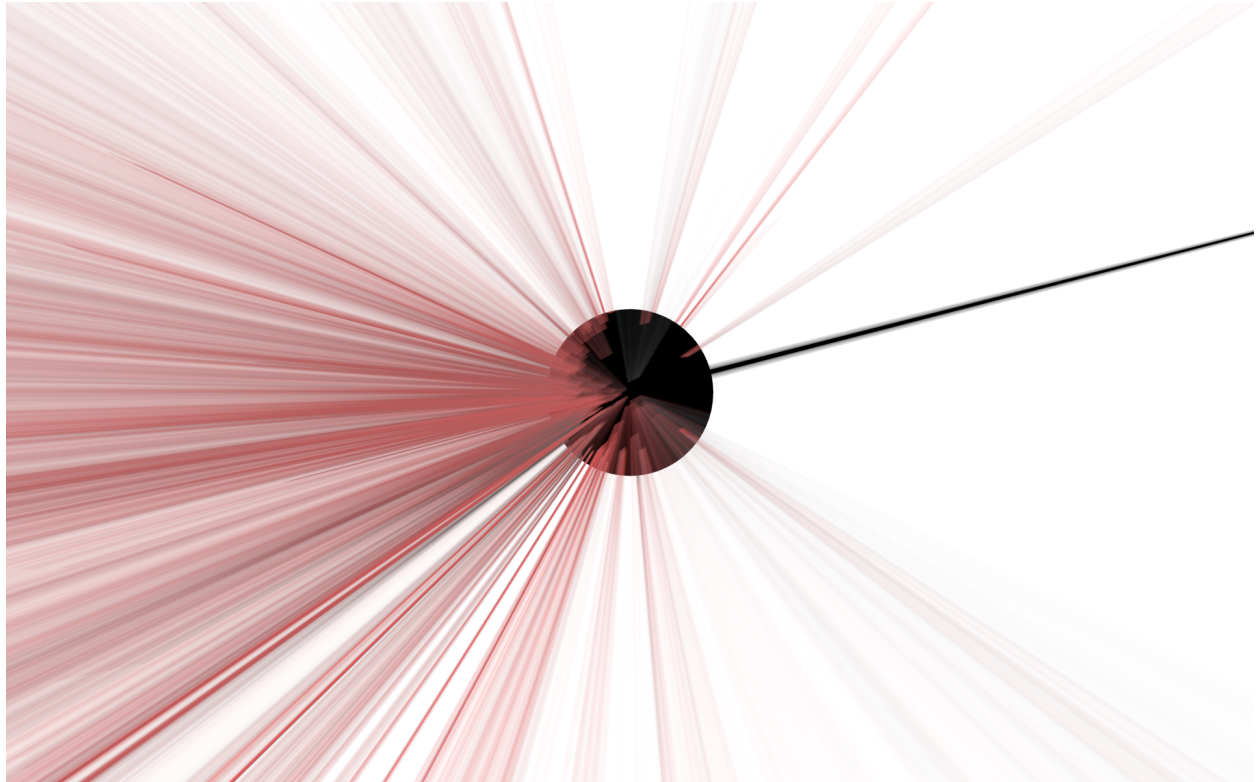




CHALMERS



# Implementering av Ab Initio-härledd Kärnresponsmodell för Kärn-neutrinointeraktion i NuWro

Kandidatarbete inom Fysik

TIFX11-VT26-08A:

LIAM FALK, MOSTAFA GHASEMI, GRIFFIN HISCOKE, JOSEFINE JÄRVDALÉN, HAMPUS MORIN

**INSTITUTIONEN FÖR FYSIK OCH ASTRONOMI**

CHALMERS TEKNISKA HÖGSKOLA  
Göteborg, Sverige 2026  
[www.chalmers.se](http://www.chalmers.se)

Implementering av Ab Initio-härledd Kärnresponsmodell för Kärn-neutrinointeraktion i NuWro

Implementation of an Ab Initio-derived Nuclear Response Model for Nuclei-neutrino Interaction in NuWro

LIAM FALK, MOSTAFA GHASEMI, GRIFFIN HISCOKE, JOSEFINE JÄRVDALEN & HAMPUS MORIN

© LIAM FALK, MOSTAFA GHASEMI, GRIFFIN HISCOKE, JOSEFINE JÄRVDALEN & HAMPUS MORIN, 2026.

Handledare: Joanna Ewa Sobczyk och Christian Forssén

Examinator: Jan Swenson

Kandidatarbete 2026

Institutionen för Fysik och Astronomi

Avdelningen för Subatomär, Högenergi- och Plasmafysik

Chalmers Tekniska Högskola

SE-412 96 Göteborg

Telefon: +46 31 772 1000

Framsida: Visualisering av simulerade utgående myoner från kvasielastisk neutrin-kärninteraktion, genererade av modellen implementerad i rapporten.

Skriven i L<sup>A</sup>T<sub>E</sub>X

Göteborg, Sverige 2026

# Sammandrag

Neutrinoexperiment är beroende av tillförlitliga modeller för neutrino-kärninteraktioner eftersom den inkommande neutrinons energi och själva reaktionsförloppet inte kan observeras direkt. I detta arbete undersöks om neurala nätverk kan användas för att interpolera ab initio-beräknade kärnresponsfunktioner, som på grund av höga beräkningskostnader endast finns tillgängliga i ett begränsat antal kinematiska punkter, och därmed möjliggöra användning av mer detaljerade kärnmodeller i händelsegeneratoren NuWro. Fokus ligger på kvasielastisk neutrino-kärnspridning där kärnresponsfunktionerna beskriver hur atomkärnan påverkas av den svaga växelverkan.

En neural interpolationsmodell utvecklades och utvärderades genom systematiska tester av nätverksarkitektur, aktiveringsfunktioner, indatarepresentationer, normalisering och förlustfunktioner. Den valda modellen användes därefter för att generera tabellerade responsfunktioner vilka implementerades i NuWro som en hybridmodell, där den ab initio-baserade modellen används vid låga rörelsemängdsöverföringar och NuWro:s inbyggda lokala fermigasmodell används vid högre värden. De resulterande tvärsnitten jämfördes med teoretiska beräkningar, den inbyggda lokala fermigasmodellen och experimentella resultat från T2K-experimentet i Japan.

Resultaten visar att det neurala nätverket kan interpolera responsfunktionerna väl trots ett begränsat antal ab initio-beräknade datapunkter. Den implementerade hybridmodellen kunde dock inte påvisas ge bättre överensstämmelse med T2K-data än NuWro:s inbyggda lokala fermigasmodell. Arbetet demonstrerar ändå en fungerande beräkningskedja för att implementera mer avancerade kärnmodeller i NuWro och utgör en grund för vidare studier.

Nyckelord: neutrino-kärnspridning, kvasielastisk spridning, ab initio, kärnresponsfunktioner, neurala nätverk, NuWro, T2K, tvärsnitt

# Abstract

Neutrino experiments depend on reliable models of neutrino–nucleus interactions since the incoming neutrino energy and the interaction process itself cannot be observed directly. This work investigates whether neural networks can be used to interpolate ab initio-calculated nuclear response functions, which due to high computational costs are only available at a limited number of kinematic points, and thereby enable the use of more detailed nuclear models in the event generator NuWro. The focus is on quasielastic neutrino–nucleus scattering where the nuclear response functions describe how the atomic nucleus is affected by the weak interaction.

A neural interpolation model was developed and evaluated through systematic tests of network architecture, activation functions, input representations, normalization and loss functions. The selected model was then used to generate tabulated response functions, which were implemented in NuWro as a hybrid model, where the ab initio-based model is used at low momentum transfers and NuWro’s built-in local Fermi gas model is used at higher values. The resulting cross sections were compared with theoretical calculations, the built-in local Fermi gas model and experimental results from the T2K experiment in Japan.

The results show that the neural network can interpolate the response functions well despite a limited number of ab initio-calculated data points. However, the implemented hybrid model could not be shown to give better agreement with T2K data than NuWro’s built-in local Fermi gas model. Nevertheless, the work demonstrates a functional computational pipeline for implementing more advanced nuclear models in NuWro and provides a basis for further studies.

Keywords: neutrino–nucleus scattering, quasielastic scattering, ab initio, nuclear response functions, neural networks, NuWro, T2K, cross sections

---

## Tillkännagivande

Vi vill tacka våra handledare Joanna Ewa Sobczyk och Christian Forssén för deras stöd och vägledning under arbetets gång.

Vi har även använt LLM-modeller som stöd i arbetet, främst för att söka efter information, generera kod samt få förslag på språkliga förbättringar i rapporten.

Andra och tredje stycket i avsnitt 2.3 har hämtats från planeringsrapporten.



# Innehållsförteckning

|          |                                                                         |           |
|----------|-------------------------------------------------------------------------|-----------|
| <b>1</b> | <b>Inledning</b>                                                        | <b>1</b>  |
| <b>2</b> | <b>Bakgrund</b>                                                         | <b>2</b>  |
| 2.1      | Neutriner och neutrinooscillation . . . . .                             | 2         |
| 2.2      | T2K-experimentet . . . . .                                              | 4         |
| 2.3      | Neutrino-kärnspridning . . . . .                                        | 5         |
| <b>3</b> | <b>Metodologi</b>                                                       | <b>9</b>  |
| 3.1      | Monte Carlo-metoder . . . . .                                           | 9         |
| 3.2      | Händelsegeneratorn NuWro . . . . .                                      | 10        |
| 3.3      | Neurala nätverk som interpolationsmetod . . . . .                       | 11        |
| <b>4</b> | <b>Utveckling av neural interpolationsmetod och tillämpning i NuWro</b> | <b>13</b> |
| 4.1      | Neurala nätverk . . . . .                                               | 14        |
| 4.2      | Prediktioner med den valda modellen . . . . .                           | 18        |
| 4.3      | Propagerad teoretisk osäkerhet . . . . .                                | 21        |
| 4.4      | Dubbeldifferentiellt tvärsnitt . . . . .                                | 23        |
| 4.5      | Implementation av ab initio-baserad modell i NuWro . . . . .            | 24        |
| 4.6      | Tvärsnitt från NuWro-händelser jämfört med teoretiskt . . . . .         | 26        |
| 4.7      | Jämförelse med data från T2K-experimentet . . . . .                     | 27        |
| <b>5</b> | <b>Slutsats</b>                                                         | <b>29</b> |
| 5.1      | Felkällor . . . . .                                                     | 29        |
| 5.2      | Vidare studier . . . . .                                                | 31        |
| 5.3      | Slutsats . . . . .                                                      | 32        |



# 1

## Inledning

Människans strävan att förstå universums uppbyggnad har under lång tid drivit högteknologisk utveckling och teoretisk modellbyggande inom kärn- och partikelfysik. Dessa forskningsområden har gjort det möjligt att identifiera ett stort antal elementarpartiklar och beskriva deras inbördes växelverkan med hög precision. Trots dessa framsteg återstår flera viktiga frågor. Egenskaper hos neutriner utgör ett sådant exempel. Dessa partiklar är svåra att studera experimentellt eftersom de mycket sällan växelverkar med materia. Trots detta spelar de en viktig roll inom partikelfysik, bland annat eftersom deras egenskaper kan ge information om fysik som ännu inte är fullständigt förstådd. För att kunna undersöka neutriner används storskaliga experiment där man försöker observera de partiklar som bildas när en neutrino växelverkar med kärnmateria i detektormaterialet [1].

En central svårighet i neutrino-experiment är att neutrino inte observeras direkt, utan att informationen om den måste återskapas från de partiklar som produceras i växelverkanprocessen. Det innebär att tolkning av experimentell data blir beroende av teoretiska modeller för hur neutriner växelverkar med atomkärnor. Problemet är att befintliga modeller bygger på kraftigt förenklade antaganden (med stora systematiska osäkerheter) eller att de är mer fundamentala men beräkningsmässigt mycket krävande.

För att beskriva neutrino-kärninteraktioner används ofta simuleringsprogram som genererar möjliga händelser utifrån teoretiska modeller. Detta kallas för en händelsegenerator. Ett sådant program är NuWro. Programmet kan användas för att simulera neutrinointeraktioner utifrån en modell som kan jämföras med faktisk experimentell data. En utmaning är dock att modellerna som används i sådana simuleringar behöver vara både tillräckligt noggranna men samtidigt snabba nog för att kunna användas med de Monte Carlo-metoder som NuWro bygger på.

Det finns avancerade beräkningsmetoder för att få en mer detaljerad beskrivning av hur atomkärnan påverkas vid neutrinointeraktion. Ett forskarlag på Chalmers har utvecklat en sådan modell baserad på ab initio-principer [2]. Denna modell är dock beräkningstung och därav tidskrävande, vilket gör den orealistisk att implementera direkt i NuWro.

Modellen är rentav så krävande att den bara går att evaluera i ett begränsat antal kinematiska punkter. För att kunna använda den i ett simuleringsprogram behöver mellanliggande värden approximeras på ett effektivt och tillförlitligt sätt. Modellen måste alltså interpoleras.

I detta arbete undersöks om neurala nätverk kan användas för att utföra denna interpolering efter att ha tränats med ab-initio beräknade punkter. Målet är därmed att skapa en modell som kan beskriva de beräknade kärnresponsfunktionerna även mellan de punkter där data finns tillgängligt. Resultatet av denna modell kommer sedan användas för att möjliggöra beräkningen av syntetisk data i NuWro.

Arbetet har två huvudmål. Det första är att konstruera och utvärdera en maskininlärningsmodell

som kan interpolera mellan kinematiska punkter där beräknade responsfunktioner är tillgängliga. Det andra är att implementera den interpolerade modellen i NuWro och undersöka hur detta påverkar beräknade tvärsnitt för neutrino-kärninteraktioner. Resultaten jämförs både med beräkningar givet en enklare kärnmodell samt med experimentell data.

Arbetet avgränsas till den kvasielastiska delen av neutrino-kärninteraktionen. Det innebär att fokus ligger på processer där neutrinen huvudsakligen växelverkar med en enskild nukleon i kärnan. Det innebär att mer komplicerade processer vid högre energier inte behandlas i detta arbete.

Rapporten är strukturerad på följande sätt. Kapitel 2 ger den teoretiska bakgrunden som sätter arbetet i rätt sammanhang från ett fysikperspektiv. Där behandlas neutriner, neutrino-kärnspridning och tvärsnitt. I kapitel 3 beskrivs de metoder och verktyg som används i arbetet, vilket inkluderar Monte Carlo-metoder, händelsegeneratorn NuWro och neurala nätverk som interpolationsmetod. Kapitel 4 redovisar arbetets genomförande och resultat. Där beskrivs hur den neurala interpolationsmetoden utvecklades och utvärderades, hur de resulterande responsfunktionerna implementerades i NuWro och hur beräknade tvärsnitt jämfördes med andra modellförutsägelser och med experimentell data. Arbetet avslutas med kapitel 5 som diskuterar de framtagna resultaten, arbetets felkällor och möjliga fortsatta studier.

# 2

## Bakgrund

Detta kapitel syftar till att introducera viktiga fysikaliska begrepp som är relevanta för studien samt motivera varför implementeringen av ab-initio beräkningar i NuWro är av intresse.

### 2.1 Neutriner och neutrinooscillation

Neutriner tillhör gruppen leptoner, som är elementarpartiklar med spinn  $\frac{1}{2}$ . Leptoner delas in i laddade partiklar; elektronen  $e$ , myonen  $\mu$  och tauonen  $\tau$ , samt neutrala partiklar; elektron-neutrinon  $\nu_e$ , myonneutrinon  $\nu_\mu$  och tauneutrinon  $\nu_\tau$ . Dessa tre typer av neutriner kallas neutrinosmaker, vilket är ett centralt begrepp för att förstå neutrinooscillationer. Neutriner och laddade leptoner delas även in i tre så kallade generationer där varje neutrino paras ihop med motsvarande laddade lepton [1].

Neutrinooscillationer är ett kvantmekaniskt fenomen där en neutrino kan ändra *smak* ( $\nu_e$ ,  $\nu_\mu$ ,  $\nu_\tau$ ) under sin utbredning. Neutriner produceras och detekteras som smaktillstånd via den svaga växelverkan, medan deras fria utbredning naturligt beskrivs i en väldefinierad bas av masstillstånd. Eftersom smaktillstånd och masstillstånd inte sammanfaller kan en producerad neutrino representeras som en superposition av flera masstillstånd. Dessa komponenter ackumulerar olika faser när neutrinen färdas, och interferensen mellan dem gör att sannolikheten att observera

en viss smak varierar med neutrinoenergin  $E$  och färdsträckan  $L$ . Oscillationer utgör därmed ett centralt verktyg för att undersöka neutrinoegenskaper och implicerar att neutriner inte kan vara helt masslösa (minst en masskillnad måste vara skild från noll) [1].

Fenomenet att neutriners smaktillstånd kan uttryckas som en linjärkombination av tillstånd med bestämd massa kallas neutrinomixning [1]. I ekvation 2.1 kopplas smaktillstånden  $\nu_e, \nu_\mu, \nu_\tau$  ihop med masstillstånden  $\nu_1, \nu_2, \nu_3$ . Matrisen i ekvation 2.1 är den så kallade unitära Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrisen. Elementen i PMNS matrisen beskriver relationen mellan neutriners masstillstånd och smaktillstånd och denna neutrinomixning ger upphov till neutrinooscillation. Matriselementen är således parametrar som beskriver neutrinooscillationernas egenskaper [3].

