

令和4年度 修士論文

Ξハイパー核分光のための  
機械学習を用いた運動量解析手法の開発



京都大学大学院 理学研究科  
物理学・宇宙物理学専攻 物理学第二教室  
原子核・ハドロン物理学研究室  
江端 健悟





## 概要

2023年より J-PARC ハドロン実験施設にて、 $\Xi$  ハイパー核の高分解能欠損質量分光実験が始まる (J-PARC E70 実験)。この実験では  $K^-$  中間子を炭素標的に照射し、 $^{12}\text{C}$  原子核内部の陽子を ( $K^-, K^+$ ) 反応を介して  $\Xi^-$  粒子に変換することにより、実験室中に  $\Xi$  ハイパー核  $^{12}_{\Xi}\text{Be}$  (コア核  $^{11}\text{B}$  に  $\Xi^-$  が束縛したハイパー核) を生成させる。この  $\Xi$  ハイパー核のエネルギーピーク構造を欠損質量分光によって測定することで、 $\Xi$ -N 間相互作用を調べ、核力の拡張である  $S = -2$  多体系でのバリオン間力の解明を進める。理論計算によって示唆されたピーク構造から、ピーク弁別のために欠損質量を  $2 \text{ MeV}/c^2$  (半値全幅) の分解能で分光測定する。

この実験では、( $K^-, K^+$ ) 反応の欠損質量分布から  $\Xi$  ハイパー核のエネルギーピーク構造を求める。この中で、 $K^+$  中間子の運動量は新設された S-2S 磁気スペクトロメータによって測定される。反応後の  $K^+$  は、S-2S 電磁石を通過し、運動量の大きさによって軌道が曲げられる。S-2S 電磁石の入り口と焦点面付近に設置したドリフトチェンバーから得られた飛跡情報から、運動量を再構成する。このためにエネルギー校正手法を含めた運動量解析手法の開発と実装を行う必要がある。運動量を再構成する場合、ルンゲクッタ法 (RK 法) や逆輸送行列を用いた再構成方法 (BTM 法) が従来手法としてよく利用される。ただし、RK 法においては、運動量再構成の後に現象論的な関数による補正が用いられることが多い。また BTM 法に関しても、アクセプタンスが広い場合困難となる。そこで本修論では、従来手法のように運動方程式やビーム光学に基づく方程式といった決まった型を取らない、自由度の高い機械学習 (ニューラルネットワーク) を再構成手法に採用する。これにより、例えば RK 法にて問題となっていた現象論的な関数による補正も自動的に取り払われることが予想される。また、従来手法では取りきれなかった飛跡情報と運動量の高次の相関を掬いとすることで、より高効率で高分解能な運動量解析が期待される。さらに、従来手法では入出力を飛跡情報と運動量と固定されていたものに対して、ドリフトチェンバーの各ワイヤー面での local hit 情報などを入力としてトラッキングなしに運動量構成を行う、といった機械学習の自由度の高さゆえの解析の拡張性の高さも利点としてあげられる。

本研究では機械学習を用いた運動量解析手法を開発するために、GEANT4 によるシミュレーションデータを用いた。シミュレータ上の磁場マップには実測した QQD 磁石の磁場を再現した磁場計算による結果を用いた。シミュレータ上に S-2S 磁気スペクトロメータを設置し標的位置から  $K^+$  を照射する。この  $K^+$  を S-2S の上下流に設置した検

出器で捉え、位置  $(x, y)$  や角度  $(u = \partial x / \partial z, v = \partial y / \partial z)$  といった飛跡情報を取得する。なお  $x$  とは偏向電磁石で偏向される方向で、 $z$  とはビーム軸方向である。これら飛跡情報を訓練データとして、反応直後の  $K^+$  の運動量  $p_S^0$  を教師データとして用いて学習を行った。飛跡情報から運動量を再現できるような機械学習モデルを構築し、基礎的な運動量解析フレームワークを開発した。

その結果、機械学習モデルで計算された運動量分解能は、半値全幅で  $1.0 \times 10^{-3}$  程度となり、 ${}_{12}^{\text{Be}}$  のミッシングマス分解能としては  $2.3 \text{ MeV}/c^2$  が達成された。これにより  ${}_{12}^{\text{Be}}$  のピークの弁別も確認された。

さらに、構築したモデルの校正可能性についての検証を行った。シミュレーション空間上の検出器セットアップは必ずしも現実のセットアップと同等のものであるとは限らない。したがって本実験中に取得した Beam through、 $\Xi^-$ 、 $\Sigma^-$  事象を用いた校正が必要となるであろう。この校正の可能性を検証するために、シミュレーション上で意図的にずらしたセットアップを作成し、取得した Beam through と  $\Xi^-$ 、 $\Sigma^-$  データによってモデルの校正が可能かを確認した。上流検出器を  $x$  方向に 1 mm ずらしたセットアップで検証を行った結果、シフトしていたミッシングマス分布が、Beam through と  $\Xi^-$ 、 $\Sigma^-$  データによりズレを抑えることができた。これにより、機械学習による運動量解析モデルのエネルギー校正可能性が確認された。

# 目次

<b>第 1 章</b>	<b>導入</b>	<b>4</b>
1.1	ハイパー核研究 . . . . .	4
1.2	Ξ ハイパー核研究 . . . . .	8
<b>第 2 章</b>	<b>J-PARC E70 実験</b>	<b>12</b>
2.1	実験の物理的動機 . . . . .	12
2.2	実験セットアップ . . . . .	15
2.2.1	J-PARC . . . . .	15
2.2.2	K1.8 ビームラインスペクトロメータ . . . . .	17
2.2.3	実験手法 . . . . .	18
2.2.4	S-2S 磁気スペクトロメータ . . . . .	20
2.2.5	粒子識別法 . . . . .	26
2.2.6	アクティブファイバー標的 . . . . .	28
2.3	期待される物理的結果 . . . . .	29
<b>第 3 章</b>	<b>運動量再構成手法</b>	<b>33</b>
3.1	従来解析手法 . . . . .	33
3.1.1	ルンゲクッタ法 . . . . .	33
3.1.2	逆輸送行列による手法 . . . . .	34
3.2	本研究の動機 . . . . .	35
3.3	先行研究	
	機械学習を用いた J-PARC E05 実験の運動量解析 . . . . .	36
3.4	本研究の目的 . . . . .	37
3.5	運動量解析への機械学習の導入 . . . . .	38
3.5.1	機械学習 . . . . .	38

	3.5.2	誤差逆伝播法 . . . . .	42
	3.5.3	ミニバッチ学習 . . . . .	43
	3.5.4	過学習 . . . . .	43
<b>第 4 章</b>		<b>機械学習モデルの構築と最適化</b>	<b>47</b>
4.1		モデル構築に用いるシミュレーションデータ . . . . .	47
4.2		機械学習ライブラリ . . . . .	51
4.3		機械学習モデルの構築 . . . . .	51
	4.3.1	データの前段階処理 . . . . .	52
	4.3.2	ノード数の最適化 . . . . .	54
	4.3.3	隠れ層の数の最適化 . . . . .	55
	4.3.4	活性化関数の最適化 . . . . .	56
	4.3.5	最適化関数の最適化 . . . . .	60
	4.3.6	損失関数の最適化 . . . . .	65
4.4		訓練データの最適化 . . . . .	66
	4.4.1	訓練データの次数最適化 . . . . .	66
	4.4.2	訓練データの統計量 . . . . .	67
4.5		最適化した訓練パラメータセットの評価 . . . . .	69
4.6		$\cos \theta$ を出力層に追加した機械学習モデルの構築とその評価 . . . . .	73
<b>第 5 章</b>		<b>機械学習による S-2S の性能評価</b>	<b>81</b>
5.1		運動量分解能の運動量依存性 . . . . .	81
5.2		多重散乱の運動量分解能への影響 . . . . .	83
5.3		検出器分解能の運動量分解能への影響 . . . . .	84
5.4		ミッシングマス分解能の評価 . . . . .	87
	5.4.1	$\Xi^-$ 、 $\Sigma^-$ のミッシングマス分解能評価 . . . . .	88
	5.4.2	$^{12}\text{Be}$ のミッシングマス分解能および分布の評価 . . . . .	89
<b>第 6 章</b>		<b>local hit 情報を訓練データとして採用した運動量再構成手法</b>	<b>97</b>
6.1		教師データに $p_S$ と $\cos \theta$ を与えた場合 . . . . .	98
6.2		教師データに飛跡情報を加えた場合 . . . . .	103
<b>第 7 章</b>		<b>機械学習モデルの校正手法</b>	<b>106</b>
7.1		Beam through、 $\Xi^-$ ピークデータを用いた校正 . . . . .	107

---

7.2	Beam through、 $\Xi^-$ 、 $\Sigma^-$ ピークデータを用いた校正 . . . . .	115
第 8 章	結論	121
第 9 章	今後の展望	123
9.1	機械学習モデルの拡張 . . . . .	123
9.2	本解析手法の運用について . . . . .	123
謝辞		125
参考文献		127

# 第 1 章

## 導入

### 1.1 ハイパー核研究

身の回りに存在する通常の原子核は、アップクォーク (u) とダウンクォーク (d) のみで構成された陽子と中性子が結合したクォーク多体系である。通常の原子核に関しては、これまで約 4000 種の核種について多くの散乱実験などを基にして核力、核構造のモデルの構築がなされてきた。これに対して u, d クォーク以外のクォークを含んだバリオン間相互作用に関しても近年研究が盛んに行われている。このような研究は、核力の拡張としてのバリオン間相互作用、量子色力学 (Quantum ChromoDynamics, QCD) の深い理解につながるであろう。中でも特に、u, d クォークの次に軽いストレンジクォーク (s) を対象としたバリオン間相互作用の研究が進められている。ストレンジ (s) クォークを含んだバリオンは「ハイペロン」と呼ばれる。ラムダ ( $\Lambda$ ) 粒子、シグマ ( $\Sigma$ ) 粒子、グザイ ( $\Xi$ ) 粒子といった粒子がハイペロン (Y) に分類される。ハイペロンが束縛した原子核はハイパー核と呼ばれる。u, d, s クォークから 3 体系としてのバリオンを構成する場合、 $3 \otimes 3 \otimes 3 = 10 \oplus 8 \oplus 8 \oplus 1$  より 10 重項と 8 重項に分類される (図 1.1)。バリオン 8 重項間の相互作用は、 $8 \otimes 8 = 27 \oplus 10^* \oplus 10 \oplus 8_a \oplus 8_s \oplus 1$  になる。27 重項、 $10^*$  重項以外の相互作用に関してはストレンジネスを含むことによってはじめて現れる相互作用である。格子 QCD 計算によると、 $V(1)$  が強い引力、 $V(10^*)$  と  $V(27)$  が引力 (NN 相互作用) を示し、 $V(10)$  には斥力が見られるというような大きな違いが見られる (図 1.2)。このような計算と、実験情報を比較してバリオン間相互作用の理解を深めることが重要である。

ハイパー核研究を進めることは、バリオン間力の解明につながるだけではない。例えば中性子星のような極限天体の構造理解にもつながる。中性子星内部のような超高密度環境においては、ハイペロンが発現すると予想される。ハイペロンの発現によって内部の状態

方程式は柔化し、中性子星の質量は太陽質量の 1.6 倍以下に制限されると予想されていた。ところが、太陽質量の 2 倍の中性子星が見つかり、想定より硬い天体であることがわかった [2]。これはハイペロンと核子の間の相互作用に斥力的な成分があることが解決の一つであることが理論面から示唆されている。この候補としては、 $\Lambda NN$  3 体力が挙げられる。ハイペロン・ハイパー核の研究はこのような高密度核物質の物理にも関連する。

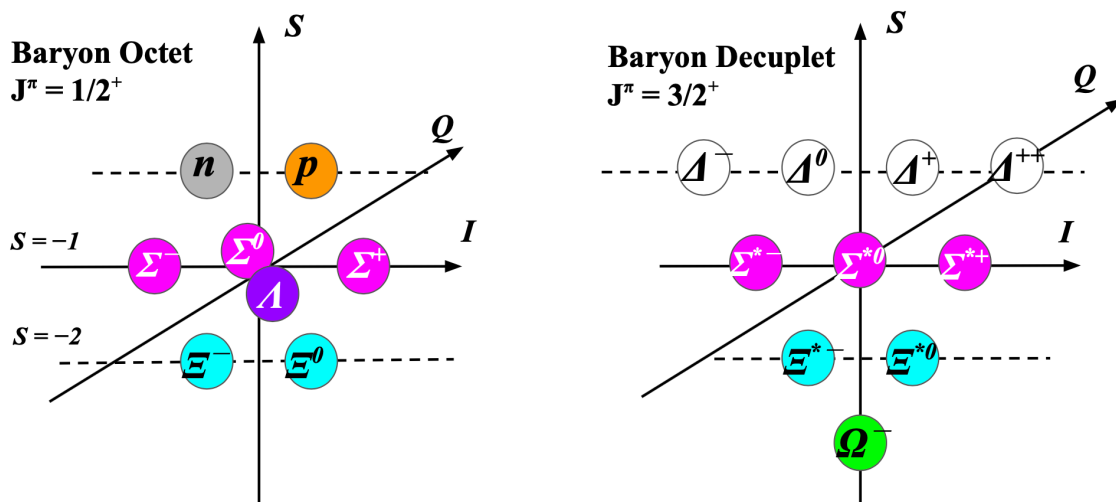


図 1.1: バリオン 8 重項と 10 重項。

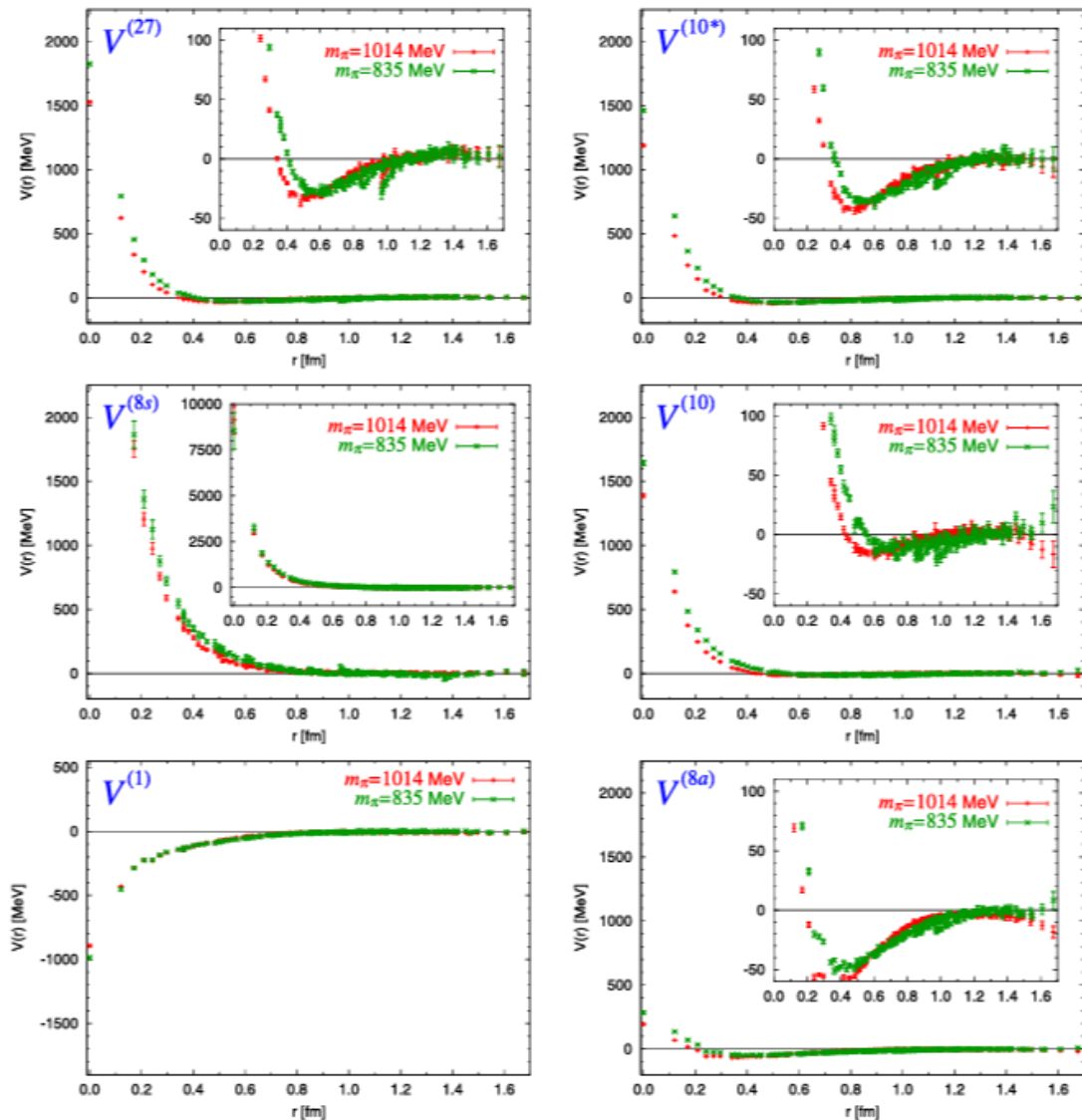


図 1.2: バリオン間相互作用。図は [1] より。

ハイペロン-核子間 (YN) 相互作用やハイペロン-ハイペロン間 (YY) 相互作用を調べる場合、直接的にハイペロンと核子の散乱事象を捉えるのが有用である。近年では J-PARC にて  $\Sigma p$  散乱 [3, 4] の測定 (J-PARC E40 実験) が行われた。しかしハイペロンの寿命が極端に短い ( $\tau_{\Lambda} = 263$  ps,  $\tau_{\Xi} = 164$  ps 程度 [5]) ため、散乱実験による直接観測は容易ではない。そこで、ハイペロンを原子核に束縛させハイパー核を実験室中に作り、その核構造を調べることによって相互作用を間接的に調べる手法が用いられてきた。

s クォークを1つだけ含んだ  $S = -1$  多体系 ( $\Lambda$  ハイパー核や  $\Sigma$  ハイパー核) に関して



は、複数の実験施設での分光実験により  $A = 208$  までの約 40 核種について多くの知見が得られてきた。 $\Lambda$  ハイパー核を生成するためには、主に  $(K^-, \pi^-)$ 、 $(\pi^+, K^+)$ 、 $(e, e'K^+)$  反応が主に用いられる。日本の KEK の 12 GeV シンクロトロン施設や J-PARC、米国の BNL-AGS では、 $\pi$  中間子や K 中間子ビームでハイパー核を生成し、その励起スペクトルを分光測定する実験が行われてきた。JLab (米国) や MAMI (ドイツ) では、電子ビームを用いた仮想光子を介した  $(e, e'K^+)$  反応による  $\Lambda$  ハイパー核分光実験が盛んである。これら施設での測定から、例えば  $\Lambda$  のポテンシャルの深さが 29 MeV 程度であることがわかった。また、ゲルマニウム検出器を用いた  $\gamma$  線分光により、 $\Lambda$  ハイパー核の準位が精密に測定されている [6]。一方、 $\Sigma$  ハイパー核に関しては、 $(K^-, \pi^-)$  反応による  $\frac{4}{2}\text{He}$  の一例のみ束縛状態が見つかる [7] にとどまり、 $\Sigma$  のポテンシャルは斥力的であることがわかっている。

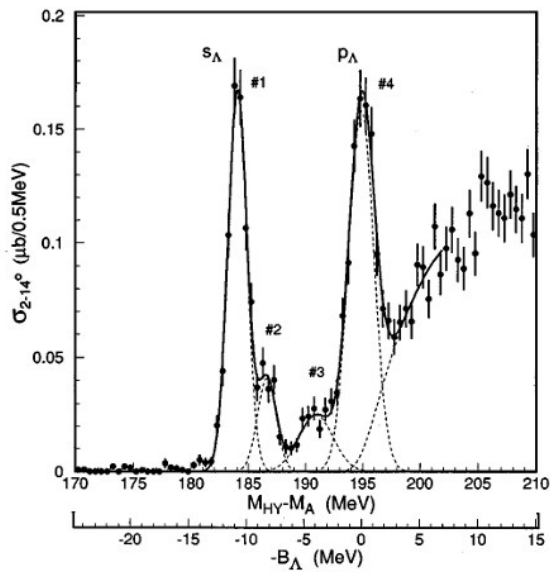


図 1.3: KEK-PS E369 分光実験で得られた  $^{12}_{\Lambda}\text{C}$  スペクトル。 [8]

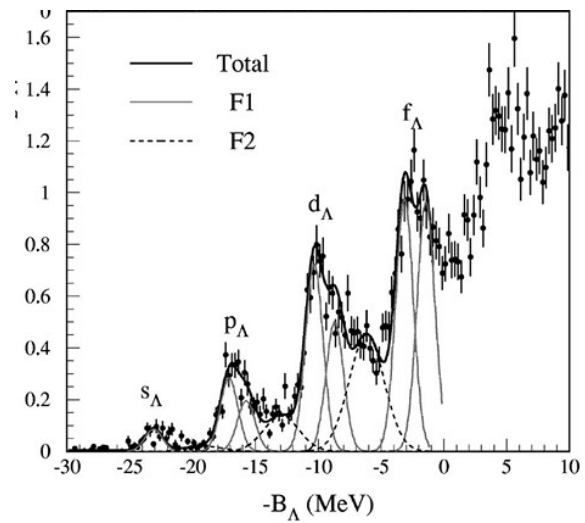


図 1.4: KEK-PS E369 分光実験で得られた  $^{89}_{\Lambda}\text{Y}$  スペクトル。 [9]

## 1.2 E ハイパー核研究

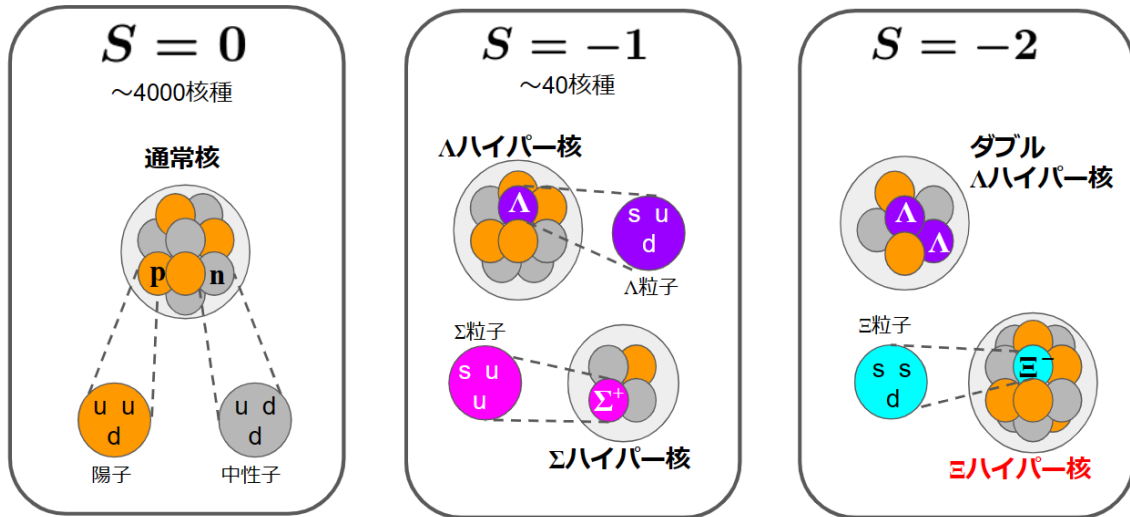


図 1.5: 様々なハイパー核。s クォークを一つだけ含んだものは  $S = -1$  多体系、二つ含んだものは  $S = -2$  多体系と呼ばれる。

一方、s クォークを 2 つ含んだ  $S = -2$  多体系 (図 1.5) (ダブル  $\Lambda$  ハイパー核や  $\Xi$  ハイパー核) は、 $S = -1$  の系で見られなかった  $\Lambda\Lambda$ ,  $\Xi N$  間相互作用が現れる。バリオン間相互作用の拡張を考える上で、 $S = -2$  多体系の実験は非常に重要な意味をもつ。これまでに原子核乾板などを使った探索で  $\text{NAGARA}(\Lambda\Lambda\text{He})$  (図 1.8) [10],  $\text{KISO}(\Xi^- + {}^{14}\text{N})$  イベント [11] などの特徴的な事象が報告されている。近年でも J-PARC E07 実験での探索で、MINO [12], IRRAWADDY, IBUKI [13] 事象などが観測された。E07 実験の原子核乾板のハイパー核事象探索には、機械学習の一種である敵対的生成ネットワーク (Generative Adversarial Network, GAN [14]) による画像認識が用いられ、これによりダブル  $\Lambda$  ハイパー核事象などが次々と発見されている。これら原子核乾板による測定から  $\Xi N$  相互作用は引力的であることが分かった。また LHC での重イオン衝突実験 (ALICE) においても、Femtoscopy によって陽子- $\Xi$  間相互作用は強い引力を持つことが示唆されている [20]。

ただし、原子核乾板によるデータは、統計精度が不足しており束縛エネルギーや自然幅を精度よく決定できていないのが現状である。また、 $\Xi$  ハイパー核では、 $\Xi$  が  $\Lambda$  と違い

アイソスピンを持つため、 $\Lambda$  ハイパー核には存在しなかったアイソスピン依存力が発現する。またスピン依存相互作用により、複雑なスペクトルが得られると予想されている。これらのスペクトルを弁別するためには、高統計で高分解能な分光測定が必要である。

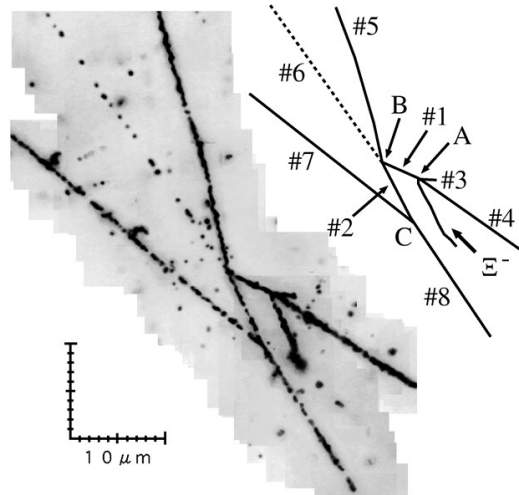
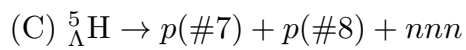
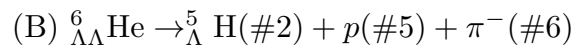
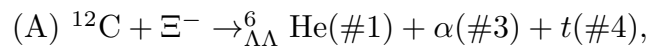


図 1.6: KEK でのエマルジョン実験による NAGARA 事象の写真と模式図。



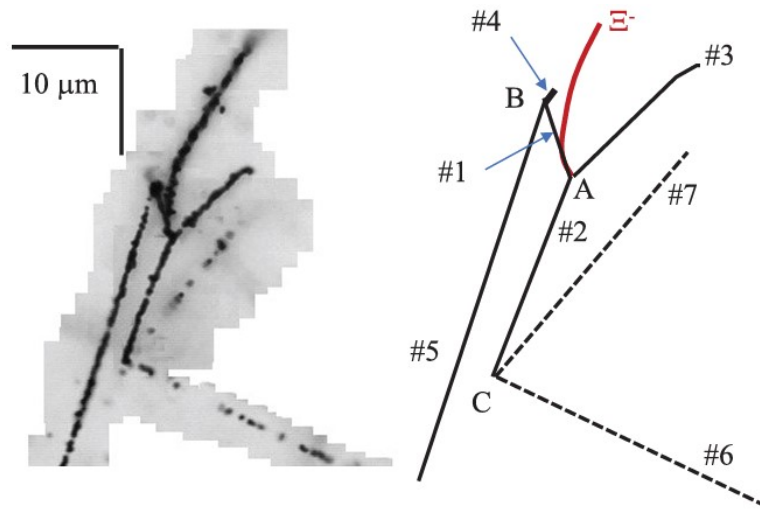
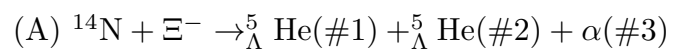


図 1.7: J-PARC E07 実験でのエマルジョン実験による IRRAWADDY 事象の写真と模式図。



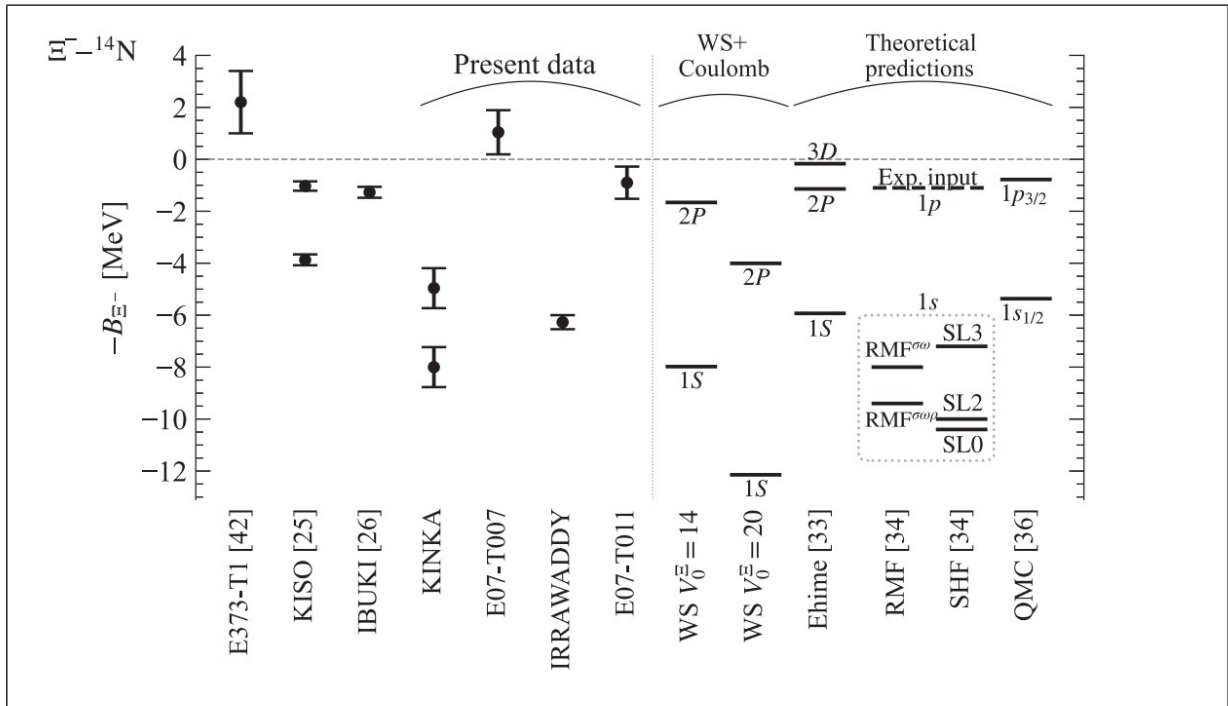


図 1.8: エマルジョン実験 (KEK E373-T1[15], KISO[11], IBUKI[13], J-PARC E07[19]) で得られた  $E^-$ - $^{14}\text{N}$  系の束縛エネルギー (左) と、Woods-Saxon、Coulomb ポテンシャルのもとで理論計算 (Ehime[16], RMF[17], SHF, QMC[18]) で与えられた束縛エネルギー (右)。  
図は [19] より。

## 第2章

# J-PARC E70 実験

### 2.1 実験の物理的動機

( $K^-$ ,  $K^+$ ) 反応を用いた  $\Xi$  ハイパー核の分光実験は KEK (KEK E224[3]) や BNL (BNL E885 [4]) にて過去に行われてきた。これらの実験では準自由散乱の形状から  $\Xi$  のポテンシャルが  $V_{0\Xi} < 14$  MeV であることや、 $\Xi$  核の生成断面積が  $42 \pm 5$  nb/sr ( $\theta < 14$  deg),  $89 \pm 14$  nb/sr ( $\theta < 8$  deg) であることがわかった。ところが質量欠損の分解能はそれぞれ 22 MeV, 14 MeV (FWHM) 程度しかなく、束縛領域でのエネルギーピーク構造を捉えるには十分な精度ではなく、また統計量も十分ではなかった (図 2.1)。2015 年には J-PARC E05 実験において、SKS 電磁石を用いた  $\Xi$  ハイパー核,  ${}_{\Xi}^{12}\text{Be}$  ( ${}^{11}\text{B}$  に  $\Xi^-$  が束縛した核) 分光をおこなったが分解能は 6 MeV 程度でこちらも十分なものではなかった。スピン依存力をも含めた  $\Xi N$  相互作用を詳細に調べるには、エネルギーピーク構造をより精度よく調べるのが重要である。例えば ESC08a[28] と呼ばれる中間子交換モデルに基づく相互作用モデルで計算した理論計算 (DWIA 計算)[23] に基づき実験的誤差を考えると、エネルギーピーク構造を弁別して捉えるには  $\sim 2$  MeV/ $c^2$  以下の分解能が必要と予想されている (図 2.2, 2.3)。 $\Xi$  ハイパー核  ${}_{\Xi}^{12}\text{Be}$  のコアの励起があった場合でも、弁別可能であると見積もられている。

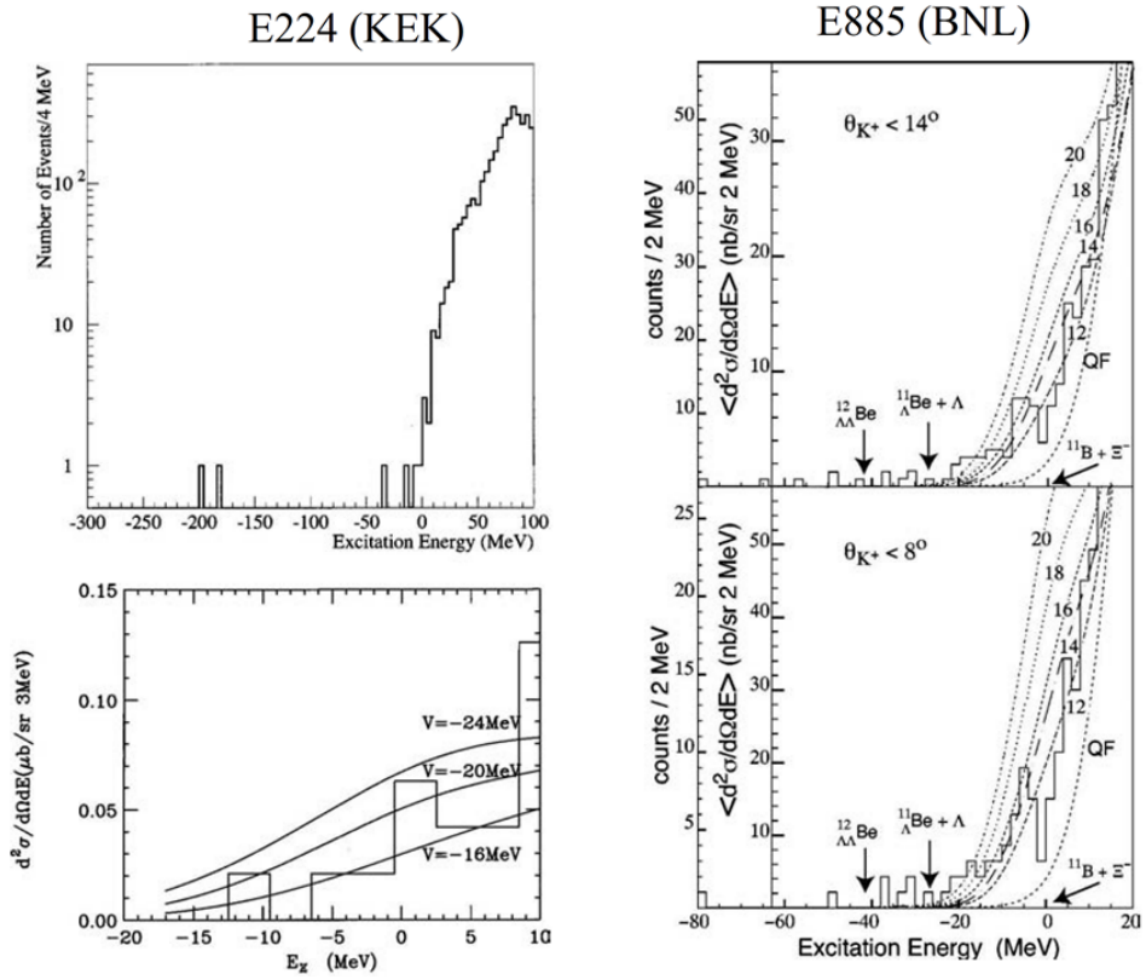


図 2.1: KEK E224[21], BNL E885[22] 実験でのミッシングマススペクトル。

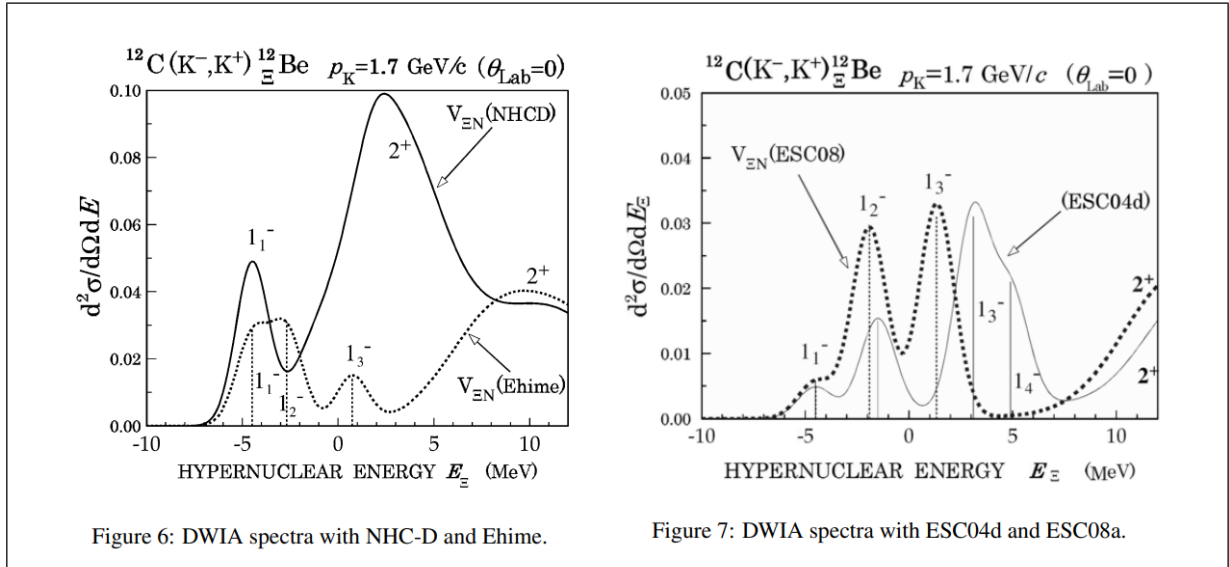


図 2.2: 様々な理論モデル (Ehime[16, 24], NHCD[25, 26], ESC04d[27], ESC08a[28]) を基にして計算された  $^{12}_{\Xi}\text{Be}$  のピーク構造。図は、[23] より。

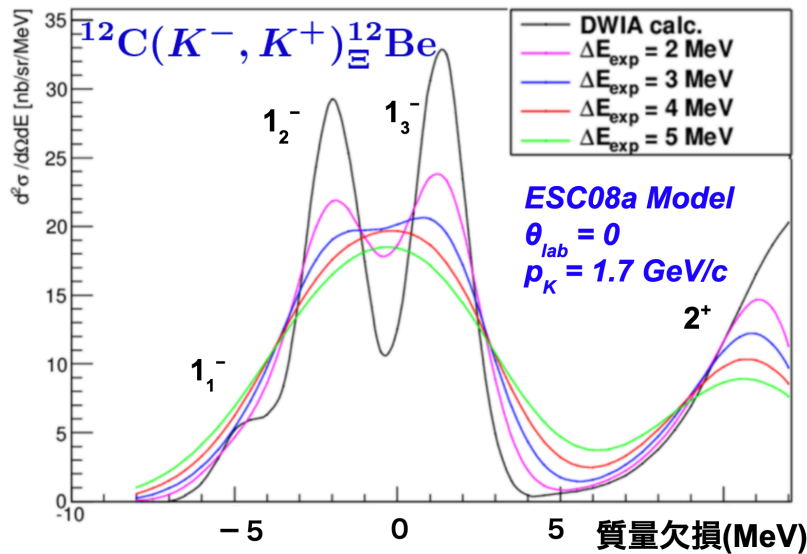


図 2.3: 理論計算 (ESC08a)(図中：黒) を基に実験的誤差 (2–5 MeV) を与えた場合のミッシングマスペクトル。2 MeV 程度でピークが弁別可能であることが分かる。

このため、我々は J-PARC E70 実験において  $\Xi$  ハイパー核 ( $^{12}_{\Xi}\text{Be}$ ) のミッシングマスを  $2 \text{ MeV}/c^2$  以下の精度で測定することを目指している。 $^{12}_{\Xi}\text{Be}$  は、 $^{11}\text{B}$  に  $\Xi^-$  が束縛した、陽子 5 つ、中性子 6 つ、 $\Xi^-$  ハイペロン 1 つから構成されるハイパー核である。(こ



の場合、電荷量が  $Z = 5 + 0 - 1 = 4$  となるため、化学的表記としてベリリウム (Be) となることに注意せよ。) これにより、束縛エネルギーピークを 100 keV, 崩壊幅を 500 keV 程度の精度で決定する。分光実験で得られたエネルギーピーク構造によって、例えば  $\Xi$  の光学ポテンシャルのパラメータ値に大きく制限が与えられるなど、 $\Xi N$  間相互作用の解明に大きく貢献するであろう。

## 2.2 実験セットアップ

### 2.2.1 J-PARC

現在我々は、J-PARC K1.8 ビームラインにおいて新たに  $\Xi$  ハイパー核の高分解能欠損質量分光実験の準備を行なっている (J-PARC E70 実験)[29]。この実験は、大強度陽子シンクロトロン加速器施設 J-PARC 内のハドロン実験施設で行われる。まず、J-PARC では、イオン源で発生させた負水素イオンを線形加速器 (図 2.4 中、LINAC) で 400 MeV まで加速させる。次に RCS と呼ばれる初段のシンクロトロンに入射される。この入射時において負水素イオンビームは荷電交換用の炭素薄膜を通して陽子ビームに転換され、3 GeV まで加速される。最後にメインリング (MR) のシンクロトロンで 30 GeV まで加速される (図 2.4)。この加速した陽子をハドロン実験施設内の金 (Au) 標的 (図 2.5, T1 target) に衝突させ、2 次粒子 ( $\pi$ , K 中間子) を生成する p+Au 反応を起こす。ハドロン実験施設では、この反応後に生成される  $\pi$ , K 中間子を 2 次粒子ビームとして用いて、様々なハドロン実験を行っている。

生成後の 2 次粒子は偏向電磁石によって偏向されたのち、中間像スリットを通過する。次に 1 段目の静電セパレータ (ESS1) と質量スリット (MS1) を通過する。ESS では、通過する粒子に高い電場を印加させることにより、粒子の質量により粒子弁別をおこなう。質量スリット通過後には偏向電磁石が設けてあり、ここで K1.8 ビームライン、K1.8BR ビームラインへの粒子の振り分けがなされる。K1.8 ビームラインへ輸送されるビームはさらに 2 段目の粒子弁別にかかる。静電セパレータ (ESS2) と質量スリット (MS2) を通過する [30]。

これらの 2 段の粒子弁別装置により、K1.8 ビームラインでの K 中間子の純度  $K/\pi$  比は 0.8 が達成される。高純度・大強度の K 中間子ビームはこうした粒子弁別機構を介して、磁気スペクトロメータにより、K1.8 ビームラインへ輸送される。今回の実験では、ビームラインのアクセプタンス中で  $\Xi^-$  の生成率が最も高い 1.8 GeV/c の運動量を持つ  $K^-$  中間子ビームを輸送し、標的に照射させる (図 2.5)。

