

1

2022年度 修士論文

2

電子陽電子ヒッグスファクトリーのための ジェット測定技術の研究

3

4

九州大学大学院 理学府物理学専攻
粒子物理学分野素粒子実験研究室

5

6

尾上 友紀

7

指導教員 末原大幹 助教 川越清以 教授

8

2023年2月3日

9



11

概要

12 2012 年に発見されたヒッグス粒子は、宇宙の物質の起源の解明につながる粒子であり、そ
 13 の性質は謎に包まれている。そのため、ヒッグスファクトリーによってヒッグス粒子を大量に
 14 生成し精密測定することで、暗黒物質などの標準理論を超える物理を切り開くことは、現代
 15 の物理学における重要事項となっている。本研究で念頭に置いている国際リニアコライダー
 16 (ILC) 計画は次世代の電子陽電子衝突型加速器であり、ヒッグスファクトリーとしての運転を
 17 期待されている。ヒッグスファクトリーにおけるヒッグス粒子生成事象は、多数のハドロンの
 18 束であるジェットを終状態に複数含むことが多いため、ジェットを高い精度で検出・再構成す
 19 ることは物理解析の性能に直結する。そのため、本論文では ILC のジェット測定技術に関する
 20 2 つの研究を行なった。

21 ILC で重要な事象にはジェットを含むものが多くあるため、ILC の物理目標を達成す
 22 るためには高いジェットエネルギー分解能が必要である。そのため、ILC では Particle Flow
 23 Algorithm と呼ばれる粒子識別のアルゴリズムによって高い分解能の達成を目指しており、こ
 24 れには非常に高精細な電磁カロリメータが求められる。シリコンタングステン電磁カロリメー
 25 タは、1 つの読み出しセルの大きさが $5 \text{ mm} \times 5 \text{ mm}$ と非常に細分化されており、国際協力に
 26 よって技術プロトタイプの開発が行われている。本研究では、15 層の検出層を持つプロトタ
 27 イプを構築し、欧州原子核研究機構 (CERN) の Super Proton Synchrotron (SPS) 加速器に
 28 て 150 GeV の高エネルギービームを用いたテストビーム実験を行った。そして実験結果から、
 29 読み出しシステムの性能やシリコンセンサーの振る舞いについて調査し、将来の技術プロトタ
 30 イプに向けた改善策を講じた。特に読み出しにおいてトリガー情報が誤って収集されてしまう
 31 Re-triggering 現象について調査を行った。

32 さらに、ILC のジェット再構成におけるフレーバー識別アルゴリズムの開発を、深層学習
 33 技術を用いて行った。フレーバー識別では、ジェットの構成粒子の種類や運動量、崩壊点に関
 34 する情報から、ジェットの元となる夸克のフレーバーを識別する。現在は従来の機械学習
 35 手法である Boosted Decision Trees が用いられているが、本研究では更なる識別性能の向上
 36 を目的に深層学習を用いたアルゴリズムを開発した。本アルゴリズムの最大の特徴は、フレ
 37 バー識別のためにグラフ構造のデータを構築し、グラフデータの学習が可能なニューラルネッ
 38 トワークであるグラフニューラルネットワークで学習を行った点である。グラフデータはデ
 39 タの関係をグラフ形式で表現することで、データの相互関係を考慮することができる。これに
 40 よって実際の物理現象と比較した際の情報損失を減らし、より高い精度で識別を行うことが出
 41 来ると考えた。本アルゴリズムの結果は従来技術と比較して b フレーバーの識別のみの改善と
 42 なったが、グラフデータの構築によって従来は別々にプロセスを行なっていた崩壊点検出のア
 43 ルゴリズムを統合することができた。

目次

46	第1章 序論	12
47	1.1 素粒子物理学	12
48	1.1.1 標準模型	12
49	1.1.2 ヒッグス機構	14
50	1.2 国際リニアコライダー計画; ILC	15
51	1.3 ILC の物理	16
52	1.3.1 ヒッグス生成過程と質量精密測定	16
53	1.3.2 ヒッグス結合定数の精密測定	17
54	1.3.3 ヒッグス自己結合	20
55	1.3.4 階層性問題	20
56	1.3.5 その他の新物理	21
57	1.4 ILC の検出器	21
58	1.4.1 ILD で検出器可能な粒子群	21
59	1.4.2 Particle Flow Algorithm; PFA	22
60	1.4.3 International Large Detector: ILD	23
61	崩壊点検出器	23
62	中央飛跡検出器	24
63	カロリメータ	24
64	ミューオン検出器	25
65	1.5 ILC のソフトウェア	25
66	1.5.1 イベントジェネレータと検出器シミュレーション	25
67	1.5.2 事象再構成	25
68	飛跡再構成	26
69	崩壊点検出	26
70	ジェットクラスタリング	27
71	フレーバー識別	27
72	1.6 本研究の目的	28
73	1.6.1 高エネルギービームによる SiW-ECAL の性能評価	28

