Modeling Gravitational-wave (GW) waveforms based on Machine learning (ML)

Expanding GW waveforms of generic-spin binary black hole (BBH) mergers to new physical parameter coverage using Deep Neural Network (DNN)

Kexuan Wu
Sukanta Bose
Praveer Tiwari
Shen, Zhuoran

Washington State University

WASHINGTON STATE UNIVERSITY

What does it mean NEW physical parameter coverage?

What is generic-spin?

What GW waveforms?

What parameter are you extending?

Why binary black hole?

Why expanding?

Why Deep Learning?

Why Machine learning?

Choose approximants for training

GWSurrogate

NRSur7dq4

A surrogate model for **generic spin numerical relativity** waveforms

Training Set: 1528 NR simulation from Spectral Einstein Code (SpEC) developed by the SXS collaboration (1464 precessing NR simulations + 64 aligned-spin simulations)

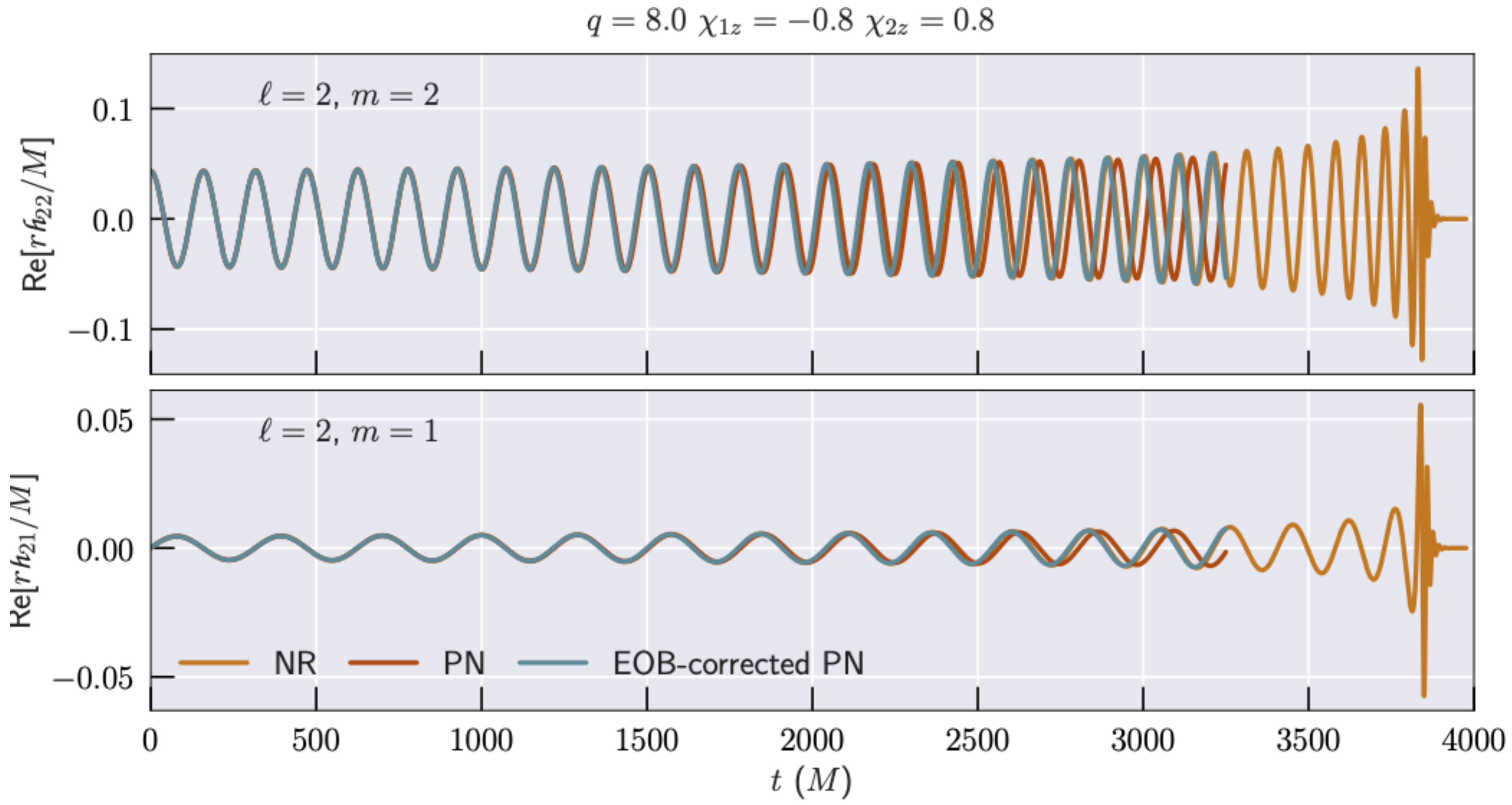
Free parameters:

$\chi_{x1}, \chi_{y1}, \chi_{z1}, \chi_{x2}, \chi_{y2}, \chi_{z2}, q(mass\ ratio\ q \leq 4)$

NRHybSur3dq8

A surrogate model for **hybridized nonprecessing numerical relativity** waveforms, that is valid for the entire LIGO band (starting at 20 Hz).