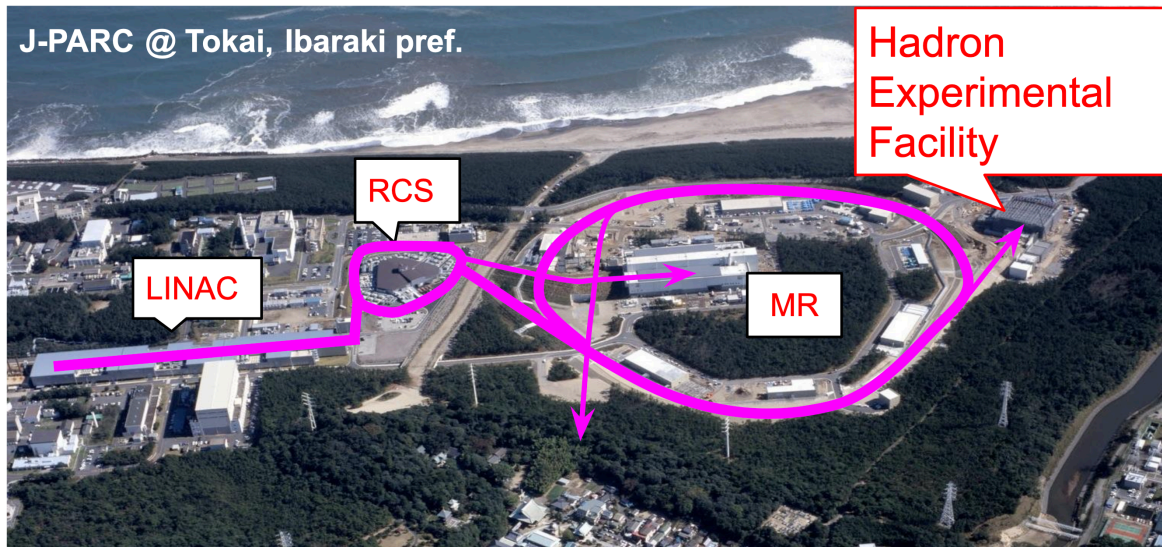


図 2.4: J-PARC の鳥瞰写真。3 段の加速器から構成される。初段の線形加速器 (LINAC) で、負水素イオンを 400 MeV まで加速させ、RCS と呼ばれるシンクロトロンに輸送する。ここで荷電交換膜により陽子ビームに変換し、RCS で 3 GeV まで加速させる。さらに MR と呼ばれるシンクロトロンに輸送する。ここで 30 GeV まで加速された陽子をハドロン実験施設まで輸送し様々な実験に利用される。

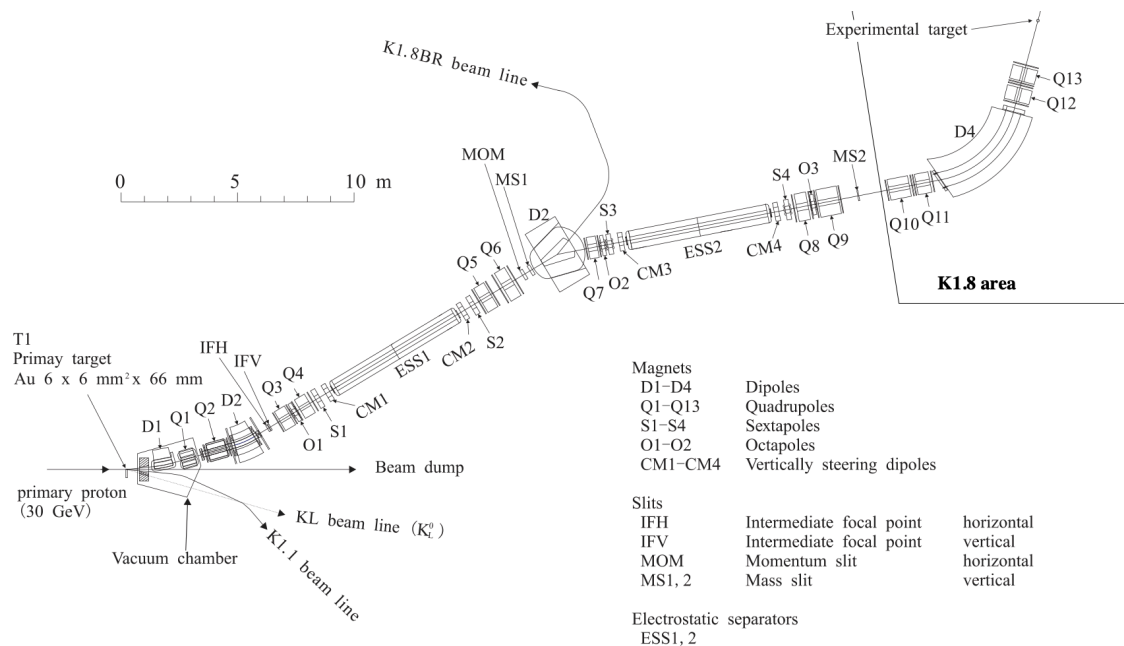


図 2.5: K1.8 ビームラインの模式図。30 GeV シンクロトロンにより加速された陽子はハドロン実験施設内の図中の左より輸送され、金 (Au) 標的と衝突させられる。このときの反応で生じた 2 次粒子は各ビームラインに輸送され、様々なハドロン実験で利用される。

E70 実験は図中右上の K1.8 area で行われる。[30]

### 2.2.2 K1.8 ビームラインスペクトロメータ

先に述べたように静電セパレータで弁別された  $K^-$  中間子ビームは、ビームラインスペクトロメータで輸送され、標的に照射される。ビームラインスペクトロメータは、K1.8 実験エリアに設置されており QQDQQ (収束電磁石 4 台 + 偏向電磁石 1 台) で構成される。ビームラインスペクトロメータの運動量測定には、後述の逆輸送行列を用いた手法が用いられている。

ビームライン側のその他検出器群に関しては、スペクトロメータの上流・下流にトリガータイミングなどを決定するためのホドスコープ (BH1, BH2)[31] が設置される。また、 $K^-$  の運動量測定のために、上流にはファイバートラッカー (BFT)[32]、下流には 2 台のドリフトチェンバー (BC3, 4)[33] といった位置検出器が設置されている。さらに下流の標的の直前には  $\pi^-$  と  $K^-$  を識別するためのエアロゲルチェレンコフカウンター (BAC)[34] が設置される (表 2.1)。

表 2.1: ビームライン検出器群の性能

検出器	構成	有感領域 (mm)	読み出し	備考
BFT	$xx'$	$160^x \times 80^y$	MPPC $\times 160 \times 2$	
BC3	$xx'vv'uu'$	$192^x \times 100^y$	ASD card	
BC4	$uu'vv'xx'$	$192^x \times 100^y$	ASD card	
BH1	11 セグメント	$170^x \times 66^y$	PMT $\times 11 \times 2$	
BH2	8 セグメント	$118^x \times 60^y$	PMT $\times 8 \times 2$	$\sigma = 30 - 40$ ps
BAC	シリカエアロゲル	$170^x \times 70^y \times 46^z$	PMT $\times 2$	$\pi$ 抑制 ( $n = 1.03$ )

### 2.2.3 実験手法

ビームラインに沿って輸送された  $K^-$  は炭素標的に照射される (図 2.6, 2.7)。この時、標的中での ( $K^-, K^+$ ) 反応によって  $\Xi$  ハイパー核 ( ${}_{\Xi}^{12}\text{Be}$ ) が生成される。反応前後の粒子である、 $K^-, K^+$  中間子の運動量  $p_B, p_S$  を測定し、エネルギー・運動量保存則から、 $\Xi$  ハイパー核の質量  $M_{\text{HYP}}$  を測定する (式 2.1)。この手法をミッシングマス (質量欠損) 法と呼ぶ。この欠損質量のスペクトルから  $\Xi$  ハイパー核の探索を行う (図 2.8)。

$$\begin{aligned}
 M_{\text{HYP}} &= \sqrt{E_{\text{HYP}}^2 - p_{\text{HYP}}^2} \\
 &= \sqrt{(E_B + M_{\text{target}} - E_S)^2 - p_B^2 - p_S^2 + 2p_B p_S \cos \theta} \quad (2.1)
 \end{aligned}$$

$K^-$  の運動量  $p_B$  は、K1.8 ビームラインに既に設置されているスペクトロメータによって測定される。一方  $K^+$  の運動量  $p_S$  の測定には新たに開発・設置された S-2S 磁気スペクトロメータを用いる。これらの運動量分解能はそれぞれ  $\Delta p/p = 3.3 \times 10^{-4}$ ,  $6.0 \times 10^{-4}$  (FWHM、設計値) 程度と見積もられている。運動量の測定手法としては、まずスペクトロメータが生み出す磁場で荷電粒子の軌道を曲げる。スペクトロメータの入り口と電磁石の焦点面付近に、計 5 台のドリフトチェンバーを設置する。ドリフトチェンバーで検出した粒子のヒット位置で軌道を再構成し、この軌道の曲率から運動量を逆算して求める (図 2.6)。

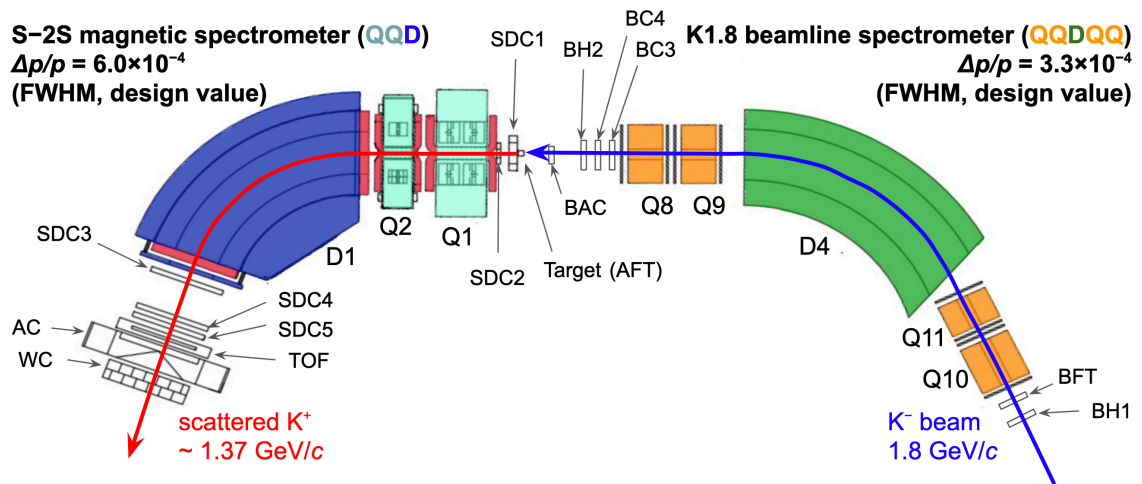


図 2.6: J-PARC E70 実験の K1.8 実験エリアでのセットアップ図。ビームラインスペクトロメータで輸送された  $1.8 \text{ GeV}/c$  の運動量を持つ  $K^-$  中間子ビームを炭素 12 標的に照射する。標的中で  $^{12}\text{C}(K^-, K^+)$  反応を起こし、三ハイパー核  $^{12}_{\text{Be}}$  を生成する。反応後の散乱粒子である  $K^+$  は、下流の S-2S によって運動量が測定される。

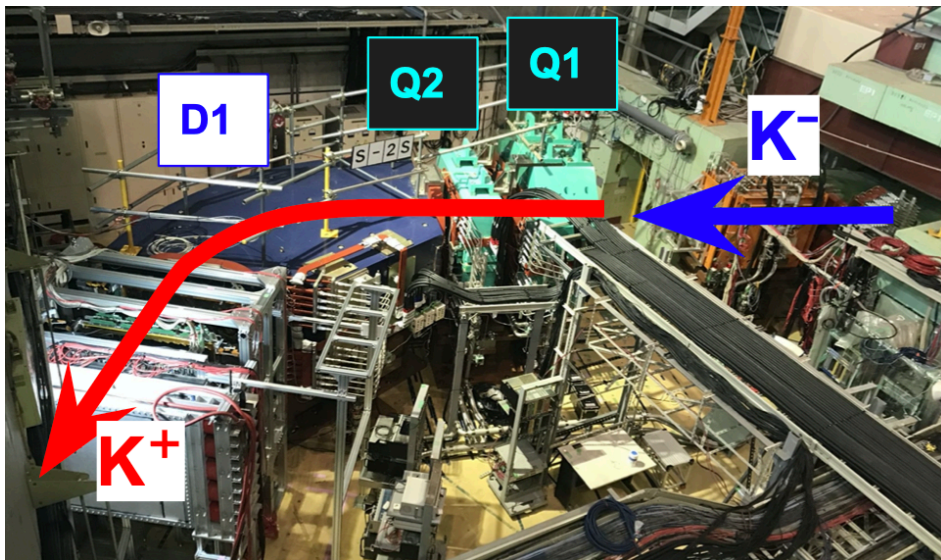


図 2.7: K1.8 エリアでの S-2S の写真。

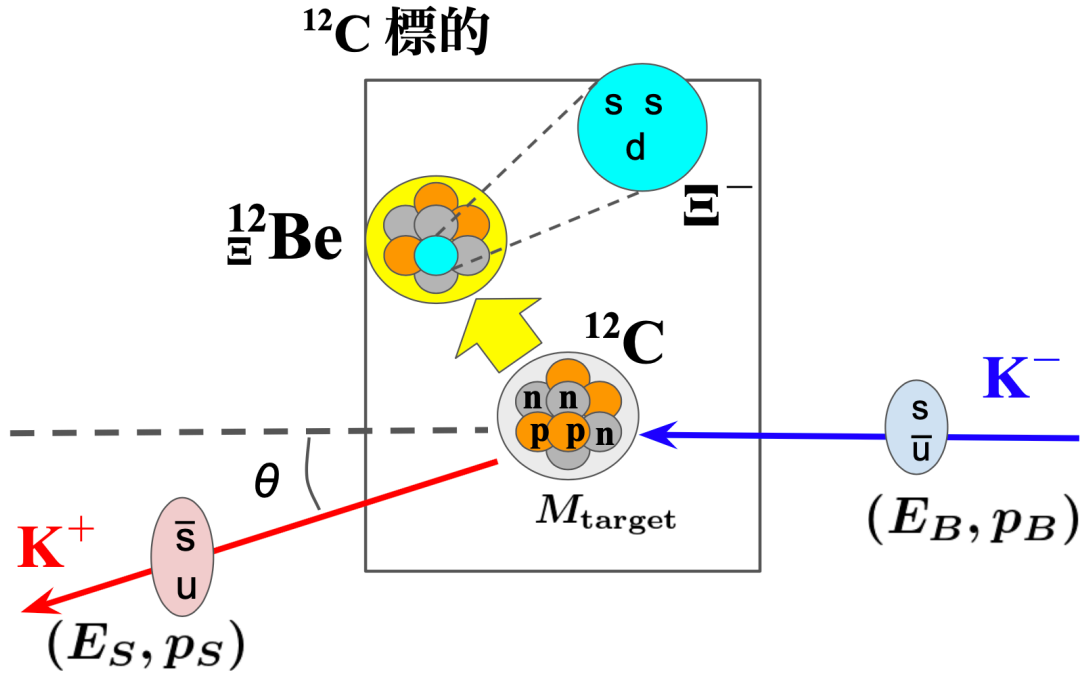


図 2.8: 炭素標的中の  $^{12}\text{C}(K^-, K^+)_{\Xi}^{12}\text{Be}$  反応の模式図。炭素原子核中の陽子を  $\Xi^-$  粒子に変換させることで  $\Xi$  ハイパー核を生成する。 $^{12}\text{C} + K^- \rightarrow K^+ + {}_{\Xi}^{12}\text{Be}$  ( $^{11}\text{B} + \Xi^-$ )。

#### 2.2.4 S-2S 磁気スペクトロメータ

S-2S (Strangeness  $-2$  Spectrometer) 電磁石は、双極（偏向）磁石 1 台と四重極（収束）磁石 2 台から構成される常伝導磁石系である。S-2S は、 $\Delta p/p = 6 \times 10^{-4}$  (FWHM) になるように設計されている。先行実験の J-PARC E05 実験で使用された SKS 電磁石と比較すると、アクセプタンスが小さい (SKS: 110 msr, S-2S: 55 msr) もの、 $R_{11}$  (magnification,  $\langle x|x \rangle$ ) と  $R_{12}$  ( $\langle x|\theta \rangle$ ) が小さく、 $R_{16}$  (dispersion,  $\langle x|p \rangle$ ) が大きいため、より高分解能な運動量測定が可能となる (式 2.2)。

$$\Delta p/p = \frac{\sqrt{(1 + R_{11}^2)\sigma_x^2 + R_{12}^2\sigma_\theta^2}}{R_{16}} \quad (2.2)$$

S-2S の性能と、S-2S と J-PARC E05 実験で用いられた SKS 電磁石のパラメータを下の表に示した (表 2.2)。



表 2.2: S-2S 磁石の性能

Q1 (図 2.9)	収束電磁石
y-conversion	
最大磁場勾配	9.06 T/m
磁極長	880 mm
コイル巻き数	82 turns/coil
定格	2500 A - 145 V
荷重	37 t
Q2 (図 2.10)	収束電磁石
x-conversion	
最大磁場勾配	4.91 T/m
磁極長	540 mm
コイル巻き数	32 turns/coil
定格	2500 A - 64 V
荷重	12 t
D1 (図 2.11)	偏向電磁石
磁場	1.5 T
曲率半径 (中心軌道)	3000 mm
偏向角度	70 deg
磁極長	3665 mm
コイル巻き数	84 turns/coil
定格	2500 A - 200 V
荷重	86 t

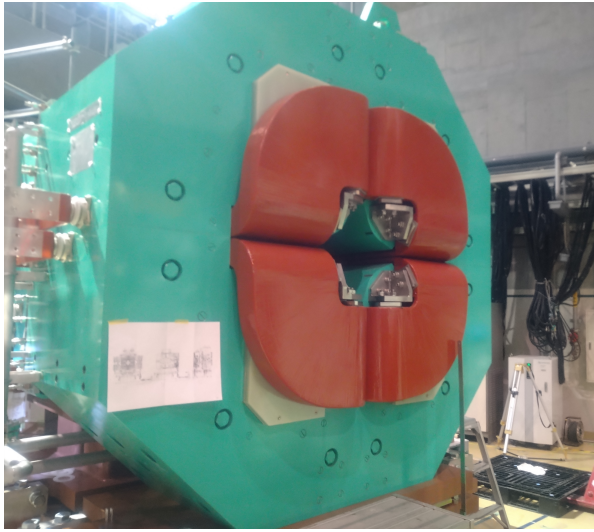


図 2.9: Q1 電磁石の写真。



図 2.10: Q2 電磁石の写真。

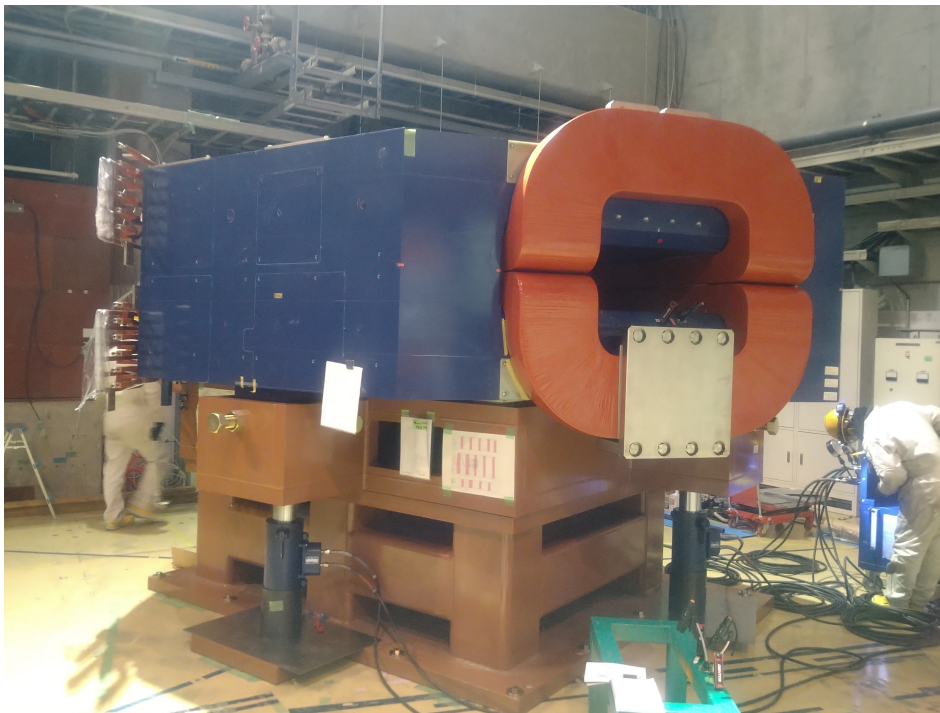


図 2.11: D1 電磁石の写真。



表 2.3: S-2S と SKS の比較

スペクトロメータ	S-2S (E70)	SKS (E05)
configuration	QQD	D
アクセプタンス (msr)	55	110
$\Delta p/p$ (設計値)	$6 \times 10^{-4}$	$3 \times 10^{-3}$
ミッシングマス分解能 (MeV/ $c^2$ )	$\lesssim 2$	6
曲げ角 (deg)	70	56.7
$\rho$ (m)	3	1.85
$\langle x x \rangle$	-0.727	0.549
$\langle x \theta \rangle$	0	0.209
$\langle x \delta \rangle$	-3.847	-0.834

この S-2S は、次節にて述べる  $\Xi$  ハイパー核の精密分光実験 (E70 実験) だけでなく、今後の K1.8 ビームラインでの実験 (J-PARC E75[35], E90[36], E94[37] など) においても利用される。

S-2S における飛跡位置検出のために、ドリフトチェンバーが S-2S 電磁石の上流に 2 台 (SDC1,2)、下流に 3 台 (SDC3-5) 設置される (表 2.4)。ドリフトチェンバーは、中にガスが満たされ、高電圧が印加されたワイヤが多数張られた検出器である。ワイヤには、検出器内に電場を与える「ポテンシャルワイヤ」と、粒子の信号を読み出すための「センスワイヤ (陽極ワイヤ)」がある。ポテンシャルワイヤが生み出したポテンシャルによって、ワイヤの近くで荷電粒子が通過すると、ガス中の分子が電離する。この時に生じた電子がセンスワイヤに吸い寄せられ、それが信号として読み出される (図 2.12)。分子が電離しセンスワイヤに信号として吸われるまでの時間はドリフト時間と呼ばれる。このドリフト時間を測定することによって粒子の飛跡位置を数百  $\mu\text{m}$  の高い位置精度で測定することができる。[38]

SDC1 は  $x, u, v$  方向にそれぞれワイヤが 2 面ずつ、SDC2 は  $u, v$  方向に 2 面ずつ張り巡らされている。なお  $u, v$  ワイヤは  $x$  ワイヤに対して  $\pm 15$  deg 傾けて張られたものである。SDC3-5 については、 $x, y$  方向にそれぞれワイヤが 2 面ずつ張り巡らされている。

SDC1, 3-5 は、ハニカム状にワイヤが張り巡らされ (図 2.12)、SDC2 はポテンシャルワイヤとセンスワイヤが交互に張られたアノードプレートと炭素マイラーのカソードプレートが交互に積み重なった構造になっている。

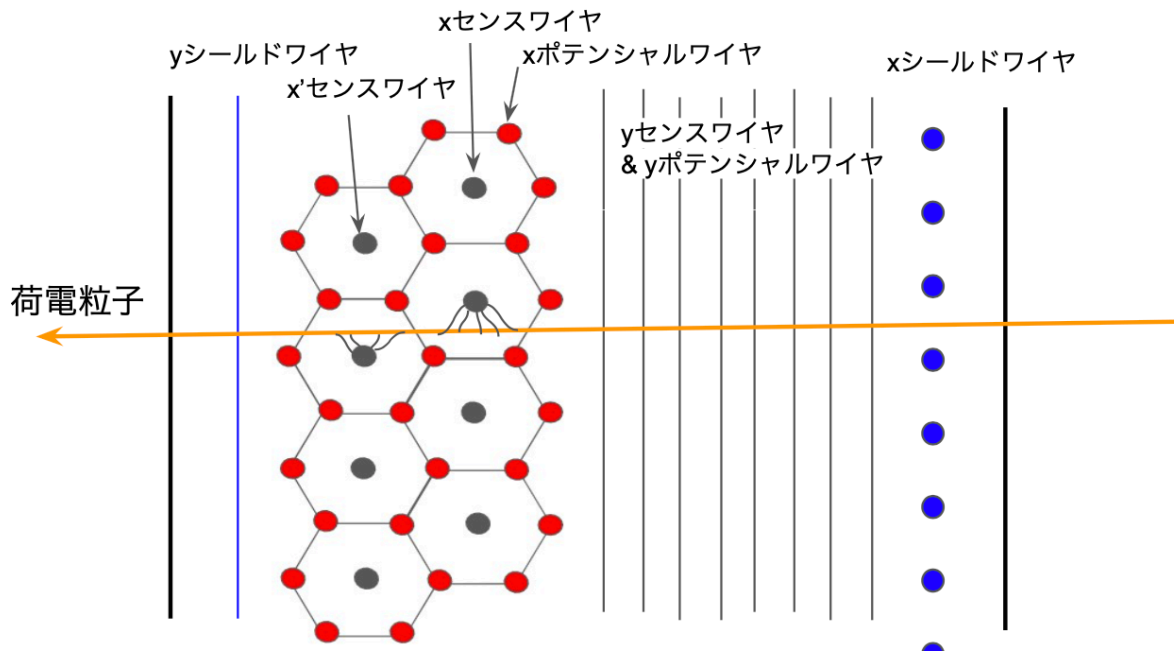


図 2.12: ドリフトチェンバー（ハニカム型）の模式図。

表 2.4: ドリフトチェンバーの性能

検出器	ワイヤ	ガス	有感領域 (mm)
SDC1 (図 2.13)	$uu'xx'vv'$	Ar + iso-C <sub>4</sub> H <sub>10</sub> + methylal (0.76 : 0.20 : 0.04)	$394^x \times 264^y$
SDC2 (図 2.14)	$vv'uu'$	Ar + iso-C <sub>4</sub> H <sub>10</sub> + methylal (0.76 : 0.20 : 0.04)	$160^x \times 300^y$
SDC3 (図 2.15)	$xx'yy'$	Ar + C <sub>2</sub> H <sub>6</sub> (1:1)	$1170^x \times 1170^y$
SDC4 (図 2.15)	$yy'xx'$	Ar + C <sub>2</sub> H <sub>6</sub> (1:1)	$1170^x \times 1170^y$
SDC5 (図 2.15)	$yy'xx'$	Ar + C <sub>2</sub> H <sub>6</sub> (1:1)	$1170^x \times 942^y$

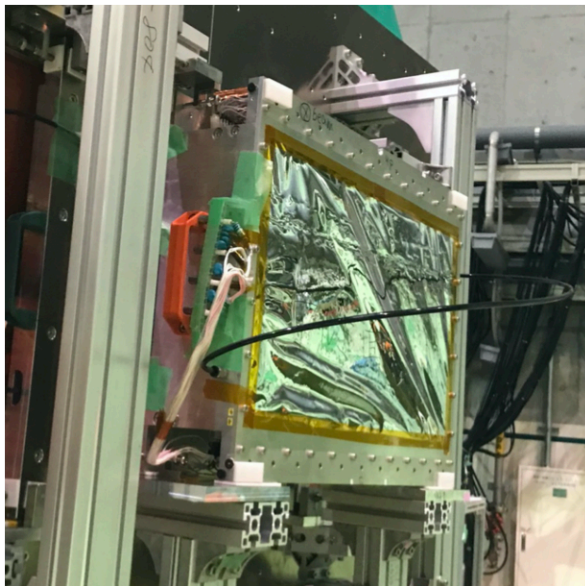


図 2.13: SDC1 の写真。

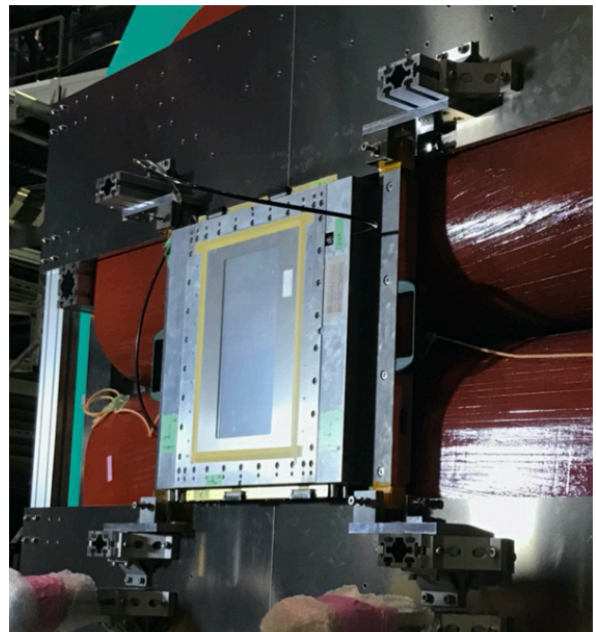


図 2.14: SDC2 の写真。



図 2.15: SDC3-5 の写真。

### 2.2.5 粒子識別法

また、S-2S を通過する粒子は  $K^+$  だけでなく、100 ~ 1000 倍程度の陽子、 $\pi^+$  中間子がバックグラウンドとして存在する。この大量のバックグラウンド事象を抑制するために粒子識別カウンターを設置する。 $\pi^+$  事象の抑制のためのエアロゲルチェレンコフ検出器 (AC)、陽子事象抑制用の水チェレンコフ検出器 (WC) [39] が設置される。また Time-of-Flight 情報を得るためにシンチレーションカウンターを設置する。ビームライン側のホドスコープ (BH2) と S-2S 下流に設置した TOF カウンターの間での粒子の飛行時間からオフラインで粒子識別を行う (表 2.5)。

表 2.5: 粒子識別カウンター群の性能

TOF (図 2.16)	
Time-of-Flight	
有感領域	$1192^x \times 600^y$ mm
時間分解能	$\sigma = 90$ ps
読み出し	36 PMTs
AC (図 2.17)	
$\pi^+$ 抑制	
媒質	シリカエアロゲル
屈折率 $n$	1.05
有感領域	$1450^x \times 995^y$ mm
$\pi^+$ 除去率	95% (simulation)
読み出し	18 PMTs
WC (図 2.18)	
陽子抑制	
溶媒	純水
屈折率 $n$	1.33
有感領域	$1495^x \times 730^y$ mm
陽子除去率	90% (simulation)
読み出し	24 PMTs





図 2.16: TOF カウンターの写真。

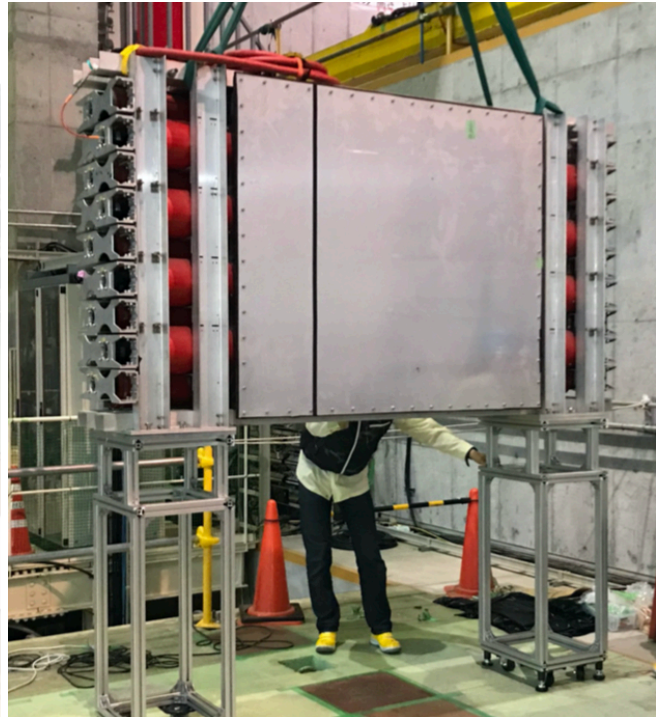


図 2.17: エアロゲルチェレンコフ検出器の写真。



図 2.18: 水チェレンコフ検出器の写真。

### 2.2.6 アクティブファイバー標的

また標的には炭素 12 を用いるが、標的物質中において粒子はエネルギー損失を起こす。この損失由来のエネルギーストラグリングによりミッシングマス分解能は悪化してしまう。特に、 $\Xi$  ハイパー核分光実験においては、まず素過程である  $p(K^-, K^+)\Xi^-$  反応の断面積が  $35 \mu\text{b}/\text{sr}$  である上に、運動量移行が  $\sim 500 \text{ MeV}/c$  もあるため、断面積が数十  $\text{nb}/\text{sr}$  程度に制限される。このため標的厚を十分厚くする必要があり、E70 実験での炭素標的は  $9 \text{ g}/\text{cm}^2$  であるため、より分解能の悪化が懸念される。分解能悪化を防ぐために、E70 実験では、炭素標的としてアクティブファイバー標的 (AFT) を採用した [40][41]。AFT は、 $\phi 3\text{mm}$  の  $\text{CH}_2$  のシンチレーションファイバーが縦横に格子状に整列した状態がビーム軸方向に 9 層重なった状態にある (表 2.6)。

$$\Delta M^2 = \left(\frac{\partial M}{\partial p_B}\right)^2 \Delta p_B^2 + \left(\frac{\partial M}{\partial p_S}\right)^2 \Delta p_S^2 + \left(\frac{\partial M}{\partial \theta}\right)^2 \Delta \theta^2 + \Delta E_{\text{strag}}^2 \quad (2.3)$$

表 2.6: AFT の性能

AFT (図 2.19)	
fiber config.	$xx'yy' \times 9 \text{ layers}$
媒質	CH $\phi 3 \text{ mm}$ ファイバー
読み出し	MPPC + VME-EASIROC
有感領域	$100^x \times 50^y \times 100^z$
Thickness	$9 \text{ g}/\text{cm}^2$

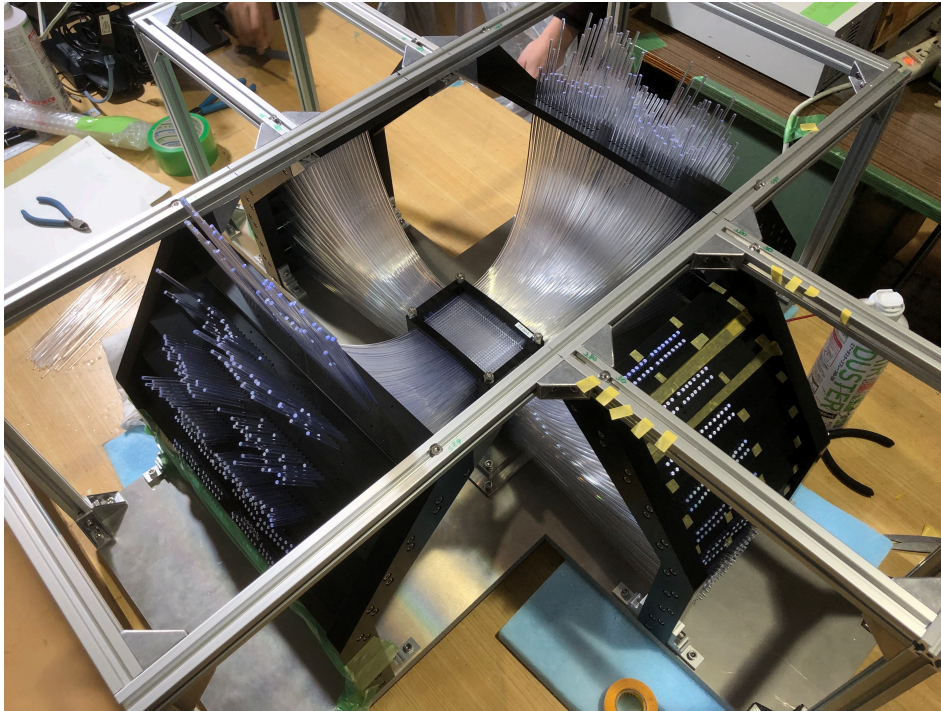


図 2.19: AFT の写真。

紙面表側からビーム粒子が中央の黒い枠内に向かって照射される。

## 2.3 期待される物理的結果

J-PARC E70 実験では  $\Xi$  ハイパー核 ( ${}_{\Xi}^{12}\text{Be}$ ) のミッシングマスを、 $2 \text{ MeV}/c^2$  の精度で測定することを目指している。20 日間の Physics run で約 100 イベントの  $\Xi$  ハイパー核事象をとらえる予定である (表 2.7)。これにより、束縛エネルギーピークを 100 keV, 崩壊幅を 500 keV 程度の精度で決定する。

表 2.7: E70 実験の収量見積もり

E70	
$K^+$ survival ratio	0.46 ( $\beta\gamma c\tau = 10.2$ m, $p_K = 1.37$ GeV/c)
Solid angle (msr)	55
Cross Section $\sigma$ (nb/sr)	60 ( $\theta < 10$ deg)
Target Thickness (g/cm <sup>2</sup> )	9 (AFT made of CH)
Intensity ( M Kaons / spill)	0.8
Spill cycle (sec)	4.2
total efficiency	0.5
physics run time	20 days
$^{12}_{\Xi}\text{Be}$ event	$\sim 100$ counts

高分解能なミッシングマス分光は、前述の S-2S 磁気スペクトロメータによる高分解能な散乱  $K^+$  の運動量測定や、アクティブファイバー標的ストラグリングによる分解能悪化の抑制によって達成される。 $K^+$ ,  $K^-$  の運動量分解能  $\Delta p_B, \Delta p_S$  と、散乱角度  $\Delta\theta$ 、エネルギーストラグリングによる分解能からミッシングマス分解能  $\Delta M_X$  への寄与  $E_{\text{strag}}$  は、下の式のようなになる。

$$M_X = \sqrt{(E_B + M_{\text{target}} - E_S)^2 - (p_B - p_S)^2} \quad (2.4)$$

$$\Delta M_X^2 = \left(\frac{\partial M_X}{\partial p_B}\right)^2 \Delta p_B^2 + \left(\frac{\partial M_X}{\partial p_S}\right)^2 \Delta p_S^2 + \left(\frac{\partial M_X}{\partial \theta}\right)^2 \Delta \theta^2 + \Delta E_{\text{strag}}^2 \quad (2.5)$$

$$\begin{cases} \frac{\partial M}{\partial p_B} = \frac{1}{M_X} \left( \frac{p_B}{E_B} (M_{\text{target}} - E_S) + p_S \cos \theta \right) \\ \frac{\partial M}{\partial p_S} = \frac{1}{M_X} \left( -\frac{p_S}{E_S} (M_{\text{target}} + E_B) + p_B \cos \theta \right) \\ \frac{\partial M_X}{\partial \theta} = \frac{1}{M_X} p_B p_S \sin \theta \end{cases} \quad (2.6)$$

これらの式のもとに計算すると、以下の条件で各項の寄与や分解能を計算すると表 2.8 のようになる。

$$\begin{cases} p_B = 1.8 \text{ GeV}/c \\ \Delta p_B/p_B = 1.0 \times 10^{-3} \\ p_S = 1.37 \text{ GeV}/c \\ \Delta p_S/p_S = 6.0 \times 10^{-4} \text{ (E70)}, 3.0 \times 10^{-3} \text{ (E05)} \\ M_B = M_S = M(\text{Kaon}) \\ M_T = M(^{12}\text{C}) \\ M_X = M(^{11}\text{B} + \Xi^-) \end{cases} \quad (2.7)$$



表 2.8: ミッシングマス分解能と、各項の寄与を示す。  
 先行実験である E05 実験と、E70 実験での設計値 (designed E70) を  
 AFT によるストラグリングの補正の有無も併せて比較する。

(MeV)	$\frac{\partial M}{\partial p_B} \Delta p_B$	$\frac{\partial M}{\partial p_S} \Delta p_S$	$\frac{\partial M}{\partial \theta} \Delta \theta$	$\Delta E_{\text{strag}}$	$\Delta M$ (total)
designed E05	1.67	3.73	0.04	3.0	5.1
designed E70 (w/o AFT)	1.67	0.74	0.04	2.4	3.0
designed E70 (w/ AFT)	1.67	0.74	0.04	0.9	2.0

AFT を考慮したミッシングマス分解能は表のように  $2.0 \text{ MeV}/c^2$  が達成されると期待されている。下の図に、ESC08a モデルに基づく理論計算 [23] で予想されるスペクトルがそれぞれの実験的分解能でどのように観測されるかを示した。BNL E885 実験 [22] での分解能  $14 \text{ MeV}$  で測定したときに予想されるミッシングマススペクトル (図 2.20) と、J-PARC E05 実験での分解能  $6 \text{ MeV}$  で測定したときに予想されるスペクトル (図 2.21) を載せる。また、本実験 (J-PARC E70 実験) において AFT を考慮しない場合 (図 2.22) と、考慮した場合 (図 2.23) の予想されるスペクトルも示した。

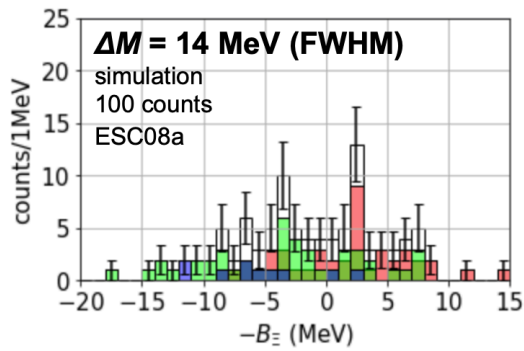


図 2.20: 分解能が 14 MeV である時に予想されるミッシングマス分布。

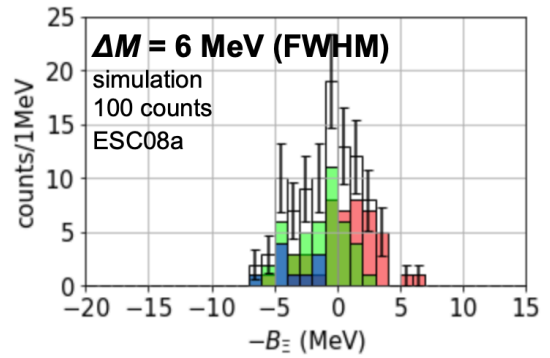


図 2.21: 分解能が 6 MeV である時に予想されるミッシングマス分布。

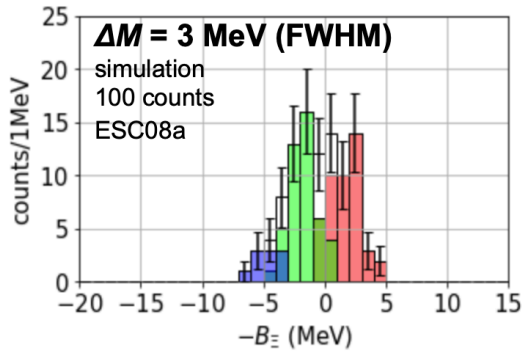


図 2.22: 分解能が 3 MeV である時に予想されるミッシングマス分布。

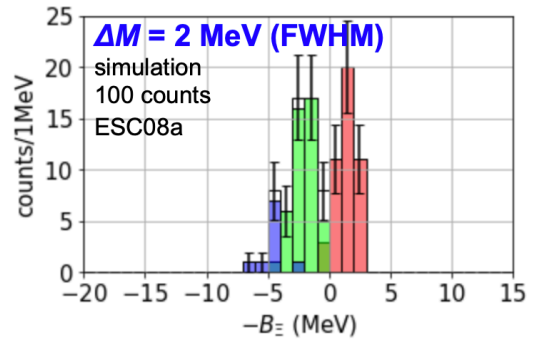


図 2.23: 分解能が 2 MeV である時に予想されるミッシングマス分布。

世界最高分解能での  $\Xi$  ハイパー核の欠損質量分光により  $\Xi N$  間相互作用に決定的な情報が与えられることが期待される。

## 第3章

# 運動量再構成手法

### 3.1 従来解析手法

運動量を飛跡情報から再構成する場合、従来方法として下記のような手法が一般的に用いられる。

#### 3.1.1 ルンゲクッタ法

スペクトロメータの磁場中の運動方程式から、ルンゲクッタ法を用いて粒子の飛跡を計算する。測定中にドリフトチェンバーのワイヤ面で捉えた hit 位置から、電荷  $q$  を持つ粒子の飛跡を磁場  $\vec{B}$  中の荷電粒子の運動方程式に従って、飛跡上の各点で  $\vec{B}$  を積分し (図 3.1)、元の運動量  $\vec{p}$  を逆算して求めていく。運動量分解能は磁場の測定精度に直接依存し、例えば飛跡に沿って  $\vec{B}$  を  $10^{-4}$  の精度で測定できれば、 $\Delta p/p \sim \Delta B/B \simeq 10^{-4}$  の分解能で運動量測定できる。したがって、ルンゲクッタを運用する場合、事前にスペクトロメータの磁場分布を詳細に知っておく必要がある。