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu1} & U_{\mu2} & U_{\mu3} \\ U_{\tau1} & U_{\tau2} & U_{\tau3} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \quad (2.1)$$

Som exempel betraktas en elektronneutrino. Tidsevolutionen av dess vågfunktion ges i ekvation 2.2 där  $\phi_i = E_i t - \mathbf{p}_i \cdot \mathbf{x}$  är fasen associerad med planvågen som beskriver utbredningen av masstillstånd  $\nu_i$ .

$$\begin{aligned} |\psi(x, t)\rangle = & \left( U_{e1}^* U_{e1} e^{-i\phi_1} + U_{e2}^* U_{e2} e^{-i\phi_2} + U_{e3}^* U_{e3} e^{-i\phi_3} \right) |\nu_e\rangle \\ & + \left( U_{e1}^* U_{\mu1} e^{-i\phi_1} + U_{e2}^* U_{\mu2} e^{-i\phi_2} + U_{e3}^* U_{\mu3} e^{-i\phi_3} \right) |\nu_\mu\rangle \\ & + \left( U_{e1}^* U_{\tau1} e^{-i\phi_1} + U_{e2}^* U_{\tau2} e^{-i\phi_2} + U_{e3}^* U_{\tau3} e^{-i\phi_3} \right) |\nu_\tau\rangle \end{aligned} \quad (2.2)$$

Sannolikheten att en neutrinos smaktillstånd ändras från  $\nu_e$  till  $\nu_\mu$  ges av  $P(\nu_e \rightarrow \nu_\mu) = |\langle \nu_\mu | \psi(x, t) \rangle|^2 = |U_{e1}^* U_{\mu1} e^{-i\phi_1} + U_{e2}^* U_{\mu2} e^{-i\phi_2} + U_{e3}^* U_{\mu3} e^{-i\phi_3}|^2$  och eftersom PMNS matrisen är unitär ( $UU^\dagger = I$ ) kommer sannolikheten att bli noll om det går att faktorisera ut fasfaktorerna. Detta sker då alla fasfaktorer är lika. Samma resonemang kan appliceras på sannolikheten  $P(\nu_e \rightarrow \nu_\tau)$ . Faserna är olika om och endast om minst två massor till massegentillstånden skiljer sig och det följer att neutrinooscillation endast kan ske då alla tre massor inte är identiska. Detta innebär också att neutriner inte kan vara masslösa [3]. I partikelfysikens standardmodell som etablerades under 1970-talet antas neutriner vara masslösa och upptäckten att neutriner inte är masslösa ger information bortom standardmodellen.

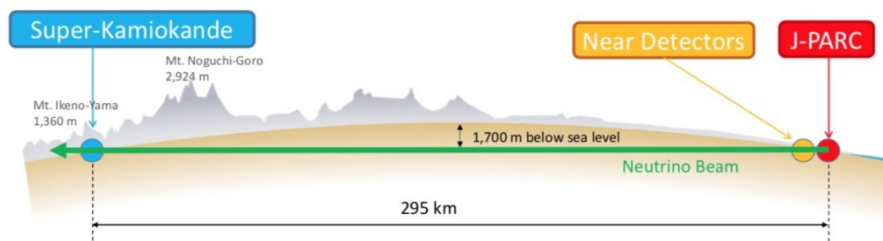
I dagsläget är de absoluta massorna hos masstillstånden inte fastställda men information om masskvadratskillnaderna har erhållits via neutrinooscillationsexperiment [4]. Huruvida  $m_3 > m_2$  eller  $m_3 < m_2$  har ännu inte fastslagits och det finns alltså två möjliga masshierarkier. Det är möjligt att massorna har en påverkan på universums utveckling [3] och studier som ökar förståelsen för neutrino-oscillation kan bidra till kunskap om neutrinernas massor och deras betydelse för fysiken som beskriver universum.

En annan oscillationsparameter av central betydelse för att förstå universum är CP-fasen. I universum finns mer materia än antimateria och bruten laddningsparitetssymmetri (CP-brott) är ett nödvändigt villkor för att förklara denna observation [3]. Paritet är en egenskap som beskriver hur ett fysikaliskt tillstånd förändras vid rumslik spegling [1]. Operationen då en partikel byts mot sin antipartikel kallas laddningssymmetri [1]. CP-brott i den svaga sektorn uppstår då PMNS-matrisen innehåller en komplex fas [4] och innebär att paritet och laddningskonjugation inte bevaras [1].

Kunskap om parametrar som beskriver fenomenet neutrinooscillation ger alltså information om universums utveckling och egenskaper. Osäkerheter i neutrino kärntvårsnitt leder till systematiska osäkerheter i neutrinooscillationsparametrar och ett mål inom partikelfysik är att minska osäkerheten hos parametrar som beskriver neutrinooscillation [5]. Implementering av ab initio-beräkningarna i NuWro kan potentiellt minska osäkerheten i tvärsnittet jämfört med nuvarande teoretiska modeller.

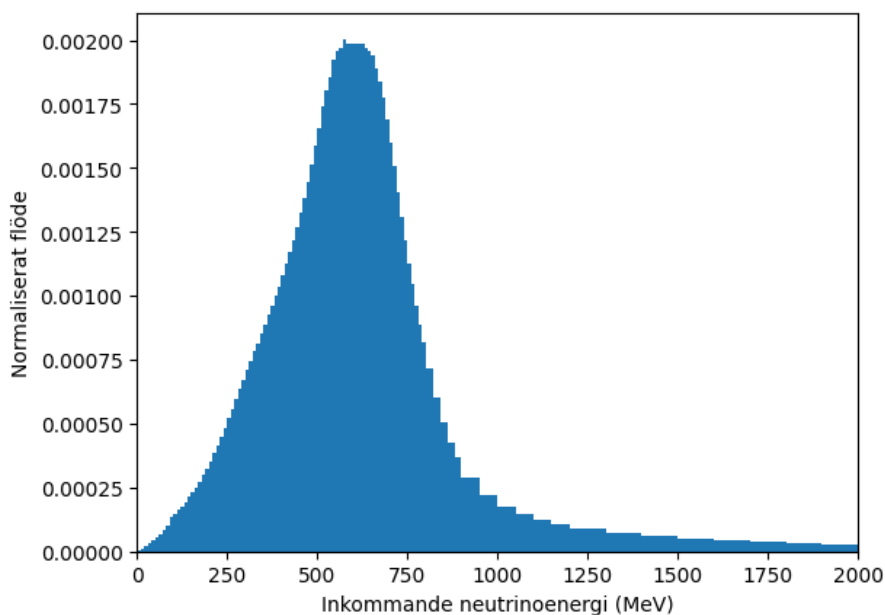
## 2.2 T2K-experimentet

För att undersöka neutrinooscillation har bland annat T2K-experimentet (Tokai to Kamioka) konstruerats. Experimentet utfördes i Japan och består av en strålkälla som först passerar genom en närbelägen detektor (på sådant avstånd att oscillationen är obetydlig). Strålen färdas därefter genom marken till en detektor belägen nästan 30 mil bort, så att oscillationen kan observeras genom den uppmätta skillnaden mellan den närbelägna och den bortre detektorn [6]. Experimentuppställningen visas i figur 2.1.



**Figur 2.1:** Experimentuppställning för T2K-experimentet. Figur hämtad från [6]

Strålen som genereras är inte monoenergetisk; den består av en fördelning av olika energier centrerat kring 600 MeV. Detta visas i figur 2.2. Förutom den breda toppen har distributionen en lång svans som sträcker sig uppåt 5 GeV [7]. Neutrinooscillation beror av både energi och avstånd. Oscillation sker över kortare avstånd för neutriner med låg energi jämfört med neutriner med högre energi [6].



**Figur 2.2:** Neutrinostrålens energifördelning i T2K-experimentet.

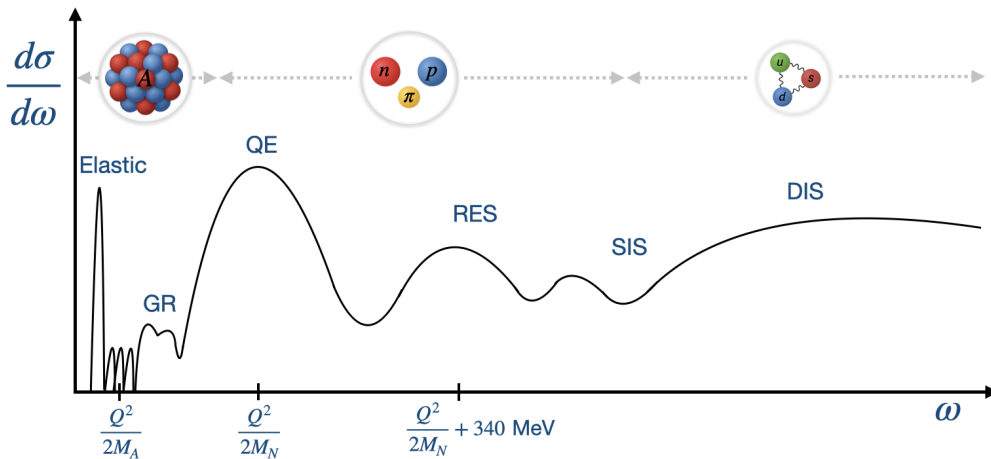
## 2.3 Neutrino-kärnspridning

Neutrino-kärnspridningsexperiment och därtill relaterade händelsegeneratorer kräver förståelse för hur inkommande neutriner växelverkar med en atomkärna vid olika energier. Neutrino-kärninteraktioner delas in i kategorierna laddad ström (CC) och neutral ström (NC) beroende på om den utgående partikeln är en laddad lepton eller en neutrino. Vid CC gäller att utgående laddad lepton tillhör samma generation som inkommande neutrino och vid NC gäller att utgående neutrino har samma smak som inkommande neutrino [8].

Neutrino-kärnspridning sker genom ett antal olika interaktioner. Vilken process som är mest betydande beror främst på den inkommande neutrinons energi ( $\epsilon$ ) samt hur mycket energi ( $\omega$ ) och rörelsemängd ( $q$ ) som förs över till kärnan. Dessa storheter avgör neutrino-interaktionens upplösning, det vill säga hur detaljerat atomkärnans inre struktur kan sonderas. I figur 2.3 visualiseras vid vilka energiöverföringar olika spridningsmekanismer dominerar. Vid mycket liten rörelsemängdsöverföring kan spridningen vara elastisk och koherent. Neutrinen växelverkar då med kärnan som en helhet, och kärnan förblir i sitt initialtillstånd (bortsett från en liten rekyl). När energi- och rörelsemängdsöverföringen ökar kan spridningen bli inelastisk, så att kärnan hamnar i ett exciterat tillstånd [9].

Så kallade kvasielastisk spridning dominerar i ett stort kinematiskt område. Denna process innebär att neutrinen approximativt växelverkar med en enskild nukleon i kärnan. I det kvasielastiska området är dock kärnfysiken fortfarande avgörande, nukleonerna är bundna i kärnan och kan dessutom vara korrelerade med varandra, vilket påverkar både de möjliga sluttillstånden och hur energi och rörelsemängd fördelas i processen. Därför beskrivs ofta växelverkan med kärnresponsfunktioner. De sammanfattar hur kärnan reagerar på den svaga växelverkan och fungerar som en länk mellan den del av händelsen som rör den inkommande och den utgående partikeln [9].

Att den inkommande partikeln vid kvasielastisk spridning växelverkar med en bunden nukleon kan resultera i att nukleonerna i fråga lämnar sitt bundna tillstånd. Det är den processen som står för huvuddelen av utfallen vid energier upp till 1 GeV. Det är vanligt förekommande att neutriners energier i experiment ligger i det område där kvasielastisk spridning dominerar [10]. Vid ännu större energiöverföring kan neutrinon i stället skapa pioner, och vid tillräckligt hög upplösning blir den avgörande mekanismen djupt inelastisk spridning där neutrinen sonderar nukleonens inre struktur i form av kvarkar och gluoner [9].



**Figur 2.3:** Differentiellt tvärsnitt som funktion av energiöverföring. Figur från [9].

Kärnmodeller som beskriver den kärnfysik som påverkar neutrino-kärnväxelvekan implementeras i neutrinohändelsegeneratorer. Flera moderna neutrinohändelsegeneratorer, däribland NuWro, antar att neutrinen träffar en enskild nukleon vid växelverkan med kärnan. En kärnmodell appliceras för att tilldela denna nukleon en initial rörelsemängd samt en bindningsenergi [11]. Vanligt förekommande val av kärnmodell är olika varianter av fermigasmodellen.

Paulis uteslutningsprincip uttrycker att fler än en fermion inte kan ockupera samma kvanttillstånd [9]. I neutrinohändelsegeneratorer kan Pauli-blockering tillämpas vid simulering av neutrino-kärninteraktioner. Detta innebär att övergångar till redan ockuperade tillstånd förbjuds. Vid kvasielastisk spridning förbjuder Pauli-blockering alla händelser där den utslagna nukleonen får lägre rörelsemängd än  $p_F$ . Effekten minskar antalet möjliga sluttillstånd och vid låga rörelsemängdsöverföringar är det fler nukleoner som blockeras. När en nukleon exciteras från ett tillstånd med rörelsemängd mindre än  $p_F$  till ett tillstånd med rörelsemängd större än  $p_F$  uppstår ett hål i den tidigare fyllda potentialbrunnen. Denna typ av excitation kallas en partikel-hål-excitation (1p-1h) [9]. En approximativ metod som används för att beskriva kärnexcitationer är random phase approximation (RPA) som beskriver dessa excitationer som linjära superpositioner av 1p-1h excitationer, där mer komplexa konfigurationer försummas [12].

## Fermigasmodellen

Fermigasmodellen är en modell som behandlar neutroner och protoner i atomkärnan som två oberoende system vilka var för sig uppfyller Paulis uteslutningsprincip. Nukleonerna antas röra sig fritt inom en begränsad volym som representeras av en ändlig potentialbrunn. I modellen betraktas en neutronpotential och en protonpotential där respektive potential är en superposition av enskilda nukleoners bidrag [1]. Eftersom två identiska nukleoner inte kan ockupera samma tillstånd samtidigt kommer tillstånden i potentialbrunnen att fyllas på successivt från det med lägst energi upp till det med högst energi. Nukleonerna antas fylla alla möjliga (diskreta) rörelsemängdstillstånd upp till den så kallade fermirörelsemängden  $p_F$  [9].

Fermigasmodellen har begränsningar vilka blir betydliga framförallt vid lägre energinivåer. Vid låga energiöverföringar får excitationer av kärnan större betydelse. Detta är en effekt som inte beskrivs väl av fermigasmodellen [13]. Vid energier så låga som endast ett fåtal MeV domineras växelverkan av övergångar till diskreta låga energinivåer i kärnan [11]. Vid energier som är något högre, men fortfarande relativt låga förekommer kollektiva excitationer, så kallad resonans då flera nukleoner rör sig tillsammans. Detta fenomen fångas inte heller väl av fermigasmodellen [11]. Dessa effekter avtar med energi men excitation till låga energinivåer förväntas vid framåtspridning ha en märkbar påverkan även vid energier runt 1 GeV [11]. Detta motiverar behovet av mer avancerade ab initio kärnmodeller vid lägre energinivåer för att bättre fånga dessa effekter. Vid högre energier blir det relativa bidraget från excitationer till låga energinivåer och kollektiva excitationer lägre vilket leder till att beräkningar baserade på fermigasmodeller blir mer tillförlitliga.

I NuWro erbjuds den lokala fermigasmodellen (LFG) som ett möjligt val av kärnmodell. LFG är en utveckling av fermigasmodellen där kärnans densitet antas bero på avståndet till kärnans centrum. Istället för att behandla kärnan som homogen antas Fermirörelsemängden vara positionsberoende [14]. Detta innebär att nukleoners maximala rörelsemängd beror på positionen i kärnan.

## Ab initio

Ab initio-modeller är betydligt mer avancerade modeller konstruerade från fysikens grundläggande principer. Den nya modellen som användes i projektet behandlar kärnan som ett mångkroppssystem och utgår från den starka växelverkan mellan nukleoner för att beskriva kärnans struktur och dynamik. Kärnans energi modelleras med en hamiltonian som inkluderar nukleonernas kinetiska energi samt två- och trenukleoninteraktioner. Ab initio-beräkningar syftar till att numeriskt lösa kärnans mångkroppsproblem, det vill säga Schrödingerekvationen där för ett system av växelverkande nukleoner. En fördel med ab initio-beräkningar är att det är möjligt att kvantifiera den teoretiska osäkerheten, något som är av största vikt när teoretiska resultat jämförs med experimentella [9].

## Tvärsnitt

Inom subatomär fysik är tvärsnitt en effektiv area vilken kvantifierar sannolikheten att en viss reaktion ska ske då en inkommande partikel skickas mot en målpartikel. Tvärsnitt anges normalt i enheten barn (b) där en barn motsvarar  $10^{-24}$  kvadratcentimeter [15]. När det gäller neutrino-kärn-reaktioner kvantifierar det totala tvärsnittet  $\sigma$  sannolikheten att en inkommande neutrino ska växelverka med en atomkärna [1].