74	1.6.2 深層学習を用いたフレーバー識別アルゴリズムの開発	29
75	第 2 章 シリコンタングステン電磁カロリメータ	30
76	2.1 入射粒子と物質の相互作用	30
77	2.1.1 荷電粒子	30
78	電離損失	30
79	制動放射	32
80	2.1.2 光子	33
81	2.1.3 ハドロン	34
82	2.2 粒子検出器の動作原理	34
83	2.2.1 ガス検出器	35
84	2.2.2 半導体検出器	35
85	2.2.3 シンチレーション検出器	36
86	2.3 シリコンタングステン電磁カロリメータ SiW-ECAL	36
87	2.3.1 SiW-ECAL の全体構造	36
88	2.3.2 シリコン半導体検出器	36
89	2.3.3 読み出しシステム	37
90	2.3.4 タングステン吸収層	39
91	2.3.5 技術プロトタイプ	40
92	第 3 章 ビームテストによる評価試験	42
93	3.1 ビームライン	42
94	3.2 実験セットアップ	43
95	3.2.1 測定機器のセットアップ	43
96	3.2.2 信号読み出し	45
97	3.3 実験結果	46
98	3.3.1 検出器応答	46
99	3.3.2 ペデスタル	47
100	3.4 まとめと考察	49
101	第 4 章 深層学習	50
102	4.1 ニューラルネットワーク	50
103	4.1.1 パーセプトロン（単層ニューラルネットワーク）	50
104	4.1.2 多層パーセプトロン（多層ニューラルネットワーク）	51
105	活性化関数	52
106	出力層の設計	53
107	損失関数	53
108	誤差逆伝播法	54

109	前処理	54
110	ミニバッチ処理	55
111	最適化アルゴリズム	55
112	過学習	57
113	4.1.3 ディープニューラルネットワーク	58
114	4.2 グラフニューラルネットワーク	58
115	4.2.1 メッセージパッシング	59
116	4.2.2 Graph Convolution Network (GCN)	60
117	Spectral Graph Convolution	60
118	Spatial Graph Convolution	61
119	4.2.3 Graph Attention Network (GAT)	61
120	第 5 章 深層学習を用いたジェットフレーバー識別	63
121	5.1 イベントサンプルと先行研究	63
122	5.1.1 LCFIPlus におけるフレーバー識別	64
123	5.2 ディープニューラルネットワークによる実装	66
124	5.2.1 実装目的	66
125	5.2.2 入力変数とネットワークアーキテクチャ	66
126	5.2.3 ハイパーパラメータの最適化	67
127	5.2.4 学習評価	69
128	5.3 グラフニューラルネットワークによる実装	72
129	5.3.1 実装目的	72
130	5.3.2 飛跡によるグラフデータセット	72
131	5.3.3 ネットワークアーキテクチャ	75
132	5.3.4 ハイパーパラメータの最適化	77
133	5.3.5 学習評価	79
134	第 6 章 まとめと今後の展望	83
135	付録 A 付録 A	86
136	A.1 LCIO parameter	86
137	参考文献	89

