\sqrt{-\bar{G}_4} \bar{G}_5 h^c_c h^d_d \partial_b \delta A^a \partial^e \partial_e h_a^b}{2\sqrt{2} \bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} + \frac{\sqrt{-\bar{G}_4} \bar{G}_5 h^c_b h^d_d \partial^b \delta A^a \partial^e \partial_e h_{ac}}{\sqrt{2} \bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} \\
& - \frac{\sqrt{2}\sqrt{-\bar{G}_4} \bar{G}_5 h^c_b h^d_c \partial^b \delta A^a \partial^e \partial_e h_{ad}}{\bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} - \frac{\sqrt{-\bar{G}_4} \bar{G}_5 h^{bc} \delta A^a \partial_a h^d_d \partial^e \partial_e h_{bc}}{\sqrt{2} \bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} \\
& + \frac{\sqrt{2}\sqrt{-\bar{G}_4} \bar{G}_5 h^{bc} \delta A^a \partial_a h^b_d \partial^e \partial_e h_{cd}}{\bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} - \frac{\sqrt{-\bar{G}_4} \bar{G}_5 h^b_b \delta A^a \partial_a h^{cd} \partial^e \partial_e h_{cd}}{2\sqrt{2} \bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}}
\end{aligned}$$

$$+ \frac{\sqrt{2}\sqrt{-\bar{G}_4} \bar{G}_5 h_b^d h^{bc} \partial^a \delta A_a \partial^e \partial_e h_{cd}}{\bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}} - \frac{\sqrt{-\bar{G}_4} \bar{G}_5 h^b{}_b h^{cd} \partial^a \delta A_a \partial^e \partial_e h_{cd}}{\sqrt{2} \bar{G}_4^2 \sqrt{\bar{G}_2^{(0,1)}}}. \quad (\text{A.9})$$

$$\begin{aligned} \mathcal{L}_6^{(4)} = & - \frac{\partial_b \delta A_a \partial^c \delta A_d \partial_c \delta A^a \partial^d \delta A^b \bar{G}'_6}{(\bar{G}_2^{(0,1)})^2} + \frac{\partial^a \delta A_a \partial_c \delta A_b \partial^d \delta A^c \partial_d \delta A^b \bar{G}'_6}{(\bar{G}_2^{(0,1)})^2} \\ & + \frac{\partial^a \delta A_d \partial_b \delta A_a \partial_c \delta A^b \partial^d \delta A^c \bar{G}'_6}{(\bar{G}_2^{(0,1)})^2} - \frac{\partial^a \delta A_a \partial^b \delta A_d \partial_c \delta A_b \partial^d \delta A^c \bar{G}'_6}{(\bar{G}_2^{(0,1)})^2} \\ & + \frac{1}{2} \frac{\partial^a \delta A_a \partial^b \delta A_b \partial^c \delta A_d \partial^d \delta A_c \bar{G}'_6}{(\bar{G}_2^{(0,1)})^2} - \frac{1}{2} \frac{\partial^a \delta A_b \partial^b \delta A_a \partial^c \delta A_d \partial^d \delta A_c \bar{G}'_6}{(\bar{G}_2^{(0,1)})^2} \\ & + \frac{1}{2} \frac{\partial^b \delta A_a \partial_b \delta A^a \partial^c \delta A_d \partial^d \delta A_c \bar{G}'_6}{(\bar{G}_2^{(0,1)})^2} - \frac{1}{2} \frac{\partial^a \delta A_a \partial^b \delta A_b \partial^d \delta A_c \partial^c \delta A_d \bar{G}'_6}{(\bar{G}_2^{(0,1)})^2} \\ & - \frac{\bar{G}_6 \partial^a h_{de} \partial^b \delta A_c \partial_b \delta A_a \partial^c h^{de}}{2 \bar{G}_4 \bar{G}_2^{(0,1)}} - \frac{\bar{G}_6 h_{de} \partial^b \delta A_c \partial_b \delta A_a \partial^c \partial^a h^{de}}{\bar{G}_4 \bar{G}_2^{(0,1)}} \\ & + \frac{1}{2} \frac{\bar{G}_6 h_d^d \partial^b \delta A_c \partial_b \delta A_a \partial^c \partial^a h^e{}_e}{\bar{G}_4 \bar{G}_2^{(0,1)}} - \frac{1}{2} \frac{\bar{G}_6 \partial^b h_{de} \partial_b \delta A_a \partial^c h^{de} \partial_c \delta A^a}{\bar{G}_4 \bar{G}_2^{(0,1)}} \\ & - \frac{\bar{G}_6 h_{de} \partial_b \delta A_a \partial^c \partial^b h^{de} \partial_c \delta A^a}{\bar{G}_4 \bar{G}_2^{(0,1)}} + \frac{1}{2} \frac{\bar{G}_6 h_d^d \partial_b \delta A_a \partial^c \partial^b h^e{}_e \partial_c \delta A^a}{\bar{G}_4 \bar{G}_2^{(0,1)}} \\ & + \frac{\bar{G}_6 \partial^a h_{de} \partial_b \delta A_a \partial^c h^{de} \partial_c \delta A^b}{\bar{G}_4 \bar{G}_2^{(0,1)}} + \frac{2 \bar{G}_6 h_{de} \partial_b \delta A_a \partial^c \partial^a h^{de} \partial_c \delta A^b}{\bar{G}_4 \bar{G}_2^{(0,1)}} \\ & - \frac{\bar{G}_6 h_d^d \partial_b \delta A_a \partial^c \partial^a h^e{}_e \partial_c \delta A^b}{\bar{G}_4 \bar{G}_2^{(0,1)}} - \frac{\bar{G}_6 \partial^b \delta A_c \partial_b \delta A_a \partial^c h^a{}_d \partial^d h^e{}_e}{\bar{G}_4 \bar{G}_2^{(0,1)}} \\ & - \frac{\bar{G}_6 \partial_b \delta A_a \partial^c h^b{}_d \partial_c \delta A^a \partial^d h^e{}_e}{\bar{G}_4 \bar{G}_2^{(0,1)}} + \frac{\bar{G}_6 \partial^a h^c{}_d \partial_b \delta A_a \partial_c \delta A^b \partial^d h^e{}_e}{\bar{G}_4 \bar{G}_2^{(0,1)}} \\ & + \frac{\bar{G}_6 \partial_b \delta A_a \partial^c h^a{}_d \partial_c \delta A^b \partial^d h^e{}_e}{\bar{G}_4 \bar{G}_2^{(0,1)}} + \frac{2 \bar{G}_6 h^c{}_d \partial_b \delta A_a \partial_c \delta A^b \partial^d \partial^a h^e{}_e}{\bar{G}_4 \bar{G}_2^{(0,1)}} \\ & + \frac{2 \bar{G}_6 \delta A_a \partial^c h^a{}_d \partial_c \delta A_b \partial^d \partial^b h^e{}_e}{\bar{G}_4 \bar{G}_2^{(0,1)}} - \frac{\bar{G}_6 h_{cd} \partial^a \delta A_b \partial_b \delta A_a \partial^d \partial^c h^e{}_e}{\bar{G}_4 \bar{G}_2^{(0,1)}} \\ & + \frac{\bar{G}_6 h_{cd} \partial^b \delta A_a \partial_b \delta A^a \partial^d \partial^c h^e{}_e}{\bar{G}_4 \bar{G}_2^{(0,1)}} - \frac{2 \bar{G}_6 h^b{}_d \partial_b \delta A_a \partial_c \delta A^a \partial^d \partial^c h^e{}_e}{\bar{G}_4 \bar{G}_2^{(0,1)}} \\ & - \frac{2 \bar{G}_6 \delta A_a \partial^b h^a{}_d \partial_c \delta A_b \partial^d \partial^c h^e{}_e}{\bar{G}_4 \bar{G}_2^{(0,1)}} + \frac{2 \bar{G}_6 \delta A_a \partial_c \delta A_b \partial^d \partial^c h^e{}_e \partial_d h^{ab}}{\bar{G}_4 \bar{G}_2^{(0,1)}} \\ & + \frac{1}{2} \frac{\bar{G}_6 \partial^b \delta A_c \partial_b \delta A_a \partial^d h^e{}_e \partial_d h^{ac}}{\bar{G}_4 \bar{G}_2^{(0,1)}} - \frac{\bar{G}_6 \partial_b \delta A_a \partial_c \delta A^b \partial^d h^e{}_e \partial_d h^{ac}}{\bar{G}_4 \bar{G}_2^{(0,1)}} \end{aligned}$$