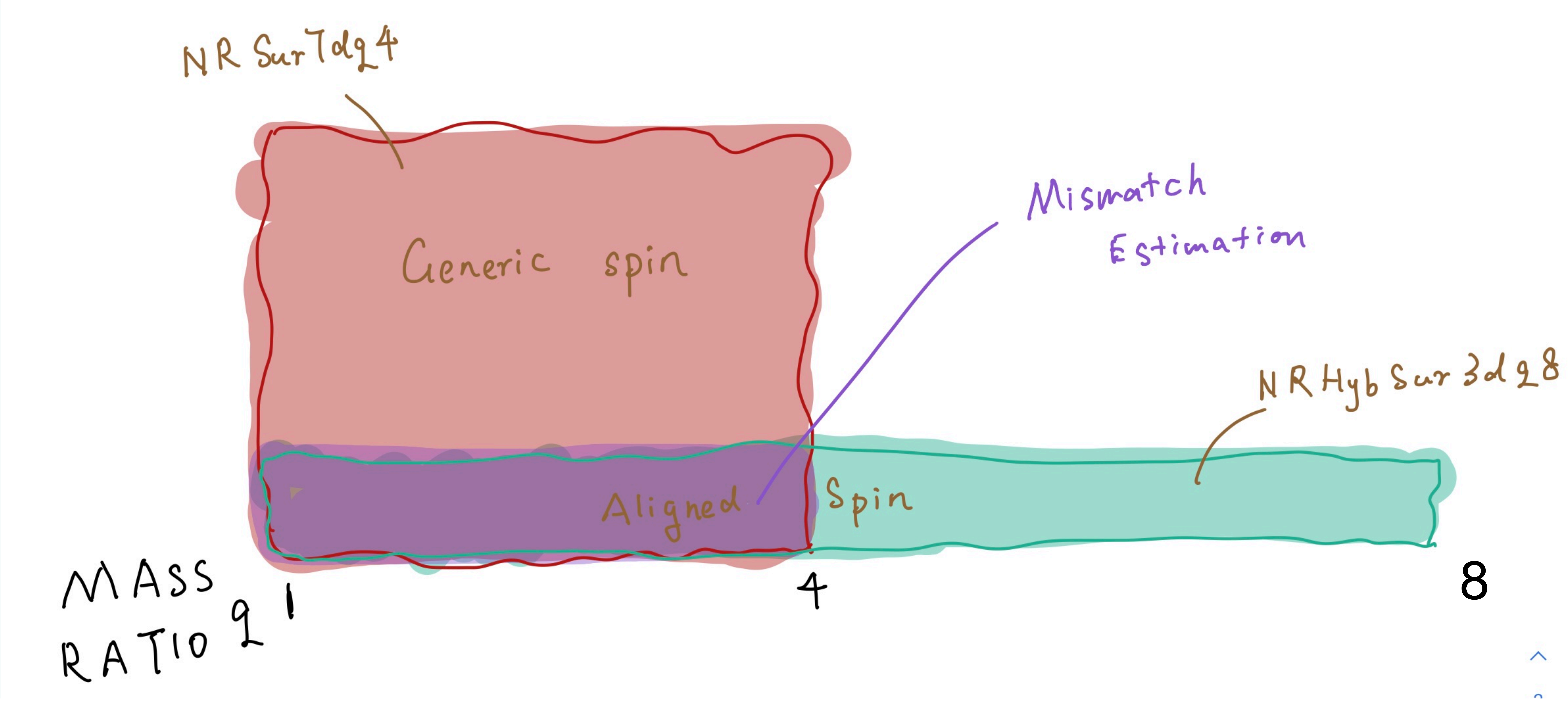
Free parameters: $\chi_{z1}, \chi_{z2}, q(mass\ ratio\ q \leq 8)$



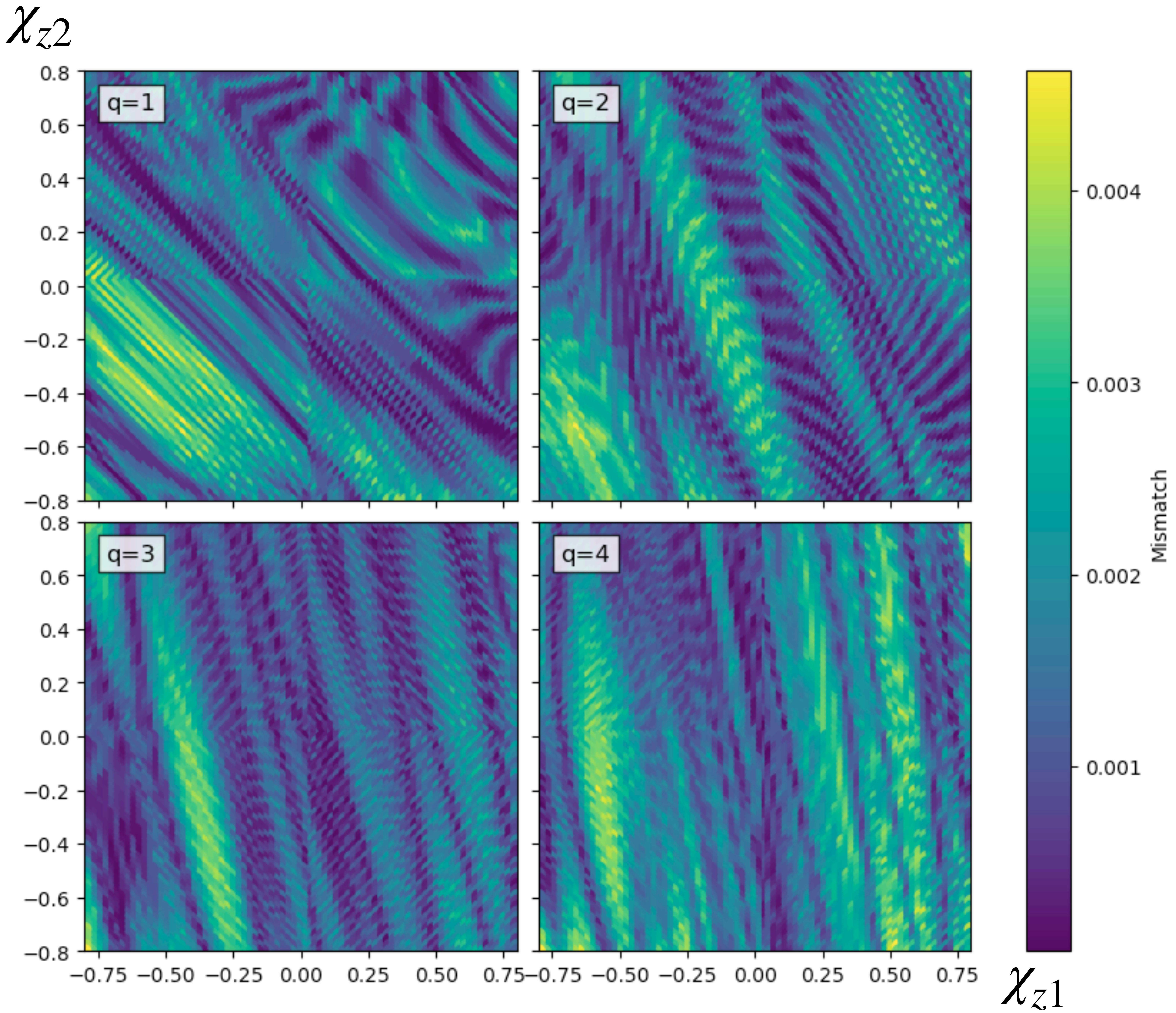
Reference: https://pypi.org/project/gwsurrogate/

Data preparation

Is that reasonable to use two models in one training?