図目次

138

139	1.1	素粒子の標準模型（数値は質量 [GeV/c ²]）	13
140	1.2	ヒッグスボテンシャル [3]	14
141	1.3	ILC の概略図	15
142	1.4	Zh 随伴生成、WW 融合反応、ZZ 融合反応におけるファインマンダイアグラム [4]。	17
143	1.5	ILC の重心系エネルギーに対するヒッグス生成断面積。ヒッグス粒子の質量が $m_h = 125 \text{ GeV}$ であるとして、Zh 随伴生成、WW 融合反応、ZZ 融合反応をそれぞれ赤、青、緑線で示している [4]。	17
144	1.6	標準模型におけるヒッグス粒子の質量と崩壊分岐比の関係	18
145	1.7	2つの新物理のモデルにおける、標準模型のヒッグス粒子結合定数とのずれ。(左) 超対称性 (SUSY) モデル (右) 複合ヒッグスモデル。誤差棒は 1σ の範囲を表している。	19
151	1.8	ILC と HL-LHC におけるヒッグス粒子の各粒子に対する結合定数の測定精度予測。誤差棒は 1σ の範囲を表している。(紫 : LHC (300 fb^{-1})、青 : ILC1 (250 GeV)、臙脂 : ILC (500 GeV)、薄橙 : ILCTeV (1 TeV)) [5]	19
154	1.9	ヒッグス自己結合 $e^+e^- \rightarrow Zhh$	20
155	1.10	ヒッグス粒子の質量補正となるフェルミオンループ	21
156	1.11	ILD と SiD の全体図	22
157	1.12	ILD の断面図	23
158	1.13	ILD 崩壊点検出器の構造	24
159	1.14	ILC におけるシミュレーションと事象再構成、物理解析の流れとソフトウェア	26
160	1.15	ジェットの崩壊の様子。赤線が 1 次崩壊点、青線が 2 次崩壊点、緑線が 3 次崩壊点を由来とする飛跡を表す。	28
162	2.1	水素 (液体)、ヘリウム (気体)、炭素、アルミニウム、鉄、スズ、および鉛における平均エネルギー損失と、入射粒子 (ミューオン、パイ中間子、陽子) の速度の関係。	32
165	2.2	光子と物質の相互作用	33

166	2.3	光子と鉛の相互作用における断面積と光子のエネルギーの関係。 $\sigma_{p.e.}$ は光電 効果を、 $\sigma_{Compton}$ はコンプトン散乱を、 κ_e, κ_{nuc} は電子陽電子対生成を指す。	35
167	2.4	ILD および ECAL の全体図	37
168	2.5	SiW-ECAL の構造	37
169	2.6	シリコンパッドセンサー	37
170	2.7	SKIROC2A のアナログ部の回路図	39
171	2.8	short slab の構造	41
172	2.9	FEV13, COB	41
173	2.10	多層読み出しのための CORE モジュール	41
174	3.1	CERN の全体図	43
175	3.2	セットアップの全体図。黄枠 : AHCAL、青枠 : SiW-ECAL、赤 : ビーム位置	44
176	3.3	SiW-ECAL のプロトタイプモジュールと電源供給、読み出しシステム . . .	44
177	3.4	SiW-ECAL 読み出しソフトウェアにおけるイベントモニター (左) ヒット マップ (右) 各層あたりのヒット数	46
178	3.5	各ビーム種類におけるヒットマップ。左上から 0 層右上が 3 層、右下が 14 層 と順に並んでいる。	47
179	3.6	40/60/100/150GeV の電子ビームにおけるペデスタル	48
180	3.7	ペデスタルをガウス関数フィットした際のパラメータ。横軸がチップを、縦 軸がチャンネル番号を表しており、それぞれ全層、全 SCA での可視化のた め、層数や SCA をかけた値となっている。(左) ガウス関数の中央値 (右) ガ ウス関数の幅の大きさ	48
181	3.8	ダブルペデスタルの例	49
182	4.1	パーセプトロン	50
183	4.2	多層パーセプトロン (ニューラルネットワーク)	51
184	4.3	活性化関数のグラフ。Forward は順方向、backward は逆方向 (後に誤差逆伝 播法で説明) の際の演算を表す。	53
185	4.4	ディープニューラルネットワーク	58
186	4.5	(左) 4 つのノードを持つ全結合グラフニューラルネットワーク。 h_i はノード 表現を表す。(右) メッセージパッシングの処理。	59
187	4.6	GraphSAGE における処理 (1. サンプリング, 2. 隣接からの集約, 3. 学習結 果による推論)	61
188	4.7	(左) 重みベクトル a を用いた Attention 処理。(右) 1 つのノード h_1 に対す る隣接ノード $h_{\neq 1}$ の Attention と、ノード特徴量の更新 h'_1	62
189	5.1	フレーバー識別のためのディープニューラルネットワークの概略図	67