$$\begin{aligned}
& - \frac{2\bar{G}_6 \delta A_a \partial_c \delta A_b \partial^d \partial^b h^e_e \partial_d h^{ac}}{\bar{G}_4 \bar{G}_2^{(0,1)}} + \frac{1}{2} \frac{\bar{G}_6 \partial_b \delta A_a \partial_c \delta A^a \partial^d h^e_e \partial_d h^{bc}}{\bar{G}_4 \bar{G}_2^{(0,1)}} \\
& + \frac{1}{4} \frac{\bar{G}_6 \partial^a \delta A_b \partial_b \delta A_a \partial^d h^e_e \partial_d h^c_c}{\bar{G}_4 \bar{G}_2^{(0,1)}} - \frac{1}{4} \frac{\bar{G}_6 \partial^b \delta A_a \partial_b \delta A^a \partial^d h^e_e \partial_d h^c_c}{\bar{G}_4 \bar{G}_2^{(0,1)}} \\
& - \frac{\bar{G}_6 \partial^b h^{de} \partial_b \delta A_a \partial^c h^a_e \partial_d \delta A_c}{\bar{G}_4 \bar{G}_2^{(0,1)}} + \frac{1}{2} \frac{\bar{G}_6 \partial^a h^{de} \partial_b \delta A_a \partial^c h^b_e \partial_d \delta A_c}{\bar{G}_4 \bar{G}_2^{(0,1)}} \\
& - \frac{\bar{G}_6}{2\bar{G}_4 \bar{G}_2^{(0,1)}} h^e_e \partial_b \delta A_a \partial^c \partial^a h^b_d \partial^d \delta A_c + \frac{2\bar{G}_6}{\bar{G}_4 \bar{G}_2^{(0,1)}} h_b^e \partial^b \delta A_a \partial^c \partial^a h_{de} \partial^d \delta A_c \\
& - \frac{\bar{G}_6}{\bar{G}_4 \bar{G}_2^{(0,1)}} h_{bd} \partial^b \delta A_a \partial^c \partial^a h^e_e \partial^d \delta A_c + \frac{\bar{G}_6}{\bar{G}_4 \bar{G}_2^{(0,1)}} h^e_e \partial^b \delta A_a \partial^c \partial^b h^a_d \partial^d \delta A_c \\
& + \frac{\bar{G}_6}{2\bar{G}_4 \bar{G}_2^{(0,1)}} \partial^b h_{ce} \partial_b \delta A_a \partial^d h^{ae} \partial_d \delta A^c + \frac{\bar{G}_6}{\bar{G}_4 \bar{G}_2^{(0,1)}} \partial^b \delta A_a \partial^c h_{be} \partial^d h^{ae} \partial_d \delta A^c \\
& - \frac{\bar{G}_6}{\bar{G}_4 \bar{G}_2^{(0,1)}} \partial^b \delta A_a \partial^c h^{ae} \partial^d h_{be} \partial_d \delta A^c - \frac{2\bar{G}_6}{\bar{G}_4 \bar{G}_2^{(0,1)}} h_b^e \partial^b \delta A_a \partial^d \partial^a h_{ce} \partial_d \delta A^c \\
& - \frac{\bar{G}_6}{2\bar{G}_4 \bar{G}_2^{(0,1)}} h^e_e \partial^b \delta A_a \partial^d \partial^b h^a_c \partial_d \delta A^c + \frac{\bar{G}_6}{\bar{G}_4 \bar{G}_2^{(0,1)}} h_{ac} \partial^b \delta A^a \partial^d \partial^b h^e_e \partial_d \delta A^c \\
& + \frac{\bar{G}_6}{2\bar{G}_4 \bar{G}_2^{(0,1)}} \partial^b \delta A^a \partial^c h_{ae} \partial^d \delta A_c \partial^e h_{bd} + \frac{2\bar{G}_6}{\bar{G}_4 \bar{G}_2^{(0,1)}} \partial^c \delta A^b \partial_b \delta A_a \partial_c h^{ad} \partial_e h^d_e \\
& + \frac{\bar{G}_6}{\bar{G}_4 \bar{G}_2^{(0,1)}} \partial^a \delta A_b \partial^b \delta A_a \partial^c h_{cd} \partial^e h^d_e - \frac{\bar{G}_6}{\bar{G}_4 \bar{G}_2^{(0,1)}} \partial^b \delta A_a \partial^c \delta A^b \partial^d h^{ae} \partial^e h_{cd} \\
& + \frac{2\bar{G}_6}{\bar{G}_4 \bar{G}_2^{(0,1)}} \partial_b \delta A_a \partial_c \delta A^b \partial_d h^{ac} \partial^d h^e_e - \frac{\bar{G}_6}{\bar{G}_4 \bar{G}_2^{(0,1)}} \partial_b \delta A_a \partial_c \delta A^a \partial_d h^{bc} \partial^d h^e_e \\
& - \frac{2\bar{G}_6}{\bar{G}_4 \bar{G}_2^{(0,1)}} h_{de} \partial_b \delta A^a \partial_c \delta A^b \partial^e \partial_a h^{cd} + \frac{\bar{G}_6}{\bar{G}_4 \bar{G}_2^{(0,1)}} h_d^d \partial_b \delta A^a \partial_c \delta A^b \partial^e \partial_a h^c_e \\
& - \frac{2\bar{G}_6}{\bar{G}_4 \bar{G}_2^{(0,1)}} h^c_d \partial_b \delta A^a \partial_c \delta A^b \partial^e \partial_a h^d_e - \frac{2\bar{G}_6}{\bar{G}_4 \bar{G}_2^{(0,1)}} \delta A^a \partial^c h_{ad} \partial_c \delta A^b \partial^e \partial_b h^d_e \\
& + \frac{2\bar{G}_6}{\bar{G}_4 \bar{G}_2^{(0,1)}} \delta A^a \partial^c \delta A^b \partial^d h_{ac} \partial^e \partial_b h_{de} + \frac{2\bar{G}_6}{\bar{G}_4 \bar{G}_2^{(0,1)}} h_{de} \partial^b \delta A^c \partial_b \delta A^a \partial^e \partial^c h_{ad} \\
& - \frac{2\bar{G}_6}{\bar{G}_4 \bar{G}_2^{(0,1)}} h_{de} \partial_b \delta A^a \partial_c \delta A^b \partial^e \partial^c h_{ad} - \frac{2\bar{G}_6}{\bar{G}_4 \bar{G}_2^{(0,1)}} h_b^e \partial^b \delta A^a \partial^d \delta A^c \partial^e \partial_c h_{ad} \\
& - \frac{\bar{G}_6}{\bar{G}_4 \bar{G}_2^{(0,1)}} h_d^d \partial^b \delta A^c \partial_b \delta A^a \partial^e \partial_c h_{ae} + \frac{\bar{G}_6}{\bar{G}_4 \bar{G}_2^{(0,1)}} h_{cd} \partial^a \delta A^b \partial_b \delta A_a \partial^e \partial^d h^c_e.
\end{aligned}
\tag{A.10}$$

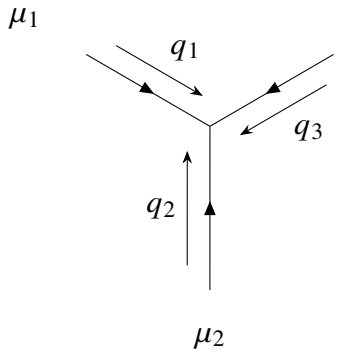
## A.2 Feynman Rules for Flat-Space Generalized Proca Theory

In this appendix, we provide the complete set of Feynman rules derived from the Lagrangian presented in Eq. (4.34), which describes the dynamics of the generalized Proca theory in flat spacetime.

The propagator:

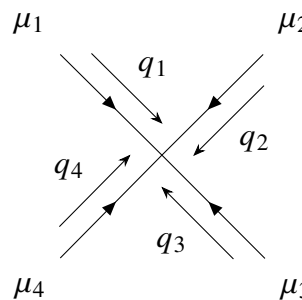
$$\mu \quad \xrightarrow{\quad p \quad} \quad \nu = \frac{1}{p^2 - m^2 + i\epsilon} \left( -\eta_{\mu\nu} + \frac{p_\mu p_\nu}{m^2} \right) \quad (\text{A.11})$$

$L_3$  contribution to the cubic vertex:



$$= \frac{2J_3 m^2 \left( q_3^{\mu_3} \eta_{\mu_1 \mu_2} + q_2^{\mu_2} \eta_{\mu_1 \mu_3} + q_1^{\mu_1} \eta_{\mu_2 \mu_3} \right)}{\Lambda_2^2} \quad (\text{A.12})$$

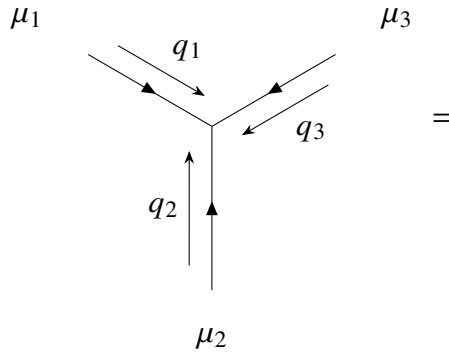
$L_4$  contribution to the quartic vertex:



$$= \quad (\text{A.13})$$

$$\begin{aligned}
= & -\frac{4im^2}{\Lambda_2^4} \left( C_4 q_2^{\mu_4} q_3^{\mu_3} \eta_{\mu_1\mu_2} - C_4 q_2^{\mu_3} q_3^{\mu_4} \eta_{\mu_1\mu_2} - 2\tilde{C}_4 q_2^{\mu_3} q_3^{\mu_4} \eta_{\mu_1\mu_2} - C_4 q_2^{\mu_4} q_4^{\mu_3} \eta_{\mu_1\mu_2} - 2\tilde{C}_4 q_2^{\mu_4} q_4^{\mu_3} \eta_{\mu_1\mu_2} \right. \\
& + C_4 q_2^{\mu_3} q_4^{\mu_4} \eta_{\mu_1\mu_2} - C_4 q_1^{\mu_2} q_2^{\mu_4} \eta_{\mu_1\mu_3} - 2\tilde{C}_4 q_1^{\mu_2} q_2^{\mu_4} \eta_{\mu_1\mu_3} - C_4 q_2^{\mu_4} q_3^{\mu_2} \eta_{\mu_1\mu_3} - 2\tilde{C}_4 q_2^{\mu_4} q_3^{\mu_2} \eta_{\mu_1\mu_3} \\
& + 2\tilde{C}_4 q_1^{\mu_2} q_3^{\mu_4} \eta_{\mu_1\mu_3} + C_4 q_2^{\mu_2} q_3^{\mu_4} \eta_{\mu_1\mu_3} - C_4 q_1^{\mu_2} q_3^{\mu_3} \eta_{\mu_1\mu_4} - C_4 q_2^{\mu_3} q_4^{\mu_2} \eta_{\mu_1\mu_4} - 2\tilde{C}_4 q_2^{\mu_3} q_4^{\mu_2} \eta_{\mu_1\mu_4} \\
& + C_4 q_3^{\mu_3} q_4^{\mu_2} \eta_{\mu_1\mu_4} + 2\tilde{C}_4 q_1^{\mu_2} q_4^{\mu_3} \eta_{\mu_1\mu_4} + C_4 q_2^{\mu_2} q_4^{\mu_3} \eta_{\mu_1\mu_4} - C_4 q_3^{\mu_2} q_4^{\mu_3} \eta_{\mu_1\mu_4} - 2\tilde{C}_4 q_3^{\mu_2} q_4^{\mu_3} \eta_{\mu_1\mu_4} \\
& + C_4 q_1^{\mu_1} q_2^{\mu_4} \eta_{\mu_2\mu_3} + 2\tilde{C}_4 q_2^{\mu_4} q_3^{\mu_1} \eta_{\mu_2\mu_3} + C_4 q_1^{\mu_1} q_3^{\mu_4} \eta_{\mu_2\mu_3} + 2\tilde{C}_4 q_2^{\mu_1} q_3^{\mu_4} \eta_{\mu_2\mu_3} - C_4 q_2^{\mu_4} q_4^{\mu_1} \eta_{\mu_2\mu_3} \\
& - 2\tilde{C}_4 q_2^{\mu_4} q_4^{\mu_1} \eta_{\mu_2\mu_3} - C_4 q_3^{\mu_4} q_4^{\mu_1} \eta_{\mu_2\mu_3} - 2\tilde{C}_4 q_3^{\mu_4} q_4^{\mu_1} \eta_{\mu_2\mu_3} + C_4 q_2^{\mu_1} q_4^{\mu_4} \eta_{\mu_2\mu_3} + C_4 q_3^{\mu_1} q_4^{\mu_4} \eta_{\mu_2\mu_3} \\
& + C_4 q_1^{\mu_1} q_2^{\mu_3} \eta_{\mu_2\mu_4} - C_4 q_2^{\mu_3} q_3^{\mu_1} \eta_{\mu_2\mu_4} - 2\tilde{C}_4 q_2^{\mu_3} q_3^{\mu_1} \eta_{\mu_2\mu_4} + C_4 q_2^{\mu_1} q_3^{\mu_3} \eta_{\mu_2\mu_4} + 2\tilde{C}_4 q_2^{\mu_1} q_3^{\mu_3} \eta_{\mu_2\mu_4} \\
& + C_4 q_3^{\mu_3} q_4^{\mu_1} \eta_{\mu_2\mu_4} + C_4 q_1^{\mu_1} q_4^{\mu_3} \eta_{\mu_2\mu_4} + 2\tilde{C}_4 q_2^{\mu_1} q_4^{\mu_3} \eta_{\mu_2\mu_4} - C_4 q_3^{\mu_1} q_4^{\mu_3} \eta_{\mu_2\mu_4} - 2\tilde{C}_4 q_3^{\mu_1} q_4^{\mu_3} \eta_{\mu_2\mu_4} \\
& + q_1^{\mu_3} \left( 2\tilde{C}_4 q_2^{\mu_4} \eta_{\mu_1\mu_2} - (C_4 + 2\tilde{C}_4) q_3^{\mu_4} \eta_{\mu_1\mu_2} + C_4 q_4^{\mu_4} \eta_{\mu_1\mu_2} + C_4 q_2^{\mu_2} \eta_{\mu_1\mu_4} - C_4 q_3^{\mu_2} \eta_{\mu_1\mu_4} - 2\tilde{C}_4 q_3^{\mu_2} \eta_{\mu_1\mu_4} \right) \\
& + 2\tilde{C}_4 q_1^{\mu_2} q_3^{\mu_1} \eta_{\mu_3\mu_4} + C_4 q_2^{\mu_2} q_3^{\mu_1} \eta_{\mu_3\mu_4} + C_4 q_1^{\mu_1} q_2^{\mu_3} \eta_{\mu_3\mu_4} - C_4 q_2^{\mu_1} q_3^{\mu_2} \eta_{\mu_3\mu_4} - 2\tilde{C}_4 q_2^{\mu_1} q_3^{\mu_2} \eta_{\mu_3\mu_4} \\
& - C_4 q_1^{\mu_2} q_4^{\mu_1} \eta_{\mu_3\mu_4} - 2\tilde{C}_4 q_1^{\mu_2} q_4^{\mu_1} \eta_{\mu_3\mu_4} + C_4 q_2^{\mu_2} q_4^{\mu_1} \eta_{\mu_3\mu_4} + 2\tilde{C}_4 q_3^{\mu_2} q_4^{\mu_1} \eta_{\mu_3\mu_4} + C_4 q_1^{\mu_1} q_4^{\mu_2} \eta_{\mu_3\mu_4} \\
& - C_4 q_2^{\mu_1} q_4^{\mu_2} \eta_{\mu_3\mu_4} - 2\tilde{C}_4 q_2^{\mu_1} q_4^{\mu_2} \eta_{\mu_3\mu_4} + 2\tilde{C}_4 q_3^{\mu_1} q_4^{\mu_2} \eta_{\mu_3\mu_4} + 2\tilde{C}_4 \eta_{\mu_1\mu_4} \eta_{\mu_2\mu_3} (q_1 \cdot q_2) + 2\tilde{C}_4 \eta_{\mu_1\mu_3} \eta_{\mu_2\mu_4} (q_1 \cdot q_2) \\
& + 2\tilde{C}_4 \eta_{\mu_1\mu_4} \eta_{\mu_2\mu_3} (q_1 \cdot q_3) + 2\tilde{C}_4 \eta_{\mu_1\mu_2} \eta_{\mu_3\mu_4} (q_1 \cdot q_3) + 2\tilde{C}_4 \eta_{\mu_1\mu_3} \eta_{\mu_2\mu_4} (q_1 \cdot q_4) + 2\tilde{C}_4 \eta_{\mu_1\mu_2} \eta_{\mu_3\mu_4} (q_1 \cdot q_4) \\
& + 2\tilde{C}_4 \eta_{\mu_1\mu_3} \eta_{\mu_2\mu_4} (q_2 \cdot q_3) + 2\tilde{C}_4 \eta_{\mu_1\mu_2} \eta_{\mu_3\mu_4} (q_2 \cdot q_3) + 2\tilde{C}_4 \eta_{\mu_1\mu_4} \eta_{\mu_2\mu_3} (q_2 \cdot q_4) + 2\tilde{C}_4 \eta_{\mu_1\mu_2} \eta_{\mu_3\mu_4} (q_2 \cdot q_4) \\
& + 2\tilde{C}_4 \eta_{\mu_1\mu_4} \eta_{\mu_2\mu_3} (q_3 \cdot q_4) + 2\tilde{C}_4 \eta_{\mu_1\mu_3} \eta_{\mu_2\mu_4} (q_3 \cdot q_4) \Big)
\end{aligned} \tag{A.14}$$

$L_5$  contribution to the cubic vertex:



$$= \tag{A.15}$$

$$\begin{aligned}
& \frac{C_5}{\Lambda_2^2} \left( 2q_1^{\mu_1} q_2^{\mu_3} q_3^{\mu_2} - 2q_3^{\mu_3} \eta_{\mu_1\mu_2} (q_1 \cdot q_2) + q_3^{\mu_2} \eta_{\mu_1\mu_3} (q_1 \cdot q_2) + q_3^{\mu_1} \eta_{\mu_2\mu_3} (q_1 \cdot q_2) + q_2^{\mu_3} \eta_{\mu_1\mu_2} (q_1 \cdot q_3) \right. \\
& - 2q_2^{\mu_2} \eta_{\mu_1\mu_3} (q_1 \cdot q_3) + q_2^{\mu_1} \eta_{\mu_2\mu_3} (q_1 \cdot q_3) - 2q_1^{\mu_1} \eta_{\mu_2\mu_3} (q_2 \cdot q_3) + q_1^{\mu_3} \left( 2q_2^{\mu_2} q_3^{\mu_1} - 3q_2^{\mu_1} q_3^{\mu_2} + \eta_{\mu_1\mu_2} (q_2 \cdot q_3) \right) \\
& \left. + q_1^{\mu_2} \left( -3q_2^{\mu_3} q_3^{\mu_1} + 2q_2^{\mu_1} q_3^{\mu_3} + \eta_{\mu_1\mu_3} (q_2 \cdot q_3) \right) \right)
\end{aligned} \tag{A.16}$$

$L_6$  contribution to the quartic vertex:



$$(A.17)$$

$$\begin{aligned}
& \frac{2iC_6}{\Lambda_2^4} \left( 3q_1^{\mu_2} q_2^{\mu_4} q_3^{\mu_3} q_4^{\mu_1} - 4q_1^{\mu_2} q_2^{\mu_3} q_3^{\mu_4} q_4^{\mu_1} - 2q_1^{\mu_1} q_2^{\mu_4} q_3^{\mu_3} q_4^{\mu_2} + 3q_1^{\mu_1} q_2^{\mu_3} q_3^{\mu_4} q_4^{\mu_2} - 4q_1^{\mu_2} q_2^{\mu_4} q_3^{\mu_1} q_4^{\mu_3} \right. \\
& + 3q_1^{\mu_1} q_2^{\mu_4} q_3^{\mu_2} q_4^{\mu_3} + 4q_1^{\mu_2} q_2^{\mu_1} q_3^{\mu_4} q_4^{\mu_3} - 2q_1^{\mu_1} q_2^{\mu_2} q_3^{\mu_4} q_4^{\mu_3} + 3q_1^{\mu_2} q_2^{\mu_3} q_3^{\mu_1} q_4^{\mu_4} - 2q_1^{\mu_1} q_2^{\mu_2} q_3^{\mu_3} q_4^{\mu_4} \\
& - 2q_1^{\mu_2} q_2^{\mu_1} q_3^{\mu_3} q_4^{\mu_4} - 2q_3^{\mu_4} q_4^{\mu_3} \eta_{\mu_1\mu_2}(q_1 \cdot q_2) + 2q_3^{\mu_3} q_4^{\mu_4} \eta_{\mu_1\mu_2}(q_1 \cdot q_2) + q_3^{\mu_4} q_4^{\mu_4} \eta_{\mu_1\mu_3}(q_1 \cdot q_2) \\
& - q_3^{\mu_2} q_4^{\mu_4} \eta_{\mu_1\mu_3}(q_1 \cdot q_2) - q_3^{\mu_3} q_4^{\mu_2} \eta_{\mu_1\mu_4}(q_1 \cdot q_2) + q_3^{\mu_2} q_4^{\mu_3} \eta_{\mu_1\mu_4}(q_1 \cdot q_2) + q_2^{\mu_3} q_4^{\mu_4} \eta_{\mu_1\mu_2}(q_1 \cdot q_3) \\
& - 2q_2^{\mu_4} q_4^{\mu_2} \eta_{\mu_1\mu_3}(q_1 \cdot q_3) + 2q_2^{\mu_2} q_4^{\mu_4} \eta_{\mu_1\mu_3}(q_1 \cdot q_3) + q_2^{\mu_3} q_4^{\mu_2} \eta_{\mu_1\mu_4}(q_1 \cdot q_3) - q_2^{\mu_2} q_4^{\mu_3} \eta_{\mu_1\mu_4}(q_1 \cdot q_3) \\
& + q_2^{\mu_4} q_4^{\mu_1} \eta_{\mu_2\mu_3}(q_1 \cdot q_3) - q_2^{\mu_1} q_4^{\mu_4} \eta_{\mu_2\mu_3}(q_1 \cdot q_3) - q_2^{\mu_2} q_4^{\mu_1} \eta_{\mu_3\mu_4}(q_1 \cdot q_3) + q_2^{\mu_1} q_4^{\mu_2} \eta_{\mu_3\mu_4}(q_1 \cdot q_3) \\
& - q_2^{\mu_4} q_3^{\mu_3} \eta_{\mu_1\mu_2}(q_1 \cdot q_4) + q_2^{\mu_3} q_3^{\mu_4} \eta_{\mu_1\mu_2}(q_1 \cdot q_4) + q_2^{\mu_2} q_3^{\mu_2} \eta_{\mu_1\mu_3}(q_1 \cdot q_4) - q_2^{\mu_2} q_3^{\mu_4} \eta_{\mu_1\mu_3}(q_1 \cdot q_4) \\
& - 2q_2^{\mu_3} q_3^{\mu_2} \eta_{\mu_1\mu_4}(q_1 \cdot q_4) + 2q_2^{\mu_2} q_3^{\mu_3} \eta_{\mu_1\mu_4}(q_1 \cdot q_4) + 2q_1^{\mu_1} q_4^{\mu_4} \eta_{\mu_2\mu_3}(q_2 \cdot q_3) - q_1^{\mu_1} q_4^{\mu_3} \eta_{\mu_2\mu_4}(q_2 \cdot q_3) \\
& + q_1^{\mu_2} q_4^{\mu_1} \eta_{\mu_3\mu_4}(q_2 \cdot q_3) - q_1^{\mu_1} q_4^{\mu_2} \eta_{\mu_3\mu_4}(q_2 \cdot q_3) - q_1^{\mu_2} q_3^{\mu_3} \eta_{\mu_1\mu_4}(q_2 \cdot q_4) - q_1^{\mu_1} q_3^{\mu_4} \eta_{\mu_2\mu_3}(q_2 \cdot q_4) \\
& + 2q_1^{\mu_1} q_3^{\mu_3} \eta_{\mu_2\mu_4}(q_2 \cdot q_4) + q_1^{\mu_2} q_3^{\mu_1} \eta_{\mu_3\mu_4}(q_2 \cdot q_4) - q_1^{\mu_1} q_3^{\mu_2} \eta_{\mu_3\mu_4}(q_2 \cdot q_4) + q_1^{\mu_2} q_2^{\mu_4} \eta_{\mu_1\mu_3}(q_3 \cdot q_4) \\
& + q_1^{\mu_1} q_2^{\mu_3} \eta_{\mu_1\mu_4}(q_3 \cdot q_4) - q_1^{\mu_1} q_2^{\mu_4} \eta_{\mu_2\mu_3}(q_3 \cdot q_4) - q_1^{\mu_1} q_2^{\mu_3} \eta_{\mu_2\mu_4}(q_3 \cdot q_4) - 2q_1^{\mu_2} q_2^{\mu_1} \eta_{\mu_3\mu_4}(q_3 \cdot q_4) \\
& + 2q_1^{\mu_1} q_2^{\mu_2} \eta_{\mu_3\mu_4}(q_3 \cdot q_4) + q_1^{\mu_4} \left( 3q_2^{\mu_1} q_3^{\mu_3} q_4^{\mu_2} + 4q_2^{\mu_2} \left( q_3^{\mu_2} q_4^{\mu_1} - q_3^{\mu_1} q_4^{\mu_2} \right) - 4q_2^{\mu_1} q_3^{\mu_2} q_4^{\mu_3} + q_4^{\mu_3} \eta_{\mu_1\mu_2}(q_2 \cdot q_3) \right. \\
& + q_4^{\mu_2} \eta_{\mu_1\mu_3}(q_2 \cdot q_3) - 2q_4^{\mu_1} \eta_{\mu_2\mu_3}(q_2 \cdot q_3) - q_3^{\mu_3} \eta_{\mu_1\mu_2}(q_2 \cdot q_4) + q_3^{\mu_1} \eta_{\mu_2\mu_3}(q_2 \cdot q_4) + q_2^{\mu_1} \eta_{\mu_2\mu_3}(q_3 \cdot q_4) \\
& \left. - q_2^{\mu_2} \left( 2q_3^{\mu_3} q_4^{\mu_1} - 3q_3^{\mu_1} q_4^{\mu_3} + \eta_{\mu_1\mu_3}(q_3 \cdot q_4) \right) \right) \\
& + q_1^{\mu_3} \left( -4q_2^{\mu_1} q_3^{\mu_4} q_4^{\mu_2} + q_2^{\mu_4} \left( -4q_3^{\mu_2} q_4^{\mu_1} + 4q_3^{\mu_1} q_4^{\mu_2} \right) + 3q_2^{\mu_1} q_3^{\mu_2} q_4^{\mu_4} - q_4^{\mu_4} \eta_{\mu_1\mu_2}(q_2 \cdot q_3) \right. \\
& + q_4^{\mu_1} \eta_{\mu_2\mu_4}(q_2 \cdot q_3) + q_3^{\mu_4} \eta_{\mu_1\mu_2}(q_2 \cdot q_4) + q_3^{\mu_2} \eta_{\mu_1\mu_4}(q_2 \cdot q_4) - 2q_3^{\mu_1} \eta_{\mu_2\mu_4}(q_2 \cdot q_4) + q_2^{\mu_1} \eta_{\mu_2\mu_4}(q_3 \cdot q_4) \\
& \left. + q_2^{\mu_2} \left( 3q_3^{\mu_4} q_4^{\mu_1} - 2q_3^{\mu_1} q_4^{\mu_4} - \eta_{\mu_1\mu_4}(q_3 \cdot q_4) \right) \right) \Big)
\end{aligned}
\tag{A.18}$$