E05 実験での、SKS スペクトロメータはこの手法を適用したスペクトロメータである。

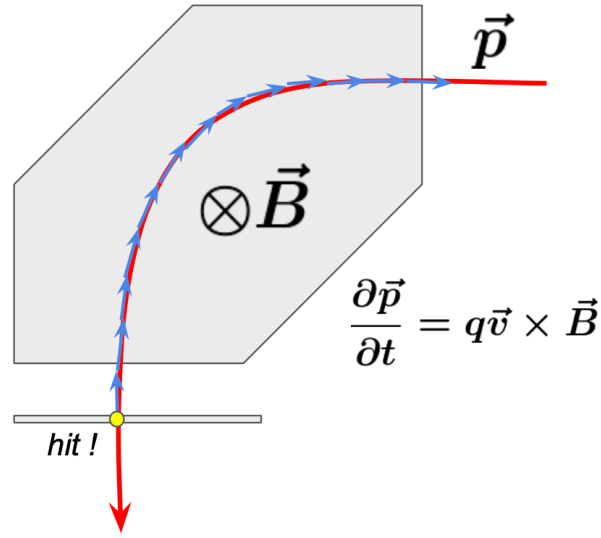


図 3.1: ルンゲクッタ法概念図。

### 3.1.2 逆輸送行列による手法

もう一つの従来の再構成手法としては、個々の電磁石の効果を中心軌道の周りの磁場を展開することで近似する、輸送行列法である。磁場分布をうまくモデル化し磁場を展開することができれば、出入り口における焦点面間でのビームの写像を近似することができる。

磁気スペクトロメータを通過した粒子は、電磁石の輸送行列で光学的にその軌道を見積もることができる。ここで、粒子の飛跡情報として、粒子位置  $(x, y)$  と、粒子の角度  $(u = \partial x / \partial z, v = \partial y / \partial z)$ 、粒子の運動量の中心運動量  $p_0$  からの分散  $\delta = (p - p_0) / p_0$  を要素に持つベクトル  $\vec{x}$  を定義する。磁石の出口での  $\vec{x}_{\text{out}}$  は、入り口での位置・角度ベクトル  $\vec{x}_{\text{in}}$  と、電磁石固有の輸送行列  $M$  から求めることができる。

例えば、中心軌道の曲げ角を  $\theta$ 、曲率半径を  $\rho$  とした偏向（双極）電磁石（粒子の飛跡に対して N 極と S 極が向かい合っている電磁石で、粒子の軌道を曲げる能力を持つ）の輸送行列は、以下のようにかける

$$M_D = \begin{pmatrix} \cos \theta & \rho \sin \theta & 0 & 0 & (1 - \cos \theta) \rho \\ -\sin \theta / \rho & \cos \theta & 0 & 0 & \sin \theta \\ 0 & 0 & 1 & \rho \theta & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (3.1)$$

純粋な収束（四重極）電磁石（粒子の軌道を中心軌道に収束させる働きを持つ）の輸送行列は以下のようにかける。磁極長を  $L$  とし、 $k$  は中心磁場勾配と中心運動量の比に比例する量である。 $k^2 = e/p_0 \cdot (\partial B/\partial x)_{x=y=0}$  である。

$$M_Q = \begin{pmatrix} \cos(kL) & \sin(kL)/k & 0 & 0 & 0 \\ -k \sin(kL) & \cos(kL) & 0 & 0 & 0 \\ 0 & 0 & \cosh(kL) & \sinh(kL)/k & 0 \\ 0 & 0 & -k \sinh(kL) & \cosh(kL) & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (3.2)$$

このような輸送行列の逆行列  $M^{-1}$  を用いて、粒子の飛跡位置  $\vec{x}_{\text{out}}$  を入力として輸送方程式を解くことで、運動量  $\delta_{\text{in}}$  を求めることができる。

$$\vec{x}_{\text{in}} = M^{-1} \vec{x}_{\text{out}} \quad (3.3)$$

この手法は、K1.8 ビームラインスペクトロメータに適用されている。

## 3.2 本研究の動機

粒子の飛跡から高分解能で運動量を測定するのは、単純な手法では難しい。今回の実験では、欠損質量分解能 2 MeV 以下を目指している。

高分解能なミッシングマス分光のためには、高分解能な運動量測定が求められる。E70 実験においては、反応後の散乱粒子  $K^+$  の運動量を S-2S 磁気スペクトロメータを用いて測定する。運動量分解能として  $\Delta p/p \simeq 6.0 \times 10^{-4}$  が必要とされている。

S-2S 電磁石は、単一 D 磁石で構成される SKS 電磁石と異なり、QQD 磁石系で構成されている。SKS 電磁石とは J-PARC E05 実験で用いられた散乱側の磁気スペクトロメータである。先にも述べたが、四重極磁石を採用し磁極長を長くしたことなどによって、 $R_{11}$  (magnification,  $\langle x|x \rangle$ ) と  $R_{12}$  ( $\langle x|\theta \rangle$ ) が小さく、 $R_{16}$  (dispersion,  $\langle x|p \rangle$ ) が大きい。そのため低次のパラメータに対しては SKS よりも補正の必要性が少なくなると予想される。これに対して SKS よりも磁場マップは Q 磁石の存在や磁石間の磁場分布の不定性などにより複雑となると予想される。このため、運動量やミッシングマスを再構成したときに、従来手法ではとりきれないような高次の効果が現れると予想される。

また、ルンゲクッタ法においては、磁場の測定誤差が直接運動量誤差に寄与するため、磁場分布を事前に詳細に知っておく必要がある。しかし  $10^{-4}$  レベルの分解能で測定するのは比較的難しい。さらに運動量再構成後に現象論的な関数を用いた補正を行う場合が多いことが問題となっている。

一方、逆輸送行列による手法に関しては、磁場分布の初期の理解が不十分でも校正データを用いたパラメータ最適化により、最終的に高分解能・高確度が達成できる可能性がある。実際に JLab ではこの手法で S-2S が目指す分解能以上の高分解能が達成されている [42]。しかし、S-2S のようにアクセプタンスの広い場合は解析は難しいとされる。

そこで、運動量解析手法に対して、従来手法で用いられるような磁場中の運動方程式やビーム光学に基づく所定の方程式ではなく、自由度の高いニューラルネットワークを用いた機械学習を採用する。これにより、例えばルンゲクッタ法で指摘されていた再構成後の補正が自動的に行われ、さらに従来手法では取りきれない高次の相関を掬いとることにより、高効率で高分解能、高確度な運動量解析が期待される。また、ニューラルネットワークの拡張性の高さから、例えば飛跡情報から運動量を再構成するのではなく、トラッキングなしにドリフトチェンバーで得た各ワイヤ面でのヒット位置から運動量を再構成する、などといった柔軟な解析も可能となる。

### 3.3 先行研究

#### 機械学習を用いた J-PARC E05 実験の運動量解析

本修論の先行研究である大橋氏の修論 [43] では、E70 の先行実験である J-PARC E05 実験のデータを用いて機械学習によるスペクトロメータの運動量解析の有効性について調べられた。

E05 実験での散乱粒子の運動量測定には、SKS 磁気スペクトロメータが使用された。SKS 電磁石は D 磁石のみで構成された超伝導電磁石である。E05 の解析においては、Beam through run で、粒子電荷の正負によって K18 スペクトロメータと SKS から与えられる解析運動量の差が異なるということがあった。また、解析運動量と角度  $(u, v)$  に依存関係があることが分かった。このため、E05 の従来手法 (ルンゲクッタ法) を用いた運動量解析では、まず Beam Through データから  $p_S - p_B = 0$  となるように補正する。そののち、素過程  $p(K^-, K^+)E^-$  での  $u_{SKS} = v_{SKS} = 0$  の領域で欠損質量を計算する。この欠損質量の値が PDG[5] での値と合致するようにフィッティングする。最後に、素過程  $p(K^-, K^+)E^-$  での  $u, v$  での相関がなくなるように補正を行った。このような補正手法は解析者に依存する。また  $x, y, u, v$  といったパラメータと  $p$  の間の相関を人の手で限なく探し、運動量解析をおこなうのは困難である。

そこで、広範に及ぶ効果を所定の方程式を置かずに考慮し、従来のモデリングで考慮していない効果を取り入れられる解析手法として、機械学習による運動量解析が適用可能か

を調べた。機械学習の手法としては、まず Beam Through run でのビーム運動量と散乱運動量の差が小さくなるように学習する。その後ミッシングマスで  $\Xi$  ピーク、 $\Sigma$  ピークを構成し、PDG との差が小さくなるように学習した。

結果として  $p(K^-, K^+)\Xi^-$  と  $p(K^-, \pi^+)\Sigma^-$  データを用いた機械学習による解析により、欠損質量ピークが観測された。 $\Sigma$  ピークに対しては、従来解析の分解能である  $7.12 \pm 0.06 \text{ MeV}/c^2$  と同程度の分解能  $7.06 \pm 0.06 \text{ MeV}/c^2$  を得ることができた。これらの結果より、運動量解析に対して機械学習の有用性が証明された。

### 3.4 本研究の目的

本研究においては、J-PARC E70 実験で用いる S-2S 磁気スペクトロメータを用いた散乱粒子  $K^+$  の運動量解析フレームワークを、機械学習を用いて構築する。

運動量分解能  $\Delta p_S/p_S$  の大きさ、過学習の有無、飛跡情報  $(x, y, u, v)$  との相関、また解析運動量に余計な相関が与えられていないかなどの観点から、機械学習モデルの開発を行う。さらにドリフトチェンバーのヒット情報を入力とした機械学習モデルを構築し、トラッキングなしに運動量再構成できるかについて検証する。

また構築したモデルが Beam through や  $\Xi^-$ ,  $\Sigma^-$  事象のデータによって校正可能かどうかについても検証する必要がある。

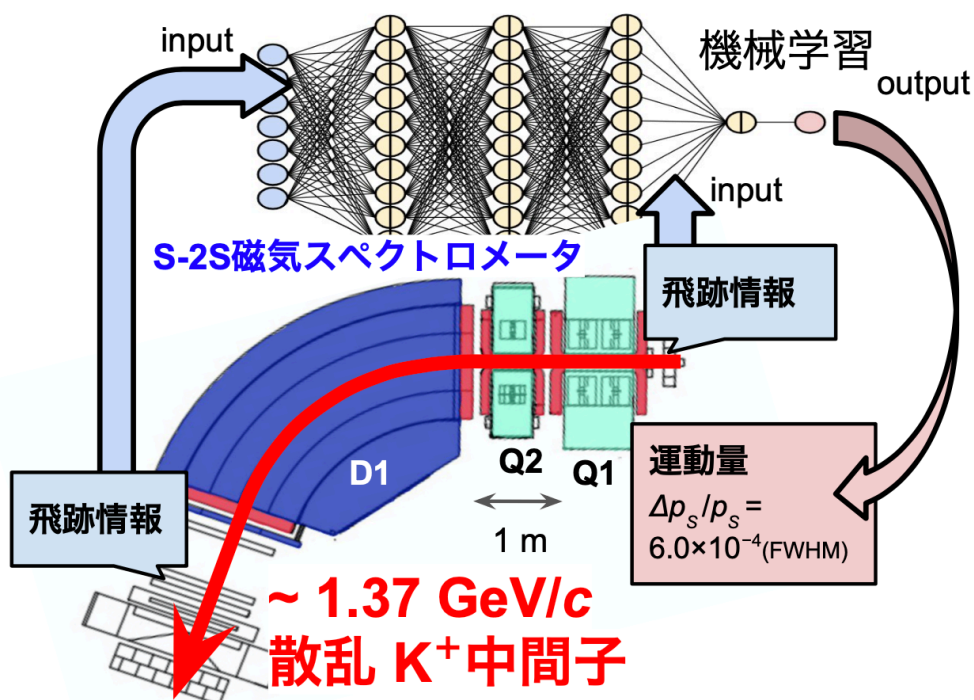


図 3.2: 本修論における機械学習モデルの概念図。

## 3.5 運動量解析への機械学習の導入

### 3.5.1 機械学習

機械学習とは、大量のデータを用いた経験的学習により自動的にデータ間の相関や分類、予測などの問題を解決するアルゴリズムの一種である。近年では J-PARC E07 実験での原子核乾板中の事象探索や、加速器制御 [44]、LHC や Belle2 などのコライダー実験での信号識別 [45, 46] や飛行検出 [47]、データ圧縮など、大量のデータを扱う素粒子原子核実験の測定に機械学習が適用されてきている。機械学習には大きく分けて、教師あり学習、教師なし学習、強化学習に分類される。本修論で扱うのは教師あり学習であり、回帰問題を扱う。教師つき回帰問題においては多くのデータ集団から、入力 (訓練) データ  $x$  と、出力 (教師) データ  $y$  間に存在する相関を見つけることが目的となる。訓練に用いる標本データを、相関  $y = f(x)$  が与えられた母集団から抽出する場合、抽出した標本集団は適当なランダムノイズ  $\epsilon$  が与えられた相関  $f + \epsilon$  を持つ。

$$y = f(x) + \epsilon \quad (3.4)$$



この時膨大なデータ  $x$  から、 $y$  に可能な限り近い予想値  $\hat{y}$  を出力できる相関  $\hat{f}(x)$  を求めることが学習の目的となる。

$$\hat{y} = \hat{f}(x) \quad (3.5)$$

したがって、 $y, \hat{y}$  の差分（損失関数）をいかに小さく抑えるかを指標として学習を行う。

$$Loss = |y - \hat{y}| \text{ or } (y - \hat{y})^2 \quad (3.6)$$

機械学習では、「パーセプトロン」と呼ばれる入出力ユニットを多数組み合わせた数理模型を用いる。これは「ニューラルネットワーク」と呼ばれる。膨大なデータを用いて模型内のパラメータを  $Loss = |y - \hat{y}|$  を基に何度も更新を繰り返すことによって、入力データ  $x$  から  $y$  に限りなく近い値を出力できるような模型を構築する。

パーセプトロンは脳神経回路を真似て作られたものである。例えば図 3.3 のように複数の入力信号  $x_1, x_2, \dots$  に対して、出力  $y$  を与えるものを考える。入力  $x_1, x_2$  に対してパラメータ  $w_1, w_2$  で重み付けし、和をとる。これにバイアス  $b$  を加えた値が或る閾値  $a$  を超えた場合に出力が与えられ、超えない場合は出力されない。

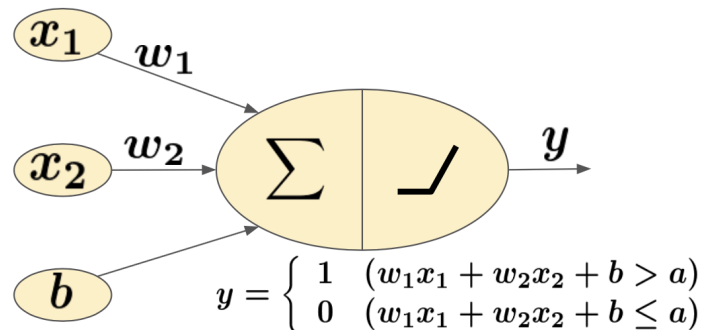


図 3.3: パーセプトロンの例。複数の入力信号に対して単一の信号を出力する機構。重み付けした和に対して或る閾値  $a$  を超えた場合出力を行う。

このパーセプトロン一個だけでは単純な値しか出力されないが、これを多層つなげることで線形分離できないような複雑なデータの識別などが可能になる。この機構は、多層パーセプトロンと呼ばれる。

先述したパーセプトロンでは、閾値  $a$  を設けて階段関数にしたがった出力を与えていた。ここで階段関数ではなく他の非線形関数を導入したニューロン（またはノードと呼ぶ）を用いることで、より複雑なニューラルネットワークを構築することが可能になる。

この非線形関数は一般に活性化関数と呼ばれている。活性化関数は生物の脳細胞の信号の入出力からヒントを得て作られた。活性化関数にはよく sigmoid 関数や ReLU 関数（一般的にはランプ関数と呼ばれる）が用いられる。

ニューラルネットワークにおいては、活性化関数を用いたノードを多層つなげることにより複雑な出力が可能になる。入力データ（本修論では飛跡情報  $x, y, u, v$  を用いる）を入力層に与える。入力層の出力をいくつかの隠れ層（中間層）に与え、結合層を経て最終的に出力層に出力される。層間の入出力においては、例えば  $l$  番目の層のノードから  $l+1$  番目の層の全ノードへ重みづけされた出力が伝播される。この重みパラメータは全ノード間に与えられる。機械学習とは、これらの層の間の重みパラメータを更新することである。

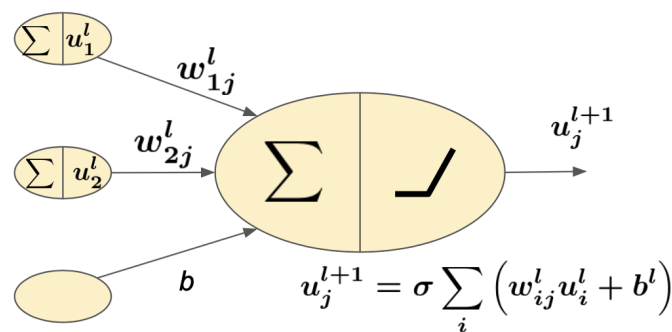


図 3.4: 単層パーセプトロンの概念図。複数の入力信号に対して単一の信号を出力する機構。入力の重み付け和をとった信号が、或る閾値  $a$  を超えた場合出力を行う。

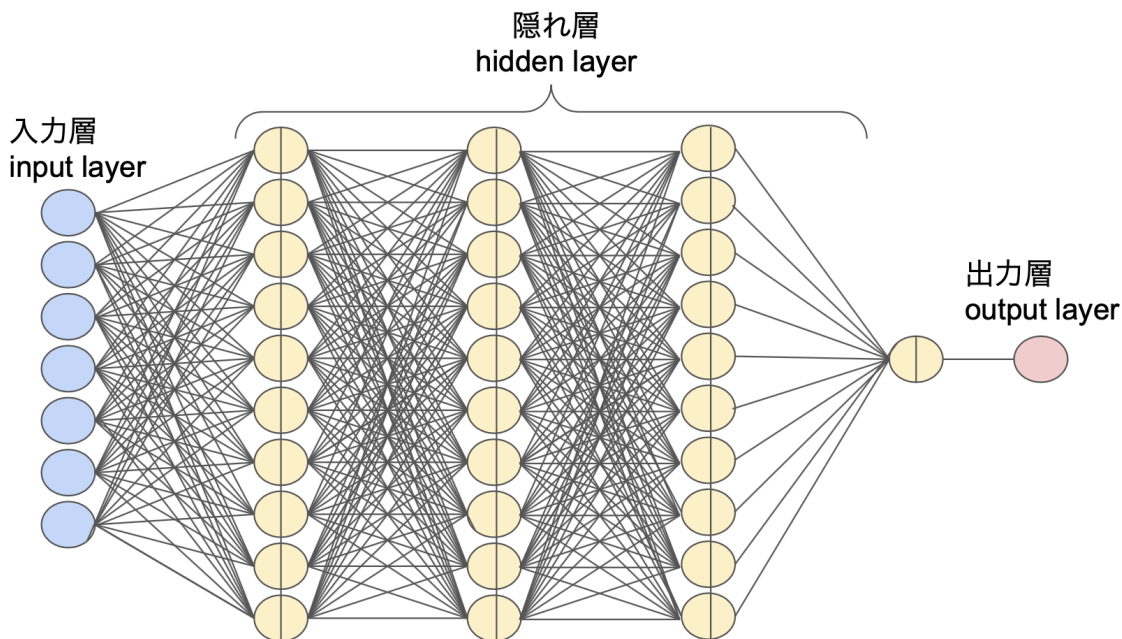


図 3.5: ニューラルネットワークの概念図。

重みパラメータの更新には損失関数を用いる。更新した重みパラメータを用いたニューラルネットワークでまた新たな損失関数を導出し、これを基に重みパラメータを再度更新する。これを繰り返すことによって訓練データから教師データが再現できるようなモデルを構築する。

回帰問題での損失関数には、平均絶対誤差 (Mean Absolute Error) や平均二乗誤差 (Mean Squared Error) がよく用いられる。重みパラメータの更新には、SGD (Stochastic Gradient Descent : 確率的勾配降下法)、RMSProp、Adam などの最適化関数と呼ばれる関数が用いられる。

$$w_{ij}^l \leftarrow w_{ij}^l - \eta \nabla_{w_{ij}^l} L \quad (3.7)$$

重みパラメータの更新方法は上の式のように、基本的に損失関数の重みについての勾配分更新する。 $w_{ij}^l$  は、 $l-1$  番目の層の  $i$  ノードから  $l$  層の  $j$  ノードに値が伝播するときの重みである (図 3.6)。損失関数の重みについての勾配分  $\nabla_w L$  と学習率 (learning rate)  $\eta$  の積の分だけ重みの更新を行っている。

ただし、パラメータ更新の際にそのまま微分を用いると、かなり実行時間がかかってしまう。このため誤差逆伝播法と呼ばれる手法が用いられる。

### 3.5.2 誤差逆伝播法

機械学習では、損失関数  $L$  の分だけ重みパラメータ  $w$  を更新する処理を繰り返すことであると述べた。例えば重みパラメータ  $w_{ij}^l$  は、次の式に従って値を更新する。

$$w_{ij}^l \leftarrow w_{ij}^l - \eta \frac{\partial L}{\partial w_{ij}^l} \quad (3.8)$$

この  $\partial L / \partial w_{ij}^l$  について考える。 $w_{ij}^l$  は、 $l$  番目のレイヤーの  $j$  ノードの出力  $u_j^l$  の関数として表されるため、下のように書き換えられる。

$$\frac{\partial L}{\partial w_{ij}^l} = \frac{\partial L}{\partial u_j^l} \frac{\partial u_j^l}{\partial w_{ij}^l} = \delta_j^l \frac{\partial u_j^l}{\partial w_{ij}^l} \quad (3.9)$$

$$\delta_j^l = \frac{\partial L}{\partial u_j^l} \quad (3.10)$$

この  $\delta_j^l$  について、 $u_j^l$  が下位レイヤー  $j+1$  の出力  $u_k^{l+1}$  の関数としても表されるため、次のように書ける。

$$\delta_j^l = \frac{\partial L}{\partial u_j^l} = \sum_k \frac{\partial L}{\partial u_k^{l+1}} \frac{\partial u_k^{l+1}}{\partial u_j^l} = \sum_k \delta_k^{l+1} \frac{\partial u_k^{l+1}}{\partial u_j^l} \quad (3.11)$$

式中の  $\delta_j^l = \partial L / \partial u_j^l$  はこのように下位層の  $\delta^{l+1}$  の関数として書くことができる。最後の出力層の  $\delta^{l+N}$  に関しては、

$$\delta_j^{l+N} = \frac{\partial L}{\partial u_j^{l+N}} \quad (3.12)$$

となり、 $L$  が  $u_j^{l+N}$  に対して陽な関数であるため、損失関数  $L$  の値から直接決まる。したがって、任意の  $\delta_j^l$  の値は、出力層の  $\delta^{l+N}$  から  $\delta^{l+N} \rightarrow \delta^{l+N-1} \rightarrow \dots \rightarrow \delta^0$  のように伝播するように決まる。

このように下位層から順に伝播的に  $\delta$  を求めることにより、勾配  $\partial L / \partial w$  を求めることができる。このアルゴリズムを誤差逆伝播法と呼び、よく機械学習での重み更新に使用される。

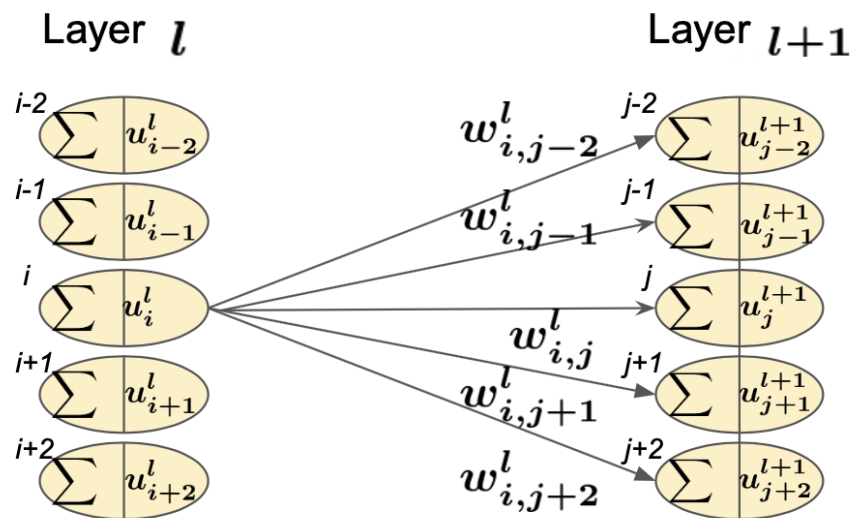


図 3.6: ノード間の重みパラメータ伝播。

### 3.5.3 ミニバッチ学習

訓練データの統計量が巨大である場合に、それら全てのデータに対して損失関数を計算するのは計算コストがかかり、効率が悪い。そこで機械学習の実装においては、よく「ミニバッチ学習」と呼ばれる手法がとられる。ミニバッチ学習では、訓練データから無作為に一部のデータセット（ミニバッチ）を抽出し、この抽出したデータセットを用いて重みパラメータの更新を行う。

また、ミニバッチ学習において、訓練データを学習で全て使い切った時の回数に対応する量を「エポック (epoch)」と呼ぶ。つまり、例えば  $10^5$  の統計量の訓練データがあり、100 個のデータセットをミニバッチとして一回のパラメータ更新に使う場合、更新が  $10^5/100$  回繰り返された時に全て使い切った場合に対応する。従って、この場合の 1 epoch はパラメータ更新  $10^3$  回に対応する。

### 3.5.4 過学習

機械学習においては、やみくもに訓練データを学習させて教師データを完全に再現させることが良いとは限らない。過学習によって抽出した訓練・教師データにのみ適用され、母集団全体には適合しない機械学習モデルが構築されてしまう。つまり標本集団の相関  $y = f(x)$  に生じたランダムノイズ  $\epsilon$  を、母集団の相関の一つとして誤って認識してしま

うことである。過学習が生じる原因としては、モデル内のパラメータが大量で複雑であるため、過剰に表現力の高いモデルとなってしまうことと、訓練データの統計量が不足していることが挙げられる。

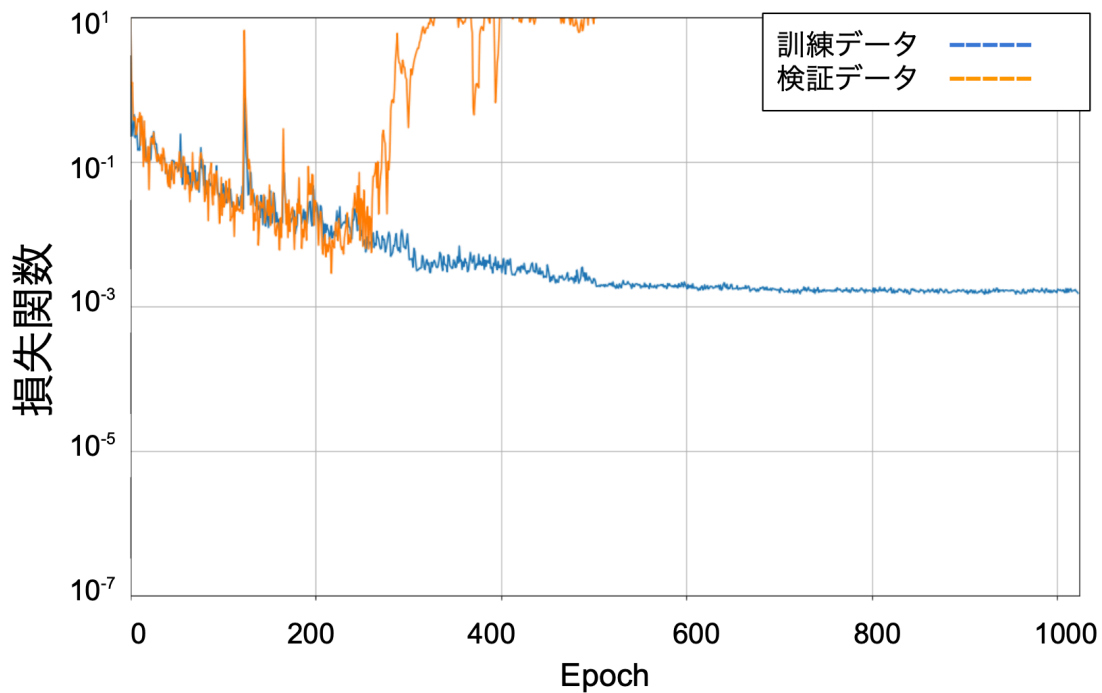


図 3.7: 過学習の例。横軸が学習におけるパラメータ更新の繰り返しの回数で、縦軸は損失関数である。青が訓練データで計算した損失関数の推移を示し、橙が過学習検証データで計算した損失関数の推移を示す。図のように訓練が進むと検証データによる損失関数が訓練データに対して悪化しているのがわかる。これはモデル内のパラメータが、母集団ではなく、抽出した訓練データのみに見える特徴に特化した値に収束している状態である。

過学習の有無を確認するために訓練データとは別に検証用データを用意し、学習時に検証データの損失関数をモニターする手法をとる (図 3.7)。また学習後に学習には全く関与しないデータ (テストデータ) を用いて、訓練データ、検証データ、テストデータの相関や計算誤差が一貫しているかを検証する。過学習を防ぐには様々な方策が用いられる。

- early stopping

損失関数が悪化する前に学習を早期に打ち切る手法。検証データによる損失関数をモニターし、損失関数が一定期間改善されない場合に学習を終了する。それにより過学習の防止につながる。また実行時間の削減にもつながる。

- drop out

過学習の原因のひとつとしてはパラメータ数の過多によるモデルの複雑化である。学習モデルの中のノードを一定の割合で使用しないことにより、学習のパラメータ数を減らし過学習を防ぐ。[48]

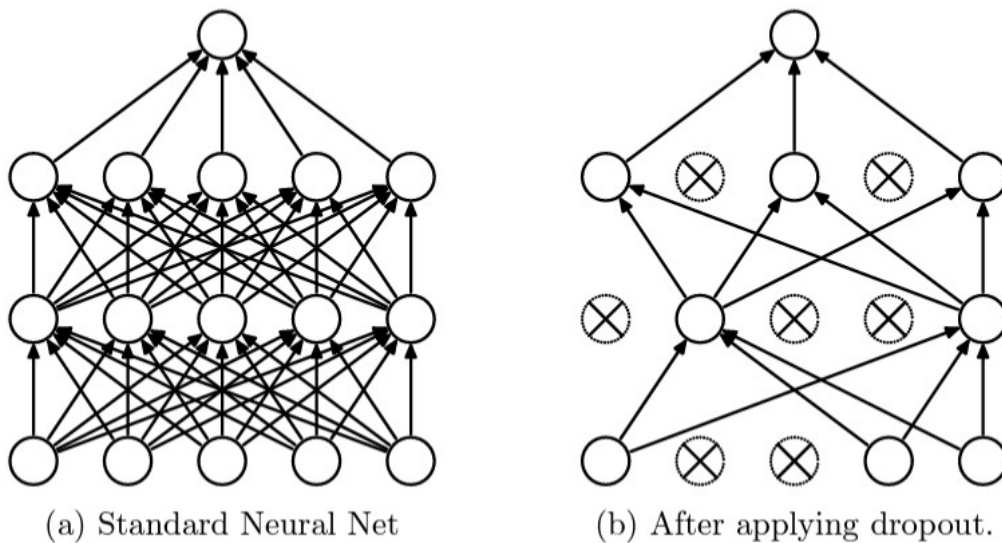


図 3.8: dropout の概念図。[48] より。

- L1, L2 正則化

重みパラメータを更新する際に、先に述べた損失関数に制限を与える。この制限によって学習に制約がかかり、過学習の防止につながる。L1 正則では損失関数に重み  $w_i$  の絶対値を加え (式 3.13)、L2 正則では  $w_i$  の二乗を加える (式 3.14)。

$$L(w) \rightarrow L(w) + \alpha \sum_i |w_i| \quad (3.13)$$

$$L(w) \rightarrow L(w) + \alpha \sum_i (w_i)^2 \quad (3.14)$$

- Batch Normalization

学習中に各層の出力値を正規化する。正規化することで各層の出力の分布が適切な広がりを持たせる。これにより過学習の抑制、学習速度が上がる。また学習の進捗が初期値に依存しにくくなる。[49]



$$\mu_B \leftarrow \frac{1}{m} \sum_i^m x_i \quad (3.15)$$

$$\sigma_B^2 \leftarrow \frac{1}{m} \sum_i^m (x_i - \mu_B)^2 \quad (3.16)$$

$$\hat{x}_i \leftarrow \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \quad (3.17)$$

入力  $x$  を正規化した  $\hat{x}$  に変換する (式 3.17)。さらに適当なスケールパラメータ  $\beta, \gamma$  で  $\hat{x}$  を変換したものを出力する。

$$y_i \leftarrow \gamma \hat{x}_i + \beta \quad (3.18)$$

## 第 4 章

# 機械学習モデルの構築と最適化

本章においては機械学習で運動量解析フレームワークを構築するために、訓練に用いるニューラルネットワークの隠れ層の数、隠れ層 1 層あたりのノード数、活性化関数、最適化関数、損失関数といったパラメータ (ハイパーパラメータ) の最適化を行った。

### 4.1 模型構築に用いるシミュレーションデータ

機械学習に用いる訓練データおよび教師データは、CERN が提供するシミュレーションソフトである GEANT4 から出力されるデータを使用した。シミュレーションデータを用いることで、先行研究 [43] で問題点となっていた統計量の問題が解消できると期待される。GEANT4 上に E70 実験の検出器セットアップを構成し、標的位置から  $K^+$  中間子を照射する。S-2S の上下流においた検出器で、 $K^+$  の位置・角度情報を取得し、これを訓練データとして扱う。S-2S の QQD 磁石系の磁場分布に関しては、2016 年に実測定 [50] した磁場に基づき TOSCA で計算した磁場マップを用いている。

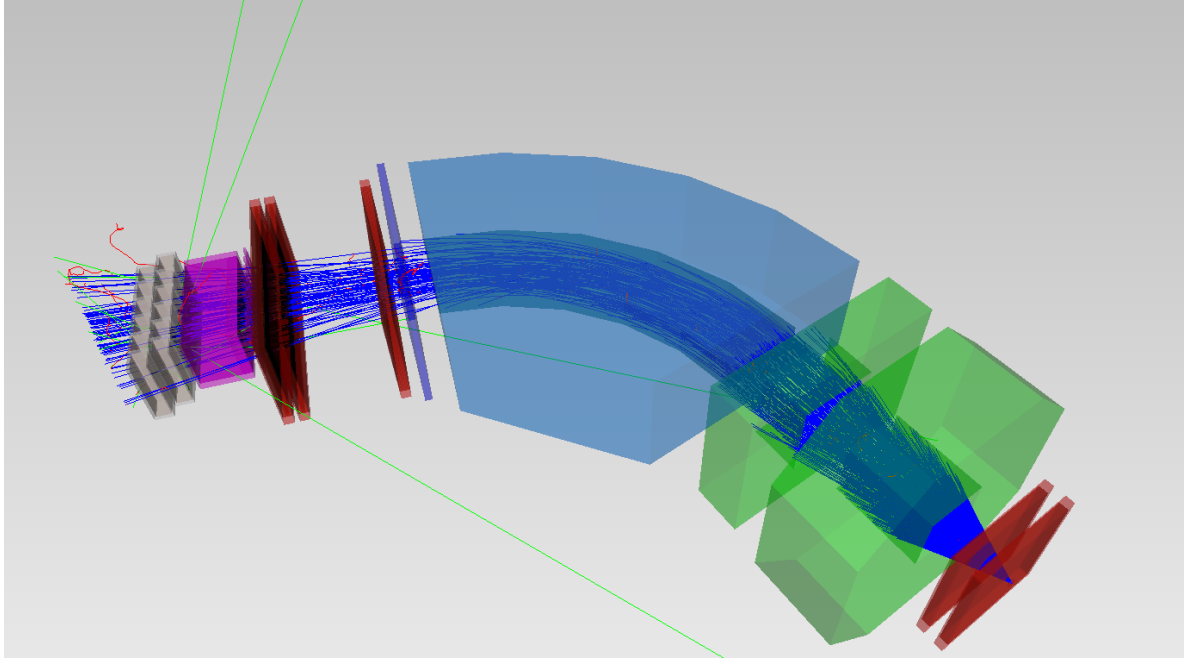


図 4.1: GEANT4 シミュレータ上で S-2S に  $K^+$  を射出した様子。  
 図中の青い線が  $K^+$  の軌跡を示す。

散乱粒子の飛跡情報  $x, y, u, v$  は、シミュレータ上の S-2S の上下流に設置した検出器から与えられる。磁場で粒子が曲げられる方向は  $x$  で、 $y$  は  $x$  とビーム軸  $z$  と垂直な鉛直成分である。角度  $u, v$  に関しては検出器で測定された運動量のベクトル比率より計算する。

$$u = \frac{p_x}{p_z} \quad (4.1)$$

$$v = \frac{p_y}{p_z} \quad (4.2)$$

また local hit 情報を訓練データとして用いる場合、複数の検出器を現実のドリフトチェンバーのワイヤ面に見立ててデータ取得を行った。

シミュレーション上での  $K^+$  の入射領域、つまり  $(K^-, K^+)$  反応が生じる領域は、標的のサイズと K1.8 ビームラインスペクトロメータから射出される  $K^-$  ビームの広がり具合から決定した。まず炭素標的 (AFT) のサイズが  $100^x \times 50^y \times 100^z$  mm であることと、E05 実験での  $K^-$  ビームの広がり (図 4.2) から、 $K^+$  の生成領域は  $x, y$  に関してはガウス分布に従い、 $z$  (ビーム軸方向) に関しては  $-50 \text{ mm} \leq z \leq +50 \text{ mm}$  の範囲で一様に分布するとした。 $x, y$  のガウス分布の幅は  $\sigma_x = 17.0 \text{ mm}$ ,  $\sigma_y = 3.4 \text{ mm}$  とした (図 4.3)。

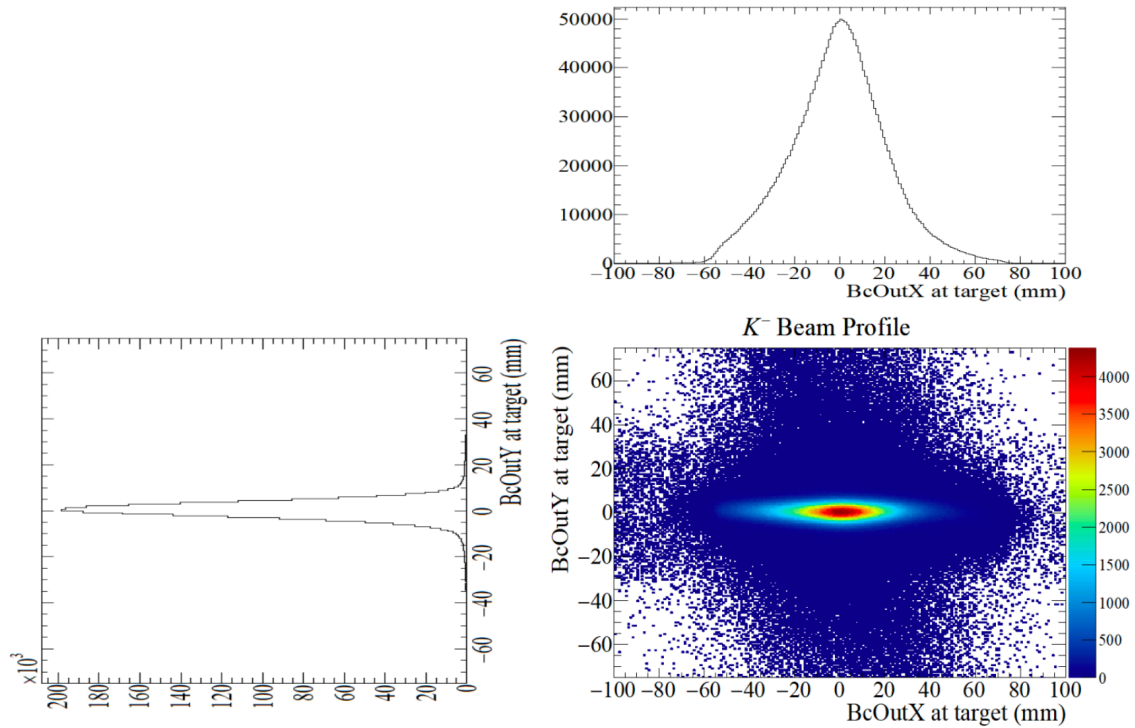


図 4.2: E05 実験での標的位置での  $K^-$  ビームの空間的広がり。

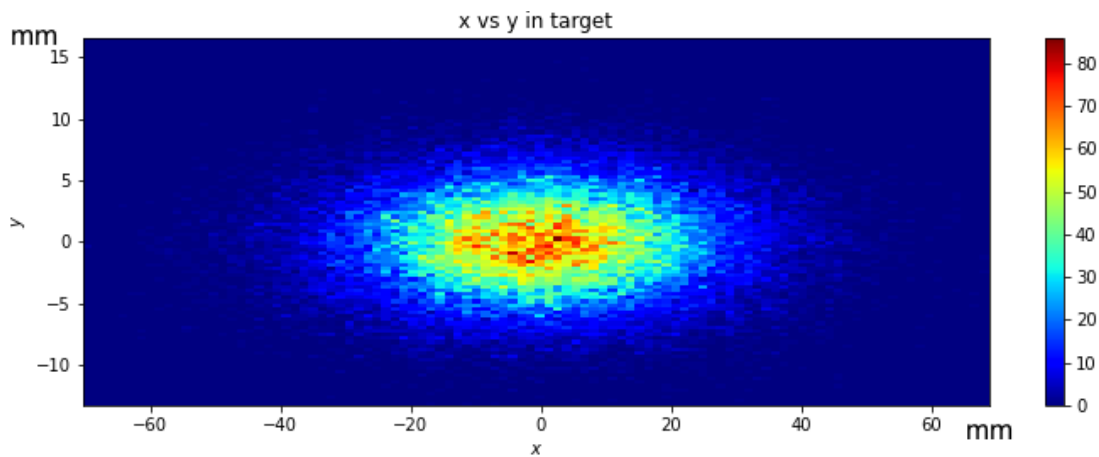


図 4.3: 本研究で用いるシミュレータの標的位置での  $K^+$  の生成点の空間的広がり。

散乱角度  $\theta$  を  $0 \sim 20$  deg の範囲として射出する。また、 $K^+$  の運動量は、 $1.2\text{--}1.4$  GeV/ $c$  および  $1.2\text{--}1.6$  GeV/ $c$  の範囲を一様にとる。これは、後にシミュレーションデータで構築した機械学習模型を現実のデータで校正するときに、 $< 1.31$  GeV/ $c$  の  $\Xi^-$  ピークのデータと、 $< 1.58$  GeV/ $c$  の  $\Sigma^-$  ピークのデータを用いるためである。(図 4.4)