200	5.2	各試行における評価関数の値。値が最小になるようなアルゴリズムを実行している。	68
201	5.3	w_{node} (図中の α), w_{edge} (図中の β)に対する評価関数の2次元ヒストグラム。	69
202	5.4	各パラメータの重要度	69
203	5.5	(左) 学習の経過における損失関数。(右) 学習の経過における学習精度	70
204	5.6	ディープニューラルネットワークの学習における混合行列。縦軸が実際の答えを、横軸が学習結果を表している。	70
205	5.7	LCFIPlus(左)とディープニューラルネットワーク(右)によるbフレーバージェットの識別効率の比較。緑:bジェットに対するcジェットの識別効率、青:bジェットに対するudsジェットの識別効率	71
206	5.8	LCFIPlus(左)とディープニューラルネットワーク(右)によるcフレーバージェットの識別効率の比較。赤:cジェットに対するbジェットの識別効率、黒:cジェットに対するudsジェットの識別効率	71
207	5.9	今回の学習に用いたグラフデータの一例。ノードは飛跡を、グラフ全体が1つのジェットに対応する。	73
208	5.10	ノードラベル毎のデータ数	74
209	5.11	フレーバー識別のためのグラフデータを用いたネットワークの概略図	76
210	5.12	各試行における評価関数の値。値が最小になるようなアルゴリズムを実行している。	78
211	5.13	w_{node} (図中の α), w_{edge} (図中の β)に対する評価関数の2次元ヒストグラム。	79
212	5.14	各パラメータの重要度。 β_2 はRAdamにおける学習の進行度のパラメータ、ドロップアウトは学習におけるアテンション係数の隣接ノードへの参照率を表す。	79
213	5.15	(上) 学習の経過における損失関数。(左下) 学習の経過におけるノードの学習精度、(中央下) リンクの学習精度、(右下) グラフの学習精度。	80
214	5.16	(左) ノード分類の混合行列、(中央) リンク予測の混合行列、(右) グラフ分類の混合行列。	80
215	5.17	LCFIPlus(左)とグラフニューラルネットワーク(右)によるbフレーバージェットの識別効率の比較。緑:bジェットに対するcジェットの識別効率、青:bジェットに対するudsジェットの識別効率。	81
216	5.18	LCFIPlus(左)とグラフニューラルネットワーク(右)によるcフレーバージェットの識別効率の比較。赤:cジェットに対するbジェットの識別効率、黒:cジェットに対するudsジェットの識別効率。	82
217	A.1	xy平面	86
218	A.2	sz平面	87

表目次

235	1.1	$m_h = 125 \text{ GeV}$ における SM ヒッグス粒子の崩壊分岐比	18
236	1.2	ジェットを占める各粒子と対応する検出器	23
237	2.1	シリコンセンサーの仕様	37
238	2.2	物質量の大きい吸収層の候補物質 (λ は相互作用長、 L_R は放射長、 R_M はモリエール半径を示す。)	40
239	3.1	H2A ビームラインにおけるビームパラメータ	43
240	3.2	SiW-ECAL のレイヤー構成 (0 層目がビーム上流側)	45
241	5.1	シミュレーションデータのパラメータ	63
242	5.2	LCFIPplus におけるフレーバー識別の入力変数	65
243	5.3	ディープニューラルネットワークにおけるハイパーパラメータ	67
244	5.4	最適化を行ったハイパーパラメータとその値の範囲。	68
245	5.5	識別効率 (Tagging Efficiency) に対するジェット誤認率 (Mis-id fraction)	72
246	5.6	各ノード (飛跡) が持つ特徴量。詳細については Appendix を参照。	73
247	5.7	ノード分類の答えラベル	74
248	5.8	リンク予測の答えラベル	74
249	5.9	グラフ分類の答えラベル	74
250	5.10	グラフニューラルネットワークにおけるハイパーパラメータ	77
251	5.11	最適化を行ったハイパーパラメータとその値の範囲。	78
252	5.12	識別効率 (Tagging Efficiency) に対するジェット誤認率 (Mis-id fraction)	82

255 第 1 章

256 序論

257 本章では、はじめに 1.1 節で素粒子とそれらに働く相互作用を説明する標準模型 (The
 258 Standard Model; SM) について述べる。そして 1.2 節にて将来の電子陽電子ヒッグスファク
 259 トリーである、国際リニアコライダー計画 (International Linear Collider; ILC) の概要に触
 260 れたのち、1.3 節で ILC が探索する物理について、特にヒッグス粒子に関係した事項を中心に
 261 言べる。ILC の実現において必要な ILC の検出器について 1.4 節で、ソフトウェアについて
 262 1.5 節でまとめたのち、1.6 節で本研究の目的を述べる。

