

2020 年度 修士論文

Ξ ハイパー核分光の為の
機械学習を用いた運動量解析手法



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

大橋 翼

概要

通常原子核を構成する核力については、現在に至るまでに高統計で系統的な数多くの核子散乱実験データが存在し、中間子交換モデルを基にした現実的な核力モデルが構築されてきている。しかし、これをストレンジネス自由度 (S) を持ったハイペロンと核子との相互作用に拡張するのは単純ではない。ハイペロンが短寿命であるために核子との散乱データを取得するのは困難であり、これまで $S = -1$ の系については、 Λ , Σ ハイパー核の構造研究により相互作用の実験情報を引き出し、中間子交換モデルをフレーバー $SU(3)$ に拡張することで理解されてきた。 $S = -2$ の系では、実験情報はさらに不足しており、 Ξ 粒子と核子の相互作用は未解明である。現在準備が進んでいる J-PARC E70 実験では、 $^{12}\text{C}(K^-, K^+)$ 反応により $^{12}_{\Xi}\text{Be}$ の欠損質量スペクトルを 2 MeV(FWHM) の高分解能で測定することを目標としている。

E70 に先行して行われた J-PARC E05 実験では、素過程 $p(K^-, K^+)\Xi^-$ の欠損質量スペクトルが 5.4 MeV(FWHM) の分解能で得られた。これはこれまでで世界一のエネルギー分解能である。この測定では散乱運動量が散乱角度などの観測量と相関を持ち、運動量を補正する必要があることがわかった。現在、この補正手法は解析者に依存しており多項式による補正のように現象論的に行われている。本研究では、機械学習の導入により、解析者に依存しない広範の効果を考慮した解析を検討した。このような運動量解析に機械学習を適用した例は乏しく、手法としての適用可能性から調べる必要がある。

本研究では、 CH_2 標的からの $K^-p \rightarrow \pi^+\Sigma^-$, $K^-p \rightarrow K^+\Xi^-$ 反応を使って生成されるハイペロン (Σ^- と Ξ^-) の質量を欠損質量として測定し、これを PDG で知られている質量と比較する。損失関数や隠れ総数などのパラメータチューンを行った上で、学習データの選び方の問題、学習データの統計量の問題、学習データの範囲などについて性能評価を行う。機械学習を適用した運動量解析を新たに導入し、E05 実験のデータに対して適用する。 $K^-p \rightarrow \pi^+\Sigma^-$, $K^-p \rightarrow K^+\Xi^-$ 反応を使って、欠損質量の解析値と PDG value との平均二乗誤差や平均絶対誤差で与えられる損失関数が最小値をとるように学習させることで欠損質量スペクトルを得る。この手法を様々な条件下で行い、学習条件に対する本手法の依存性を評価する。

目次

第 1 章	はじめに	1
1.1	Ξ ハイパー核研究の背景	1
1.1.1	ストレンジネス核物理	1
1.1.2	$S = -2$ の物理	5
1.1.3	$^{12}\text{C}(K^-, K^+)$ 反応を用いた $^{12}_{\Xi}\text{Be}$ 分光実験の理論的要請	8
1.2	J-PARC E05/E70 実験	10
1.2.1	J-PARC	10
1.2.2	J-PARC E05 実験	13
1.2.3	J-PARC E70 実験	17
1.3	E05 実験における運動量解析	19
1.3.1	入射運動量	19
1.3.2	散乱運動量	21
1.3.3	運動量補正	22
1.3.4	従来手法の問題点	25
1.4	研究目的	25
第 2 章	機械学習	27
2.1	教師あり学習	27
2.1.1	クラス分類問題	28
2.1.2	回帰問題	28
2.1.3	損失関数	29
2.1.4	K-分割交差検証	29
2.2	ニューラルネットワーク	31
2.2.1	パーセプトロン	31

2.2.2	ニューラルネットワーク	31
2.2.3	活性化関数	32
2.2.4	誤差逆伝播法	34
2.2.5	ソフトウェア	35
2.3	勾配ブースティング決定木	35
2.3.1	決定木	35
2.3.2	勾配ブースティング決定木	36
第3章	機械学習を用いた運動量解析の最適化	37
3.1	運動量最適化の為の事象選別	37
3.2	機械学習と運動量最適化の対応関係	39
3.3	モデルの実装	40
3.3.1	従来解析フロー	40
3.3.2	モデル構成	42
3.3.3	学習手法	43
3.4	学習の進み方	44
3.4.1	early stopping	44
3.4.2	交差検証	45
3.5	性能評価	46
3.5.1	学習データの種類に対する依存性	47
3.5.2	学習データの統計に対する依存性	51
3.5.3	M_{Ξ^-} , M_{Σ^-} カット依存性	54
3.5.4	隠れ層数依存性	57
3.5.5	損失関数依存性	58
3.5.6	1変数 p_S チューニングのまとめ	60
3.5.7	チューニング変数依存性	61
第4章	まとめと今後の展望	63
付録A	学習時データ詳細	67
参考文献		77

表目次

1.1	K1.8 ビームラインのデザインパラメータ	12
1.2	E05 実験で取得されたデータの一覧	14
1.3	E05/E70 実験におけるスペクトロメータ分解能	16
1.4	$^{12}\text{C}(K^-, K^+)_{\Xi}^{12}\text{Be}$ での欠損質量分解能と各項の寄与	17
1.5	2つのスペクトロメータ S-2S と SKS の比較	18
3.1	性能評価に使用したデータ	38
3.2	位置や角度座標に付随する記号と検出器の対応	41
3.3	従来解析の性能	46
3.4	調査する学習条件	46
3.5	本節で基本的に使用するニューラルネット (model1)	47
3.6	p_S をチューニングした際の学習データの種類依存性	48
3.7	欠損質量 $3\sigma_{cut}$ の下での各 fold での学習データ数	51
3.8	p_S チューニングの学習データ統計依存性	53
3.9	p_S をチューニングした際の M_{Ξ^-} , M_{Σ^-} カット依存性	54
3.10	p_S をチューニングした際の隠れ層数依存性	57
3.11	p_S をチューニングした際の損失関数依存性	58
3.12	1変数 p_S のチューニングにおいて、依存性の調査より定まった学習条件	60
3.13	p_S をチューニング, 損失関数 MAE の場合の隠れ層数依存性	61
3.14	チューニング変数依存性	62
A.1	小節 3.5.2 で検討したデータ数の詳細	68
A.2	小節 3.5.3 で検討したデータ数の詳細	73

目次

1.1	バリオン 8 重項とバリオン 10 重項	2
1.2	格子 QCD 計算によるバリオン間相互作用ポテンシャル	3
1.3	$\Lambda 1$ 粒子状態の質量数依存性	4
1.5	長良事象	6
1.6	先行実験で得られた $^{12}\text{C}(K^-, K^+)$ 欠損質量スペクトル	8
1.7	$^{12}\text{C}(K^-, K^+)$ 反応より生成される $^{12}_{\Xi}\text{Be}$ の励起エネルギースペクトル	9
1.8	ESC08a を仮定した際, 実験により得られる予想励起エネルギースペクトル	10
1.9	J-PARC の加速器及び周辺施設	11
1.10	ハドロン実験施設の全景	12
1.11	J-PARC E05 実験のセットアップ	13
1.12	素過程 $p(K^-, K^+)\Xi^-$ の欠損質量スペクトル	15
1.13	素過程 $p(K^-, K^+)\Xi^-$ における散乱断面積の入射運動量依存性	15
1.14	S-2S スペクトロメータとその周辺検出器	18
1.15	S-2S スペクトロメータの運動量アクセプタンス	18
1.16	AFT の概念図	19
1.17	BT run における運動量の差	23
1.18	入射運動量 1.8 GeV/c における $p(K^-, K^+)$ 散乱運動量の角度依存性	24
2.1	クラス分類の概念図	28
2.2	回帰分析の概念図	28
2.3	K-分割交差検証	30
2.4	パーセプトロン	31
2.5	ニューラルネットワーク	32
2.6	決定木	35

2.7	勾配ブースティング決定木	36
3.1	入射 K^+ 運動量 $1.8 \text{ GeV}/c$ のもとでの ,素過程 $p(K^-, K^+) \Xi^-, p(K^-, \pi^+) \Sigma^-$ における散乱角と散乱運動量の関係	38
3.2	検証に使用する素過程の欠損質量スペクトル	39
3.3	M_{Ξ^-}, M_{Σ^-} カットを $1, 2, 3, 4\sigma$ 範囲とした時の学習データ	39
3.4	従来手法による欠損質量計算フロー	41
3.5	ローカルトラックの概念図	42
3.6	p_S のみを機械学習に置換した欠損質量計算フロー	43
3.7	損失関数の推移と early stopping の例	45
3.8	欠損質量 3σ cut 下の p_S^{pre}	48
3.9	Ξ^- データのみを学習に用いた結果	49
3.10	Σ^- データのみを学習に用いた結果	49
3.11	Ξ^- データと Σ^- データの両方を学習に用いた結果	50
3.12	p_S チューニングの学習データ統計依存性	52
3.13	p_S チューニングの inner train 数依存性	53
3.14	p_S チューニングの M_{Ξ^-}, M_{Σ^-} カット依存性	55
3.18	p_S チューニングの隠れ層数依存性	58
3.19	p_S チューニングの損失関数依存性	59
3.20	MAE を使用した場合の欠損質量	60
3.21	p_S チューニングの隠れ層数依存性	61
A.1	小節 3.5.2 で学習データ 80% を用いた際の欠損質量スペクトル	69
A.2	小節 3.5.2 で学習データ 60% を用いた際の欠損質量スペクトル	69
A.3	小節 3.5.2 で学習データ 40% を用いた際の欠損質量スペクトル	70
A.4	小節 3.5.2 で学習データ 20% を用いた際の欠損質量スペクトル	70
A.5	小節 3.5.2 で学習データ 16% を用いた際の欠損質量スペクトル	70
A.6	小節 3.5.2 で学習データ 12% を用いた際の欠損質量スペクトル	71
A.7	小節 3.5.2 で学習データ 8% を用いた際の欠損質量スペクトル	71
A.8	小節 3.5.2 で学習データ 4% を用いた際の欠損質量スペクトル	72
A.9	小節 3.5.3 で欠損質量カット 1σ の場合の欠損質量スペクトル	74
A.10	小節 3.5.3 で欠損質量カット 1.5σ の場合の欠損質量スペクトル	74
A.11	小節 3.5.3 で欠損質量カット 2σ の場合の欠損質量スペクトル	75

A.12	小節 3.5.3 で欠損質量カット 2.5σ の場合の欠損質量スペクトル	75
A.13	小節 3.5.3 で欠損質量カット 3.5σ の場合の欠損質量スペクトル	76
A.14	小節 3.5.3 で欠損質量カット 4σ の場合の欠損質量スペクトル	76

第 1 章

はじめに

1.1 Ⅲ ハイパー核研究の背景

1.1.1 ストレンジネス核物理

陽子 (p) と中性子 (n) からなる通常原子核を構成する核力については、現在に至るまでに数多くの p - p , p - n 散乱実験が行われ、得られた $\frac{d\sigma}{d\Omega}$ や偏極量を元に、中間子交換モデルを基にした現実的な核力モデルが構築されてきている。しかし、これをストレンジネス自由度をもったハイペロンと核子との相互作用に拡張するのは単純ではない。特に短距離部分の振る舞いには曖昧さが大きい。陽子と中性子はそれぞれアップクォーク (u , $m_u = 2.2_{-0.4}^{+0.5}$ MeV) とダウンクォーク (d , $m_d = 4.7_{-0.3}^{+0.5}$ MeV) から構成されているが、その他のクォークも含んだ系も考えることができる。そのような系についての情報を通じて、核力の拡張としてのバリオン間力の理解を深めることができる。中でも特にストレンジクォーク (s , $m_s = 95_{-3}^{+9}$ MeV) については、 u , d クォークと質量が比較的近く、 $SU(3)_f$ の対称性の下で統一的に扱うことができる。以上 3 つのクォークから 3 体系であるバリオンを構成すると、 $3 \otimes 3 \otimes 3 = 10 \oplus 8 \oplus 8 \oplus 1$ より、10 重項と 8 重項に分類される (図 1.1)。この中でストレンジクォークを含むバリオンをハイペロン (Y) と呼び、このハイペロンが、原子核に束縛した系をハイパー核と呼ぶ。ハイパー核が束縛状態として存在するかは、ハイペロンと核子間の相互作用 (YN 相互作用) の強さによる。

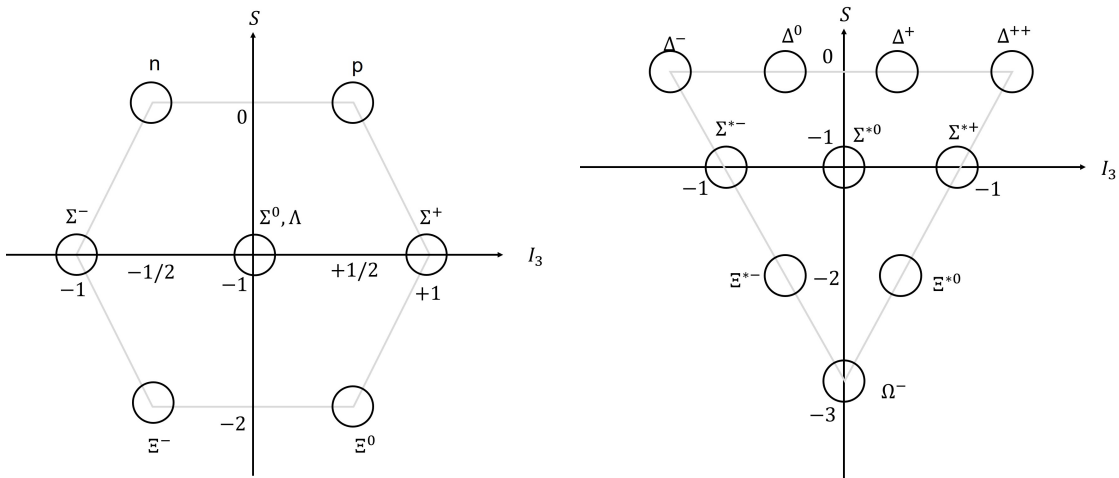


図 1.1: スピン 1/2 であるバリオン 8 重項 (左図) とスピン 3/2 であるバリオン 10 重項 (右図). アイソスピン I の量子化軸を第三成分にとっている. S はストレンジネス量子数である.

バリオン 8 重項間相互作用を $SU(3)_f$ で分類すると, 次のようになる.

$$\mathbf{8} \otimes \mathbf{8} = \mathbf{27} \oplus \mathbf{8}_s \oplus \mathbf{1} \oplus \mathbf{10}^* \oplus \mathbf{10} \oplus \mathbf{8}_a$$

このうち, $\mathbf{27}$ 重項, $\mathbf{10}^*$ 重項以外はストレンジネスを含むことで初めて現れる. 3 フレーバー基底でのポテンシャルの形は, 図 1.2 に見られるように, 格子 QCD 計算で, $V^{(1)}$ が強い引力, $V^{(10^*)}$ と $V^{(27)}$ が引力 (NN 相互作用) であり, $V^{(10)}$ には斥力芯が見られるというような大きな違いが見られる [1]. 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.

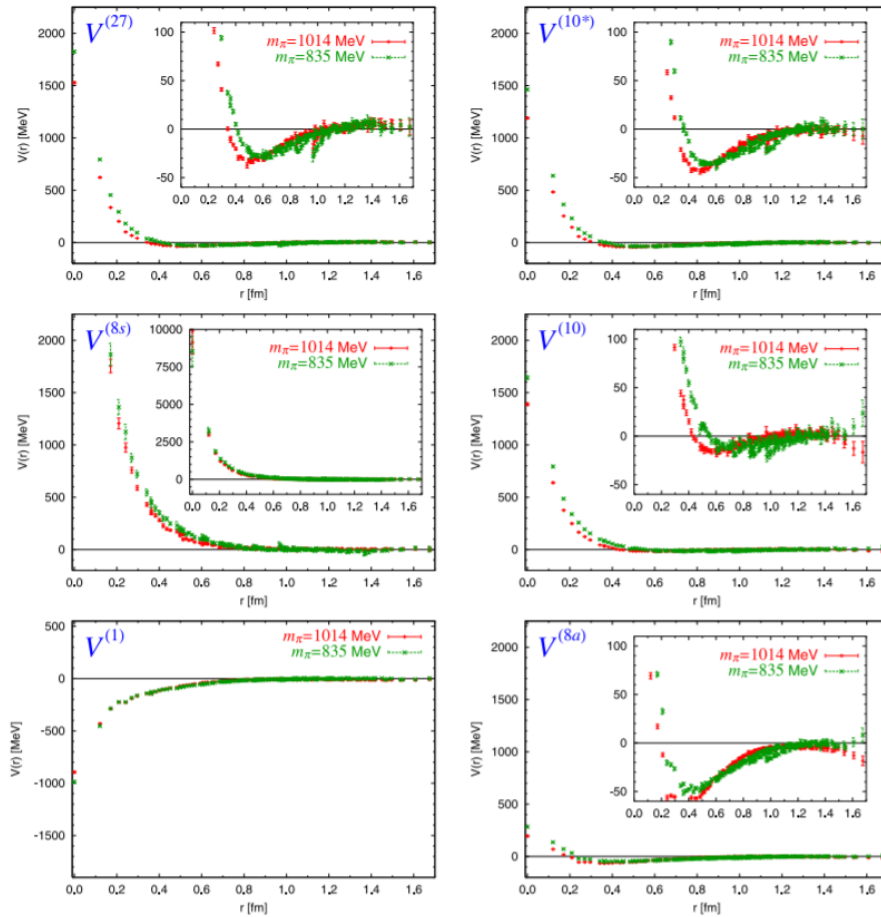


図 1.2: 格子 QCD 計算によるバリオン間相互作用ポテンシャル [1].

核力の場合同様，バリオン間相互作用の直接の理解には YN, YY 散乱データが有用である．しかしながら，ハイペロンの寿命が短い ($c\tau \sim$ 数 cm) 為に，S 波散乱が支配的となる低エネルギー領域での散乱データの取得は困難である．そこで，ハイペロンが原子核に束縛されたハイパー核の構造から，相互作用情報を引き出す手法が取られてきた．

ΛN 有効相互作用は

$$V_{\Lambda N}(r) = V_0(r) + V_\sigma(r)\vec{s}_N \cdot \vec{s}_\Lambda + V_\Lambda(r)\vec{l}_{\Lambda N} \cdot \vec{s}_\Lambda + V_N(r)\vec{l}_{\Lambda N} \cdot \vec{s}_N + V_T S_{12} \quad (1.1)$$

と書き表され，それぞれ，中心力，スピン-スピン相互作用， Λ スピン軌道相互作用，N スピン軌道相互作用，テンソル相互作用に対応する．これをもとに，p-殻のハイパー核の微細構造が理論的に計算されている．

これまで，KEK の 12 GeV シンクロトロン (KEK-PS) や米国ブルックヘブン研究所の Alternating Gradient Synchrotron (BNL-AGS)，J-PARC において π 中間子や K 中間子を

用いた実験が数多く行われてきた．また近年では，米国のジェファソン研究所 (JLab) やドイツのマインツ大学の MAMI-C 施設において，直流状の高エネルギー電子ビームを用いた $(e, e'K^+)$ 反応によるハイパー核分光も盛んになっている．

これらの実験により， $S = -1$ の系 (Λ, Σ ハイパー核) については多くの知見が得られてきた．例えば，ラムダハイパー核については，KEK の SKS スペクトロメータを用いた (π^+, K^+) 分光実験 [2][3] などにより， $A = 7 \sim 208$ に及ぶラムダハイパー核 1 粒子束縛エネルギーが測定されており (図 1.3)，結果として核物質中での Λ 粒子の中心力ポテンシャルの深さは 29 ± 1 MeV と見積もられた．これは，通常原子核のポテンシャルの深さ 50 MeV の約 60% の大きさである．また， Λ 粒子はスピン 1/2 を持っているので，スピン依存力が存在する．式 (1.1) で表される ΛN 相互作用のうち，スピン-スピン相互作用の大きさは軽いラムダハイパー核の詳細なレベル構造に反映される． ΛN 相互作用のスピン依存成分は，ゲルマニウム検出器群 Hyperball を用いて， ${}^7_\Lambda\text{Li}$ 等からの γ 線を 2keV のエネルギー分解能で測定し，その大きさが得られている [4][5]．中でも ${}^9_\Lambda\text{Be}$ (E2) γ 線の測定により，非常に小さなスピン軌道分離が観測された．

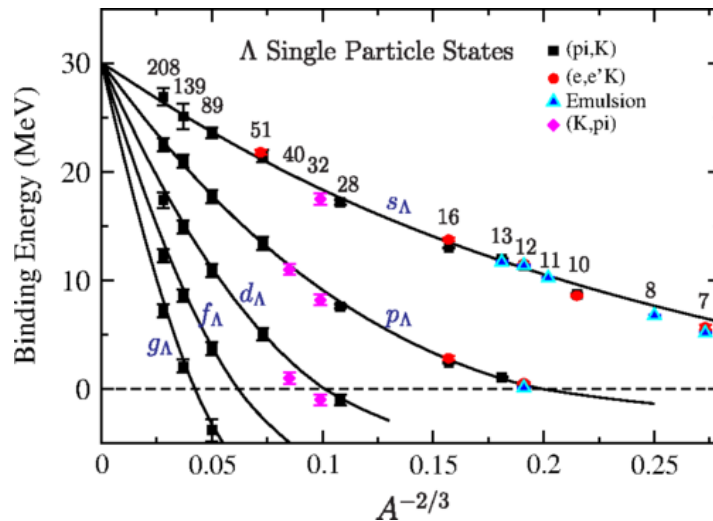


図 1.3: Λ 1 粒子状態の質量数依存性 [6].

また， ΣN 相互作用については，当初 Σ^- 原子の X 線分光により Σ ポテンシャルの実部は約 20 MeV，虚部は約 20 MeV 程度と見積もられ，結果， Σ ハイパー核の幅 (20-30 MeV 程度) は実部の深さと同程度であり，その束縛状態をピークとして観測することは困難だと考えられた．これは， $\Sigma N \rightarrow \Lambda N$ という強い相互作用による転換幅の存在のためである．KEK PS における ${}^4\text{He}(K_{stop}^-, \pi^-)$ 反応により ${}^4_\Sigma\text{He}$ の束縛状態が報告され，さらに

BNL AGS における高統計精度の実験によりこの存在が確認された [7] . この束縛状態は (K^-, π^+) のスペクトルには見られないことから , 強いアイソスピン依存性が明らかになった . また , KEK において得られた $\text{Si}(\pi^-, K^+)$ スペクトル [8] から , Σ 粒子のポテンシャルが斥力的で , 大きさは 30 MeV 程度であることがわかっている [9] .

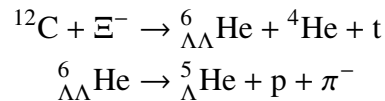
このように , バリオン 8 重項において $\Lambda N, \Sigma N$ 相互作用は明らかにされてきた .

1.1.2 $S = -2$ の物理

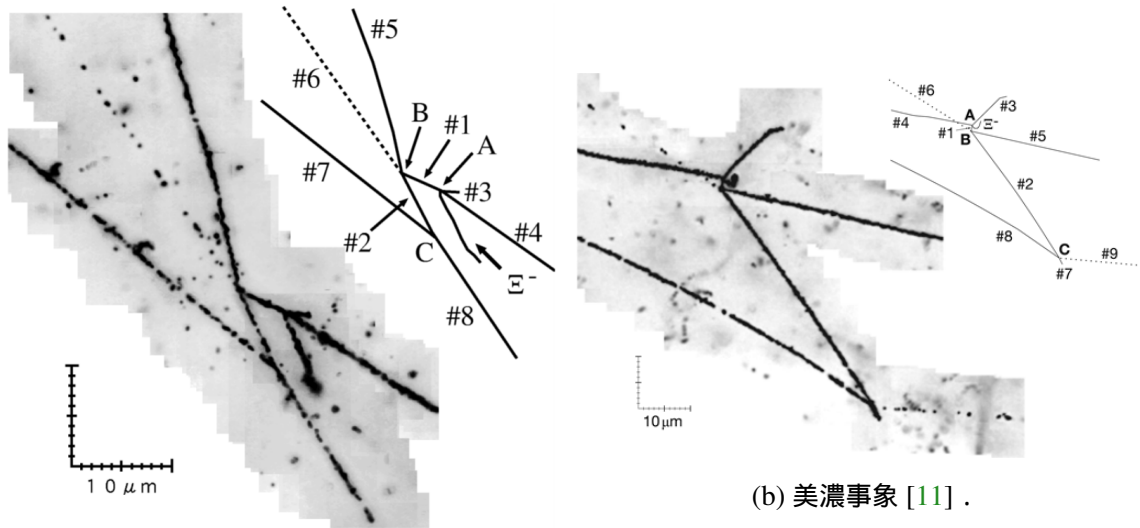
一方で , $S = -2$ の系である Ξ ハイパー核や $\Lambda\Lambda$ ハイパー核の生成については , 生成面積が小さい為 , 実験データが比較的限られている . $S = -2$ の系では , $S = -1$ の系では現れなかった YY 相互作用が現れ , これは , さらなるバリオン間力の拡張を考える上で重要な情報となる .