Det förekommer även differentiella tvärsnitt samt dubbeldifferentiella tvärsnitt. Dessa beskriver tvärsnittets beroende på olika differentiella variabler, exempelvis  $\frac{d\sigma}{d\Omega}$  som kvantifierar sannolikheten att en utgående partikel observeras inom rymdvinkeln  $\Omega$  [9]. Integration av det differentiella tvärsnittet ger det totala tvärsnittet  $\sigma = \int \frac{d\sigma}{d\Omega} d\Omega$  där  $d\Omega = d\cos\theta d\phi$  [3].

Det inklusiva dubbeldifferentiella tvärsnittet  $\left. \frac{d\sigma}{d\cos\theta d\epsilon'} \right|_{\nu/\bar{\nu}}$  för neutrino  $\nu$  eller antineutrino  $\bar{\nu}$  beräknas enligt

$$\left. \frac{d\sigma}{d\cos(\theta)d\epsilon'} \right|_{\nu/\bar{\nu}} = \frac{G^2}{8\pi^2} \frac{k^2}{\epsilon} [v_{00}R_{00} + v_{zz}R_{zz} - v_{0z}R_{0z} + v_{xx}R_{xx} \mp v_{xy}R_{xy}], \quad (2.3)$$

där Fermikonstanten  $G = 1.1663787 \times 10^{-11} \text{MeV}^{-2}$ ,  $\epsilon$  är ingående partikelns energi och  $\epsilon'$  utgående,  $k$  är storlek av den ingående partikelns trerörelsemängdsvektor,  $R_{00}, R_{zz}, R_{0z}, R_{xx}, R_{xy}$  responsfunktioner och  $v_{00}, v_{zz}, v_{0z}, v_{xx}, v_{xy}$  leptoniska kinematiska faktorer [16]. Att dubbeldifferentiella tvärsnittet är inklusivt innebär att det summeras över de observerbara tillstånd som det inte differentieras med avseende på. De kinematiska leptoniska faktorerna beräknas enligt

$$v_{00} = 2\epsilon\epsilon' \left( 1 + \frac{k'}{\epsilon'} \cos(\theta) \right), \quad (2.4)$$

$$v_{zz} = \frac{\omega^2}{q^2} (m_l^2 + v_{00}) + \frac{m_l^2}{q^2} [m_l^2 + 2\omega(\epsilon + \epsilon') + q^2], \quad (2.5)$$

$$v_{0z} = \frac{\omega}{q} (m_l^2 + v_{00}) + m_l^2 \frac{\epsilon + \epsilon'}{q}, \quad (2.6)$$

$$v_{xx} = Q^2 + \frac{Q^2}{2q^2} (m_l^2 + v_{00}) - \frac{m_l^2}{q^2} \left[ \frac{m_l^2}{2} + \omega(\epsilon + \epsilon') \right], \quad (2.7)$$

$$v_{xy} = Q^2 \frac{\epsilon + \epsilon'}{q} - m_l^2 \frac{\omega}{q}, \quad (2.8)$$

där  $k'$  är storlek av den utgående partikelns trerörelsemängdsvektor,  $q$  är beloppet av den tredimensionella rörelsemängdsöverföringen,  $\omega$  är energiöverföringen,  $m_l$  massan hos utgående

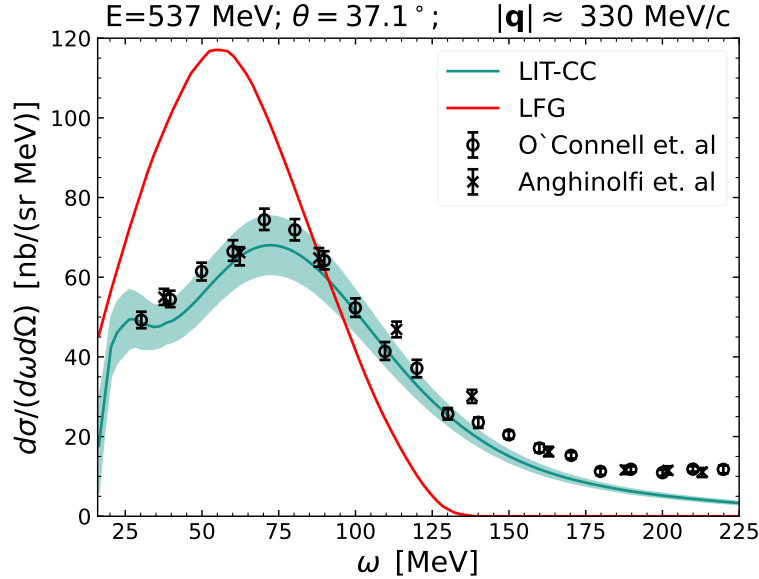
lepton och  $Q^2 = q^2 - \omega^2$  [17]. Ytterligare förhåller sig  $\cos(\theta)$  och  $k, k', q$  enligt

$$\begin{aligned} q &= |\vec{k} - \vec{k}'| \\ q^2 &= k^2 + k'^2 - 2kk' \cos(\theta) \\ \cos(\theta) &= \frac{k^2 + k'^2 - q^2}{2kk'}. \end{aligned} \quad (2.9)$$

Dubbelintegration av det differentiella tvärsnittet med avseende på cosinus av spridningsvinkeln och den utgående energin ger det totala tvärsnittet.

Det inklusiva dubbeldifferentiella tvärsnittet kan delas upp i faktorer som beskriver inkommande och utgående leptons kinematik samt faktorer som beskriver kärnans respons. Kärnresponsfunktioner  $R_{ij}$  (NRF:er, *Nuclear Response Functions*) beskriver hur kärnan svarar på svaga växelverkan vid neutrino-kärnspridning och det är vid beräkningar av dessa funktioner som kärnmodeller används. Det är alltså vid beräkningen av NRF:er som ab initio-metoden implementeras [9].

I figur 2.4 visas det dubbeldifferentiella tvärsnittet  $\frac{d\sigma}{d\omega d\Omega}$  för elektron-kärnspridning vid energinivåer lägre än 225 MeV. Den röda grafen är beräknad med responsfunktioner framtagna med LFG-modellen, den turkosa grafen är framtagen med ab initio-beräkningar och de svarta datapunkterna representerar experimentell mätdata. Det kan observeras att ab initio resultatet stämmer bättre överens med observerad data än LFG resultatet. Även om jämförelsen gäller elektron-kärnspridning kan den, eftersom fokus på kärnmodellerna snarare än den inkommande partikeln, ge insikter även för neutrino-kärnspridning.



**Figur 2.4:** Dubbeldifferentiellt tvärsnitt för elektron-kärnspridning. LIT-CC kurvan är framtagen från ab initio-beräkningar, de svarta datapunkterna representerar experimentell data och LFG-kurvan är beräknad utifrån LFG-modellen. Ab initio-beräkningarna är baserade på och den experimentella datan är hämtad från [16]. Det dubbeldifferentiella tvärsnittet är beräknat enligt teori i [17].

# 3

## Metodologi

En utmaning i neutrinoexperiment är att varken den inkommande neutrinos energi eller detaljer vid efterföljande växelverkan med kärnan är känd [8]. Detta skapar ett behov av neutrino-händelsegeneratorer som utifrån experimentella observationer kan generera en sannolikhetsfördelning för växelverkans kinematik [8].

### 3.1 Monte Carlo-metoder

Monte Carlo-metoder (MC) är en familj av metoder där probabilistiska tillvägagångssätt används för att uppskatta olika storheter. De har fördelen att de är generella och fungerar på komplicerade, högdimensionella och icke-analytiska funktioner med relativt svaga krav på funktionens egenskaper [18]. I detta projekt används MC-sampling och MC-integration.

Problemet MC-sampling löser är att göra slumpdragningar ur en täthetsfunktion  $f(x)$  med parameter  $x \in \mathbf{A}$ , där  $\mathbf{A}$  är parameterrymden belägen i  $\mathbf{R}^n$  med volym  $A$ . Ett utfall  $x$  föreslås, och sannolikheten för att det utfallet skulle ha skett utvärderas utifrån täthetsfunktionen. Denna sannolikhet avgör probabilistiskt om den dragna händelsen godtas eller förkastas. Processen upprepas till dess att det önskade antalet händelser godtagits. Dessa händelser kommer då vara distribuerade enligt  $f$ .

Förutom att sampla en fördelning kan MC-metoder användas för att integrera komplicerade, icke-analytiska funktioner. Vi antar en funktion  $g(x)$ , som till skillnad från  $f$  inte måste vara normerad och icke-negativ. Istället för en sannolikhet beräknas för varje händelse en vikt  $w$ , som bland annat beror på funktionens värde i den dragna punkten.

Värdet för en funktion över integrationsområdet  $\mathbf{A}$  beräknas analytiskt enligt

$$G = \int_{\mathbf{A}} g(x) dx. \quad (3.1)$$

En för de flesta funktioner dålig approximation  $\hat{G}$  är att välja en punkt  $x$  och anta att  $g$  är konstant på hela integrationsspannet.

$$\hat{G} = g(x) \cdot A = w(x) \quad (3.2)$$

För att förbättra approximationen  $\hat{G}$  kan många slumpmässigt valda punkter  $x_n$  snittas över

$$\hat{G} = \frac{1}{N} \sum_{n=1}^N g(x_n) A = \frac{1}{N} \sum_{n=1}^N w(x_n). \quad (3.3)$$

Som kan ses är skälet att en vikt används istället för det "råa" värdet givet av  $g$  att kompensera för sannolikheten att samplen som valdes blev vald. Detta avgörs av fördelningen i  $\mathbf{A}$  för slumpdragningen av parametern  $x$ . För fallet med en uniform fördelning måste utfallets vikt skalas med storleken av  $A$ : om gränserna dubblas kommer hälften så många händelser genereras i ett visst spann. Genom att då dubbla vikten behålls samma totala vikt i detta spann med samma totala antal dragningar.

Vissa områden av  $\mathbf{A}$  kan innehålla mer detalj och ha större inverkan på integralen. Exempelvis erhålls ingen information från att göra dragningar i ett område där funktionen är lika med noll. Förutom att använda snävare integrationsgränser kan en transformation appliceras på dragningsfördelningen för att ge en större andel utfall i områdena av störst vikt, vilket gör samplingen mer effektiv. För att kompensera den förändrade dragningsfördelningen behöver jacobianen  $J(x)$  av transformationen beräknas och appliceras på vikten. Samma metod kan även användas för att göra dragningar i en annan parametrisering än den funktionen är uttryckt i.

Vid MC-integration är standardavvikelsen för felet  $G - \hat{G}$  proportionell mot  $1/\sqrt{N}$ , så att approximationen blir allt mer exakt med fler dragningar. Detta är inte särskilt bra jämfört med till exempel Simpsons regel som konvergerar med  $1/N^4$  - men den antar kontinuerliga, endimensionella funktioner. Konvergensen för MC-integration gäller även för diskontinuerliga och högdimensionella funktioner [18].

I de fall en funktion både ska integreras och göras dragningar ur kan först MC-integrationen göras enligt ovan. Därefter används vikterna som sannolikheter genom att normera dem mot den största observerade vikten, så att inga sannolikheter är  $> 1$ . Detta kräver dock till exempel att funktionen som undersöks är icke-negativ. Det är av intresse att sannolikheterna som används för acceptans/förkastning använder hela spannet  $[0, 1]$ , eftersom onödigt låga sannolikheter leder till att många fler dragningar måste göras för att uppnå det önskade antalet godtaga händelser.

## 3.2 Händelsegeneratoren NuWro

NuWro är en Monte Carlo-händelsegenerator med öppen källkod skriven i C++, utvecklad vid Wrocławsk universitet [19]. NuWro kompileras med beroende på programbiblioteket PYTHIA 6 samt ROOT. PYTHIA6 innehåller datastrukturer, modeller och teori för arbete med elementarpartiklar och högenergikrockar mellan dem [20]. ROOT är ett av CERN utvecklat ramverk för datahantering.

NuWro kan användas för att simulera neutrinointeraktioner med två olika typer av mål, neutronkärnväxelverkan samt neutronukleonväxelverkan [19]. Olika kärnmodeller kan väljas i NuWro, däribland lokala fermigasmodellen vilken används i detta arbete. Med hjälp av NuWro är det möjligt att utvärdera hur mätbara storheter såsom tvärsnitt påverkas av olika teoretiska implementationer [19]. I NuWro finns olika dynamiska modeller implementerade, däribland kvasielastisk spridning för laddad ström [19]. Processen att generera händelser med NuWro inleds med att välja vilka interaktionsmekanismer som ska användas. Eftersom olika mekanismer är olika betydliga vid olika delar av parameterrymden är det användbart att kunna kombinera flera modeller för att få en komplett bild av interaktionerna.

För att göra detta används MC-sampling. De olika modellerna är uttryckta på olika former, bland annat som rena matematiska funktioner, förberäknade tabellvärden eller andra mer procedurella metoder. Dessa kombineras med för- och efterbearbetningssteg i form av kompensationsfaktorer

och filtrering av fysikaliskt otillåtna utfall. Med analytiska metoder hade detta krävt att hela processen analyseras för varje parameteruppsättning. MC fungerar istället som ett universellt gränssnitt för att sammanföra hela beräkningskedjan, där varje steg har möjlighet att påverka resultatet genom att justera händelsens vikt. Detta gör systemet flexibelt och tillåter det att utökas enkelt.

NuWro arbetar i en test- och genereringsfas. I testfasen samlas information in om de olika interaktionsmekanikerna så som deras relativa tvärsnitt och maximala vikt. Detta används sedan i genereringsfasen för att tilldela antalet händelser per mekanism. Den maximala vikten används för att normalisera de genererade vikterna, för att erhålla sannolikheter. Genom detta behöver inte modellerna alstra normerade sannolikhetsfördelningar, vilket hade varit svårt att uppehålla då flera justeringar kedjas efter varandra.

För att generera en händelse slumpas först en punkt i parameterrymden ("phase space"). Tvärsnittet beräknas för denna punkt och justeras enligt ovan för dragningsfördelningen. För att öka effektiviteten bör parameterrymdens gränser vara så snäva som möjligt så att inte punkter där tvärsnittet är 0 samplas, då dessa händelser alltid kommer slängas och ger då ingen information.

### 3.3 Neurala nätverk som interpolationsmetod

Interpolering är metoder för att approximera en funktion mellan kända diskreta datapunkter [21]. Några av dessa är linjär interpolering, bilinjär interpolering samt splineinterpolering. Vid linjär interpolering antas det okända värdet mellan två kända datapunkter ligga på en rät linje som skär de två kända punkterna [21]. Bilinjär interpolering kan användas för approximering av en punkt som beror av två variabler och innebär linjär interpolering i först den ena riktningen och därefter i den andra.

I detta arbete används ett neuralt nätverk för att interpolera ab initio-beräknade responsfunktioner. Eftersom ab initio-beräkningarna endast finns tillgängliga för ett begränsat antal diskreta värden på rörelsemängdsöverföringen behövs en metod som kan uppskatta responsfunktionerna även mellan dessa punkter. Det neurala nätverket används därför som en icke-linjär interpolator, vilket är relevant eftersom responsfunktionerna kan ha snabbt varierande struktur kring den kvasielastiska toppen.

Det neurala nätverk som används i detta arbete är ett djupt framåtriktat nätverk, även kallat multilayer perceptron (MLP) [22], [23]. Nätverket består av ett inmatningslager, ett antal dolda lager samt ett utmatningslager. Inmatningslagret tar emot de variabler som beskriver punkter där responsfunktionerna ska beräknas, exempelvis  $q$  och  $\omega$ , eller transformerade varianter av dessa. Utmatningslagret ger nätverkets prediktion av responsfunktionerna. I detta arbete kan nätverket betraktas som en funktion  $\hat{\mathbf{R}} = f_{\theta}(q, \omega)$ , där  $\hat{\mathbf{R}}$  är nätverkets predikterade responsvektor och  $\theta$  betecknar nätverkets träningsbara parametrar. Dessa parametrar består framför allt av vikter och biastermer i nätverkets lager [22], [23].

I varje lager utförs en transformation av föregående lagers utdata. Matematiskt kan detta skrivas som

$$h^{(l)} = \sigma(W^{(l)}h^{(l-1)} + b^{(l)}), \quad (3.4)$$

där  $h^{(l)}$  är utdata från lager  $l$ ,  $W^{(l)}$  och  $b^{(l)}$  är lagrets vikter respektive bias, och  $\sigma$  är en

aktiveringsfunktion [22], [23]. Först appliceras alltså en affin transformation, följt av en icke-linjär aktiveringsfunktion. Denna icke-linearitet är avgörande för att nätverket ska kunna representera komplexa samband i data [22].