Mismatch Analysis of Aligned-Spin BBHs for Various q



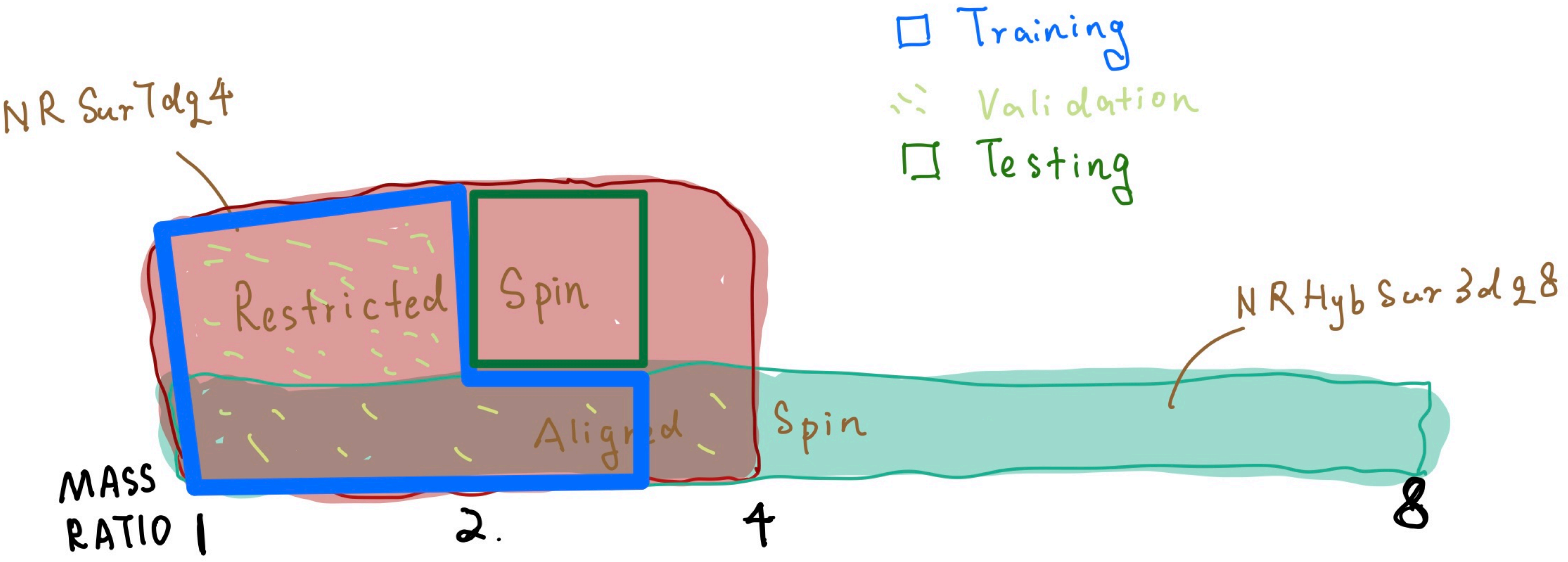
Mean Mismatch:
1.53e-03

Standard Deviation:
8.92e-05

Max Mismatch:
5.52 e-03

Training plan

From big picture to practical first step



Data preparation

Seeding, train/test in/out

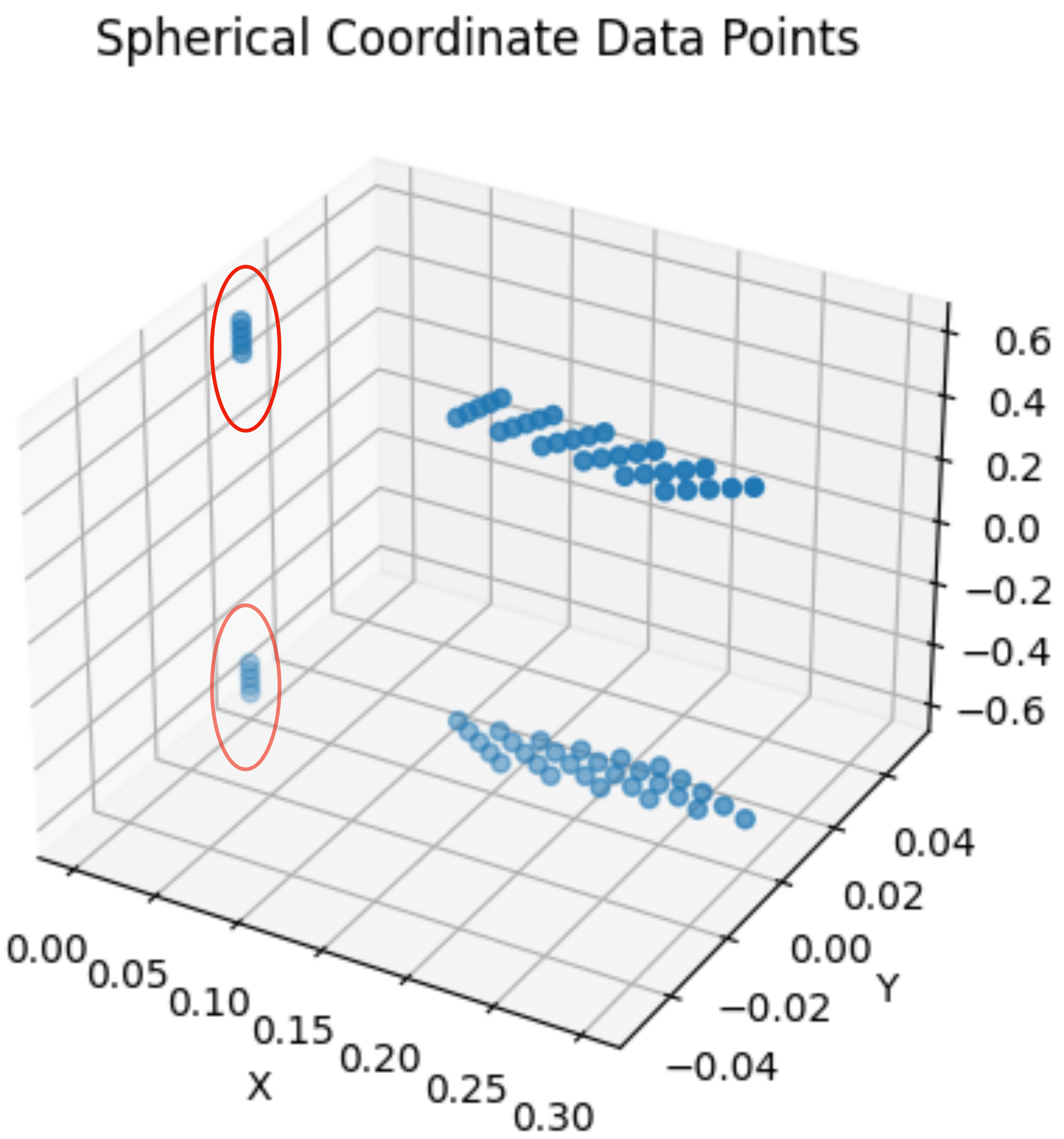
NRSur7dq4: $\pi/12 \leq \theta \leq \pi/6, \theta = 0, \pi$; $0.4 \leq r \leq 0.5$; $q = 1.0, 1.5, 2.0$ $for\ train$ $2.5, 3.0$ $for\ test$, Mode: (2,2)