$S = -2$ の系の生成には , $K^-p \rightarrow K^+\Xi^-$ 反応により $S = -2$ の Ξ^- を核内に持ち込むことから始まる . この反応は 2 体反応なので効率よく $S = -2$ の生成をタグすることができる . 一方で , $S = -2$ の多体系の束縛状態の検出は容易ではない . そこで古くから利用されてきたのが原子核乾板による検出である . エマルジョンは $1\mu\text{m}$ オーダーでの飛跡の測定が可能であり , ハイパー核の基底状態からの弱い相互作用による崩壊パターンを観測できる . 2 度の弱崩壊を分離して観測できれば , $S = -2$ のハイパー核生成の裏付けとなる . 崩壊粒子を全て検出できれば , 基となったハイパー核の質量がわかる . この手法の欠点は , $K^-p \rightarrow K^+\Xi^-$ 反応の Ξ^- の反跳運動量が $500 \text{ MeV}/c$ と高いのでなかなか静止せず , その前に崩壊してしまうことにある .

これまで , このような原子核乾板を用いた実験により Ξ ハイパー核やダブル Λ ハイパー核についてのデータが得られてきた . KEK-PS E373 実験では , 2 重ラムダハイパー核の核種の同定と束縛エネルギーの測定に成功した . この事象は長良事象 (図 1.4a) と呼ばれ ,



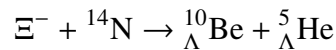
により ${}^6_{\Lambda\Lambda}\text{He}$ が生成されたことを示しており , 2 重ラムダハイパー核内の 2 個のラムダ粒子間の結合エネルギーの大きさ ($\Delta B_{\Lambda\Lambda} = B_{\Lambda\Lambda} - 2 \times B_{\Lambda}$) は $0.67 \pm 0.17 \text{ MeV}$ で弱い引力であることがわかった [10] . さらに近年 , J-PARC E07 実験において , ${}_{\Lambda\Lambda}\text{Be}$ が発見され , 美濃事象と呼ばれている (図 1.4b)[11] .



(a) 長良事象 [10].

(b) 美濃事象 [11].

また、KEK-PS E373 実験からは、木曾事象（図 1.5）



も発見されている [12]. この事象から Ξ^- と ${}^{14}\text{N}$ 間の束縛エネルギーが推定された結果、この吸収が $\Xi^- - {}^{14}\text{N}$ 原子状態からの吸収ではなく、強い相互作用による束縛状態からの吸収だと考えられており、 Ξ 粒子と原子核の束縛状態の存在を裏付ける結果となった。

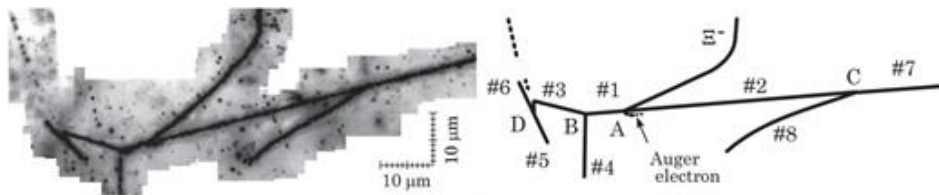


図 1.5: 長良事象 [12].

また、LHC で行われた ALICE 実験の p-Pb 衝突 ($\sqrt{s_{NN}} = 5.02$ TeV) データからは、p- Ξ^- の強い引力的相互作用が確認され、HAL-QCD で得られた p- Ξ^- 相互作用と矛盾なく計算結果をしていることがわかった [13]. HAL-QCD のポテンシャルでは $\Xi\text{N} - \Lambda\Lambda$ の結合が非常に弱く、 Ξ 粒子ポテンシャルの虚部が小さいことを示唆している。

主に行われてきた原子核乾板を用いた実験のもう 1 つの欠点は、弱崩壊を観測してハイパー核を同定しているため、基底状態の観測に留まることである。励起エネルギースペクトルを得るためには、直接反応が必要とされる。そこでもう 1 つの生成法として

(K^- , K^+) 直接反応法がある．これは (K^- , K^+) 反応により， Ξ ハイパー核を生成し，イベントごとに入射 K^- と散乱 K^+ の運動量を測定し，欠損質量 (ミッシングマス，Missing Mass) 法により Ξ ハイパー核のエネルギースペクトルを得る手法である．この手法での先行実験として，KEK-PS E224 実験 [14]，BNL-E885 実験 [15] がある．これらの実験では， $^{12}\text{C}(\text{K}^-, \text{K}^+)$ 反応を用いて， Ξ^- と ^{11}B の束縛系としての $^{12}_{\Xi}\text{Be}$ の探索が行われた ($\Xi^- ^{11}\text{B}_6$, $Z = 4$, $A = 12$ を $^{12}_{\Xi}\text{Be}$ と記す)．それぞれの実験で得られた Ξ ハイパー核のエネルギースペクトルを図 1.6 に示す．それぞれのエネルギー分解能 ΔM は，22 MeV，14 MeV(FWHM) であり，束縛領域における統計量は 3 イベント，42-67 イベントであった．この分解能と統計量の不足により， Ξ ハイパー核の束縛状態をピーク構造として観測するには至っていない．しかし， ΞN 相互作用のスピンのアイソスピン非依存の部分についてウッズ-サクソン型のポテンシャルを仮定し， Ξ の 1 体ポテンシャルの値を仮定して得られる理論計算のスペクトルと，実験のスペクトルを形状比較することにより， Ξ の 1 体ポテンシャルの深さ V_0^{Ξ} が見積もられた．その結果，KEK-PS E224 では $V_0^{\Xi} < 20$ MeV，BNL-E885 では $V_0^{\Xi} \simeq 14$ MeV とされた．また，BNL-E885 では $^{12}_{\Xi}\text{Be}$ の生成断面積が見積もられ，入射 K^- の運動量が 1.8 GeV/c の場合，実験室系散乱角 $\theta < 8^\circ$ で 89 ± 14 nb/sr， $\theta < 14^\circ$ で 42 ± 5 nb/sr という結果が示された．この手法でも Ξ^- の反跳運動量が大きい (~ 500 MeV/c) ことがわざわざいして，束縛する確率 (sticking probability) が小さい (1/100 程度) ことがあげられる．

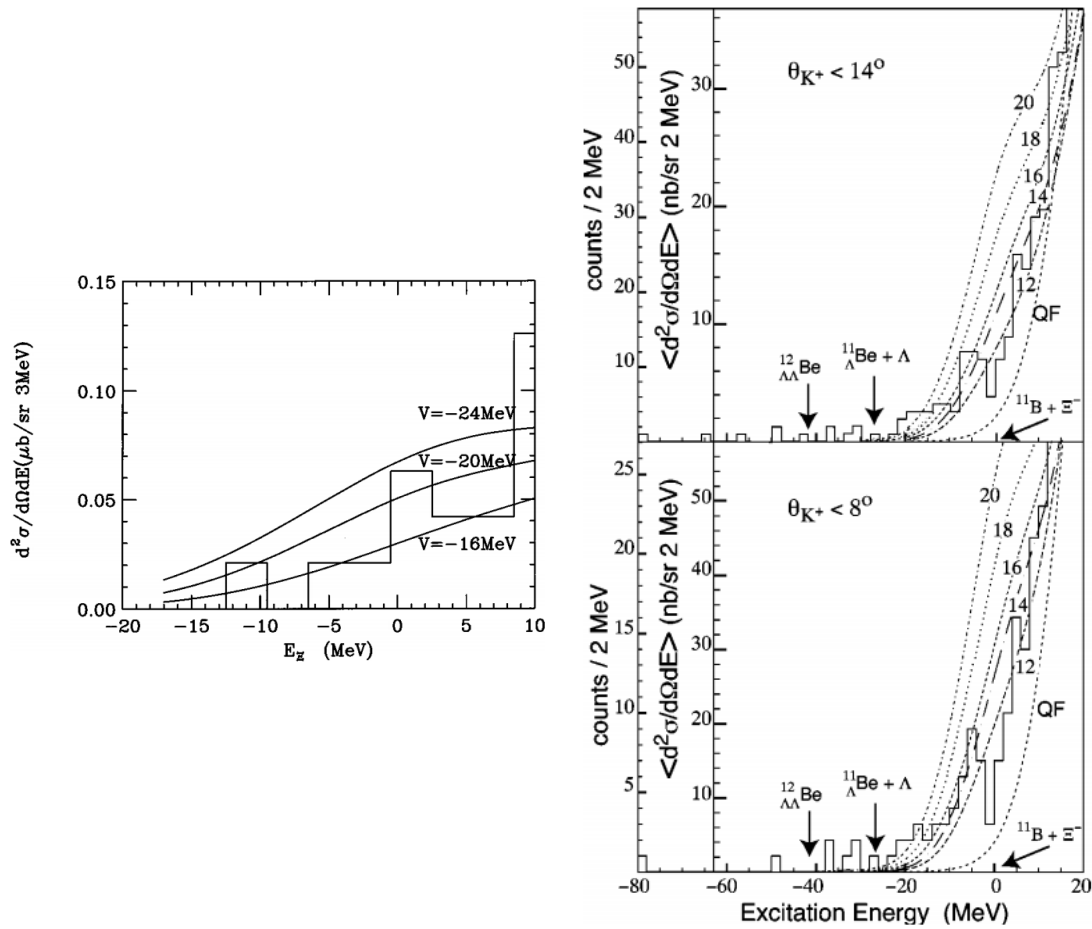


図 1.6: 先行実験で得られた $^{12}C(K^-, K^+)$ 欠損質量スペクトル．左が KEK-PS E224[14]，右が BNL-E885[15] で得られたものである．

1.1.3 $^{12}C(K^-, K^+)$ 反応を用いた $^{12}_{\Xi}Be$ 分光実験の理論的要請

小節 1.1.2 で述べたように， $S = -2$ の系における YN, YY 相互作用は $S = -1$ の系に比較して不確定な部分が未だ多く残っている． $^{12}C(K^-, K^+)^{12}_{\Xi}Be$ 反応では，束縛状態のピークが観測できれば，そのピーク位置より束縛エネルギーの大きさ（ポテンシャル実部の大きさ）が決定できる．またそのピーク幅より，ポテンシャル虚部の大きさが決まる．しかし，先行実験においては，エネルギー分解能と統計量の不足によりピーク構造を観測することができなかった．即ち，高分解能，高統計で上記のような分光実験を実施することにより，バリオン間相互作用に対して強い制限をかけることが期待される．

現在， $^{12}C(K^-, K^+)$ 反応において BNL E885 と矛盾しない $^{12}_{\Xi}Be$ が生成されるような 4 種

類のバリオン間相互作用モデル (NHC-D[16][17] Ehime[18][19] ESC04d[20] ESC08a[21]) が提案され, これを使ってコア原子核 ^{11}B の励起も考慮した $^{12}_{\Xi}\text{Be}$ 励起エネルギースペクトルの理論計算が存在している [22]. その結果を図 1.7 に示す. この計算では, BNL-E885 の計算を元に, 1_1^- 状態で $B_{\Xi} = 4.5$ MeV となるようにポテンシャルの深さを決めてある. 束縛エネルギーが 4.5 MeV 程度と予想されるので, $\Delta M \sim 1 - 3$ MeV が必要となるであろう.

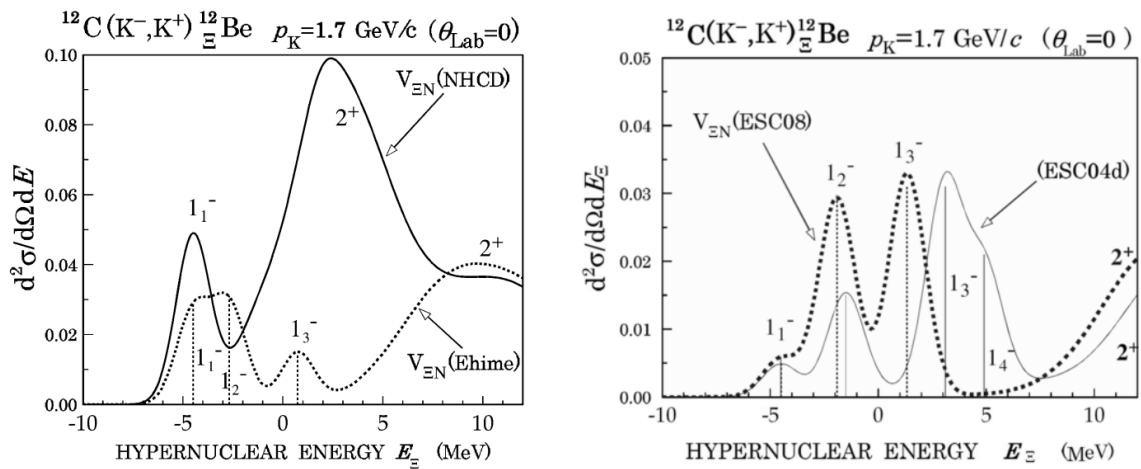


図 1.7: $^{12}\text{C}(K^-, K^+)$ 反応より生成される $^{12}_{\Xi}\text{Be}$ の励起エネルギースペクトル. 左: NHC-D, Ehime, 右: ESC04d, ESC08a の相互作用モデルにおける結果である.

また, スピン-スピン相互作用によるコア原子核の励起も考慮すると図 1.8 のように, $1_1^-, 1_2^-, 1_3^-$ といった構造が表れると予想され, 2 MeV 程度のエネルギー分解能が必要と考えられる.

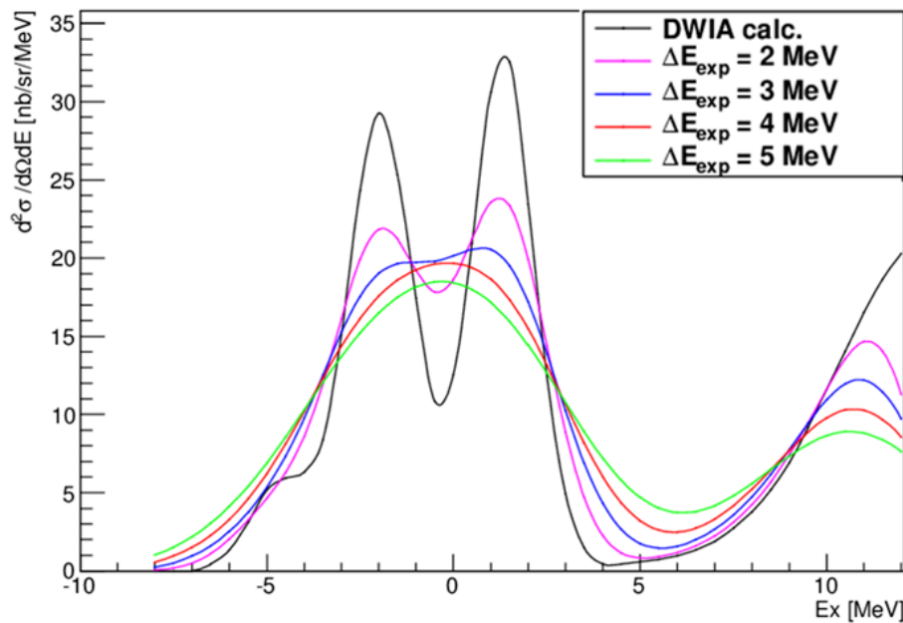


図 1.8: ESC08a を仮定した時，実験により得られる励起エネルギースペクトルを，実験の分解能ごとにシミュレートした図．

1.2 J-PARC E05/E70 実験

小節 1.1.3 で述べたように，高分解能かつ高統計で $^{12}\text{C}(K^-, K^+)$ 反応を用いた Ξ ハイパー核分光実験を行うことにより，バリオン間相互作用に制限を与えることが期待される．束縛状態のピークを観測する目的で 2015 年 10 月から 11 月にかけて J-PARC E05 実験が行われた．また，さらなる高エネルギー分解能下でデータ取得を行う J-PARC E70 実験が準備段階にある．本節ではこれらの実験の詳細について述べる．

1.2.1 J-PARC

大強度陽子加速器施設 Japan Proton Accelerator Research Complex(J-PARC) は，高エネルギー加速器研究機構 (KEK) と日本原子力研究開発機構 (JAEA) により建設された，共同利用実験施設である [23]．図 1.9 に，J-PARC の加速器及び周辺施設を示す．図のように，J-PARC は線形加速器 (LINAC)，3 GeV シンクロトロン (Rapid Cycling Synchrotron, RCS)，30 GeV シンクロトロン (Main Ring, MR) の 3 つの大強度陽子加速器を有している．イオン源から取り出した負水素イオン H^- ビームを初段の LINAC で 400 MeV まで

加速し，荷電交換ホイルを通過させて陽子ビームに荷電交換してから RCS で 3 GeV まで加速する．その後，物質・科学実験施設 (MLF) と MR に輸送される．MR に輸送された陽子は更に 30 GeV まで加速され，ニュートリノ実験施設とハドロン実験施設で取り出して利用されている．



図 1.9: J-PARC の加速器及び周辺施設 [24] .

ハドロン実験施設の全体図を図 1.10 に示す．ハドロン実験施設では 5.52 秒周期中に約 2.2 秒かけて MR から陽子ビームを取り出す．取り出された陽子ビームは一次標的 (T1 Target) に照射され，二次ビームとして大量の中間子が生成される．我々はこの二次粒子を利用して実験を行う．二次粒子は電磁石系からなる磁気光学系や静電セパレータを用いて運動量選別や粒子識別され，各ビームラインで中間子ビームとして利用できる．

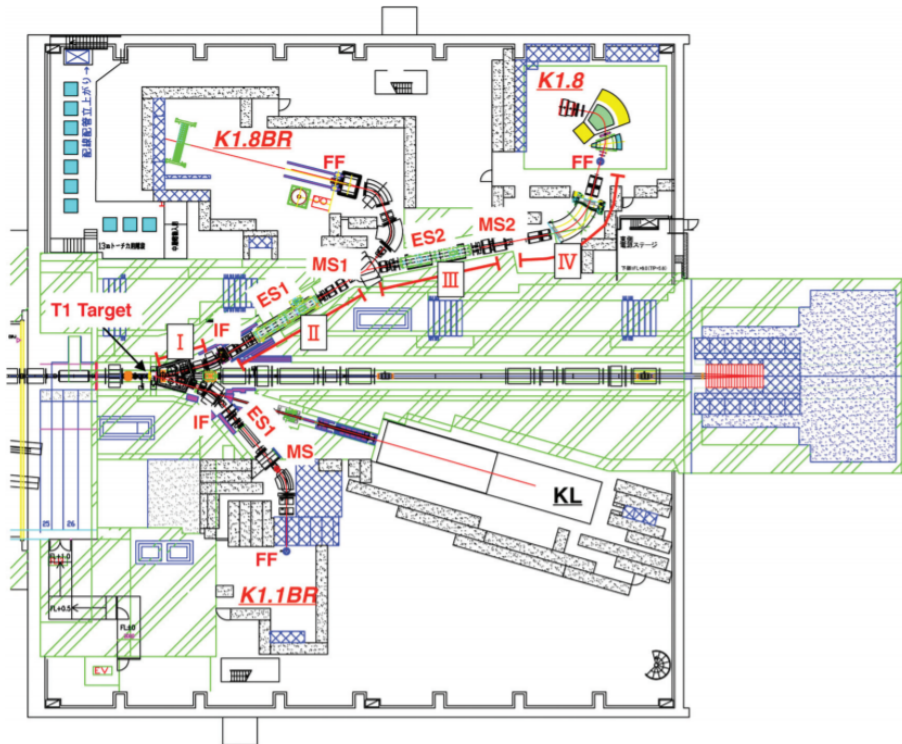


図 1.10: ハドロン実験施設の全景 [25] .

ハドロン実験施設内には数本のビームラインが存在しており，そのうち K1.8 ビームラインでは (K^- , K^+) 反応を用いた Ξ ハイパー核分光実験を主な目標としている．K1.8 ビームラインのデザインパラメータを表 1.1 に示す．1.8 GeV/c の K^- ビームに対して最適化されており，T1 で生成された二次粒子を，二段の静電セパレータにより粒子識別を行うことで K^-/π^- 比 ~ 1 を実現している． K^- ビームは QDQD からなる K1.8 ビームラインスペクトロメータにより運動量が測定される．ユーザーはその下流に標的や検出器を設置し，実験を行う．

表 1.1: K1.8 ビームラインのデザインパラメータ．MR が 30 GeV/c, $9\mu\text{A}$ で運転される場合の値を示す．

最大運動量	2.0 GeV/c
全長	45.97 m
静電セパレータ	750 kV/10 cm, 6 m \times 2
K^- 強度 [/spill] at 1.8 GeV/c	1.4×10^6
K^-/π^-	6.9

1.2.2 J-PARC E05 実験

$^{12}\text{C}(K^-, K^+)$ 反応を用いた高統計・高エネルギー分解能での三ハイパー核分光実験を目的とした J-PARC E05 実験は K1.8 ビームラインで実施された。E05 実験のセットアップを図 1.11 に示す。

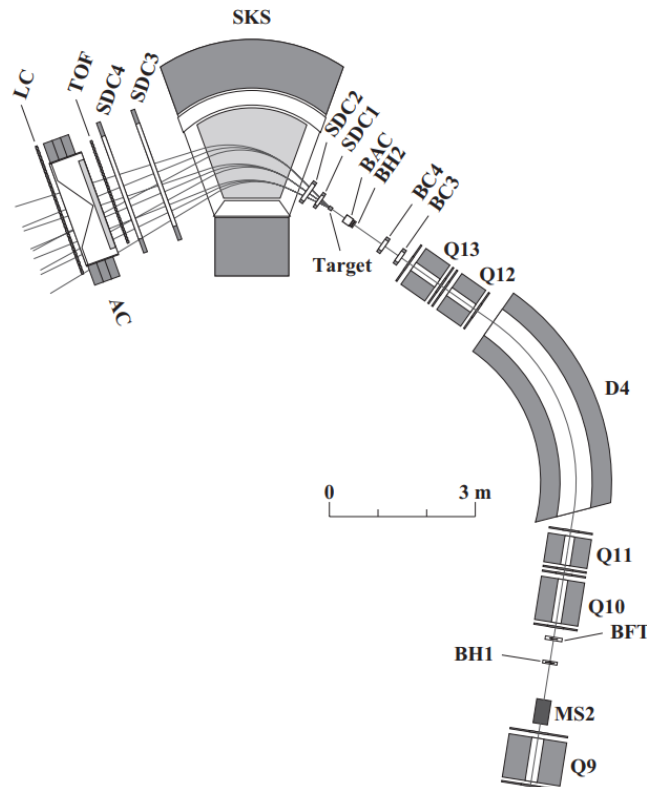


図 1.11: J-PARC E05 実験のセットアップ

図 1.11 下方から、中心運動量 $1.8 \text{ GeV}/c$ の K^- ビームが入射される。入射 K^- は 2 つのプラスチックシンチレータホドスコープ (BH1, BH2) 間の飛行時間差で識別され、 π^- はエアロジェルチェレンコフカウンター (BAC) を用いて、オンライントリガーレベルで抑制される。 K^- の運動量は飛跡検出器 (BFT, BC3, BC4) を用いてイベントごとに解析される。運動量解析の手法は 1.3 で述べる。K1.8 ビームラインスペクトロメータは 4 台の四重極磁石 (Q10-Q13) と 1 台の双極電磁石 (D4) で構成され、 $\Delta p/p = 1 \times 10^{-3}$ (FWHM) の分解能を持つ。典型的な K^- のビーム強度は $6 \times 10^5 K^-$ per spill (5.52 秒周期, 2.2 秒)

であった．また， K^-/π^- 比は約 0.8 である．

データは， 9.36-g/cm^2 の炭素標的に対して約 10 日間， 9.54-g/cm^2 の CH_2 に対して約 2 日間取得された． CH_2 標的は $1.5\text{-}1.9\text{ GeV}/c$ の 5 点の異なる入射運動量に対して，素過程 $p(K^-, K^+)\Xi^-$ 反応を測定するために使用された．表 1.2 に，E05 で取得されたデータの一覧を示す．表中 Beam Through については，1.3 で述べる．

表 1.2: E05 実験で取得されたデータの一覧.