Under träningsprocessen justeras nätverkets parametrar  $\theta$  för att minimera skillnaden mellan nätverkets prediktion och de kända responsvärdena. För varje träningspunkt jämförs nätverkets prediktion  $\hat{R}_i$  med den givna responsen  $R_i$ . Skillnaden mellan dessa kvantifieras med hjälp av en förlustfunktion, som anger hur väl modellen beskriver träningsdata. En vanlig förlustfunktion för regressionsproblem är *mean squared error* (MSE). För ett skalärt mål, eller för en enskild responskomponent, kan den skrivas som

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2, \quad (3.5)$$

där  $N$  är antalet datapunkter,  $\hat{y}_i$  är modellens prediktion och  $y_i$  är det sanna värdet. MSE bestraffar stora avvikelser relativt hårt eftersom felet kvadreras [22], [23].

Eftersom valet av förlustfunktion styr vilka egenskaper modellen prioriterar under träningen kan den anpassas till det aktuella fysikaliska problemet. Detta kan exempelvis göras genom att inkludera fysikalisk modellkunskap i förlustfunktionen [23]. I detta arbete är det särskilt viktigt att modellen beskriver fysikaliskt relevanta områden väl, exempelvis området kring den kvasielastiska toppen. Detta motiverar användningen av en viktad förlustfunktion, där fel i sådana områden ges större betydelse.

Som komplement till MSE kan modellens noggrannhet även beskrivas med *mean absolute error* (MAE), där absolutbeloppet av skillnaden mellan prediktion och sant värde används. MAE kan användas både som förlustfunktion vid träning och som utvärderingsmått vid jämförelse mellan modeller. Eftersom MAE anger den genomsnittliga absoluta avvikelsen i samma skala som responsfunktionerna själva är måttet särskilt användbart för att tolka modellens prediktionsfel [22], [23].

För att minimera förlustfunktionen används en optimeringsalgoritm, exempelvis Adam. Optimeringsalgoritmens uppgift är att justera modellens parametrar  $\theta$  så att förlustfunktionen minskar. Detta görs genom gradientbaserad optimering, där gradienten av förlustfunktionen med avseende på nätverkets parametrar beräknas med hjälp av felbakåtspridning. Parametrarna uppdateras därefter iterativt i den riktning som minskar förlusten [22], [23]. Hur stora uppdateringar som görs styrs bland annat av inlärningshastigheten. Inlärningshastigheten kan vara konstant under träningen, men kan också ändras enligt ett schema, exempelvis genom att minska när träningen fortskrider. Valet av inlärningshastighet och schema påverkar därför både hur snabbt modellen tränas och hur stabil optimeringen blir [22].

En viktig del av träningen är att modellen inte bara ska beskriva träningsdata väl, utan också ge tillförlitliga prediktioner för punkter som inte ingått i träningen. Detta är särskilt viktigt i detta arbete, eftersom syftet är att interpolera responsfunktionerna mellan de diskreta punkter där ab initio-beräkningar finns tillgängliga. Två vanliga problem under träningen av neurala nätverk är underanpassning och överanpassning. Underanpassning innebär att modellen är för enkel, eller inte har tränats tillräckligt väl, för att fånga det underliggande sambandet i data. Detta leder ofta till högt fel både på träningsdata och på validerings- eller testdata. Överanpassning innebär i stället att modellen beskriver träningsdata mycket väl, men samtidigt lär sig detaljer eller brus som inte generaliserar till nya datapunkter. Detta visar sig ofta genom att träningsfelet

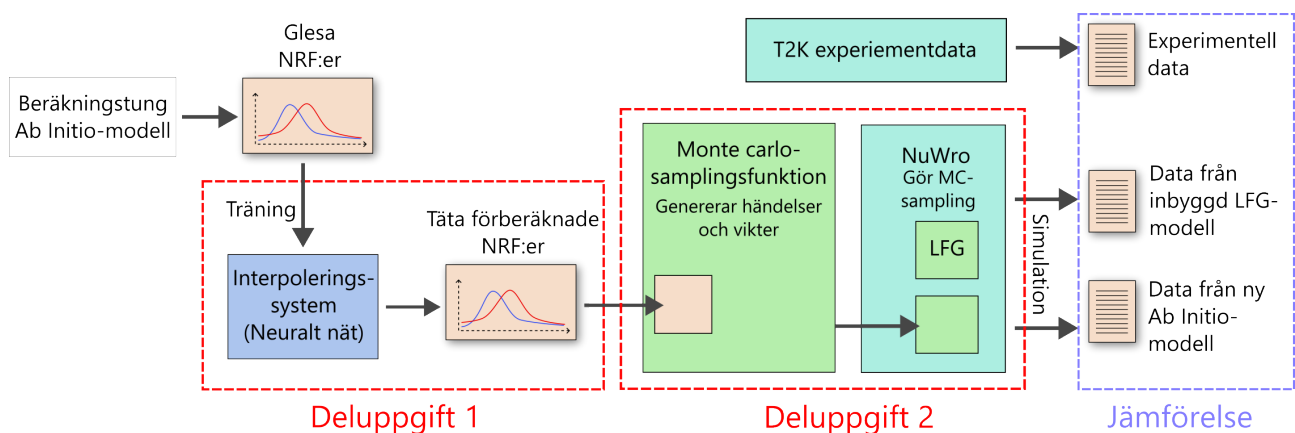
är lågt medan felet på validerings- eller testdata är betydligt högre [22], [23]. Överanpassning är problematiskt eftersom modellen då kan ge missvisande interpolationer mellan de kända beräkningspunkterna. För att bedöma modellens användbarhet är det därför nödvändigt att jämföra modellens fel på träningsdata med dess fel på validerings- eller testdata.

# 4

## Utveckling av neural interpolationsmetod och tillämpning i NuWro

Det finns flera möjliga metoder för att utföra interpoleringen av den högkvalitativa datan. I en inledande fas testades olika interpoleringsmodeller, däribland neurala nätverk, linjär interpolering, interpolering efter bastransformation (ortogonal dekomposition) och koordinattransformation. Interpoleringsmetoden utvärderades genom jämförelse mellan interpolerade kurvor samt ursprungliga kurvor. Efter den utforskande fasen av projektet gjordes valet att gå vidare med interpolering med neuralt nätverk.

Projektet delades därefter upp i två olika delar enligt figur 4.1 vilka bearbetades parallellt. Den ena deluppgiften innefattade att välja en metod för att utvärdera prestandan hos ett neuralt nätverk med syfte att interpolera mellan ab initio beräkningar. Därefter utvärderades val av olika arkitekturer och hyperparametrar systematiskt. Deluppgiften innefattade också att undersöka ifall det var gynnsamt att interpolera med hjälp av två olika nätverk där det ena tränades till att interpolera responsfunktioner beräknade med en enklare modell (så kallade spektralfunktioner). Det andra neurala nätverket skulle sedan interpolera residualerna mellan responsfunktionerna från ab initio-modellen och spektralmodellen. I slutändan togs beslutet att endast använda ett



**Figur 4.1:** Projektets arkitektur och deluppgifter

neuralt nätverk som direkt interpolerade ab initio-responsfunktionerna.

Den andra deluppgiften innefattade att implementera en modell för kvasielastiskt spridning i NuWro baserad på ab initio-kärnresponsfunktionerna. Detta utgjordes av en händelsegenerationsfunktion som NuWro använde för att göra dragningar till dess MC-sampling. För att beräkna en händelses vikt används det differentiella tvärsnittet som berodde på tabellerade NRF:er genererade av interpolatorn i deluppgift 1. Det differentiella tvärsnittet beräknat från av NuWro-genererade händelser jämfördes till sist med experimentell data. Detta gjordes för både den inbyggda LFG-modellen samt den nya ab initio-baserade modellen.

## 4.1 Neurala nätverk

För att identifiera en lämplig modell för interpolering av de ab initio-beräknade responsfunktionerna genomfördes en systematisk undersökning av flera av modellens hyperparametrar. De undersökta hyperparametrarna omfattade nätverksarkitektur, aktiveringsfunktion, basinlärningshastighet, schema för inlärningshastigheten, förlustfunktion, indatavarianter, normalisering och enhetssystem. Målet var att förutsäga de fem responskomponenterna

$$R_{00}, \quad R_t, \quad R_{xy}, \quad R_{zz}, \quad R_{0z}$$

som funktion av  $q$  och  $\omega$ , med hög noggrannhet samtidigt som överanpassning till träningsdatan undveks.

Eftersom de ab initio-beräknade responsfunktionerna har en tillhörande teoretisk osäkerhet anges varje responskurva som ett intervall. För varje responskomponent  $R_\alpha$  motsvarar  $R_\alpha^{\min}(q, \omega)$  den nedre gränsen och  $R_\alpha^{\max}(q, \omega)$  den övre gränsen för detta osäkerhetsband vid givna värden på rörelsemängdsöverföringen  $q$  och energiöverföringen  $\omega$ . Som central målkurva definierades därför medelresponsen

$$R_\alpha^{\text{med}}(q, \omega) = \frac{1}{2} \left( R_\alpha^{\min}(q, \omega) + R_\alpha^{\max}(q, \omega) \right). \quad (4.1)$$

Hyperparametersökningen i detta avsnitt gjordes med  $R_\alpha^{\text{med}}$  som träningsmål.

Den ursprungliga datamängden bestod av åtta ab initio-beräknade responskurvor för  $q$ -värden från 50 till 400 MeV i steg om 50 MeV, där varje kurva var tätt tabellerad i energiöverföringen  $\omega$ . Senare tillkom även en responskurva för  $q = 75$  MeV, vilken användes vid träningen av det slutgiltiga nätverket men inte ingick i hyperparametersökningen.

En central hyperparameter var nätverksarkitekturen, där nätverk med olika bredd och djup testades. Mindre nätverk har färre träningsbara parametrar och är därför mer stabila i fall där mängden träningsdata är begränsad, men samtidigt har de inte förmågan att lära sig mer komplicerade samband i datan. Större nätverk har fler träningsbara parametrar och kan därmed lära sig mer avancerade samband i träningsdatan. Nackdelen är att det finns en större risk för överanpassning med större nätverk. Att testa olika nätverksarkitekturer hjälper till att hitta en balans mellan förmågan att lära sig av datan och överanpassning till datan.

Nätverket tränades genom att ensamt prediktera samtliga fem responskomponenter,

$$R_{00}, \quad R_t, \quad R_{xy}, \quad R_{zz}, \quad R_{0z},$$

för varje indatapunkt  $(q, \omega)$ . På så sätt fick modellen möjlighet att utnyttja gemensamma strukturer hos responsfunktionerna, istället för att fem separata nätverk tränades oberoende av varandra.

För att minska risken för överanpassning användes tidigt stopp. Varje modell tränades i högst 700 epoker, men träningen avbröts om valideringsmättet inte förbättrades med minst  $10^{-6}$  under 70 efterföljande epoker. Den modellversion som gav bäst valideringsvärde under träningen sparades och användes vid den slutliga utvärderingen.

Även olika aktiveringsfunktioner undersöktes. Funktionerna GELU, SiLU, SELU och Tanh testades eftersom de har olika egenskaper och därmed påverkar hur informationen transformeras mellan nätverkets lager.

Som optimeringsmetod användes AdamW, medan olika metoder för att variera inlärningshastigheten jämfördes. En fast inlärningshastighet användes som utgångspunkt för jämförelse. Därtill undersöktes ett cosinusschema samt ett schema där inlärningshastigheten minskades när förbättringen av valideringsfelet avstannade. Syftet var att avgöra om en varierande inlärningshastighet kunde ge bättre slutlig prestanda. Hastigheten av träning och inferens var av mindre intresse eftersom modellen enbart behöver generera de slutliga responskurvorna en gång, vilka sedan sparas i tabell.

Olika förlustfunktioner testades också. MAE (*mean absolute error*) och MSE (*mean squared error*) användes eftersom de uppskattar modellfelet på olika sätt. Dessutom testades en viktad variant av MAE, där fel i områden med större respons gavs större betydelse. Detta motiverades av att området kring den kvasielastiska toppen är mer fysikaliskt relevant än övriga delar av responsfunktionen.

Den viktade MAE:n definierades som

$$w_i^{(c)} = 1 + \alpha \left( \frac{|y_i^{(c)}|}{\max |y^{(c)}|} \right)^p, \quad \mathcal{L}_{\text{wMAE}} = \frac{1}{5} \sum_{c=1}^5 \frac{\sum_i w_i^{(c)} |\hat{y}_i^{(c)} - y_i^{(c)}|}{\sum_i w_i^{(c)}},$$

där  $\alpha = 4$  och  $p = 1$ . Viktningen gjordes alltså responsvis, vilket innebär att varje responskomponent normaliserades med sitt eget maximala absolutvärde. På så sätt gavs punkter nära toppen större vikt inom respektive responsfunktion, utan att responskomponenter med större absolut amplitud automatiskt dominerade förlustfunktionen. Detta var särskilt viktigt eftersom de fem responsfunktionerna har olika storleksordning. Viktningen ändrade alltså inte vilken respons som prioriterades mellan komponenterna, utan vilka delar av varje enskild responskurva som betonades under träningen.

En annan viktig del av undersökningen var valet av indatavariabler. Förutom de direkta variablerna  $q$  och  $\omega$  testades även transformerade variabler, exempelvis  $q - \omega$ ,  $\omega/q$ ,  $\log(1 + q)$  och  $\log(1 + \omega)$ . Anledningen var att nätverket inte nödvändigtvis lär sig mest effektivt från de ursprungliga variablerna  $q$  och  $\omega$ . Genom att tillföra kombinationer och transformationer av indata kunde det undersökas om sambandet mellan indata och respons blev enklare för nätverket att approximera. Transformationerna  $\log(1 + q)$  och  $\log(1 + \omega)$  användes i stället för  $\log(q)$  och  $\log(\omega)$ , eftersom additionen av 1 gör uttrycken väldefinierade även då variabeln är noll.

Innan träningen modifierades träningsdatan så att det för varje givet  $q$ -värde lades till extra datapunkter från  $\omega = 0$  upp till det lägsta  $\omega$ -värde som fanns i den ursprungliga responskurvan. I dessa punkter sattes samtliga responsfunktioner till noll. Detta var fysikaliskt motiverat för att svara mot att det krävs en minsta energi för att övervinna bindningsenergin hos nukleonen.

Slutligen undersöktes effekten av normalisering och val av enhetssystem. Normalisering kan förbättra träningsprocessen genom att placera indatavariablerna på jämförbara numeriska skalor.

Eftersom  $q$  och  $\omega$  kan uttryckas i både MeV och GeV testades även båda enhetssystemen. Detta gjordes för att undersöka om den numeriska skalan på indata hade betydelse för nätverkets prestanda.

När normalisering användes beräknades medelvärde och standardavvikelse endast från träningsdatan i respektive korsvalideringsdel. Samma skalning applicerades därefter på både tränings- och valideringsdatan inom folden. Eftersom valideringskurvan inte användes vid beräkningen av normaliseringsparametrarna undveks informationsläckage från valideringsdatan till träningsprocessen. Detta var viktigt för att valideringsfelet skulle ge ett så oberoende mått som möjligt på modellens interpolationsförmåga.

Indatavariablerna normaliserades innan de gavs till nätverket. Nätverkets utdata behandlades också i en skalad representation, men prediktionerna transformerades tillbaka till responsfunktionernas ursprungliga skala innan förlustfunktionen och valideringsfelet beräknades. Felmåtten uttrycktes därför fortfarande i samma skala som de ursprungliga responsfunktionerna.

Modellerna utvärderades med sexfaldig korsvalidering, där en responskurva i taget användes för validering medan resterande kurvor användes för träning. Den första och sista responskurvan i  $q$  användes inte som valideringskurvor, eftersom målet var att undersöka interpolering mellan kända  $q$ -värden snarare än extrapolering. Nedan sammanfattas de parametrar som varierades i modellselektionen. För nätverksarkitekturen anger varje lista antalet dolda lager och antalet neuroner i varje dolt lager. Exempelvis betyder  $[64, 64]$  två dolda lager och 64 neuroner i varje dolt lager.

- **Nätverksarkitektur:**

- $[64, 64]$
- $[256, 256]$
- $[256, 256, 128, 128, 64]$
- $[128, 128, 128, 128, 128, 128, 128, 128, 128, 128]$
- $[128, 128, 128, 128, 128, 128]$
- $[128, 128, 64]$

- **Aktiveringsfunktion:** GELU, SiLU, SELU och Tanh.

- **Schema för inlärningshastigheten:** fast värde, cosineschema och halvering vid plåtå.

- **Förlustfunktion:** MAE, MSE och viktad MAE.

- **Indatavarianter:**

- *Bas*:  $q$  och  $\omega$
- *Avstånd*:  $q - \omega$  och  $\omega/q$
- *Logaritmer*:  $\log(1 + q)$  och  $\log(1 + \omega)$
- *Kombinationer*: *Bas*, *Bas+Avstånd*, *Bas+Logaritmer* och *Bas+Avstånd+Logaritmer*

- **Normalisering:** med respektive utan normalisering.
- **Enhetssystem:** MeV och GeV.

Totalt testades 3 456 olika nätverkskonfigurationer. Eftersom varje konfiguration utvärderades med sexfaldig korsvalidering genomfördes sammanlagt 20 736 träningskörningar för att identifiera de bäst presterande hyperparametrarna.