NRHybSur3dq8: $\theta = 0, \pi$; $0.4 \leq r \leq 0.5$; $q = 1.0, 1.5, 2.0, 2.5, 3.0$, Mode:(2,2)

Train/Test in: $\chi_{x1}, \chi_{y1}, \chi_{z1}, \chi_{x2}, \chi_{y2}, \chi_{z2}, q$

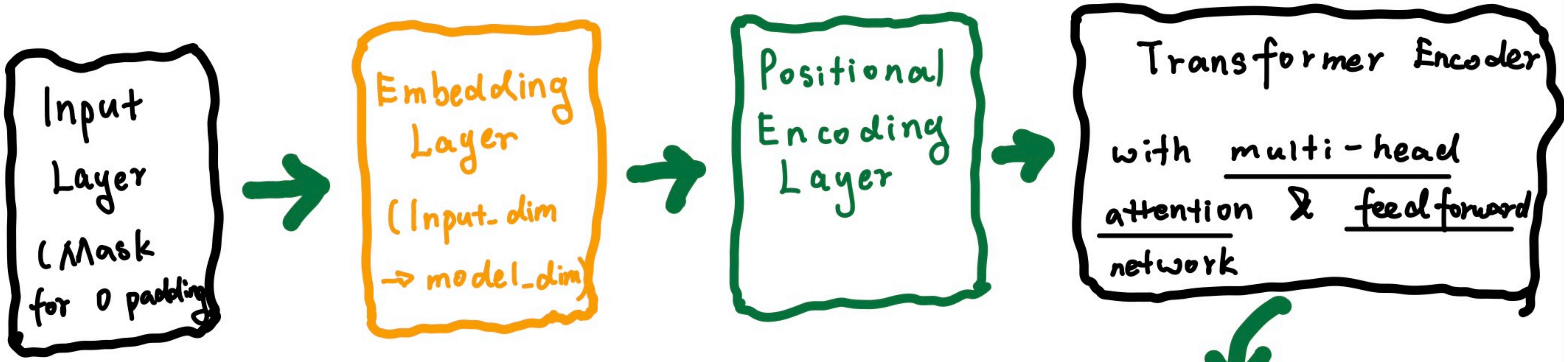
Train/Test out: amplitudes and phases

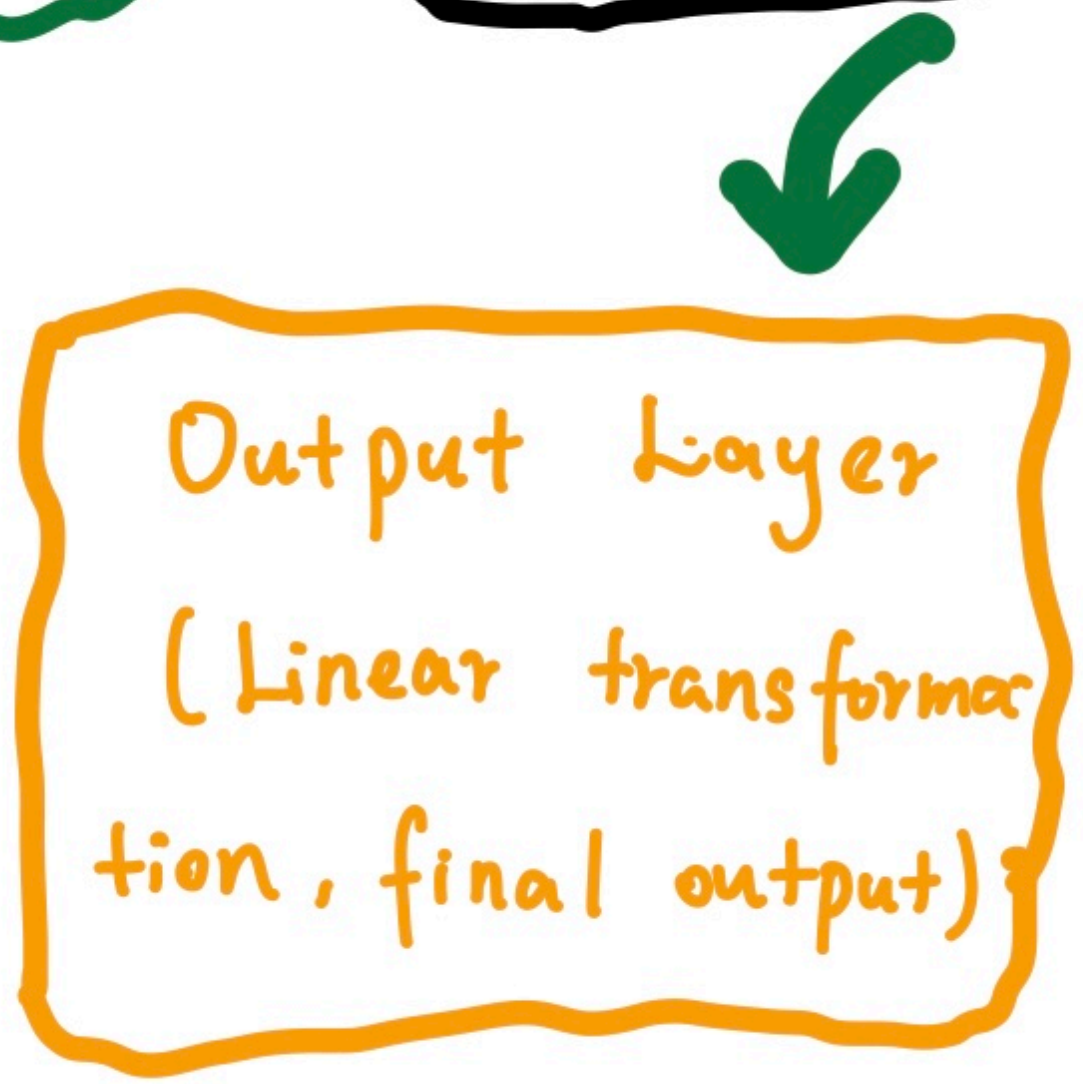