Title	Target	Beam momentum (GeV/c)	Number of K^- 's
ハイペロン生成 Ξ^-, Σ^-	CH_2	1.5	3.17×10^9
		1.6	3.23×10^9
		1.7	3.05×10^9
		1.8	1.10×10^{10}
		1.9	1.34×10^9
Beam through	Empty	1.2, 1.35, 1.5, 1.8 (K^-) 1.1, 1.2, 1.35, 1.5, 1.8, 2.0 (K^+)	-
$^{12}_{\Xi}\text{Be}$ production	Graphite	1.8	8.43×10^{10}

$^{12}\text{C}(K^-, K^+)$ 反応により散乱した K^+ の運動量は，4 台のドリフトチェンバー (SDC1-4) と SKS スペクトロメータを用いて解析される．SKS の中心運動量は $1.3\text{ GeV}/c$ に設定され，約 110 msr の広い立体角を覆っている．SKS スペクトロメータの運動量分解能は $\Delta p/p = 3 \times 10^{-3}$ (FWHM) であり，これは過去に行われた (K^-, K^+) 反応に用いられたスペクトロメータに比較して大幅に改善されている．

SKS で検出された粒子はプラスチックシンチレーションカウンター (図 1.11 中"TOF") で測定された飛行時間と，飛跡検出器により得られた飛行距離，運動量を用いて識別される．

入射運動量 p_B ，散乱運動量 p_S が求めれば，欠損質量 (Missing Mass) M_X は以下のように計算される．

$$M_X = \sqrt{(E_B + m_A - E_S)^2 - p_B^2 - p_S^2 + 2p_B p_S \cos \theta} \quad (1.2)$$

ここで， E, p, m は各粒子のエネルギー，運動量，質量であり，下付き文字 B, S, A はそれぞれ入射粒子，散乱粒子，標的粒子を表す．さらに， θ は実験室系での散乱角である．

図 1.12 に入射中心運動量 $1.8\text{ GeV}/c$ ， CH_2 標的で取得された $p(K^-, K^+)\Xi^-$ 反応に対する欠損質量スペクトルを示す． $p(K^-, K^+)\Xi^-$ の質量分解能は 5.4 MeV (FWHM) となり，

これは BNL E885 において CH_2 標的で取得された際の分解能である 10 MeV に比較して 2 倍ほど良い結果となった。また、図 1.13 に、異なる入射運動量 (1.5, 1.6, 1.7, 1.8, 1.9 GeV/c) について得られた $p(K^-, K^+)\Xi^-$ の散乱断面積を示す。この結果から、 Ξ^- の生成断面積は 1.8 GeV/c でピークを持つことが従来の泡箱による測定データに比較して約 20 倍の統計で確認された。これより、分光実験も 1.8 GeV/c で実施することが最適だと結論づけられる。

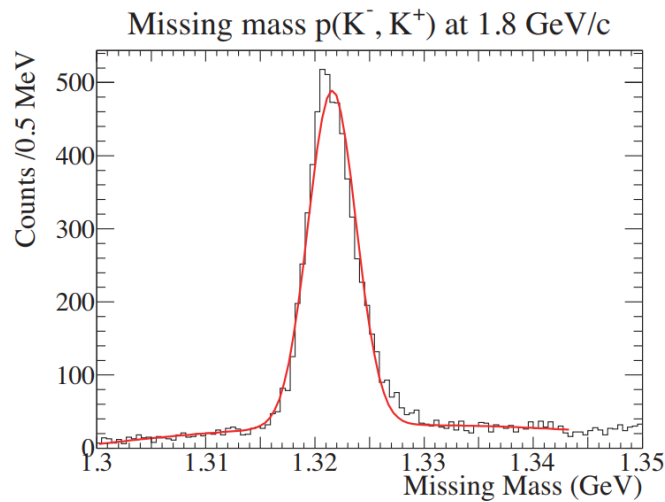


図 1.12: 素過程 $p(K^-, K^+)\Xi^-$ の欠損質量スペクトル。ピーク下に見られる連続的なバックグラウンドは ^{12}C 由来の準弾性 Ξ 生成によるもの。

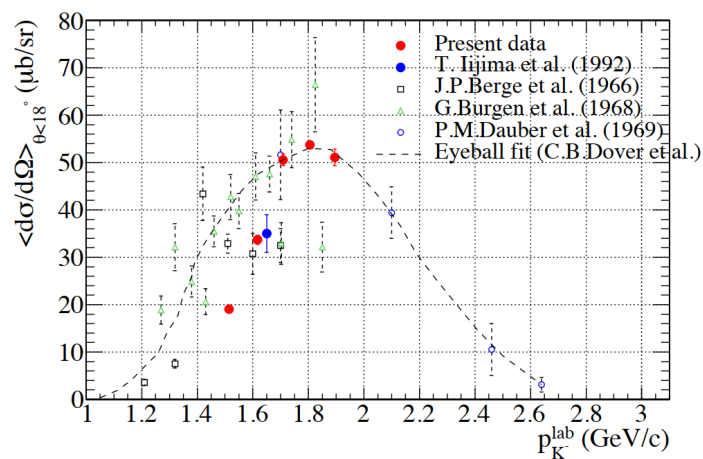


図 1.13: 素過程 $p(K^-, K^+)\Xi^-$ における散乱断面積の入射運動量依存性。赤点が E05 実験のデータである。

エネルギー分解能に対する各項の寄与

式 (1.2) において, p_B, p_S, θ について偏微分すると,

$$\frac{\partial M_X}{\partial p_B} = \frac{1}{M_X} \left[\frac{p_B}{E_B} (m_A - E_S) + p_S \cos \theta \right] \quad (1.3)$$

$$\frac{\partial M_X}{\partial p_S} = \frac{1}{M_X} \left[-\frac{p_S}{E_S} (E_B + m_A) + p_B \cos \theta \right] \quad (1.4)$$

$$\frac{\partial M_X}{\partial \theta} = \frac{p_B p_S \sin \theta}{M_X} \quad (1.5)$$

である．欠損質量の分解能に対しての各変数の寄与は, 式 (1.3) (1.5) から計算でき,

$$\Delta M_X^2 = \left(\frac{\partial M_X}{\partial p_B} \right)^2 \Delta p_B^2 + \left(\frac{\partial M_X}{\partial p_S} \right)^2 \Delta p_S^2 + \left(\frac{\partial M_X}{\partial \theta} \right)^2 \Delta \theta^2 + \Delta E_{straggle}^2 \quad (1.6)$$

$$= \Delta_B^2 + \Delta_S^2 + \Delta_\theta^2 + \Delta E_{straggle}^2 \quad (1.7)$$

となる．

$^{12}\text{C} (K^-, K^+) \Xi^{12}\text{Be}$ の運動学を考え, $p_B = 1.8$ [GeV/c], $\theta = 5^\circ$, 束縛閾値を仮定すると,

$$\frac{\partial M_X}{\partial p_B} = 0.93 \quad (1.8)$$

$$\frac{\partial M_X}{\partial p_S} = -0.91 \quad (1.9)$$

$$\frac{\partial M_X}{\partial \theta} = -0.019 \text{ [MeV/mrad]} \quad (1.10)$$

それぞれのスペクトロメータの運動量分解能が, 表 1.3 のようになっているとすると, 質量分解能は表 1.4 中 E05 のようになり, E05 実験では $\Delta_S, \Delta E_{straggle}$ が分解能の主たる要因となっている．ビームラインスペクトロメータは E05 と E70 では同じものを用いる．

表 1.3: E05/E70 実験におけるスペクトロメータ分解能．E05 は実験より得られた分解能, E70 の散乱側は S-2S スペクトロメータの設計値を用いている．尚, ビームラインスペクトロメータの設計値は 5×10^{-4} である．

$\Delta p/p$	ビームラインスペクトロメータ	散乱スペクトロメータ
E05	1×10^{-3} (測定値)	3×10^{-3} (SKS 測定値)
E70	1×10^{-3} (測定値)	5×10^{-4} (S-2S 設計値)

表 1.4: $^{12}\text{C}(K^-, K^+)_{\Xi}^{12}\text{Be}$ での欠損質量分解能と各項の寄与。E05, E70 のスペクトロメータ分解能は, 表 1.3 中の値を仮定した。単位は MeV/c^2 である。

Conditions	Δ_B	Δ_S	Δ_θ	$\Delta E_{straggle}$	ΔM
E05	1.67	3.74	0.04	3	5.1
E70	1.67	0.62	0.04	0.9	2

1.2.3 J-PARC E70 実験

表 1.4 のように, E05 実験では Δ_S と $\Delta E_{straggle}$ が質量分解能悪化の主要因であった。J-PARC E70 実験では, この 2 項を小さくし, 質量分解能を向上させる為に, 主に以下の 2 点が E05 実験から改善される。

S-2S スペクトロメータ

E70 実験では, 散乱粒子の運動量解析に新たに建設された Strangeness -2 Spectrometer(S-2S) を用いる。図 1.14 にその構成を示す。S-2S スペクトロメータは 2 つの四重極磁石(Q1,Q2) と 1 つの双極磁石(D1) から構成される。散乱された K^+ は 4 台のドリフトチェンバー(SDC) で飛跡検出され, 散乱運動量を, $\Delta p/p = 5 \times 10^{-4}$ (FWHM) の分解能で解析できる(設計値)。また, 運動量 $1.37 \text{ GeV}/c$ 付近に大きなアクセプタンスを持つ(図 1.15)。表 1.5 に, SKS スペクトロメータ(E05) との比較を示している。また, K^+ 粒子を π^+ や陽子と識別するために, 飛行時間を測定する TOF カウンター, π^+ 抑制のためのエアロジェルチェレンコフカウンター(AC, $n = 1.055$), 陽子抑制のための水チェレンコフカウンター(WC, $n = 1.33$) を用いる。

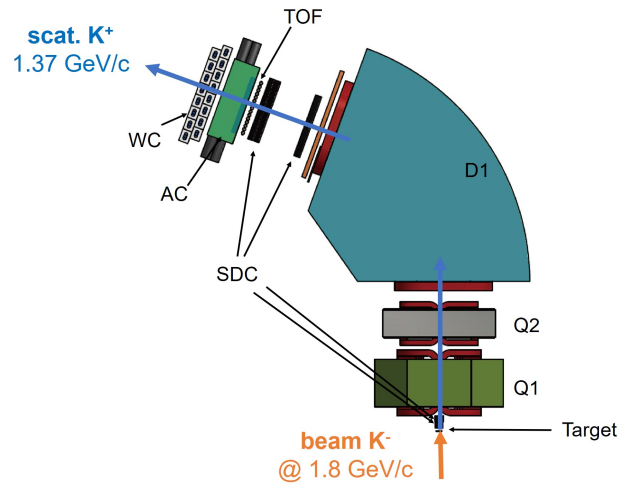


図 1.14: S-2S スペクトロメータとその周辺検出器

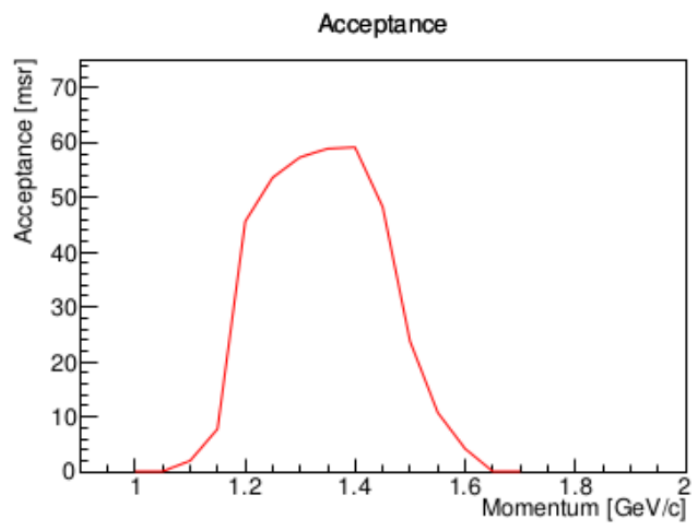


図 1.15: S-2S スペクトロメータの運動量アクセプタンス

表 1.5: 2 つのスペクトロメータ S-2S と SKS の比較

	S-2S	SKS
立体角 (msr)	55	110
$\Delta p/p$ (FWHM)	5×10^{-4}	3×10^{-3}
磁石配置	QQD	D

アクティブファイバー標的

表 1.3 にある運動量分解能を用いて表 1.7 から質量分解能を見積もると, $\Delta M < 2$ MeV(FWHM) の達成には, $\Delta E_{straggle} < 0.8$ MeV(FWHM) が要求されることがわかる.

高統計を維持したまま, $\Delta E_{straggle}$ の要求性能を満たす為に, 新たな標的としてアクティブファイバー標的 (AFT) が開発されている. 図 1.16 に AFT の概念図を示す. AFT は直径 3 mm のシンチレーションファイバーから構成され, 標的と同時にシンチレーション検出器としての役割を果たし, 標的中での K 中間子のエネルギー損失をイベントごとに直接測定することができる. 測定されたエネルギー損失を用いて反応点での K 中間子の運動量を補正することで, エネルギー分解能の大幅な改善を見込んでいる.

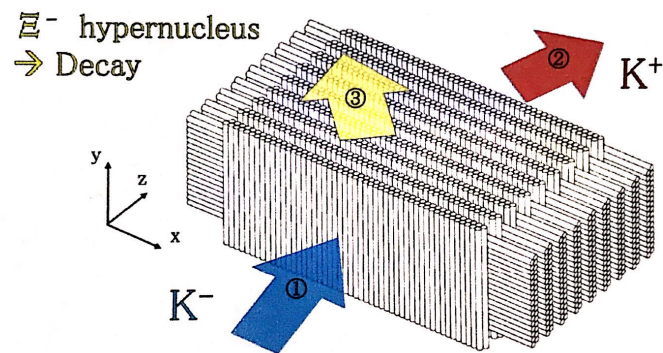


図 1.16: AFT の概念図.

以上 2 点の改善により, E70 実験でのエネルギー分解能は表 1.4 のように見積もられており, 目標分解能 $\Delta M < 2$ MeV を達成することができる.

1.3 E05 実験における運動量解析

小節 1.2.2 では, E05 実験のセットアップと, 運動量解析の概要を述べた. 本節では, E05 実験で行われている運動量解析の詳細を述べる.

1.3.1 入射運動量

入射 K^- の運動量は, K1.8 ビームラインスペクトロメータの上流, 下流の飛跡情報と, QQDQQ システムの磁場分布から計算して得られる 3 次の輸送行列によって求められる.

一般に，輸送行列を M_t ，スペクトロメータ上流での位置・角度ベクトルを $\vec{X}_{in} = (x_{in}, y_{in}, \frac{dx_{in}}{dz_{in}}, \frac{dy_{in}}{dz_{in}})$ ，下流での位置・角度ベクトルを $\vec{X}_{out} = (x_{out}, y_{out}, \frac{dx_{out}}{dz_{out}}, \frac{dy_{out}}{dz_{out}})$ ，とすると，以下の式によって運動量は計算される．

$$\vec{X}_{out} = M_t \left(\vec{X}_{in}, \delta \right) \quad (1.11)$$

ここで， δ は D 磁石の磁場から計算される中心軌道の運動量 p_0 と粒子の運動量 p を用いて， $\delta = \frac{p-p_0}{p_0}$ と書ける．この時，

$$\chi^2 = \frac{1}{n-5} \left(\sum_{i=1}^{N_{in}} H_i \left(P_i - f_i(\vec{X}_{in})\omega_i \right)^2 + \sum_{i=1}^{N_{out}} H_i \left(P_i - g_i(\vec{X}_{in})\omega_i \right)^2 \right) \quad (1.12)$$

$$n = \sum_{i=1}^{N_{in}+N_{out}} H_i$$

$$H_i = \begin{cases} 1 & i \text{ 番目の面に hit があった時} \\ 0 & i \text{ 番目の面に hit がなかった時} \end{cases}$$

$$\left\{ \begin{array}{l} i : \text{チェンバ面の番号} \\ P_i : i \text{ 番目の面の hit 位置} \\ \omega_i : i \text{ 番目の面の位置分解能} \\ f_i(\vec{X}_{in}), g_i(\vec{X}_{out}) : \text{輸送行列で計算した } i \text{ 番目の面の hit 位置} \\ N_{in}, N_{out} : \text{上流側, 下流側のチェンバの面の数} \end{array} \right.$$

で計算される χ^2 が最小値をとる 5 次元ベクトル (\vec{X}_{in}, δ) を求め，運動量 p を求める．

E05 実験では，上流側の飛跡検出器として BFT[26] が用いられている．BFT は直径 1 mm のシンチレーションファイバーを 2 層 (xx') に並べた構造をとり， x 方向の hit 位置 x_{BL-in} のみを取得する．

式 (1.11) において，下流から上流への輸送行列を M'_t とすると，

$$\vec{X}_{BL-in} = M'_t \left(\vec{X}_{BL-out}, \delta \right) \quad (1.13)$$

となり， M'_t を三次まで取ることで，

$$x_{BL-in} = \sum_{a+b+c+d+e=1}^3 C_{(a,b,c,d)} (x_{BL-out})^a (y_{BL-out})^b (u_{BL-out})^c (v_{BL-out})^d (\delta)^e \quad (1.14)$$

$$\text{ただし, } u_{BL-out} = \frac{dx_{BL-out}}{dz_{BL-out}}, v_{BL-out} = \frac{dy_{BL-out}}{dz_{BL-out}}$$

といった関係式が得られる．ここで $C_{(a,b,c,d)}$ は輸送行列 M'_l の要素であり，BL-in, BL-out はそれぞれ，K1.8 ビームラインスペクトロメータの上流側と下流側の座標であることを意味する．E05 実験では，式 (1.14) を δ についての 3 次方程式とみなし，その解を求めることで運動量 p を求めている．また，輸送行列は計算コード ORBIT[27] を使用して計算されている．

1.3.2 散乱運動量

散乱運動量の解析には，Runge-Kutta 法を用いている [28]．これは，運動方程式

$$\frac{d\vec{p}}{dt} = q\vec{v} \times \vec{B} \quad (1.15)$$

$$\begin{cases} \vec{p}: \text{運動量ベクトル}, \left(\vec{p} = \frac{m\vec{v}}{\sqrt{1-\beta^2}} \right) \\ \vec{v}: \text{速度ベクトル} \\ \vec{B}: \text{磁場} \end{cases}$$

に従って Runge-Kutta 法により飛跡を計算し，SKS の上流側 (SDC1,2) の hit 位置，下流側 (SDC3,4) の hit 位置に対して，式 (1.12) で表される， χ^2 を最小にする 5 次元ベクトル (\vec{X}_{SKS-in}, p) を求めるものである．SKS スペクトロメータは立体角が大きく，K1.8 スペクトロメータに比較すると磁場が均一的でないため，輸送行列ではなく Runge-Kutta 法を現在用いている．

E05 実験で用いられている運動量解析手法は，前述の通りだが，他に代表的には，以下のような手法が存在している．

スプライン法

スプライン法 (cubic spline 法) は，粒子の飛跡をスプライン曲線で近似する手法である．構成された飛跡は実際の磁場の下での運動方程式は満たしておらず，数値積分より解析速度は速いものの，運動量分解能は悪化する．また，実験で検出された全てのトラックに fit することはできない可能性がある．これは特に，磁場の変化が大きい領域の飛跡や，湾曲の強い飛跡で起こることが知られている [29]．しかしながら，飛跡の測定誤差が小さく，多重散乱が少ない場合にはよく成功している．高エネルギーの実験で，それほどの分解能を必要としない場合に使用されている．

RCNP Grand Raiden (スペクトログラフ法)

大阪大学核物理研究センター (Research Center for Nuclear Physics, RCNP) のサイクロトロン施設では, 高分解能ビーム分析ラインと組み合わせて, 運動量分散整合技術 (dispersion matching) を用いることで 300 MeV の陽子に対して 15 keV の非常に高い分解能を達成している [30][31]. 輸送行列の高次の項も含めて δ と収束面での位置 x が 1 対 1 対応し ($\langle x|\theta \rangle = 0, \langle y|\phi \rangle = 0$), ビーム側と散乱側の分散 $\langle x|\delta \rangle$ を一致させるような磁場を求め, 実験中はその磁場を実現するように磁気光学系を制御する. 磁場測定の結果から飛跡の fitting を行うのではなく, 計算磁場を実現し, x の測定値のみから高分解能での運動量解析を実現している.

我々は, ここで挙げられている Runge-Kutta 法, 輸送行列法, スプライン法, スペクトログラフ法

1.3.3 運動量補正

入射運動量 p_B と散乱運動量 p_S は小節 1.3.1, 1.3.2 で述べたように解析される.

式 (1.2) を用いて欠損質量を求めるが, 実験特有のものとして, 以下 2 点の運動量補正を行う必要が生じている.

Beam Through 補正

E05 実験では, 同じ運動量であるにも拘わらず, 粒子の電荷の正負によってスペクトロメータで解析される運動量に違いが出る現象が生じている. さらに, 入射運動量の大きさに依存して, K1.8 スペクトロメータと SKS スペクトロメータで解析される運動量にも差が生じている. (K^-, K^+) 反応において, 入射 K^- と散乱 K^+ の運動量に差が生じるのを補正する必要があり, この補正を本論文では Beam Through 補正 (BT 補正) とする. その詳細を以下に記す.

E05 実験では, 標的を設置しない状態, つまり入射粒子が K1.8 スペクトロメータで解析された後, 反応を起こさずそのまま SKS スペクトロメータで解析されるデータが取得されている. これを Beam Through run (BT) と呼ぶ. Beam Through データは表 1.2 中 Beam Through で示される入射運動量, 入射粒子に対して取得された.

同一の粒子に対して運動量解析を行うため, K1.8, SKS で得られる運動量の差 $dp_{BS} = p_B - p_S$ は, 0 を中心に分布することが期待される. しかしながら, 図 1.17 が示すように, dp_{BS} のピーク位置は粒子の電荷と入射運動量に依存した値となっている.

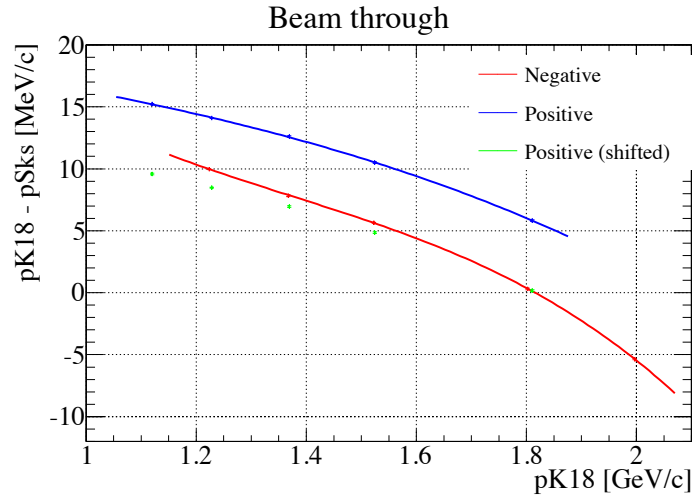


図 1.17: BT run における運動量の差 $dp_{BS} = p_B - p_S$ のピーク位置を，入射運動量に対してプロットした図．縦軸は dp_{BS} のピーク位置，横軸は p_B である． dp_{BS} は p_B に依存して 0 でない値をとっている．また，入射粒子は赤線が K^- ，青線が K^+ であり，これらの比較により，粒子の電荷の正負によって dp_{BS} が異なることがわかる．

UV 補正

散乱運動量の解析値と，散乱粒子の飛跡の角度に相関があることが経験的に知られている．この相関を除き，散乱粒子運動量を補正することを，本論文では UV 補正と呼ぶ．以下にその詳細を示す．

CH_2 標的で取得したデータでは，素過程 $p(K^-, K^+)\Xi^-$ による Ξ^- の欠損質量ピークが観測される (図 1.12)．E05 実験では，欠損質量のピーク位置が $u_{SKS-in} = \frac{dx_{SKS-in}}{dz_{SKS-in}}$ ， $v_{SKS-in} = \frac{dy_{SKS-in}}{dz_{SKS-in}}$ と相関を持つ．実際，観測されたビーム運動量と散乱角から $p(K^-, K^+)\Xi^-$ を仮定して二体の運動学より計算される散乱運動量と観測された散乱運動量の差は 1.18 に示すように u_{SKS-in} , v_{SKS-in} と相関を持つ．

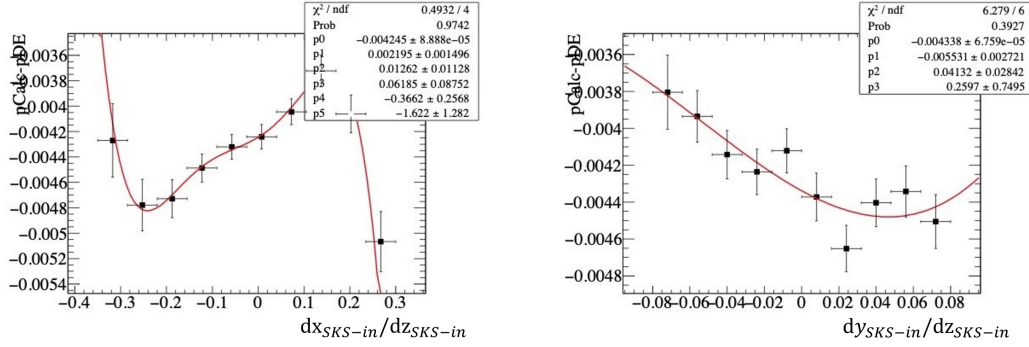


図 1.18: 入射運動量 1.8 GeV/c における $p(K^-, K^+)$ 散乱運動量の角度依存性．縦軸は素過程 $p(K^-, K^+)\Xi^-$ を仮定して運動学より計算される散乱運動量 pCalc と解析散乱運動量 pDE の差，横軸は散乱 K^+ の飛跡の角度 (左: $u_{SKS-in} = \frac{dx_{SKS-in}}{dz_{SKS-in}}$, 右: $v_{SKS-in} = \frac{dy_{SKS-in}}{dz_{SKS-in}}$) である．