263 1.1 素粒子物理学

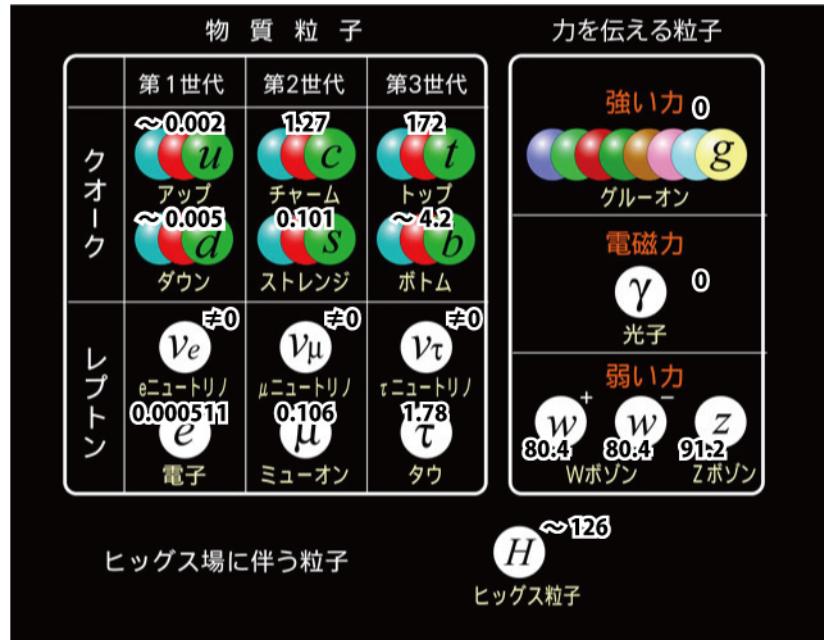
264 1.1.1 標準模型

265 素粒子とは、物質を構成している究極要素をさす名称である。そして素粒子物理学は、それ
 266 ら構成要素とその間に働く相互作用の性質を解明する学問である。現代の素粒子物理学では、
 267 すべての現象を説明するための基本的な枠組みとして図 1.1 のような標準模型を掲げており、
 268 これは現時点の実験データと高い精度で一致することが確認されている。

269 標準模型は、主に次に挙げる 2 つの基本的な前提に沿って記述されている。1 つ目に、物質
 270 の究極要素である素粒子はクォークとレプトンというスピン $1/2$ のフェルミオンである。2 つ
 271 目に、素粒子の相互作用はゲージ理論によって記述され、標準模型における相互作用は電磁相
 272 互作用・弱い相互作用・強い相互作用の 3 つである。

273 物質の化学的性質を失わない最小単位は分子であり、分子はさらに原子の組み合わせによっ
 274 て構成されている。そして原子は原子核と電子によって構成されており、原子核は陽子と中性
 275 子のような核子からなっている。この核子を構成するものがクォークであり、標準模型におい
 276 ては 6 種類存在する。一方で、電子のような核力といった強い相互作用をしないものをレプト
 277 ンと呼び、同様に 6 種類存在する。クォーク・レプトンともに 3 つの世代と 2 つの電荷タイプ
 278 をもっており、世代の大きい粒子ほど重いため弱い相互作用により小さい世代のクォーク・レ
 279 プトンへと崩壊する。

280 場の量子論では量子場 ϕ が素粒子と関連して記述されており、ラグランジアンによって相

図 1.1: 素粒子の標準模型 (数値は質量 [GeV/c²])

互作用の性質が記述される。ラグランジアンが量子場の局所的なゲージ対称性のもとで不变であると仮定すると、ゲージ場と呼ばれるベクトル場が現れるが、このゲージ場と量子場の積によって相互作用を表す理論をゲージ理論という。ゲージ粒子はこのゲージ場が粒子として現れたもので、素粒子の相互作用を媒介するとされているスピン 1 のゲージ粒子には、グルーオン・光子・ W ボソン・ Z ボソンの 4 種類がある。クォークとグルーオンの相互作用である強い相互作用は、量子色力学に基づき $SU(3)$ 対称性をもつ。また荷電粒子と光子の相互作用である電磁相互作用と W ・ Z ボソンを介する弱い相互作用は、グラショウ=ワインバーグ=サラム理論 (GWS 理論) [1] によって統一され電弱相互作用と呼ばれており、 $SU(2) \times U(1)$ 対称性をもつ。これに加えて重力相互作用が存在するが、他の 3 つの相互作用と比較して非常に弱いため標準模型では扱われない。