Aligning, Normalization, Zero padding…



Deep Neural Network

An attention-based transformer deep neural network

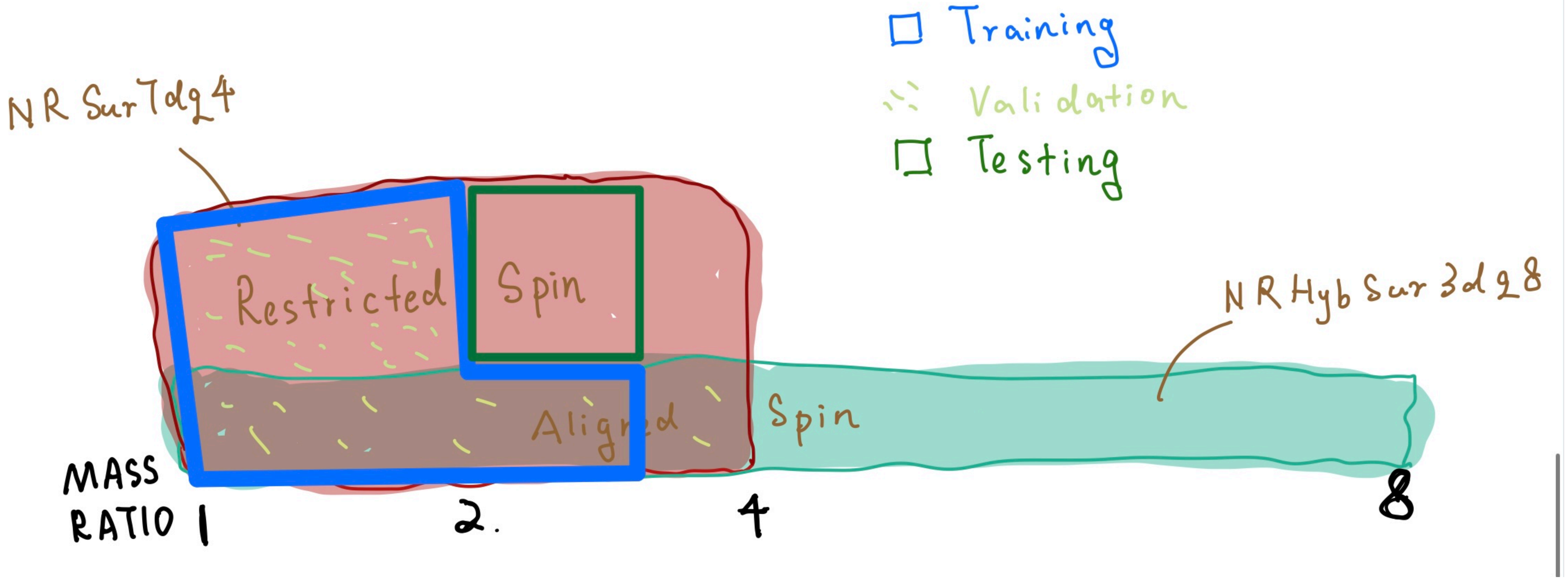


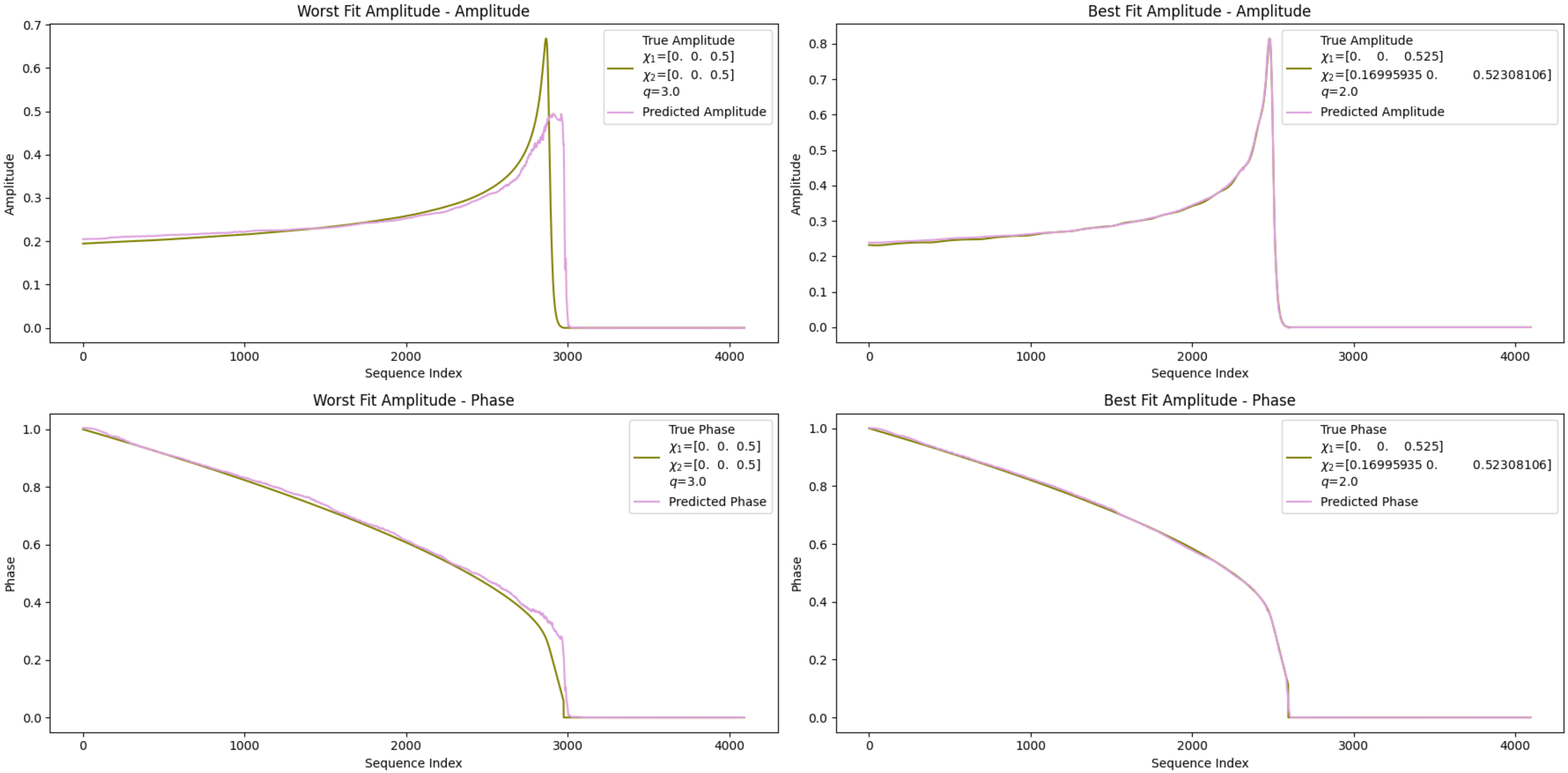


training epoch = 16
batch size = 10
model dim = 32
num of head = 8
num of encoder layer = 4
num of decoder layer = 4
feed forword dim = 128