従来手法

以上 2 点を考慮し，従来手法では以下の関数を置き運動量の絶対値を定めていた．

$$p_B^{cor} = ap_B + b \quad (1.16)$$

$$p_S^{cor} = \alpha p_S^3 + \beta p_S^2 + \gamma p_S + \delta \quad (1.17)$$

$$p_S^{cor2} = p_S^{cor} + \sum_{i=1}^5 p_i^u u_{SKS-in}^i + \sum_{i=1}^3 p_i^v v_{SKS-in}^i \quad (1.18)$$

ここで， $a, b, \alpha, \beta, \gamma, \delta, p_i^u, p_i^v$ は以下のように決められている．

1. BT データより， $p_B^{cor} - p_S^{cor} = 0$ となるように fitting する．これより， $a - \gamma, b - \delta, \alpha, \beta$ を決める．
2. 素過程 $p(K^-, K^+)\Xi^-$ ， $p(K^-, \pi^+)\Sigma^-$ について $u_{SKS-in} = v_{SKS-in} = 0$ の領域で p_B^{cor}, p_S^{cor} を用いて式 (1.2) から欠損質量を計算する．欠損質量のピークがそれぞれ Ξ^- ， Σ^- の質量 (PDG value) に合うように fitting することで a, b, γ, δ の値を決定する．
3. 図 1.18 で見られる素過程 $p(K^-, K^+)\Xi^-$ での UV 相関を補正する．図 1.18 の赤線のように fitting し，その係数から p_i^u, p_i^v が決まる．

以上のような補正を経て，最終的に得られた p_B^{cor}, p_S^{cor2} を欠損質量の計算に用いている．

1.3.4 従来手法の問題点

このような補正手法は現在，解析者に依存しており現象論的に行われている．従来手法においては，式 (1.16)-(1.18) のような補正関数を置いていたが，どのような関数形を定めるかは任意である．また，経験則的，現象論的に見られる相関に対しての補正にとどまっており，実験データに現れている相関をくまなく探すのは困難である．

1.4 研究目的

前小節 1.3.4 で述べたように，従来の解析手法は補正手法が解析者に依存している．本研究では，機械学習を適用した運動量解析を新たに導入し，解析者に依存しない解析手法を検討する．

機械学習は，

1. 広範に及ぶ効果を所定の方程式を置かずに考慮できる
2. 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可能性がある

といった利点を持つ．その為，運動量補正に機械学習を適用できれば，解析者に依存せず，従来考慮できていなかった相関を補正できる可能性がある．特に未知の相関を補正することは，欠損質量分解能の向上にも繋がる．しかしながら，現在に至るまで，分光実験における運動量解析へ機械学習をこのように適用した事例は乏しい．本研究では，E05 実験データに対して機械学習を適用して，補正の自動化並びに分解能が向上するかを確かめる．具体的な適用方法は 3 で述べる．

E05 実験で見られる相関は実験特有のものである．SKS スペクトロメータは，実験前に非常に精密に磁場測定が行われたが，コイルにクエンチが起きる事故により測定時と実験時では磁場マップが異なっている．

しかしながら，補正に機械学習を適用する意義は E05 実験の解析に留まらない，例えば，E70 実験で新しく用いられる S-2S スペクトロメータは QQD で構成されており，その磁場マップは SKS より複雑である．精密な磁場測定とシミュレーションにより，現在 $\Delta p/p = 5 \times 10^{-4}$ が達成できると見込まれている．また，ビームラインスペクトロメータでは， $\Delta p/p \sim 1 \times 10^{-3}$ しか達成されておらず，S-2S でこの分解能が達成できれば，ビーム側も 5×10^{-4} という設計性能が得られるかもしれない．しかし，シミュレーション状況と実験状況の違いによりどのような相関が現れるかは，実験データから初めてわかる．そ

のため，将来行われる E70 実験の運動量解析においても，相関の補正は分解能達成の鍵となる．E05 実験のセットアップは，散乱スペクトロメータと標的を除いて同じである．また，SKS(E05) は，S-2S(E70) に 1 桁程度劣るものの，非常に高い分解能をもつ．これらのことから，E05 実験データに対しての本手法の適用可能性は E70 実験データに対しての指標となり得る．

さらに一般に，2 次粒子を用いる分光実験においては，入射粒子の運動量の拡がりが大きく，入射粒子と散乱粒子の運動量分解能が有限であることから E05/E70 実験に限らず粒子の運動量絶対値を一意に定めるのは困難である．この点で，本手法は様々な分光実験に対しての適用も考えられる．

第2章

機械学習

機械学習とは、明示的なプログラムをすることなく、与えられたデータに基づき、機械が自動的にその背後に存在する規則性を見つけ出す（学習する）ものである。一般的な関数フィッティングとは異なり、所定の式を置くことなく解析することができるので、解析者が従来まで考慮しきれていなかった効果を考慮することができる可能性がある。機械学習の中でも、教師あり学習、教師なし学習、強化学習など様々な種類のものが存在するが、本章では中でも教師あり学習を取り上げ、説明する。

2.1 教師あり学習

教師あり学習とは、入力と出力のペア群から学習する機械学習アルゴリズムである。教師あり学習では、以下のようなデータセットの存在を想定する。

$$(X, y) = \{(X_i, y_i)\}_{i=1,2,\dots,N} \quad (2.1)$$

ここで、 X_i は i 番目のデータ値（一般にベクトル値）、 y_i はそのデータに対しての教師信号を表す。 N はデータ数である。

教師あり学習とはこのデータセット (2.1) に対してよく適合する関数 $f: X \rightarrow y$ を求めるものであり、 X を説明変数や特徴量、 y を目的変数と呼ぶ。

教師あり学習の中でも扱う問題により、クラス分類問題と回帰問題に分けられる。式 (2.1) の y が離散値である場合はクラス分類問題、連続値である場合は回帰問題と呼ぶ。

2.1.1 クラス分類問題

クラス分類問題は、データを複数のクラスに分類する問題である。クラス数が 2 つのものは 2 値分類、3 つ以上のものは多クラス分類と呼ばれる。図 2.1 にクラス分類の概念図を示す。この図においては、特徴量が $X = (x_0, x_1)$ で表されるような 2 次元の場合に、このパラメータ空間内で境界線を決定することでデータ群を 3 クラスに分類する様子を示している。実際には、この境界線の決定方法は用いるアルゴリズムによって異なる。高エネルギー物理学実験においては、イベントセレクション等への応用が進んでいる [32, 33]。

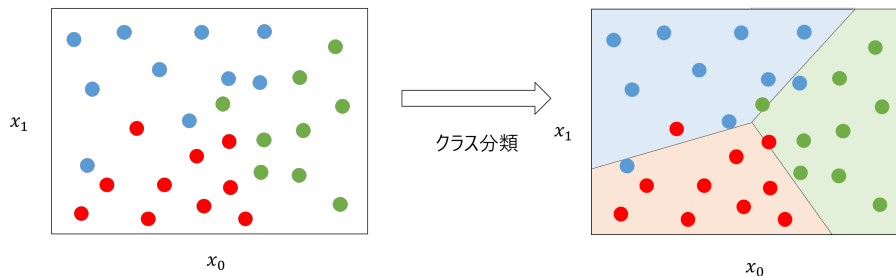


図 2.1: クラス分類の概念図。図では、青、赤、緑の 3 つのクラスに分類されている。

2.1.2 回帰問題

回帰問題は、特徴量と目的変数の間の関係を求める問題である。図 2.2 に回帰問題の概念図を示す。この図においては、特徴量が $X = (x_0)$ のように 1 次元で表される場合に、 $f(X) = ax_0 + b$ と線形関数を置き、学習した場合の例を示す。

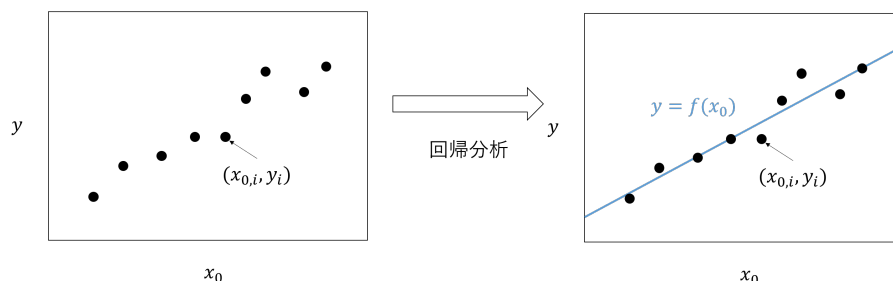


図 2.2: 回帰分析の概念図。図では、線形関数 $f(x_0) = ax_0 + b$ を置いた場合の学習結果を表す。

学習後の予測の際には、特徴量 $X_j = (x_{0,j})$ をもつデータに対して出力値 \hat{y}_j を $\hat{y}_j = f(X_j)$ として推定する。

上記の例では、 f の関数系をあらかじめ仮定して学習を行っていた。このように例えば、多項式の字数を fix すると、単純な関数フィッティングとみなせる。一方で多くの機械学習アルゴリズムでは、複雑な効果を取り入れる為に、非線形関数を用いる。その為、機械学習は非線形関数のフィッティングと捉えることもできる。

2.1.3 損失関数

教師あり学習の際には、学習データに対しての非適合度を損失関数 (Loss Function) として定義し、損失関数が最小値をとる関数 f を探索する。

損失関数は解くタスクにより適切なものを選択する必要がある。ここでは、回帰問題で使用される代表的な損失関数について述べる。以下、学習に用いたデータの目的変数を y 、関数 f による推定値 (予測値) を $\hat{y} = f(X)$ とする。

Mean Squared Error

$$MSE(y, \hat{y}) = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (2.2)$$

Mean Squared Error(平均二乗誤差, MSE) は回帰問題においてよく使用される損失関数である。MSE を最小化した場合に求まる解は、誤差が正規分布に従う前提のもとでの最尤解と一致し、式 (2.2) を N 倍すると最小二乗法で最小化する関数と一致する。

Mean Absolute Error

$$MAE(y, \hat{y}) = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (2.3)$$

Mean Absolute Error(平均絶対誤差, MAE) は、式 (2.3) で表され、MSE と比較して外れ値に対しての影響を受けにくい特徴がある。

2.1.4 K-分割交差検証

機械学習を用いた解析の妥当性を評価するには、学習に用いたデータ以外のデータに対して同じ解析を実行することでテストする必要がある。この時学習に用いられるデータは

「学習データ（訓練データ，training set）」と呼ばれ，テストに用いられるデータは「テストデータ（testing set）」と呼ばれるこの手法は交差検証と呼ばれ，解析手法の母集団への適用妥当性を評価することに用いられる．

交差検証の一種に，K-分割交差検証 (K-Fold cross-validation) が上げられる．図 2.3 に $K = 4$ の際の例を示す．K-分割交差検証では，

1. データを K 個に分割し，そのうち 1 つをテストデータとし，残る $K - 1$ 個を学習データとしたテストを実行する．
2. K 個に分割されたデータそれぞれをテストデータとしこのテストを K 回繰り返す．
3. 最終的に，そうして得られた K 回の結果を平均することで 1 つの推定を得る．

一方で，手順 1 のように，テストを 1 度だけ実行する手法をホールドアウト検証という．ホールドアウト検証では，ランダムに分割した際に，学習データ，テストデータに偏りが生じ，それに伴い適切な評価を得られない可能性がある．K-分割交差検証では，2, 3 の手順のようにホールドアウト検証を K 回繰り返し，ホールドアウト検証よりロバストな結果を得ることができる．分割数 K が多いほどロバストなテスト結果を得られるが，計算量が K 倍されるので，分割数はそのトレードオフを考慮して選択しなければならない．

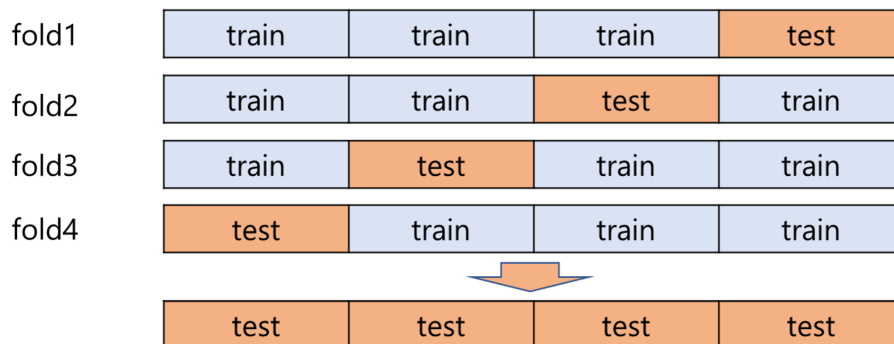


図 2.3: K-分割交差検証．図では， $K = 4$ の場合を示す．各回のホールドアウト検証は fold とよばれる．

層化 K-分割交差検証

分類タスクの場合，K-分割交差検証において，fold ごとに含まれるクラスの割合を等しくすることが行われる．これは層化 K-分割交差検証 (stratified K-Fold cross-validation) と呼ばれ，分割の際の各クラスの割合の違いによる評価のばらつきを抑える目的で行わ

れる。

本節 2.1 では、教師あり学習におけるタスクの種類や、検証方法を述べた。次節以降は、機械学習で用いられる具体的なアルゴリズムや、学習方法について説明する。

2.2 ニューラルネットワーク

2.2.1 パーセプトロン

パーセプトロン (perceptron) とは、脳の神経細胞から着想を得て開発されたアルゴリズムであり、1 つ以上の入力値から 1 つ値を出力する。図 2.4 に 2 つの入力値と 1 つの出力値を持つパーセプトロンの例を示す。

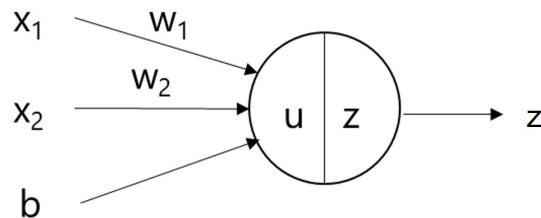


図 2.4: 入力 2 つ (x_1, x_2), 出力 1 つ z のパーセプトロン。

一般に、 p 個の入力値 x_1, x_2, \dots, x_p がパーセプトロンに与えられた時、

$$u = \sum_{i=1}^p w_i x_i + b \quad (2.4)$$

が総入力として与えられる。ここで、 w_i は重み、 b はバイアスと呼ばれる。その後、このパーセプトロンは

$$z = \varphi(u) \quad (2.5)$$

を出力する。ここで、 φ は活性化関数 (activation function) と呼ばれる。活性化関数については、2.2.2 で述べる。また、図 2.4 の丸の部分は、式 (2.4), (2.5) の計算を行うが、この部分を特にユニットと呼ぶ。

2.2.2 ニューラルネットワーク

パーセプトロンは多層に並べることで、図 2.5 のようなネットワークを構成することができる。

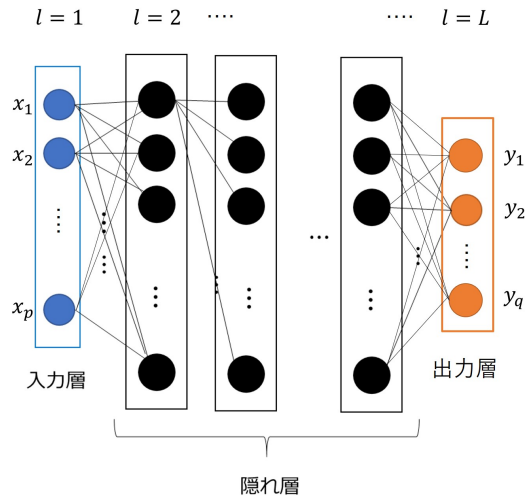


図 2.5: 入力値 p 個, 出力値 q 個の L 層からなるニューラルネットワーク. 丸はユニット, 線は重みを表す.

このようにして作られたネットワークをニューラルネットワーク (Neural Network, NN) という. ニューラルネットは入力層 (input layer), 中間層 (hidden layer), 出力層 (output layer) からなり, L 層からなる場合に 1 層目は入力層, 2, 3, ..., $L-1$ 層目は中間層, L 層目は出力層となる. 中間層の数やユニット数は任意に定めることができ, それらの数が多いほど複雑な関数を表現できることが知られている. 中間層を多数持ち, 深層化したニューラルネットワークを深層ニューラルネットワーク (Deep Neural Network, DNN) と言う. いくつ以上の中間層数から深層ニューラルネットと呼ぶかは様々な定義があるが, 本論文では本章以降, 隠れ層数によらず, ニューラルネット (NN) と記述する.

ここで, このニューラルネットにおける入力が $x = (x_1, x_2, \dots, x_p)$, 出力が $\hat{y} = (\hat{y}_1, \hat{y}_2, \dots, \hat{y}_q)$ であるとき, l 層目の出力は

$$x_i^{(l)} = \begin{cases} x_i^{(0)} & (l = 1) \\ \varphi^{(l)} \left(\sum_{j=1}^{\#unit_{(l-1)}} W_{ij}^{(l)} x_j^{(l-1)} + b_i^{(l)} \right) & (l = 2, 3, \dots, L) \end{cases} \quad (2.6)$$

で表される. ここで $x = x^{(0)}$, $\hat{y} = x^{(L)}$ であり, $\#unit_{(l)}$ は l 層目のユニット数である. $\varphi^{(l)}$ は l 層目の活性化関数である.

2.2.3 活性化関数

2.2.1, 2.2.2 で触れたように, パーセプトロンは出力に活性化関数 φ を用いる. この活性化関数には, 目的に応じて種々の関数を使用されている. 以下にそのうち数例を取り上

げ，特徴を述べる．

階段関数

$$\varphi(u) = \begin{cases} 0 & (u \leq 0) \\ 1 & (u > 0) \end{cases} \quad (2.7)$$

階段関数 (ステップ関数, step function) は閾値を 0 とし，入力値に応じて 0 か 1 を出力する関数であり．パーセプトロンが開発された当初，実装に用いられていた．しかし，2.2.4 で述べるように，誤差逆伝播法に基づくニューラルネットの学習の際には，活性化関数の微分が必要であるので，現在ではあまり使用されていない．

シグモイド関数

$$\varphi(u) = \frac{1}{1 + e^{-u}} \quad (2.8)$$

誤差逆伝播法が提案された後，階段関数と似たふるまいをし，かつ微分値を得ることができるので使用されるようになった．現在は，中間層の活性化関数に使用されることは少なくなかったが，二値分類問題の際の出力層の活性化関数に使用されている．

ReLU 関数

$$\varphi(u) = \max(0, u) = \begin{cases} 0 & (u \leq 0) \\ u & (u > 0) \end{cases} \quad (2.9)$$

ReLU(Rectified Linear Unit) 関数は，シグモイド関数に比較して計算量が少なく，より良い結果が得られることもある．その為，中間層の活性化関数として現在よく使用されている．

恒等関数

$$\varphi(u) = u \quad (2.10)$$

恒等関数は，多次元である隠れ層の出力を 1 次元にする目的で，回帰問題の際の出力層の活性化関数として主に使用されている．

2.2.4 誤差逆伝播法

2.1.3 で触れたように、機械学習における学習とは、損失関数 L を最小化する関数を求めることである。特にニューラルネットにおいては、式 2.6 で表されるパラメータ（重み $W_{ij}^{(l)}$ とバイアス $b_i^{(l)}$ ）を

$$W_{ij}^{(l) \text{ new}} = W_{ij}^{(l)} - \epsilon \frac{\partial L}{\partial W_{ij}^{(l)}} = W_{ij}^{(l)} + \Delta W_{ij}^{(l)} \quad (2.11)$$

$$b_i^{(l) \text{ new}} = b_i^{(l)} - \epsilon \frac{\partial L}{\partial b_i^{(l)}} = b_i^{(l)} + \Delta b_i^{(l)} \quad (2.12)$$

に従って徐々に更新（勾配降下法，gradient descent method）し、損失関数が最小になるパラメータに到達させる。ここで ϵ は学習率と呼ばれ、微小な値を用いる。各層のパラメータの更新は、出力層側から入力層側へ順に以下の式で更新される。

$$\Delta W_{ij}^{(l)} = \epsilon \delta_i^{(l)} x_j^{(l+1)} \quad (2.13)$$

$$\Delta b_i^{(l)} = \epsilon \delta_i^{(l)} \quad (2.14)$$

これを誤差逆伝播法という。ここで

$$\delta_i^{(l)} = \begin{cases} -\frac{\partial \varphi^{(L)}(u_i^{(L)})}{\partial u_i^{(L)}} \frac{\partial L}{\partial \hat{y}_i} & (l = L) \\ \frac{\partial \varphi^{(l)}(u_i^{(l)})}{\partial u_i^{(l)}} \sum_{k=1}^{\#unit_{(l+1)}} \left(\delta_k^{(l+1)} W_{ki}^{(l+1)} \right) & (l = 1, 2, \dots, L-1) \end{cases} \quad (2.15)$$

$$u_i^{(l)} = \sum_{j=1}^{\#unit_{(l-1)}} W_{ij}^{(l)} x_j^{(l-1)} + b_i^{(l)} \quad (2.16)$$

であり、 $\frac{\partial \varphi^{(l)}(u_i^{(l)})}{\partial u_i^{(l)}}$ が含まれることから、活性化関数 $\varphi^{(l)}$ が微分可能であることが要求される。

ここではパラメータ更新が式 (2.12)、(2.12) に則るとしているが、学習がより効率的に進むような最適化のアルゴリズムとして、例として Momentum, RMSProp[34], Adam[35] などが考案されている。

また、NN の学習では、用いている学習データの全てを用いて 1 度パラメータ更新した段階を 1 epoch と呼ぶ。

2.2.5 ソフトウェア

本研究では，ニューラルネットの実装に TensorFlow[36] と Keras[37] を用いた．TensorFlow は Google により開発された機械学習のオープンソースライブラリである．また，Keras は Python で書かれたニューラルネットのライブラリであり，TensorFlow を含む様々なライブラリの上で動作する．Keras はニューラルネットを平易に実装する為に開発され，特に TensorFlow では Keras を高レベル API として利用することができる．

2.3 勾配ブースティング決定木

勾配ブースティング決定木は，イベントセレクション等で成功を収めている [32]．ここでは，その概要に触れる．

2.3.1 決定木

決定木 (Decision Tree) は，図 2.6 のような構造を持つ．データセット $(X, y) = (X_i, y_i)_{i=1,2,\dots,N}$ が与えられたとき，各データは X の条件に従って図 2.6 最深部のいずれかの葉 (ノード) に属する．各々のノードには 1 つの予測値が割り当てられており，予測の際にはデータ X_i が属するノードに割り当てられている予測値がそのまま \hat{y}_i として用いられる．

学習時には，損失関数の減少が最大となる条件を随時探索し，木を構築する．

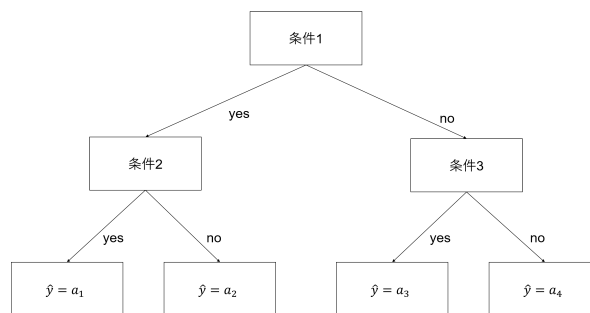


図 2.6: 深さ 2，葉の数 4 の決定木．もしデータ X_i に対して条件 1=yes，条件 2=no ならば $\hat{y}_i = a_2$ となる．

2.3.2 勾配ブースティング決定木

勾配ブースティング決定木 (Gradient Boosting Decision Tree, GBDT) は決定木の集合であり, 図 2.7 のような構造をとる .

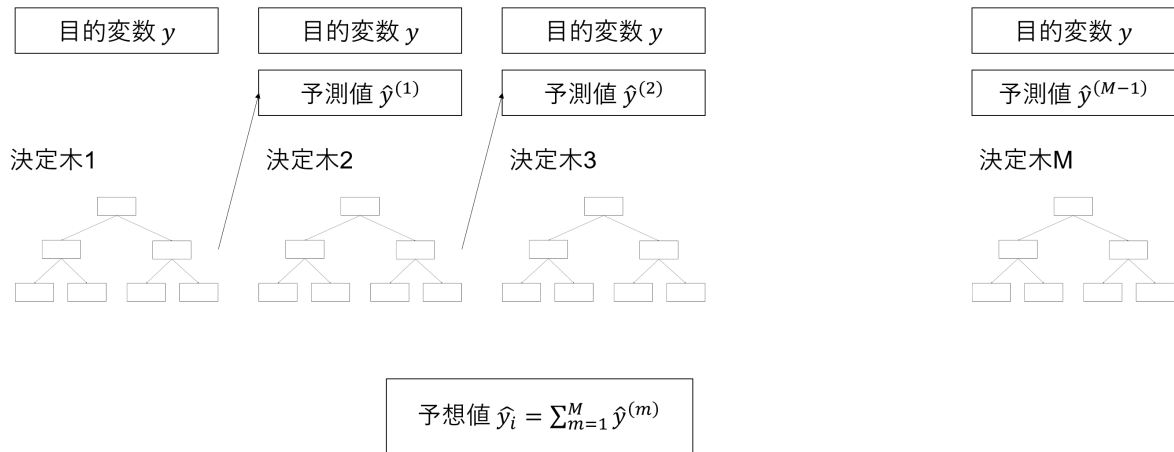


図 2.7: 木の数が M である勾配ブースティング決定木 .