### A.3 Feynman Rules from Curved-Space Lagrangian (Expanded on Flat Background)

This section presents the Feynman rules derived from the generalized Proca theory in curved spacetime, as defined in Equation 5.18.

The propagators:

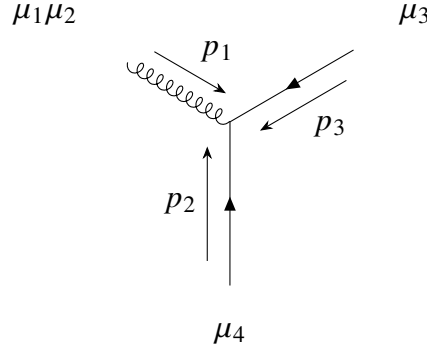
$$\mu \xrightarrow{p} \nu = D_{\mu\nu}(p) = \frac{1}{p^2 - m^2 + i\epsilon} \left( -\eta_{\mu\nu} + \frac{p_\mu p_\nu}{m^2} \right) \quad (\text{A.19})$$

for the vector field and

$$\mu\nu \xrightarrow{p} \rho\sigma = D_{\mu\nu\rho\sigma}(p) = \frac{1}{2(p^2 + i\epsilon)} (\eta_{\mu\rho}\eta_{\nu\sigma} + \eta_{\mu\sigma}\eta_{\nu\rho} - \eta_{\mu\nu}\eta_{\rho\sigma}) \quad (\text{A.20})$$

For the graviton.

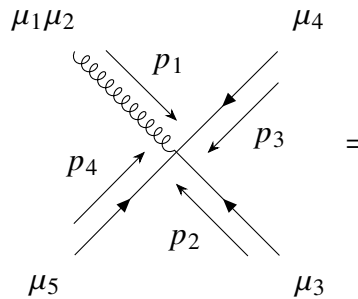
The mixed cubic vertex:



$$= \quad (\text{A.21})$$

$$\begin{aligned} & -\frac{2i\sqrt{2}\pi}{M_{\text{pl}}} \left[ m^2(-\eta_{\mu_1\mu_3})\eta_{\mu_2\mu_4} + m^2\eta_{\mu_1\mu_2}\eta_{\mu_3\mu_4} \right. \\ & \quad + (m^2 + p_1 \cdot p_2 + 2p_1 \cdot p_3) \eta_{\mu_1\mu_4}\eta_{\mu_2\mu_3} + (p_1 \cdot p_2) \eta_{\mu_1\mu_3}\eta_{\mu_2\mu_4} \\ & \quad + (p_1 \cdot p_2) \eta_{\mu_1\mu_2}\eta_{\mu_3\mu_4} + 2(p_2 \cdot p_3) \eta_{\mu_1\mu_3}\eta_{\mu_2\mu_4} \\ & \quad - \eta_{\mu_1\mu_2}p_1^{\mu_3}p_2^{\mu_4} - 2\eta_{\mu_1\mu_3}p_3^{\mu_2}p_2^{\mu_4} \\ & \quad \left. - p_1^{\mu_4} \left( \eta_{\mu_1\mu_2}p_2^{\mu_3} + 2\eta_{\mu_2\mu_3}p_3^{\mu_1} \right) - \eta_{\mu_3\mu_4}p_1^{\mu_2}p_2^{\mu_1} \right] \quad (\text{A.22}) \end{aligned}$$

The mixed quartic vertices:



$$= \quad (\text{A.23})$$

$$\begin{aligned}
\frac{2\sqrt{2}\pi J_3 m^2}{\Lambda^2 M_{\text{pl}}} \Big[ & \eta_{\mu_1\mu_5}\eta_{\mu_2\mu_4} p_3^{\mu_3} + \eta_{\mu_1\mu_4}\eta_{\mu_2\mu_5} p_3^{\mu_3} + \eta_{\mu_1\mu_5}\eta_{\mu_3\mu_4} p_2^{\mu_2} \\
& + \eta_{\mu_2\mu_5}\eta_{\mu_3\mu_4} p_1^{\mu_1} + \eta_{\mu_1\mu_4}\eta_{\mu_3\mu_5} p_2^{\mu_2} + \eta_{\mu_2\mu_4}\eta_{\mu_3\mu_5} p_1^{\mu_1} \\
& + \eta_{\mu_1\mu_2}\eta_{\mu_4\mu_5} p_3^{\mu_3} + \eta_{\mu_1\mu_2}\eta_{\mu_4\mu_5} p_4^{\mu_3} + \eta_{\mu_1\mu_3}\eta_{\mu_4\mu_5} p_2^{\mu_2} \\
& + \eta_{\mu_1\mu_3}\eta_{\mu_4\mu_5} p_4^{\mu_2} + \eta_{\mu_2\mu_3}\eta_{\mu_4\mu_5} p_1^{\mu_1} + \eta_{\mu_2\mu_3}\eta_{\mu_4\mu_5} p_4^{\mu_1} \Big] \quad (\text{A.24})
\end{aligned}$$