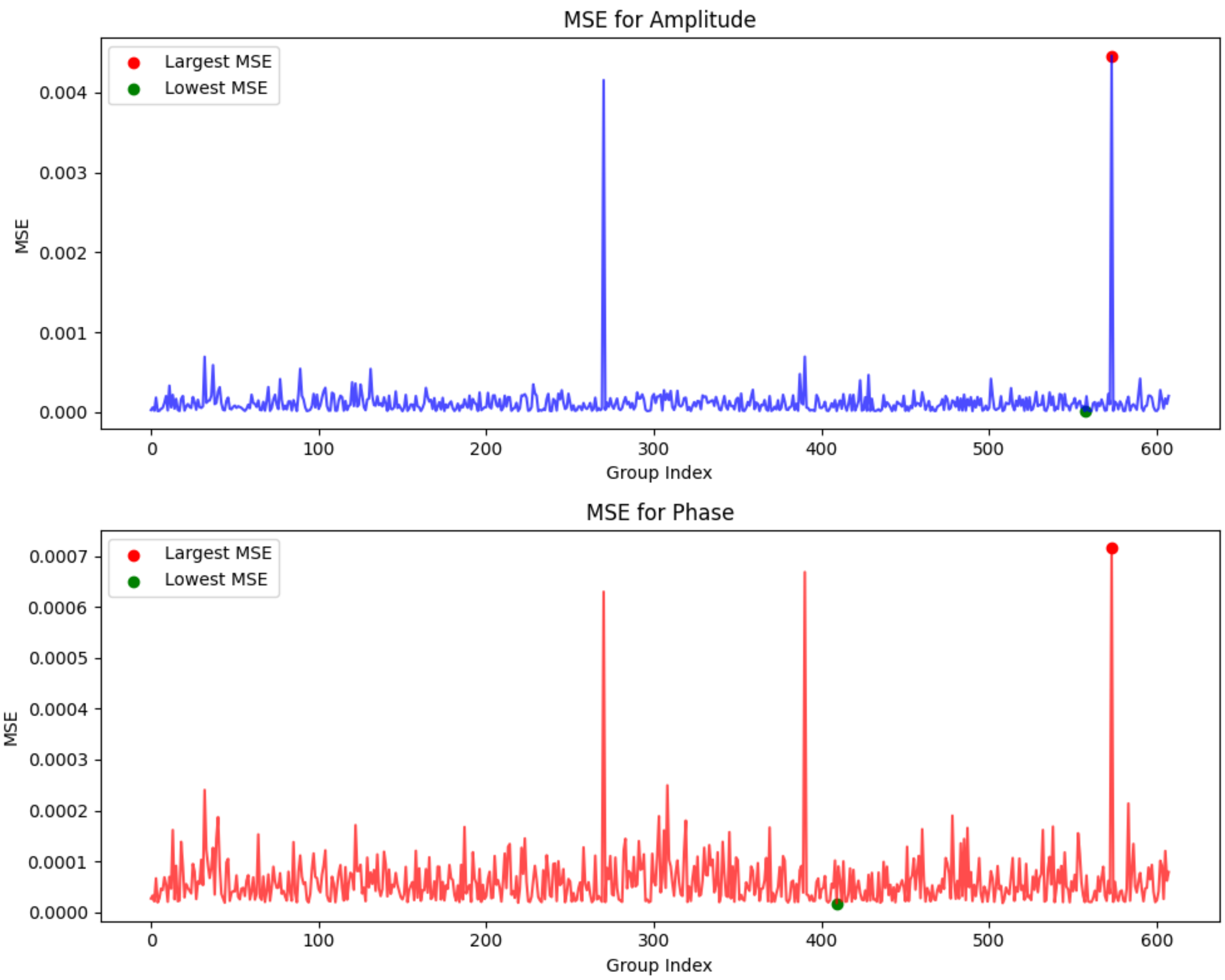
Validation

Interpolation: testing within training physics parameter coverage





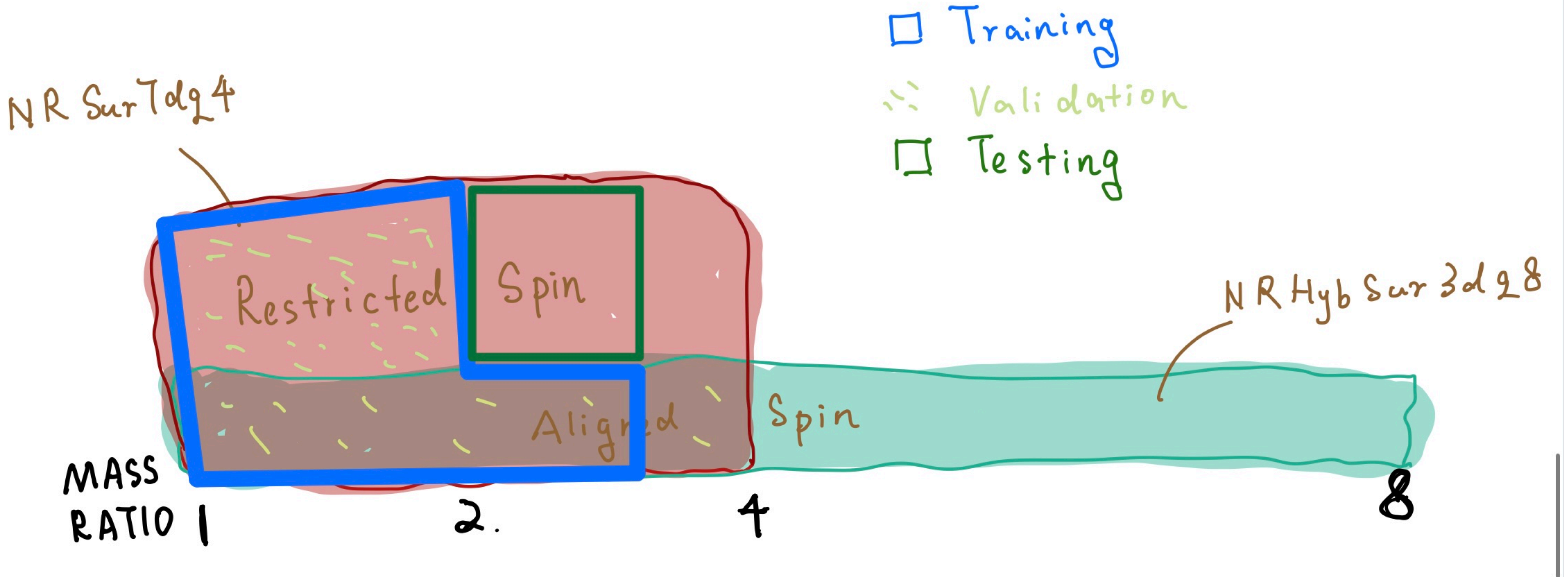
MSE loss for amplitude: 0.0001
MSE loss for phase: 0.0001