具体的には, 以下のように GBDT は構築される .

1. まず, 損失関数を減少させるように決定木を 1 つ構築する .
2. 次の決定木は, 教師データ y と, 前の決定木の予測値の差に対して学習を行う .
3. 2 の手順を, ハイパーパラメータとして指定した決定木の数だけ逐次的に繰り返す .

予測値 \hat{y}_i は, ハイパーパラメータとして与える決定木の総数を M とすると, それぞれの決定木の予測値 $\hat{y}_i^{(m)}$, ($m = 1, 2, \dots, M$) を用いて

$$\hat{y}_i = \sum_{m=1}^M \hat{y}_i^{(m)} \quad (2.17)$$

とする .

第 3 章

機械学習を用いた運動量解析の最適化

従来の運動量解析の問題と、機械学習で期待される改善は 1.4 節で述べた通りである。本研究では、2 章で説明した機械学習手法を用いた運動量解析を実装した。本章ではその実装内容と性能を述べる。

3.1 運動量最適化の為の事象選別

1.2.2 節で述べたように、E05 実験では、 CH_2 標的でのデータが取得されている。このデータについては、入射 K^- が CH_2 中の H と反応し、 M_{Ξ^-} や M_{Σ^-} といったよく知られている欠損質量のピークを観測できる。素過程 $p(K^-, K^+)\Xi^-$, $p(K^-, \pi^+)\Sigma^-$ における、入射運動量 $p_B = 1.8 \text{ GeV}/c$ のもとでの実験室系散乱角 θ と散乱運動量 p_S の関係を図 3.1 に示す。 $\theta = 0^\circ$ では $p(K^-, K^+)\Xi^-$ において $p_S \sim 1.3 \text{ GeV}/c$, Ξ^- への運動量移行 $\sim 0.5 \text{ GeV}/c$ となり、 $p(K^-, \pi^+)\Sigma^-$ において $p_S \sim 1.6 \text{ GeV}/c$, Σ^- への運動量移行 $\sim 0.2 \text{ GeV}/c$ となる。

$p_B = 1.8 \text{ GeV}/c$ での $^{12}\text{C}(K^-, K^+)\Xi^-$ ^{12}Be における散乱 K^+ の運動量は $\sim 1.37 \text{ GeV}/c$ であり、近い運動量領域であるこれら 2 つの素過程を用いて、図 3.1 で表される運動学を満たすように較正を行う。

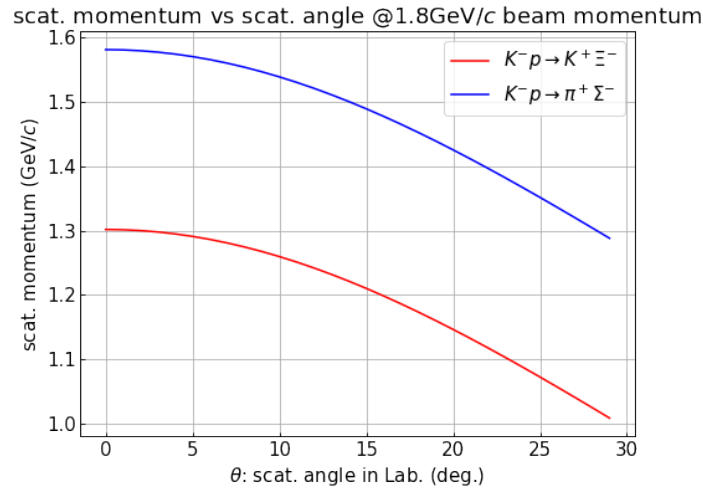


図 3.1: 入射 K^+ 運動量 1.8 GeV/c のもとの、素過程 $p(K^-, K^+) \Xi^-, p(K^-, K^+) \Sigma^-$ における散乱角と散乱運動量の関係

本研究では、これらの素過程のイベントを用いて、機械学習で運動学最適化を行う。表 3.1 に、それぞれのデータ数を示す。ここで、 Ξ^- 生成、 Σ^- 生成は散乱粒子が K^-, π^- であり、や反応点の位置条件、欠損質量が正になる条件を課したデータである。散乱粒子の識別は、飛行時間の測定により行われている (1.2.2 節)。

表 3.1: 性能評価に使用したデータ

	Ξ^- 生成	Σ^- 生成
データ数	1.1×10^4	2.2×10^5

しかし、図 3.2 に見られるように、 CH_2 標的を使った学習データ (表 3.1) には水素からの信号に加えて ^{12}C 由来のバックグラウンドが存在している。その為、 M_{Ξ^-} 、 M_{Σ^-} ピークの周辺を選択しカットをかけて、水素と反応した素過程のイベントとみなして学習する。例として、図 3.3 に M_{Ξ^-} 、 M_{Σ^-} カットの範囲を $1-4\sigma$ まで様々に変えた場合を示す。カット範囲を広げるほど統計は増加するが、 ^{12}C の混入も多くなり、適切なカット範囲はこのトレードオフで決まると考えられる。その検討は 3.5.3 で行う。

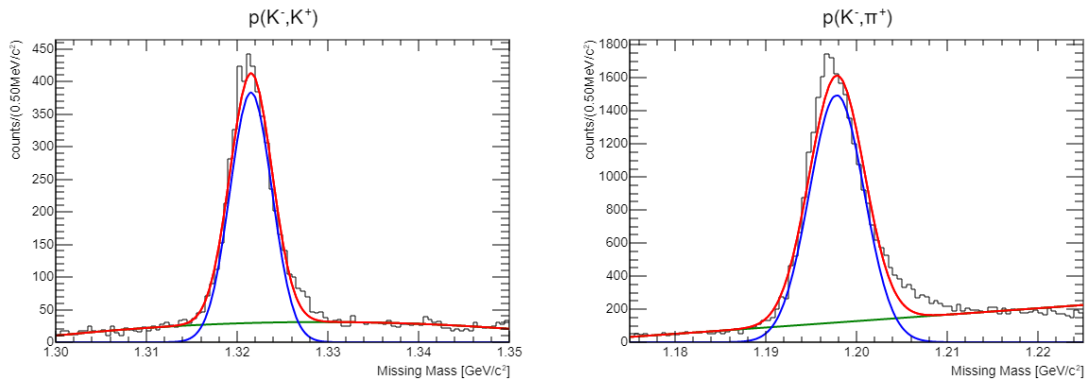


図 3.2: 検証に使用する素過程の欠損質量スペクトル . 左が $p(K^-, K^+)\Xi^-$, 右が $p(K^-, \pi^+)\Sigma^-$ である . H との reaction で現れる Ξ^- , Σ^- のピークを正規分布 (青), ^{12}C 由来のバックグラウンドを二次多項式 (緑) を使ってあてはめた .

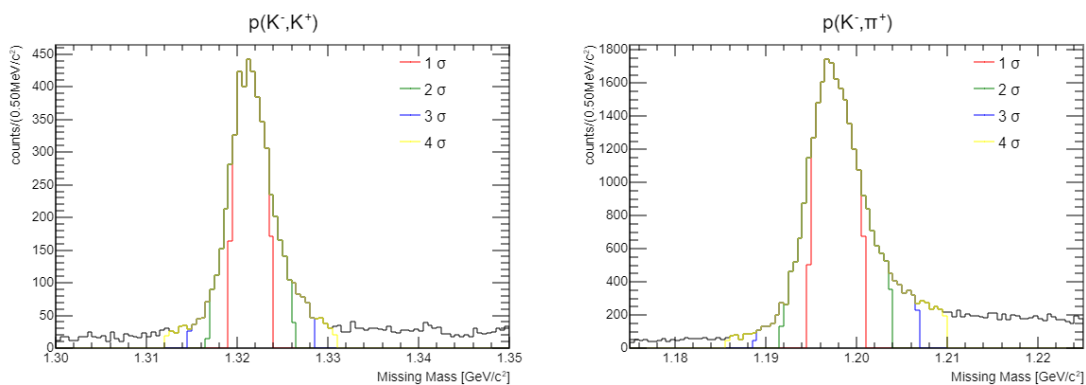


図 3.3: M_{Ξ^-} , M_{Σ^-} カットを $1, 2, 3, 4\sigma$ 範囲とした時の学習データ .

3.2 機械学習と運動量最適化の対応関係

1.3.3 で述べた補正とは, 運動量 p_B, p_S が以下の拘束条件に従うように, p_B, p_S の関数を定めることであった .

1. Beam Through run において $p_B = p_S$.
2. CH_2 target run での素過程 $p(K^-, K^+)\Xi^-$, $p(K^-, \pi^+)\Sigma^-$ において, それぞれの式 (1.2) で表される M_X が, Ξ^- と Σ^- の質量 (M_{Ξ^-} , M_{Σ^-}) にピークを持つ .

これを教師あり学習の回帰問題に落とし込むと

1. Beam Through run において ,

$$Loss(p_B, p_S) \quad (3.1)$$

が最小値をとる $p_B(X_B), p_S(X_S)$ を求める .

2. CH₂target run での素過程 $p(K^-, K^+) \Xi^- , p(K^-, \pi^+) \Sigma^-$ において ,

$$Loss(M_X(p_B, p_S) - M_{\Xi^-, \Sigma^-}) \quad (3.2)$$

が最小値をとる $p_B(X_B), p_S(X_S)$ を求める .

となる . Loss 関数は式 (2.2)

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

で表される MSE もしくは , 式 (2.3)

$$MAE(y, \hat{y}) = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|$$

で表される MAE を用いる . ただしここで , X_B, X_S はそれぞれビームラインスペクトロメータ側の飛跡検出器 BFT, BC3, 4 と SKS 側の飛跡検出器 SDC1-4 で得られた hit 情報を表す .

ここで , $p_S(X_S)$ は X_S の関数であり , X_B を入力変数に持たないことに留意する . なぜなら , p_S を求める際に X_B, X_S を両方用いれば , 内部的に運動学計算をするだけのモデルとなるからである . 詳細は 3.3.3 で述べる .

3.3 モデルの実装

3.3.1 従来解析フロー

従来手法 (小節 1.3.3) による欠損質量の計算の流れを , 図 3.4 のように表す . 解析は図の左から右の順に行われる .

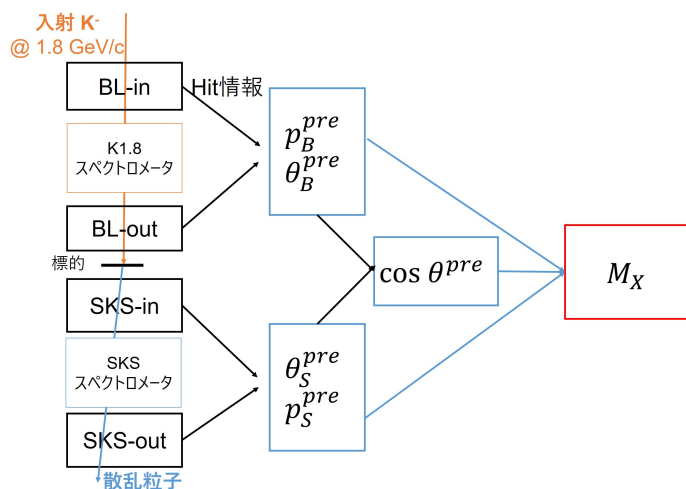


図 3.4: 従来手法による欠損質量計算フロー．左列の hit 情報は，それぞれ表 3.2 で表す検出器から得られたものである．中列の青い四角はそれぞれの変数 p_B , $\cos \theta$, p_S を従来手法によって定めることを意味し，右列の M_X は，1.2 に従って計算される欠損質量を表す．

表 3.2: 位置や角度座標に付随する下付き文字と検出器の対応

記号	飛跡検出器	構成
$BL - in$	BFT	xx'
$BL - out$	BC3,4	$xx'vv'uu', uu', vv', xx'$
$SKS - in$	SDC1,2	$xx'vv'uu', uu'xx'$
$SKS - out$	SDC3,4	$vxuv'x'u', vxuv'x'u'$

図中上付き文字 pre は，従来手法を意味する．入射粒子側は，BL-out のヒット情報から局所的な飛跡 (ローカルトラック) を構成する．ローカルトラックとは，図 3.5 で示されるような，ヒット情報から最小二乗法より決定される 1 本の直線トラックである．

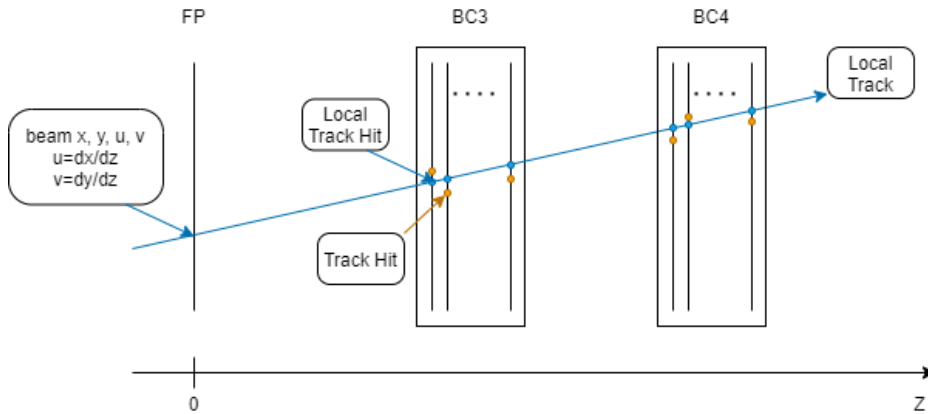


図 3.5: ローカルトラックの概念図．一組の検出器 (図は BL-out の BC3,4) のヒット情報から一本の直線トラックを定め、これをローカルトラックとする．

BL-out でのローカルトラックが決まれば、BFT の情報から 1.3.1 節で述べた手法で運動量 p_B^{pre} が求まる．ここで、 p_B^{pre}, p_S^{pre} はスカラーである．また、BL-out でのローカルトラックから、標的への入射角 θ_B が決まる．散乱粒子側は、SKS-in, SKS-out のヒット情報より、1.3.2 節で述べた手法で運動量構成を行う．結果として、散乱運動量 p_S^{pre} と SKS 入口での粒子飛跡角度 θ_S^{pre} が求まる．入射側と散乱側で得られた飛跡角度 $\theta_B^{pre}, \theta_S^{pre}$ から散乱角 θ^{pre} が求まり、 $p_B^{pre}, \cos \theta^{pre}, p_S^{pre}$ から欠損質量を計算する．

3.3.2 モデル構成

本研究では、図 3.4 中の中列に現れている、従来手法による $p_B, \cos \theta, p_S$ の解析のうち 1 変数を、機械学習によるものに置換する．その際、どの変数を機械学習に置換するかは自由である．例えば、 p_S のみを置換した場合は図 3.6 のように表される．

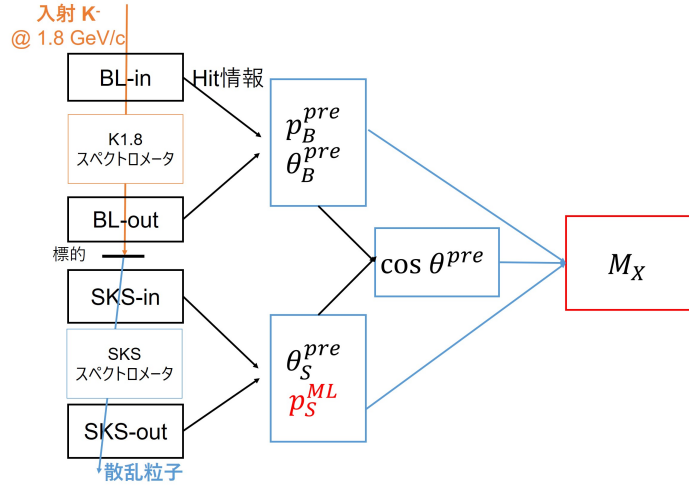


図 3.6: p_S のみを機械学習に置換した欠損質量計算フロー

3.3.3 学習手法

3.1 節で述べた実験データに対して式 (3.2) で表される損失関数が最小値をとるような $p_B(X_B), \cos \theta(X_C), p_S(X_S)$ を構築する．ここで新たに, X_C は BL-out と SKS-in とした．本研究では, モデルのインプットとして BL-out, SKS-in, SKS-out のヒット情報をそのまま与えるのではなく, ローカルトラックの, 焦点面での座標を用いた．これは BL-out では, 図 3.5 における x, y, u, v で表されている量である．

$p_B, \cos \theta, p_S$ のいずれか 1 変数を最適化 (チューニング) する際, 残りの 2 変数は従来解析の結果をそのまま用いる．例えば, 図 3.6 のように p_S のみをチューニングする場合, 残り 2 変数 $p_B, \cos \theta$ は従来解析の結果 $p_B^{pre}, \cos \theta^{pre}$ をそのまま使用する．ここで素過程 $p(K^-, K^+) \Xi^-$ または $p(K^-, \pi^+) \Sigma^-$ を仮定すれば, 式 (1.2) を p_S について陽に解くことができる．つまり, その解を p_S^{kine} とすると

$$M_X(p_B^{pre}, \cos \theta^{pre}, p_S^{kine}) = M_{\Xi^-(\Sigma^-)} \quad (3.3)$$

が成立している．この値をイベントごとに求め, $p_{S,i}^{kine}$, ($i = 1, 2, \dots, N$) とする (N はイベント数)．すると, 式 (3.2) で表される損失関数の最小化は

$$Loss(p_S^{ML}, p_S^{kine}) \quad (3.4)$$

の最小化に帰着する．学習は, この損失関数を最小化するように進められる．

以上の手順は $p_B, \cos \theta$ の最適化でも全く同様に行うことができ、本研究では、このように、式 (3.2) で表されている損失関数を式 (3.4) のように、最適化対象の変数に関する損失関数に帰着させてから学習を行った。それは、以下2点の理由からである。

- 式 (1.2) で M_X が虚数になるのを防ぐ。 $p_B, \cos \theta, p_S$ の非現実的な組み合わせでは、 M_X は虚数になり得る。NN の学習時は、小節 2.2.4 で述べたようにパラメータが更新されるので、Loss 関数が虚数となれば、学習が進まなくなってしまう。
- 学習時間の短縮。学習時に用いる損失関数を単純化することで、学習時間の短縮が見込まれる。

ここで、 $p_S(X_S)$ の入力変数に X_B を加えてはならない理由を述べる。今、 p_S の教師データ p_S^{kine} は式 (3.3) で作成されている。もし p_S のモデルに X_B を追加すると、モデルが内部的に $p_B^{pre}, \cos \theta^{pre}$ を再現して、式 (3.3) を計算することで損失関数を非常に小さくする。これは実質的には $p_B, \cos \theta$ から式 (3.3) の運動学を計算した結果を出力することとなり、SKS での運動量観測量という意味から逸脱してしまう。その為、 p_S の入力変数は X_S のみにすべきであり、同様に p_B は X_B 、 $\cos \theta$ は X_C を入力変数とする必要がある。

3.4 学習の進み方

本研究では、前節 3.3 で実装したモデルについて、以下のようにその学習させる。

3.4.1 early stopping

学習回数については、early stopping という手法で決めている。early stopping とは、損失関数の収束を判断し、学習を終了する手法である。学習開始前に、学習に用いるデータの一部を検証データ (inner test data, validation data) として分割しておく。分割データ以外のデータを内部学習データ (inner train data) と呼び、学習 (モデルの内部パラメータ更新) は、inner train data に対して行われる。学習が 1 epoch 進むごとに、検証データに対する損失関数が計算される。学習を進めていき、検証データの損失関数が一定の epoch 数連続で減少しなければ、検証データの損失関数が最小値をとった状態のモデルを最終段階のモデルとして、学習を終了する。今回は、学習データの 1/10 を検証データとし、100 epoch 連続で減少しなければ、終了とする。図 3.7 に、1 度の学習で見られる損失関数の推移と early stopping の例を示す。

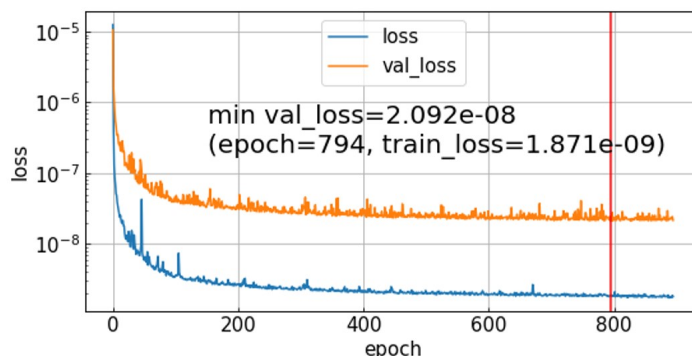


図 3.7: 損失関数の推移と early stopping の例．青 (loss) が学習に使用したデータ (inner train data) の損失関数の推移，橙 (val loss) が検証データ (validation data) の損失関数の推移を示す．この例では 794 epoch(赤) から 100 epoch 連続で val loss の改善が見られなかった為学習を終了し，794 epoch 目の状態を最終モデルとしている．

3.4.2 交差検証

1つのモデルに対しては，前節のように学習を終了させる．同じ学習を異なるデータに対して実行した時，同様に収束して，似た性能を示すかを確認する為に，性能評価には，4分割交差検証(小節 2.1.4 参照)を用いる．各 fold において，全体のうち 3/4 を学習に用い，残りの 1/4 をテストに用いる．データの分割の際には，データの種類の比 ($\Xi^- : \Sigma^-$) が等しくなるようにした．

ただし，従来解析のスペクトル図 3.2 と比較するため，各 fold のテスト結果の平均を取るのではなく，各 fold で得た欠損質量スペクトルを 1つのヒストグラムにまとめ，従来解析のスペクトルと比較する．このように機械学習で得た欠損質量スペクトルを，従来解析のスペクトルと同じように fitting し正規分布関数のピーク位置 μ と， σ を得る．ピーク位置の，PDG 値 [38] からのずれ

$$\langle M_{\Xi^-(\Sigma^-)} \rangle = \mu - M_{\Xi^-(\Sigma^-)}^{PDG} \quad (3.5)$$

と，ピーク幅

$$\sigma(M_{\Xi^-(\Sigma^-)}) = 2.35 \times \sigma \quad (3.6)$$

を指標としてモデルの性能を評価する．式 (3.6) は，FWHM で議論するために 2.35 倍している．この 2つの指標のうち特に重視するのは，分解能に対応する $\sigma(M_{\Xi^-(\Sigma^-)})$ で

ある．機械学習による最適化で，従来解析で考慮できていない相関を考慮できた場合， $\sigma(M_{\Xi^-}(\Sigma^-))$ は従来解析の値よりも小さくなる可能性がある．従来解析の結果である図 3.2 でこの 2 つの指標は表 3.3 となる．この表中 $\sigma(M_{\Xi^-}(\Sigma^-))$ より小さな値をとる運動量解析を最適化により目指す．

表 3.3: 従来解析の性能．単位は MeV/c^2 ．

	$\langle M_{\Xi^-} \rangle$	$\sigma(M_{\Xi^-})$	$\langle M_{\Sigma^-} \rangle$	$\sigma(M_{\Sigma^-})$
pre	-0.18 ± 0.05	5.45 ± 0.10	0.38 ± 0.03	7.12 ± 0.06

3.5 性能評価

本節では，前節 3.4 で述べた手法を様々な学習条件下で適用し，その依存性を調査する．調査する学習条件は，以下の通りである．

表 3.4: 調査する学習条件

学習条件	備考
学習データの種類	二種類のデータ (Ξ^- , Σ^-) をどう使用するか
学習データの統計量	学習に十分な統計量が存在しているか
M_{Ξ^-} , M_{Σ^-} カット	どの範囲を学習データに用いるか
隠れ層数	適した隠れ総数
損失関数	MSE か MAE か

本節では特に指定のない限り，表 3.5 で示すニューラルネットを用いた結果を述べる．以下，このモデルを `modell` と呼ぶ．

表 3.5: 本節で基本的に使用するニューラルネット (model1)

隠れ層の数	3
隠れ層のユニット数	64
隠れ層の活性化関数	ReLU
最適化アルゴリズム	Adam
batch size	64
学習回数	early stopping で決定 (3.4.1 節参照)
損失関数	MSE

また, 指定のない限り, 学習データの M_{Ξ^-} , M_{Σ^-} カットは $\mu \pm n\sigma$ としている (図 3.3 参照). このような M_{Ξ^-} , M_{Σ^-} カットを以下, $n\sigma$ cut と呼ぶ.