機械学習の訓練データには、以上の条件でシミュレーション空間に射出され、S-2S 電磁石中を通過した  $K^+$  のみ用いる。

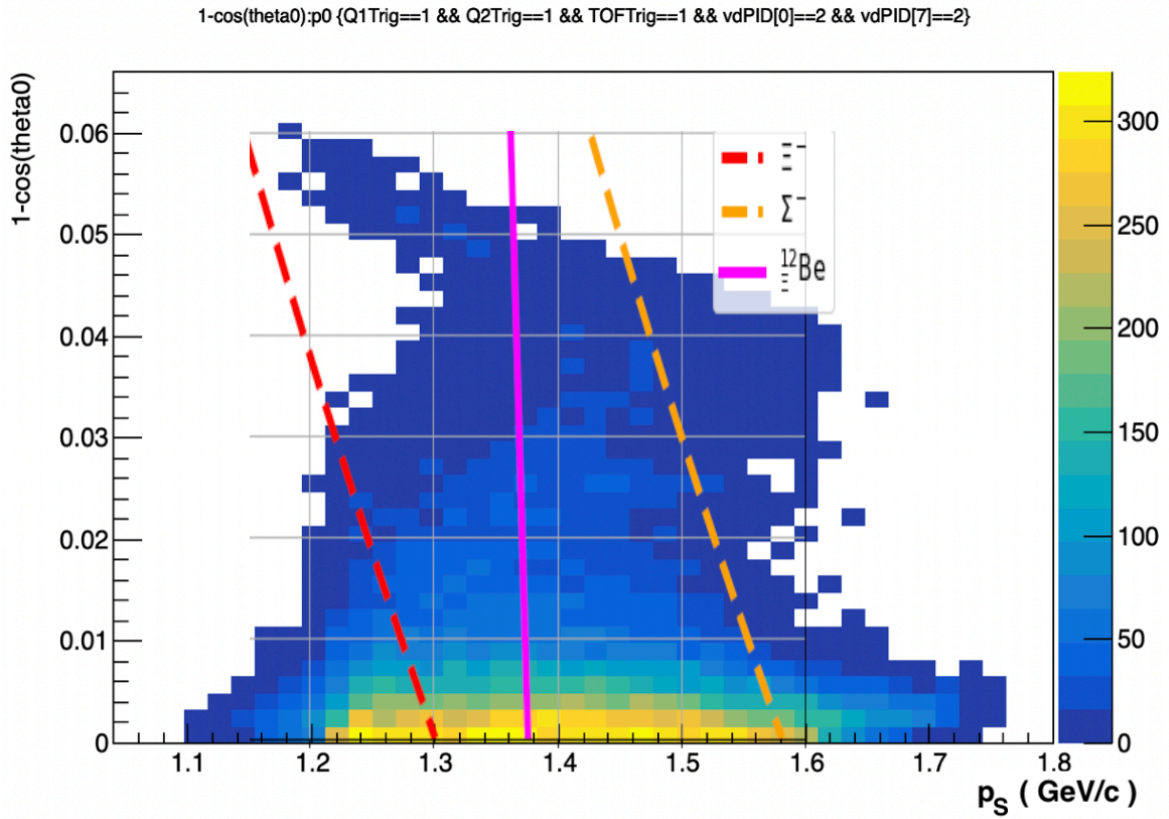


図 4.4: S-2S の運動量アクセプタンス領域と  $p(K^-, K^+)\Xi^-$ ,  $p(K^-, \pi^+)\Sigma^-$  (点線) と  $^{12}\text{C}(K^-, K^+)\Xi^-$  の運動学に基づいた運動量-角度 ( $p_S - (1 - \cos \theta)$ ) 分布。

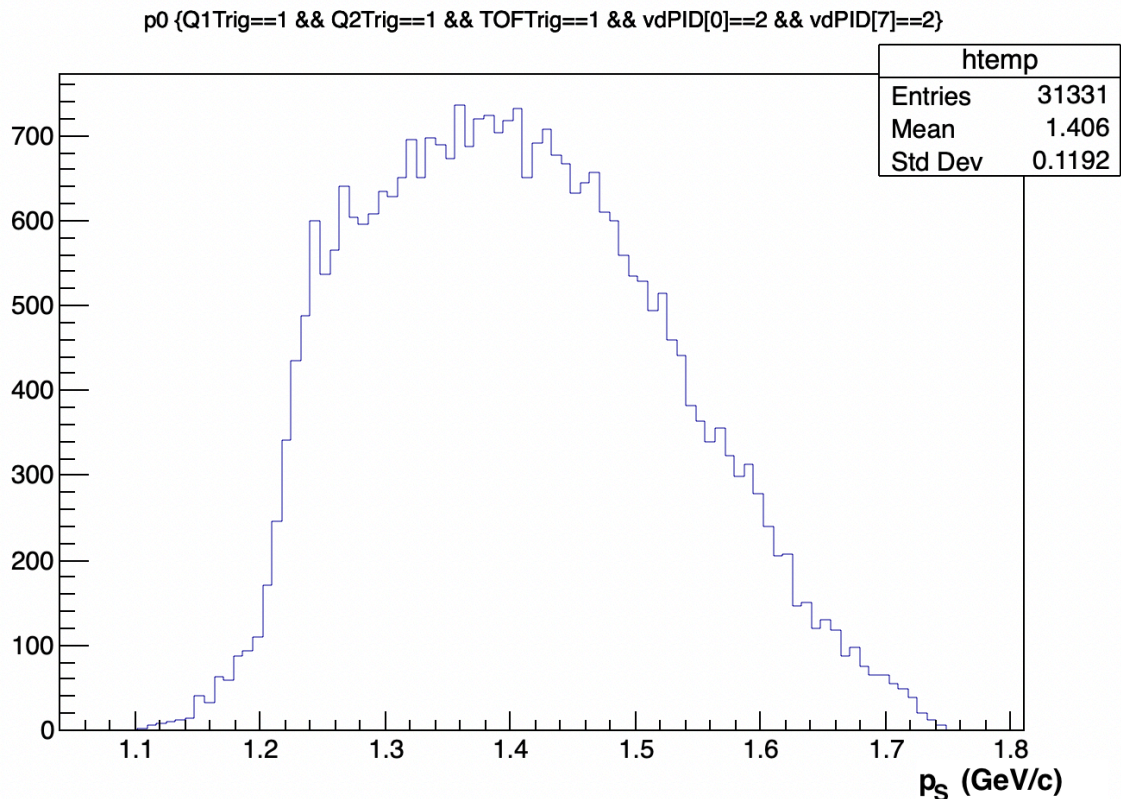


図 4.5: 最大磁場での S-2S 磁気スペクトロメータの運動量のアクセプタンス。  
アクセプタンス領域は、1.2–1.6 GeV/ $c$  にわたる。

## 4.2 機械学習ライブラリ

本研究においては、機械学習モデルの構築のために TensorFlow, keras ライブラリを使用した。なおバージョンについては TensorFlow2.9.2[51]、keras 2.9.0[52] である。TensorFlow は Google により開発された機械学習のオープンライブラリである。また、Keras は Python で書かれたニューラルネットワークに関するライブラリであり、TensorFlow を含む 様々なライブラリの上で動作する。なお実行は Google Colab 上で行った。

## 4.3 機械学習モデルの構築

機械学習モデルにおいては 1 層あたりのノード数、隠れ層の数、活性化関数、最適化関数など多くのパラメータが存在する。これらのパラメータを変えながら散乱粒子の飛跡情報

から求めた計算運動量  $p_S^{\text{ML}} - p_S^0$  の誤差が小さくなる最適な訓練パラメータを探索し、機械学習モデルの構築を行った。

訓練パラメータは、以下の3つの条件で選定を行う。

- 運動量分解能が小さく抑えられていること。
- 過学習が抑えられていること。
- 可能な限り単純なパラメータ値を選ぶこと。

単に運動量誤差が小さくなるだけでなく、損失関数の振る舞いや訓練データとテストデータから計算した運動量分解能の差分から、過学習が生じていないか判断し最適なパラメータを探索する。また、パラメータ値は可能な限り小さく複雑ではない値とし、機械学習モデルが過剰に複雑にならないように注意する。これは、後のエネルギー校正の際に Beam Through と  $p(K^-, K^+)E^-$  データから機械学習モデルを校正し、 $E^-$  ハイパー核ピークをそれらのデータで外挿して求めるが、モデルが複雑すぎる場合、外の運動量空間に外挿できず十分に校正できない可能性があるためである。

#### 4.3.1 データの前段階処理

機械学習モデルの構築に用いる訓練データに関しては、入力層に入れる前にデータ処理を施している。シミュレータで取得した  $x, y, u, v$  に対してまず検出器誤差で歪ませる。その後2次以上の飛跡情報の項を追加するための処理（高次項追加）を行い、最後に正規化処理を行った（図 4.6）。

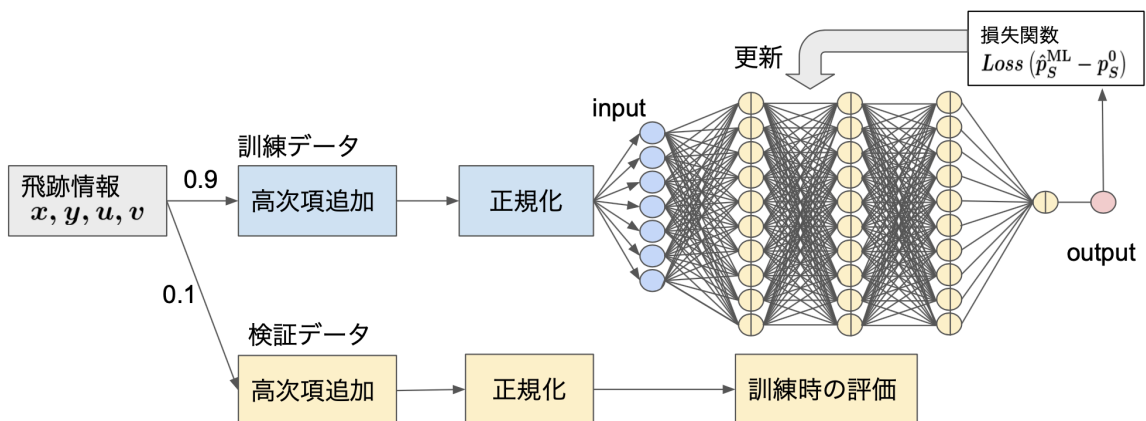


図 4.6: 機械学習の訓練の前の入力データの前段階処理のフローチャート。



また過学習を監視するため、GEANT4 より得られたデータセットを 9:1 の割合で学習用の訓練データと、評価データに無作為に分けた。訓練データを機械学習モデルの構築に用い、評価データを用いて過学習の有無を観測する。この評価データから計算した epoch ごとの損失関数の大きさから、early stopping や learning rate の値を決定する。また、これらとは別に学習に全く関与しないデータ（テストデータ）も用意する。運動量分解能に関してはこれらのデータにより評価する。

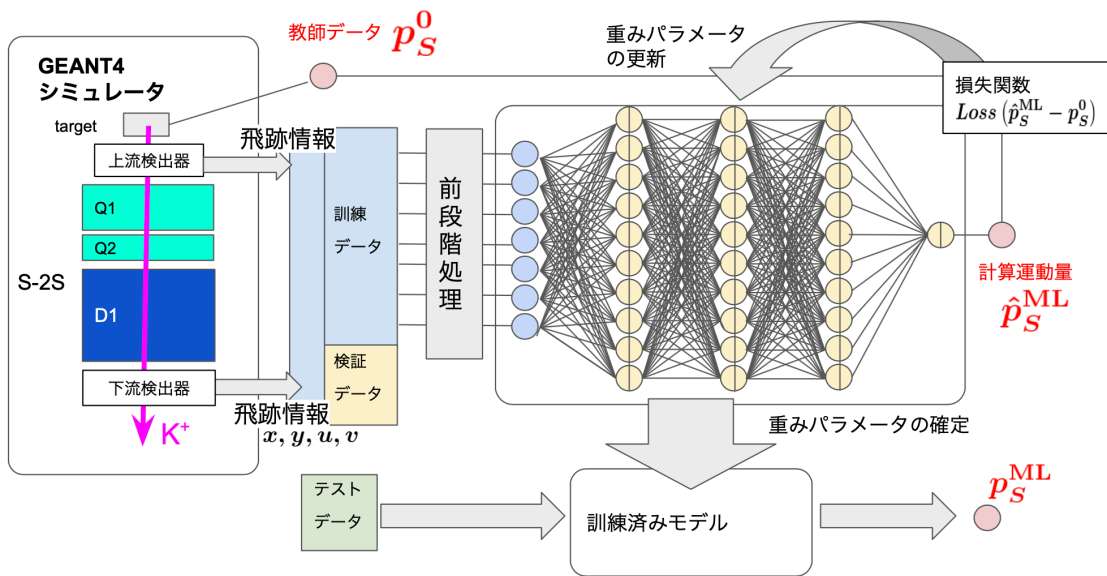


図 4.7: 本修論における機械学習モデルの概念図。訓練データの他に、過学習対策のために訓練中の損失関数を監視するための評価データと、訓練後に過学習が生じているか評価するためのテストデータを与える。

なお  $K^+$  の運動量は  $1.2 - 1.4 \text{ GeV}/c$  の範囲で射出させる。 $K^+$  の散乱角度は  $0 \leq \theta \leq 20 \text{ deg}$  の範囲でとる。この運動量-角度範囲で射出した  $K^+$  のうち、QQD 電磁石を通過し検出器の有感領域に入ったイベントのみを訓練データに用いる。

学習率 (learning rate) および epoch (繰り返し数) は、値を固定せず次の様にスケジューリング設定した。学習率に関しては、初期値を  $1.0 \times 10^{-3}$  とし学習が進むに伴い学習率を小さくするように設定した。10 epoch の間損失関数の値が、平均絶対誤差を採用した場合  $1.0 \times 10^{-6}$ 、平均二乗誤差を採用した場合は  $1.0 \times 10^{-9}$ 、の分改善されない場合は、学習率を 0.8 倍になるように設定する。epoch は early stopping を採用する。50 epoch の間学習を繰り返し損失関数に改善が見られない場合、学習を途中で強制終了するように設定した。



また、過学習対策として他にドロップアウトや L1, L2 正則について述べたが、損失関数の訓練時の推移や、訓練データとテストデータのそれぞれで計算した運動量分布から、過学習は認められないと判断し、今回の最適化のための訓練パラメータの中には加えなかった。

### 4.3.2 ノード数の最適化

まず隠れ層 1 層あたりのノード数の最適化を行った。隠れ層 1 層あたりのノード数を変えながら、 $K^+$  の運動量を求め、最適なパラメータ値を探索した。なお他のパラメータの値や関数の種類は固定して探索を行った。訓練データについて、統計量は  $9.53 \times 10^5$ 、入力データには 3 次の項まで追加した。訓練データに用いた運動量領域は  $1.2 - 1.4 \text{ GeV}/c$  とした。他のモデルパラメータに関して、活性化関数は swish、隠れ層は 3 層、バッチサイズは 512、損失関数は絶対平均誤差、最適化関数は Adam を用いた。

最適化の指標としては、訓練データと過学習評価用のテストデータの 2 つのデータセットで計算した運動量分解能を用いる。これら 2 種類の運動量分解能を調べることで、運動量を高分解能で見積もられているか、と過学習の有無を評価し、最適なパラメータ値を決定する。

この 2 つのデータセットを用いた検証は、以降のパラメータ最適化においても同様に用いる。

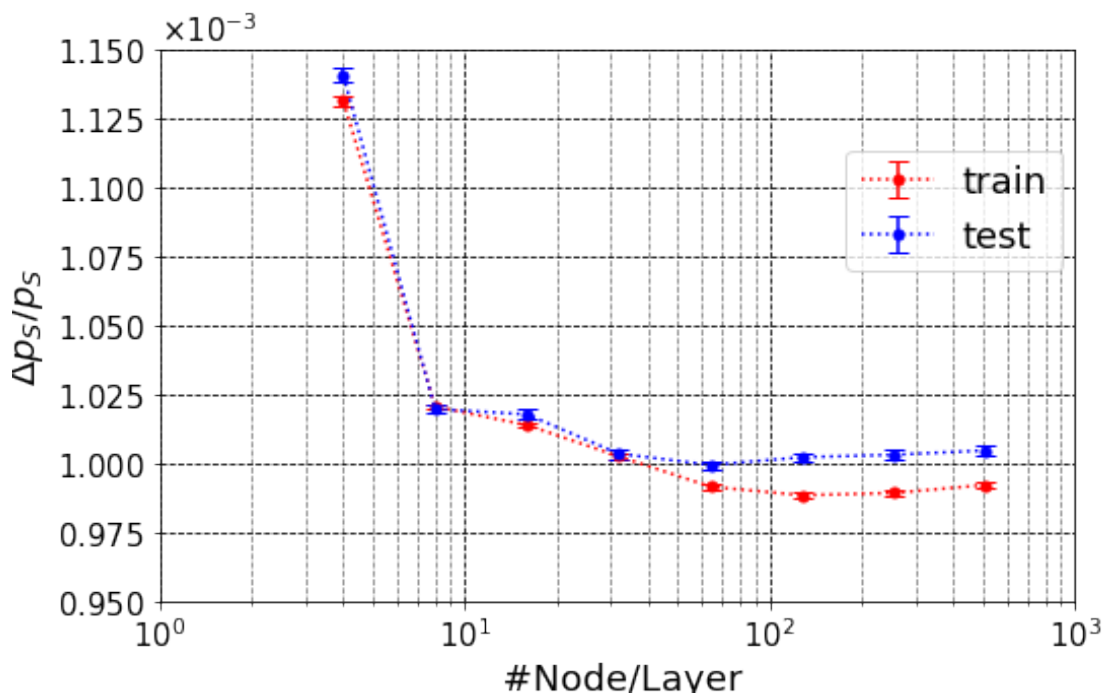


図 4.8: 運動量分解能 (半値全幅、FWHM) の隠れ層 1 層あたりのノード数依存性。

ノード数を増やすとパラメータが複雑化し、図 4.8 のようにノード数が 64 以上になると訓練データで計算した運動量分解能 (赤) と、評価データで計算した運動量分解能 (青) の間に差が生じる。これはある種の過学習である。逆にパラメータを単純化しすぎた場合は、訓練の進捗が悪くなり最適な重みパラメータに収束しなくなる。

今回の検証によると、1 隠れ層あたりのノード数が 32 の場合に、過学習があまり認められず、また重みパラメータが十分最適化されていることが認められた。

### 4.3.3 隠れ層の数の最適化

次に隠れ層の数の最適化を行った。ノード数の最適化においても述べたが、パラメータを複雑化しすぎると過学習が生じ、逆に単純化しすぎると学習が進まず最適な重みパラメータに収束しなくなり、結果として運動量分解能が悪くなる。これらの事象が生じにくい最適な層数を選定するために、層の数を 1 ~ 5 まで変化させて学習を行い、運動量分解能を評価した。

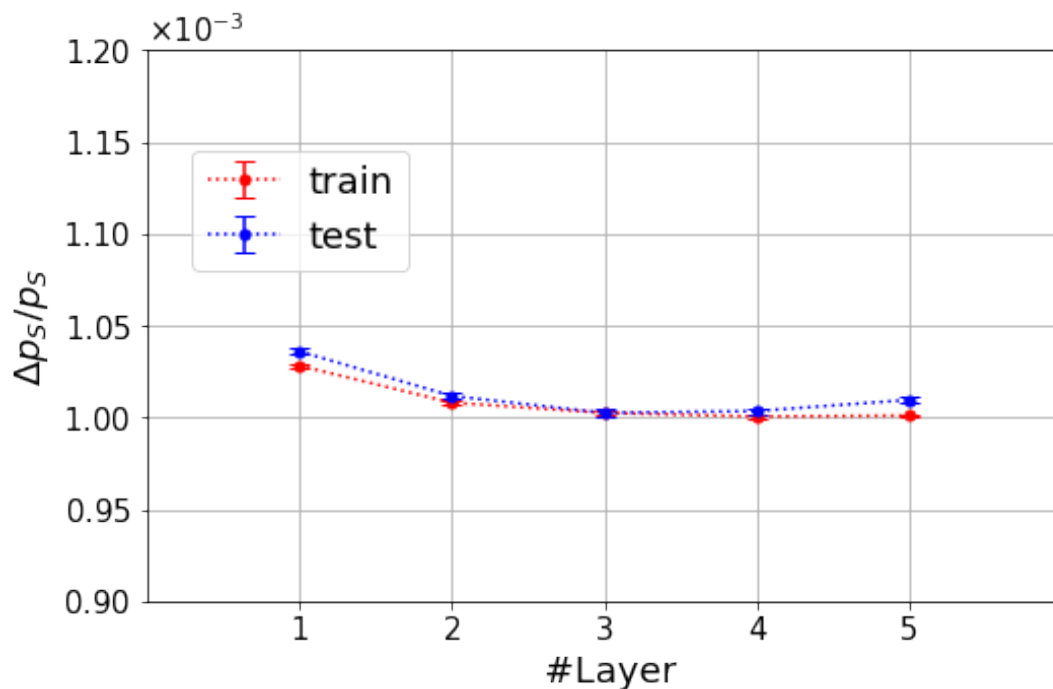


図 4.9: 運動量分解能 (半値全幅、FWHM) の隠れ層数依存性。

隠れ層数の検証においても、層数を増やすとパラメータが複雑化し、図 4.9 のように 4 以上になると訓練データで計算した運動量分解能 (赤) と、評価データで計算した運動量分解能 (青) の間に差が生じ、過学習が発生する。逆にパラメータを単純化しすぎた場合は、訓練の進捗が悪くなり最適な重みパラメータに収束しなくなる。

今回の検証によると、隠れ層数が 3 の場合に、過学習があまり認められず、また重みパラメータが十分最適化されていることが認められた。

#### 4.3.4 活性化関数の最適化

活性化関数は、よく ReLU(Rectified Linear Unit) (式 4.3) や sigmoid 関数 (式 4.4) と呼ばれる関数が用いられる。

$$y = \begin{cases} x & (x \geq 0) \\ 0 & (x < 0) \end{cases} \quad (4.3)$$

$$y = \frac{1}{1 + e^{-x}} \quad (4.4)$$

しかし、誤差逆伝播でパラメータ更新をしていくと、下位層で勾配が緩やかになっていく。このため更新は進まず、パラメータは最適値に収束しづらくなる。これを勾配消失と呼ばれる。逆に勾配が急になりパラメータが発散する、勾配爆発と呼ばれる問題も存在する。特に sigmoid 関数や tanh 関数は、入力が 0 近傍においては線形な振る舞いをするが、入力の絶対値が大きくなると出力が 0, 1 に飽和する (図 4.10)。誤差逆伝播では活性化関数の導関数を用いたため、導関数の値が限りなく 0 に近づき、下位層の勾配はほぼ 0 になり更新がなされない。

そのため非飽和関数である ReLU がよく用いられるようになった。この関数は入力が負の場合は 0 を出力し、正の場合はそのまま入力値を出力する。しかし、ReLU は負の値で 0 しか出力しないため、多くのニューロンが訓練中に 0 しか出力しない問題が生じる場合がある。これを防ぐために負入力に対して、ハイパーパラメータ  $\alpha$  を用いて  $\alpha x$  を出力する関数を用いることがある (Leakly ReLU 関数)(式 4.5)。またこれを応用し学習中に  $\alpha$  をも訓練し最適化する parametric leakly ReLU 関数 (PReLU) も用いられる。

$$y = \begin{cases} x & (x \geq 0) \\ \alpha x & (x < 0) \end{cases} \quad (4.5)$$

また、ReLU 関数は、 $x = 0$  について微分可能ではない。この問題を負入力に対して指数関数を用いた出力を用いることで、あらゆる点で滑らかな出力を与えることができる (ELU, Exponential Linear Unit)(式 4.6)。これらの ReLU 関数の派生は勾配消失や、負入力での問題を緩和するものだが、同時に過学習を起ししやすい問題がある。

$$y = \begin{cases} x & (x \geq 0) \\ \alpha(e^x - 1) & (x < 0) \end{cases} \quad (4.6)$$

swish 関数は [53]、前述の sigmoid 関数に  $x$  をかけた、滑らかで非単調な関数である (式 4.7)。入力が正の場合においては ReLU と同様の挙動を示し、負方向に対しては 0 に収束する。ニューラルネットワーク上では一貫して ReLU と一致するかそれを上回る性能を持つとされている。

$$y = \frac{x}{1 + e^{-\beta x}} \quad (4.7)$$

なお下の図 4.10 に各活性化関数の入出力の関係を示す。

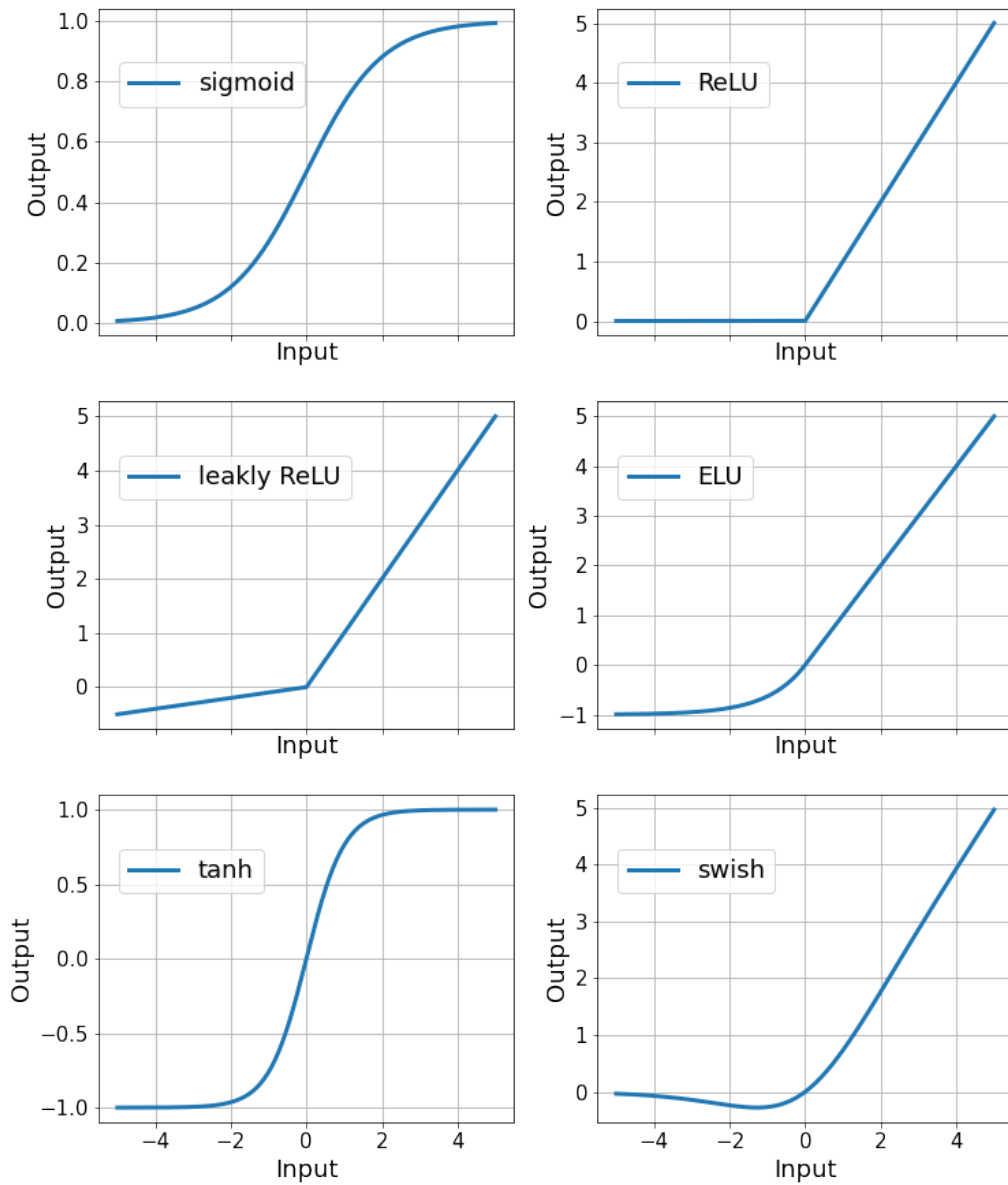


図 4.10: 様々な活性化関数の出力。

活性化関数の最適化においては、他パラメータを次の表の値で固定し検証した。

表 4.1: 活性化関数最適化のためのパラメータ設定。

他パラメータ条件			
#Nodes/Layer	#Layers	損失関数	最適化関数
32	3	MSE	Adam

表 4.2: 活性化関数を変化させた場合に、機械学習モデルで計算される運動量分解能。

活性化関数	$\Delta p_S/p_S(\text{train})$	$\Delta p_S/p_S(\text{test})$
sigmoid	$1.06 \times 10^{-3}$	$1.07 \times 10^{-3}$
tanh	$1.02 \times 10^{-3}$	$1.02 \times 10^{-3}$
ReLU	$1.06 \times 10^{-3}$	$1.06 \times 10^{-3}$
Leakly ReLU	$1.04 \times 10^{-3}$	$1.04 \times 10^{-3}$
PReLU	$1.05 \times 10^{-3}$	$1.06 \times 10^{-3}$
ELU	$1.02 \times 10^{-3}$	$1.02 \times 10^{-3}$
Swish	$1.02 \times 10^{-3}$	$1.02 \times 10^{-3}$

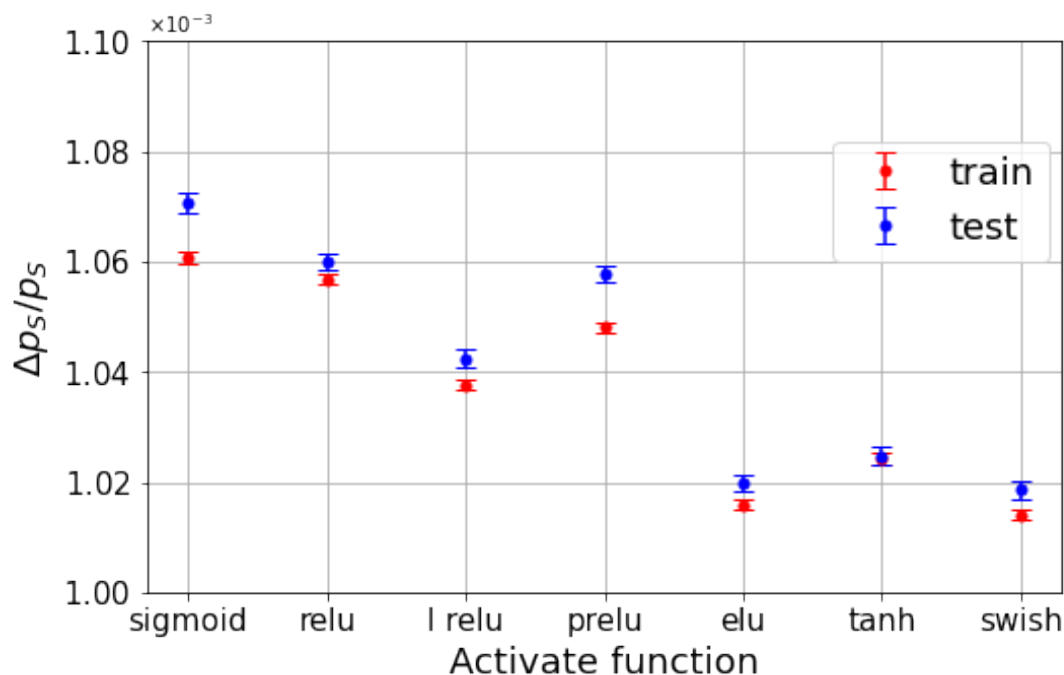


図 4.11: 運動量分解能 (半値全幅、FWHM) の活性化関数依存性。

検証の結果、過学習が生じず、運動量分解能に関しても収束化している swish 関数を選定した。

#### 4.3.5 最適化関数の最適化

重みパラメータ  $w$  を更新する際には、損失関数の微分  $\nabla_w L$  だけ更新を行う。この損失関数に対してどのような微分を行い、どのようにパラメータ更新を行うかによって、学習の進捗具合または結果が異なる。例えば最も基本的な最適化関数として SGD (Stochastic Gradient Descent: 確率的勾配降下法) が挙げられる。この関数は次のような式を基にパラメータ更新を行う。

$$t \leftarrow t - 1 \quad (4.8)$$

$$w_t \leftarrow w_{t-1} - \eta \nabla_w L \quad (4.9)$$

$t$  は訓練のイテレーションを表す。ミニバッチ学習においては、(1 エポックにおけるミニバッチ数) × (エポック数) = (イテレーション数) であり、何番目のイテレーションかを指す。式 4.9 場合、以前の勾配の値を考慮していないため、局所的な勾配が緩やかである場

合学習の進捗が遅くなってしまう。これを抑制するため Momentum 法と呼ばれる方法を用いる。この方法では、慣性ベクトル  $\vec{m}$  と呼ぶものを用いて  $w$  を更新する。まず  $\vec{m}$  は、局所的な勾配  $\nabla_w L$  で更新される。この更新された  $m$  を用いて  $w$  を更新する (式 4.13)。

$$\beta \in [0, 1] \quad (4.10)$$

$$t \leftarrow t - 1 \quad (4.11)$$

$$\vec{m}_t \leftarrow \beta \vec{m}_{t-1} - (1 - \beta) \eta \nabla_w L \quad (4.12)$$

$$w_t \leftarrow w_{t-1} + \vec{m}_{t-1} \quad (4.13)$$

$\beta$  は、0 から 1 までの範囲をとる。 $\beta = 0$  の時単純な勾配降下法に相当する。

またこの Momentum 法を改善した Nesterov の加速法 (Nesterov's accelerated gradient, NAG) と呼ばれるアルゴリズムがある [54]。Nesterov 加速法では、Momentum 法よりもさらに収束への加速を増したものである。勾配の値を評価する位置を  $w_t + \beta w_{t-1}$  としている (式 4.16)。後述の Nadam にて実装されている。

$$t \leftarrow t - 1 \quad (4.14)$$

$$\vec{m}_t \leftarrow \beta \vec{m}_{t-1} - (1 - \beta) \eta \nabla_{w_t + \beta w_{t-1}} L \quad (4.15)$$

$$w_t \leftarrow w_{t-1} + \vec{m}_{t-1} \quad (4.16)$$

Momentum 法や NAG は振動抑制と収束加速に関するアルゴリズムであったが、収束方向に関しては考慮していない。ニューラルネットワークにおいては、学習率  $\eta$  が小さすぎると収束に時間がかかり、大きすぎても値が発散してしまう。そこで機械学習においては、学習率の初期値を大きく見積もり、徐々に小さくしていく方策がよく取られる。AdaGrad[55] では、この学習率の最適化をパラメータ一つ一つに対して行う。具体的には次の式 4.18 に従い更新を行う。過去の勾配を要素ごとに 2 乗和  $s$  として保存し、パラメータ更新の際に  $1/\sqrt{s}$  をかけることで学習をパラメータごとに抑制している。例えばパラメータ空間において、独立な  $w_1, w_2$  に対して勾配  $\nabla_{w_1} L < \nabla_{w_2} L$  である時、 $w_1$  よりも  $w_2$  の学習率が抑制される。これにより重みパラメータが全空間に対する極小値へ向かって更新されていき、効率よく収束される。

$$\vec{s}_t \leftarrow \vec{s}_{t-1} + \nabla_w L \otimes \nabla_w L \quad (4.17)$$

$$w_t \leftarrow w_{t-1} - \eta \frac{\nabla_w L}{\sqrt{s_{t-1} + \epsilon}} \quad (4.18)$$



しかし、Adagrad は勾配の最小値に到達する前に学習率が低下し、収束しきれない問題がある。これを改善したものが RMSProp[56] と呼ばれるアルゴリズムである。

$$\vec{s}_t \leftarrow \mu \vec{s}_{t-1} + (1 - \mu) \nabla_w L \otimes \nabla_w L \quad (4.19)$$

$$w_t \leftarrow w_{t-1} - \eta \frac{\nabla_w L}{\sqrt{\vec{s}_{t-1} + \epsilon}} \quad (4.20)$$

式 4.20 中の  $\epsilon$  は、分母が 0 になることで発散するのを防ぐための定数である。RMSProp では、以前のイテレーションの勾配を考慮しそれを足し込むことで、最小値に収束しきれない問題を解決している。また、RMSProp では式 4.20 から分かるように両辺で次元数の整合が取れていない。Adadelta はこの問題を無次元化により解消した関数である [57]。

この RMSProp と Momentum 法の手法を組み合わせたアルゴリズムが、Adam(Adaptive Moment Estimation) である (式 4.25)。

$$\vec{m}_t \leftarrow \beta \vec{m}_{t-1} - (1 - \beta) \eta \nabla_w L \quad (4.21)$$

$$\vec{s}_t \leftarrow \mu \vec{s}_{t-1} + (1 - \mu) \nabla_w L \otimes \nabla_w L \quad (4.22)$$

$$\vec{m}_t \leftarrow \frac{\vec{m}_{t-1}}{1 - \beta^T} \quad (4.23)$$

$$\vec{s}_t \leftarrow \frac{\vec{s}_{t-1}}{1 - \mu^T} \quad (4.24)$$

$$w_t \leftarrow w_{t-1} + \eta \frac{\vec{m}_{t-1}}{\sqrt{\vec{s}_{t-1} + \epsilon}} \quad (4.25)$$

Adam のハイパーパラメータの  $\beta, \mu$  はよく  $\beta = 0.9, \mu = 0.999$  で設定されることが多い。今回の検証、それ以降の評価に関してもこの値を採用する。この Adam に Momentum 法ではなく Nesterov の加速法を適用させた勾配法が、Nadam[59] と呼ばれるアルゴリズムである。

本検証に関しては、SGD, Adagrad, RMSProp, Adam, Adadelta, Adamax, Nadam の最適化関数を用いて運動量分解能評価を行った。

他パラメータ条件

活性化関数	#Nodes/Layer	#Layers	損失関数
Swish	32	3	mse
最適化関数	$\Delta p_S/p_S$ (FWHM,train)	$\Delta p_S/p_S$ (FWHM, test)	実行時間
SGD	$2.65 \times 10^{-2}$	$2.64 \times 10^{-2}$	39m 5s
RMSProp	$1.05 \times 10^{-3}$	$1.05 \times 10^{-3}$	33m 55s
Adagrad	$2.42 \times 10^{-2}$	$2.40 \times 10^{-2}$	46m 57s
Adam	$1.02 \times 10^{-3}$	$1.02 \times 10^{-3}$	34m 5s
Adadelta	$2.89 \times 10^{-2}$	$2.90 \times 10^{-2}$	44m 58s
Adamax	$1.08 \times 10^{-3}$	$1.08 \times 10^{-3}$	36m 43s
Nadam	$1.02 \times 10^{-3}$	$1.02 \times 10^{-3}$	28m 30s

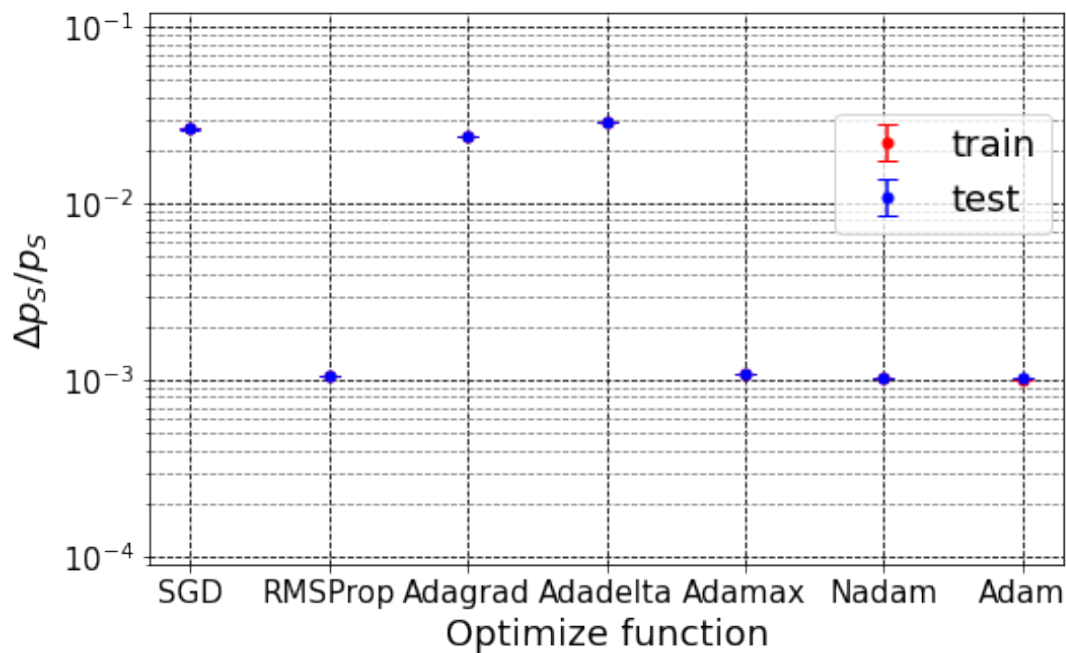


図 4.12: 運動量分解能 (半値全幅、FWHM) の最適化関数依存性。

最適化関数の分解能の挙動は二極化しており、RMSProp, Adam, Adamax, Nadam が計算誤差を比較的抑えることができることがわかった。次に学習時の損失関数の推移を

示す。

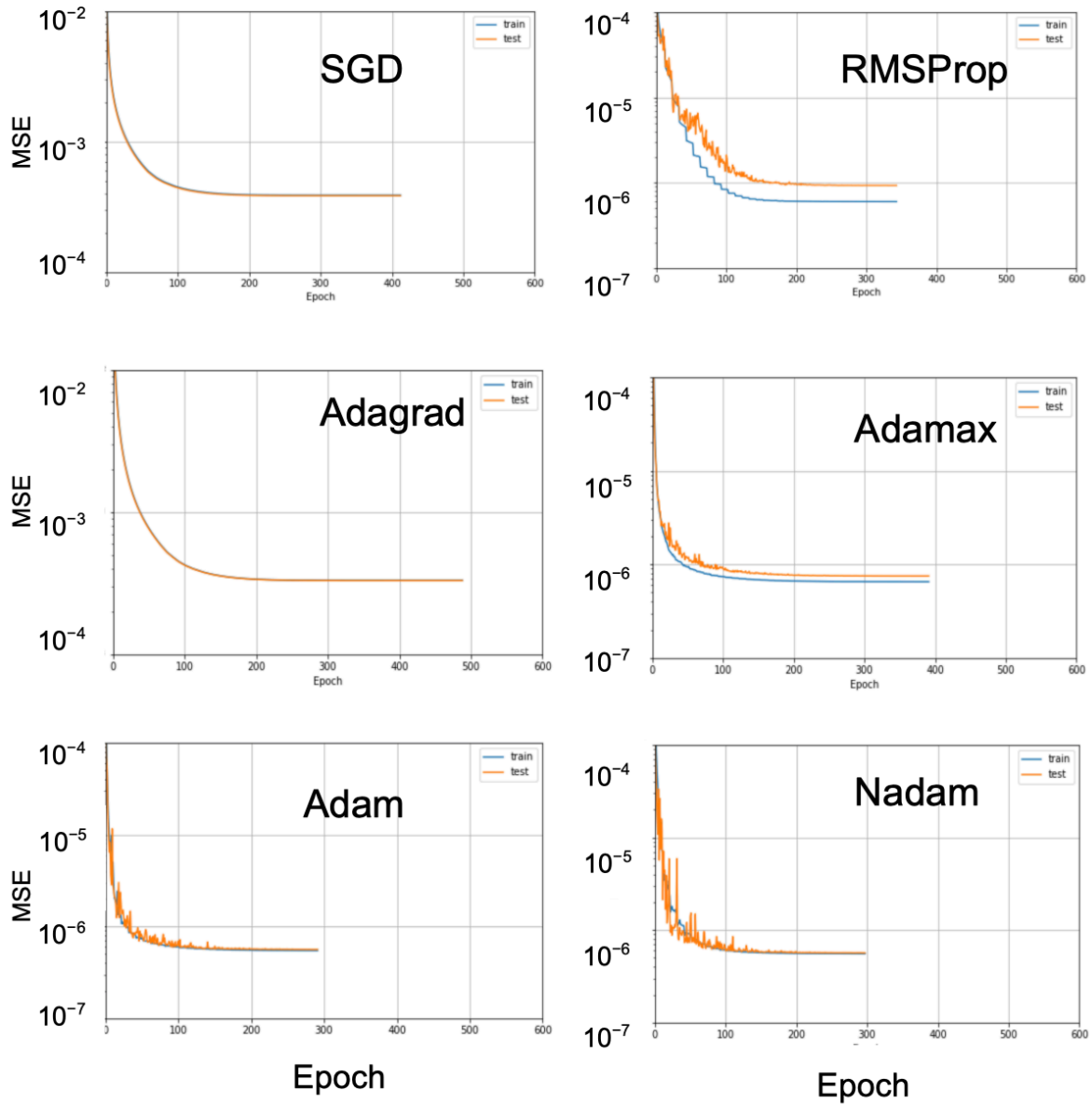


図 4.13: 最適化関数ごとの損失関数 (MSE) の推移。青が訓練データから計算された損失関数の推移、橙が検証データから計算された損失関数の推移を示す。

図 4.13 は、各最適化関数を用いた訓練における損失関数の推移を示す。RMSProp と Adamax を損失関数として使用した場合、訓練データによる損失関数と検証データによる損失関数の値が訓練の終了時に比較的差が生じている。つまり過学習が生じている。