**Tabell 4.1:** De tio bästa modellkonfigurationerna från hyperparametersökningen, rangordnade efter lägst medelvärde av validerings-MAE över sexfaldig korsvalidering.

| #  | Nätverksarkitektur      | Aktiveringsfunktion | Inlärningsschema    | Förlustfunktion | Indatavarianter | Norm. | Enhet | Val.-MAE |
|----|-------------------------|---------------------|---------------------|-----------------|-----------------|-------|-------|----------|
| 1  | 128-128-128-128-128-128 | GELU                | Fast                | MAE             | Base+logs       | Ja    | MeV   | 0,00049  |
| 2  | 256-256-128-128-64      | GELU                | Fast                | MAE             | Base+logs       | Ja    | MeV   | 0,00050  |
| 3  | 256-256-128-128-64      | GELU                | Fast                | MAE             | Base+dist+logs  | Ja    | MeV   | 0,00051  |
| 4  | 128-128-128-128-128-128 | GELU                | Fast                | MAE             | Base+dist+logs  | Ja    | MeV   | 0,00053  |
| 5  | 256-256-128-128-64      | GELU                | Halvering vid platå | MAE             | Base+logs       | Ja    | MeV   | 0,00054  |
| 6  | 256-256-128-128-64      | SiLU                | Fast                | MAE             | Base+dist+logs  | Ja    | MeV   | 0,00057  |
| 7  | 256-256-128-128-64      | GELU                | Fast                | Viktad MAE      | Base+logs       | Ja    | MeV   | 0,00059  |
| 8  | 256-256-128-128-64      | GELU                | Fast                | MAE             | Base            | Ja    | GeV   | 0,00059  |
| 9  | 256-256-128-128-64      | GELU                | Fast                | MAE             | Base            | Ja    | MeV   | 0,00060  |
| 10 | 256-256-128-128-64      | GELU                | Coschema            | MAE             | Base+dist+logs  | Ja    | MeV   | 0,00060  |

Tabell 4.1 visar de tio modellkonfigurationer som gav lägst genomsnittligt valideringsfel över den sexfaldiga korsvalideringen. För de tre bäst rankade konfigurationerna beräknades även standardavvikelsen i validerings-MAE över de sex korsvalideringsdelningarna. För modell 1, 2 och 3 var dessa standardavvikelser  $8,6 \cdot 10^{-5}$ ,  $6,7 \cdot 10^{-5}$  respektive  $1,0 \cdot 10^{-4}$ . Skillnaderna mellan de tre lägsta medelvärdena är därmed små i förhållande till spridningen mellan korsvalideringsdelningarna. Modellvalet ska därför inte tolkas som att den bäst rankade modellen är entydigt överlägsen, utan som att den gav lägst genomsnittligt valideringsfel enligt det valda jämförelsemåttet.

Som jämförelsemått användes MAE, oberoende av vilken förlustfunktion modellen hade tränats med. Detta val gjordes för att samtliga modellkonfigurationer skulle kunna jämföras med ett gemensamt och direkt tolkningsbart felmått. Till skillnad från MSE, där större avvikelser förstärks kvadratisk, ger MAE ett mått på det genomsnittliga absoluta felet i samma skala som responsfunktionerna själva. Detta gör MAE mindre känsligt för enstaka stora lokala avvikelser, exempelvis nära tröskelområden eller i snabbt varierande delar av kurvan.

Ett viktat felmått hade också kunnat användas för valideringen, eftersom vissa delar av responsfunktionen, särskilt området kring den kvasielastiska toppen, är fysikaliskt viktiga. En sådan viktning introducerar dock ytterligare antaganden om vilka delar av kurvan som ska prioriteras. Det finns därför en risk att ett viktat valideringsmått ger en alltför smal bild av modellens kvalitet, där god prestanda i ett begränsat område kan dölja större fel i andra delar av responsfunktionen. Eftersom syftet med modellselektionen var att välja en generellt användbar interpolationsmodell bedömdes MAE vara ett mer neutralt och robust jämförelsemått.

Resultaten i tabell 4.1 visar att de bäst presterande modellerna har flera gemensamma egenskaper. Samtliga tio använder normalisering, vilket tyder på att skalningen av indata och utdata hade stor betydelse för träningsstabiliteten. Vidare domineras de bäst rankade modellerna av GELU som aktiveringsfunktion och ett fast schema för inlärningshastigheten. Detta antyder att mer avancerade scheman för inlärningshastigheten inte gav någon tydlig förbättring i detta fall. De bästa modellerna använder dessutom i huvudsak indatavarianten *Base+logs* eller *Base+dist+logs*, vilket visar att de logaritmiska transformationerna av  $q$  och  $\omega$  bidrog till att göra sambandet

mellan indata och respons lättare för nätverket att lära sig.

Den modell som gav lägst genomsnittlig validerings-MAE, modell 1 i tabell 4.1, hade arkitekturen

$$128 - 128 - 128 - 128 - 128 - 128,$$

GELU som aktiveringsfunktion, fast inlärningshastighet, MAE som träningsförlust, indatavarianten *Base+logs*, normalisering samt MeV som enhetssystem. Det genomsnittliga valideringsfelet var

$$4.90 \times 10^{-4}.$$

Skillnaden till de närmast följande modellerna är inte stor, men valet av denna modell är ändå välmotiverat eftersom rangordningen baseras på medelvärdet över samtliga korsvalideringsdelningar. Modellen har därmed inte endast presterat väl på en enskild utelämnad kurva, utan uppvisat bäst genomsnittlig interpolationsförmåga över flera olika  $q$ -värden.

Valet av modellen med lägst valideringsmått ska därför inte tolkas som att övriga modeller i tabellen är kvalitativt sämre, utan som ett konsekvent modellval utifrån ett fördefinierat jämförelsemått. Eftersom modellselektionen omfattade många kombinationer av arkitektur, aktiveringsfunktion, indatavariabler och träningsinställningar är det rimligt att välja den konfiguration som gav lägst genomsnittlig validerings-MAE. Detta minskar inslaget av subjektiva efterhandsval och gör modellvalet reproducerbart.

Efter modellselektionen användes den valda konfigurationen i de fortsatta studierna. Vid träning användes  $q = 250$  MeV som valideringskurva och utnyttjades även för tidigt stopp. Detta val baserades på tidigare tester där denna kurva gav lägst validerings-MAE när den utelämnades ur träningen. Det tyder på att  $q = 250$  MeV är en av de kurvor som är lättast för modellen att interpolera från de övriga  $q$ -värdena. Genom att använda denna kurva för validering minskas risken att träningsstoppet styrs av en särskilt svår eller avvikande valideringskurva. Samtidigt innebär valet att en kurva som modellen redan har visat sig kunna rekonstruera väl utelämnas ur träningen, vilket förhoppningsvis påverkar den slutliga modellens förmåga att beskriva de övriga delarna av parameterområdet så lite som möjligt.

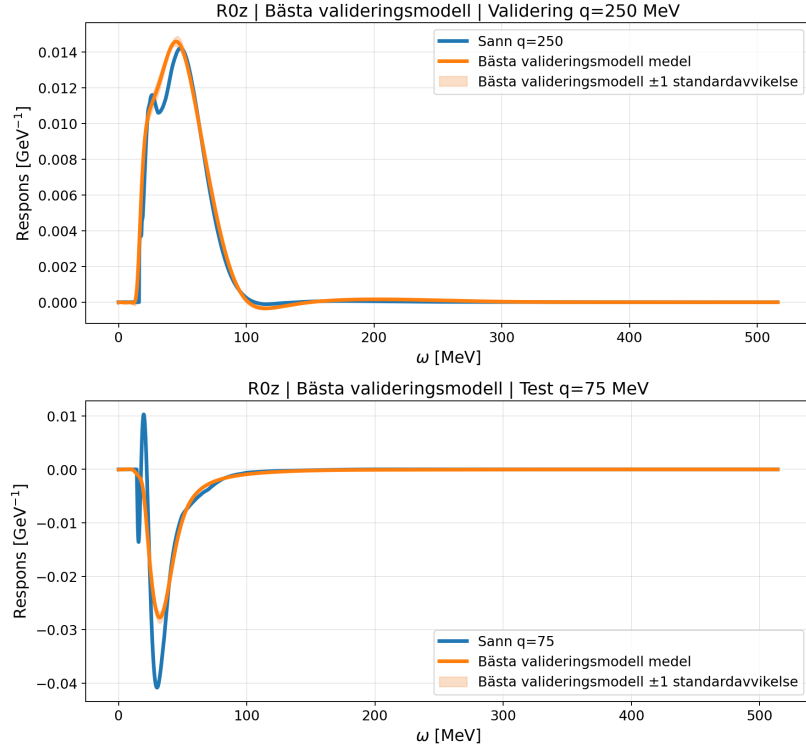
Sammantaget är det tydligt att den valda modellen ger den bästa genomsnittliga interpolationsförmågan enligt ett enkelt, transparent och gemensamt valideringsmått. Den valda modellen kombinerar låg validerings-MAE med en arkitektur som har tillräcklig kapacitet för att beskriva responsfunktionernas struktur, utan att valet bygger på ett mer komplext eller starkt viktat mått.

## 4.2 Prediktioner med den valda modellen

För att undersöka den valda modellkonfigurationens robusthet tränades modellen tio gånger. Det som skiljer körningarna åt är den slumpmässiga initialiseringen av nätverkets vikter. För varje responsfunktion beräknades sedan medelprediktionen och standardavvikelsen över de tio körningarna.

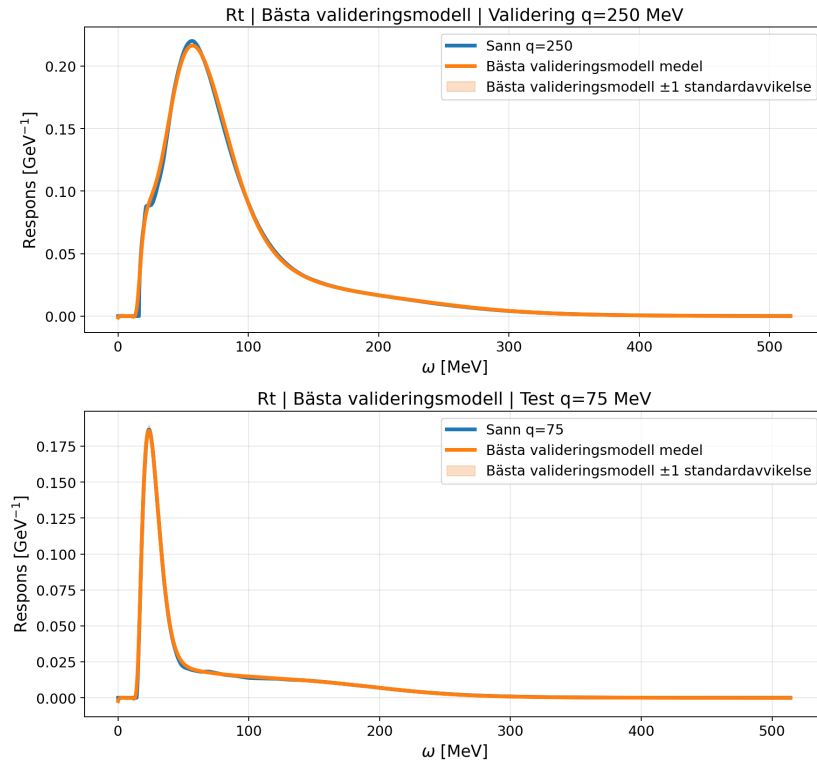
I varje figur nedan visas två fall. Den övre panelen motsvarar valideringskurvan  $q = 250$  MeV, medan den nedre panelen motsvarar testkurvan  $q = 75$  MeV. Testkurvan  $q = 75$  MeV användes inte under träningen av modellen. Den användes i stället för att undersöka hur modellen predikterar okända data och hur väl den generaliserar till ett  $q$ -värde som inte ingick i träningsmängden. Den blå kurvan visar referensdata, den orange kurvan visar medelprediktionen och det skuggade området visar  $\pm 1$  standardavvikelse beräknad från de tio träningarna.

R0z: sann kurva vs Bästa valideringsmodell medel  $\pm 1$  standardavvikelse över 10 körningar



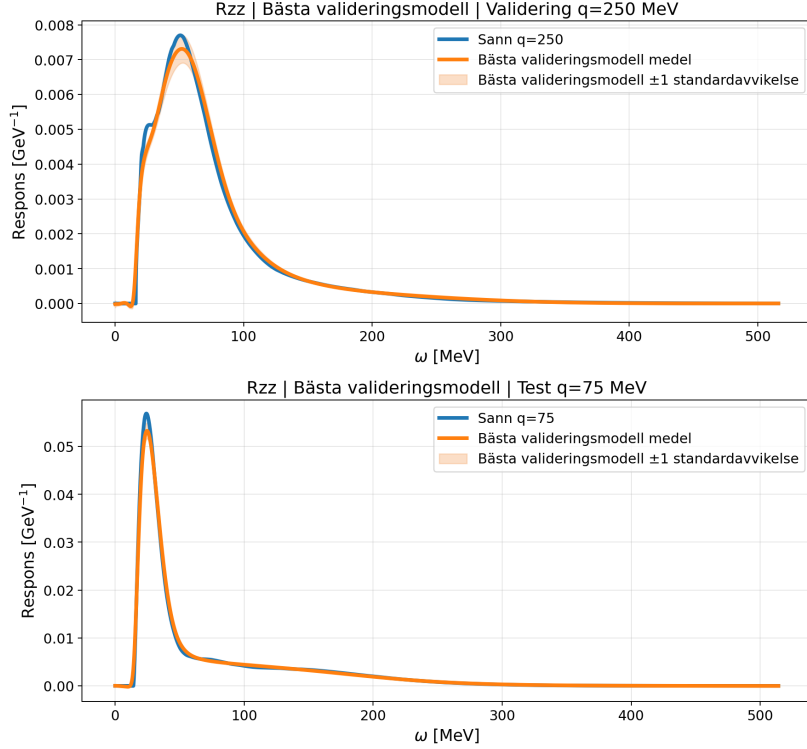
**Figur 4.2:** Förutsägelse av  $R_{0z}$  med den valda modellen vid  $q = 250$  MeV och  $q = 75$  MeV.

Rt: sann kurva vs Bästa valideringsmodell medel  $\pm 1$  standardavvikelse över 10 körningar



**Figur 4.3:** Förutsägelse av  $R_t$  med den valda modellen vid  $q = 250$  MeV och  $q = 75$  MeV.

Rzz: sann kurva vs Bästa valideringsmodell medel  $\pm 1$  standardavvikelse över 10 körningar



**Figur 4.4:** Förutsägelse av  $R_{zz}$  med den valda modellen vid  $q = 250$  MeV och  $q = 75$  MeV.

Figurerna 4.2–4.4 visar tre representativa exempel på hur den valda modellen återger responsfunktionernas form. I de visade fallen följer medelprediktionen referenskurvan väl både för valideringsfallet och för testfallet. Detta tyder på att modellen inte enbart har anpassats till  $q = 250$  MeV, utan även kan ge rimliga prediktioner för ett separat  $q$ -värde.

För  $R_{0z}$ , figur 4.2, är bilden mer varierad. I valideringsfallet återges den positiva toppen relativt väl, även om små skillnader syns nära den första snabba ökningen och vid övergången mot små negativa värden. I testfallet är avvikelsen tydligare, särskilt kring den skarpa strukturen vid låga  $\omega$ -värden, där modellen inte fullt ut följer referenskurvas snabba positiva och negativa variation. Trots detta är standardavvikelsen smal, vilket innebär att de olika initialiseringarna leder till liknande prediktioner.

Responsen  $R_t$ , figur 4.3, predikteras mycket stabilt. Medelprediktionen följer referenskurvan väl både vid toppen och i svansen. Det gäller för både  $q = 250$  MeV och  $q = 75$  MeV. Standardavvikelsen är mycket smal jämfört med responsens storlek, vilket visar att resultatet är stabilt med avseende på den slumpmässiga initialiseringen.

För  $R_{zz}$ , figur 4.4, följer medelprediktionen referenskurvan på ett övergripande bra sätt. I valideringspanelen syns en viss underskattning nära toppen, men modellen återger fortfarande både toppens läge och kurvans avtagande del rimligt väl. I testpanelen är överensstämmelsen god efter det första toppområdet, även om mindre skillnader förekommer vid toppen. Av de olika responsfunktionerna är det  $R_{zz}$  som har det tydligaste standardavvikelsebandet. Det visar att denna respons är mer känslig för vilken initialisering som används vid träningen, jämfört med de övriga responserna.

Resultaten visar sammantaget att den valda modellen ger konsekventa prediktioner för de olika

responsfunktionerna. I de flesta fall är spridningen mellan de tio träningskörningarna liten, vilket innebär att modellen ger robusta prediktioner trots olika initialiseringar. Den största variationen mellan körningarna ses för  $R_{zz}$ , där standardavvikelsebandet är tydligare än för de övriga responserna. De största avvikelserna mellan modellens medelvärde och referensdata uppträder främst i områden där responsen ändras snabbt, särskilt för  $R_{0z}$  i testfallet  $q = 75$  MeV. Eftersom  $q = 75$  MeV inte ingick i träningen ger detta testfall en separat kontroll av modellens generaliseringsförmåga. Överlag visar figurerna att modellen har lärt sig ett robust samband mellan  $(q, \omega)$  och de olika responsfunktionerna.