3.5.1 学習データの種類に対する依存性

1 変数 p_S のチューニングを考える. 使用モデルは model1, M_{Ξ^-} , M_{Σ^-} カットは 3σ cut とする. 素過程 $p(K^-, K^+) \Xi^-$, $p(K^-, \pi^+) \Sigma^-$ では, p_S の分布する領域が図 3.8 のように異なるので, M_{Ξ^-} を最適化する際には Ξ^- データ, M_{Σ^-} を最適化する際には Σ^- データのみを学習に使用すると, 対象領域の散乱運動量に絞って p_S のチューニングができ, Ξ^- と Σ^- の両方のデータを学習に使用した際より良い性能を示す可能性がある. 一方で, 異なる p_S 領域のデータも使用する事で学習に用いることのできるデータ数は増加する. ニューラルネットは一般に, 学習データ数が大量に必要な為, 異なる p_S 領域のデータであってもそこから全 p_S 領域に共通する情報が得られるのであれば, Ξ^- と Σ^- の両方を学習に使用の方が良い性能を示す可能性がある. そこで, 実際に学習に用いるデータを変えて, モデル性能を比較する.

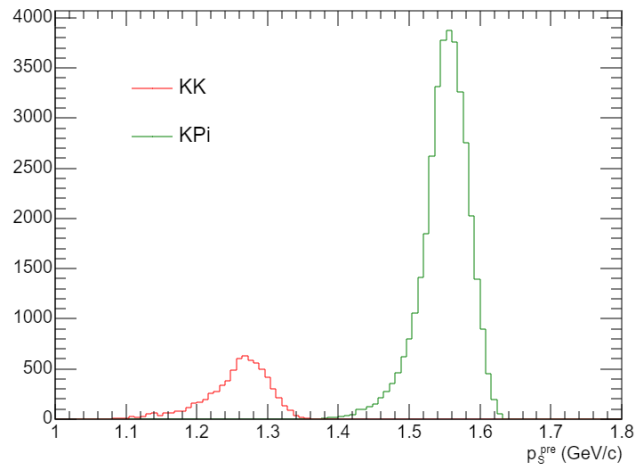


図 3.8: 欠損質量 3σ cut 下の p_S^{pre} . Ξ^- と Σ^- で散乱運動量領域は異なる .

表 3.6 に , 使用した学習データの種類とモデルの性能を示す .

表 3.6: p_S をチューニングした際の学習データの種類依存性 .

単位は MeV/c^2 . - は , 欠損質量のピークが得られない事を示す .

学習データの種類	$\langle M_{\Xi^-} \rangle$	$\sigma(M_{\Xi^-})$	$\langle M_{\Sigma^-} \rangle$	$\sigma(M_{\Sigma^-})$
Ξ^-	-0.18 ± 0.07	8.4 ± 0.2	-	-
Σ^-	-	-	-0.12 ± 0.03	8.15 ± 0.07
Ξ^- and Σ^-	-0.05 ± 0.06	6.7 ± 0.2	-0.22 ± 0.03	8.08 ± 0.07

また , 図 3.9 - 3.11 は , 表 3.6 で示したモデルで得られた Ξ^- , Σ^- の欠損質量を , 従来解析と比較したものである .

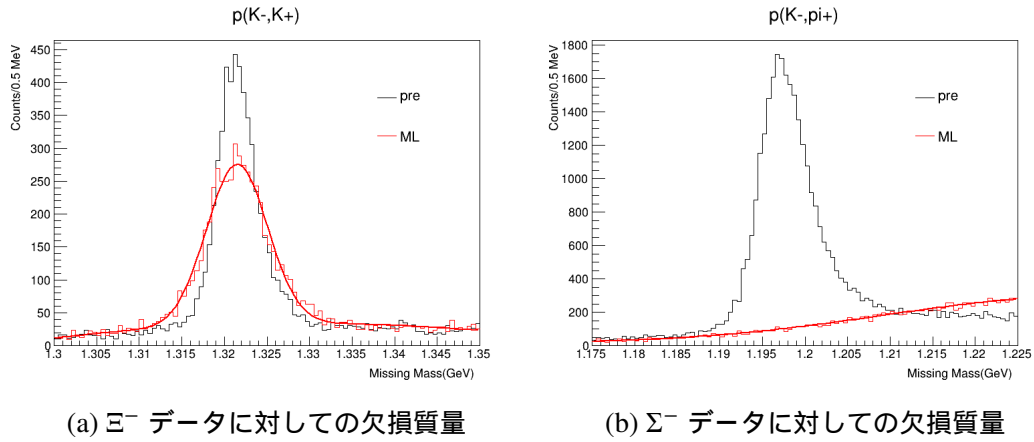


図 3.9: Ξ^- データのみを学習に用いた結果．黒ヒストグラムは従来解析である． M_{Σ^-} の欠損質量ピークは見られない．

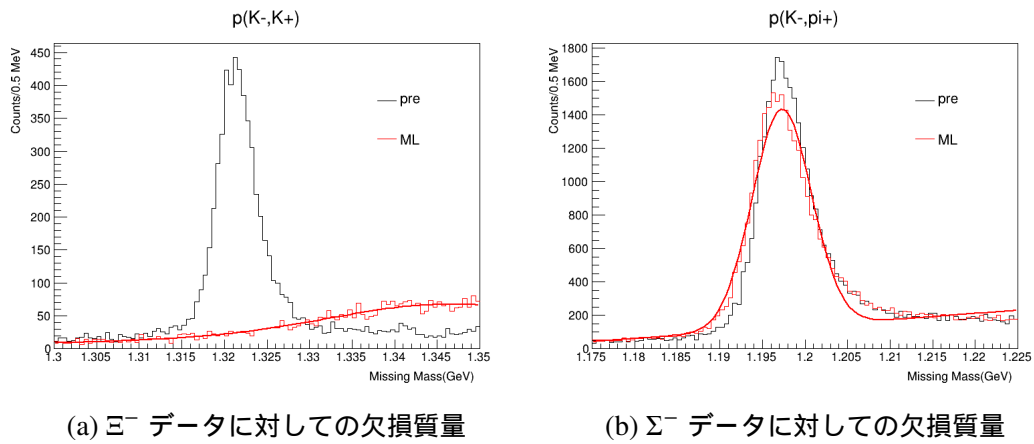


図 3.10: Σ^- データのみを学習に用いた結果．黒ヒストグラムは従来解析である． M_{Ξ^-} の欠損質量ピークは見られない．

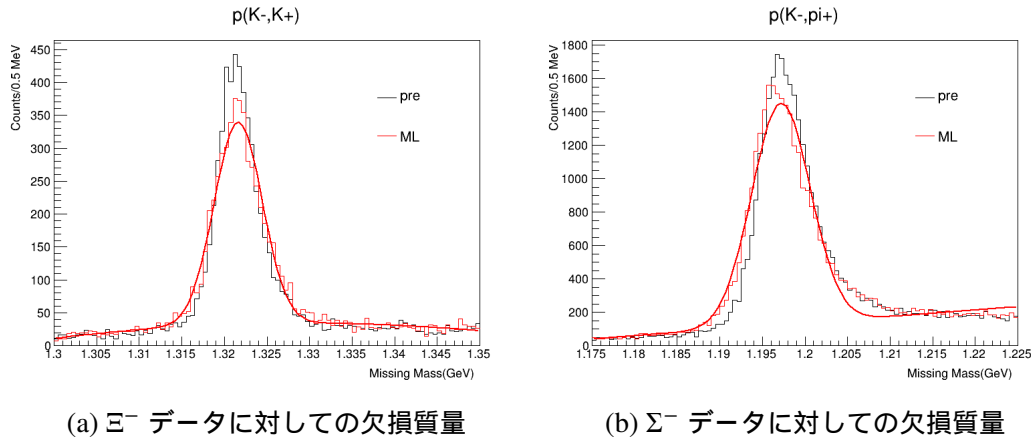
(a) Ξ^- データに対しての欠損質量(b) Σ^- データに対しての欠損質量

図 3.11: Ξ^- データと Σ^- データの両方を学習に用いた結果．黒ヒストグラムは従来解析である． M_{Ξ^-} と M_{Σ^-} 両方のピークが見られる．

Ξ^- データのみ、 Σ^- データのみを学習に用いた場合は、それぞれの運動量領域に限られた範囲で学習しているので、学習に用いていないデータに対しての運動量は適切に解析できず、結果として図 3.9b、3.10a のように、それぞれ、 M_{Σ^-} や M_{Ξ^-} の位置にピークは見られない。

表 3.6 において、分解能に対応する $\sigma(M_{\Xi^-})$ 、 $\sigma(M_{\Sigma^-})$ を比較すると、 M_{Ξ^-} 、 M_{Σ^-} とともに Ξ^- データと Σ^- データの両方を学習に用いた場合が最も小さくなった．特に、 M_{Ξ^-} については、その違いが顕著に見られる． $\langle M_{\Xi^-} \rangle$ も改善した．これより、 $p(K^-, K^+) \Xi^-$ の解析については、 Ξ^- データと Σ^- データの両方を学習に使用するのが良いと結論付けた．異なる p_S 領域のデータであるにも拘わらず、学習に使用した方が良い結果を示す理由として、学習データ数の依存性が強いことが考えられる．それは以下の理由からである．

今回学習に用いたデータ数は欠損質量 3σ cut の下で、表 3.7 のようであった． Ξ^- の学習データ数は Σ^- の 0.2 倍程度であることがわかる．表 3.6 において、 M_{Σ^-} については Σ^- データのみを使用した場合と両方のデータを使用した場合で顕著に差が現れなかったことから、 Σ^- データは Ξ^- データに比較すると、学習に十分な量のデータが存在していることが示唆される．

結果として、 $p(K^-, K^+) \Xi^-$ の解析については Ξ^- と Σ^- 両方のデータを用いるべきであるが、 $p(K^-, \pi^+) \Sigma^-$ については、明確な差がみられない．そこで以下、 $p(K^-, K^+) \Xi^-$ 、 $p(K^-, \pi^+) \Sigma^-$ 両方の結果を得られる利点から、 Ξ^- と Σ^- の両方を学習に用いて、他の依存性を調べる．

表 3.7: 欠損質量 $3\sigma_{cut}$ の下での各 fold での学習データ数 .

		fold1	fold2	fold3	fold4
inner train	Ξ^-	3620	3628	3561	3659
	Σ^-	18689	18829	18730	18823
valid	Ξ^-	393	409	410	391
	Σ^-	2005	2123	2013	2041

3.5.2 学習データの統計に対する依存性

前小節 3.5.1 のように、 Ξ^- について学習データの特徴が不足している一方で、 Σ^- は比較的十分存在していることが示唆された。本小節では、学習データを意図的に減らし、性能の変化からデータの統計に対しての依存性を調べる。

統計の依存性については、E70 実験での本手法の使用を検討する際、較正データ (CH₂ 標的データ) をどの程度必要とするかの手がかりにもなる。

前小節同様、 p_S を model1 でチューニングする。学習データは Ξ^- と Σ^- を両方使い、 M_{Ξ^-} 、 M_{Σ^-} カットは 3σ とする。モデルの学習前に、一定の割合で学習データをランダムにサンプリングする。サンプリングの割合は、100% - 20% まで 20% ずつ、さらに 20% - 4% まで 4% ずつ変化させ、性能を比較する。統計の詳細は A.1 に示す。図 3.13 に、得られた欠損質量スペクトルを示す。

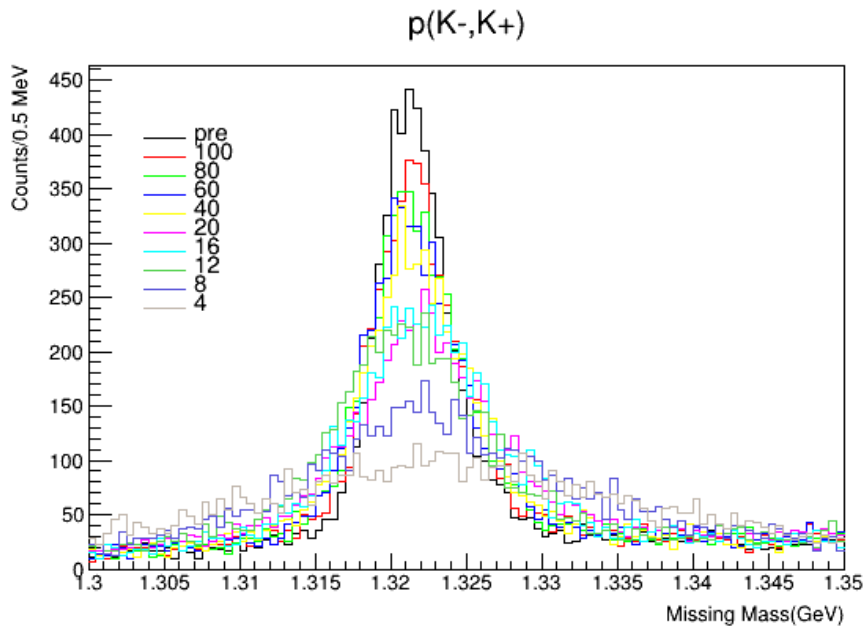
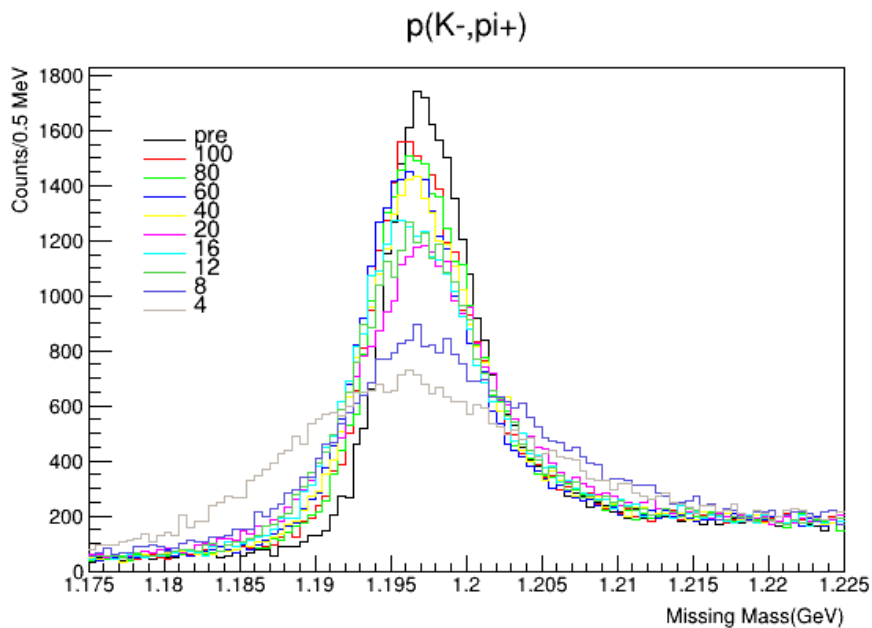
(a) Ξ^- データに対するの欠損質量(b) Σ^- データに対するの欠損質量

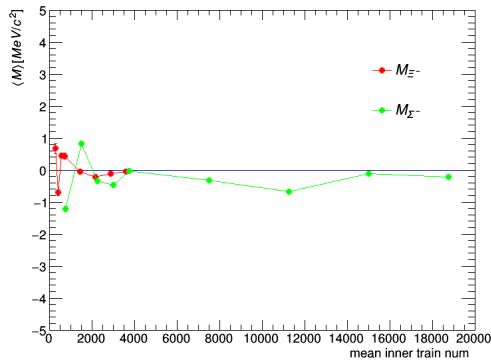
図 3.12: p_S チューニングの学習データ統計依存性．凡例の数字は，サンプリング割合 (%) を示す．サンプリング割合の増加に従い，欠損質量ピークが明確に見られるようになる．

図 3.13 を fit して得られた結果を表 3.8 に示す．fit の詳細は図 A.1-A.8 に示した．表中の平均 inner train 数は，1fold あたりの inner train data 数の平均をとったものである（表 A.1 参照）．また，図 3.13 は， $\langle M_{\Xi^-} \rangle, \sigma(M_{\Xi^-})$ の平均 inner train 数依存性をプロットしたものである．

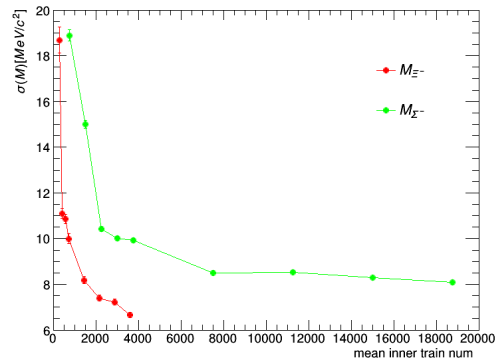
表 3.8: p_S チューニングの学習データ統計依存性．

単位は MeV/c^2 ． - は適切な fit が出来なかった事を示す．

学習データ統計 サンプリング割合 (%)	平均 inner train 数		$\langle M_{\Xi^-} \rangle$	$\sigma(M_{\Xi^-})$	$\langle M_{\Sigma^-} \rangle$	$\sigma(M_{\Sigma^-})$
	Ξ^-	Σ^-				
100	3617	18768	-0.05 ± 0.06	6.7 ± 0.2	-0.22 ± 0.03	8.08 ± 0.07
80	2894	15014	-0.11 ± 0.06	7.2 ± 0.2	-0.10 ± 0.03	8.29 ± 0.07
60	2170	11260	-0.21 ± 0.06	7.4 ± 0.2	-0.67 ± 0.04	8.53 ± 0.07
40	1447	7507	-0.05 ± 0.07	8.2 ± 0.2	-0.33 ± 0.04	8.50 ± 0.08
20	724	3753	0.43 ± 0.10	10.0 ± 0.3	-0.02 ± 0.05	9.92 ± 0.10
16	579	3003	0.44 ± 0.09	10.9 ± 0.2	-0.47 ± 0.04	10.01 ± 0.09
12	434	2252	-0.71 ± 0.09	11.1 ± 0.3	-0.35 ± 0.04	10.4 ± 0.09
8	289	1501	0.7 ± 0.2	18.7 ± 0.6	0.83 ± 0.07	15.0 ± 0.2
4	145	751	-	-	-1.21 ± 0.08	18.9 ± 0.3



(a) $\langle M_{\Xi^-} - \Sigma^- \rangle$ の inner train 数依存性



(b) $\sigma(M_{\Xi^-} - \Sigma^-)$ の inner train 数依存性

図 3.13: p_S チューニングの inner train 数依存性

図 3.13b より， $\sigma(M_{\Sigma^-})$ は約 7000 の学習データがあれば十分であることが推察される．また， $\sigma(M_{\Xi^-})$ は， $\lesssim 3500$ の統計領域で変化の傾向が $\sigma(M_{\Sigma^-})$ と似ている．これから， $\sigma(M_{\Xi^-})$ についても 7000 程度の統計が必要であることが示唆される．これは，本小節冒

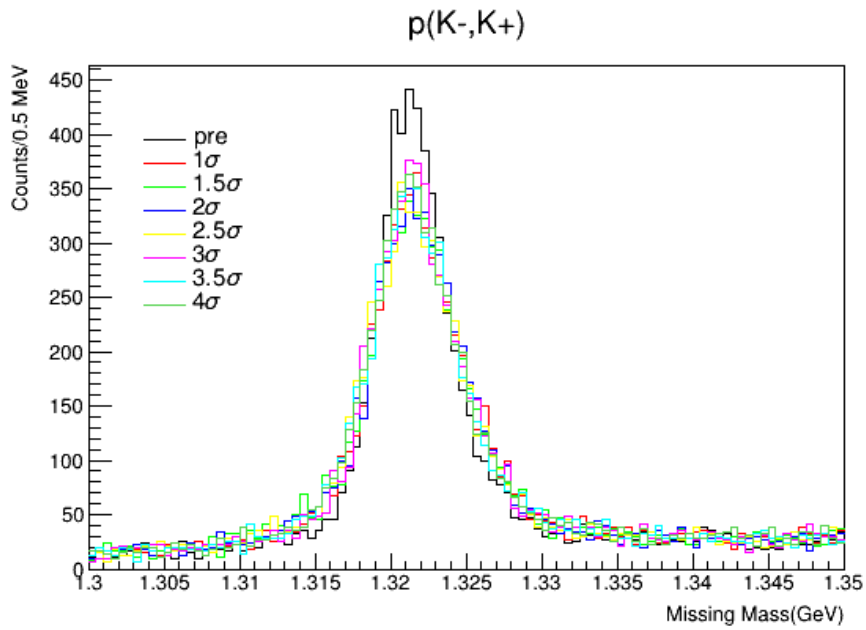
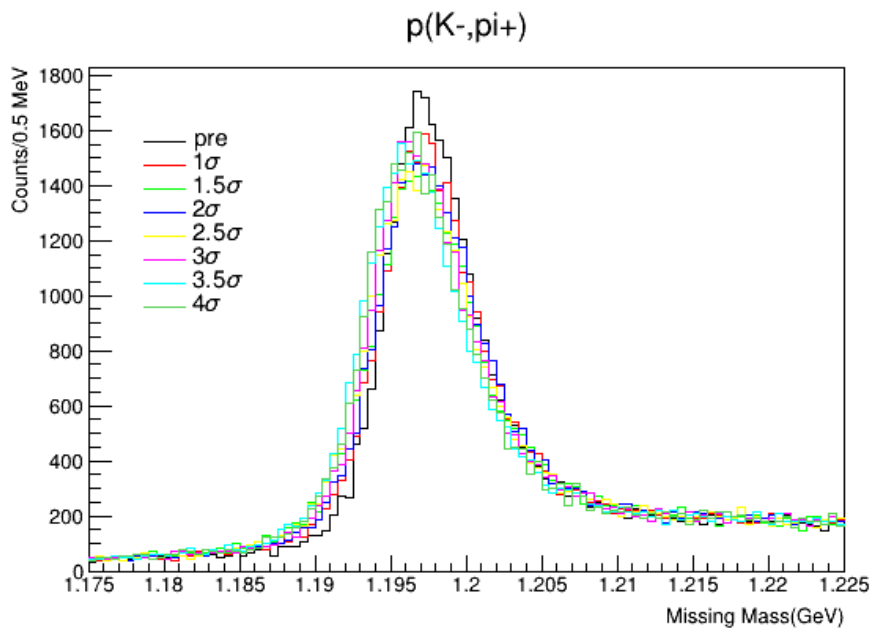
頭で述べた推察を支持する結果となっている。

3.5.3 M_{Ξ^-} , M_{Σ^-} カット依存性

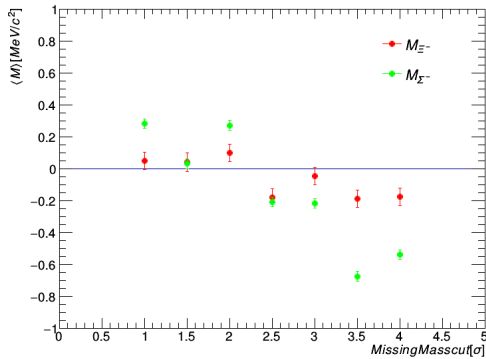
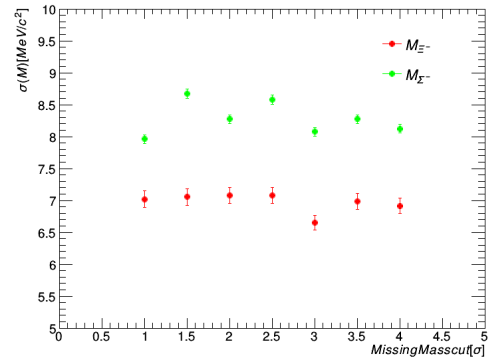
3.1 節で述べたように, M_{Ξ^-} , M_{Σ^-} カットの適切な範囲はデータの純度と統計のトレードオフで決まると考えられる. p_S のチューニング, model1, 学習に Ξ^- と Σ^- の両方を用いた場合において, M_{Ξ^-} , M_{Σ^-} カットの幅を $1\sigma - 4\sigma$ の範囲で変更した場合の結果を表 3.9, 図 3.14 に示す. それぞれのカットにおける学習データの統計は表 A.2 に示した.

表 3.9: p_S をチューニングした際の M_{Ξ^-} , M_{Σ^-} カット依存性. 単位は MeV/c^2 .