本論文のテーマである電子・陽電子はレプトンに部類される素粒子であり、粒子・反粒子の関係にある。反粒子とはある粒子の符号の異なる電荷を持つが、質量やスピン角運動量など電荷以外は同じ特性を持つ粒子を指しており、陽電子は電子の反粒子である。粒子・反粒子の関係にある粒子はその質量エネルギーを放出して対消滅を起こす性質があり、ヒッグスファクタリーにおいてもこの現象によって別の素粒子を生成する。

また、標準模型は実験データと高い精度で一致しているが、標準模型だけでは説明しきれない物理も数多く存在しており、標準理論を越えた新しい物理 (Beyond the Standard Model; BSM) を検証する実験が世界中で行われている。

299 1.1.2 ヒッグス機構

GWS理論においてゲージ粒子はゲージ対称性により質量項が禁止されているが、先述のW・Zボソンはそれぞれ $80.4\text{ GeV}/c^2$ 、 $91.2\text{ GeV}/c^2$ の質量を持っている。標準模型ではこれを説明するためにヒッグス機構を導入し、ゲージ対称性が自発的に破れることで質量を獲得している。このヒッグス機構によると、宇宙の膨張によって真空は冷却されてヒッグスが凝縮した状態に相転移が起き、この相転移によって対称性が破れる。そのため真空に複素2次元のスカラー場としてヒッグス場 Φ を導入し、これとゲージ場との相互作用によって質量を持つとしている。またヒッグス場の存在と同時に、対応する粒子としてヒッグス粒子の存在が必要となる。ヒッグス場のポテンシャル[2]は

$$V(\phi) = \mu^2(\Phi^\dagger\Phi) + \lambda(\Phi^\dagger\Phi)^2 \quad (\text{ただし } \mu^2 < 0) \quad (1.1)$$

300 と書くことができ、ポテンシャルの概形は図1.2のような形状になる。

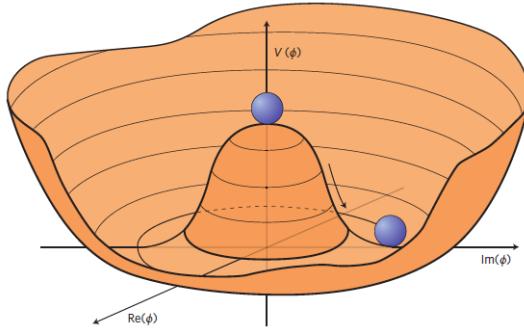


図1.2: ヒッグスポテンシャル[3]

図1.2のようなポテンシャルでは、 $V(\phi)$ 軸周りに真空がリング状に縮退している。また真空のポテンシャルは最小値をとる点で安定するが、ヒッグスポテンシャルにおいては一つの点をとると位相回転対称性が破れてしまう。このときヒッグス場に、

$$\langle\Phi\rangle = \frac{v}{2} = \sqrt{\frac{-\mu^2}{2\lambda}} \quad (1.2)$$

301 の有限の真空期待値が現れる。これによってゲージ粒子が質量を獲得する。より直感的には、
302 真空に凝縮されたヒッグス場の中でゲージ粒子を加速しようとした時に、ヒッグス場から抵抗
303 を受ける。この抵抗は、ゲージ場が1個のヒッグス粒子と衝突する頻度を意味する結合定数
304 と、真空のヒッグスの密度に比例することから、質量は加速されにくさを表す量と考えること
305 ができる。

306 このヒッグス粒子は、2012年7月に欧洲原子核研究機構(Conseil Européen pour la
307 Recherche Nucléaire; CERN)の大型ハドロン衝突型加速器(Large Hadron Collider; LHC)
308 におけるATLAS、CMS実験によって発見され、理論と実験との一致が確認された。本論文