### 4.3 Propagerad teoretisk osäkerhet

För att propagera osäkerhetsbandet hos responsfunktionerna genom det neurala nätverket tränades den valda modellen tre gånger. Nätverken skiljde sig endast åt genom träningsmålet:  $R_{\alpha}^{\min}$ ,  $R_{\alpha}^{\text{med}}$  respektive  $R_{\alpha}^{\max}$ , enligt definitionerna i sektion 4.1. Detta gjordes för samtliga responskomponenter

$$R_{\alpha} \in \{R_{00}, R_t, R_{xy}, R_{zz}, R_{0z}\}.$$

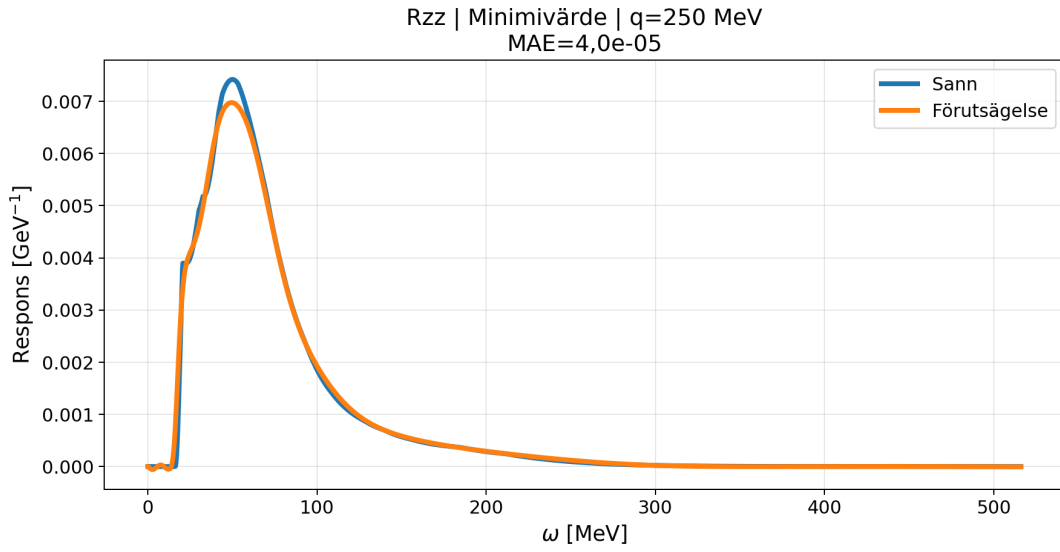
De tre resulterande nätverken beskriver därmed den nedre gränsen, den centrala responsen och den övre gränsen för det teoretiska osäkerhetsbandet. Träningen utfördes på samtliga tillgängliga  $q$ -värden förutom  $q = 250$  MeV, vilken användes som valideringskurva. På så sätt kunde nätverkets förmåga att interpolera till ett icke-tränat  $q$ -värde undersökas.

Figurerna 4.5–4.7 visar prediktionerna för  $R_{zz}$  på valideringskurvan  $q = 250$  MeV för de tre träningsmålen. Endast  $R_{zz}$  visas här, eftersom de övriga responskomponenterna inte uppvisade några större visuella skillnader mellan minimi-, medel- och maximimålet.

För medelmålet följer nätverkets prediktion den sanna responsen väl över nästan hela  $\omega$ -intervallet. Både den större kvasielastiska toppen och den långsamma avklingningen vid högre energiöverföring återges med god noggrannhet. För minimi- och maximimålet är avvikelserna något större, framför allt i området där responsen stiger snabbt mot sitt maximum. Detta är även den del av kurvan där responsfunktionen har störst lokal variation och där interpoleringen därför är mest känslig.

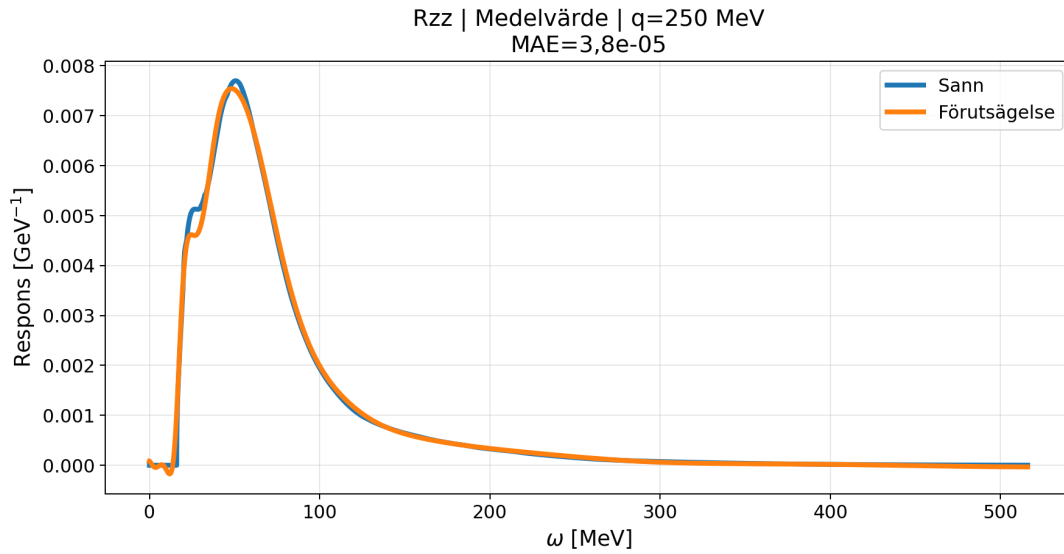
För maximimålet syns dessutom att nätverket inte fullt ut återskapar den mindre strukturen vid låga  $\omega$ -värden innan den huvudsakliga toppen. Detta tyder på att mindre och smalare strukturer i responsfunktionerna är svårare för nätverket att lära sig än den övergripande formen. Trots dessa avvikelser beskriver samtliga tre nätverk den dominerande toppen och responsens avtagande beteende väl.

Rzz: Förutsägelse jämfört med sann kurva på valideringskurvan  $q=250$  MeV för minimivärde



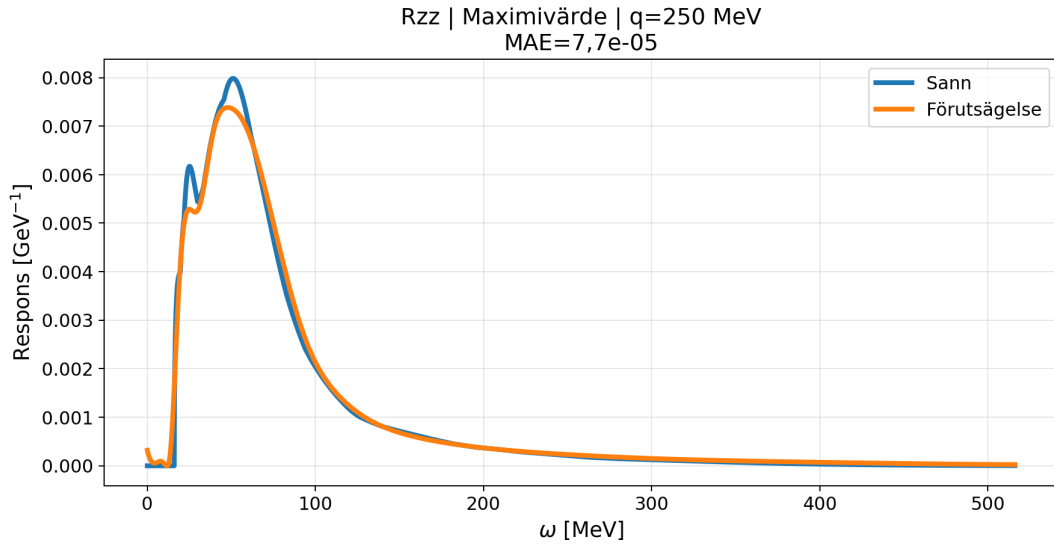
**Figur 4.5:** Prediktion och sann respons för  $R_{zz}$  vid  $q = 250$  MeV för minimimålet.

Rzz: Förutsägelse jämfört med sann kurva på valideringskurvan  $q=250$  MeV för medelvärde



**Figur 4.6:** Prediktion och sann respons för  $R_{zz}$  vid  $q = 250$  MeV för medelmålet.

Rzz: Förutsägelse jämfört med sann kurva på valideringskurvan  $q=250$  MeV för maximivärde



**Figur 4.7:** Prediktion och sann respons för  $R_{zz}$  vid  $q = 250$  MeV för maximimålet.

De numeriska felen för valideringskurvan sammanfattas i tabell 4.2. Eftersom responskomponenterna har olika storleksordning bör MAE-värdena främst jämföras inom samma responskomponent och inte direkt mellan olika komponenter.

**Tabell 4.2:** Medelabsolutfel på valideringskurvan  $q = 250$  MeV för nätverk tränade på minimi-, medel- respektive maximirespons. Raderna visar de fem responskomponenterna. Sista raden visar medelvärdet över responskomponenterna.

| Respons                              | Minimi               | Medel                | Maximi               |
|--------------------------------------|----------------------|----------------------|----------------------|
| $R_{00}$                             | $5,2 \times 10^{-4}$ | $4,9 \times 10^{-4}$ | $5,9 \times 10^{-4}$ |
| $R_{0z}$                             | $1,5 \times 10^{-4}$ | $1,3 \times 10^{-4}$ | $1,2 \times 10^{-4}$ |
| $R_t$                                | $8,9 \times 10^{-4}$ | $6,6 \times 10^{-4}$ | $7,1 \times 10^{-4}$ |
| $R_{xy}$                             | $7,8 \times 10^{-4}$ | $4,9 \times 10^{-4}$ | $4,5 \times 10^{-4}$ |
| $R_{zz}$                             | $4,0 \times 10^{-5}$ | $3,8 \times 10^{-5}$ | $7,7 \times 10^{-5}$ |
| Medelvärde över responskomponenterna | $4,8 \times 10^{-4}$ | $3,6 \times 10^{-4}$ | $3,9 \times 10^{-4}$ |

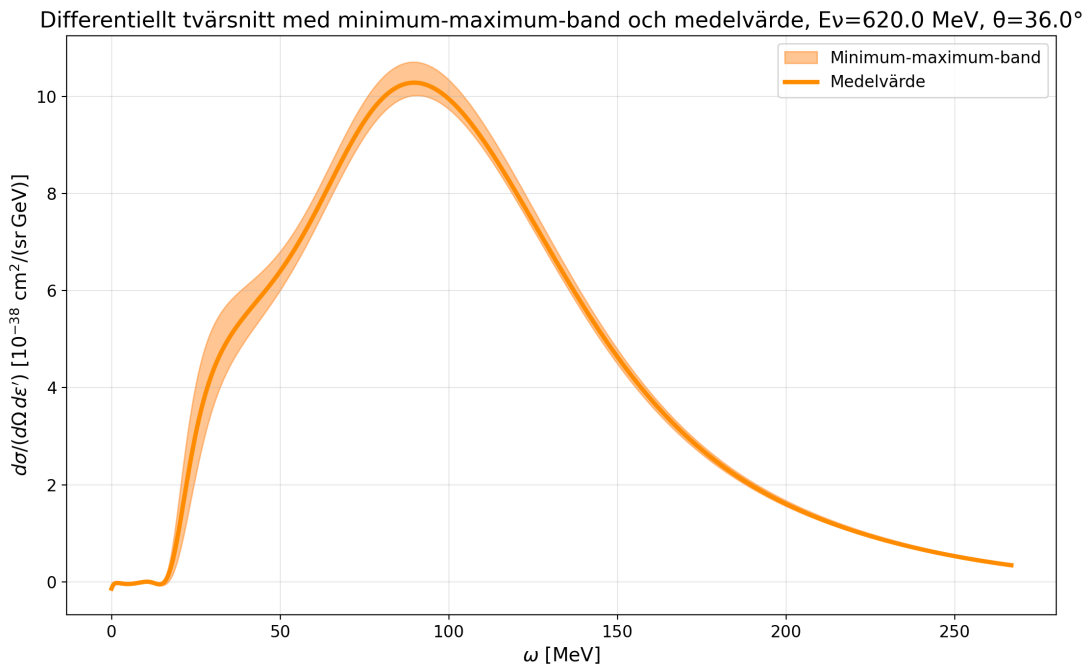
Tabellen visar att modellen som tränades på medelresponsen uppvisar det lägsta genomsnittliga valideringsfelet. Skillnaderna mellan de tre träningsfallen är dock små, och bör därför inte övertolkas. De kan delvis bero på stokastiska variationer i träningsprocessen, exempelvis initialisering av nätverkets vikter och optimeringens konvergens, snarare än på en systematisk skillnad mellan minimi-, medel- och maximiresponsen.

## 4.4 Dubbeldifferentiellt tvärsnitt

Med hjälp av den valda modellen beräknades det dubbeldifferentiella tvärsnittet  $\frac{d\sigma}{d\Omega d\epsilon'}$  för  $E_\nu = 620$  MeV och  $\theta = 36^\circ$ . Den utgående laddade leptonen är en myon. Kinematiken beräknades enligt uttrycken i avsnitt 2.3.

Kurvan beräknades punktvis. Först valdes flera diskreta värden på energiöverföringen  $\omega$ . För varje sådant värde beräknades det motsvarande  $q$ -värdet från den valda kinematiken med  $E_\nu = 620$

MeV,  $\theta = 36^\circ$  och en utgående myon. Den valda modellen användes sedan för att förutsäga responsfunktionerna i varje  $(q, \omega)$ -punkt. Dessa responsvärden sattes in i tvärsnittsformeln, vilket gav ett värde på det dubbeldifferentiella tvärsnittet för varje  $\omega$ . Figur 4.8 visar det dubbeldifferentiella tvärsnittet när den valda modellen används för att förutsäga responsfunktionerna. Den orange kurvan visar det dubbeldifferentiella tvärsnittet beräknat från medelresponsen, medan det skuggade området visar minimum–maximum-bandet. Detta gjordes genom två separata beräkningar, en med minimirespons och en med maximirespons, där minimi- och maximirespons definieras som i avsnitt 4.3. Kurvan växer snabbt vid låga  $\omega$ , når sitt maximum omkring  $\omega \approx 90$  MeV och avtar därefter mot större energiöverföring. Detta visar att det största bidraget till det dubbeldifferentiella tvärsnittet, för den valda kinematiken, kommer från området runt toppens läge. Bandet är som bredast under det dubbeldifferentiella tvärsnittets uppgång och i närheten av toppområdet, vilket visar att skillnaden mellan minimi- och maximirespons påverkar tvärsnittet mest i det område där kurvan växer snabbt och når sina största värden. Vid högre  $\omega$  blir bandet smalare, eftersom de två beräkningarna närmar sig varandra.



**Figur 4.8:** Differentiellt tvärsnitt från den valda modellen, med medelvärde och minimum–maximum-band.

## 4.5 Implementation av ab initio-baserad modell i NuWro

För att generera händelser i NuWro måste dubbeldifferentiella tvärsnittet för slumpmässigt dragna punkter i parameterrymden beräknas. Detta måste ske med hög hastighet, så att ett stort antal dragningar kan bearbetas för att minska den statistiska osäkerheten i beräkningen. Den i NuWro implementerade lokala fermigasmodellen är snabb nog för att kunna användas direkt i händelsegenerationen [24].

Den nya ab initio-baserade modellen är som tidigare nämnts mycket beräkningsintensiv, vilket motiverade användandet av det interpolerande neurala nätverket. Men även det neurala nätverket bedömdes vara för beräkningstungt för att användas direkt i NuWro. Antalet stora matrismultiplikationer som skulle behöva göras för varje dragning skulle göra händelsegenerationen extremt långsam. NuWro stödjer inte heller den grafikprocessoraccelerering som använts

under träningen av nätverket utan betydligt arbete och omstrukturering.

Ab initio-modellen implementerades därför genom uppslag i förberäknade tabeller. Tabelluppslagning användes i NuWro sedan tidigare (i *2020Valencia*-modellen) [25], vilket stödjer att det är ett giltigt tillvägagångssätt. Värdena för de fem responskomponenterna samplades från det neurala nätverket på ett 4 MeV-grid i parameterrymden  $0 \text{ MeV} < q, \omega < 1200 \text{ MeV}$ . Datan sparades som endimensionella arrays i headerfiler och kompilerades med programmet - de lästes alltså inte in under körtid. I NuWro användes sedan bilinjär interpolering för att beräkna responskomponenterna för godtyckliga  $(q, \omega)$  i utfallsrymden. Rutnätet i  $q$  och  $\omega$  för tabellvärdena är så tätt att användandet av bilinjär interpolering i detta läge inte bedöms introducera något signifikant interpolationsfel. Responsfunktionernas värden användes sedan i ekvation 2.3 för att beräkna det dubbeldifferentiella tvärsnittet.