M_{Ξ^-} , M_{Σ^-} カット範囲 [σ]	$\langle M_{\Xi^-} \rangle$	$\sigma(M_{\Xi^-})$	$\langle M_{\Sigma^-} \rangle$	$\sigma(M_{\Sigma^-})$
1	0.05 ± 0.06	7.0 ± 0.2	0.28 ± 0.03	7.96 ± 0.07
1.5	0.04 ± 0.06	7.1 ± 0.2	0.03 ± 0.04	8.67 ± 0.08
2	0.10 ± 0.06	7.1 ± 0.2	0.27 ± 0.03	8.27 ± 0.07
2.5	-0.18 ± 0.06	7.1 ± 0.2	-0.21 ± 0.04	8.57 ± 0.07
3	-0.05 ± 0.06	6.7 ± 0.2	-0.22 ± 0.03	8.08 ± 0.07
3.5	-0.19 ± 0.06	7.0 ± 0.2	-0.68 ± 0.04	8.27 ± 0.07
4	-0.17 ± 0.06	6.9 ± 0.2	-0.54 ± 0.03	8.12 ± 0.07

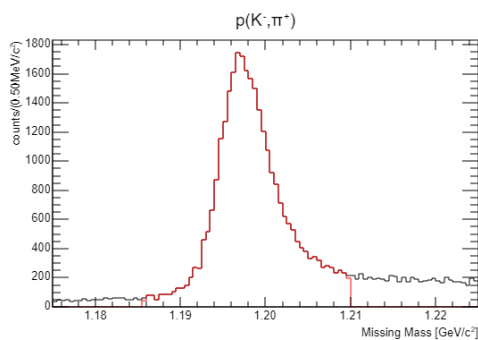
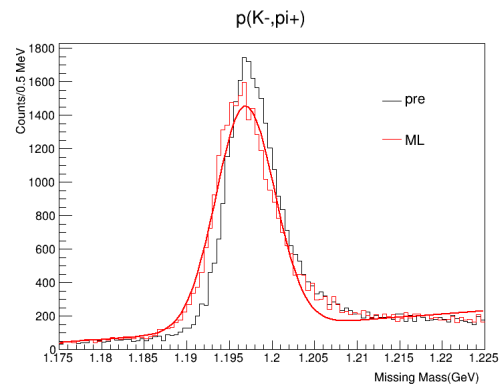
(a) Ξ^- データに対しての欠損質量(b) Σ^- データに対しての欠損質量図 3.14: p_S チューニングの M_{Ξ^-} , M_{Σ^-} カット依存性

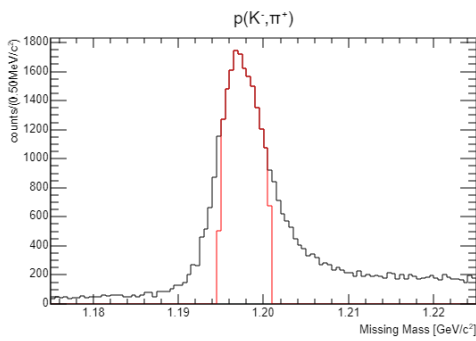
また, 図 3.15a, 3.15b は, $\langle M_{\Xi^-(\Sigma^-)} \rangle$ と, $\sigma(M_{\Xi^-(\Sigma^-)})$ の, M_{Ξ^-} , M_{Σ^-} カット依存性をプロットした図である.

(a) $\langle M_{\Xi^-(\Sigma^-)} \rangle$ の M_{Ξ^-} , M_{Σ^-} カット依存性 .(b) $\sigma(M_{\Xi^-(\Sigma^-)})$ の M_{Ξ^-} , M_{Σ^-} カット依存性 .

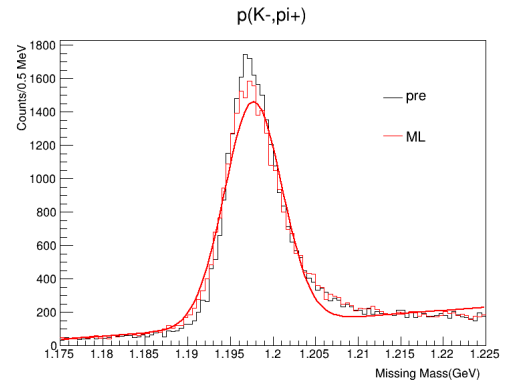
3.5 σ 以上の範囲では、 $\langle M_{\Sigma^-} \rangle$ が負の方向に 0.6 MeV/ c^2 ほどずれていて、その大きさは $\langle M_{\Xi^-} \rangle$ より大きい。これは、 M_{Σ^-} の分布が非対称であることが要因として考えられる。図 3.16a に、 M_{Ξ^-} , M_{Σ^-} カット 4σ をかけた時の $p(K^-, \pi^+)\Sigma^-$ の欠損質量分布を示す。また、図 3.16b に、その条件下で学習した際に Σ^- データに対して得た欠損質量を示す。MSE を損失関数とする学習を通じて、図 3.16a の赤の領域が正規分布に近づく。欠損質量の大きい領域に学習データが多く存在することから、学習後のピーク位置は負の方向にシフトすると考えられる。逆に、図 3.17a, 3.17b で見られる 1σ cut の場合は、欠損質量の小さな領域に学習データが多く存在することから、学習後のピーク位置は正の方向にシフトすると考えられる。

$\sigma(M_{\Xi^-(\Sigma^-)})$ が小さく、 $\langle M_{\Xi^-(\Sigma^-)} \rangle$ が 0 に近いことから、 M_{Ξ^-} , M_{Σ^-} カットは 3σ が適切であると結論づけた。

(a) 4σ cut をかけた時の $p(K^-, \pi^+)$ の欠損質量分布 (赤) .(b) 4σ cut 下で学習した際の Σ^- データに対しての欠損質量



(a) 1σ cut をかけた時の $p(K^-, \pi^+)$ の欠損質量分布 (赤).



(b) 1σ cut 下で学習した際の Σ^- データに対しての欠損質量

3.5.4 隠れ層数依存性

NN は、隠れ層の増加に従い、表現力が指数的に増加することが知られている [39]。一方で、NN が複雑であることは、過学習が起きやすくなる一因ともなる。 p_S のチューニング、 M_{Ξ^-} 、 M_{Σ^-} カット 3σ 、学習に Ξ^- と Σ^- の両方を用いた場合について、model1 の隠れ層数を変化させた場合の結果を表 3.10、図 3.18 に示す。

表 3.10: p_S をチューニングした際の隠れ層数依存性。単位は MeV/c^2 。

隠れ層数	$\langle M_{\Xi^-} \rangle$	$\sigma(M_{\Xi^-})$	$\langle M_{\Sigma^-} \rangle$	$\sigma(M_{\Sigma^-})$
1	-0.20 ± 0.05	6.4 ± 0.2	-0.18 ± 0.03	8.11 ± 0.07
3	-0.05 ± 0.06	6.7 ± 0.2	-0.22 ± 0.03	8.08 ± 0.07
5	0.03 ± 0.06	6.8 ± 0.2	0.05 ± 0.04	8.36 ± 0.06
7	-0.01 ± 0.06	6.9 ± 0.2	-0.29 ± 0.03	8.16 ± 0.06
9	-0.09 ± 0.06	6.9 ± 0.2	-0.15 ± 0.03	8.27 ± 0.06

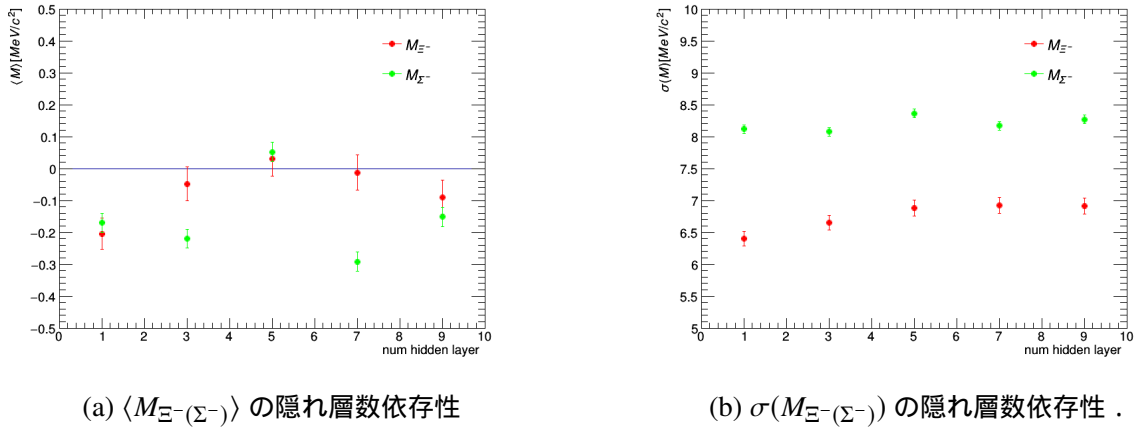
図 3.18: p_S チューニングの隠れ層数依存性

図 3.18b より，隠れ層数が増加するに従い， $\sigma(M_{\Xi^-})$ も僅かに増加傾向にある．これは， M_{Ξ^-} のチューニングについては，僅かな隠れ層数で十分であることを示している．一方で， M_{Σ^-} については M_{Ξ^-} と比較して，増加傾向は見られないことから， M_{Ξ^-} はデータ数の不足の為に過学習が行っている可能性がある．

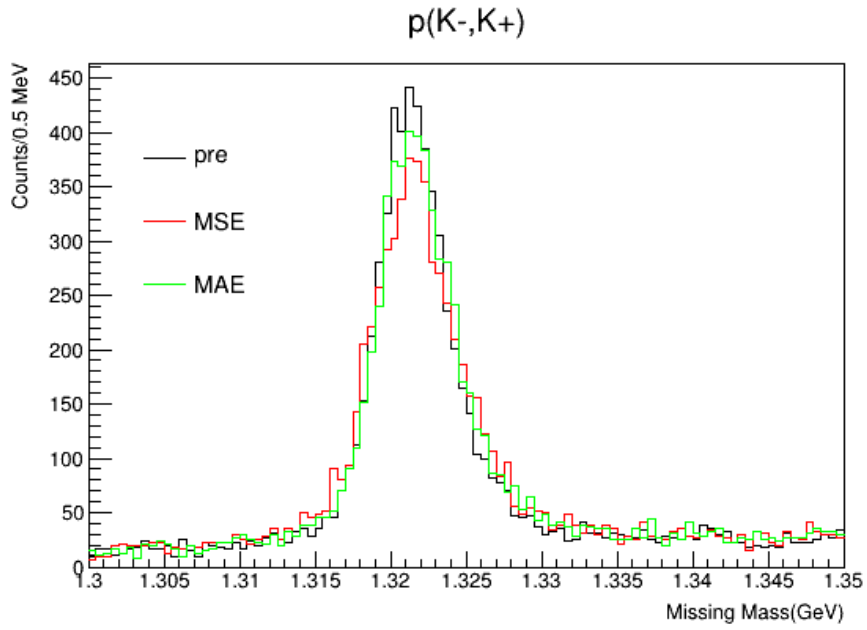
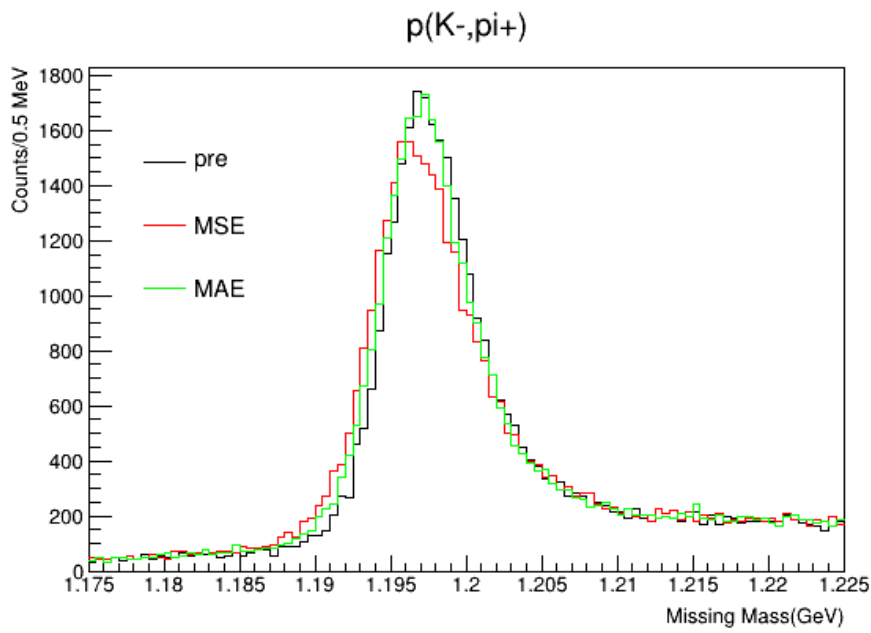
$\sigma(M)$ の大きさから，隠れ層の数は 1 層もしくは 3 層が適していると考えられる．

3.5.5 損失関数依存性

回帰問題に使用される代表的な損失関数として，MSE(式 2.2) と MAE(式 2.3) が存在する． p_S のチューニング， M_{Ξ^-} ， M_{Σ^-} カット 3σ ，学習に Ξ^- と Σ^- の両方を用いた場合，表 3.5 で表される model1 に対して，損失関数を MAE とした際の結果を比較することで，損失関数に対する結果の依存性を評価する．表 3.11，図 3.19, 3.20 にその結果を示す．

表 3.11: p_S をチューニングした際の損失関数依存性．単位は MeV/c² .

損失関数	$\langle M_{\Xi^-} \rangle$	$\sigma(M_{\Xi^-})$	$\langle M_{\Sigma^-} \rangle$	$\sigma(M_{\Sigma^-})$
MSE	-0.05 ± 0.06	6.7 ± 0.2	-0.22 ± 0.03	8.08 ± 0.07
MAE	0.03 ± 0.05	5.9 ± 0.2	0.01 ± 0.03	7.06 ± 0.06

(a) Ξ^- データに対するの欠損質量(b) Σ^- データに対するの欠損質量図 3.19: p_S チューニングの損失関数依存性 .

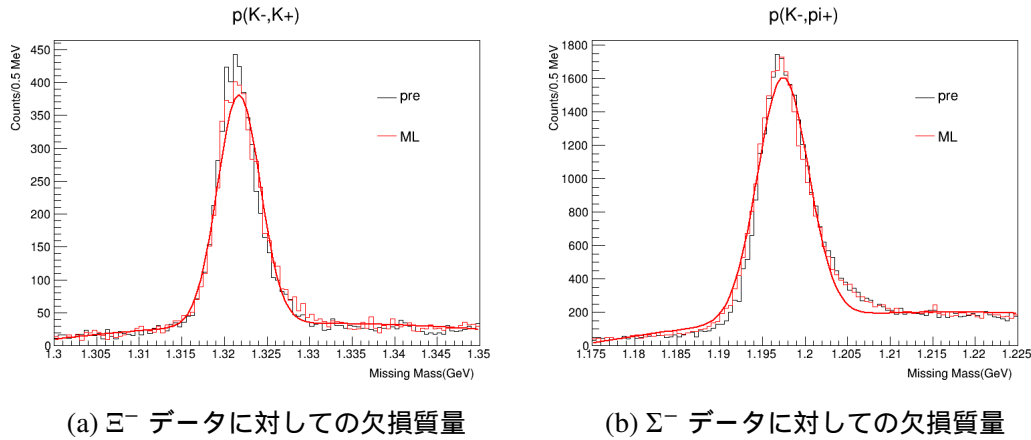


図 3.20: MAE を使用した場合の欠損質量

表 3.11 の全ての指標で、MSE より MAE を用いた方が良い結果となった。これより、損失関数として MAE を使用するのが適切だと結論付けた。現在のモデル評価手法は、ピーク fit により行っている為、MAE の外れ値の影響を受けにくいという特徴が有利にはたらいっていると推察できる。MAE での $\sigma(M_{\Sigma^-})$ に関しては、 $7.06 \pm 0.06 \text{ MeV}/c^2$ となり、従来解析 (表 3.3) での分解能 $7.12 \pm 0.06 \text{ MeV}/c^2$ に到達した。 $\sigma(M_{\Xi^-})$ は、小節 3.5.2 で見られたように、学習データが不足している為に、従来解析に届いていないと考えられる。

3.5.6 1 変数 p_S チューニングのまとめ

前小節までに考慮した依存性より、NN を用いて 1 変数 p_S のチューニングをする際は、表 3.12 の条件下で行うことが好ましいと推察される。

表 3.12: 1 変数 p_S のチューニングにおいて、依存性の調査より定まった学習条件

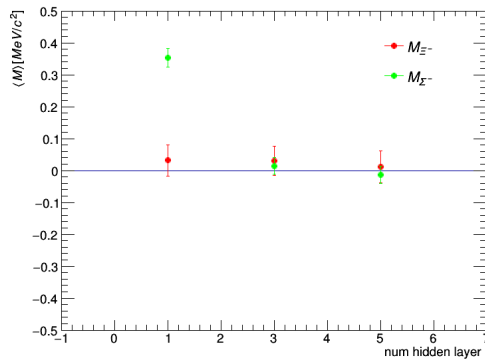
考慮した依存性	採用条件	備考
学習データの種類	Ξ^- and Σ^-	Ξ^- のデータ不足を Σ^- で補う
M_{Ξ^-} , M_{Σ^-} カット	3σ	
隠れ層数	1 or 3	MSE で検討した
損失関数	MAE	外れ値の影響を受けにくい

ここで、隠れ層数の依存性は損失関数を MSE として検討したものであった。改めて、

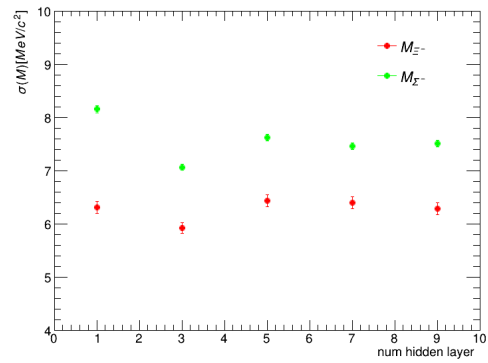
損失関数として MAE を使用した場合に，隠れ総数はどちらを採用すべきか検討すると，表 3.13，図 3.21 のような結果が得られた．

表 3.13: p_S をチューニング，損失関数 MAE の場合の隠れ層数依存性．単位は MeV/c^2 ．

隠れ層数	$\langle M_{\Xi^-} \rangle$	$\sigma(M_{\Xi^-})$	$\langle M_{\Sigma^-} \rangle$	$\sigma(M_{\Sigma^-})$
1	0.03 ± 0.05	6.3 ± 0.2	0.35 ± 0.03	8.15 ± 0.07
3	0.03 ± 0.05	5.9 ± 0.2	0.01 ± 0.03	7.06 ± 0.06
5	0.01 ± 0.06	6.4 ± 0.2	-0.01 ± 0.03	7.62 ± 0.07



(a) $\langle M_{\Xi^-(\Sigma^-)} \rangle$ の隠れ層数依存性 (損失関数 MAE)



(b) $\sigma(M_{\Xi^-(\Sigma^-)})$ の隠れ層数依存性 (損失関数 MAE)

図 3.21: p_S チューニングの隠れ層数依存性

$\sigma(M)$ が最も小さく， $\langle M \rangle$ も 0 に近いので，隠れ総数は 3 が適していると結論づけた．よって，改めて p_S の 1 変数チューニングにおいては，表 3.12 で隠れ総数 3 の条件下で行うことが適切だと推察され，その場合の結果は表 3.13 のようになった．

3.5.7 チューニング変数依存性

小節 3.3.3 で述べたように， p_S 以外にも p_B や $\cos \theta$ をチューニングすることも考えられる．現在， p_S についてのチューニングを考えた理由として，1.4 節で示していたように，E05 実験においては Δ_S の分解能への寄与が Δ_B, Δ_θ より大きいことが上げられる． p_B や $\cos \theta$ のチューニングでどれほどの性能を示すかを，3.5.6 節で同定した条件と同じ手法を

$p_B, \cos \theta$ に対して用いて評価し、最適化対象とする変数に対しての依存性を評価する。表 3.14 にその結果を示す。

$\sigma(M_{\Xi^-})$ に関しては、 p_B をチューニングした結果が最も良い性能を示した。 p_B に関しては、 M_{Ξ^-} と M_{Σ^-} で分布する領域は双方ともに 1.8 GeV/c 付近である。そのため、 M_{Ξ^-} の学習データの統計が不足していることが p_S の最適化では示唆されていたが、 p_B の場合は M_{Σ^-} の情報が M_{Ξ^-} の情報をより補っていることが示唆される。

一方で、 $\cos \theta$ に関しては、 $\Delta\theta$ が他の 2 変数に比較して小さいことから、良い性能を示さないという直感と合った結果となった。

表 3.14: チューニング変数依存性。単位は MeV/c^2 。

変数	$\langle M_{\Xi^-} \rangle$	$\sigma(M_{\Xi^-})$	$\langle M_{\Sigma^-} \rangle$	$\sigma(M_{\Sigma^-})$
p_B	-0.01 ± 0.05	5.4 ± 0.1	-0.12 ± 0.03	7.20 ± 0.06
p_S	0.03 ± 0.05	5.9 ± 0.2	0.01 ± 0.03	7.06 ± 0.06
$\cos \theta$	-0.09 ± 0.05	5.9 ± 0.2	-0.24 ± 0.03	7.81 ± 0.07

第 4 章

まとめと今後の展望

本研究では、E05 実験データに対して様々な条件下で機械学習を適用することで、 Ξ ハイパー核分光実験における機械学習を用いた運動量解析の最適化の実現可能性を検討した。

結果として、素過程 $p(K^-, K^+)\Xi^-$, $p(K^-, \pi^+)\Sigma^-$ を学習データとして用いることで、機械学習を用いた解析により、欠損質量ピークが観測された。さらに、学習条件に対する性能の依存性を調査することで、 M_{Ξ^-} , M_{Σ^-} カット 3σ をかけて学習し、 M_{Σ^-} に対して、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。これにより、最適な機械学習の条件について、一定の適用可能範囲が理解された。

また、 M_{Ξ^-} に関しては、学習に必要な統計が不足していることが判明した。現在、学習に用いた統計は 3500 程度であるが、7000 程度の統計が必要であると示唆される。倍の学習データが必要であるとすると、今回実施した 4 分割交差検証を 8 分割にすることによって学習データの統計を 7000 程度まで引き上げた際の性能を評価することが考えられる。これは今後の課題である。

また、本研究では散乱運動量 p_S を最適化する機械学習を主に行った。しかし、入射運動量 p_B や散乱角 θ に対しての最適化を行う際には、本来それぞれのモデル構成下で、今回 p_S に対して調査したように学習条件に対しての依存性を調査すべきである。

また、本研究で使用したデータは素過程として扱う学習データの中に常に ^{12}C 由来のバックグラウンドが混入していた。今回 CH_2 で学習させたモデルを炭素標的データに対して適用することで、直接的に ^{12}C の混入の寄与を調べる必要がある。

以上のような評価を通じて、 M_{Ξ^-} に対しても従来解析と同等かそれ以上の分解能に到達できるか調査することが今後の大きな課題である。

謝辞

本研究は多くの方のご助力により成立しています。指導教官である永江知文教授には、ハイパー核分光への機械学習の適用という非常に興味深いテーマを頂きました。また、右も左もわからない私に、研究の進め方、発表方法などの根本的な部分を教えてくださいました。勉強不足である私の相談や質問にも、グループに配属が決まってから修論提出の直前に至るまで何度でも非常に丁寧に、親身に時間を割いて頂き大変感謝しております。

機械学習の適用にあたり、北里大学の長谷川智之教授には、当初とりかかれていた計算の知見をご共有頂き、さらに研究の方向性のご助言もして頂きました。

また、後神利志助教には、物理から機械学習に至るまで幅広くご指導頂きました。研究に行き詰まる度に適切なご助言を頂き、何度も相談にのって頂きました。また、進捗状況のみならず心身にも気にかけてくださいました。ありがとうございました。

今回、E05 実験データを扱うにあたり、金築俊輔さんは必要なデータをご提供くださったり、非常に丁寧に長時間、長期間にわたり、解析手法のフォローをお手伝いくださいました。また、研究を進める上で起きた不明点は何度も積極的に相談にのって頂きました。

原田健志さんには、同研究グループの年齢の近い先輩として日頃から実験や研究活動の知識を数多く頂きました。特に J-PARC 出張の際には、経験が浅い私を数多くフォロー頂きました。七村拓野さんには、日々のミーティングや J-PARC での実験、生活など多くの場面でお世話になりました。また、鈴木一輝さんにはミーティングで数多くのご助言を頂きました。

同期の延與紫世さん、片山一樹くん、高浦雄大くん、豊田峻史くん、中須賀さとみさん、土方佑斗くんとは他愛もないことから研究活動に至るまで気軽に話すことができ、楽しい時間を過ごせました。

研究室の成木恵准教授、銭廣十三准教授、村上哲也講師、先輩方は研究の進捗を気にかけてくださいました。

最後に、今まで私を支えて頂いた両親、家族、京都での六年間の学生生活を支えて頂い

た祖父母，楽しい時間をともに過ごし，励ましてくれた友人に深く感謝の意を申し上げます。

付録 A

学習時データ詳細

本文中で示さなかったものの、[3.5](#) 節の検証において残すべき補助的なデータを本章に示す。

[3.5.2](#) では、学習データ数を、欠損質量カット 3σ の下に意図的に一定の比率で変更させた。そのデータ数を [A.1](#) に示す。表中の sampling rate 100% については、[表 3.7](#) と等しい。