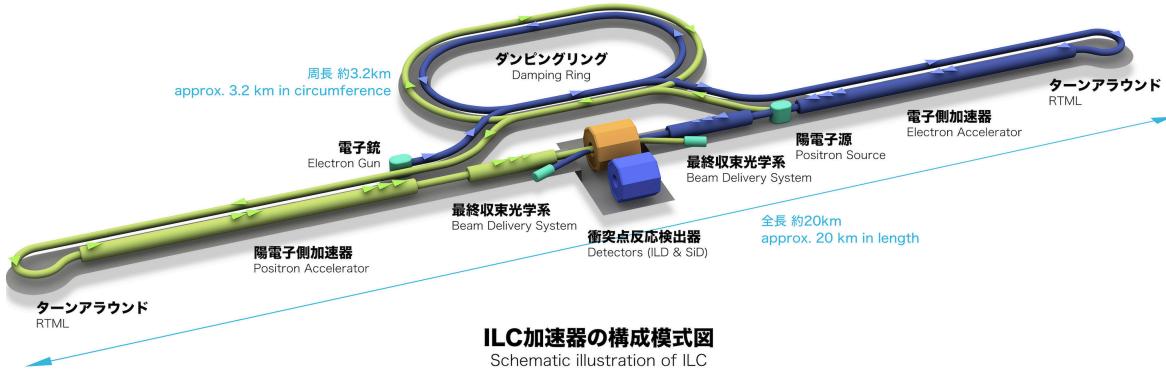


図 1.3: ILC の概略図

309 のテーマであるヒッグスファクトリーでは、このヒッグス粒子を大量に生成し詳細に研究する
310 ことを最大の目的としている。

311

312 1.2 国際リニアコライダー計画; ILC

313 国際リニアコライダー (International Linear Collider; ILC) は、岩手県北上山地に建設
314 が計画されている電子陽電子衝突型線形加速器であり、将来のヒッグスファクトリーとして
315 の稼働を期待されている。全長 20 km の線形加速器を用いて電子と陽電子を加速し、中央の
316 Interaction Point (IP) で衝突させることで様々な粒子を生成し、これを解析することでヒッ
317 グス粒子を始めとする新物理を探索することを目的としている (図 1.3)。また ILC は重心系
318 エネルギー $\sqrt{s} = 250 \text{ GeV}$ での運転開始を予定しているが、線形加速部を延長することで最
319 大 1 TeV までのアップグレードも可能になっており、各エネルギーにおける物理プログラム
320 のメインターゲットは次のようにになっている。

- 321 • $\sqrt{s} = 250 \text{ GeV}$: Zh 随伴生成過程の研究
- 322 • $\sqrt{s} = 350 \text{ GeV}$: $t\bar{t}$ 対生成、WW 融合過程によるヒッグス生成
- 323 • $\sqrt{s} = 500 \text{ GeV}$: ヒッグスの自己結合とトップ湯川結合の測定、高統計によるヒッグス
324 精密測定
- 325 • $\sqrt{s} = 1 \text{ TeV}$: ヒッグスの自己結合とトップ湯川結合の精密測定

326 ヒッグス粒子を発見した LHC と比較して ILC には以下の 4 つの利点が存在する。1 つ目は
327 LHC が複合粒子であるハドロンのコライダーであるのに対して、ILC はレプトンコライダー
328 である点である。ILC では背景事象が少ないクリーンな環境で、ヒッグス粒子を始めとした網
329 紹的な新物理探索が可能になっている。また、LHC では断面積を計算する上で QCD に基づ
330 く系統的な不確定性が存在するが、ILC では電弱相互作用のみについて考えることができる
331 ため、高精度な理論検証が可能になる。2 つ目は加速粒子である電子陽電子が粒子反粒子の関

係にある点である。粒子反粒子が対消滅することで全エネルギーを目的粒子の生成に効率的に用いることができる。加えて ILC のビームはバンチ構造をしており、全事象を記録したデータをバンチの衝突間隔を利用して転送し、オフラインで事象選択を行うことができるためトリガーレスで運転することが可能である。これによって、多重度が低く運動量が小さい軌跡だけを持つ事象のような、トリガーで捉えることが難しい事象までをも活かすことができる。3つ目はビーム起因のバックグラウンドが小さい点である。これによって崩壊点検出器をビームからおよそ 15 mm と近い距離におくことができ、フレーバーの識別において b フレーバのみではなく c フレーバの識別も可能となる。4つ目はビーム偏極が全エネルギーで可能となっている点である。ビーム偏極とは進行方向に対するスピンを意味しており、これによって測定可能な物理量を増やすことができる。