$$\text{Diagram (A.25)} = \quad (\text{A.25})$$

$$\begin{aligned}
& \frac{8i\pi m^2}{M_{\text{pl}}^2} \eta_{\mu_1\mu_6}\eta_{\mu_2\mu_5}\eta_{\mu_3\mu_4} + \frac{8i\pi m^2}{M_{\text{pl}}^2} \eta_{\mu_1\mu_5}\eta_{\mu_2\mu_6}\eta_{\mu_3\mu_4} - \frac{16i\pi m^2}{M_{\text{pl}}^2} \eta_{\mu_1\mu_2}\eta_{\mu_3\mu_5}\eta_{\mu_4\mu_6} \\
& + \frac{8i\pi m^2}{M_{\text{pl}}^2} \eta_{\mu_1\mu_4}\eta_{\mu_2\mu_3}\eta_{\mu_5\mu_6} + \frac{8i\pi m^2}{M_{\text{pl}}^2} \eta_{\mu_1\mu_3}\eta_{\mu_2\mu_4}\eta_{\mu_5\mu_6} + \frac{8i\pi m^2}{M_{\text{pl}}^2} \eta_{\mu_1\mu_2}\eta_{\mu_3\mu_4}\eta_{\mu_5\mu_6} \\
& - \frac{16i\pi}{M_{\text{pl}}^2} p_1^{\mu_6} p_2^{\mu_5} \eta_{\mu_1\mu_4}\eta_{\mu_2\mu_3} - \frac{32i\pi}{M_{\text{pl}}^2} p_1^{\mu_5} p_3^{\mu_6} \eta_{\mu_1\mu_4}\eta_{\mu_2\mu_3} - \frac{32i\pi}{M_{\text{pl}}^2} p_3^{\mu_6} p_4^{\mu_4} \eta_{\mu_1\mu_5}\eta_{\mu_2\mu_3} \\
& - \frac{16i\pi}{M_{\text{pl}}^2} p_1^{\mu_5} p_2^{\mu_6} \eta_{\mu_1\mu_3}\eta_{\mu_2\mu_4} - \frac{32i\pi}{M_{\text{pl}}^2} p_2^{\mu_5} p_3^{\mu_6} \eta_{\mu_1\mu_3}\eta_{\mu_2\mu_4} - \frac{32i\pi}{M_{\text{pl}}^2} p_3^{\mu_6} p_4^{\mu_4} \eta_{\mu_1\mu_3}\eta_{\mu_2\mu_5} \\
& - \frac{16i\pi}{M_{\text{pl}}^2} p_1^{\mu_4} p_2^{\mu_3} \eta_{\mu_1\mu_6}\eta_{\mu_2\mu_5} - \frac{32i\pi}{M_{\text{pl}}^2} p_1^{\mu_3} p_4^{\mu_4} \eta_{\mu_1\mu_6}\eta_{\mu_2\mu_5} - \frac{16i\pi}{M_{\text{pl}}^2} p_1^{\mu_3} p_2^{\mu_4} \eta_{\mu_1\mu_5}\eta_{\mu_2\mu_6} \\
& - \frac{32i\pi}{M_{\text{pl}}^2} p_2^{\mu_3} p_4^{\mu_4} \eta_{\mu_1\mu_5}\eta_{\mu_2\mu_6} - \frac{8i\pi}{M_{\text{pl}}^2} p_1^{\mu_6} p_2^{\mu_5} \eta_{\mu_1\mu_2}\eta_{\mu_3\mu_4} - \frac{8i\pi}{M_{\text{pl}}^2} p_1^{\mu_5} p_2^{\mu_6} \eta_{\mu_1\mu_2}\eta_{\mu_3\mu_4} \\
& - \frac{16i\pi}{M_{\text{pl}}^2} p_2^{\mu_6} p_4^{\mu_2} \eta_{\mu_1\mu_5}\eta_{\mu_3\mu_4} - \frac{16i\pi}{M_{\text{pl}}^2} p_1^{\mu_6} p_4^{\mu_1} \eta_{\mu_2\mu_5}\eta_{\mu_3\mu_4} + \frac{32i\pi}{M_{\text{pl}}^2} p_1^{\mu_5} p_2^{\mu_3} \eta_{\mu_1\mu_2}\eta_{\mu_4\mu_6} \\
& + \frac{32i\pi}{M_{\text{pl}}^2} p_1^{\mu_3} p_2^{\mu_5} \eta_{\mu_1\mu_2}\eta_{\mu_4\mu_6} + \frac{32i\pi}{M_{\text{pl}}^2} p_2^{\mu_5} p_3^{\mu_2} \eta_{\mu_1\mu_3}\eta_{\mu_4\mu_6} + \frac{32i\pi}{M_{\text{pl}}^2} p_2^{\mu_3} p_4^{\mu_2} \eta_{\mu_1\mu_5}\eta_{\mu_4\mu_6} \\
& + \frac{32i\pi}{M_{\text{pl}}^2} p_1^{\mu_5} p_3^{\mu_1} \eta_{\mu_2\mu_3}\eta_{\mu_4\mu_6} + \frac{32i\pi}{M_{\text{pl}}^2} p_1^{\mu_3} p_4^{\mu_1} \eta_{\mu_2\mu_5}\eta_{\mu_4\mu_6} + \frac{16i\pi}{M_{\text{pl}}^2} p_1^{\mu_2} p_2^{\mu_1} \eta_{\mu_3\mu_5}\eta_{\mu_4\mu_6} \quad (\text{A.26})
\end{aligned}$$

We do not explicitly include the pure sectors for this setup since the rules for the pure Proca sector can be directly obtained from the flat-space Feynman rules. As for the pure

graviton sector, the corresponding Feynman rules are well established in the literature and can be found, for instance, in [[39](#)].

## Appendix B

### 3.5PN Phasing Coefficients in the Stationary Phase Approximation

In this appendix, we summarize the post-Newtonian (PN) coefficients that contribute to the gravitational wave (GW) phase in the stationary-phase approximation (SPA) up to 3.5PN order, including spin effects in General Relativity [96]. These coefficients have been introduced in earlier chapters in the context of waveform modeling.

We begin by defining the following parameters:

$$M = m_1 + m_2, \tag{B.1}$$

$$\eta = \frac{m_1 m_2}{M^2}, \tag{B.2}$$

$$\delta = \frac{m_1 - m_2}{M}, \tag{B.3}$$

$$\chi_S = \frac{\chi_1 + \chi_2}{2}, \tag{B.4}$$

$$\chi_A = \frac{\chi_1 - \chi_2}{2}, \tag{B.5}$$

where  $\eta$  is the symmetric mass ratio, and  $\gamma_E$  denotes the Euler–Mascheroni constant.

The PN coefficients entering the GW phasing are given by:

$$\psi_0 = 1, \tag{B.6}$$

$$\psi_1 = 0, \tag{B.7}$$

$$\psi_2 = \frac{3715}{756} + \frac{55\eta}{9}, \tag{B.8}$$

$$\psi_3 = -16\pi + \frac{113\delta\chi_A}{3} + \left(\frac{113}{3} - \frac{76\eta}{3}\right)\chi_S, \tag{B.9}$$

$$\psi_4 = \frac{15293365}{508032} + \frac{27145\eta}{504} + \frac{3085\eta^2}{72} + \left(-\frac{405}{8} + 200\eta\right)\chi_A^2 - \frac{405\delta\chi_A\chi_S}{4} + \left(-\frac{405}{8} + \frac{5\eta}{2}\right)\chi_S^2, \quad (\text{B.10})$$

$$\psi_5 = \frac{38645\pi}{756} - \frac{65\pi\eta}{9} + \left(-\frac{732985}{2268} - \frac{140\eta}{9}\right)\delta\chi_A + \left(-\frac{732985}{2268} + \frac{24260\eta}{81} + \frac{340\eta^2}{9}\right)\chi_S, \quad (\text{B.11})$$

$$\psi_{5\log} = 3\psi_5 = \frac{38645\pi}{252} - \frac{65\pi\eta}{3} + \left(-\frac{732985}{756} - \frac{140\eta}{3}\right)\delta\chi_A + \left(-\frac{732985}{756} + \frac{24260\eta}{27} + \frac{340\eta^2}{3}\right)\chi_S, \quad (\text{B.12})$$

$$\begin{aligned} \psi_6 = & \frac{11583231236531}{4694215680} - \frac{6848\log(4)}{21} - \frac{640\pi^2}{3} + \frac{6848\gamma_E}{21} \\ & + \left(-\frac{15737765635}{3048192} + \frac{2255\pi^2}{12}\right)\eta + \frac{76055\eta^2}{1728} - \frac{127825\eta^3}{1296} \\ & + \frac{2270\pi\delta\chi_A}{3} + \left(\frac{2270\pi}{3} - 520\pi\eta\right)\chi_S + \left(\frac{75515}{288} - \frac{547945\eta}{504} - \frac{8455\eta^2}{24}\right)\chi_A^2 \\ & + \left(\frac{75515}{144} - \frac{8225\eta}{18}\right)\delta\chi_A\chi_S + \left(\frac{75515}{288} - \frac{126935\eta}{252} + \frac{19235\eta^2}{72}\right)\chi_S^2, \end{aligned} \quad (\text{B.13})$$

$$\psi_{6\log} = -\frac{6848}{21}, \quad (\text{B.14})$$

$$\begin{aligned} \psi_7 = & \frac{77096675\pi}{254016} + \frac{378515\pi\eta}{1512} - \frac{74045\pi\eta^2}{756} + \left(-\frac{25150083775}{3048192} + \frac{26804935\eta}{6048} - \frac{1985\eta^2}{48}\right)\delta\chi_A \\ & + \left(-\frac{25150083775}{3048192} + \frac{10566655595\eta}{762048} - \frac{1042165\eta^2}{3024} + \frac{5345\eta^3}{36}\right)\chi_S. \end{aligned} \quad (\text{B.15})$$

# Bibliography

- [1] Steven Weinberg. “The cosmological constant problem”. In: *Reviews of Modern Physics* 61.1 (1989), pp. 1–23. doi: [10.1103/RevModPhys.61.1](https://doi.org/10.1103/RevModPhys.61.1).
- [2] Michael Edward Peskin and Daniel V. Schroeder. *An introduction to quantum field theory*. eng. The advanced book program. Boca Raton London New York: CRC Press, Taylor & Francis Group, 2019. ISBN: 978-0-367-32056-0 978-0-201-50397-5.
- [3] John David Jackson. *Classical electrodynamics*. 3rd ed. New York: Wiley, 1999. ISBN: 978-0-471-30932-1.
- [4] Jun John Sakurai and Jim Napolitano. *Modern quantum mechanics*. eng. 3rd ed. Cambridge: Cambridge University Press, 2021. ISBN: 978-1-108-47322-4.
- [5] Mark Allen Srednicki. *Quantum field theory*. OCLC: ocm71808151. Cambridge ; New York: Cambridge University Press, 2007. ISBN: 978-0-521-86449-7.
- [6] Matthew Dean Schwartz. *Quantum field theory and the standard model*. eng. New York: Cambridge university press, 2014. ISBN: 978-1-107-03473-0.
- [7] A. Zee. *Quantum field theory in a nutshell*. 2nd ed. In a nutshell. OCLC: ocn318585662. Princeton, N.J: Princeton University Press, 2010. ISBN: 978-0-691-14034-6.
- [8] Lavinia Heisenberg. “A systematic approach to generalisations of General Relativity and their cosmological implications”. en. In: *Physics Reports* 796 (Mar. 2019). arXiv:1807.01725 [astro-ph, physics:gr-qc, physics:hep-th], pp. 1–113. ISSN: 03701573. doi: [10.1016/j.physrep.2018.11.006](https://doi.org/10.1016/j.physrep.2018.11.006). URL: <http://arxiv.org/abs/1807.01725>.
- [9] E. Wigner. “On Unitary Representations of the Inhomogeneous Lorentz Group”. In: *The Annals of Mathematics* 40.1 (Jan. 1939), p. 149. ISSN: 0003486X. doi: [10.2307/1968551](https://doi.org/10.2307/1968551). URL: <https://www.jstor.org/stable/1968551?origin=crossref>.