Före det dubbeldifferentiella tvärsnittet kan beräknas krävs en parameterdragning. Ekvation 2.3 är uttryckt i  $\epsilon'$  och  $\cos \theta$ , vilket gör det naturligt att sampla i den parametreringen. Den utgående energin  $\epsilon'$  har dock ett linjärt beroende till  $\omega$ , och då responsfunktionerna är uttryckta i  $\omega$  samplades  $\omega$  istället. För varje iteration i MC-processen slumpas ett  $\cos \theta$  från en likformig fördelning i intervallet  $[-1, 1]$  och ett  $\omega$  i intervallet  $[\omega_{min}, \omega_{max}]$ . Eftersom  $\omega$  är energiöverföring fås  $\omega_{max} = \epsilon - \epsilon'_{min}$ .  $\epsilon$  fås i sin tur från den inkommande leptonens energi och  $\epsilon'_{min}$  är den minsta möjliga utgående energin. Detta inträffar då den utgående leptonen har 0 rörelsemängd. I naturliga enheter har vi  $E^2 = k^2 + m^2$ , vilket för detta fall ger

$$\epsilon'^2_{min} - k^2 = m_l^2 \implies \epsilon'_{min} = m_l$$

där  $\omega_{min}$  blir bindningsenergin för att slå loss en nukleon från kärnan, vilket är vad som sker vid kvaselastisk spridning. Parametrar utanför dessa spann är icke-fysikaliska och ska ge 0 i tvärsnitt. Dragningen begränsas ändå till spannet för att effektivisera processen samt undvika oförutsedda bieffekter.

Det differentiella tvärsnittet  $\frac{d\sigma^2}{d\cos\theta d\omega}$  beräknas sedan för varje dragen parameter med ekvation 2.3 och värden på responsfunktioner uppslagna i tabellerna. Eftersom responsfunktionerna är uttryckta i  $q$  istället för  $\cos \theta$  användes ekvation 2.9 för att konvertera koordinaten.

Varje händelse tilldelas en vikt  $w = \frac{d\sigma}{d\cos\theta d\omega} \Delta\omega \Delta\cos\theta$  där  $\Delta\omega = \omega_{max} - \omega_{min}$  och  $\Delta\cos\theta = 2$  är samplingsintervallerna.

Efter testfasen som beskrivs i kapitel 3.2 genereras de faktiska händelserna. De slumpdragna händelserna med tilldelade vikter  $w_i$  accepteras eller förkastas med sannolikheten  $\frac{w_i}{w_{max}}$  där  $w_{max}$  är maximala vikten från testfasen. De accepterade händelserna sparas som oviktade händelser i filen "eventsout.root". Det totala tvärsnittet  $\sigma$  beräknas även. Det ges analytiskt av

$$\sigma = \int_{\epsilon'_{min}}^{\epsilon'_{max}} \int_{-1}^1 \frac{d\sigma}{d\cos\theta d\epsilon'} d\cos\theta d\epsilon' = \int_{\omega_{min}}^{\omega_{max}} \int_{-1}^1 \frac{d\sigma}{d\cos\theta d\omega} d\cos\theta d\omega,$$

men vid MC-integration i NuWro uppskattas det totala tvärsnittet med  $\sigma = \frac{1}{N} \sum_{i=1}^N w_i$  där  $N$  är antalet accepterade händelser. Det totala tvärsnittet ges alltså av medelvärdet av vikterna, vilket är en mycket enkel operation jämfört med det tidigare integrationsuttrycket.

Till sist ska även den utgående myonens rörelsemängdsvektor  $k'$  beräknas.  $k'$  ska uppfylla

ekvationssystemet

$$\begin{cases} \vec{k} \cdot \vec{k}' = kk' \cos \theta \\ |\vec{k} - \vec{k}'| = q \end{cases} \quad (4.2)$$

vilket har lösningen

$$\begin{aligned} |k'| &= \epsilon \cos \theta + \sqrt{q^2 - \epsilon^2 \sin^2 \theta} \\ k'_z &= |k'| \sin \theta \\ k'_r &= |k'| \cos \theta \end{aligned} \quad (4.3)$$

där komponenterna  $k'_z$  och  $k'_r$  är komponenten längsmed respektive ortogonal mot den infallande axeln. Interaktionen är radiellt symmetrisk.

Vid körning ställdes önskade parametervärden in genom redigering av textfilen params i mappen data. Den inkommande strålens energi sattes till 600 MeV och Pauli-blockering stängdes av. Alla dynamiska kanaler utom neutrino-elektron-spridning samt kvasielastisk laddad ström avaktiverades. Användningen av spektralfunktioner som kärnmodell samt RPA-korrelationer (Random Phase Approximation) stängdes av.

För de slutliga resultaten implementerades en hybridmodell som valde den nya ab initio-modellen eller den ursprungliga LFG-modellen beroende på interaktionens  $q$ . Den nya modellen bedömdes ge tillförlitliga resultat enbart för  $q < 400$  MeV, så för alla händelser med ett högre  $q$  användes istället LFG-modellen.

## 4.6 Tvärsnitt från NuWro-händelser jämfört med teoretiskt

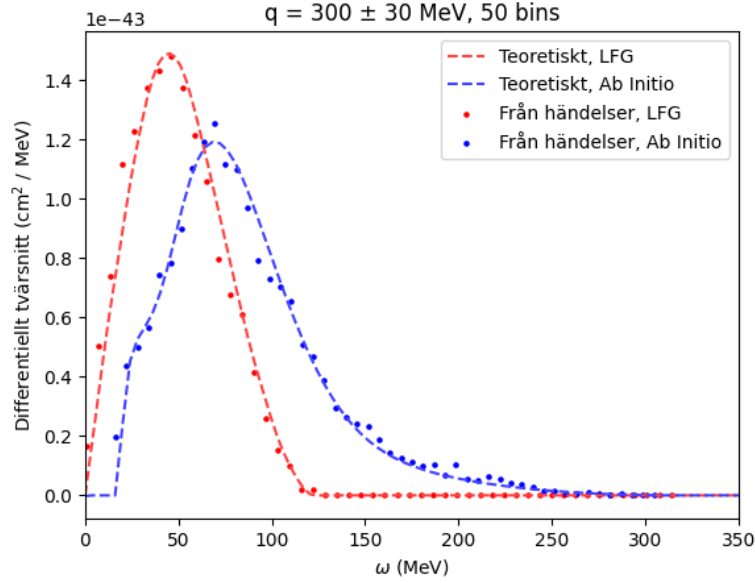
Vid körning av NuWro sparas de genererade händelserna i ROOT-filen "eventsout.root". Varje händelse är ett objekt innehållande information om en enskild neutrino-kärninteraktion, där framförallt de ingående och utgående partiklarnas typ och 4-rörelsemängd är av intresse. För att validera att LFG-modellen implementerats korrekt jämfördes det teoretiska dubbeldifferentiella tvärsnittet som funktion av  $\omega$  beräknat i Python med det dubbel-differentiella tvärsnittet beräknat från de genererade händelserna. Det teoretiska tvärsnittet  $\frac{d\sigma}{d\cos\theta d\epsilon'}$  beräknades med ekvation 2.3. För att kunna jämföra differentiella tvärsnitt med varandra och jämsides NRF:erna är det användbart att göra en graf i  $\omega$  för ett specifikt  $q$ . Det behövs då göras ett variabelbyte  $(\cos\theta, \epsilon') \rightarrow (q, \omega)$ . Från ekvation 2.9 kan jakobianen  $J$  härledas

$$\begin{aligned} 2q dq &= -2kk' d\cos(\theta) \\ J &= \frac{d\cos(\theta)}{dq} = \frac{q}{kk'} \end{aligned} \quad (4.4)$$

vilket multipliceras med det teoretiska tvärsnittet.

För att beräkna det differentiella tvärsnittet från genererade händelser antogs tvärsnittet vara konstant i ett litet tvådimensionellt intervall  $A = \Delta q \Delta \omega$ . Detta gav

$$\iint_A \frac{d\sigma}{dq d\omega} dq d\omega \approx \frac{d\sigma}{dq d\omega} \Delta q \Delta \omega \quad (4.5)$$



**Figur 4.9:** Differentiella tvärsnittet beräknat från NuWro-genererade händelser samt teoretiskt tvärsnitt beräknat från tabellerade responsfunktioner. LFG här är kurvan i figur 2.4.

och

$$\frac{d\sigma}{dq d\omega} \approx \frac{1}{\Delta q \Delta \omega} \iint_A \frac{d\sigma}{dq d\omega} dq d\omega \approx \frac{1}{\Delta q \Delta \omega} \frac{\# \text{ events i } \Delta q \Delta \omega}{\text{totalt } \# \text{ events}} \cdot \sigma \quad (4.6)$$

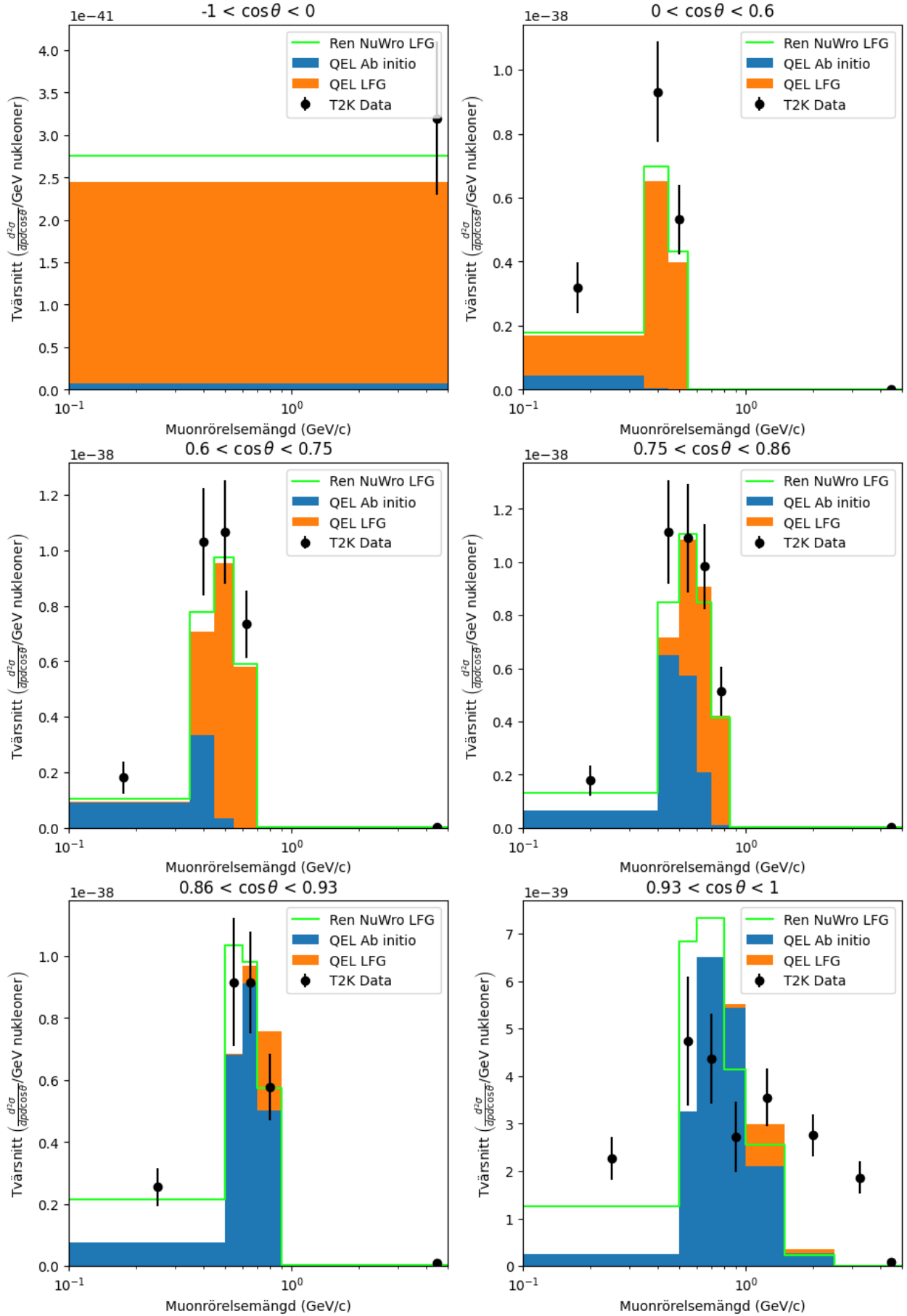
där  $\sigma$  är det totala tvärsnittet.

Figur 4.9 visar det differentiella tvärsnittet beräknat från händelser genererat i NuWro med den ovanstående metoden. Händelserna genereras för en monoenergetisk stråle med  $E = 600$  MeV eftersom det teoretiska tvärsnittet beräknas för en ingångsenergi. För LFG-modellen är Pauli blockering och RPA-korrigerad avstängt för att kunna observera modellens okorrigerade tvärsnitt. Detta jämförs med det teoretiska tvärsnittet beräknat med ekvation 2.3 från samma tabeller som laddats i NuWro. Detta indikerar att tvärsnittsberäkningen är korrekt implementerad. Händelserna genererade med den nya modellen ger även ett differentiellt tvärsnitt med liknande form och magnitud till NuWro:s inbyggda LFG-modell, vilket ytterligare stärker en korrekt implementering. Spridningen i punkterna beror på statistisk osäkerhet, orsakat av små bins med händelser i varje bin. Om binstorleken ökas blir dock istället topparna utsmetade, då data från ett större spann av  $q$ -värden används.

## 4.7 Jämförelse med data från T2K-experimentet

I figur 4.10 visas hur hybridmodellen påverkar det beräknade tvärsnittet. ”Ren NuWro LFG” syftar till NuWro:s inbyggda LFG-modell, medan hybridmodellen är den tidigare beskrivna kombinationen av den nya ab initio-modellen och den inbyggda LFG-modellen. Notera att tvärsnittet här uttrycks i  $dp/d\cos(\theta)$  eftersom  $\omega$  inte är direkt mätbart i neutrinoexperiment. Händelserna genererades för en stråle med energifördelningen satt till den uppmätt i T2K-experimentet [7]. För LFG-modellen aktiverades Pauli blocking, då det utgör en del av den kompletta LFG-modellen [24]. Ungefär 25% av händelserna tilldelades den nya ab initio-baserade modellen medan resterande 75% beräknades med den inbyggda LFG-modellen.

Det kan ses att ab initio-bidraget är störst för  $\cos\theta$  nära 1. Detta är väntat då  $\theta$  nära noll innebär liten spridning och därmed låga  $q$ , vilket är där den ab initio-baserade modellen används.



**Figure 4.10:** Det differentiella värsnittet för syre beräknat från händelser för hybridmodellen, med de separata bidragen från ab initio-datan och NuWro:s inbyggda LFG-modell synliga. Detta jämförs med den rena LFG-modellen samt med tvärsnitt uppmätt i T2K-experimentet.

4.10 replikerar även delar av figur 13 från [7], där tvärsnittet för syre uppmätt från T2K-experimentet jämförs med modeller från olika neutrinohändelsegeneratorer. Det beräknade värdet för den rena LFG-modellen är nära det som presenteras i rapporten, men skiljer framförallt för  $\cos(\theta)$  nära 1. Detta tros bero på att en annan variant av modellen används i rapporten. Både fördelningen och magnituden stämmer dock för det mesta väl överens, så metoden för att beräkna det differentiella tvärsnittet bedöms vara korrekt.

4.11 visar ett preliminärt resultat för bidragen från de olika interaktionsmekanismer som är relevanta för T2K-experimentet. Detta tros dock inte vara tillförlitligt - troligtvis används fel parametrar för NuWro samt felaktig efterbearbetning (bland annat filtrering av alstrade pioner som inte kan detekteras i verkligheten) för att återskapa vad som sker i T2K-experimentet.