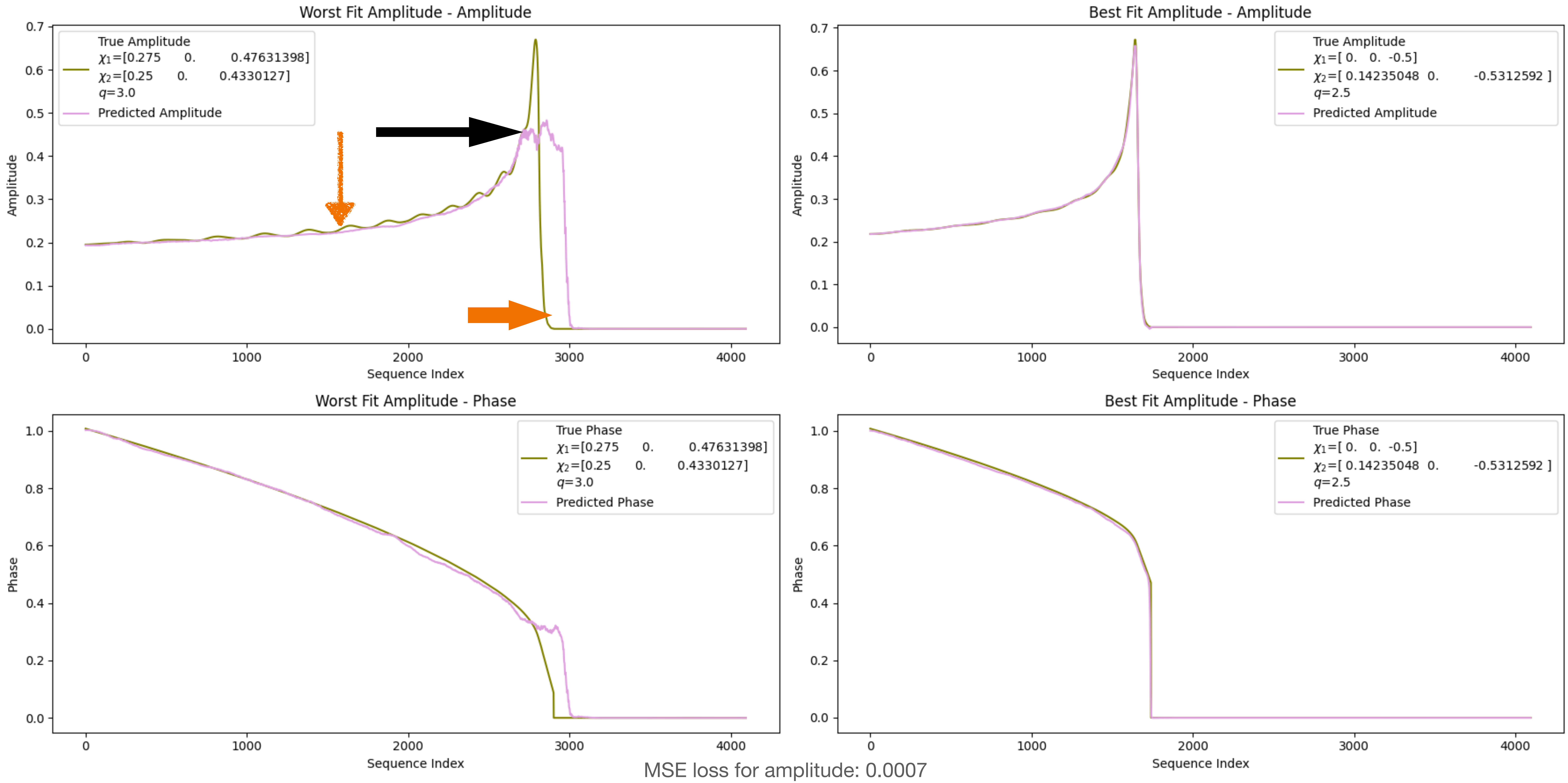


MSE for Amplitude
Largest MSE
Lowest MSE
MSE
0.004
0.003
0.002
0.001
0.000
0
100
200
300
400
500
600
Group Index
MSE for Phase
Largest MSE
Lowest MSE
MSE
0.0007
0.0006
0.0005
0.0004
0.0003
0.0002
0.0001
0.0000
0
100
200
300
400
500
600
Group Index

Testing

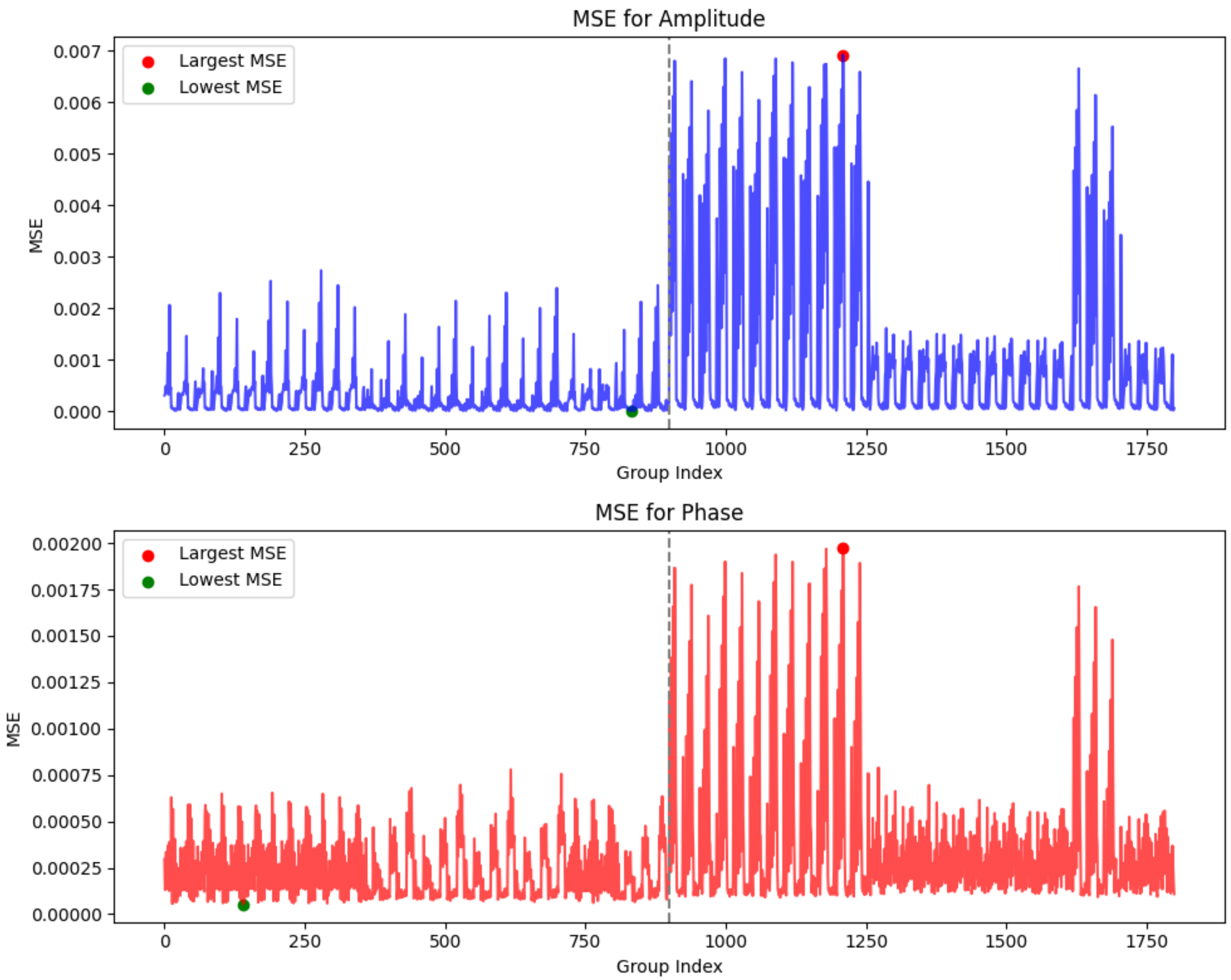
Extrapolation: testing within new physics parameter coverage