- [10] Steven Weinberg. “The quantum theory of fields. 1: Foundations”. eng. In: Paperback ed., 10. print. Num Pages: 609. Cambridge: Cambridge Univ. Press, 2014. ISBN: 978-0-521-67053-1.
- [11] E. C. G. Stueckelberg. “Die Wechselwirkungskräfte in der Elektrodynamik und in der Feldtheorie der Kernkräfte. Teil I”. In: (Apr. 1938). Medium: text/html,application/pdf,text/html Publisher: Birkhäuser. DOI: [10.5169/SEALS-110852](https://doi.org/10.5169/SEALS-110852). URL: <https://www.e-periodica.ch/digbib/view?pid=hpa-001:1938:11::636>.
- [12] Henri Ruegg and Marti Ruiz-Altaba. “The Stueckelberg Field”. In: *International Journal of Modern Physics A* 19.20 (Aug. 2004). arXiv:hep-th/0304245, pp. 3265–3347. ISSN: 0217-751X, 1793-656X. DOI: [10.1142/S0217751X04019755](https://doi.org/10.1142/S0217751X04019755). URL: <http://arxiv.org/abs/hep-th/0304245>.
- [13] S. Coleman, J. Wess, and Bruno Zumino. “Structure of Phenomenological Lagrangians. I”. en. In: *Physical Review* 177.5 (Jan. 1969), pp. 2239–2247. ISSN: 0031-899X. DOI: [10.1103/PhysRev.177.2239](https://doi.org/10.1103/PhysRev.177.2239). URL: <https://link.aps.org/doi/10.1103/PhysRev.177.2239>.
- [14] Curtis G. Callan et al. “Structure of Phenomenological Lagrangians. II”. en. In: *Physical Review* 177.5 (Jan. 1969), pp. 2247–2250. ISSN: 0031-899X. DOI: [10.1103/PhysRev.177.2247](https://doi.org/10.1103/PhysRev.177.2247). URL: <https://link.aps.org/doi/10.1103/PhysRev.177.2247>.
- [15] Nima Arkani-Hamed, Howard Georgi, and Matthew D. Schwartz. “Effective Field Theory for Massive Gravitons and Gravity in Theory Space”. In: *Annals of Physics* 305.2 (June 2003). arXiv:hep-th/0210184, pp. 96–118. ISSN: 00034916. DOI: [10.1016/S0003-4916\(03\)00068-X](https://doi.org/10.1016/S0003-4916(03)00068-X). URL: <http://arxiv.org/abs/hep-th/0210184>.
- [16] Tsutomu Kobayashi. “Horndeski theory and beyond: a review”. In: *Reports on Progress in Physics* 82.8 (Aug. 2019). arXiv:1901.07183 [gr-qc], p. 086901. ISSN: 0034-4885, 1361-6633. DOI: [10.1088/1361-6633/ab2429](https://doi.org/10.1088/1361-6633/ab2429). URL: <http://arxiv.org/abs/1901.07183>.
- [17] Claudia de Rham and Andrew Matas. “Ostrogradsky in Theories with Multiple Fields”. In: *Journal of Cosmology and Astroparticle Physics* 2016.06 (June 2016). arXiv:1604.08638 [hep-th], pp. 041–041. ISSN: 1475-7516. DOI: [10.1088/1475-7516/2016/06/041](https://doi.org/10.1088/1475-7516/2016/06/041). URL: <http://arxiv.org/abs/1604.08638>.



- [18] R. P. Woodard. *The Theorem of Ostrogradsky*. arXiv:1506.02210 [hep-th]. Aug. 2015. DOI: [10.48550/arXiv.1506.02210](https://doi.org/10.48550/arXiv.1506.02210). URL: <http://arxiv.org/abs/1506.02210> (visited on 03/22/2025).
- [19] Alexander Ganz and Karim Noui. *Reconsidering the Ostrogradsky theorem: Higher-derivatives Lagrangians, Ghosts and Degeneracy*. arXiv:2007.01063 [hep-th]. July 2022. DOI: [10.48550/arXiv.2007.01063](https://doi.org/10.48550/arXiv.2007.01063). URL: <http://arxiv.org/abs/2007.01063> (visited on 03/22/2025).
- [20] Claudia de Rham. “Galileons in the Sky”. In: *Comptes Rendus. Physique* 13.6-7 (July 2012). arXiv:1204.5492 [astro-ph], pp. 666–681. ISSN: 1878-1535. DOI: [10.1016/j.crhy.2012.04.006](https://doi.org/10.1016/j.crhy.2012.04.006). URL: <http://arxiv.org/abs/1204.5492>.
- [21] Remko Klein, Mehmet Ozkan, and Diederik Roest. “Galileons as the Scalar Analogue of General Relativity”. In: *Physical Review D* 93.4 (Feb. 2016). arXiv:1510.08864 [hep-th], p. 044053. ISSN: 2470-0010, 2470-0029. DOI: [10.1103/PhysRevD.93.044053](https://doi.org/10.1103/PhysRevD.93.044053). URL: <http://arxiv.org/abs/1510.08864>.
- [22] *Galileons as the scalar analogue of general relativity* | *Phys. Rev. D*. URL: <https://journals.aps.org/prd/abstract/10.1103/PhysRevD.93.044053>.
- [23] Arthur Lue. “The Phenomenology of Dvali-Gabadadze-Porrati Cosmologies”. In: *Physics Reports* 423.1 (Jan. 2006). arXiv:astro-ph/0510068, pp. 1–48. ISSN: 03701573. DOI: [10.1016/j.physrep.2005.10.007](https://doi.org/10.1016/j.physrep.2005.10.007). URL: <http://arxiv.org/abs/astro-ph/0510068>.
- [24] Lavinia Heisenberg. *Generalised Proca Theories*. en. arXiv:1705.05387 [hep-th]. May 2017. DOI: [10.48550/arXiv.1705.05387](https://doi.org/10.48550/arXiv.1705.05387). URL: <http://arxiv.org/abs/1705.05387>.
- [25] Antonio De Felice et al. “Effective gravitational couplings for cosmological perturbations in generalized Proca theories”. en. In: *Physical Review D* 94.4 (Aug. 2016), p. 044024. ISSN: 2470-0010, 2470-0029. DOI: [10.1103/PhysRevD.94.044024](https://doi.org/10.1103/PhysRevD.94.044024). URL: <https://link.aps.org/doi/10.1103/PhysRevD.94.044024> (visited on 04/12/2025).
- [26] Lavinia Heisenberg. “Generalization of the proca action”. In: *Journal of Cosmology and Astroparticle Physics* 2014.05 (2014), pp. 015–015. DOI: [10.1088/1475-7516/2014/05/015](https://doi.org/10.1088/1475-7516/2014/05/015).

- [27] Alberto Nicolis, Riccardo Rattazzi, and Enrico Trincherini. “Galileon as a local modification of gravity”. In: *Physical Review D* 79.6 (2009), p. 064036. DOI: [10.1103/PhysRevD.79.064036](https://doi.org/10.1103/PhysRevD.79.064036).
- [28] Lavinia Heisenberg and Jann Zosso. “Quantum Stability of Generalized Proca Theories”. en. In: *Classical and Quantum Gravity* 38.6 (Mar. 2021). arXiv:2005.01639 [gr-qc, physics:hep-th], p. 065001. ISSN: 0264-9381, 1361-6382. DOI: [10.1088/1361-6382/abd680](https://doi.org/10.1088/1361-6382/abd680). URL: <http://arxiv.org/abs/2005.01639>.
- [29] Vladyslav Shtabovenko, Rolf Mertig, and Frederik Orellana. “FeynCalc 9.3: New features and improvements”. In: *Computer Physics Communications* 256 (Nov. 2020). arXiv:2001.04407 [hep-ph], p. 107478. ISSN: 00104655. DOI: [10.1016/j.cpc.2020.107478](https://doi.org/10.1016/j.cpc.2020.107478). URL: <http://arxiv.org/abs/2001.04407>.
- [30] Adam Alloul et al. “FeynRules 2.0 - A complete toolbox for tree-level phenomenology”. In: *Computer Physics Communications* 185.8 (Aug. 2014). arXiv:1310.1921 [hep-ph], pp. 2250–2300. ISSN: 00104655. DOI: [10.1016/j.cpc.2014.04.012](https://doi.org/10.1016/j.cpc.2014.04.012). URL: <http://arxiv.org/abs/1310.1921>.
- [31] T. Hahn. “Generating Feynman Diagrams and Amplitudes with FeynArts 3”. In: *Computer Physics Communications* 140.3 (Nov. 2001). arXiv:hep-ph/0012260, pp. 418–431. ISSN: 00104655. DOI: [10.1016/S0010-4655\(01\)00290-9](https://doi.org/10.1016/S0010-4655(01)00290-9). URL: <http://arxiv.org/abs/hep-ph/0012260>.
- [32] David Brizuela, Jose M. Martin-Garcia, and Guillermo A. Mena Marugan. “xPert: Computer algebra for metric perturbation theory”. In: *General Relativity and Gravitation* 41.10 (Oct. 2009). arXiv:0807.0824 [gr-qc], pp. 2415–2431. ISSN: 0001-7701, 1572-9532. DOI: [10.1007/s10714-009-0773-2](https://doi.org/10.1007/s10714-009-0773-2). URL: <http://arxiv.org/abs/0807.0824>.
- [33] Joshua Ellis. “TikZ-Feynman: Feynman diagrams with TikZ”. In: *Computer Physics Communications* 210 (Jan. 2017). arXiv:1601.05437 [hep-ph], pp. 103–123. ISSN: 00104655. DOI: [10.1016/j.cpc.2016.08.019](https://doi.org/10.1016/j.cpc.2016.08.019). URL: <http://arxiv.org/abs/1601.05437>.
- [34] Eugeny Babichev and Cedric Deffayet. “An introduction to the Vainshtein mechanism”. In: *Classical and Quantum Gravity* 30.18 (Sept. 2013). arXiv:1304.7240 [gr-qc], p. 184001. ISSN: 0264-9381, 1361-6382. DOI: [10.1088/0264-9381/30/18/184001](https://doi.org/10.1088/0264-9381/30/18/184001). URL: <http://arxiv.org/abs/1304.7240>.