SGD と AdaGrad に関しては、損失関数が比較的大きいまま最適パラメータに収束しきれていない。

したがって、最適化関数においては、過学習が比較的生じず、分解能が小さく抑えられた Adam 関数を選定した。

### 4.3.6 損失関数の最適化

回帰問題での損失関数は主に下記の平均絶対誤差と平均二乗誤差が用いられる。

- 平均絶対誤差 (Mean Absolute Error)

$$MAE = \sum_i |p_i^0 - \hat{p}_i^{ML}| \quad (4.26)$$

- 平均二乗誤差 (Mean Squared Error)

$$MSE = \sum_i (p_i^0 - \hat{p}_i^{ML})^2 \quad (4.27)$$

なお他ハイパーパラメータの値は次を採用した。

活性化関数	#Nodes/Layer	#Layers	最適化関数
Swish	32	3	Adam

訓練データの運動量領域に 1.2–1.4 GeV/ $c$  を用いた場合、機械学習モデルで計算した運動量分解能は表 4.3 のようになる。

表 4.3: 1.2–1.4 GeV/ $c$  の範囲での損失関数による運動量分解能 (FWHM) の差。

損失関数	$\Delta p_S/p_S$ (train)	$\Delta p_S/p_S$ (test)
MAE( $\times 10^{-3}$ )	1.00 $\pm$ 0.00	1.00 $\pm$ 0.00
MSE( $\times 10^{-3}$ )	1.02 $\pm$ 0.00	1.02 $\pm$ 0.00

訓練データの運動量領域に 1.2–1.6 GeV/ $c$  を用いた場合は表 4.4 のようになる。

表 4.4: 1.2–1.6 GeV/ $c$  の範囲での損失関数による運動量分解能 (FWHM) の差。

損失関数	$\Delta p_S/p_S$ ( train)	$\Delta p_S/p_S$ ( test)
MAE( $\times 10^{-3}$ )	$1.08 \pm 0.00$	$1.09 \pm 0.00$
MSE( $\times 10^{-3}$ )	$1.09 \pm 0.00$	$1.10 \pm 0.00$

どちらの運動量領域でも、平均絶対誤差の方が高分解能で運動量を再構成できることがわかった。

## 4.4 訓練データの最適化

### 4.4.1 訓練データの次数最適化

本修論では、機械学習の訓練データとして  $x, y, u = \partial x/\partial z, v = \partial y/\partial z$ 、教師データとして  $K^+$  の初期運動量である  $p_S^0$  を用いている。 $x, y, u = \partial x/\partial z, v = \partial y/\partial z$  について単に1次の項だけを訓練データとして用いても、複雑な相関は得られにくい。従来手法である逆輸送行列を用いた運動量解析 (3.1.2 節) においても、 $x^2, \dots, xy^2, \dots, uv^4, v^5, \dots$  のような3次、5次といった高次の効果を加えることにより、運動量を高分解能で求めることができる。しかし、高次の項を加えすぎても入力項の数が指数関数的に増加する (表 4.5) ことで、パラメータが複雑化し過学習が生じることが懸念されるため、どの程度の次数の項を加えると最適かを調べる必要がある。

最適化においては、シミュレーションで得られた  $x, y, u, v$  のデータに高次の項を加えたデータセットを訓練データとして用いる。1~5 次までの項を順に加え、それにより運動量分解能がどう振舞うかを調べた。

表 4.5: 次数と入力パラメータ数の関係。

最大次数	1	2	3	4	5
項数	8	45	165	495	1287

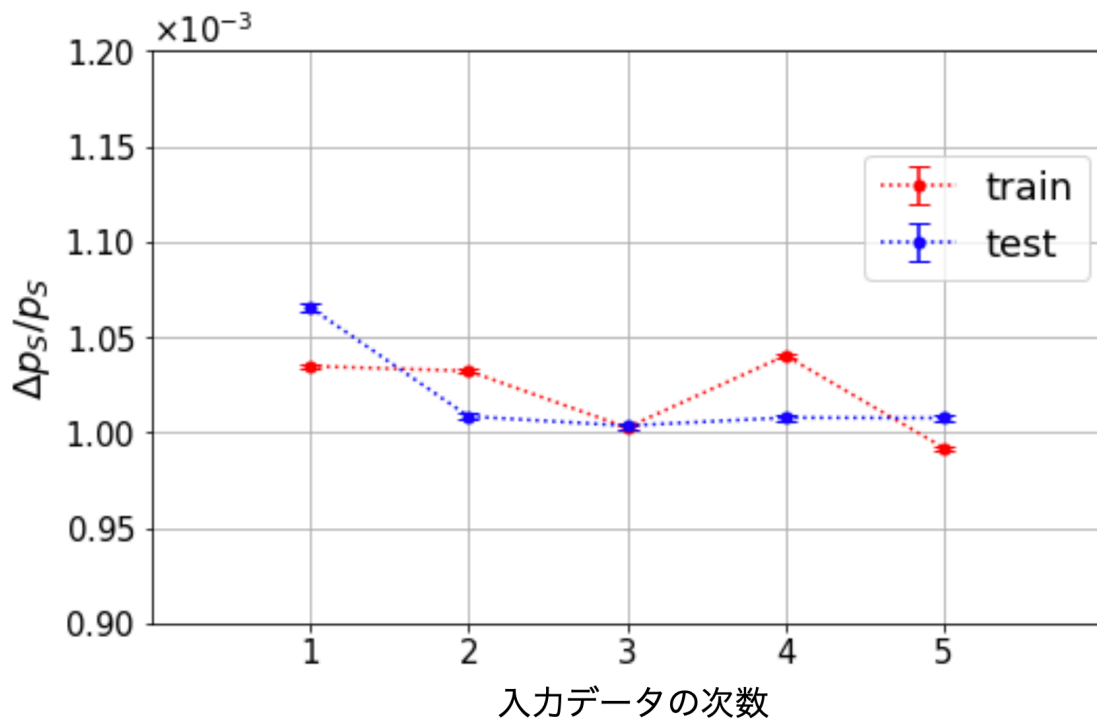


図 4.14: 運動量分解能 (半値全幅、FWHM) の入力データの次数依存性。

その結果、図 4.14 に見られるように、テストデータの分解能の推移に注目した場合、 $dim \geq 3$  で分解能が改善されなくなった。また  $dim = 3$  においては訓練データとテストデータから計算した運動量分解能の間の差が小さく、過学習が抑えられていると見られる。したがって、最適な入力データとしては 3 次の項まで取り入れるれば妥当であると考えられる。

#### 4.4.2 訓練データの統計量

次に訓練データの統計量の検証を行った。機械学習では、訓練データの統計量が大きいほど過学習は生じにくく、また学習の進捗が大きくなる。しかし、統計量を上げすぎるとプログラムの実行に膨大な時間がかかるなどソフトウェア面での制約が課せられる。

本検証では、 $9.53 \times 10^5$ ,  $5.50 \times 10^5$ ,  $1.37 \times 10^5$ ,  $4.98 \times 10^4$ ,  $1.85 \times 10^4$  の統計量の訓練データで学習を行い、訓練データ、テストデータから計算される運動量分解能の誤差を評価した。この評価によって要求される統計量を決定する。

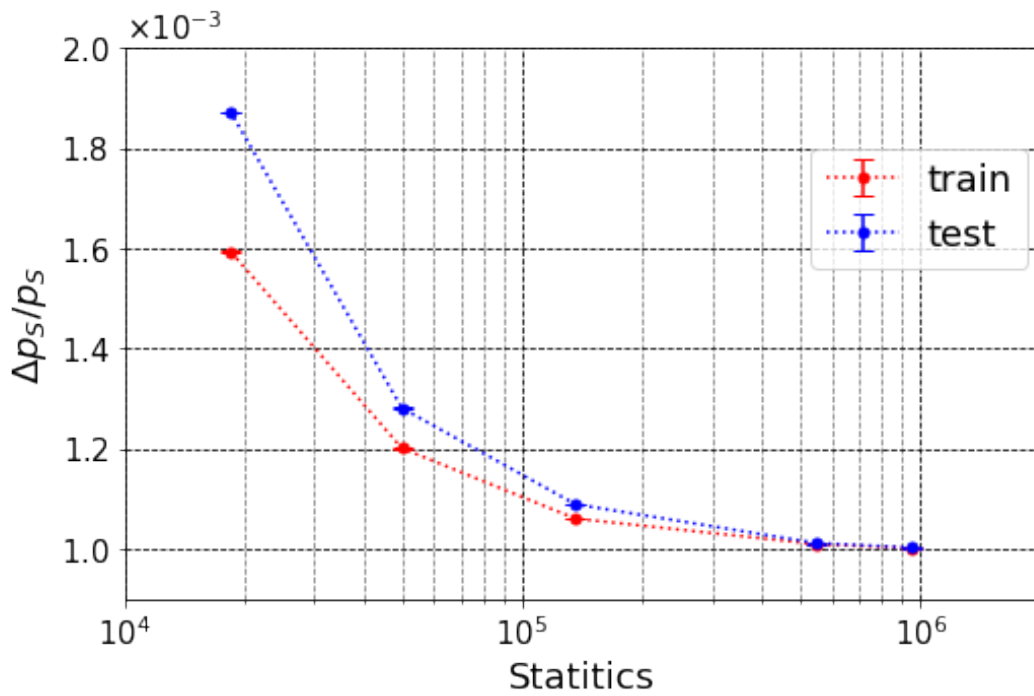


図 4.15: 運動量分解能 (半値全幅、FWHM) の統計量依存性。

各統計量での学習の際にかかった実行時間を下の表に示す。

表 4.6: 入力データの統計量と運動量分解能 (半値全幅、FWHM)。

統計量	$\Delta p_S/p_S$ (train)	$\Delta p/p$ (test)	実行時間
$1.85 \times 10^4$	$1.59 \times 10^{-3}$	$1.87 \times 10^{-3}$	24m45s
$4.98 \times 10^4$	$1.20 \times 10^{-3}$	$1.28 \times 10^{-3}$	17m33s
$1.37 \times 10^5$	$1.06 \times 10^{-3}$	$1.09 \times 10^{-4}$	5m2s
$5.50 \times 10^5$	$1.01 \times 10^{-3}$	$1.01 \times 10^{-3}$	2m
$9.53 \times 10^5$	$1.00 \times 10^{-3}$	$1.00 \times 10^{-3}$	54.4s

図 4.15 に見られるように、統計量を  $1 \times 10^6$  程度要求すれば、運動量分解能は  $\Delta p/p \simeq 1.00 \times 10^{-3}$  (半値全幅、FWHM) に収束することがわかった。

## 4.5 最適化した訓練パラメータセットの評価

以上の検証の結果、以下のパラメータセットにより学習した機械学習モデル (表 4.7) を採用した。

表 4.7: 機械学習モデルのパラメータ設定の検証により構築したモデルパラメータ群。今後の検証には、基本的にこのパラメータセットでスペクトロメータの評価等を行う。

入力データ	
要求データセット統計量	$\gtrsim 10^6$
モデルパラメータ	
活性化関数	Swish
ノード数/レイヤー	32
隠れ層数	3
入力次数	3
入力パラメータ数	165
総 weight パラメータ数	7788
損失関数	平均絶対誤差 (MAE)
最適化関数	Adam

この時の運動量分解能は訓練データに 1.2–1.4 GeV/ $c$  の運動量領域を用いた場合、 $\Delta p_S/p_S = (1.00 \pm 0.00) \times 10^{-3}$  (半値全幅、FWHM) となった。

下に選定したパラメータでの学習の推移を示す (図 4.16)。図の横軸が epoch であり、全訓練データを用いて何回パラメータの更新を繰り返したかを示す。縦軸は各 epoch 時点での損失関数 (MAE) の値を示す。図中の青で示された線は訓練データで計算した損失関数 (MAE) の推移を示しており、橙で示した線は過学習モニター用の評価データで計算した損失関数 (MAE) の推移を示す。



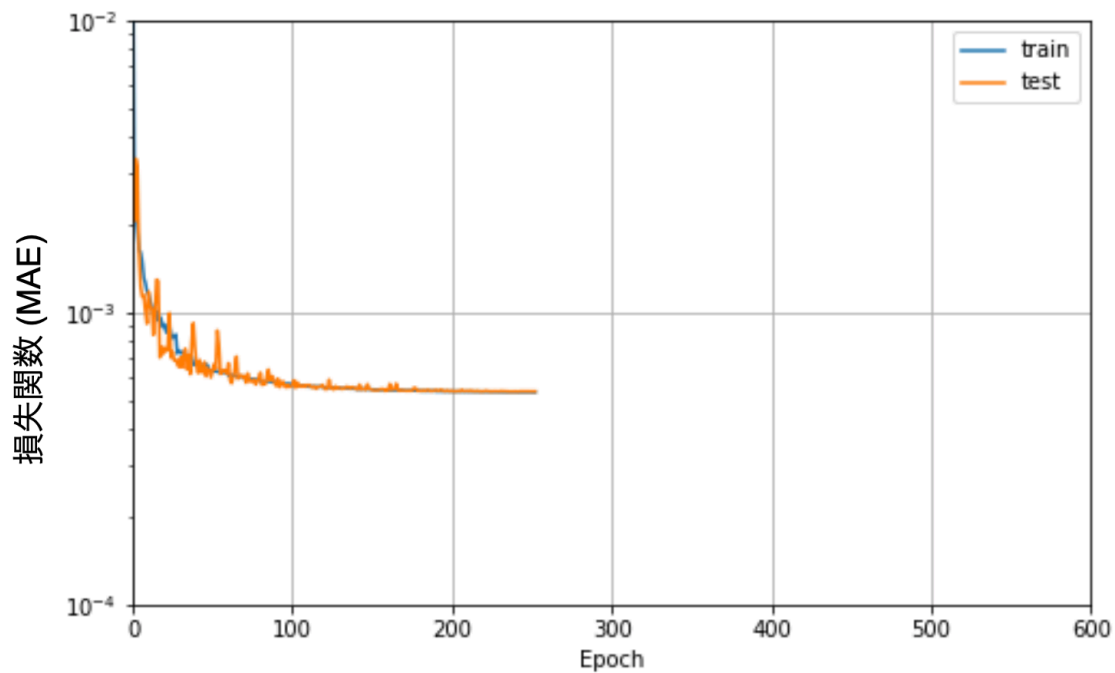


図 4.16: 損失関数の訓練時の推移。青は訓練データで、橙は検証データでそれぞれ計算した損失関数である。

この訓練データによる損失関数 (青) と、評価データによる損失関数 (橙) が大きく差が開いている場合、過学習が生じている証左である。しかし図で示したように差は十分小さく、最終的な損失関数の値としては、訓練データによる損失関数 (MAE) が  $5.36 \times 10^{-4}$ 、評価データによる損失関数 (MAE) は、 $5.38 \times 10^{-4}$  であった。またこの学習においては、250 epoch 目で early stopping している。

また、選定したパラメータのもとで学習させた機械学習モデルで計算した運動量  $p_S^{\text{ML}}$  と、教師データ  $p_S^0$  の差分  $\Delta p_S (= p_S^{\text{ML}} - p_S^0)$  分布を下に示した (図 4.17)。

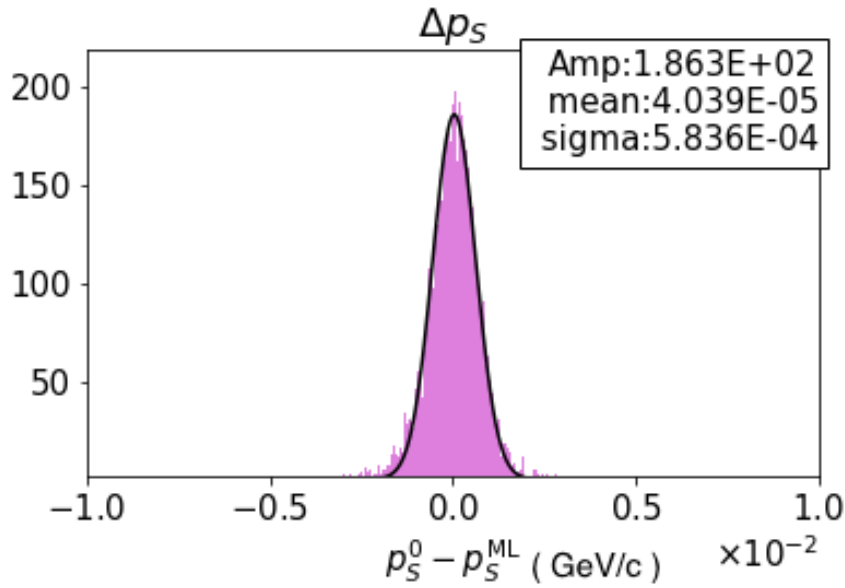


図 4.17: 機械学習で計算した運動量  $p_S^{\text{ML}}$  と教師データ  $p_S^0$  の差の分布。

$\Delta p$  分布は、フィッティングの結果  $\text{FWHM} = (1.00 \pm 0.00) \times 10^{-3}$  に対して、分布のシフトは  $\langle p_S^{\text{ML}} - p_S^0 \rangle$  は  $(4.04 \pm 0.01) \times 10^{-5}$  となった。

次に、 $p_S^{\text{ML}}, p_S^0$  と下流の飛跡情報  $x$  の相関を下図 4.18 に示した。S-2S においては  $x$  方向に磁場によって粒子が曲げられるため、運動量  $p_S$  と  $x$  は強い相関を持つ。教師データ  $p_S^0$  と機械学習模型で計算した運動量  $p_S^{\text{ML}}$  が同じ分布を示すことを確かめた。

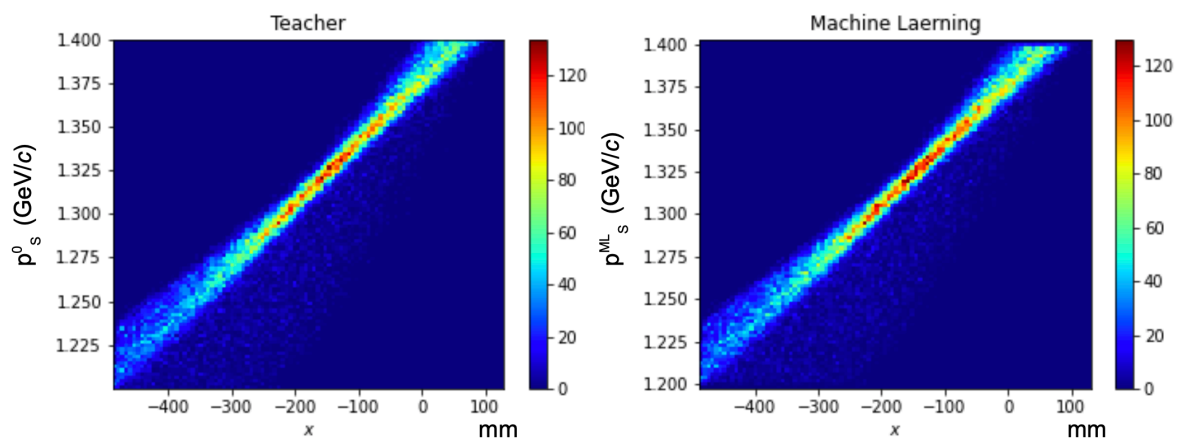


図 4.18: 下流飛跡情報  $x$  と、機械学習で計算した運動量  $p_S^{\text{ML}}$  (右)、教師データ  $p_S^0$  (左) との間の相関。

次に、計算運動量と教師データの差  $p_S^{\text{ML}} - p_S^0$  が、訓練データとして用いた飛跡情報  $x, y, u, v$  のそれぞれに対してどのように分布しているかを確認した。

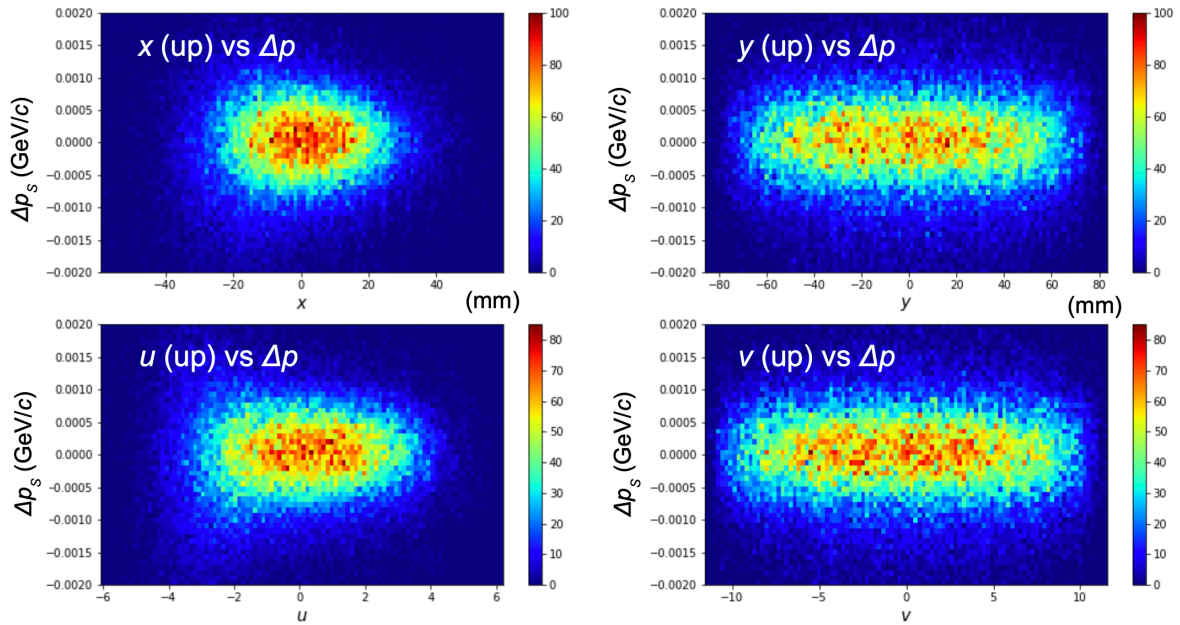


図 4.19: 上流飛跡情報 ( $x$ (左上)、 $y$ (右上)、 $u$ (左下)、 $v$ (右下)) と、機械学習で計算した運動量誤差  $p_S^{\text{ML}} - p_S^0$ 。

上流について、図 4.19 のように、いずれのパラメータにも特に相関は存在せず一様に  $\Delta p_S$  は  $\pm 5 \times 10^{-4}$  程度に抑えられている。下流の飛跡情報についても同様の結果となった (図 4.20)。

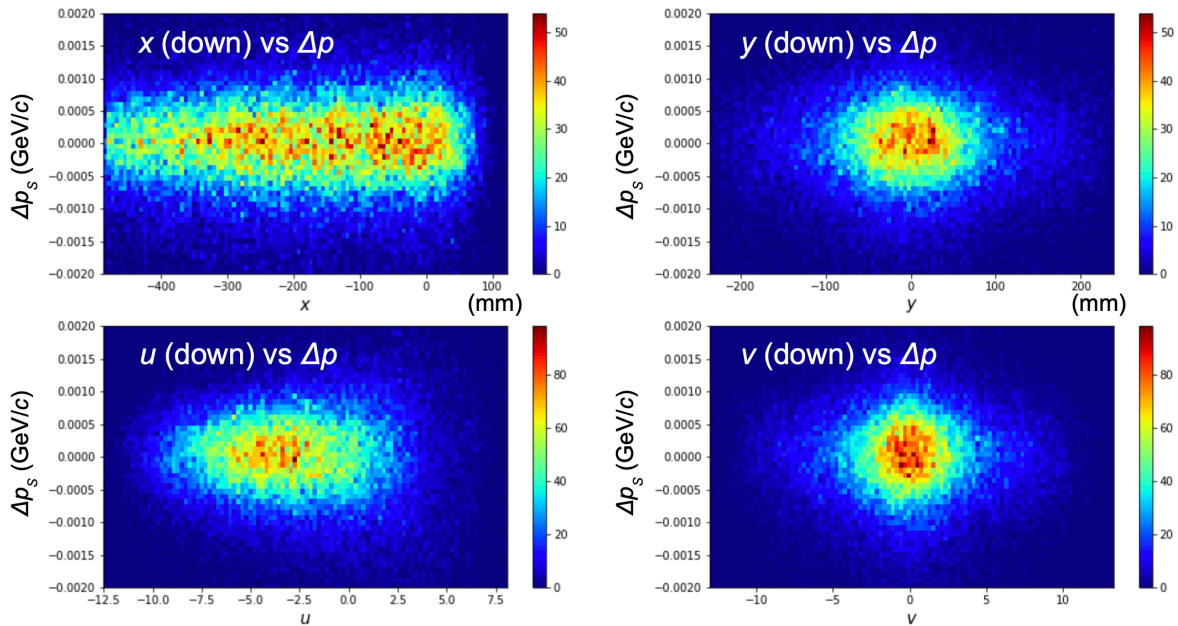


図 4.20: 下流飛跡情報 ( $x$ (左上)、 $y$ (右上)、 $u$ (左下)、 $v$ (右下)) と、  
機械学習で計算した運動量誤差  $p_S^{\text{ML}} - p_S^0$ 。

結果として、機械学習モデルで計算した運動量  $p_S^{\text{ML}}$  と入力データである飛跡情報  $x, y, u, v$  の間には本来ないはずの相関などが加えられることなどはなく、教師データとしての運動量  $p_S^0$  を適切に再現できていることが確認された。

## 4.6 $\cos \theta$ を出力層に追加した機械学習モデルの構築とその評価

ここまでの 4 節などでの議論では教師データを  $p_S$  のみとして散乱角度  $\theta$  については言及してこなかった。しかし、ミッシングマスを再構成するためには、 $p_S$  だけでなく、 $\cos \theta$  の情報も必要である。もちろん上流の飛跡情報から直接  $\cos \theta$  を導出することはできるが、7 章で述べるように機械学習モデルを校正するにあたり散乱角度についても校正する余地がある。そのため、 $\cos \theta$  を出力層として追加した機械学習モデルを新たに構築した。

モデルとしては、今まで  $p_S$  のみを出力層としていたモデルに  $\cos \theta$  を出力する層を追加する (図 4.21)。損失関数についても  $\cos \theta$  の項を追加し、(式 4.28) のように重みづけした損失関数でニューラルネットワーク中の重みパラメータを更新した。損失関数中の重み (0.8, 0.2) に関しては、 $\partial M_X / \partial p_S$ ,  $\partial M_X / \partial (\cos \theta)$  値に合わせたものである。

$$\begin{aligned} Loss &= 0.8MAE(p_S) + 0.2MAE(\cos \theta) \\ &= 0.8 |p_S^0 - \hat{p}_S^{ML}| + 0.2 |\cos \theta^0 - \cos \hat{\theta}^{ML}| \end{aligned} \quad (4.28)$$

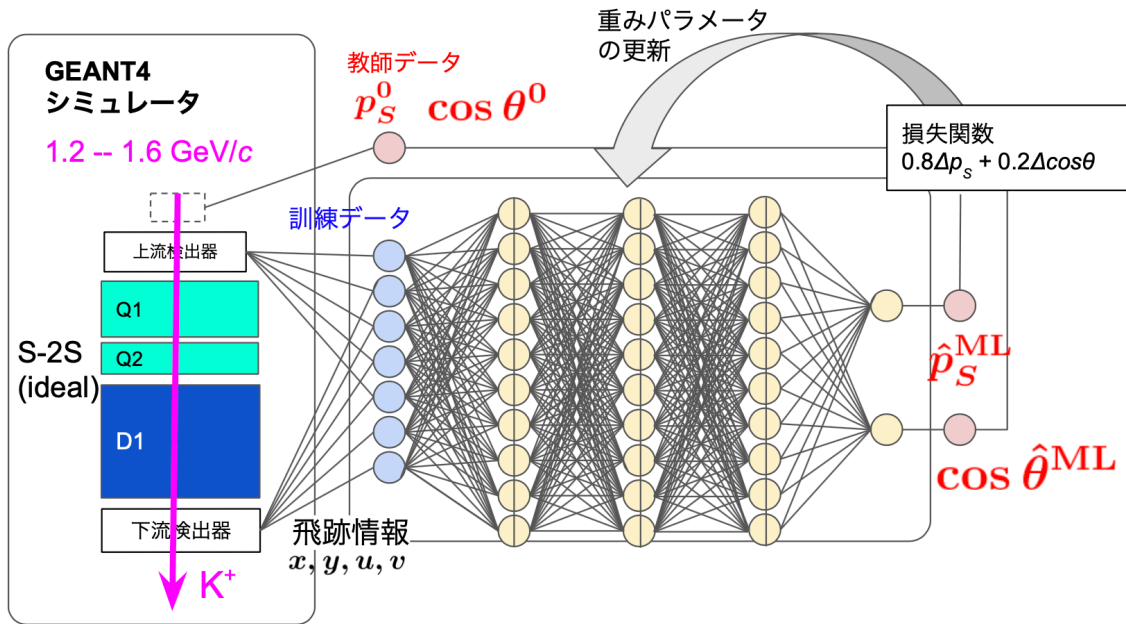


図 4.21: 機械学習モデルの出力層に  $\cos \theta$  を追加した模型。

模型構築においては訓練パラメータを以下の表 4.8 のように設定した。ミッシングマス分解能  $\Delta M$  への寄与は、 $\Delta \cos \theta$  よりも  $\Delta p_S$  の方が大きいため、今回は厳密に訓練パラメータの最適化は行わず、 $p_S$  のみを教師データとして用いていた模型の最適訓練パラメータとほぼ同様のパラメータ設定を選んだ。

表 4.8:  $\cos \theta$  層を追加した機械学習モデルで採用した訓練パラメータ群。

入力データ	
要求データセット統計量	$\gtrsim 10^6$
モデルパラメータ	
活性化関数	Swish
ノード数/レイヤー	32
隠れ層数	3
入力次数	3
入力パラメータ数	165
総 weight パラメータ数	7821
損失関数	平均絶対誤差 (MAE)
最適化関数	Adam

表 4.8 のもとで訓練した結果、 $\Delta p_S$ ,  $\Delta \cos \theta$  は図 4.9 のような値に収束した。なお比較として、 $p_S$  のみを教師データとした場合のモデルで計算した運動量分解能  $\Delta p_S/p_S$  も併せて載せる。 $\Delta p_S/p_S$  に関してはどちらもほぼ性能は同じで  $1.1 \times 10^{-3}$  程度の分解能（半値全幅、FWHM）、 $\Delta \cos \theta$  についても  $9 \times 10^{-5}$  の精度で計算することができた。

表 4.9:  $p_S$  のみを教師データとしたモデルと、 $p_S$ ,  $\cos \theta$  を教師データとしたモデルでの分解能  $\Delta p_S/p_S$  (半値全幅、FWHM)、 $\Delta \cos \theta$  (半値全幅、FWHM) の比較。

教師データ	$p_S$ のみ	$p_S, \cos \theta$
$\Delta p_S/p_S$ ( $\times 10^{-3}$ , train)	$1.09 \pm 0.00$	$1.08 \pm 0.00$
$\Delta p_S/p_S$ ( $\times 10^{-3}$ , test)	$1.10 \pm 0.00$	$1.08 \pm 0.00$
$\Delta \cos \theta$ ( $\times 10^{-5}$ , train)	-	$9.22 \pm 0.03$
$\Delta \cos \theta$ ( $\times 10^{-5}$ , test)	-	$9.05 \pm 0.03$

下の図 4.22 には、損失関数 (MAE) の訓練時の推移を載せる。結果的に  $p_S$  も  $\cos \theta$  に対しても過学習は特に認められなかった。

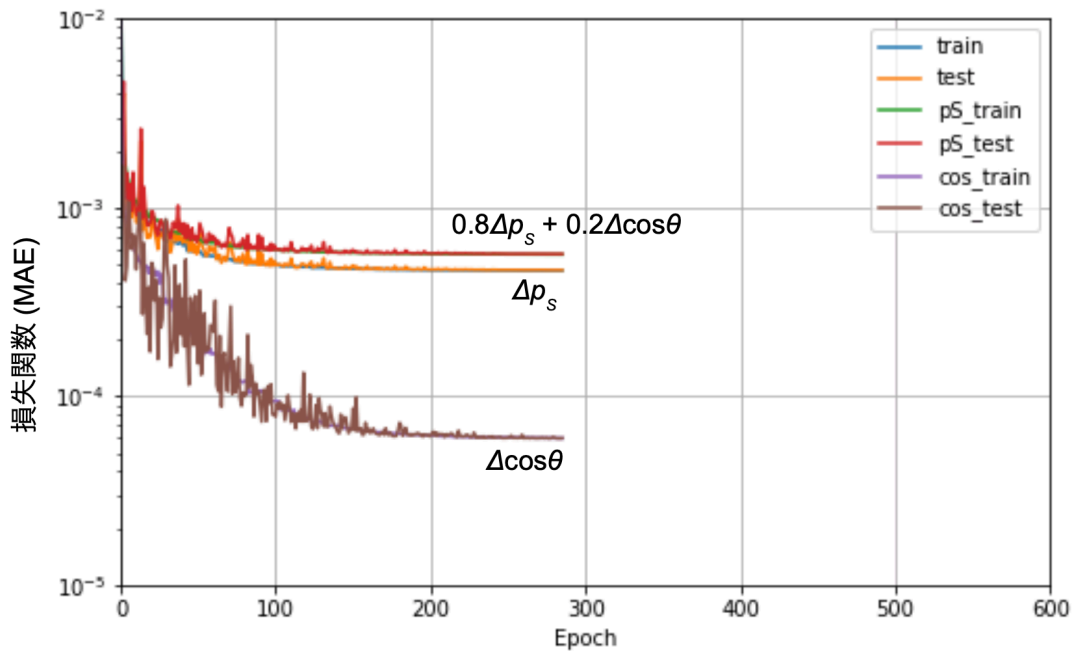


図 4.22:  $0.8MAE(p_S) + 0.2MAE(\cos \theta)$  と  $p_S$ 、 $\cos \theta$  の損失関数 (MAE) の訓練時の推移を示す。

また、下の図 4.23, 4.24 には、誤差  $\Delta \cos \theta$  と  $\Delta p_S$  の分布を載せる。

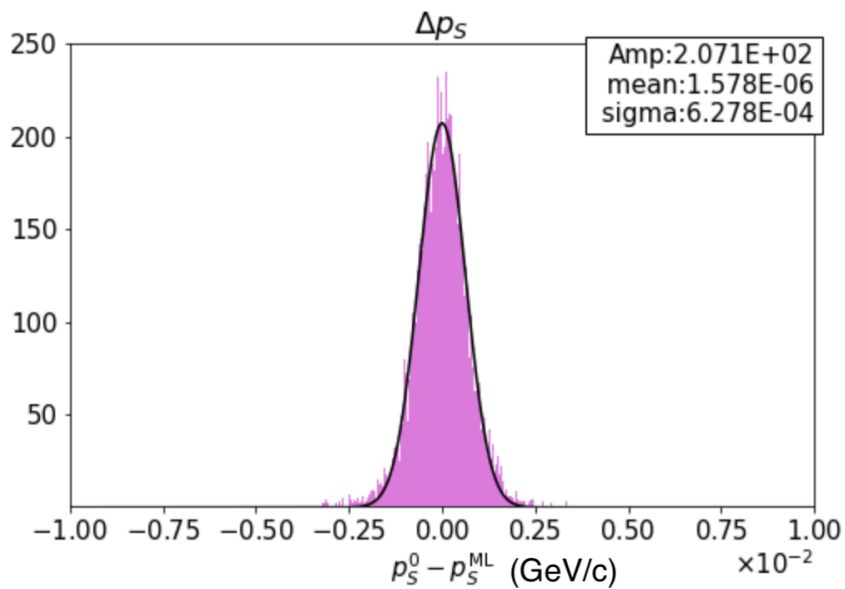


図 4.23: 誤差  $\Delta p_S/p_S$  の分布。

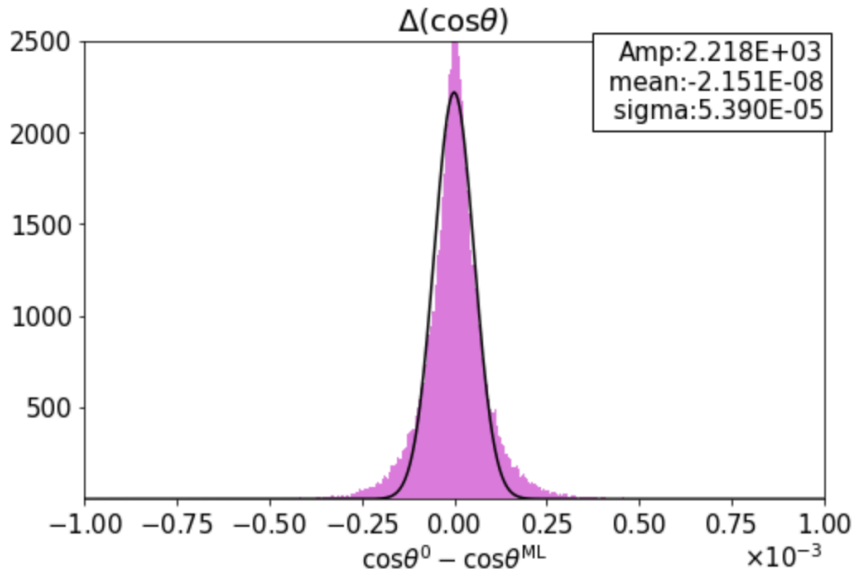


図 4.24: 誤差  $\Delta \cos \theta = \cos \theta^{\text{ML}} - \cos \theta^0$  の分布。

次に、計算運動量、計算角度と教師データが一对一对応しているかについて評価した (図 4.25, 4.26)。縦軸が教師データで、横軸が  $\cos \theta$  を追加した機械学習モデルで計算した  $p_S, \cos \theta$  である。いずれも一对一对応していることが確認された。

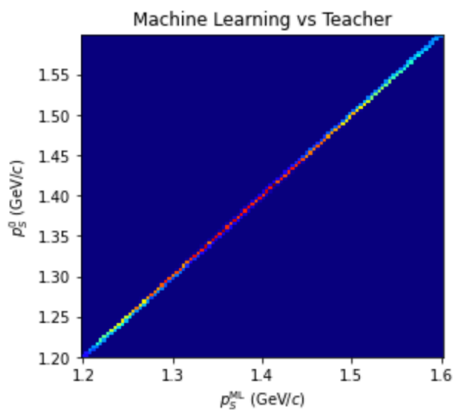


図 4.25:  $p_S^{\text{ML}}$  vs  $p_S^0$ 。

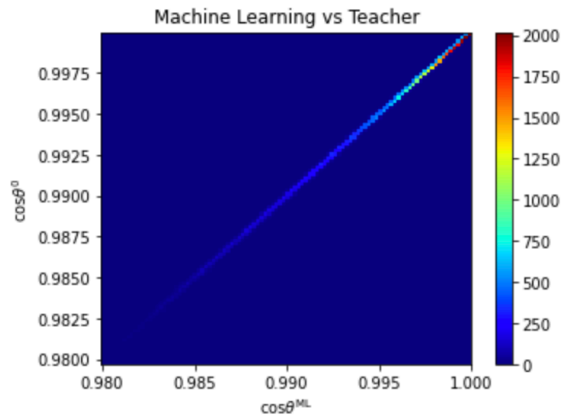


図 4.26:  $\cos \theta^{\text{ML}}$  vs  $\cos \theta^0$ 。

次に、計算運動量と教師データの差  $p_S^{\text{ML}} - p_S^0$  および  $\cos \theta^{\text{ML}} - \cos \theta^0$  が、訓練データとして用いた飛跡情報  $x, y, u, v$  のそれぞれに対してどのように分布しているかを確認した。



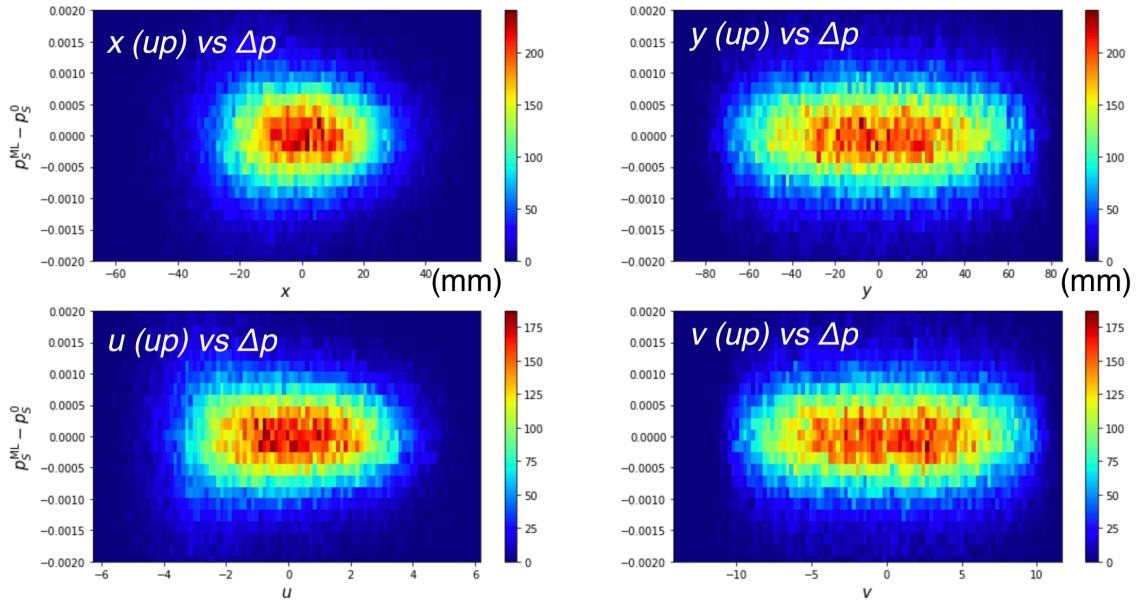


図 4.27: 上流飛跡情報 ( $x$ (左上)、 $y$ (右上)、 $u$ (左下)、 $v$ (右下)) と、  
機械学習で計算した運動量誤差  $p_S^{\text{ML}} - p_S^0$ 。

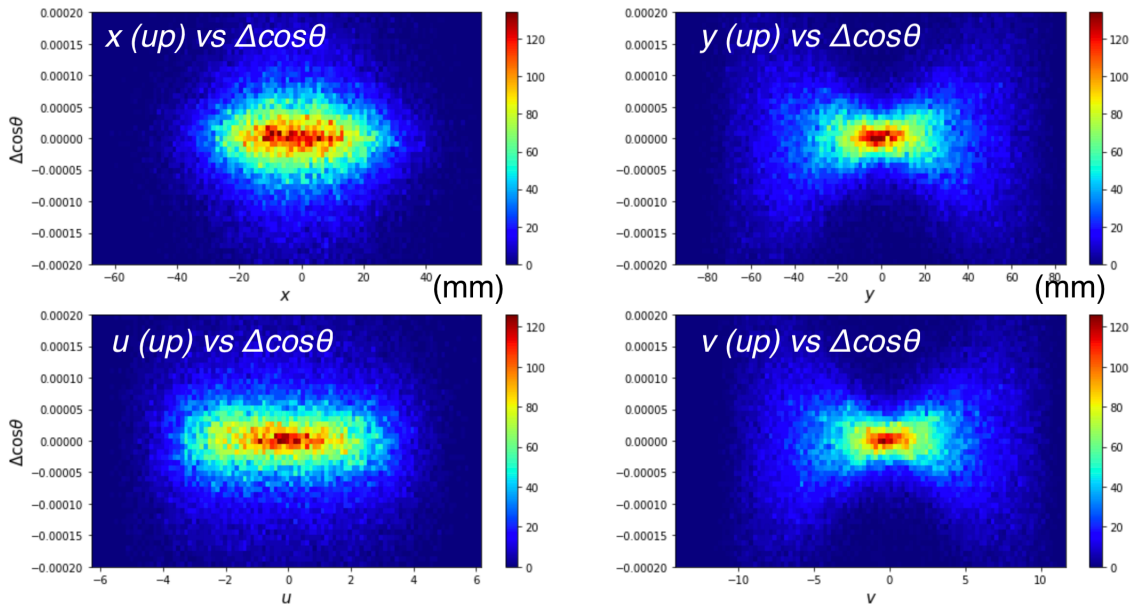


図 4.28: 上流飛跡情報 ( $x$ (左上)、 $y$ (右上)、 $u$ (左下)、 $v$ (右下)) と、  
機械学習で計算した運動量誤差  $\cos \theta^{\text{ML}} - \cos \theta^0$ 。