MSE loss for amplitude: 0.0007

MSE loss for phase: 0.0003



MSE for Amplitude
Largest MSE
Lowest MSE
MSE
0.007
0.006
0.005
0.004
0.003
0.002
0.001
0.000
0
250
500
750
1000
1250
1500
1750
Group Index
MSE for Phase
Largest MSE
Lowest MSE
MSE
0.00200
0.00175
0.00150
0.00125
0.00100
0.00075
0.00050
0.00025
0.00000
0
250
500
750
1000
1250
1500
1750
Group Index

What is the next step?

Computer scientist VS Physicist

Computer Science

Improve data preparation and NN architecture

Fine-tune hyperparameter

Switch study strategy: predict length of waveforms/ merging moment…

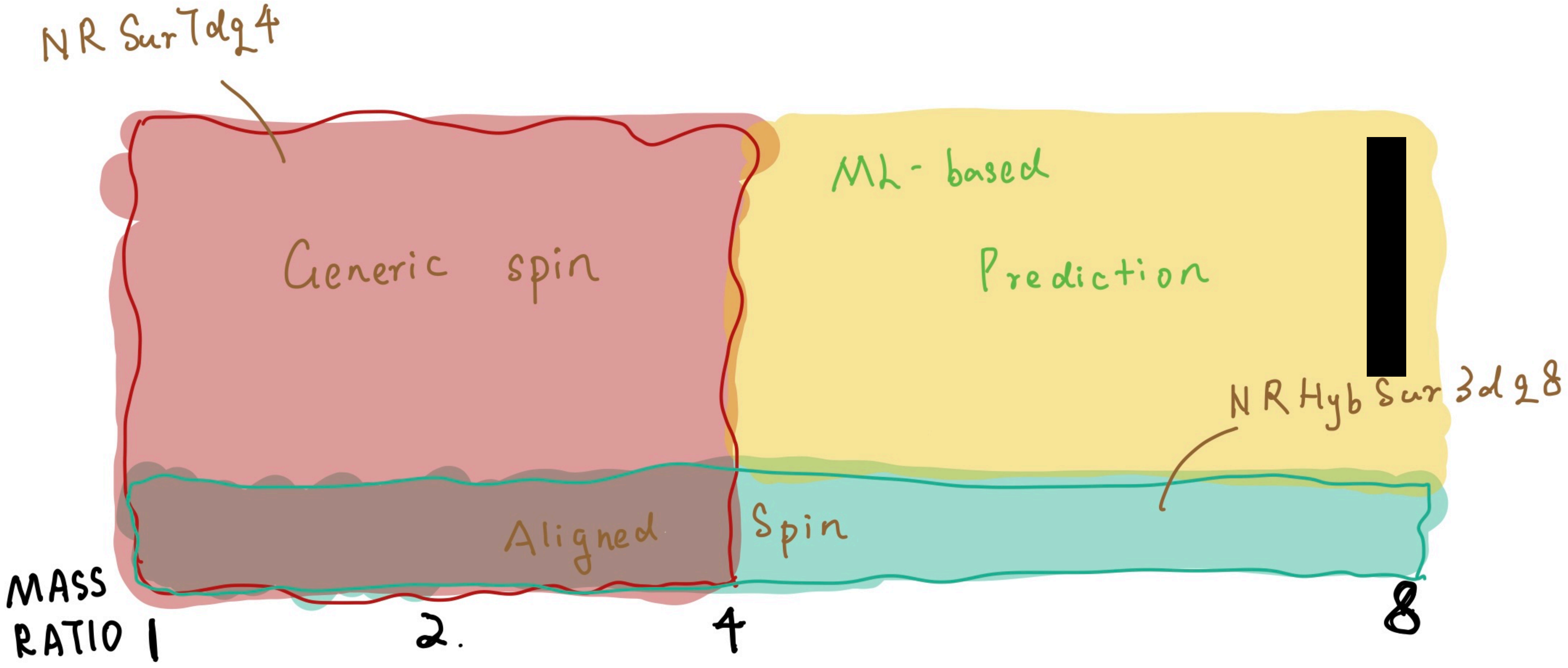
When data become larger, find the proper way for training

Physics

Include total mass as a new parameter

Proper step to take

Using mismatch to evaluate test/validation result



Model Binary GW waveforms by Machine Learning

Machine learning architecture

Mismatch:

In order to estimate the difference between two complexified waveforms, h_1 and h_2, we use:

$$MM = 1 - \max_{t_{\min} \le t \le t_{\max}} \left| \frac{\langle h_1, h_2 \rangle}{\sqrt{\langle h_1, h_1 \rangle \langle h_2, h_2 \rangle}} \right|$$

$$\langle h_1, h_2 \rangle = 4\mathrm{Re} \int_{f_{\min}}^{f_{\max}} \frac{\tilde{h}_1(f) \tilde{h}_2^*(f)}{S_n(f)} df,$$