- [35] Charles W. Misner et al. *Gravitation*. OCLC: on1006427790. Princeton, N.J: Princeton University Press, 2017. ISBN: 978-0-691-17779-3.
- [36] Steven Weinberg. *Gravitation and cosmology: principles and applications of the general theory of relativity*. New York: Wiley, 1972. ISBN: 978-0-471-92567-5.
- [37] Sean Carroll. *Spacetime and geometry: an introduction to general relativity*. eng. Cambridge: Cambridge university press, 2019. ISBN: 978-1-108-77038-5.
- [38] Mike Guidry. *Modern general relativity: black holes, gravitational waves, and cosmology*. eng. Cambridge New York: Cambridge university press, 2019. ISBN: 978-1-107-19789-3.
- [39] John F. Donoghue, Mikhail M. Ivanov, and Andrey Shkerin. *EPFL Lectures on General Relativity as a Quantum Field Theory*. arXiv:1702.00319 [hep-th]. Feb. 2017. DOI: [10.48550/arXiv.1702.00319](https://doi.org/10.48550/arXiv.1702.00319). URL: <http://arxiv.org/abs/1702.00319>.
- [40] Lavinia Heisenberg, Shayan Hemmatyar, Nadine Nussbaumer. “Quantum Stability of Generalized Proca Theories in Curved Space-time”. (In preparation).
- [41] Talal Ahmed Chowdhury, Rakibur Rahman, and Zulfiqar Ali Sabuj. “Gravitational Properties of the Proca Field”. In: *Nuclear Physics B* 936 (Nov. 2018). arXiv:1807.10284 [hep-th], pp. 364–382. ISSN: 05503213. DOI: [10.1016/j.nuclphysb.2018.09.009](https://doi.org/10.1016/j.nuclphysb.2018.09.009). URL: <http://arxiv.org/abs/1807.10284>.
- [42] Sukanta Panda, Abbas Tinwala, and Archit Vidyarthi. “Covariant Effective Action for Generalized Proca Theories”. In: *Journal of Cosmology and Astroparticle Physics* 2022.01 (Jan. 2022). arXiv:2112.04391 [hep-th], p. 062. ISSN: 1475-7516. DOI: [10.1088/1475-7516/2022/01/062](https://doi.org/10.1088/1475-7516/2022/01/062). URL: <http://arxiv.org/abs/2112.04391>.
- [43] Yu-Qi Dong, Yu-Qiang Liu, and Yu-Xiao Liu. *Polarization modes of gravitational waves in generalized Proca theory*. arXiv:2305.12516. Feb. 2024. URL: <http://arxiv.org/abs/2305.12516> (visited on 11/04/2024).
- [44] Antonio De Felice et al. “Cosmology in generalized Proca theories”. en. In: *Journal of Cosmology and Astroparticle Physics* 2016.6 (June 2016), p. 048. ISSN: 1475-7516. DOI: [10.1088/1475-7516/2016/06/048](https://doi.org/10.1088/1475-7516/2016/06/048). URL: <https://ui.adsabs.harvard.edu/abs/2016JCAP...06..048D/abstract>.

- [45] Lavinia Heisenberg, Guangzi Xu, and Jann Zosso. *Unifying Ordinary and Null Memory*. arXiv:2401.05936 [gr-qc]. Dec. 2024. DOI: [10.48550/arXiv.2401.05936](https://doi.org/10.48550/arXiv.2401.05936). URL: <http://arxiv.org/abs/2401.05936>.
- [46] Shaon Sahoo. *Inverse Vector Operators*. arXiv:0804.2239 [math-ph]. Aug. 2010. DOI: [10.48550/arXiv.0804.2239](https://doi.org/10.48550/arXiv.0804.2239). URL: <http://arxiv.org/abs/0804.2239>.
- [47] Paolo Creminelli et al. “Gravitational Wave Decay into Dark Energy”. In: *Journal of Cosmology and Astroparticle Physics* 2018.12 (Dec. 2018). arXiv:1809.03484 [astro-ph], pp. 025–025. ISSN: 1475-7516. DOI: [10.1088/1475-7516/2018/12/025](https://doi.org/10.1088/1475-7516/2018/12/025). URL: <http://arxiv.org/abs/1809.03484>.
- [48] Cyril Pitrou, Xavier Roy, and Obinna Umeh. “xPand: An algorithm for perturbing homogeneous cosmologies”. In: *Classical and Quantum Gravity* 30.16 (Aug. 2013). arXiv:1302.6174 [astro-ph], p. 165002. ISSN: 0264-9381, 1361-6382. DOI: [10.1088/0264-9381/30/16/165002](https://doi.org/10.1088/0264-9381/30/16/165002). URL: <http://arxiv.org/abs/1302.6174>.
- [49] Pierre Baldi, Peter Sadowski, and Daniel Whiteson. “Searching for Exotic Particles in High-Energy Physics with Deep Learning”. In: *Nature Communications* 5.1 (July 2014). arXiv:1402.4735 [hep-ph], p. 4308. ISSN: 2041-1723. DOI: [10.1038/ncomms5308](https://doi.org/10.1038/ncomms5308). URL: <http://arxiv.org/abs/1402.4735>.
- [50] *Machine learning and the physical sciences* | *Rev. Mod. Phys.* URL: <https://journals.aps.org/rmp/abstract/10.1103/RevModPhys.91.045002#>.
- [51] Alexander Radovic et al. “Machine learning at the energy and intensity frontiers of particle physics”. English. In: *Nature (London)* 560.7716 (Aug. 2018). Institution: SLAC National Accelerator Laboratory (SLAC), Menlo Park, CA (United States); Fermi National Accelerator Laboratory (FNAL), Batavia, IL (United States) Number: FERMILAB-PUB-18-436-ND Publisher: Nature Publishing Group. ISSN: 0028-0836. DOI: [10.1038/s41586-018-0361-2](https://doi.org/10.1038/s41586-018-0361-2). URL: <https://www.osti.gov/pages/biblio/1469751>.
- [52] Daniel George and E. A. Huerta. “Deep Learning for Real-time Gravitational Wave Detection and Parameter Estimation: Results with Advanced LIGO Data”. In: *Physics Letters B* 778 (Mar. 2018). arXiv:1711.03121 [gr-qc], pp. 64–70. ISSN: 03702693. DOI: [10.1016/j.physletb.2017.12.053](https://doi.org/10.1016/j.physletb.2017.12.053). URL: <http://arxiv.org/abs/1711.03121>.

- [53] Daniel George and E. A. Huerta. “Deep neural networks to enable real-time multimessenger astrophysics”. en. In: *Physical Review D* 97.4 (Feb. 2018), p. 044039. ISSN: 2470-0010, 2470-0029. DOI: [10.1103/PhysRevD.97.044039](https://doi.org/10.1103/PhysRevD.97.044039). URL: <https://link.aps.org/doi/10.1103/PhysRevD.97.044039>.
- [54] Wei Wei et al. “Deep Learning with Quantized Neural Networks for Gravitational-wave Forecasting of Eccentric Compact Binary Coalescence”. English. In: *The Astrophysical Journal* 919.2 (Sept. 2021). Institution: Argonne National Laboratory (ANL), Argonne, IL (United States) Publisher: IOP Publishing. ISSN: 0004-637X. DOI: [10.3847/1538-4357/ac1121](https://doi.org/10.3847/1538-4357/ac1121). URL: <https://www.osti.gov/pages/biblio/1844212>.
- [55] Andriy Burkov. *The hundred-page machine learning book*. eng. Polen: Andriy Burkov, 2019. ISBN: 978-1-9995795-0-0 978-1-9995795-1-7.
- [56] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. Adaptive computation and machine learning. Cambridge, Massachusetts: The MIT Press, 2016. ISBN: 978-0-262-03561-3.
- [57] Yann LeCun et al. “Gradient-Based Learning Applied to Document Recognition”. en. In: (1998).
- [58] Adam: *A Method for Stochastic Optimization*. | BibSonomy. URL: <https://www.bibsonomy.org/bibtex/c14f3bd32b4636eff1d0234f08025bd5#>.
- [59] Nitish Srivastava et al. “Dropout: A Simple Way to Prevent Neural Networks from Overfitting”. en. In: ().
- [60] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. “Deep learning”. en. In: *Nature* 521.7553 (May 2015), pp. 436–444. ISSN: 0028-0836, 1476-4687. DOI: [10.1038/nature14539](https://doi.org/10.1038/nature14539). URL: <https://www.nature.com/articles/nature14539>.
- [61] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. “Learning representations by back-propagating errors”. en. In: *Nature* 323.6088 (Oct. 1986), pp. 533–536. ISSN: 0028-0836, 1476-4687. DOI: [10.1038/323533a0](https://doi.org/10.1038/323533a0). URL: <https://www.nature.com/articles/323533a0>.
- [62] Cortes, C. and Vapnik, V. (1995) *Support-Vector Networks*. *Machine Learning*, 20, 273-297. - *References - Scientific Research Publishing*. URL: <https://www.scirp.org/reference/referencespapers?referenceid=1150668>.

- [63] J. R. Quinlan. “Induction of decision trees”. en. In: *Machine Learning* 1.1 (Mar. 1986), pp. 81–106. ISSN: 0885-6125, 1573-0565. DOI: [10.1007/BF00116251](https://doi.org/10.1007/BF00116251). URL: <http://link.springer.com/10.1007/BF00116251>.
- [64] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. eng. Softcover reprint of the original 1st edition 2006 (corrected at 8th printing 2009). Information science and statistics. New York, NY: Springer New York, 2016. ISBN: 978-1-4939-3843-8.
- [65] Michele Maggiore. *Gravitational waves*. OCLC: ocn180464569. Oxford: Oxford University Press, 2008. ISBN: 978-0-19-857074-5 978-0-19-857089-9.
- [66] Ajit Kumar Mehta et al. “Accurate inspiral-merger-ringdown gravitational waveforms for nonspinning black-hole binaries including the effect of subdominant modes”. en. In: *Physical Review D* 96.12 (Dec. 2017), p. 124010. ISSN: 2470-0010, 2470-0029. DOI: [10.1103/PhysRevD.96.124010](https://doi.org/10.1103/PhysRevD.96.124010). URL: <https://link.aps.org/doi/10.1103/PhysRevD.96.124010> (visited on 04/12/2025).
- [67] Ajit Kumar Mehta et al. “Accurate inspiral-merger-ringdown gravitational waveforms for non-spinning black-hole binaries including the effect of subdominant modes”. en. In: *Physical Review D* 96.12 (Dec. 2017). arXiv:1708.03501 [gr-qc], p. 124010. ISSN: 2470-0010, 2470-0029. DOI: [10.1103/PhysRevD.96.124010](https://doi.org/10.1103/PhysRevD.96.124010). URL: <http://arxiv.org/abs/1708.03501>.
- [68] Sebastian Khan et al. “Frequency-domain gravitational waves from non-precessing black-hole binaries. II. A phenomenological model for the advanced detector era”. In: *Physical Review D* 93.4 (Feb. 2016). arXiv:1508.07253 [gr-qc], p. 044007. ISSN: 2470-0010, 2470-0029. DOI: [10.1103/PhysRevD.93.044007](https://doi.org/10.1103/PhysRevD.93.044007). URL: <http://arxiv.org/abs/1508.07253>.
- [69] Lavinia Heisenberg, Shayan Hemmatyar, Hector Villarrubia Roj. “Multi-Method Classification of Gravitational Wave Observables”. (In preparation).
- [70] Yi Pan et al. “Inspiral-merger-ringdown multipolar waveforms of nonspinning black-hole binaries using the effective-one-body formalism”. en. In: *Physical Review D* 84.12 (Dec. 2011), p. 124052. ISSN: 1550-7998, 1550-2368. DOI: [10.1103/PhysRevD.84.124052](https://doi.org/10.1103/PhysRevD.84.124052). URL: <https://link.aps.org/doi/10.1103/PhysRevD.84.124052> (visited on 04/12/2025).