上流について、図 4.27 のように  $p_S$  に関してはいずれのパラメータにも特に相関は存在

せず一様に  $\Delta p_S$  は  $\pm 5 \times 10^{-4}$  程度に抑えられている。一方、 $\cos \theta$  に関しては、図 4.28 のように、 $y, v$  の絶対値が大きい領域に関しては分解能が悪くなっているが、余計な相関は与えられていない。 $x, u$  に関しても特に相関は見られなかった。

下流の飛跡情報に関しても同様に  $\Delta p_S$ ,  $\Delta \cos \theta$  の  $x, y, u, v$  依存性を確認した (図 4.29, 4.30)。下流に関しては  $\Delta p_S$ ,  $\Delta \cos \theta$  と飛跡情報  $x, y, u, v$  の間に特に相関は見られなかった。

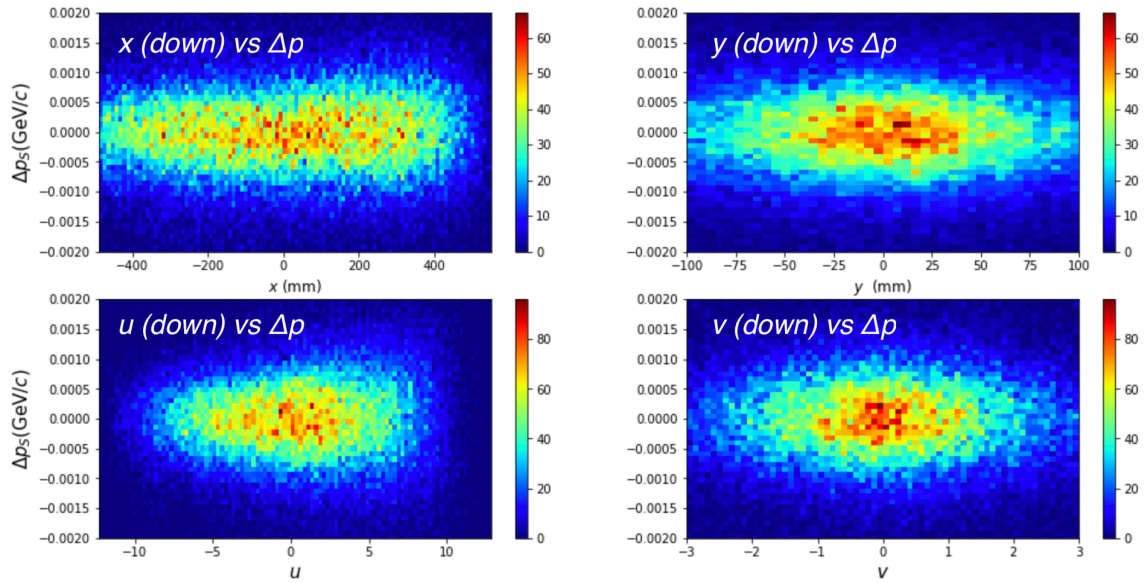


図 4.29: 下流飛跡情報 ( $x$ (左上)、 $y$ (右上)、 $u$ (左下)、 $v$ (右下)) と、機械学習で計算した運動量誤差  $p_S^{\text{ML}} - p_S^0$ 。

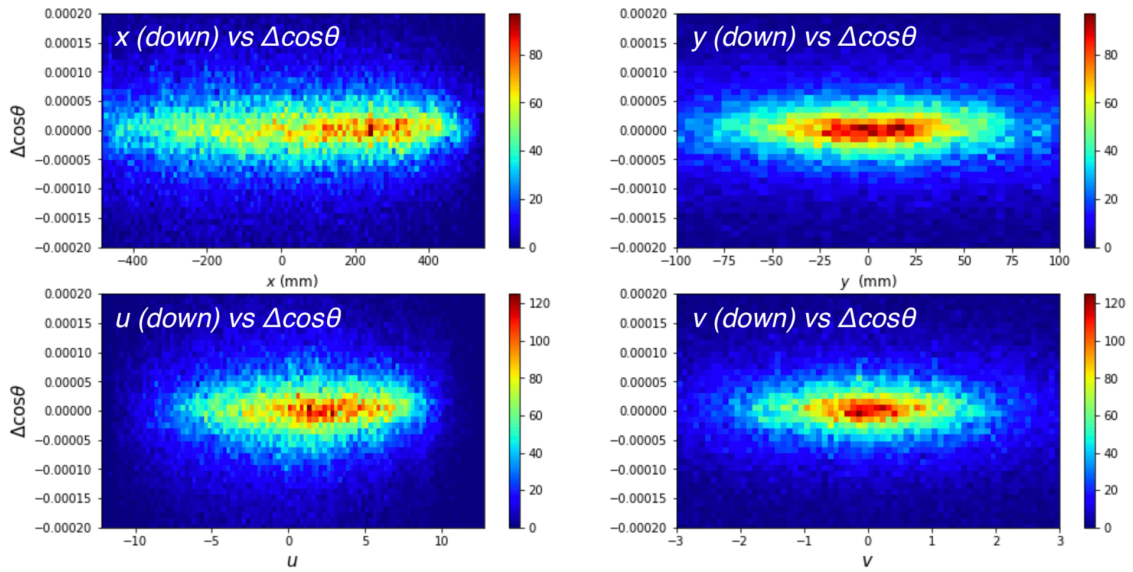


図 4.30: 下流飛跡情報 ( $x$ (左上)、 $y$ (右上)、 $u$ (左下)、 $v$ (右下)) と、  
機械学習で計算した運動量誤差  $\cos \theta^{\text{ML}} - \cos \theta^0$ 。

結果として、散乱角度  $\cos \theta$  を機械学習モデルの出力に加えても、計算した運動量や散乱角度と入力データである飛跡情報  $x, y, u, v$  の間に本来ないはずの相関などが加えられることなどはなく、教師データとしての運動量  $p_S^0$  および  $\cos \theta^0$  を適切に再現できていることが確認された。

## 第5章

# 機械学習による S-2S の性能評価

本章においては機械学習で構築した運動量解析フレームワークを用いて S-2S の性能評価、および  $\Xi^-$ ,  $\Sigma^-$ ,  ${}^{12}_{\Xi}\text{Be}$  のミッシングマス分解能の評価を行った。

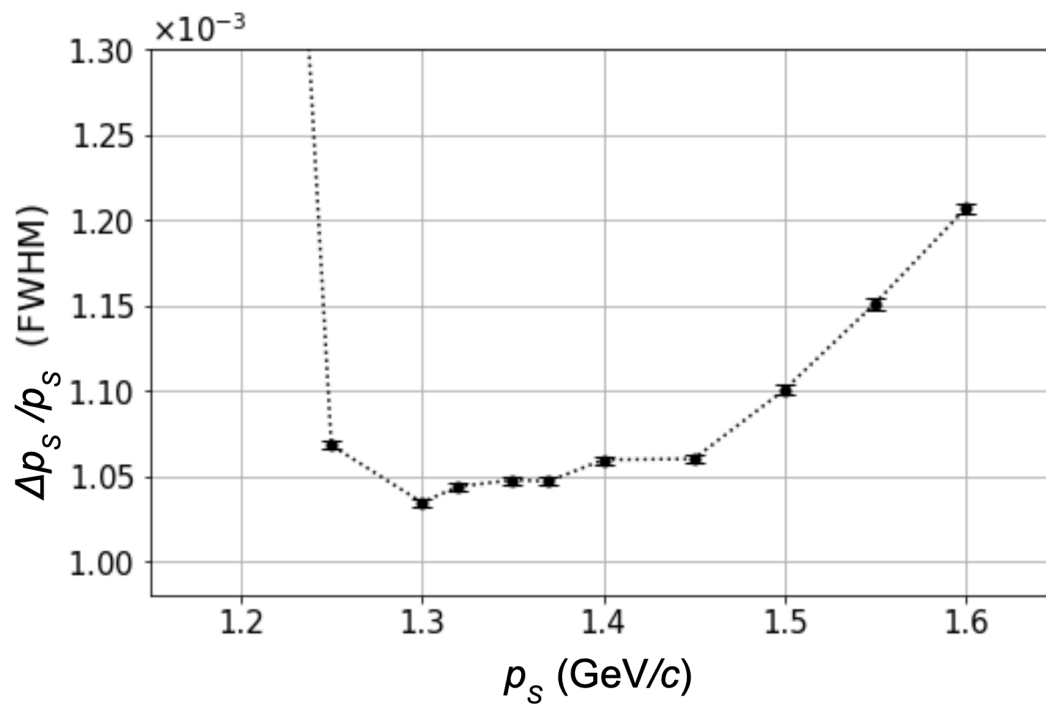
### 5.1 運動量分解能の運動量依存性

前章において、ある運動量空間内の S-2S のアクセプタンス領域での運動量・飛跡情報を教師データ・訓練データとして用いて機械学習モデルを構築していた。この時得られた全訓練運動量領域を統合した運動量分解能は、 $\sim 1 \times 10^{-3}$  程度であった。本節では、訓練済み機械学習モデルに、単色運動量の  $K^+$  で得られた飛跡情報を入力することで、各運動量での運動量分解能を評価した。

なお本検証での訓練運動量領域は  $1.2\text{--}1.6 \text{ GeV}/c$  を用いた。S-2S の全アクセプタンス領域（飽和状態）は  $1.2\text{--}1.6 \text{ GeV}/c$  を占めており、全領域での運動量分解能の推移を見るためである。

表 5.1:  $\Delta p_S/p_S$ (半値全幅、FWHM) vs  $p_S$ 。

運動量 $p_S$ (GeV/c)	$\Delta p_S/p_S$ (FWHM)
1.2	$2.01 \times 10^{-3}$
1.25	$1.07 \times 10^{-3}$
1.3	$1.03 \times 10^{-3}$
1.32	$1.04 \times 10^{-3}$
1.35	$1.05 \times 10^{-3}$
1.37	$1.05 \times 10^{-3}$
1.4	$1.06 \times 10^{-3}$
1.45	$1.06 \times 10^{-3}$
1.5	$1.10 \times 10^{-3}$
1.55	$1.15 \times 10^{-3}$
1.6	$1.21 \times 10^{-3}$

図 5.1: 運動量分解能 (FWHM) の運動量依存性 ( $\Delta p_S/p_S$  vs  $p_S$ )。

機械学習モデルで計算した運動量分解能によると、運動量が 1.3 GeV/c 付近で極小値をとり、その後運動量が大きくなるにつれて運動量分解能も悪くなることがわかった。この傾向に関しては [60] での計算と合致している。

## 5.2 多重散乱の運動量分解能への影響

S-2S は  $K^+$  を散乱させることで運動量を測定するが、上下流でのドリフトチェンバーでその飛跡情報を得るまでに、空気、チェンバーガス ( $C_2H_5$ , Ar, iso- $C_4H_{10}$ ) などといった物質を通過する。これらの物質中での多重散乱の影響で運動量分解能が悪化すると考えられる。多重散乱の影響を抑制するために、S-2S の Q1, Q2, D1 磁石内には比較的物質量が低い He ガスで満たす。これら物質量が分解能にどのくらい影響を及ぼすかを機械学習モデルによる運動量解析の視点から調べる。

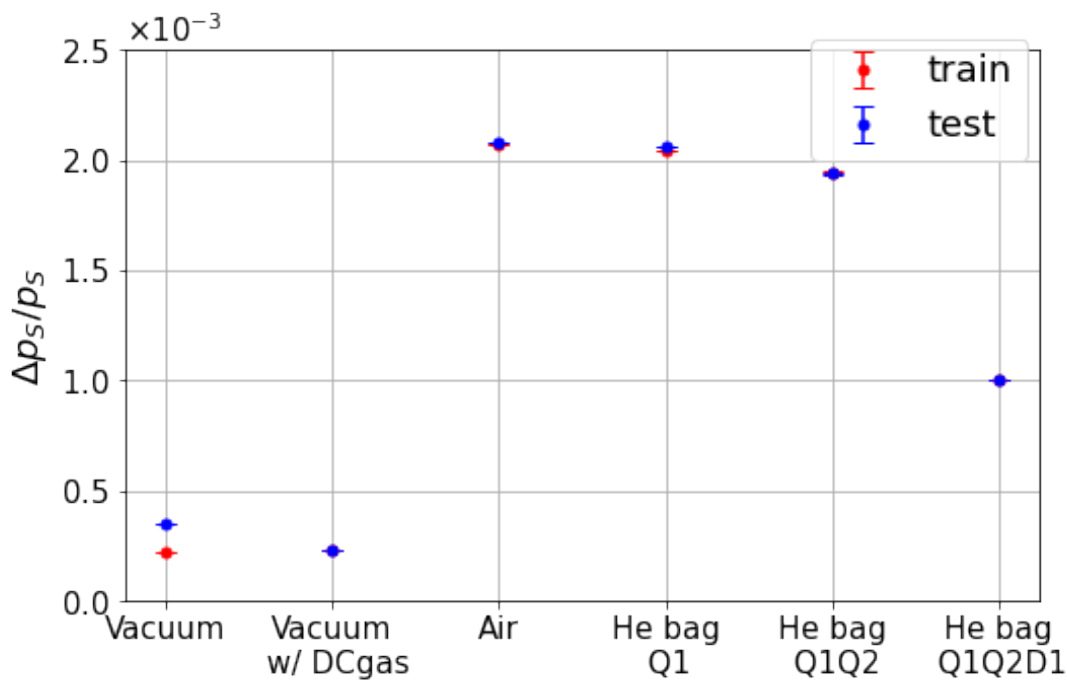


図 5.2: 物質から運動量分解能 (半値全幅、FWHM) への影響

図 5.2 は、横軸の粒子の通過領域の状態に対する運動量分解能の大きさを示す。図内の、“Vacuum” は、通過領域を全て真空に設定し物質量を考慮していない状態である。“Vacuum w/ DC gas” は真空状態にドリフトチェンバーのガスの物質量を与えた状態

“Air”は通過領域に空気を満たした状態である。さらに、“He bag Q1”はQ1電磁石内の通過領域にヘリウムバッグを設置した状態、“He bag Q1Q2”はQ1とQ2の両電磁石内にヘリウムバッグを設置した状態、“He bag Q1Q2D1”はQQD電磁石内の通過領域にヘリウムバッグを設置した状態、である。“He bag Q1Q2D1”が最も現実のセットアップに近い。

S-2S中の物質量を考慮しない、真空状態 (Vacuum) のときの運動量分解能は、 $3 \times 10^{-4}$ 程度であった。ここに空気を加えた場合には、分解能が $2 \times 10^{-3}$ 程度と8倍ほど悪化する。この物質量による分解能悪化を防ぐために、Q1, Q2, D1磁石の粒子の通過領域にヘリウムバッグを設置したが、やはり磁極長が最も長いD1のヘリウムバッグの影響が大きく効いていることがわかる。ヘリウムバッグの効果により、分解能は1/2程度まで抑えられることがわかった。

### 5.3 検出器分解能の運動量分解能への影響

E70では飛跡検出器としてドリフトチェンバーを使用する。この位置分解能は約 $200 \mu\text{m}$ と見積もられる。今回の検証では、検出器の分解能が運動量分解能にどの程度影響が出るかを検証する。また、同時に粒子の通過領域の物質量ごとの、検出器分解能による運動量分解能の推移もみた。

表 5.2:  $K^+$  の通貨領域が真空の場合と、空気で満たされた場合、電磁石にヘリウムバッグを設置した場合（現実のセットアップに最も近い）での、運動量分解能を検出器分解能を変えながら評価した。

真空		
検出器分解能 ( $\mu\text{m}$ )	$\Delta p_S/p_S$ (train, FWHM)	$\Delta p_S/p_S$ (test)
0	$1.58 \times 10^{-4}$	$1.59 \times 10^{-4}$
50	$1.81 \times 10^{-4}$	$1.83 \times 10^{-4}$
100	$2.04 \times 10^{-4}$	$2.04 \times 10^{-4}$
200	$2.65 \times 10^{-4}$	$2.65 \times 10^{-4}$
500	$5.13 \times 10^{-4}$	$5.17 \times 10^{-4}$
空気 (He バッグなし)		
検出器分解能 ( $\mu\text{m}$ )	$\Delta p_S/p_S$ (train, FWHM)	$\Delta p_S/p_S$ (test)
0	$1.17 \times 10^{-3}$	$1.17 \times 10^{-3}$
50	$2.01 \times 10^{-3}$	$2.02 \times 10^{-3}$
100	$2.05 \times 10^{-3}$	$2.05 \times 10^{-3}$
200	$2.08 \times 10^{-3}$	$2.08 \times 10^{-3}$
500	$2.18 \times 10^{-3}$	$2.17 \times 10^{-3}$
空気 (QQD 磁石中に He バッグあり)		
検出器分解能 ( $\mu\text{m}$ )	$\Delta p_S/p_S$ (train, FWHM)	$\Delta p_S/p_S$ (test)
0	$9.56 \times 10^{-4}$	$9.59 \times 10^{-4}$
50	$9.69 \times 10^{-4}$	$9.68 \times 10^{-4}$
100	$9.89 \times 10^{-4}$	$9.86 \times 10^{-4}$
200	$1.02 \times 10^{-3}$	$1.03 \times 10^{-3}$
500	$1.12 \times 10^{-3}$	$1.11 \times 10^{-3}$



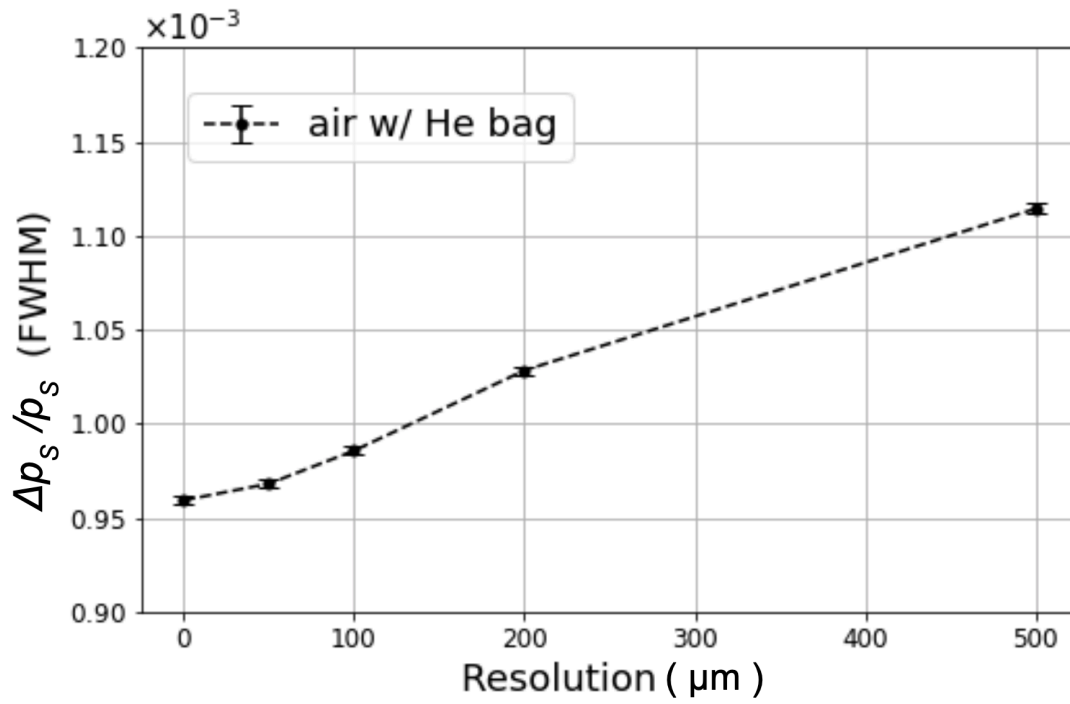


図 5.3: 散乱粒子の通過領域を空気で満たしたのち、QQD 内にヘリウムバッグ (He bag) を追加したセットアップ (最も現実に近いセットアップ) での検出位置分解能の運動量分解能 (半値全幅、FWHM) に対する影響。

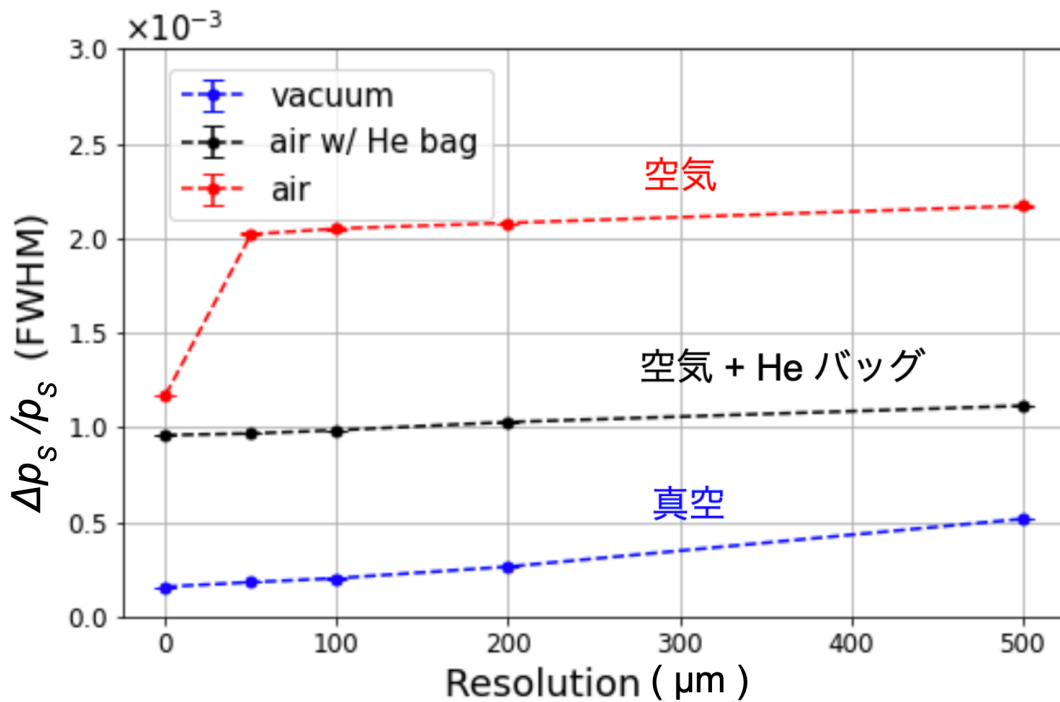


図 5.4: 粒子の通過領域を全て真空にしたセットアップ (青)、全て空気で満たした場合のセットアップ (赤)、ヘリウムバッグを QQD 電磁石に設置した場合 (黒) での運動量分解能 (半値全幅、FWHM)。

## 5.4 ミッシングマス分解能の評価

前章で構築した機械学習モデルを用いて、現実にもっと近いシミュレーションセットアップでの運動量分解能が  $dp/p \simeq 1.0 \times 10^{-3}$  程度と見積もられた。この運動量分解能でミッシングマスを組んだ場合での、 $\Xi^-$ ,  $\Sigma^-$ ,  ${}_{12}^{\text{Be}}$  のミッシングマスの分解能の評価と  $\Xi$  ハイパー核のピーク構造の評価を行った。なお今回は標的中でのエネルギーストラグリングによる分解能の悪化に関しては考慮していないことに注意せよ。

ミッシングマスを計算するためには、散乱運動量  $p_S$  の他にビーム運動量  $p_B$  と散乱角度  $\theta$  の情報が必要である。 $p_S$ ,  $\theta$  に関しては、4章で構築した機械学習モデルを用いて計算する。

一方  $p_B$  に関しては、シミュレーション上でミッシングマス分布に従った運動学の  $K^+$  を照射するとき指定する。値としては  $p_B = 1.8 \text{ GeV}/c$  で固定し、分解能も

$\Delta p_B/p_B = 0, 3.3 \times 10^{-4}, 1.0 \times 10^{-3}$  の3つの値で固定する。 $\Delta p_B/p_B$  を単一の値で評価しないのは、ビームラインスペクトロメータの運動量分解能は Beam Through により散乱側のスペクトロメータの性能で測定されているため、現時点では上限が  $1.0 \times 10^{-3}$  であることがわかっている。しかし設計値としては  $3.3 \times 10^{-4}$  であるため、今回のミッシングマスの評価では、上の3つの値で評価する。

#### 5.4.1 $\Xi^-$ 、 $\Sigma^-$ のミッシングマス分解能評価

ミッシングマス分解能を評価するために、 $p(K^-, K^+)\Xi^-$ 、 $p(K^-, \pi^+)\Sigma^-$  および  $^{12}\text{C}(K^-, K^+)\Xi^-$  の運動学に則った  $K^+$ 、 $\pi^+$  を射出する。シミュレータ上に置かれた仮想検出器より飛跡情報  $x, y, u, v$  を得て、これを訓練済み機械学習モデルに入力する。このときの出力運動量からミッシングマスを計算する。

まず  $\Xi^-$  ピークのミッシングマス分解能を評価した。ビームラインスペクトロメータ、つまり  $K^-$  の運動量分解能が  $\Delta p_B/p_B = 0, 3.3 \times 10^{-4}, 1.0 \times 10^{-3}$  (FWHM) の場合での、 $\Xi^-$  ピークのミッシングマス分解能を評価した (表 5.3)。

表 5.3:  $\Xi^-$  ピークのミッシングマス分解能。

$\Xi^-$ ピーク			
$\Delta p_B/p_B (\times 10^{-4}, \text{FWHM})$	0	3.3	10
$\Delta M_{\text{ML}} (\text{MeV}/c^2, \text{FWHM})$	0.904	0.999	1.51
$\langle M_{\text{ML}} \rangle - M_0 (\text{MeV}/c^2)$	$8.12 \times 10^{-4}$	$2.99 \times 10^{-3}$	$1.83 \times 10^{-2}$

次に  $\Sigma^-$  ピークの分解能を評価した。 $K^-$  の運動量分解能が  $\Delta p_B/p_B = 0, 3.3 \times 10^{-4}, 1.0 \times 10^{-3}$  (FWHM) の場合での、 $\Sigma^-$  ピークのミッシングマス分解能を評価した (表 5.4)。

表 5.4:  $\Sigma^-$  ピークのミッシングマス分解能。

$\Sigma^-$ ピーク			
$\Delta p_B/p_B (\times 10^{-4}, \text{FWHM})$	0	3.3	10
$\Delta M_{\text{ML}} (\text{MeV}/c^2, \text{FWHM})$	1.52	1.59	2.10
$\langle M_{\text{ML}} \rangle - M_0 (\text{MeV}/c^2)$	$-5.07 \times 10^{-2}$	$-5.07 \times 10^{-2}$	$-1.58 \times 10^{-2}$

5.4.2  ${}^{12}_{\Xi}\text{Be}$  のミッシングマス分解能および分布の評価

${}^{12}_{\Xi}\text{Be}$  のエネルギーピーク構造の再現性を評価する。まず、 ${}^{12}_{\Xi}\text{Be}$  のコア核  ${}^{11}\text{B}$  と  $\Xi^-$  の質量和に仮想的なミッシングマスピークが立つような運動学で、シミュレーション空間上に  $K^+$  を照射する。この仮想ピークのミッシングマス分解能について、各ビームラインスペクトロメータの運動量分解能ごとに評価した (表 5.5)。

表 5.5:  ${}^{11}\text{B} + \Xi^-$  のミッシングマス分解能。

${}^{11}\text{B} + \Xi^-$			
$\Delta p_B/p_B (\times 10^{-4}, \text{FWHM})$	0	3.3	10
$\Delta M_{\text{ML}} (\text{MeV}/c^2, \text{FWHM})$	1.27	1.41	2.13
$\langle M_{\text{ML}} \rangle - M_0 (\text{MeV}/c^2)$	$8.12 \times 10^{-4}$	$2.99 \times 10^{-3}$	$1.83 \times 10^{-2}$

${}^{12}_{\Xi}\text{Be}$  の束縛領域でのミッシングマス分布を評価した。横軸は  $\Xi$  ハイパー核の束縛エネルギー ( $-BE = M_{\text{HYP}} - M_{\text{core}} - M_{\Xi}$ ) であり、 $-8$  から  $8 \text{ MeV}/c^2$  までの領域内で  $2 \text{ MeV}/c^2$  ごとにミッシングマス分解能  $\Delta M$  を評価した。(図 5.5)

結果としては束縛領域に対してほぼ一様に  $\Delta M \simeq 1.4 \text{ MeV}/c^2$  の分解能を持つことがわかった。

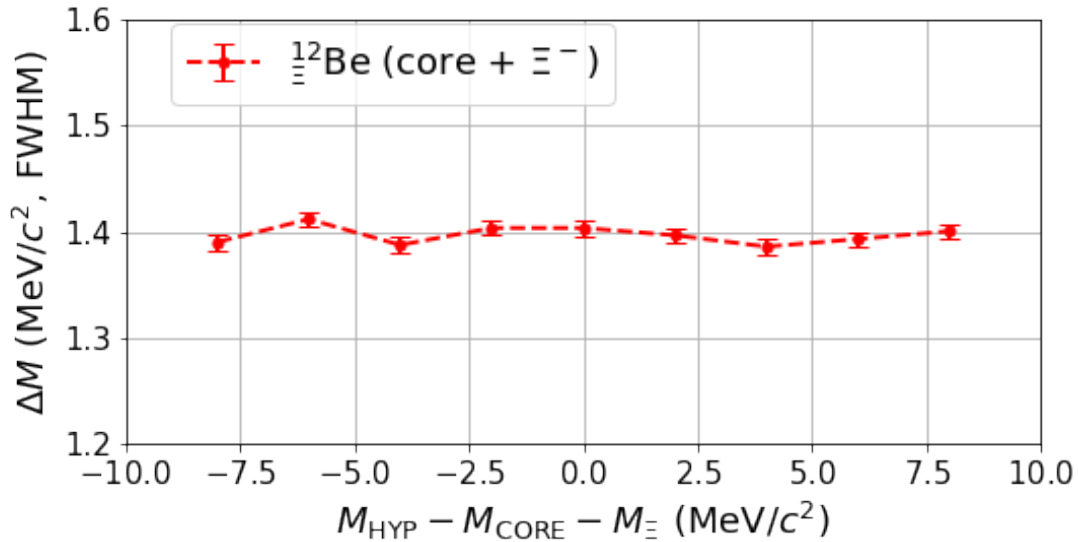


図 5.5:  $^{12}\text{Be}$  の束縛領域付近のミッシングマス  $M(= M_{\text{HYP}} - M_{\text{core}} - M_{\Xi})$  に対する分解能  $\Delta M_{\text{ML}}$ (FWHM)。

E70 実験では 100 イベントの  $^{12}\text{Be}$  事象を取得する予定である。機械学習を用いてミッシングマスを構成したときに、100 イベントの統計量で  $^{12}\text{Be}$  のエネルギーピークが弁別可能であるかについても検証した。今回もシミュレーション空間に  $^{12}\text{Be}$  のエネルギーピーク構造が立つような運動学で、 $K^+$  を射出し、検証を行った。ピーク構造は理論計算 ESC08a[23] に基づいている。ピーク位置は  $-BE = 2.5, -1.5, -4.5$  MeV とし、幅は 0.5 MeV, 各ピークの収量が約 45, 45, 10 カウントになるように、 $K^+$  を射出した (図 5.6)。なお準自由散乱に関しては今回考慮しなかった。

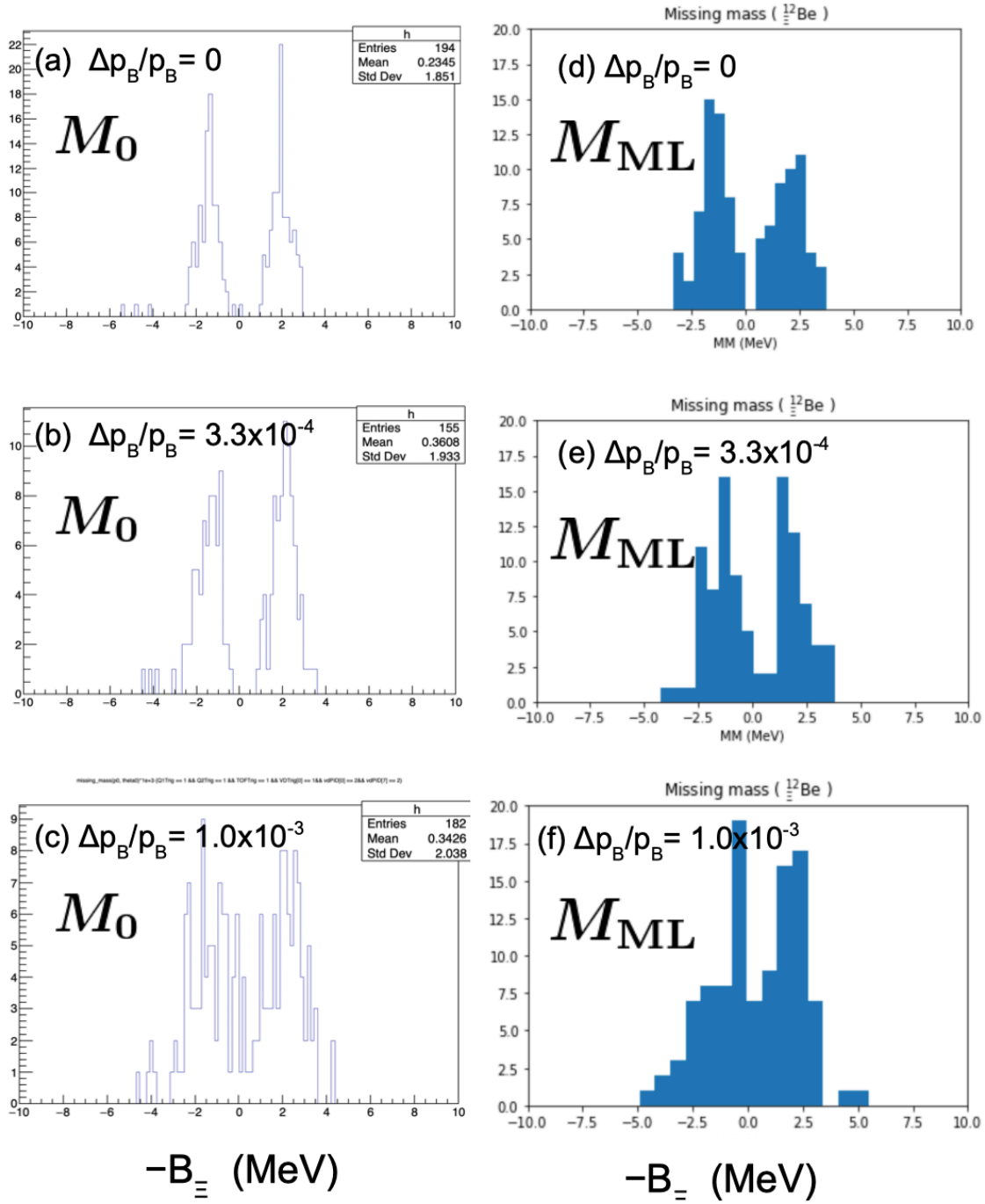


図 5.6: GEANT4 で出力された教師データ  $p_S^0$  から直接計算した  $^{12}\text{Be}$  のエネルギースペクトル ( $M_0$ ) (左図)。飛跡情報から機械学習模型で計算したスペクトル ( $M_{ML}$ ) (右図)。上から順に  $K^-$  の運動量分解能  $dp_B/p_B$  を  $0, 3.3, 10 \times 10^{-4}$  (FWHM) と変えている。エネルギーピーク位置は ESC08a[23] を仮定している。

$\Xi^-$ ,  $\Sigma^-$ ,  ${}^{12}\text{Be}$  の  $\Delta M$  の  $\Delta p_B/p_B$  依存性をまとめた図を表 5.8 に載せる。

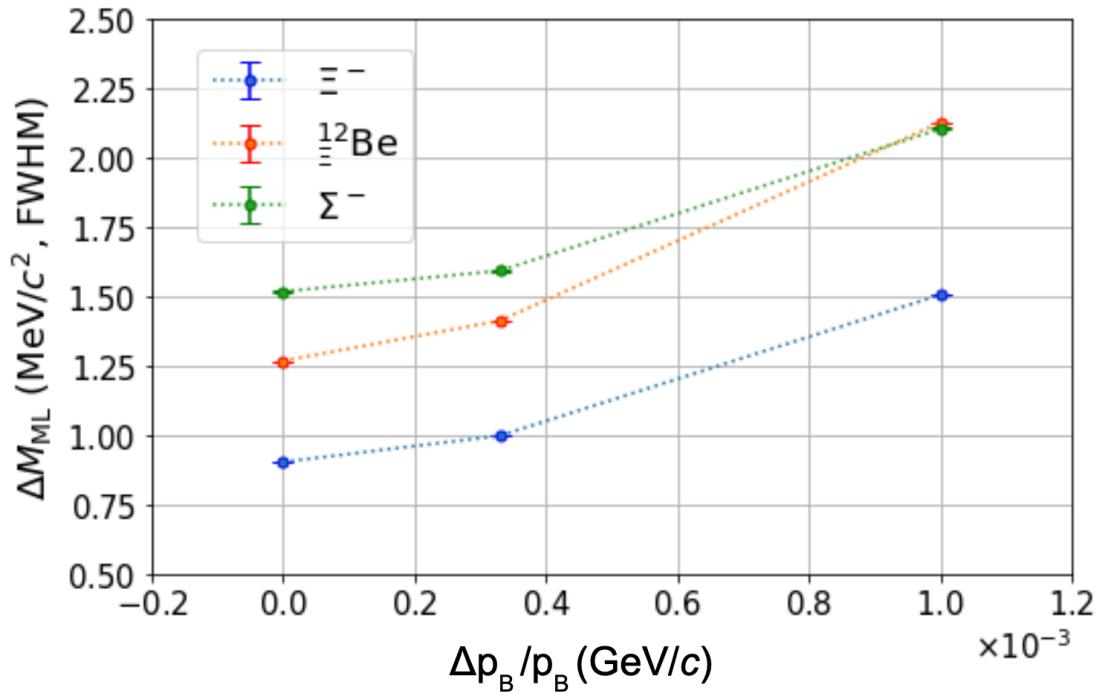


図 5.7: ミッシングマス分解能の  $K^-$  ビームの運動量分解能依存性。  $p(K^-, K^+)\Xi^-$  (青),  $p(K^-, \pi^+)\Sigma^-$  (緑),  ${}^{12}\text{C}(K^-, K^+){}^{12}\text{Be}$  (赤) ピークのミッシングマス分解能を、各  $dp_B/p_B$  の値で調べた。なおエネルギーストラグリングの効果は考慮していない。

基本的に  $\Xi^-$ ,  ${}^{12}\text{Be}$ ,  $\Sigma^-$  の順に  $\Delta M$  が悪化している。これは、 $\Xi^-$  の散乱運動量領域が  $< 1.31 \text{ GeV}/c$ 、 ${}^{12}\text{Be}$  が  $\sim 1.37 \text{ GeV}/c$ 、 $\Sigma^-$  が  $< 1.58 \text{ GeV}/c$  であり、5.1 節での結果より  $p_S$  が大きいほど  $\Delta p_S/p_S$  は悪くなるため、 $\Delta M$  についても同様の分布が示されたと考えられる。

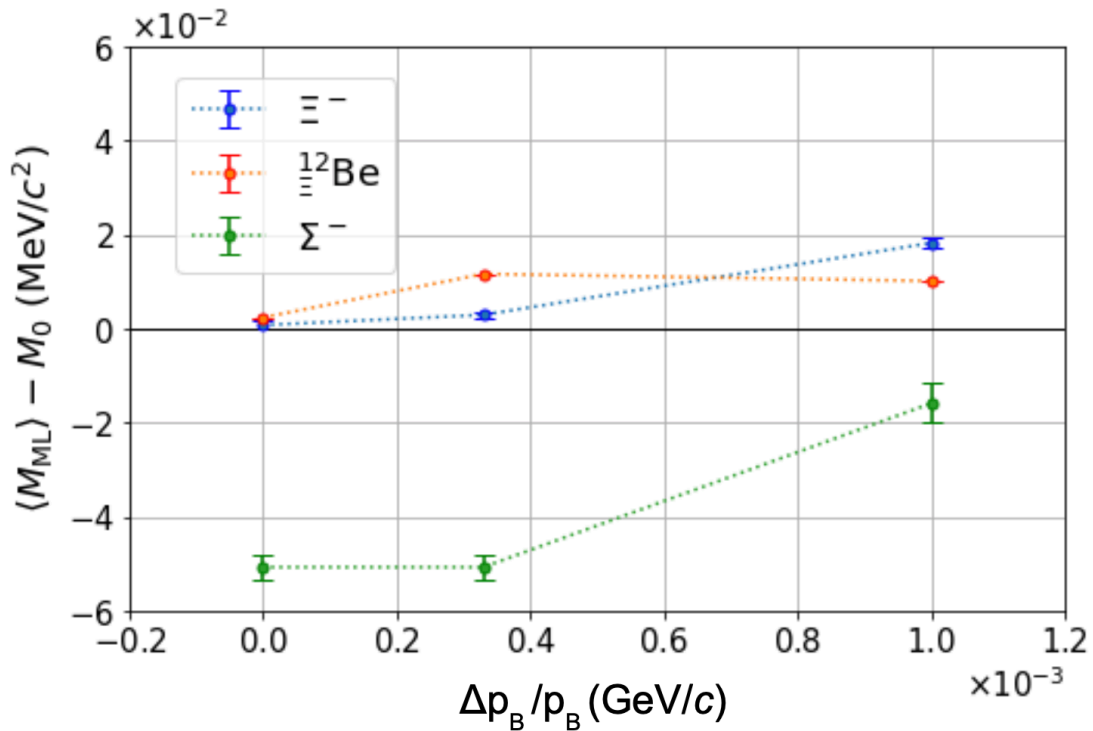


図 5.8: ミッシングマス分布のずれ ( $\langle M_{ML} \rangle - M_0$ ) の  $K^-$  ビームの運動量分解能依存性。 $p(K^-, K^+)\Xi^-$  (青),  $p(K^-, \pi^+)\Sigma^-$  (緑),  $^{12}\text{C}(K^-, K^+)\text{}^{12}\text{Be}$  (赤) ピークのミッシングマス分布のずれを、各  $dp_B/p_B$  の値で調べた。

ミッシングマス分布のシフト  $\langle M_{ML} \rangle - M_0$  に関しては、 $\Xi^-$ ,  $^{12}\text{Be}$  は、 $< 2 \times 10^{-2} \text{ MeV}/c^2$  に抑えられている。一方で、 $\Sigma^-$  についてはその 2 倍程度のシフトが見られた。

ところで、本検証においては、シミュレーション空間で  $K^+$  を生成する領域に炭素標的に相当する物質質量を入れていないので、これらの分解能は標的でのエネルギーストラグリングの効果は考慮していない。ストラグリングからの効果を含めたミッシングマス分解能を次の式 5.1 で計算し求めた。なお、式中の  $p_B$  はビーム粒子の運動量、 $p_S$  は散乱粒子の運動量、 $\theta$  は散乱角度、 $\Delta E_{\text{strag}}$  はストラグリングによる分解能の悪化分、 $\Delta_{\text{sim}}$  はシミュレーション誤差である。