# 5

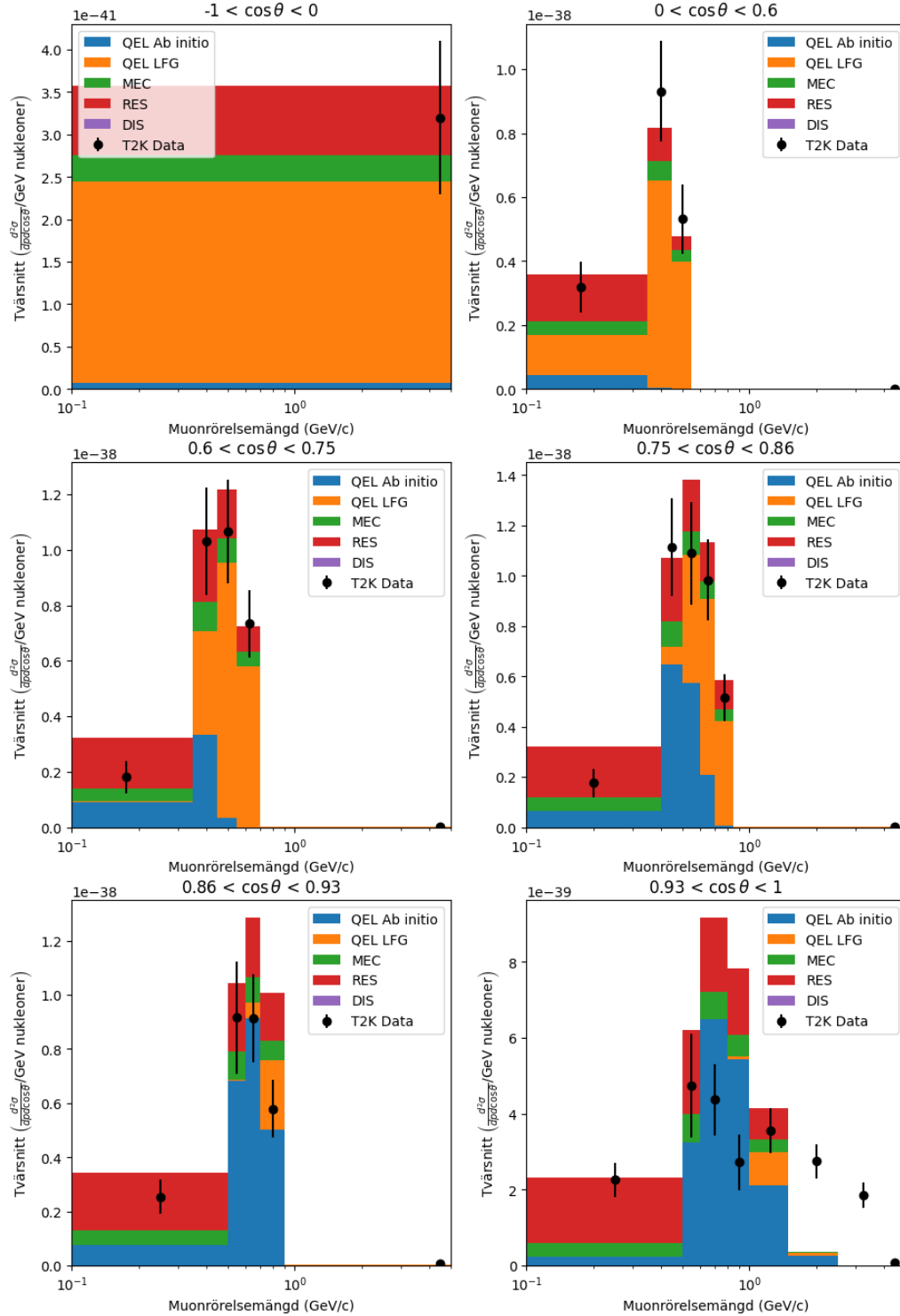
## Slutsats

### 5.1 Felkällor

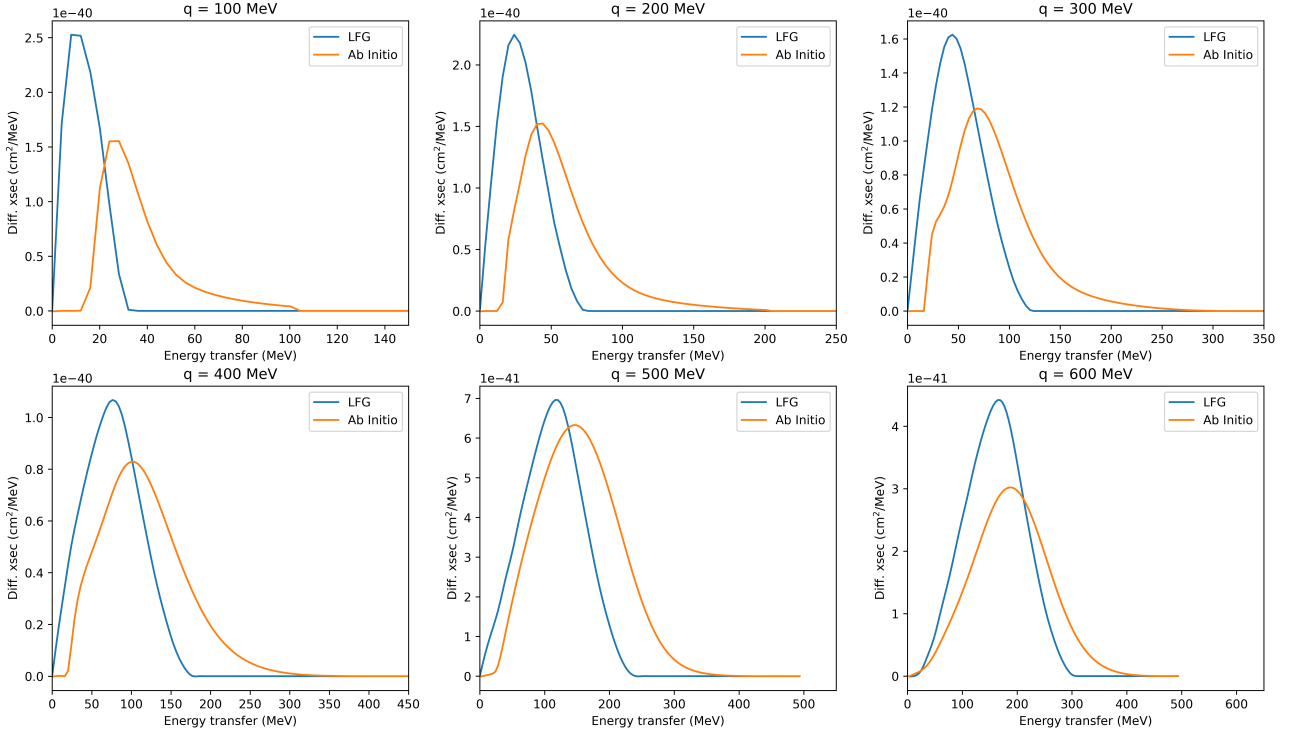
En central felkälla i arbetet var den begränsade mängden ab initio-beräknad data. Det neurala nätverket tränades på responsfunktionerna för ett begränsat antal värden på rörelsemängdsöverföringen  $q$ , inom ett intervall från 50 MeV till 400 MeV, med totalt åtta datapunkter. Det innebär att information om hur responsfunktionerna varierar mellan olika  $q$ -värden var begränsad. Detta hade särskilt stor påverkan i områden där responsfunktionernas form ändras snabbt mellan olika  $q$ -värden. Vidare hade en större mängd av ab initio-beräknad data möjliggjort fler uppdelningar i validerings- och testmängder under utvärderingen av modellen.

Ytterligare osäkerheter kommer från de ab initio-beräknade datapunkterna och interpolationen av dessa med neurala nätverket. Det neurala nätverket tränades på medelvärdet av det minsta och största värdet i energiöverföringen  $\omega$  för varje värde av respektive responsfunktion. Det innebär att själva utseendet av responsfunktionerna är endast ett medelvärde. Det resulterade i att intressanta ändringar i kurvan inte tas upp av det neurala nätverket, trots att sådana kan vara av intresse.

En potentiell felkälla är att den implementerade ab initio-modellen används i form av diskreta tabellvärden som sedan interpoleras bilinjärt. Som nämns i kapitel 4 antas denna approximation emellertid inte introducera något signifikant fel, vilket stöds av figur 5.1. Figuren visar det dubbeldifferentiella tvärsnittet  $\frac{d\sigma}{d\omega dq}$  och graferna är linjära mot responsfunktionerna implementerade i form av tabeller. Att graferna, speciellt de för lägre  $q$ , inte är släta beror på att tabellvärdena är diskreta. Detta bedöms dock ha en liten procentuell inverkan på arean under grafen, det vill säga integralen, och således påverkas inte det resulterande tvärsnittet betydligt. Om något leder det till att tvärsnittet underskattas aningen.



**Figur 4.11:** Preliminärt resultat för tvärsnitt beräknat från händelser för alla interaktionsmekanismer relevanta för T2K-experimentet, där hybridmodellen används för kvasielastisk spridning (QEL).



**Figur 5.1:** Dubbel-differentiellt tvärsnitt  $\frac{d\sigma}{d\omega dq}$  beräknat med LFG-modellen samt ab initio-modellen.

Jämförelsen med experimentell data innehåller dessutom flera osäkerheter. Skillnader kan bero av parameterval i NuWro, behandling av Pauli-blockering, histogrambinning och andra detaljer i analyskedjan. Detta gör att jämförelsen med T2K-data främst bör ses som en kvalitativ kontroll av modellens beteende, snarare än som ett exakt test av modellens precision.

Slutligen är en viktig begränsning i hur modellen implementeras som en hybridmodell. Den ab initio-baserade modellen användes endast för  $q < 400$  MeV medan NuWro:s ursprungliga LFG-modell användes för högre  $q$ . Det finns ingen exakt fysikalisk motivering när bytet mellan modellerna bör ske.

De dominerande felkällorna bedöms därmed vara den begränsade mängden ab initio-beräknad data, interpolationsprocessen med det neurala nätverket samt det begränsade  $q$ -område där den ab initio-baserade modellen används i NuWro.

## 5.2 Vidare studier

Det tydligaste området för vidare studier är att utveckla jämförelsen med T2K-datan, då det finns vissa begränsningar i verkliga experiment som inte beaktas fullt i det nuvarande resultatet. Detta skulle tillåta att utvärdera om den ab initio-baserade modellen ger ett mer korrekt totalt dubbeldifferentiellt tvärsnitt.

En alternativ metod för att interpolera ab initio-datan undersöktes kort i de tidigare faserna, men valdes att inte användas då det introducerade ytterligare komplexitet utan att ge tydligt förbättrade resultat. Tillvägagångssättet är dock inte uteslutet och skulle kunna ge fördelar i form av kortare träningstider och enklare modeller. Metoden innebär att använda en enklare och snabbare modell (i detta fall de så kallade spektralfunktionerna) för kärnresponsfunktionerna

och träna det neurala nätverket på residualen mellan *ab initio*-datan och den enklare modellen. Motivationen var att försöka minska frihetsgraderna i datan som ska interpoleras genom att använda information från en redan känd fysikaliskt motiverad approximation.

Den ytliga undersökningen indikerade att beteendet som dessa enklare modeller modellerar redan var det interpolatorn lärde sig enklast, medan de mer komplexa detaljerade delarna inte inte förenklades. Färre frihetsgrader, om än enkla, skulle dock leda till att det neurala nätverket kan förenklas och på så sätt bli snabbare.

Den största begränsande faktorn bedöms vara mängden *ab initio*-genererad data och de stora avstånden i  $q$ . Vissa av kurvorna bedömdes variera så pass mycket i  $q$  att datan som fanns inte innehöll nog information för att återge det fulla beteendet i responsfunktionerna.

## 5.3 Slutsats

En metod för att ladda tabellerade responsfunktioner i NuWro har utvecklats och visats alstra händelser med korrekt dubbeldifferentiellt tvärsnitt. Den beräkningspipeline som demonstrerats väntas vara användbar för att enkelt kunna implementera nya modeller. Av särskilt intresse är faktumet att responsfunktionerna kunnat interpoleras så väl som de gjorts, trots den mycket glesa underliggande datan. Detta tillåter modeller som är mycket kostsamma att evaluera att användas i NuWro, vilket förbättrar möjligheterna att kunna korrekt modellera vår fysiska värld.

Arbetet med jämförelsen för T2K-experimentets uppmätta tvärsnitt har inte ännu gett ett slutgiltigt resultat. För att utvärdera den *ab initio*-baserade kärnmodellens prestanda jämfört med den inbyggda LFG-modellen krävs vidare studier.

# Litteratur

- [1] B. R. Martin och G. Shaw, *Nuclear and Particle Physics*. John Wiley och Sons Ltd, 2019.
- [2] J. Sobczyk m. fl., *Ab initio-beräkningar för neutrino-kärninteraktioner*, Opublicerat arbete, Chalmers tekniska högskola.
- [3] M. Thomson, *Modern Particle Physics*. Cambridge: Cambridge University Press, 2013, ISBN: 9781107034266.
- [4] C. Giganti, S. Lavignac och M. Zito, "Neutrino oscillations: The rise of the PMNS paradigm," *Progress in Particle and Nuclear Physics*, årg. 98, s. 1–54, jan. 2018, ISSN: 0146-6410. DOI: 10.1016/j.ppnp.2017.10.001.
- [5] R. Lalnuntluanga, R. Pradhan och A. Giri, "Probing neutrino-nucleus interaction in DUNE and MicroBooNE," *Nuclear Physics B*, årg. 1008, s. 116 703, 2024, ISSN: 0550-3213. DOI: 10.1016/j.nuclphysb.2024.116703.
- [6] "About T2K," hämtad 3 maj 2026. URL: <https://t2k-experiment.org/t2k/>.
- [7] K. Abe m. fl., "Simultaneous measurement of the muon neutrino charged-current cross section on oxygen and carbon without pions in the final state at T2K," *Physical Review D*, årg. 101, nr 11, juni 2020, ISSN: 2470-0029. DOI: 10.1103/physrevd.101.112004.
- [8] L. Alvarez-Ruso m. fl., "NuSTEC White Paper: Status and challenges of neutrino–nucleus scattering," *Progress in Particle and Nuclear Physics*, årg. 100, s. 1–68, 2018, ISSN: 0146-6410. DOI: 10.1016/j.ppnp.2018.01.006.
- [9] V. Pandey, *Neutrino-Nucleus Scattering Cross Sections at Medium Energies*, 2025. DOI: 10.48550/arXiv.2511.05413. arXiv: 2511.05413 [hep-ph].
- [10] J. Wuttke, "Quasielastic Scattering," i *Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources*, ser. Lecture Notes of the 43rd IFF Spring School, M. Angst, T. Brückel, D. Richter och R. Zorn, utg., Forschungszentrum Jülich, 2012, ISBN: 978-3-89336-759-7. hämtad 30 april 2026. URL: [https://www.fz-juelich.de/ias/jsc/EN/Expertise/High-Q-Club/\\_node.html](https://www.fz-juelich.de/ias/jsc/EN/Expertise/High-Q-Club/_node.html).
- [11] S. Gardiner, "Simulating low-energy neutrino interactions with MARLEY," *Computer Physics Communications*, årg. 269, s. 108 123, 2021, ISSN: 0010-4655. DOI: 10.1016/j.cpc.2021.108123.
- [12] D. Gambacurta, "Gamow-Teller excitations in the Subtracted Second RPA approach," *Journal of Physics: Conference Series*, årg. 2453, nr 1, s. 012017, mars 2023. DOI: 10.1088/1742-6596/2453/1/012017.
- [13] V. Pandey, N. Jachowicz, T. Van Cuyck, J. Ryckebusch och M. Martini, "Low-energy excitations and quasielastic contribution to electron-nucleus and neutrino-nucleus scattering in the continuum random-phase approximation," *Physical Review C*, årg. 92, nr 2, aug. 2015, ISSN: 1089-490X. DOI: 10.1103/physrevc.92.024606.
- [14] J. E. Amaro m. fl., "Neutrino Interactions Importance to Nuclear Physics," i *AIP Conference Proceedings*, AIP, 2009, s. 24–31. DOI: 10.1063/1.3274166.
- [15] The Editors of Encyclopaedia Britannica. "Cross section (physics)," hämtad 20 april 2026. URL: <https://www.britannica.com/science/cross-section-physics>.

- [16] B. Acharya, J. E. Sobczyk, S. Bacca, G. Hagen och W. Jiang, "O16 Electroweak Response Functions from First Principles," *Phys. Rev. Lett.*, årg. 134, nr 20, s. 202 501, 2025. DOI: 10.1103/PhysRevLett.134.202501. arXiv: 2410.05962 [nucl-th].
- [17] J. D. Walecka, *Theoretical Nuclear and Subnuclear Physics*, 2. utg. World Scientific, 2004. DOI: 10.1142/5500.
- [18] T. Kennedy. "Monte Carlo Methods - a special topics course," hämtad 1 maj 2026. URL: <https://math.arizona.edu/~tgk/mc/book.pdf>.
- [19] T. Golan, J. T. Sobczyk och J. Żmuda, "NuWro: the Wrocław Monte Carlo Generator of Neutrino Interactions," *Nuclear Physics B - Proceedings Supplements*, årg. 229-232, 2012. DOI: 10.1016/j.nuclphysbps.2012.09.136.
- [20] PYTHIA. "PYTHIA," hämtad 15 maj 2026. URL: <https://www.pythia.org/>.
- [21] R. Sheposh. "Interpolation (Mathematics)." Tillgänglig via EBSCOhost, hämtad 20 april 2026. URL: <https://research.ebsco.com/linkprocessor/plink?id=cb56b49e-77ec-32a3-b6d8-6812a24ead3c>.
- [22] I. Goodfellow, Y. Bengio och A. Courville, *Deep Learning*. MIT Press, 2016. URL: <http://www.deeplearningbook.org>.
- [23] C. Forssén. "13. Neural networks, Bayesian inference and machine learning," hämtad 12 maj 2026. URL: <https://cforssen.gitlab.io/tif385-book/content/MachineLearning/NeuralNet.html>.
- [24] C. Juszczak. "Quasi elastic neutrino interactions in NuWro." URL: <https://indico.ectstar.eu/event/19/contributions/229/attachments/318/419/juszczak-QEL-in-NuWro-trento-2018.pdf>.
- [25] grufkork. "Working NuWro fork," hämtad 15 maj 2026. URL: <https://github.com/grufkork/nuwro>.