- [71] Chris Van Den Broeck and Anand S. Sengupta. “Phenomenology of amplitude-corrected post-Newtonian gravitational waveforms for compact binary inspiral. I. Signal-to-noise ratios”. en. In: *Classical and Quantum Gravity* 24.1 (Jan. 2007). arXiv:gr-qc/0607092, pp. 155–176. ISSN: 0264-9381, 1361-6382. DOI: [10.1088/0264-9381/24/1/009](https://doi.org/10.1088/0264-9381/24/1/009). URL: <http://arxiv.org/abs/gr-qc/0607092>.
- [72] *Open Data from the Third Observing Run of LIGO, Virgo, KAGRA, and GEO - IOPscience*. URL: <https://iopscience.iop.org/article/10.3847/1538-4365/acdc9f> (visited on 04/12/2025).
- [73] Alex Nitz et al. *gwastro/pycbc: v2.3.3 release of PyCBC*. Jan. 2024. DOI: [10.5281/ZENODO.10473621](https://doi.org/10.5281/ZENODO.10473621). URL: <https://zenodo.org/doi/10.5281/zenodo.10473621>.
- [74] Christopher J. Moore, Robert H. Cole, and Christopher P. L. Berry. “Gravitational-wave sensitivity curves”. en. In: *Classical and Quantum Gravity* 32.1 (Jan. 2015). arXiv:1408.0740 [gr-qc], p. 015014. ISSN: 0264-9381, 1361-6382. DOI: [10.1088/0264-9381/32/1/015014](https://doi.org/10.1088/0264-9381/32/1/015014). URL: <http://arxiv.org/abs/1408.0740>.
- [75] Martin Abadi et al. “TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems”. en. In: ().
- [76] Fabian Pedregosa et al. “Scikit-learn: Machine Learning in Python”. en. In: *MACHINE LEARNING IN PYTHON* ().
- [77] Ivo Düntsch and Günther Gediga. “Confusion matrices and rough set data analysis”. In: *Journal of Physics: Conference Series* 1229.1 (May 2019). arXiv:1902.01487 [cs], p. 012055. ISSN: 1742-6588, 1742-6596. DOI: [10.1088/1742-6596/1229/1/012055](https://doi.org/10.1088/1742-6596/1229/1/012055). URL: <http://arxiv.org/abs/1902.01487>.
- [78] Michael Ebersold et al. “Gravitational-wave amplitudes for compact binaries in eccentric orbits at the third post-Newtonian order: Memory contributions”. en. In: *Physical Review D* 100.8 (Oct. 2019), p. 084043. ISSN: 2470-0010, 2470-0029. DOI: [10.1103/PhysRevD.100.084043](https://doi.org/10.1103/PhysRevD.100.084043). URL: <https://link.aps.org/doi/10.1103/PhysRevD.100.084043>.
- [79] Cecilio García-Quirós et al. “Multimode frequency-domain model for the gravitational wave signal from nonprecessing black-hole binaries”. en. In: *Physical Review D* 102.6 (Sept. 2020), p. 064002. ISSN: 2470-0010, 2470-0029. DOI: [10.1103/PhysRevD.102.064002](https://doi.org/10.1103/PhysRevD.102.064002). URL: <https://link.aps.org/doi/10.1103/PhysRevD.102.064002>.



- [80] Lavinia Heisenberg, Hector Villarrubia-Rojó, and Jann Zosso. “Simultaneously solving the  $H_0$  and  $\sigma_8$  tensions with late dark energy”. en. In: *Physics of the Dark Universe* 39 (Feb. 2023), p. 101163. ISSN: 22126864. DOI: [10.1016/j.dark.2022.101163](https://doi.org/10.1016/j.dark.2022.101163). URL: <https://linkinghub.elsevier.com/retrieve/pii/S2212686422001364>.
- [81] Lavinia Heisenberg, Hector Villarrubia-Rojó, and Jann Zosso. “Can late-time extensions solve the  $H_0$  and  $\sigma_8$  tensions?” en. In: *Physical Review D* 106.4 (Aug. 2022), p. 043503. ISSN: 2470-0010, 2470-0029. DOI: [10.1103/PhysRevD.106.043503](https://doi.org/10.1103/PhysRevD.106.043503). URL: <https://link.aps.org/doi/10.1103/PhysRevD.106.043503>.
- [82] Ajit Kumar Mehta et al. “Tests of General Relativity with Gravitational-Wave Observations using a Flexible–Theory-Independent Method”. In: *Physical Review D* 107.4 (Feb. 2023). arXiv:2203.13937 [gr-qc], p. 044020. ISSN: 2470-0010, 2470-0029. DOI: [10.1103/PhysRevD.107.044020](https://doi.org/10.1103/PhysRevD.107.044020). URL: <http://arxiv.org/abs/2203.13937>.
- [83] Hugo Cui, Florent Krzakala, and Lenka Zdeborová. *Bayes-optimal Learning of Deep Random Networks of Extensive-width*. arXiv:2302.00375 [stat]. June 2023. DOI: [10.48550/arXiv.2302.00375](https://doi.org/10.48550/arXiv.2302.00375). URL: <http://arxiv.org/abs/2302.00375>.
- [84] Benjamin Aubin et al. *Generalization error in high-dimensional perceptrons: Approaching Bayes error with convex optimization*. en. arXiv:2006.06560 [stat]. Nov. 2020. DOI: [10.48550/arXiv.2006.06560](https://doi.org/10.48550/arXiv.2006.06560). URL: <http://arxiv.org/abs/2006.06560>.
- [85] Nicolas Yunes and Frans Pretorius. “Fundamental Theoretical Bias in Gravitational Wave Astrophysics and the Parameterized Post-Einsteinian Framework”. In: *Physical Review D* 80.12 (Dec. 2009). arXiv:0909.3328 [gr-qc], p. 122003. ISSN: 1550-7998, 1550-2368. DOI: [10.1103/PhysRevD.80.122003](https://doi.org/10.1103/PhysRevD.80.122003). URL: <http://arxiv.org/abs/0909.3328>.
- [86] Yiqi Xie et al. “Neural post-Einsteinian framework for efficient theory-agnostic tests of general relativity with gravitational waves”. In: *Physical Review D* 110.2 (July 2024). arXiv:2403.18936 [gr-qc], p. 024036. ISSN: 2470-0010, 2470-0029. DOI: [10.1103/PhysRevD.110.024036](https://doi.org/10.1103/PhysRevD.110.024036). URL: <http://arxiv.org/abs/2403.18936>.
- [87] Sharaban Tahura and Kent Yagi. “Parameterized Post-Einsteinian Gravitational Waveforms in Various Modified Theories of Gravity”. In: *Physical Review D* 98.8 (Oct.



- 2018). arXiv:1809.00259 [gr-qc], p. 084042. ISSN: 2470-0010, 2470-0029. DOI: [10.1103/PhysRevD.98.084042](https://doi.org/10.1103/PhysRevD.98.084042). URL: <http://arxiv.org/abs/1809.00259>.
- [88] Ippocratis D. Saltas et al. *Modified gravity, gravitational waves and the large-scale structure of the Universe: A brief report*. arXiv:1812.03969 [astro-ph]. Dec. 2018. DOI: [10.48550/arXiv.1812.03969](https://doi.org/10.48550/arXiv.1812.03969). URL: <http://arxiv.org/abs/1812.03969>.
- [89] Carla R. Almeida, O. Galkina, and J. C. Fabris. “Quantum and classical cosmology in the Brans-Dicke theory”. In: *Universe* 7.8 (Aug. 2021). arXiv:2106.14299 [gr-qc], p. 286. ISSN: 2218-1997. DOI: [10.3390/universe7080286](https://doi.org/10.3390/universe7080286). URL: <http://arxiv.org/abs/2106.14299>.
- [90] Claudia de Rham. “Massive Gravity”. In: *Living Reviews in Relativity* 17.1 (Dec. 2014). arXiv:1401.4173 [hep-th], p. 7. ISSN: 2367-3613, 1433-8351. DOI: [10.12942/lrr-2014-7](https://doi.org/10.12942/lrr-2014-7). URL: <http://arxiv.org/abs/1401.4173>.
- [91] Kurt Hinterbichler. “Theoretical Aspects of Massive Gravity”. In: *Reviews of Modern Physics* 84.2 (May 2012). arXiv:1105.3735 [hep-th], pp. 671–710. ISSN: 0034-6861, 1539-0756. DOI: [10.1103/RevModPhys.84.671](https://doi.org/10.1103/RevModPhys.84.671). URL: <http://arxiv.org/abs/1105.3735>.
- [92] Priscilla Canizares, Jonathan R. Gair, and Carlos F. Sopuerta. “Testing Chern-Simons Modified Gravity with Gravitational-Wave Detections of Extreme-Mass-Ratio Binaries”. In: *Physical Review D* 86.4 (Aug. 2012). arXiv:1205.1253 [gr-qc], p. 044010. ISSN: 1550-7998, 1550-2368. DOI: [10.1103/PhysRevD.86.044010](https://doi.org/10.1103/PhysRevD.86.044010). URL: <http://arxiv.org/abs/1205.1253>.
- [93] Maria Okounkova et al. “Numerical binary black hole mergers in dynamical Chern-Simons: I. Scalar field”. In: *Physical Review D* 96.4 (Aug. 2017). arXiv:1705.07924 [gr-qc], p. 044020. ISSN: 2470-0010, 2470-0029. DOI: [10.1103/PhysRevD.96.044020](https://doi.org/10.1103/PhysRevD.96.044020). URL: <http://arxiv.org/abs/1705.07924>.
- [94] Claudia de Rham et al. “Graviton Mass Bounds”. In: *Reviews of Modern Physics* 89.2 (May 2017). arXiv:1606.08462 [astro-ph], p. 025004. ISSN: 0034-6861, 1539-0756. DOI: [10.1103/RevModPhys.89.025004](https://doi.org/10.1103/RevModPhys.89.025004). URL: <http://arxiv.org/abs/1606.08462>.

- 
- [95] Ercan Kilicarslan and Bayram Tekin. “Graviton Mass and Memory”. In: *The European Physical Journal C* 79.2 (Feb. 2019). arXiv:1805.02240 [gr-qc], p. 114. ISSN: 1434-6044, 1434-6052. DOI: [10.1140/epjc/s10052-019-6636-4](https://doi.org/10.1140/epjc/s10052-019-6636-4). URL: <http://arxiv.org/abs/1805.02240>.
- [96] Ajit Kumar Mehta et al. “Tests of General Relativity with Gravitational-Wave Observations using a Flexible–Theory-Independent Method”. en. In: *Physical Review D* 107.4 (Feb. 2023). arXiv:2203.13937 [gr-qc], p. 044020. ISSN: 2470-0010, 2470-0029. DOI: [10.1103/PhysRevD.107.044020](https://doi.org/10.1103/PhysRevD.107.044020). URL: <http://arxiv.org/abs/2203.13937> (visited on 04/12/2025).