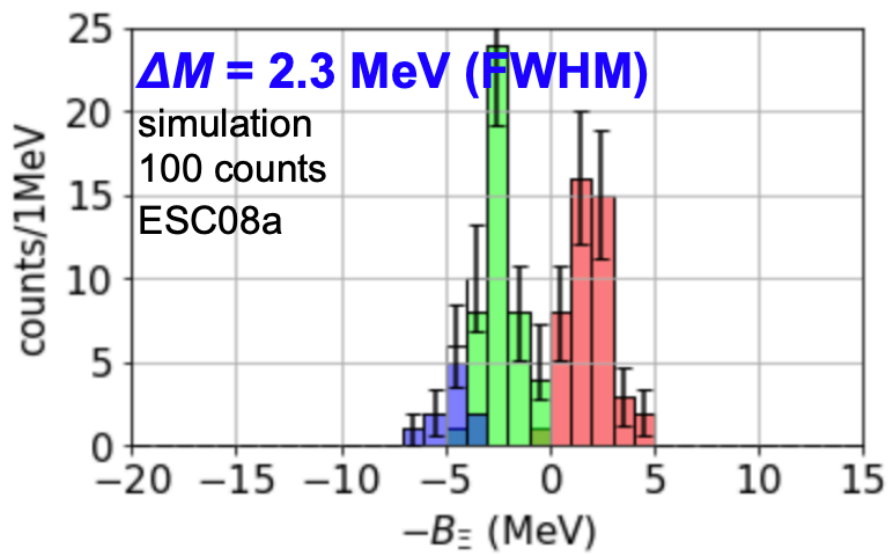
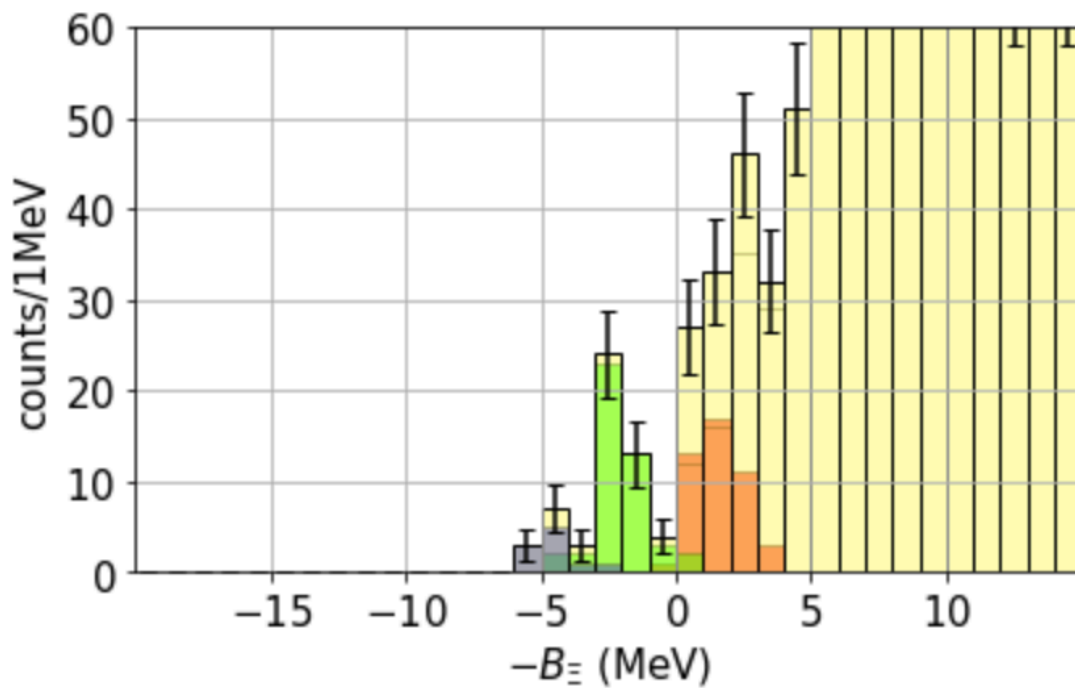


$$\Delta M^2 = \left(\frac{\partial M}{\partial p_B}\right)^2 \Delta p_B^2 + \left(\frac{\partial M}{\partial p_S}\right)^2 \Delta p_S^2 + \left(\frac{\partial M}{\partial \theta}\right)^2 \Delta \theta^2 + \Delta E_{\text{strag}}^2 + \Delta_{\text{sim}}^2 \quad (5.1)$$

このとき得られたミッシングマス分解能の値 (E70 by ML) を下の表に示す。先行実験である E05 実験と、E70 実験での設計値 (designed E70) を AFT によるストラグリングの補正の有無も併せて比較する。

表 5.6: ミッシングマス分解能と、各項の寄与を示す。先行実験である E05 実験と、E70 実験での設計値 (designed E70) と本修論で得られた値 (E70 by ML) を AFT によるストラグリングの補正の有無も併せて比較する。

(MeV)	$\frac{\partial M}{\partial p_B} \Delta p_B$	$\frac{\partial M}{\partial p_S} \Delta p_S$	$\frac{\partial M}{\partial \theta} \Delta \theta$	$\Delta E_{\text{strag}}$	$\Delta M$
E05	1.67	3.73	0.04	3.0	5.1
designed E70 (w/o AFT)	1.67	0.74	0.04	2.4	3.0
designed E70 (w/ AFT)	1.67	0.74	0.04	0.9	2.0
(MeV)	$\frac{\partial M}{\partial p_B} \Delta p_B$	$\frac{\partial M}{\partial p_S} (\Delta p_S + \Delta_{\text{sim}})$	$\frac{\partial M}{\partial \theta} (\Delta \theta + \Delta_{\text{sim}})$	$\Delta E_{\text{strag}}$	$\Delta M$
E70 by ML (w/o AFT)	1.67	1.30	0.04	2.4	3.2
E70 by ML (w/ AFT)	1.67	1.30	0.04	0.9	2.3

図 5.9:  $\Delta M = 2.3 \text{ MeV}$  でのミッシングマス分布。図 5.10: 準自由散乱を考慮した場合に予想される  $\Delta M = 2.3 \text{ MeV}$  でのミッシングマス分布。

今回構築した機械学習モデルから計算した、標的中のストラグリングの効果を含まない  $^{12}\text{Be}$  のミッシングマス分解能は、 $1.4 \text{ MeV}/c^2$  程度であった。しかしストラグリングの効果を入れた場合、 $3.2 \text{ MeV}/c^2$  まで悪化する。これを AFT で補正した場合には  $2.3 \text{ MeV}/c^2$  まで抑えられることがわかった。

$\Delta p_B/p_B = 1.0 \times 10^{-3}$  であった時のミッシングマス分解能は  $2.13 \text{ MeV}/c^2$  程度だったが、図 5.6 からミッシングマスピークを弁別できていることがわかる。この結果から AFT を考慮したミッシングマス分解能  $2.3 \text{ MeV}/c^2$  においても、ピーク弁別には十分なものであると思われる。

## 第 6 章

# local hit 情報を訓練データとして採用した運動量再構成手法

これまでの 4, 5 章での訓練データには、既に S-2S 上下流でトラッキング済みの飛跡情報  $(x, y, u, v)$  を用いていた。しかし、この飛跡情報も結局ドリフトチェンバーのワイヤ面での hit 位置から求めたものである。本節では、ドリフトチェンバーの各ワイヤ面で検出された hit 情報を訓練データとして学習させることで、自動的なトラッキングを含んだ運動量再構成が可能かを検証した。より低層の特徴量を訓練データとして選ぶことで更なる高精度な運動量解析を目指す。

E70 実験では S-2S 電磁石の上流に 2 台 (SDC1,2)、下流に 3 台 (SDC3-5) のドリフトチェンバーが設置されている。ワイヤ面は、SDC1 に 6 面、SDC2 に 4 面、SDC3 に 4 面、SDC4 に 4 面、SDC5 に 4 面の全部で 22 面存在する (表 6.1)。なおシミュレータ上のワイヤ面は、仮想検出器を現実のワイヤ面の位置に置いたものを用いる。位置分解能は  $\sigma_x = 200\mu\text{m}$  とした。

表 6.1: ドリフトチェンバーの性能 (再掲)

検出器	wire	ガス	有感領域 (mm)	陽極ワイヤ間隔 (mm)
SDC1	$uu'xx'vv'$	Ar + iso-C <sub>4</sub> H <sub>10</sub> + methylal (0.76 : 0.20 : 0.04)	$394^x \times 264^y$	6
SDC2	$vv'uu'$	Ar + iso-C <sub>4</sub> H <sub>10</sub> + methylal (0.76 : 0.20 : 0.04)	$160^x \times 300^y$	5
SDC3	$xx'yy'$	Ar + C <sub>2</sub> H <sub>6</sub> (1:1)	$1170^x \times 1170^y$	9
SDC4	$yy'xx'$	Ar + C <sub>2</sub> H <sub>6</sub> (1:1)	$1170^x \times 1170^y$	9
SDC5	$yy'xx'$	Ar + C <sub>2</sub> H <sub>6</sub> (1:1)	$1170^x \times 942^y$	9

今回の local hit を用いた機械学習モデルの開発においては、教師データに用いるデータセットを2種類用意した。それぞれで運動量  $p_S$  とミッシングマス分布を再構成可能かと、飛跡情報を訓練データとした場合と比較して分解能はどう変化したかについて検証した。

## 6.1 教師データに $p_S$ と $\cos \theta$ を与えた場合

まず、教師データとして運動量  $p_S$  と散乱角度の余弦  $\cos \theta$  を与えた場合を考える。今回も機械学習モデルの最適化のために1隠れ層あたりのノード数、隠れ層数、損失関数、入力データの最大次数を変化させ、最適な訓練パラメータを選定した。

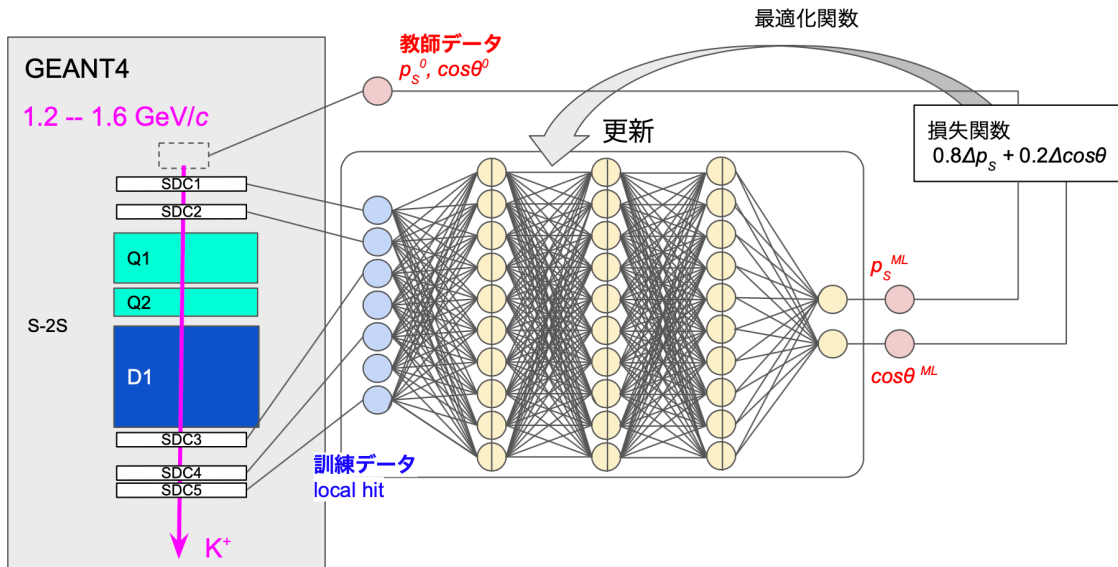


図 6.1: local-hit 情報を訓練データとして用いた機械学習模型。

- ノード数最適化

まず隠れ層 1 層あたりのノード数に関して他の訓練パラメータは以下の表のように固定して最適化を行った。

表 6.2: ノード数最適化のために用いた訓練パラメータ。

訓練パラメータ	
活性化関数	Swish
隠れ層数	3
損失関数	平均絶対誤差 (MAE)
最適化関数	Adam
訓練データの統計量	$1.07 \times 10^6$
入力運動量領域 (GeV/c)	1.2-1.4

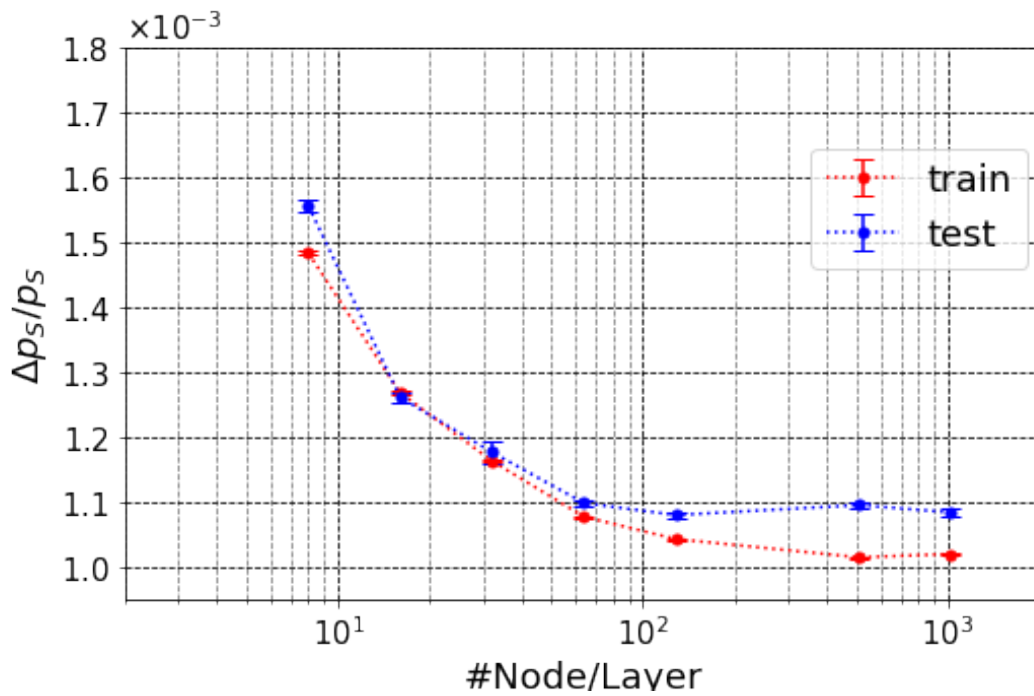


図 6.2: 運動量分解能 (半値全幅、FWHM) のノード数依存性。

こちらも訓練データに飛跡情報を用いていた場合と同じく、ノード数が  $\geq 128$  で訓練データとテストデータ間の分解能の差が広がり、 $\leq 32$  では分解能が収束しきれていない。したがって、1 隠れ層あたりのノード数には 64 を採用する。

- 隠れ層数の最適化

隠れ層数に関して他の訓練パラメータは以下の表のように固定して最適化を行った。

表 6.3: 隠れ層数最適化のために用いた訓練パラメータ。

訓練パラメータ	
活性化関数	Swish
ノード数	32
損失関数	平均絶対誤差 (MAE)
最適化関数	Adam
訓練データの統計量	$1.07 \times 10^6$
入力運動量領域 (GeV/c)	1.2–1.4

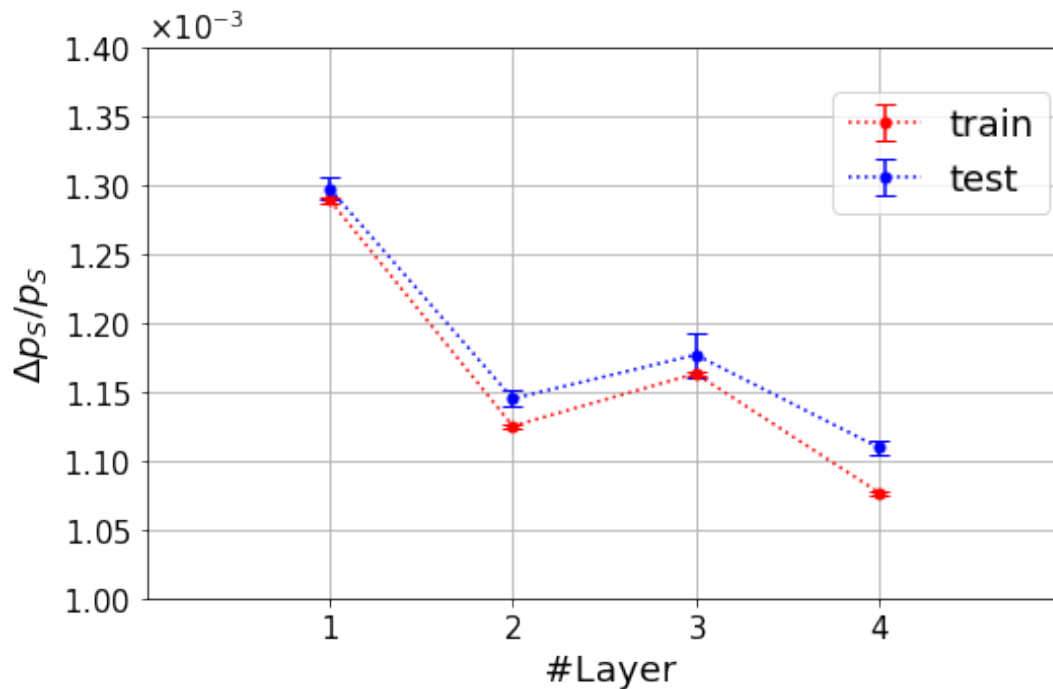


図 6.3: 運動量分解能 (半値全幅、FWHM) の隠れ層数依存性。

隠れ層数には 2 を採用した。

- 損失関数の最適化

損失関数に関して他の訓練パラメータは以下の表のように固定して最適化を行った。なお今回の最適化でも MAE と MSE の二つの損失関数を比較対象とする。



表 6.4: 損失関数最適化のために用いた訓練パラメータ。

訓練パラメータ	
活性化関数	Swish
隠れ層数	3
ノード数	32
最適化関数	Adam
訓練データの統計量	$1.07 \times 10^6$
入力運動量領域 (GeV/c)	1.2–1.4

表 6.5: 運動量分解能 (半値全幅、FWHM) の損失関数依存性。

損失関数	$\Delta p_S/p_S \times 10^{-3}$ (train)	$\Delta p_S/p_S \times 10^{-3}$ (test)
MAE	$1.14 \pm 0.00$	$1.18 \pm 0.01$
MSE	$1.32 \pm 0.00$	$1.34 \pm 0.01$

検証の結果、local hit 情報を訓練データとした場合も、損失関数に関しては絶対平均誤差 (MAE) の方が高い分解能で運動量を再構成できることがわかった。

結果として、最適化した訓練パラメータ (表 6.6) を基に学習で得られた分解能は  $\Delta p/p \sim 1.1 \times 10^{-3}$  (FWHM) となった。トラッキングした後の  $x, y, u, v$  を訓練データとした場合  $\Delta p/p \sim 1.0 \times 10^{-3}$  (FWHM) よりも分解能の悪化が見られた。

表 6.6: local hit 情報の訓練のために選定した訓練パラメータ。

訓練パラメータ	
活性化関数	Swish
ノード数/レイヤー	64
隠れ層数	2
損失関数	平均絶対誤差 (MAE)
最適化関数	Adam
訓練データの統計量	$1.48 \times 10^6$

また、機械学習模型で計算した  $p_S, \cos\theta$  の値からミッシングマスを再構成し、 $\Xi^-, \Sigma^-$ ,

$^{11}\text{B} + \Xi^-$  のミッシングマス分解能についても評価を行った。なお  $K^-$  の運動量分解能は  $3.3 \times 10^{-4}$  に固定して検証を行った。

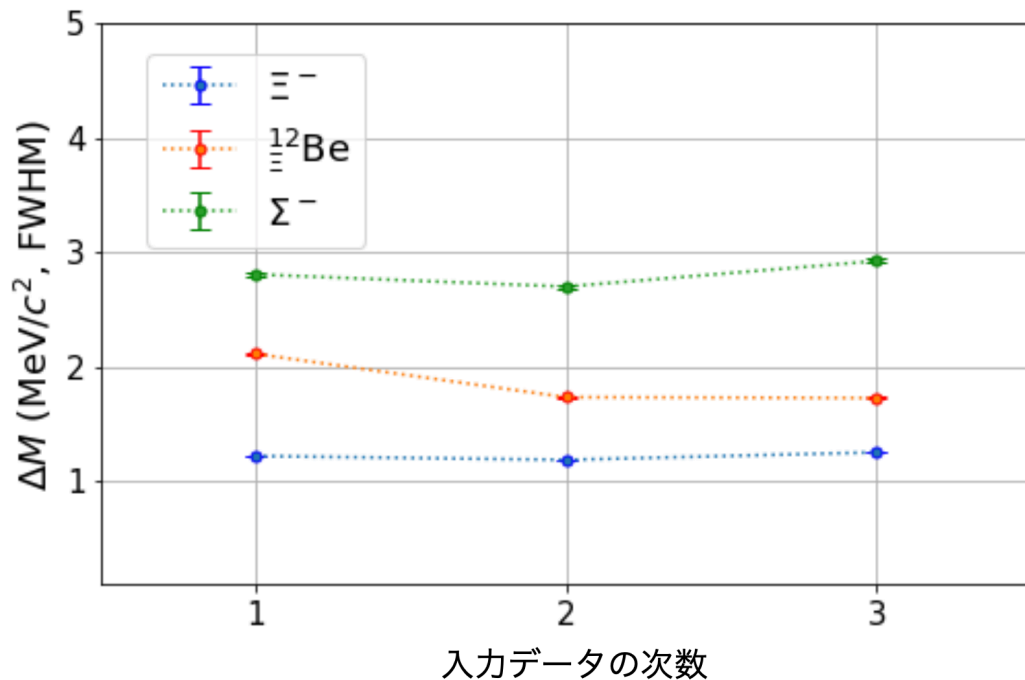


図 6.4: local hit 情報で計算した  $\Xi^-$ ,  $\Sigma^-$ ,  $^{11}\text{B} + \Xi^-$  のミッシングマス分解能の入力次数依存性。

入力データの最高次数が 1 から 3 までのモデルを構築し、それぞれでミッシングマス分解能を評価した。分解能は次数に依らず、 $\Xi^-$  のミッシングマス分解能は  $1.2 \text{ MeV}/c^2$  程度、 $^{12}\text{Be}$  の分解能は  $1.7 \text{ MeV}/c^2$ 、 $\Sigma^-$  の分解能は  $2.7 \text{ MeV}/c^2$  程度の結果となった。

ミッシングマス分解能に関しても、訓練データに飛跡情報を用いた場合よりも 1.5 倍程度悪化している。

## 6.2 教師データに飛跡情報を加えた場合

local hit 情報を直接機械学習で学習し運動量を再構成する場合、飛跡情報を訓練データとした場合よりも分解能が悪くなった。6.1 節では、教師データとして散乱運動量  $p_S$  と散乱角度  $\cos\theta$  を用いていたが、今回は訓練データをワイヤ面での local hit 情報、教師データには  $p_S$  と上流下流での飛跡情報  $x, y, u, v$  の計 9 つのデータを用いて学習を行った。教師データに reference 面での飛跡情報を加えることによって散乱粒子の飛跡がより

最適化され、local hit 情報での分解能の悪化を抑制できるかを検証した。

飛跡情報に関しては、S-2S の上下流にワイヤ面とは別に reference 面をそれぞれ設け、そこで捉えた  $x, y, u, v$  の値を教師データとして採用した。上流 reference 面については SDC1 と SDC2 の間に設置し、下流 reference 面は SDC5 の下流に設置した。

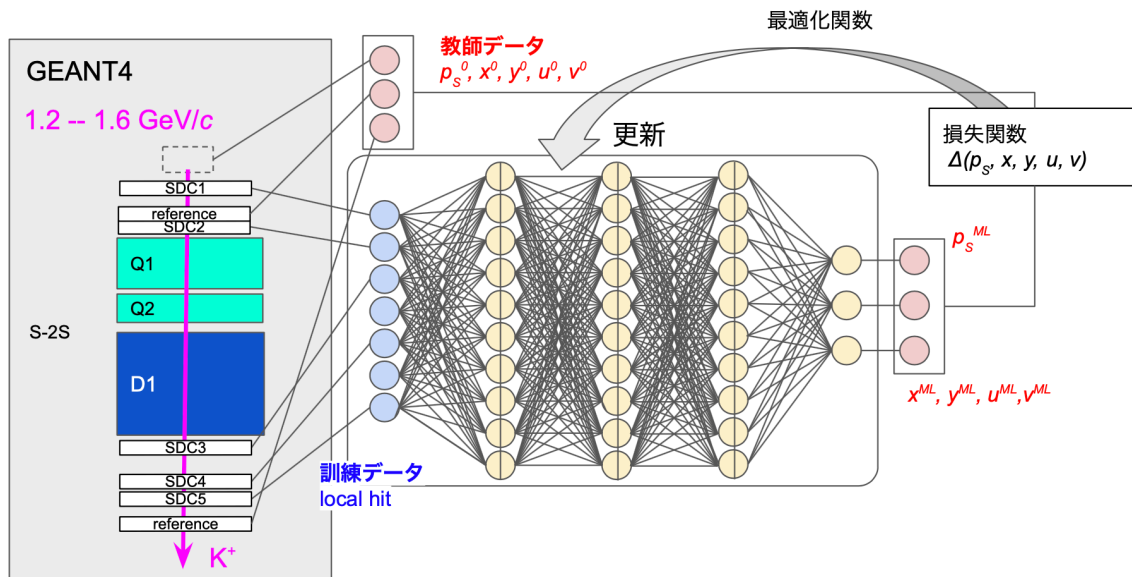


図 6.5: 飛跡情報を新たに教師データとして加え、local hit 情報を訓練データとして用いた機械学習模型。

訓練に用いたパラメータ値は以下の表 6.7 に示した。

表 6.7: local hit 情報の訓練のために選定した訓練パラメータ。

訓練パラメータ	
活性化関数	Swish
ノード数/レイヤー	32
隠れ層数	3
損失関数	平均絶対誤差 (MAE)
最適化関数	Adam
訓練データの統計量	$1.48 \times 10^6$

これらのパラメータ設定のもとで訓練データの最高次数を1、3の2パターンで検証した。ところ運動量分解能は表 6.8 のような結果が得られた。分解能は  $6.5 \sim 7.0 \times 10^{-3}$  と、教師データに飛跡情報を加えなかった場合より6倍程度悪化した。また作成した機械学習模型を用いて  $\Xi^-$ ,  $\Sigma^-$ ,  $^{12}\text{Be}$  のミッシングマスピークを計算した。散乱角度  $\cos \theta$  に関しては、機械学習模型で得られた上流の飛跡情報 ( $u^{\text{ML}}, v^{\text{ML}}$ ) から計算している。得られた  $p_S^{\text{ML}}$  と  $u^{\text{ML}}, v^{\text{ML}}$  からミッシングマスを計算した。ところが、 $u^{\text{ML}}, v^{\text{ML}}$  などの飛跡情報の値が教師データに収束されておらず、そこから計算した散乱角度  $\cos \theta$  が1を超える値をとり、ミッシングマスピークとしては再構成することができなかった。

表 6.8: 飛跡情報 ( $x, y, u, v$ ) を教師データに用いた場合に得られた運動量分解能  $\Delta p/p$ (半値全幅、FWHM)。

最大次数	$\Delta p_S/p_S \times 10^{-3}$ (train)	$\Delta p_S/p_S \times 10^{-3}$ (test)
1	$6.52 \pm 0.02$	$6.84 \pm 0.11$
3	$7.06 \pm 0.02$	$7.05 \pm 0.11$

## 第7章

# 機械学習模型の校正手法

これまでの機械学習模型の構築やスペクトロメータの性能評価では、シミュレーションデータを使用していた。しかし、シミュレーション上での実験セットアップは必ずしも現実でのセットアップと合致しているとは限らない。K1.8 実験エリアでの、2022年1月からの S-2S 電磁石や、ドリフトチェンバーのインストール作業は、セオドライトを用いて  $100\mu\text{m}$  以下の精度で行われた。しかし検出器群のズレ、シミュレーション上の磁場マップの精度の積算によって、現実の実験データでトラッキング後の  $x, y, u, v$  情報がシミュレーションで予想されるものから大きくずれる可能性がある。

したがって機械学習で運動量を見積もる場合、まずシミュレーションデータで学習した機械学習模型を、実際に得られた実験データで校正を行う必要がある。今回は実験データの代わりに、シミュレーション上で検出器を意図的にずらすことにより、シミュレーション空間での理想的なセットアップと現実との差異を表現する。なお本研究では、シミュレーション空間上の上流の検出器を  $x$  方向に  $+1.0\text{ mm}$  ずらしたセットアップで検証を行った。これにより  $\Xi$  ハイパー核の束縛エネルギー領域付近での分布は  $300\text{ keV}$  程度シフトすることがわかった。(図 7.1)

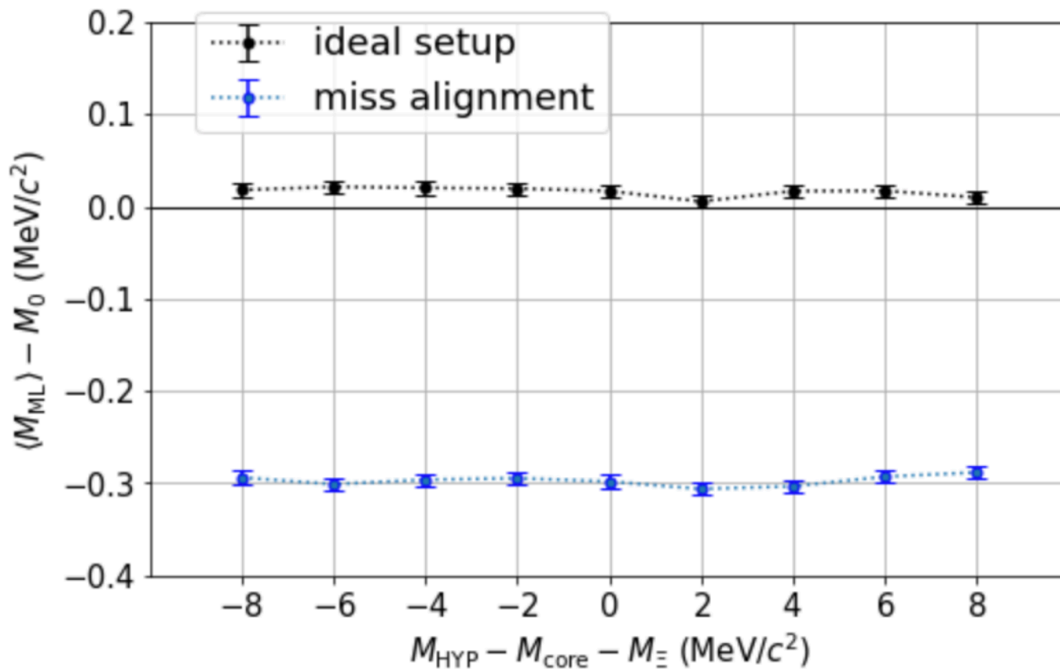


図 7.1: 上流検出器を  $x$  方向に 1 mm ずらした場合の  $^{12}\text{Be}$  の束縛領域におけるミッシングマス分布のシフト。黒がずらす前、青がずらした後の分布である。

具体的には 2 つの手法で校正できるかを検証した。1.2–1.4 GeV/ $c$  の範囲内の Beam Through データと  $\Xi^-$  ピークのデータを用いる検証と、1.2–1.6 GeV/ $c$  の範囲内の Beam Through と  $\Xi^-$ 、 $\Sigma^-$  ピークのデータを用いる検証を行った。それぞれで、 $\Xi$  ハイパー核ピークがどの程度校正できるかを検証した。なお Beam Through データとは、標的を設置しない状態、つまり入射粒子がビームラインスペクトロメータを通過し解析した後、反応を起こさずそのまま S-2S スペクトロメータで運動量解析されるデータである。実際に J-PARC E70 実験ではこの Beam Through run を行う。

## 7.1 Beam through、 $\Xi^-$ ピークデータを用いた校正

まずはシミュレーション空間と現実のセットアップの間の差を、 $\Xi^-$  ピークと Beam Through データを用いて補正できるかを検証する。

シミュレーション空間での飛跡情報で (図 7.2) 学習し、機械学習モデルを構築する。訓練データには S-2S アクセプタンス内の 1.2–1.4 GeV/ $c$  の範囲の飛跡情報を使用した。教師データには、4.6 節にて述べた運動量  $p_S$ 、散乱角度  $\cos\theta$  を使用する。構築した模型

を「ベースモデル」と呼ぶことにする。ベースモデルの訓練の際の損失関数には、 $p_S$  と  $\cos\theta$  の絶対平均誤差 (MAE) をそれぞれ 0.8, 0.2 で重み付けした関数を採用した。ここでの絶対平均誤差の重みの値は、ミッシングマス分解能での散乱運動量による誤差と散乱角度による誤差の重みの値に従い決定した。このベースモデルをずれセットアップでの  $p(K^-, K^+)\Xi^-$ 、 $p(K^-, \pi^+)\Sigma^-$  と Beam Through データで再学習し修正する。

Beam through データは、1.25, 1.3, 1.35, 1.36, 1.37, 1.375, 1.4 GeV/c の運動量を用いる。運動量分解能は K1.8 ビームラインスペクトロメータの設計値に従い  $\Delta p_B/p_B = 3.3 \times 10^{-4}$  (FWHM) とした。ずらしたセットアップの元で取得した Beam through データを用いて、 $p_S$  を出力する結合層の重みパラメータのみを更新する。Beam through データで再学習したモデルを「BT モデル」と呼ぶことにする。(図 7.3)

次に BT モデルを  $\Xi^-$  ピークで修正する。まず、ずらしたセットアップの元で  $p(K^-, K^+)\Xi^-$  の運動学に従った  $K^+$  をデータを照射し、取得した飛跡情報を BT モデルで計算する。BT モデルで出力されたミッシングマス分布をガウスフィットする。ガウスフィットした  $\Xi^-$  事象のなかで  $\pm 2\sigma$  以内の事象のみを選別した。(図 7.4) この選別した事象が、PDG[5] での値 ( $M_\Xi = 1321.71 \text{ MeV}/c^2$ ) に合うように更に BT モデルを校正する。この校正では、主に飛跡情報と散乱角度  $\theta$  間の相関関係を修正することを目標としている。したがって、 $\cos\theta$  を出力する結合層の重みパラメータのみを更新する(図 7.5)。なお訓練データ数は、 $\Xi^-$  を  $1.0 \times 10^4$  程度とした。

校正において、BT-model の 2 つの出力層  $p_S$ ,  $\cos\theta$  からミッシングマスを計算する層を下位層に加える。損失関数にはミッシングマスの 2 乗の平均絶対誤差 ( $\Delta M^2 = |M_0^2 - M_{ML}^2|$ ) を用いて、 $\cos\theta$  を出力する結合層の重みパラメータを再学習する。ミッシングマスを直接損失関数として用いると訓練中に平方根の中が負になり、損失関数が複素数になり訓練が途中で停止する可能性があるため、損失関数にはミッシングマスの 2 乗を用いた。Beam through データの後に  $\Xi^-$  で再学習したモデルを「 $\Xi$  モデル」と呼ぶことにする。

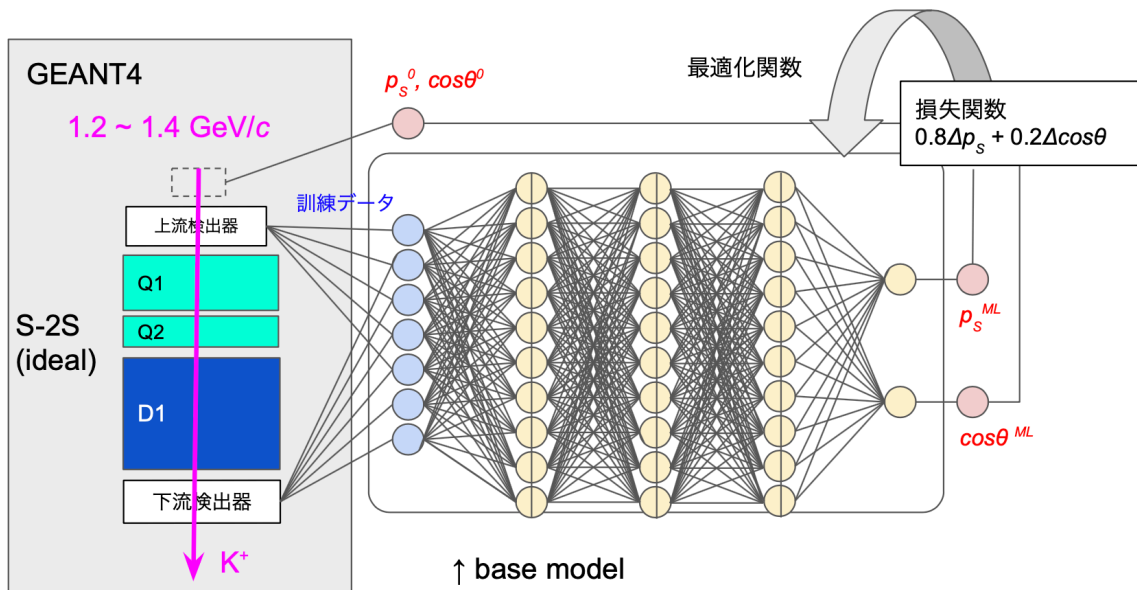


図 7.2: 理想セットアップ下の飛跡情報を訓練データとしてベースモデルを構築する。

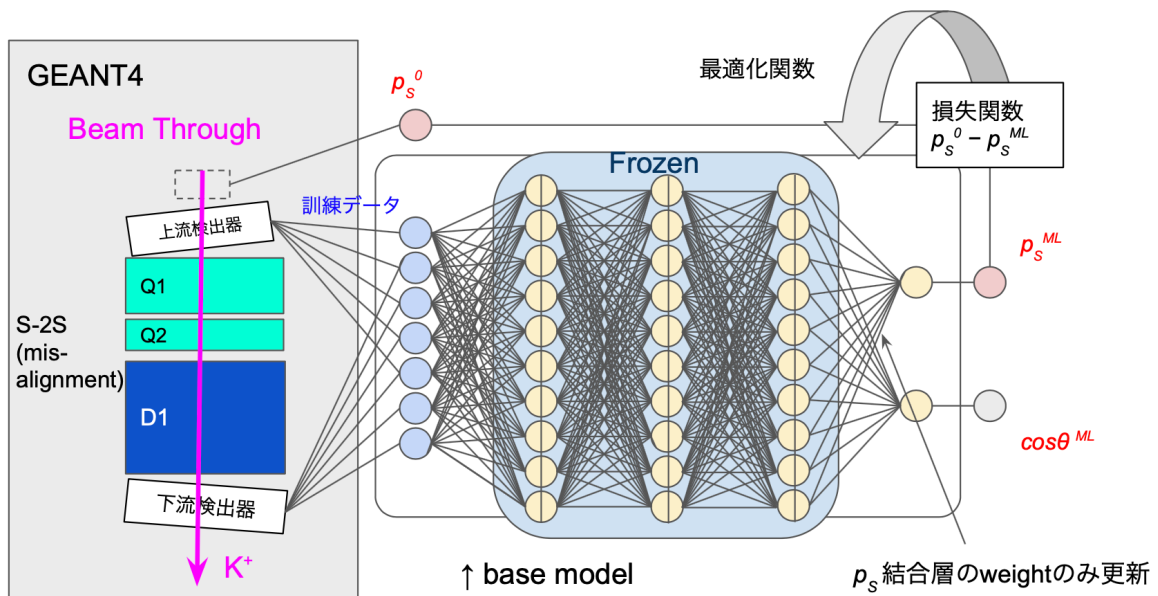


図 7.3: ベースモデルを Beam through データで再学習し修正する。



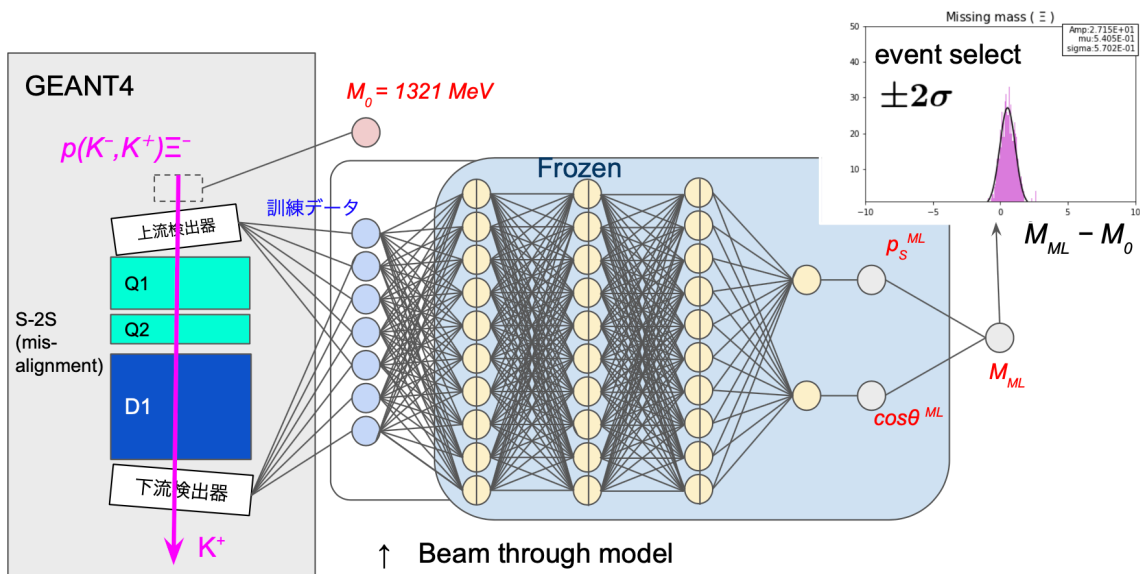


図 7.4: BT モデルで、 $\Xi^-$  のミッシングマス分布を計算する。  
得られた分布のうち、 $\pm 2\sigma$  以内のイベントを選別する。

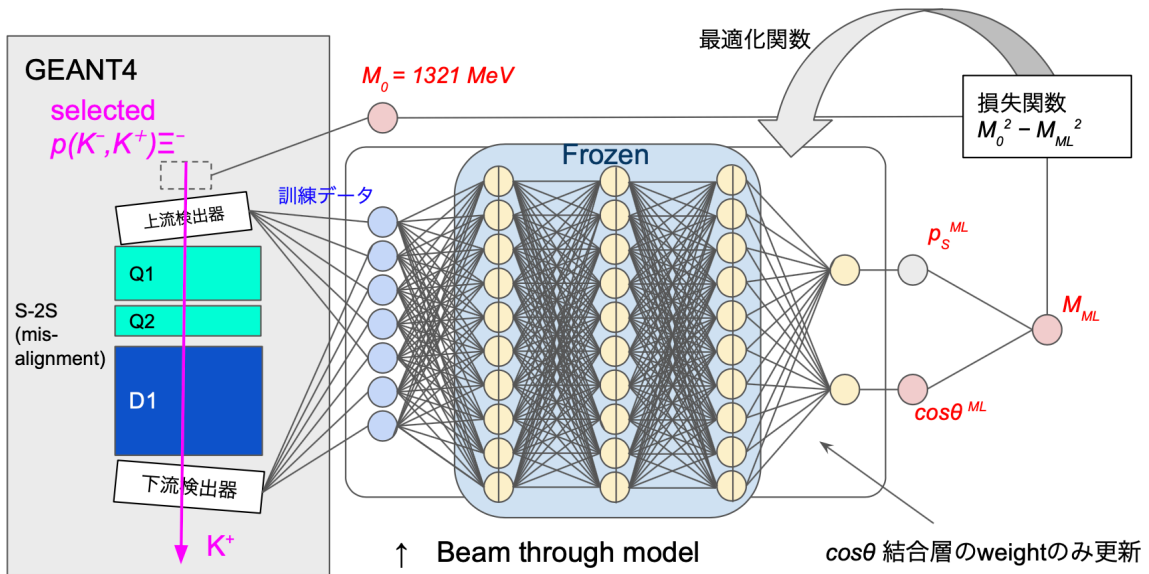


図 7.5: 最後に BT モデルを選別した  $\Xi^-$  ピークで更に再学習する。

以上の手法で修正し、base, BT,  $\Xi$  モデルのそれぞれでの  $\Xi^-$  と  $^{11}\text{B} + \Xi^-$  ピーク、 $^{12}\text{Be}$  の束縛領域のミッシングマス分解能  $\Delta M$  と、ピークがどの程度シフトしているか

$(\langle M_{ML} \rangle - M_0)$  を評価した。

まず  $\Xi^-$  と  $^{11}\text{B} + \Xi^-$  ピークの  $\Delta M$  に関しては、上流検出器をずらした場合には悪化が認められなかった。しかし Beam through で校正を行った場合、 $\Xi^-$  と  $^{11}\text{B} + \Xi^-$  のいずれも分解能の悪化が認められた。その後の  $\Xi^-$  による校正では分解能は改善された。