表 A.1: 小節 3.5.2 で検討したデータ数の詳細

sampling rate(%)			fold1	fold2	fold3	fold4	mean
100	inner train	KK	3620	3628	3561	3659	3617
		Kpi	18689	18829	18730	18823	18767.75
	valid	KK	393	409	410	391	400.75
		KPi	2005	2123	2013	2041	2045.5
80	inner train	KK	2896	2902	2849	2927	2893.5
		Kpi	14951	15063	14983	15058	15013.75
	valid	KK	314	327	328	313	320.5
		KPi	1604	1698	1610	1632	1636
60	inner train	KK	2172	2177	2137	2195	2170.25
		Kpi	11213	11297	11237	11294	11260.25
	valid	KK	236	245	246	235	240.5
		KPi	1202	1274	1207	1224	1226.75
40	inner train	KK	1448	1451	1424	1463	1446.5
		Kpi	7475	7531	7492	7529	7506.75
	valid	KK	157	163	164	156	160
		KPi	802	849	805	816	818
20	inner train	KK	724	726	712	732	723.5
		Kpi	3737	3765	3746	3764	3753
	valid	KK	79	82	82	78	80.25
		KPi	400	424	402	408	408.5
16	inner train	KK	579	580	570	585	578.5
		Kpi	2990	3013	2996	3012	3002.75
	valid	KK	63	65	65	63	64
		KPi	320	340	322	326	327
12	inner train	KK	434	435	427	439	433.75
		Kpi	2243	2259	2247	2258	2251.75
	valid	KK	47	49	49	47	48
		KPi	240	254	241	244	244.75
8	inner train	KK	289	290	285	293	289.25
		Kpi	1495	1506	1498	1505	1501
	valid	KK	31	33	33	31	32
		KPi	160	169	160	163	163
4	inner train	KK	145	145	142	146	144.5
		Kpi	747	753	749	753	750.5
	valid	KK	16	16	16	16	16
		KPi	79	85	80	81	81.25

小節 3.5.2 で検討したモデルで得られた欠損質量スペクトルと、それぞれの fit の様子を
 示す。尚、学習データを全て用いた際 (学習データ割合 100%) の結果は、3.5.1 で示した
 通りである。

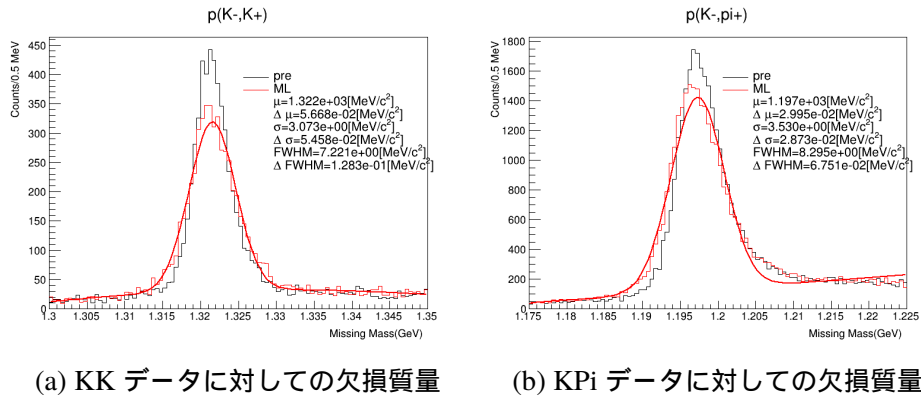


図 A.1: 学習データ 80% を用いた際の欠損質量スペクトル

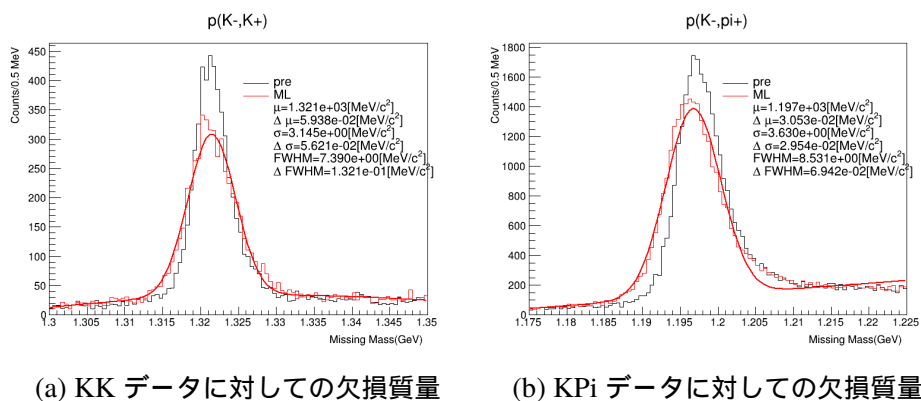
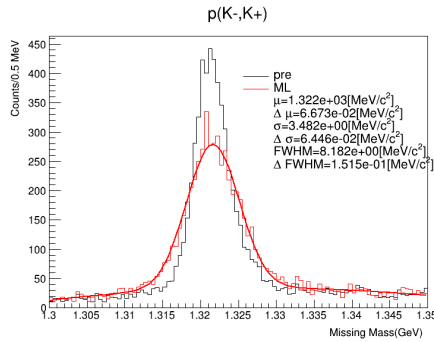
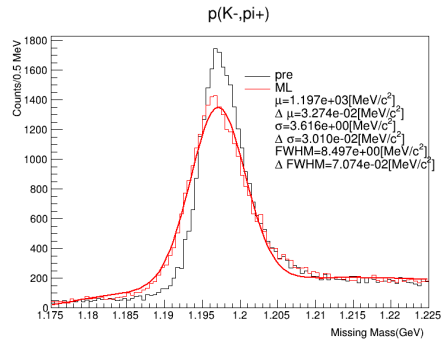


図 A.2: 学習データ 60% を用いた際の欠損質量スペクトル

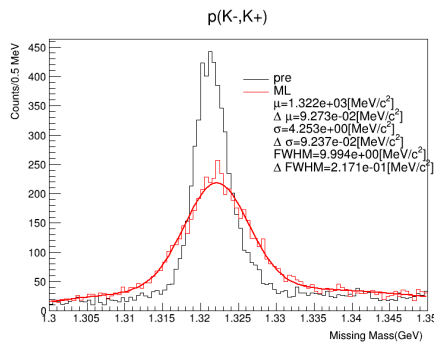


(a) KK データに対する欠損質量

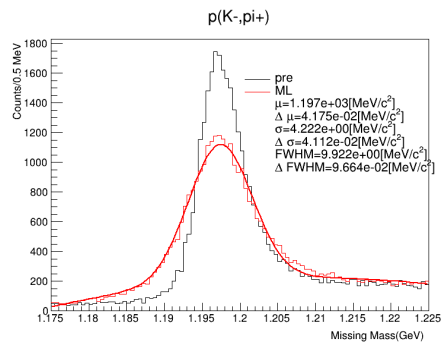


(b) KPi データに対する欠損質量

図 A.3: 学習データ 40% を用いた際の欠損質量スペクトル

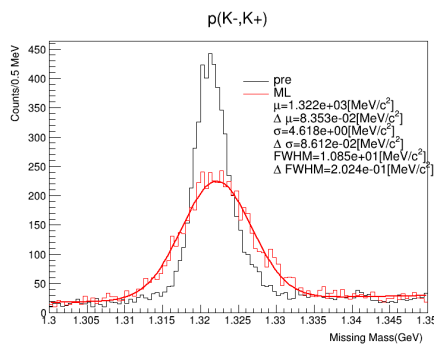


(a) KK データに対する欠損質量

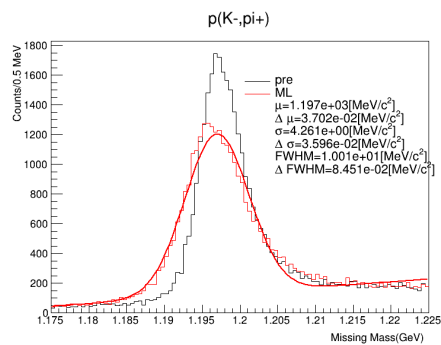


(b) KPi データに対する欠損質量

図 A.4: 学習データ 20% を用いた際の欠損質量スペクトル

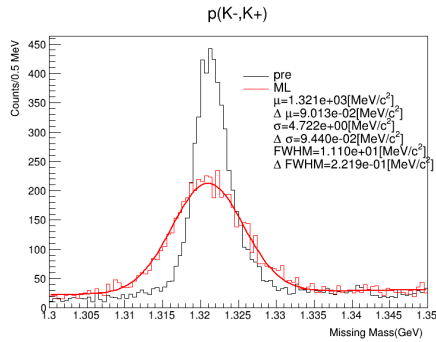


(a) KK データに対する欠損質量

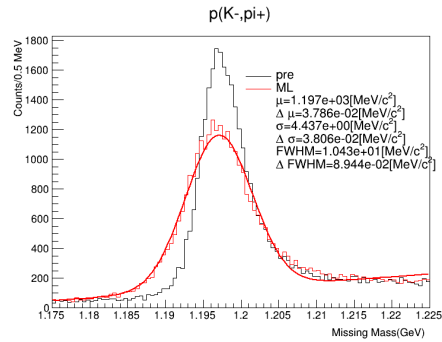


(b) KPi データに対する欠損質量

図 A.5: 学習データ 16% を用いた際の欠損質量スペクトル

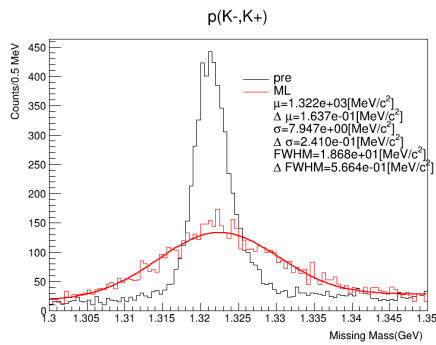


(a) KK データに対する欠損質量

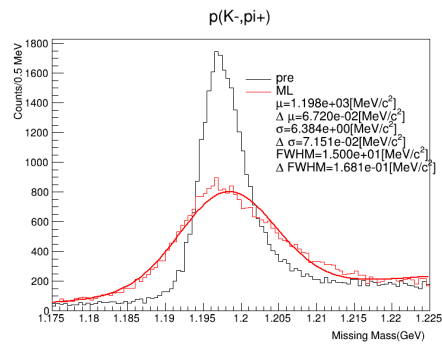


(b) KPi データに対する欠損質量

図 A.6: 学習データ 12% を用いた際の欠損質量スペクトル

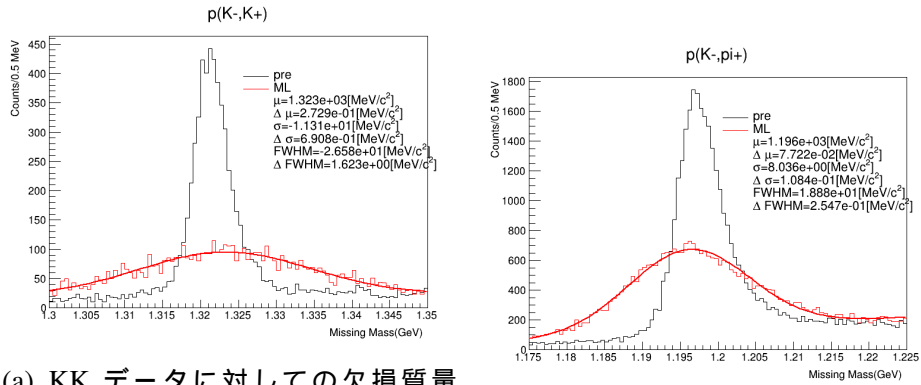


(a) KK データに対する欠損質量



(b) KPi データに対する欠損質量

図 A.7: 学習データ 8% を用いた際の欠損質量スペクトル



(a) KK データに対するの欠損質量 . FWHM が負となり，従来解析の fit と (b) KPi データに対するの欠損質量 同じ手法では適切に fit できない .

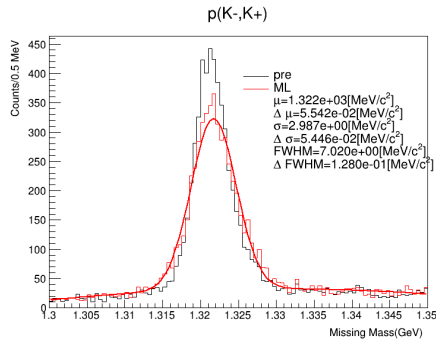
図 A.8: 学習データ 4% を用いた際の欠損質量スペクトル

3.5.3 では，学習データに対する欠損質量カット範囲を $1\sigma - 4\sigma$ で変えた．それぞれの場合のデータ数を A.2 に示す．カット範囲が 3σ の場合は，表 3.7 と等しい．

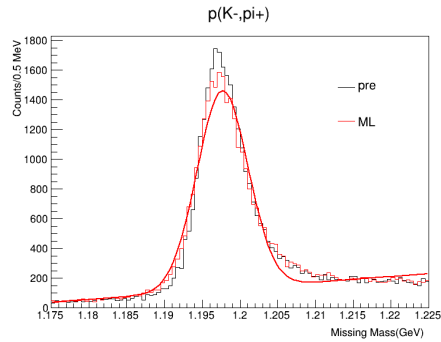
表 A.2: 小節 3.5.3 で検討したデータ数の詳細

MM cut sigma			fold1	fold2	fold3	fold4	mean
1	inner train	KK	2279	2329	2277	2330	2303.75
		Kpi	11615	11674	11651	11654	11648.5
	valid	KK	267	242	251	249	252.25
		KPi	1233	1322	1250	1303	1277
1.5	inner train	KK	2905	2913	2866	2956	2910
		Kpi	14769	14901	14799	14919	14847
	valid	KK	326	321	323	323	323.25
		KPi	1587	1671	1609	1637	1626
2	inner train	KK	3241	3248	3199	3302	3247.5
		Kpi	16622	16768	16670	16818	16719.5
	valid	KK	354	369	372	357	363
		KPi	1777	1883	1814	1835	1827.25
2.5	inner train	KK	3470	3485	3412	3520	3471.75
		Kpi	17775	17921	17839	17937	17868
	valid	KK	381	391	398	381	387.75
		KPi	1918	2022	1929	1961	1957.5
3	inner train	KK	3620	3628	3561	3659	3617
		Kpi	18689	18829	18730	18823	18767.75
	valid	KK	393	409	410	391	400.75
		KPi	2005	2123	2013	2041	2045.5
3.5	inner train	KK	3750	3761	3677	3787	3743.75
		Kpi	19387	19515	19462	19522	19471.5
	valid	KK	398	419	428	400	411.25
		KPi	2091	2197	2103	2120	2127.75
4	inner train	KK	3852	3854	3773	3882	3840.25
		Kpi	20006	20129	20096	20146	20094.25
	valid	KK	410	426	437	409	420.5
		KPi	2161	2273	2173	2191	2199.5

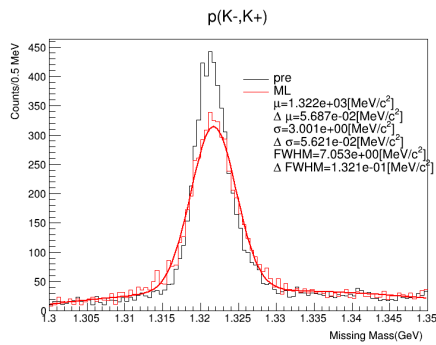
それぞれの欠損質量カット下で得られた欠損質量の fit の様子を図 A.9 - A.14 に示す .



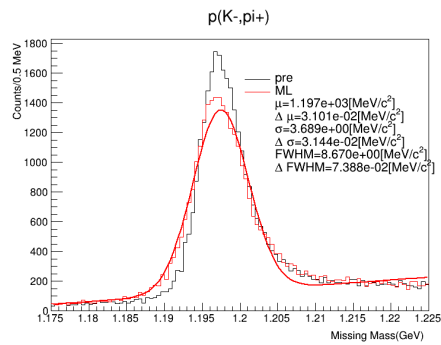
(a) KK データに対するの欠損質量 .



(b) KPi データに対するの欠損質量

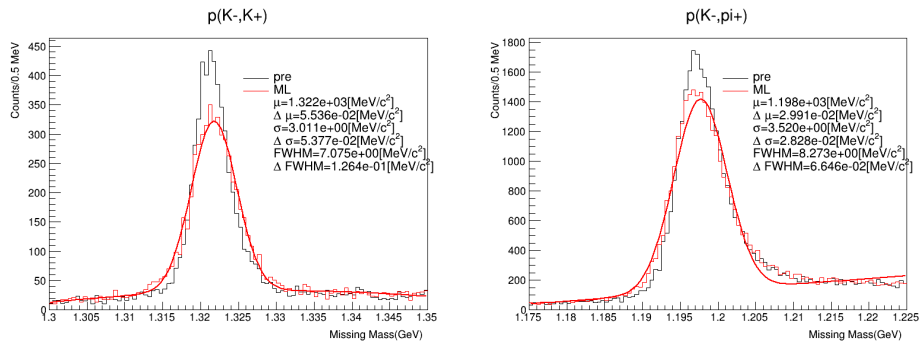
図 A.9: 小節 3.5.3 で欠損質量カット 1σ の場合の欠損質量スペクトル

(a) KK データに対するの欠損質量 .



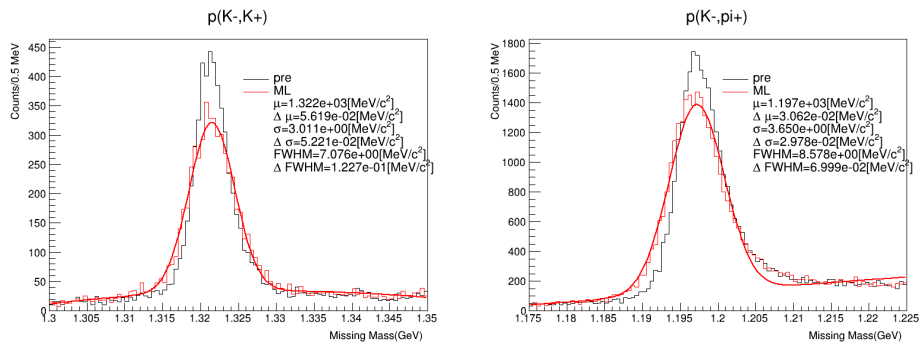
(b) KPi データに対するの欠損質量

図 A.10: 小節 3.5.3 で欠損質量カット 1.5σ の場合の欠損質量スペクトル



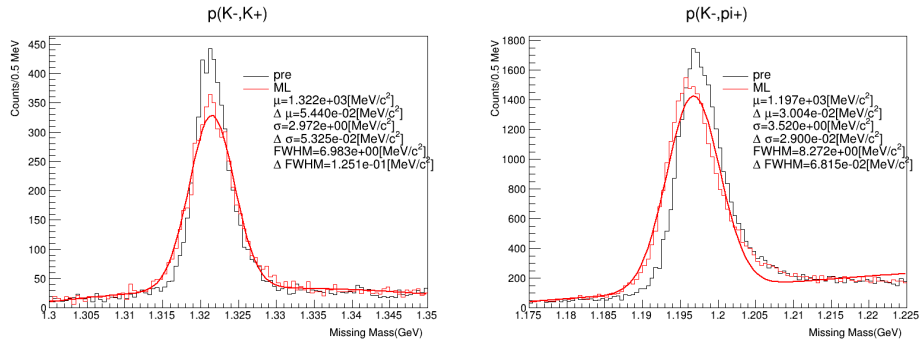
(a) KK データに対するの欠損質量 . (b) KPi データに対するの欠損質量

図 A.11: 小節 3.5.3 で欠損質量カット 2σ の場合の欠損質量スペクトル



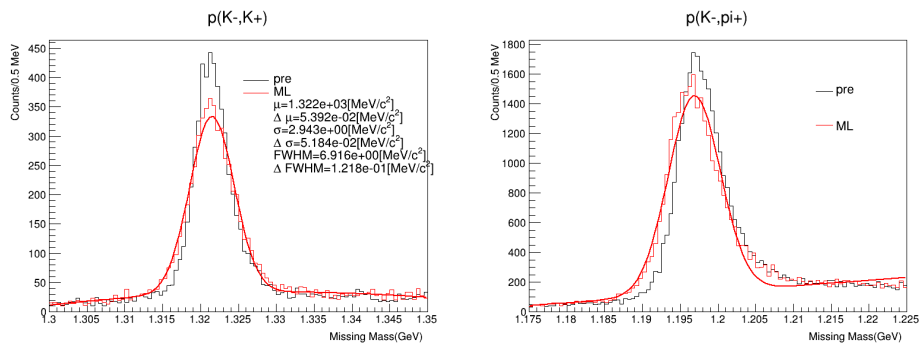
(a) KK データに対するの欠損質量 . (b) KPi データに対するの欠損質量

図 A.12: 小節 3.5.3 で欠損質量カット 2.5σ の場合の欠損質量スペクトル



(a) KK データに対するの欠損質量 . (b) KPi データに対するの欠損質量

図 A.13: 小節 3.5.3 で欠損質量カット 3.5σ の場合の欠損質量スペクトル



(a) KK データに対するの欠損質量 . (b) KPi データに対するの欠損質量

図 A.14: 小節 3.5.3 で欠損質量カット 4σ の場合の欠損質量スペクトル

参考文献

- [1] T. Inoue et al.: Progress of Theoretical Physics **124** (2010) 591.
- [2] T. Hasegawa et al.: Phys. Rev. Lett. **74** (1995) 224.
- [3] H. Hotchi et al.: Phys. Rev. C **64** (2001) 044302.
- [4] H. Tamura et al.: Phys. Rev. Lett. **84** (2000) 5963.
- [5] D. Millener: Nuclear Physics A **754** (2005) 48 . Proceedings of the Eighth International Conference on Hypernuclear and Strange Particle Physics.
- [6] A. Gal, E. V. Hungerford, and D. J. Millener: Rev. Mod. Phys. **88** (2016) 035004.
- [7] T. Nagae et al.: Phys. Rev. Lett. **80** (1998) 1605.
- [8] P. K. Saha et al.: Phys. Rev. C **70** (2004) 044613.
- [9] T. Harada and Y. Hirabayashi: Nuclear Physics A **759** (2005) 143 .
- [10] H. Takahashi et al.: Phys. Rev. Lett. **87** (2001) 212502.
- [11] H. Ekawa et al.: Progress of Theoretical and Experimental Physics **2019** (2019). 021D02.
- [12] K. Nakazawa et al.: Progress of Theoretical and Experimental Physics **2015** (2015). 033D02.
- [13] S. Acharya et al.: Phys. Rev. Lett. **123** (2019) 112002.
- [14] T. Fukuda et al.: Phys. Rev. C **58** (1998) 1306.
- [15] P. Khaustov et al.: Phys. Rev. C **61** (2000) 054603.
- [16] M. M. Nagels, T. A. Rijken, and J. J. de Swart: Phys. Rev. D **15** (1977) 2547.
- [17] M. M. Nagels, T. A. Rijken, and J. J. de Swart: Phys. Rev. D **20** (1979) 1633.
- [18] K. Tominaga et al.: Nuclear Physics A **642** (1998) 483 .
- [19] M. o. Yamaguchi: Progress of Theoretical Physics **105** (2001) 627.
- [20] T. A. Rijken and Y. Yamamoto: Phys. Rev. C **73** (2006) 044008.
- [21] Y. Yamamoto, E. Hiyama, and T. Rijken: Nuclear Physics A **835** (2010) 350 . Proceedings of the 10th International Conference on Hypernuclear and Strange Particle Physics.

- [22] T. Motoba and S. Sugimoto: Nuclear Physics A **835** (2010) 223 . Proceedings of the 10th International Conference on Hypernuclear and Strange Particle Physics.
- [23] S. Nagamiya: Progress of Theoretical and Experimental Physics **2012** (2012). 02B001.
- [24] <https://j-parc.jp/Acc/ja/index.html>. J-PARC. <https://j-parc.jp/Acc/ja/index.html>.
- [25] K. o. Agari: Progress of Theoretical and Experimental Physics **2012** (2012). 02B009.
- [26] 松本祐樹: 修士論文, 東北大学 (2012).
- [27] S.Morinobu. ORBIT. "private communication".
- [28] J. Myrheim and L. Bugge: Nuclear Instruments and Methods **160** (1979) 43 .
- [29] H. Eichinger and M. Regler: *Review of track-fitting methods in counter experiments* (CERN Yellow Reports: Monographs. CERN, Geneva, 1981), CERN Yellow Reports: Monographs.
- [30] <https://www.rcnp.osaka-u.ac.jp/Divisions/np1-a/RCF/RCNPCF-GRj.html>. グランドライデン/LAS. <https://www.rcnp.osaka-u.ac.jp/Divisions/np1-a/RCF/RCNPCF-GRj.html>.
- [31] M. Fujiwara et al.: Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **422** (1999) 484 .
- [32] B. P. Roe et al.: Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **543** (2005) 577 .
- [33] : Quark versus Gluon Jet Tagging Using Jet Images with the ATLAS Detector (2017).
- [34] G. Hinton. rmsprop: Divide the gradient by a running average of its recent magnitude, 2012.
- [35] D. P. Kingma and J. Ba. Adam: A Method for Stochastic Optimization, 2017.
- [36] M. Abadi et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, 2015. Software available from tensorflow.org.
- [37] F. Chollet et al. Keras. <https://keras.io>, 2015.
- [38] M. Tanabashi et al.: Phys. Rev. D **98** (2018) 030001.
- [39] H. Lee et al. On the ability of neural nets to express distributions, 2017.