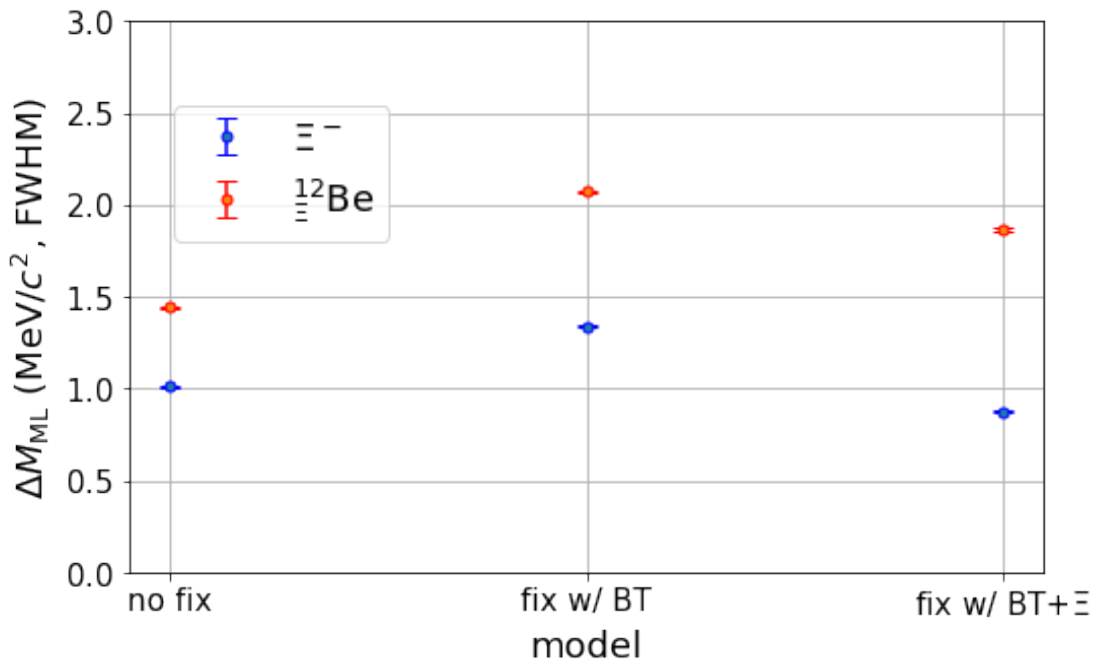


図 7.6: ベースモデル (no fix) , BT モデル (fix w/ BT) ,  $\Xi^-$  モデル (fix w/ BT+ $\Xi^-$ ) のそれぞれで計算した、 $\Xi^-$  と  $^{11}\text{B} + \Xi^-$  ピークのミッシングマス分解能  $\Delta M$ 。

$^{12}\text{Be}$  の束縛領域の分解能においても同様の結果が得られた。

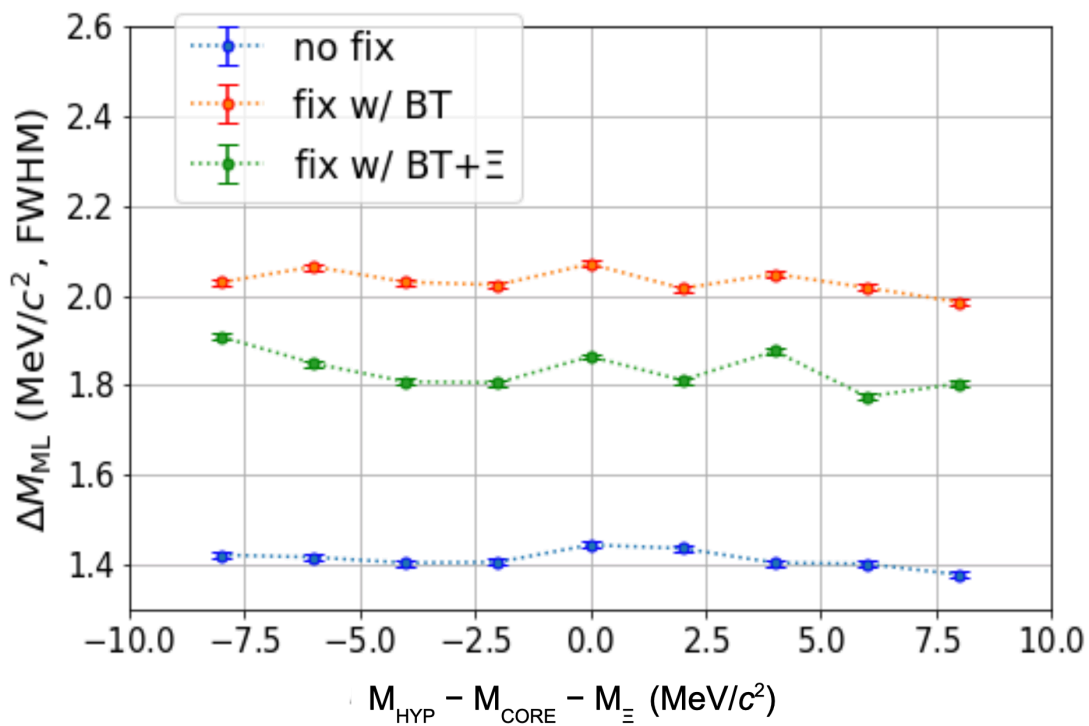


図 7.7: ベースモデル (青) , BT モデル (赤) ,  $\Xi$  モデル (緑) のそれぞれで計算した  $^{12}\text{Be}$  の束縛領域近傍におけるミッシングマス分解能。

一方でミッシングマスのシフト ( $\langle M_{\text{ML}} \rangle - M_0$ ) について評価したところ、上流検出器が  $x$  方向に 1 mm シフトさせた場合ミッシングマス分布が  $-0.3 \text{ MeV}/c^2$  程度のシフトが認められた。これを Beam through と  $\Xi^-$  によって校正を試みたところ、 $^{11}\text{B} + \Xi^-$  ピークは  $2.29 \text{ MeV}/c^2$  分シフトし、校正前と比べ悪化が見られた。

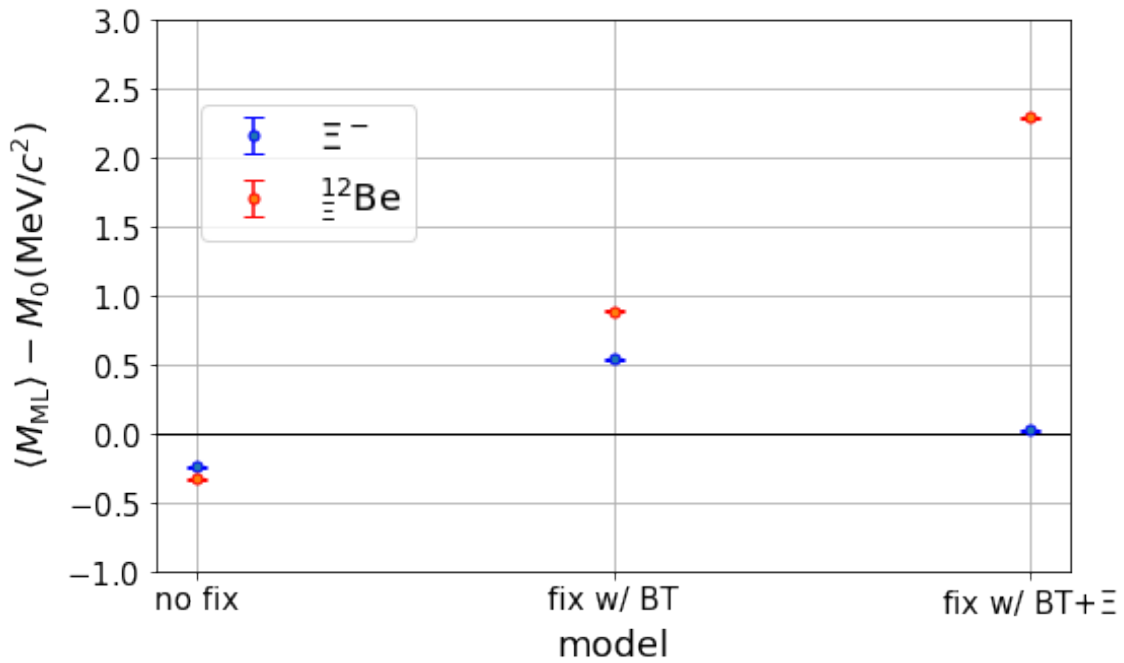


図 7.8: ベースモデル (no fix) , BT モデル (fix w/ BT) ,  $\Xi$  モデル (fix w/ BT+ $\Xi$ ) のそれぞれで計算した、 $\Xi^-$  と  $^{12}\text{Be}$  と  $^{11}\text{B} + \Xi^-$  ピークのミッシングマス分布のシフト ( $\langle M_{ML} \rangle - M_0$ )。

こちらも  $^{12}\text{Be}$  の束縛領域の分解能においても同様の結果が得られた。

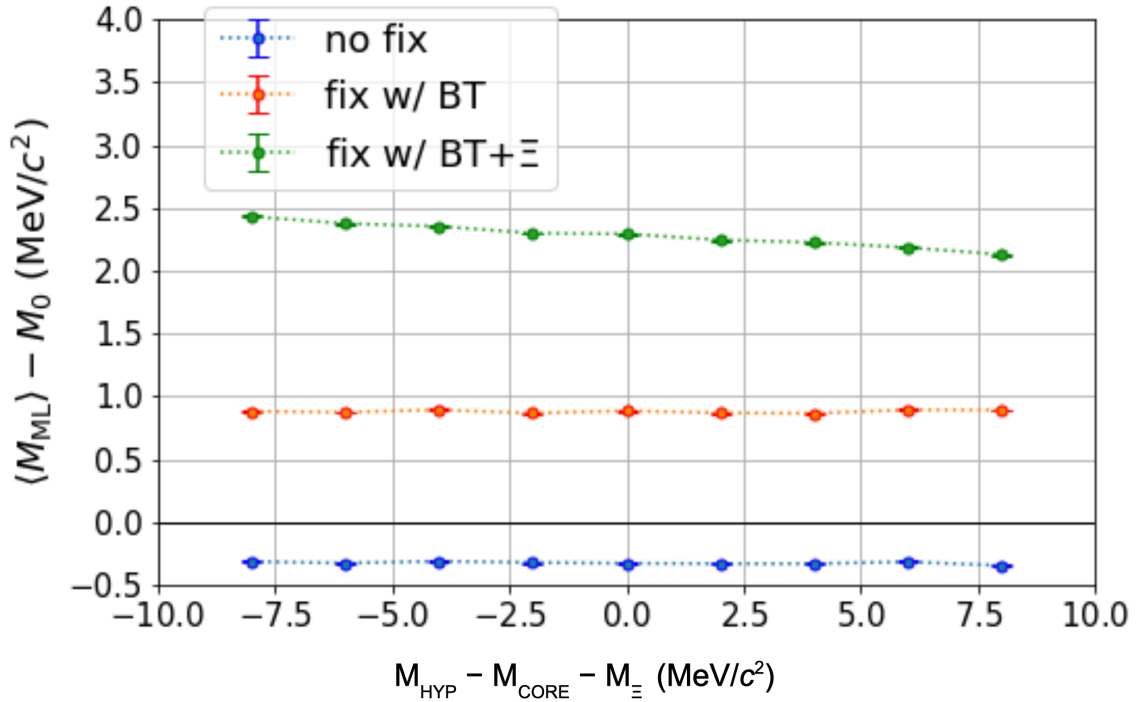


図 7.9: ベースモデル (青), BT モデル (赤),  $\Xi$  モデル (緑) で計算した  $^{12}\text{Be}$  の束縛領域におけるミッシングマス分解能。

表 7.1: エネルギー校正によるミッシングマスピークの分解能  $\Delta M$  と分布のシフト  $\langle M_{ML} \rangle - M_0$  のモデルごとの比較。

$\Xi^-$ ピーク			
model	base (修正前)	BT(1.2–1.4 GeV/c)	BT+ $\Xi$
$\Delta M(\text{MeV}/c^2, \text{FWHM})$	1.01	1.33	0.87
$\langle M_{ML} \rangle - M_0(\text{MeV}/c^2)$	-0.236	0.540	$2.75 \times 10^{-2}$
$^{11}\text{B} + \Xi^-$			
model	base (修正前)	BT(1.2–1.4 GeV/c)	BT+ $\Xi$
$\Delta M(\text{MeV}/c^2, \text{FWHM})$	1.44	2.07	1.86
$\langle M_{ML} \rangle - M_0(\text{MeV}/c^2)$	-0.328	0.885	2.29

1.2–1.4 GeV/c の Beam through、 $\Xi^-$  で校正しようとしたが、逆にミッシングマス分解能  $\Delta M$  と分布のシフト ( $\langle M_{ML} \rangle - M_0$ ) は共に悪化する結果となった。

## 7.2 Beam through、 $\Xi^-$ 、 $\Sigma^-$ ピークデータを用いた校正

次に  $\Xi^-$  ピークと Beam Through データに加えて、 $\Sigma^-$  ピークのデータを校正に用いた場合、ピーク位置が改善されるかを検証した。

$\Sigma^-$  の運動量領域は  $< 1.58 \text{ GeV}/c$  であるから、base-model の学習では S-2S アクセプタンス内の  $1.2\text{--}1.6 \text{ GeV}/c$  の範囲のデータを用いた。Beam through データは、1.25, 1.3, 1.35, 1.36, 1.37, 1.375, 1.4, 1.45, 1.5, 1.55  $\text{ GeV}/c$  の運動量を用いた。運動量分解能は K1.8 ビームラインスペクトロメータの設計値に従い  $\Delta p/p = 3.3 \times 10^{-4}$  (FWHM) とした。Beam through データを用いた校正手法は前小節と同じである。(図 7.3)

BT モデルは、 $\Sigma^-$ 、 $\Xi^-$  ピークで修正した。先に述べたように、BT モデルにずらした  $p(K^-, K^+)\Xi^-$ 、 $p(K^-, \pi^+)\Sigma^-$  データを入力し、出力されたミッシングマス分布をガウスフィットする。この時、 $\Xi^-$ 、 $\Sigma^-$  事象のなかで  $\pm 2\sigma$  以内の事象のみを選別する。(図 7.4) この選別した事象が PDG[5] での値 ( $M_{\Xi^-}$ ,  $M_{\Sigma^-} = 1321.71, 1197.45 \text{ MeV}/c^2$ ) に合うように更に BT モデルを校正する。この校正では、主に飛跡情報と散乱角度  $\theta$  間の相関関係を校正することを目標としている。したがって、出力の  $\cos\theta$  層の重みパラメータのみを更新する。 $\Sigma^-$  と  $\Xi^-$  によって校正された後のモデルを「 $\Sigma\Xi$  モデル」と呼ぶことにする。なお訓練データ数は、 $\Xi^-$  を  $1.0 \times 10^4$ 、 $\Sigma^-$  を  $5.0 \times 10^3$  程度とした。(図 7.5)

校正前と、Beam through 校正、 $\Sigma\Xi$  校正での、 $\Xi^-$ 、 $\Sigma^-$ 、 $^{12}\text{Be}$ 、およびその束縛領域でのミッシングマス分解能  $\Delta M$  と分布のシフト ( $\langle M_{\text{ML}} \rangle - M_0$ ) を表に示す。

まず、分解能  $\Delta M$  については、修正後に  $0.01 \text{ MeV}/c^2$  程度の一様な悪化が見られたが、 $\Sigma$  を用いない校正ほどの分解能悪化は見られなかった。

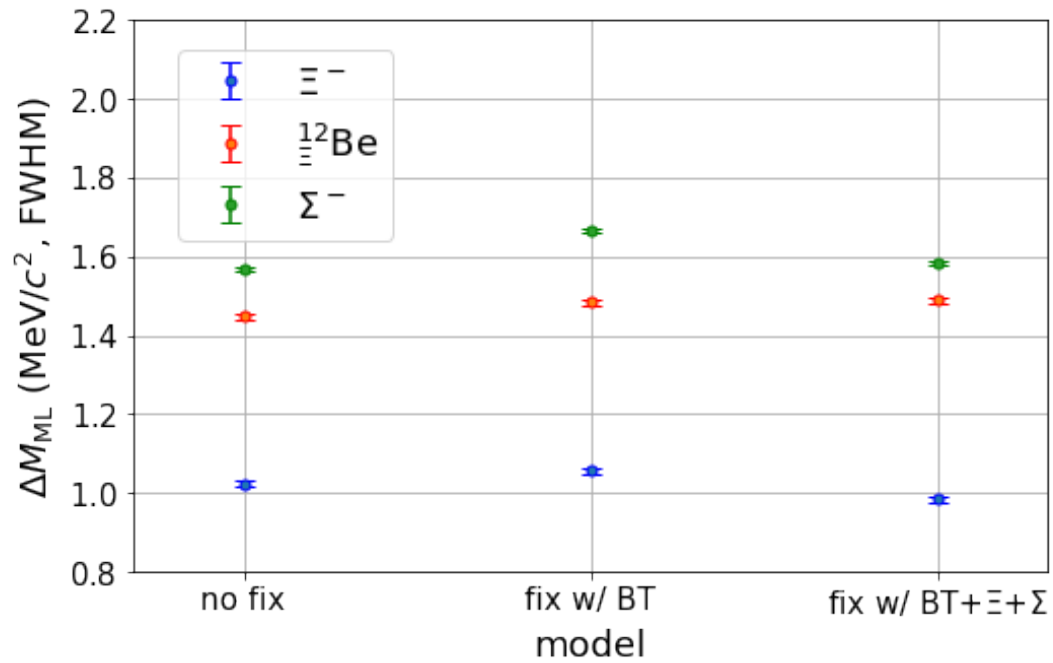


図 7.10: ベースモデル (no fix) , BT モデル (fix w/ BT) ,  $\Sigma\Xi$  モデル (fix w/ BT+ $\Xi$ ) で計算した  $\Xi^-$  ,  $\Sigma^-$ 、 $^{11}\text{B} + \Xi^-$  のミッシングマス分解能。

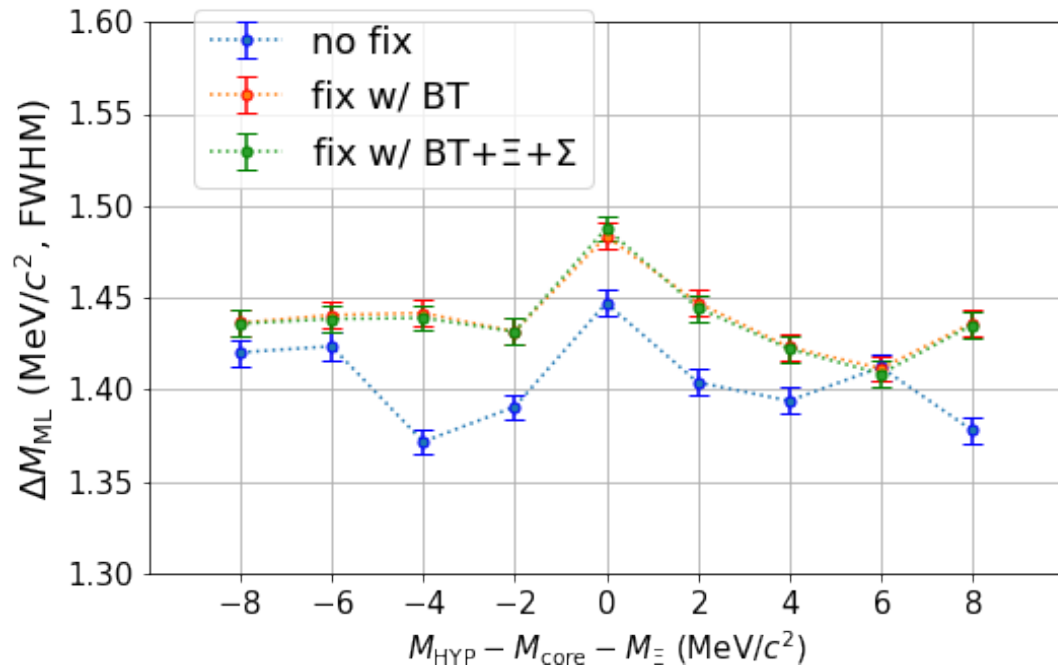


図 7.11: ベースモデル (青), BT モデル (赤),  $\Sigma\Xi$  モデル (緑) で計算した  $^{12}_{\Xi}\text{Be}$  の束縛領域におけるミッシングマス分解能。

次に校正前と、Beam through 校正、 $\Sigma\Xi$  校正での、 $\Xi^-$ ,  $\Sigma^-$ ,  $^{12}_{\Xi}\text{Be}$ , およびその束縛領域でのミッシングマス分布のシフト ( $\langle M_{\text{ML}} \rangle - M_0$ ) を図に示す。



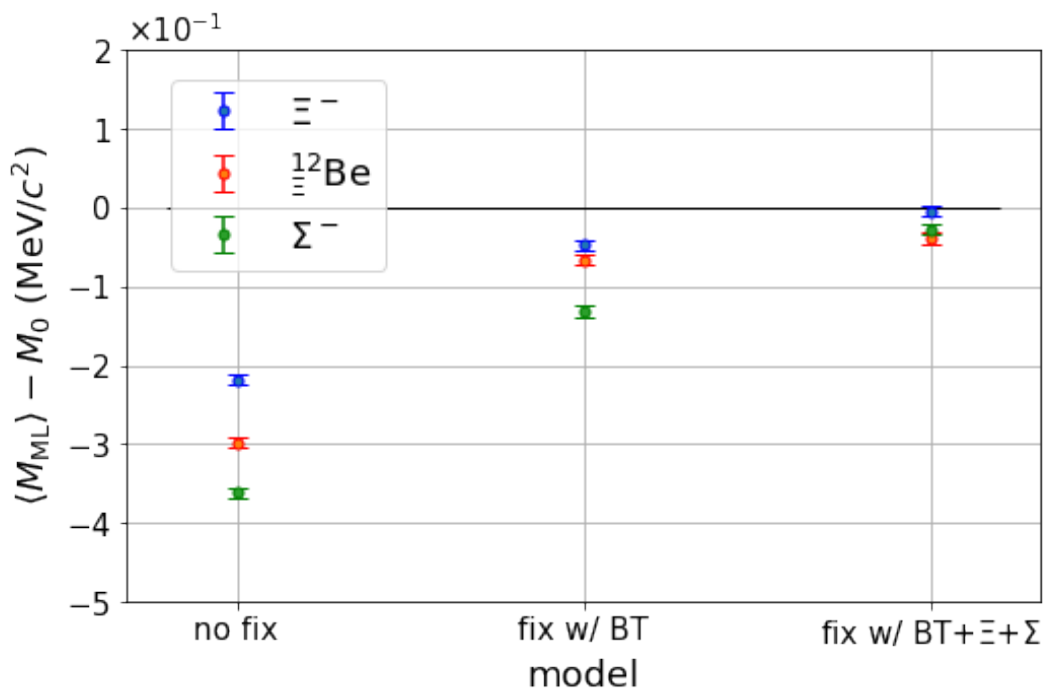


図 7.12: ベースモデル (no fix) , BT モデル (fix w/ BT) ,  $\Xi$  モデル (fix w/ BT+ $\Sigma\Xi$ ) で計算した  $\Xi^-$ ,  $\Sigma^-$ ,  $^{11}\text{B} + \Xi^-$  のミッシングマス分布のシフト  $\langle M_{ML} \rangle - M_0$ 。

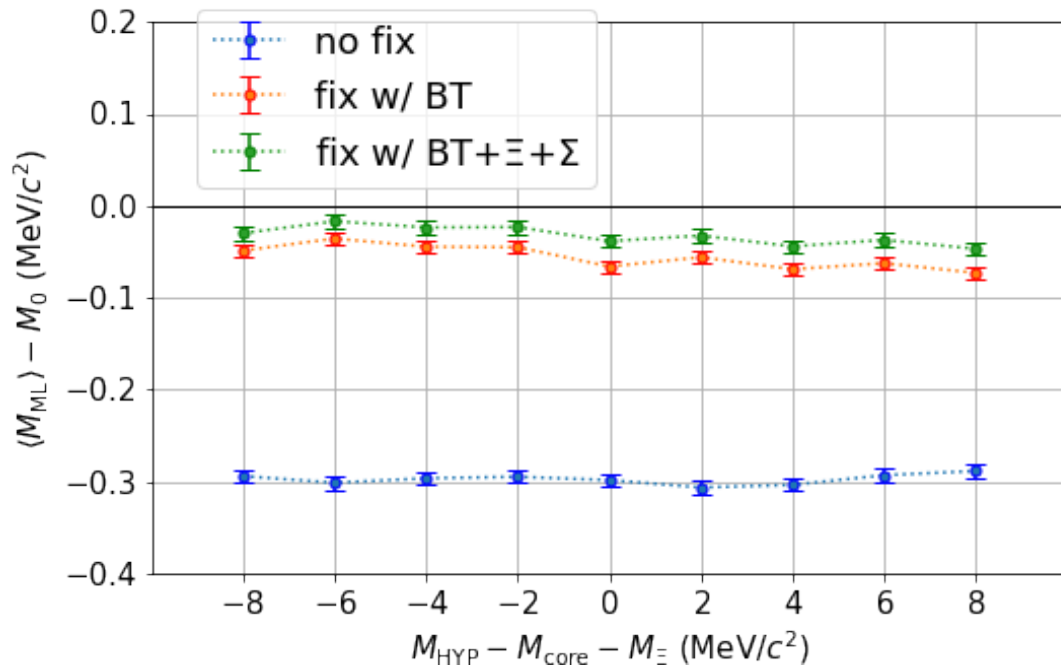


図 7.13: ベースモデル (青), BT モデル (赤),  $\Sigma\Xi$  モデル (緑) で計算した  $^{12}_{\Xi}\text{Be}$  の束縛領域におけるミッシングマス分布のシフト  $\langle M_{ML} \rangle - M_0$ 。

表 7.2: エネルギー校正によるミッシングマスピークの分解能  $\Delta M$  と分布のシフト  $\langle M_{\text{ML}} \rangle - M_0$  のモデルごとの比較。

$\Xi^-$ ピーク			
model	base (校正前)	BT (1.2–1.6 GeV/c)	BT+ $\Sigma\Xi$
$\Delta M(\text{MeV}/c^2, \text{FWHM})$	1.02	1.06	0.983
$\langle M_{\text{ML}} \rangle - M_0(\text{MeV}/c^2)$	-0.218	$-4.79 \times 10^{-2}$	$-4.48 \times 10^{-2}$
$^{11}\text{B} + \Xi^-$			
model	base (校正前)	BT(1.2–1.6 GeV/c)	BT+ $\Sigma\Xi$
$\Delta M(\text{MeV}/c^2, \text{FWHM})$	1.45	1.48	1.49
$\langle M_{\text{ML}} \rangle - M_0(\text{MeV}/c^2)$	-0.298	$-6.61 \times 10^{-2}$	$-3.85 \times 10^{-2}$
$\Sigma^-$ ピーク			
model	base (校正前)	BT(1.2–1.6 GeV/c)	BT+ $\Sigma\Xi$
$\Delta M(\text{MeV}/c^2, \text{FWHM})$	1.58	1.67	1.58
$\langle M_{\text{ML}} \rangle - M_0(\text{MeV}/c^2)$	-0.362	-0.131	$-2.76 \times 10^{-2}$

1.2–1.6 GeV/c の Beam through,  $\Xi^-$ ,  $\Sigma^-$  で校正した結果、 $^{12}\text{Be}$  のミッシングマス分布のシフト ( $\langle M_{\text{ML}} \rangle - M_0$ ) に関しては、大きく改善され、 $-0.3 \text{ MeV}/c^2$  程度シフトしていた分布が  $-0.03 \text{ MeV}/c^2$  程度まで抑制することができた。検出器をずらしていない時のセットアップでの、ミッシングマス分布のシフトの値も  $\sim 2 \times 10^{-2}$  程度であったことから、十分抑制されている。

検証の結果、校正に用いるデータにおいては、Beam through は 1.2–1.6 GeV/c の広い範囲の運動量領域をとり、ミッシングマスピークとしては  $\Xi^-$  だけでは十分ではなく  $\Sigma^-$  も加えることによって校正可能であることが示された。

## 第8章

# 結論

J-PARC E70 実験では、世界最高分解能で  $\Xi$  ハイパー核を分光することにより、そのピーク構造から  $\Xi N$  間相互作用に決定的な情報を与えられると期待されている。これにより核力の拡張である、 $S = -2$  多体系でのバリオン間力の理解につながる。本研究では、高分解能な  $\Xi$  ハイパー核のミッシングマス分光のために、機械学習 (ML) を用いた磁気スペクトロメータ S-2S のための運動量解析手法のフレームワーク開発を行った。磁気スペクトロメータの運動量解析では、スペクトロメータの上下流での飛跡情報から ML 模型を用いた計算により運動量を再構成する。ML 模型構築のための学習に用いる訓練データや教師データには、GEANT4 シミュレータから取得したシミュレーションデータを用いた。

学習においては、ML 模型を過学習がなく高精度で学習させるために、損失関数や最適化関数などの訓練パラメータの最適化を行った。最適化によって構築した ML 模型を用いて飛跡情報から運動量分解能を見積もった結果、最高で  $\Delta p/p \simeq 1.00 \times 10^{-3}$  (半値全幅, FWHM) の結果が得られた。またドリフトチェンバーのワイヤ面のヒット位置から見積もった場合、 $\Delta p/p \simeq 1.10 \times 10^{-3}$  の結果が得られた。これらの運動量分解能と標的物質によるエネルギーストラグリングの効果を考慮した場合、 ${}_{12}^{\text{Be}}$  のミッシングマス分解能は、2.3 MeV 程度であった。この分解能は、ESC08a モデルを用いた理論計算により予想されるピークを弁別するには十分なものであると確認された。

さらに、解析フレームワークのエネルギー校正の可能性についても評価した。シミュレーションデータによって構築した ML 模型を現実のデータによって校正可能かを評価するために、シミュレーション上で仮想的に検出器をずらし、校正データ (Beam through と  $\Xi^-$ ,  $\Sigma^-$  生成事象) を用いた ML 模型のエネルギー校正の試験を行った。その結果、ここで用いた ML 模型の場合、Beam through と  $\Xi^-$  生成事象のみの校正では、 $\Xi$  ハイパー

核の質量測定に対して十分なエネルギー決定精度は得られないことがわかった。そこで、Beam through と  $\Xi^-$  生成事象に加えて、S-2S のアクセプタンス範囲に混入する  $\Sigma^-$  生成事象もエネルギー校正に使用する試験を行った。その結果、シミュレーション内に仮想的に導入した検出器位置のずれによりハイパー核の質量に  $300 \text{ keV}/c^2$  程度の誤差が生じていたものが、これらの校正データにより、ずらす前の  $30 \text{ keV}/c^2$  程度に抑えることができた。従って、本 ML 模型のエネルギー校正が、Beam through と  $\Xi^-$ ,  $\Sigma^-$  生成事象を用いることで可能であることが証明された。

本研究により、実際の実験データ解析に導入できる ML を用いた運動量解析フレームワークの構築に成功した。

## 第9章

# 今後の展望

### 9.1 機械学習モデルの拡張

本研究では、運動量  $p_S$  を飛跡情報  $x, y, u, v$  から機械学習を用いて  $1.0 \times 10^{-3}$  程度の分解能で表現することができた。しかし今回確認された運動量分解能は、S-2S の設計値である  $6.0 \times 10^{-4}$  に及ばなかった。さらに輸送行列を用いて運動量を再構成した場合も同様に  $1.0 \times 10^{-3}$  程度の分解能が出ている。これらの結果からシミュレータ中の S-2S の計算磁場マップに改善の余地があると考えられる。

上流・下流での飛跡情報  $(x, y, u, v)$ 、とドリフトチェンバーのワイヤー面での local hit 位置情報という2つの階層で運動量解析フレームワークを構築し、散乱運動量を再構成できるかを検証した。local hit 情報を用いたフレームワークにおいては、飛跡情報によるフレームワークの結果よりも分解能がわずかに低くなった。教師データに飛跡情報を追加した場合は逆に分解能が悪化した。これに関しては、モデルの構築の仕方などの改善の余地がある。また、運動量再構成する際に、さらに下の階層の時間情報から運動量を導出できるモデルも将来的には構築する。local hit 位置を導出する時は、まずドリフト時間情報を位置情報に変換する必要がある。時間情報は高次の関数を用いて位置情報に変換されてきたが、本研究の延長としてはこの時間情報を直接訓練データとして、運動量が導出できるような機械学習モデルも構築する。

### 9.2 本解析手法の運用について

今回はシミュレーション空間で閉じた S-2S の運動量解析手法について述べたが、S-2S と K1.8 ビームラインスペクトロメータの間の運動量整合性 ( $p_B = p_S$ ) についても検討

する必要がある。このために、構築した運動量解析フレームワークを K1.8 ビームラインスペクトロメータを含めた模型に拡張し、両スペクトロメータで整合性の取れるフレームワークを構築する。

実際に実験データで解析する場合は、過学習が生じていないことや、4.5 節で述べたような運動量  $p_S$  と飛跡情報の間の相関を確認し、多角的に最適な運動量再構成が行われているかを吟味し運用する必要がある。また、従来手法であるルンゲクッタ法や逆輸送行列を用いた手法と比較し、機械学習による手法の利点、欠点などを定量的に評価し、手法の選定と実用を行う計画である。

# 謝辞

本修論を書き上げるにあたっては、多くの方々の協力がありました。この修論の謝辞の場を借りて感謝申し上げます。

指導教官である永江知文教授には、本修論である S-2S 磁気スペクトロメータの解析手法の開発や J-PARC E70 実験およびそれに伴う S-2S 検出器系の初期の立ち上げという貴重な研究課題を与えてくださいました。ハイパー核物理や磁気スペクトロメータの運動量解析などの知識について何も知らなかった私に、とても丁寧かつ深く教えてくださり、とても勉強になりました。本修論においては、毎週のミーティングなどで修論の方向性に関するアドバイスや議論などをして下さり、本修論を書き上げるにあたってはお忙しい中何度も添削して下さい、本当に感謝しております。今後ともご指導のほどよろしくお願いいたします。

後神利志助教には、修論での GEANT4 シミュレータの使い方、修論の議論から、実験現場での水チェレンコフ検出器や TOF カウンターの準備やテストに至るまで、実験物理に関わる様々なことを学ばせていただきました。また、その明るい性格には何度も救われたと思います。J-PARC の PAC の準備などで大変お忙しい中、夜中まで私の修論を何度も添削して下さいととても感謝しております。

同じ E70 実験グループの先輩である原田健志氏には、K1.8 実験エリアでの S-2S の検出器立ち上げなどにおいて、大変お世話になりました。ドリフトチェンバーなどの検出器インストール作業や、BFT や BH1 の検出器の立ち上げテストにおいては、未熟な私に対し、とてもわかりやすく丁寧にアドバイスをいただきとても感謝しております。またそれだけではなく、事務手続きのやり方や、修論の方向性のアドバイスであったり、何から何までお世話になったと思います。また、実験現場で先頭に立って実験準備を指揮している姿には、毎度毎度感服しており、その姿は私の目標でもあります。これからも E70 実験遂行に向けて何かとご迷惑おかけするかもしれませんが、今後ともよろしくお願いいたします。また、七村拓野氏には J-PARC での生活に関するアドバイスや、修論に関しても磁



場マップなどに関して丁寧に説明して下さいました。京都から遠く離れ J-PARC に来たばかりの私にとっては心強かったです。ありがとうございます。

JAEA の市川裕大氏には、私を JAEA の特別研究生として受け入れて下さいました。実験準備作業での的確なアドバイスを頂いたり、申請書の添削など快く引き受けてくださり感謝しております。また物理にとっても情熱を持って取り組んでいる姿は本当に尊敬いたします。JAEA の山本剛史氏や KEK の鵜養美冬氏には、ドリフトチェンバーの架台設計、製作、検出器立ち上げテスト、検出器の測量など実験現場の作業において非常にお世話になりました。実験現場での豊富な経験と知識には毎度驚かされるばかりでした。今後ともよろしく願いいたします。

また東北大、東工大の学生の皆さまには実験準備をともに盛り上げ、楽しい時間を過ごさせていただきました。東北大の坂尾珠和氏、鎌田健人氏、大浦文也氏、成済秀氏、渡辺大護氏、大橋和真氏、東工大の根岸亮輔氏、小堀匠氏、京大の高橋秀治氏とは、実験準備作業において共に切磋琢磨しあっていました。それだけではなく、よく食事に連れて行ってもらったり、他愛もない会話をしたりと、精神的な支えとなりました。皆さんのおかげで J-PARC での充実した研究生生活を送ることができたと思います。

それから当研究室の秘書の木村友理さんには、検出器架台の部品発注や旅費の申請などの手続きで大変お世話になりました。今後とも事務手続きの際はよろしく願いします。

他にも当研究室の方々や、また J-PARC で実験準備に共に携わっているの方々など、この修士課程の間とてもたくさんの人のお世話になりました。本当にありがとうございます。これからも J-PARC での実験の遂行に向けて精進してまいりますので、今後ともどうぞよろしく願いいたします。

## 参考文献

- [1] T. Inoue *et al.* (HAL QCD collaboration), Prog. Theor. Phys. **124** 591 (2010) .
- [2] P. B. Demorest *et al.*, Nature **467**, 1081 (2010).
- [3] K. Miwa *et al.* (J-PARC E40 Collaboration), Phys. Rev. C **104**, 045204 (2021).
- [4] T. Nanamura *et al.*, Prog. Theor. Exp. Phys. **9** 093D01 (2022).
- [5] P. A. Zyla *et al.* (Particle Data Group), Prog. Theor. Exp. Phys. **2020**, 083C01 (2020).
- [6] H. Tamura *et al.*, Phys. Rev. Lett. **84** 5963 (2000).
- [7] T. Nagae *et al.*, Phys. Rev. Lett. **80** 1605 (1998).
- [8] T. Hasegawa *et al.*, Phys. Rev. C **53**, 3 (1996).
- [9] T. Hotchi *et al.*, Phys. Rev. C **64**, 044302 (2001).
- [10] H. Takahashi *et al.*, Phys. Rev. Lett. **87**, 212502 (2001).
- [11] K. Nakazawa *et al.*, Prog. Theor. Exp. Phys. **2015** 033D02 (2015).
- [12] H. Ekawa *et al.*, Prog. Theor. Exp. Phys. **2019**, 021D02 (2019).
- [13] S. H. Hayakawa *et al.*, Phys. Rev. Lett. **126**, 062501 (2021).
- [14] I. J. Goodfellow *et al.*, Generative Adversarial Nets arXiv:1406.2661v1 (2014).
- [15] A. Ichikawa *et al.*, Phys. Lett. B **500**, 37 (2001).
- [16] M. Yamaguchi, K. Tominaga, Y. Yamamoto, and T. Ueda, Prog. Theor Phys. **105**, 627 (2001).
- [17] T. T. Sun, E. Hiyama, H. Sagawa, H. J. Schulze, and J. Meng, Phys. Rev. C **94**, 064319 (2016).
- [18] R. Shyam and K. Tsushima, arXiv:1901.06090 (2019).
- [19] M. Yoshimoto *et al.*, Prog. Theor. Exp. Phys., **2021**, 073D02 (2021)
- [20] The ALICE Collaboration, Phys. Rev. Lett. **123**, 112002 (2019).
- [21] T. Fukuda *et al.*, Phys. Rev. C **58**, 1306 (1998) .

- 
- [22] P. Khaustov *et al.*, Phys. Rev. C **61**, 054603 (2000).
- [23] T. Motoba, S. Sugimoto, Nucl. Phys. A **835**, 223 – 230 (2010).
- [24] K. Tominaga *et al.*, Nucl. Phys. A **642** 483 (1998).
- [25] M. M. Nagels, T. A. Rijken, and J. J. de Swart, Phys. Rev. D **15**, 2547 (1977)
- [26] M. M. Nagels, T. A. Rijken, and J. J. de Swart, Phys. Rev. D **20**, 1633 (1979).
- [27] Th. A. Rijken and Y. Yamamoto, Phys. Rev. C **73**, 044008 (2006).
- [28] Y. Yamamoto *et al.*, Nucl. Phys. A **835**, 350-353 (2010).
- [29] T. Nagae *et al.*, Proposal for the next E05 run with the S-2S spectrometer, Proposals for Nuclear and Particle Physics Experiments at J-PARC (2018).
- [30] K. Agari *et al.*, Prog. Theor. Exp. Phys. **2012**, 02B009 (2012).
- [31] R. Honda, PhD thesis, Study of  $\Sigma N$  interaction in nuclei using meson beams, Graduate School of Science, Department of Physics, Tohoku University (2014).
- [32] Y. Matsumoto, Master's thesis, Tohoku University (2013).
- [33] M. Fujita, PhD thesis, Experimental study for spectroscopy of  $\Xi^-$ -atomic X rays, Graduate School of Science, Department of Physics, Tohoku University (2019).
- [34] H. Ekawa, PhD thesis, Observation of a Double- $\Lambda$  Hypernucleus  $^{\Lambda\Lambda}\text{Be}$  with Hybrid Emulsion Method at J-PARC, Graduate School of Science, Department of Physics, Kyoto University (2019).
- [35] H. Fujioka *et al.*, Decay Pion Spectroscopy of  $^5_{\Lambda\Lambda}\text{H}$  Produced by  $\Xi$ -hypernuclear Decay ,Proposals for Nuclear and Particle Physics Experiments at J-PARC (2018).
- [36] Y. Ichikawa, K. Tanida *et al.*, High resolution spectroscopy of the “ $\Sigma N$  cusp” by using the  $d(K^-, \pi^-)$  reaction , Proposals for Nuclear and Particle Physics Experiments at J-PARC (2021).
- [37] T. Gogami *et al.*, New generation  $\Lambda$  hypernuclear spectroscopy with the  $(\pi^+, K^+)$  reaction by S-2S, Proposals for Nuclear and Particle Physics Experiments at J-PARC (2022).
- [38] R. Negishi, Master's thesis, Tokyo Institute of Technology (2023).
- [39] T. Gogami *et al.*, Nuclear Instruments and Methods in PRA **817**, 70-84 (2016).
- [40] A. Koshikawa, Master's thesis, Kyoto University (2017).
- [41] T. K. Harada, Master's thesis, Kyoto University (2020).
- [42] T. Gogami *et al.*, Nuclear Inst. and Method in Physics Research, A **900** 69-83 (2018).
- [43] T. Ohashi, Master's thesis, Kyoto University (2021).

- 
- [44] E. Iwai *et al.*, Proc. of PASJ **2021** WEOB02 (2021).
- [45] The ATLAS Collaboration, Search for neutral long-lived particles in  $pp$  collisions at  $\sqrt{s} = 13$  TeV that decay into displaced hadronic jets in the ATLAS calorimeter, JHEP06 005 (2022).
- [46] The ATLAS Collaboration, Search for non-resonant pair production of Higgs bosons in the  $b\bar{b}b\bar{b}$  final state in  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector, ATLAS-CONF-2022-035 (2022).
- [47] Y. Horii *et al.*, Fast Muon Tracking with Machine Learning Implemented in FPGA. arXiv:2202.04976v1 (2022).
- [48] N. Srivastava *et al.*, Dropout: A Simple Way to Prevent Neural Networks from Overfitting, Journal of Machine Learning Research **15** 1929-1958 (2014).
- [49] S. Ioffe and C. Szegedy, Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. arXiv:1502.03167 (2015).
- [50] T. Nanamura, Master's thesis, Kyoto University (2016).
- [51] M. Abadi *et al.* TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, Software available from tensorflow.org. (2015).
- [52] F. Chollet *et al.* Keras. <https://keras.io>, (2015).
- [53] P. Ramachandran *et al.*, arXiv:1710.05941v2 (2017).
- [54] Y. Nesterov, A method of solving a convex programming problem with convergence rate  $\mathcal{O}(1/k^2)$ , <https://goo.gl/V011vD>/(1983).
- [55] J. Duchi *et al.*, Journal of Machine Learning Research **12** 2121-2159 (2011).
- [56] G. Hinton *et al.*, [http://www.cs.toronto.edu/~tijmen/csc321/slides/lecture\\_slides\\_lec6.pdf](http://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf) (2012).
- [57] M. D. Zeiler, ADADELTA: An Adaptive Learning Rate Method. arXiv:1212.5701v1 (2012).
- [58] D. P. Kingma and J. Lei Ba, ADAM: A Method For Stochastic Optimization. arXiv:1412.6980 (2014).
- [59] T. Dozat, Incorporating Nesterov Momentum into Adam (2015).
- [60] S. Kanatsuki, Master's thesis, Kyoto University (2013).