1.3 ILC の物理

1.3.1 ヒッグス生成過程と質量精密測定

1.1 節で述べた通り、ILC はヒッグスファクトリーとしての役割を期待されている。ヒッグスファクトリーでは、ヒッグス粒子を大量に生成し崩壊過程を精密測定することで、他の粒子との結合定数を測定し標準模型を検証することができる。ILC におけるヒッグス粒子の生成断面積は図 1.5 のようになっており、運転開始で予定している $\sqrt{s} = 250 \text{ GeV}$ 付近では、主に Zh 随伴生成過程の断面積が最大となる。この Zh 随伴生成過程では、反跳粒子である Z ボソンを正確に再構成することで、ヒッグス粒子の質量を

$$M_{recoil}^2 = (\sqrt{s} - E_{ff})^2 - |\vec{p}_{ff}|^2 \quad (1.3)$$

によって、高い精度で求めることができる。 $(E_{ff}$ は Z ボソンが崩壊するフェルミオン対のエネルギーを、 \vec{p}_{ff} は運動量を表す。)

LHC のような複合粒子同士の衝突では、始状態の運動量が定まっていないことに加えバックグラウンドとなる事象も多い。しかし ILC の反応過程では、始状態の運動量がわかっているため、反跳粒子を正確に再構成することでヒッグス粒子の崩壊モードに依存せず全断面積の測定が可能となっている。このようヒッグス粒子の崩壊モードに依存しない測定によってヒッグス粒子の invisible な崩壊の存在を検証することができる。これにはヒッグス粒子が暗黒物質や後述する超対称性粒子に崩壊するモードなどが含まれており、標準模型を超える新物理の検証を行うことが可能である。

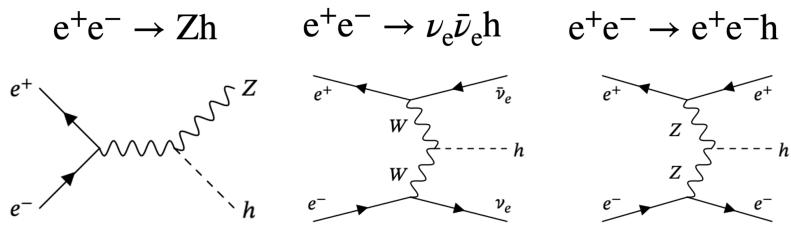


図 1.4: Zh 随伴生成、WW 融合反応、ZZ 融合反応におけるファインマンダイアグラム [4]。

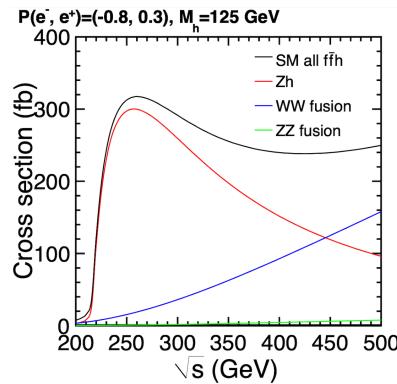


図 1.5: ILC の重心系エネルギーに対するヒッグス生成断面積。ヒッグス粒子の質量が $m_h = 125 \text{ GeV}$ であるとして、Zh 随伴生成、WW 融合反応、ZZ 融合反応をそれぞれ赤、青、緑線で示している [4]。

353 1.3.2 ヒッグス結合定数の精密測定

354 電子陽電子衝突によって生成されるヒッグス粒子は不安定であるため、より質量の小さい粒
355 子のペアに崩壊する。標準模型におけるヒッグス粒子の各質量に対する崩壊分岐比の割合を図
356 1.6 に、 $m_h = 125 \text{ GeV}$ における崩壊分岐比を表 1.1 示す。ILC ではヒッグス粒子の生成断面
357 積と全崩壊幅を精密に測定することができるため、ヒッグス粒子の崩壊分岐比を精密に決定す
358 ることができる。この崩壊分岐比はヒッグス粒子の結合定数の 2 乗に比例しており、標準模型
359 を超える新物理のシナリオにおいてヒッグス粒子の結合定数に生じるズレを検証するこ
360 ことができる。

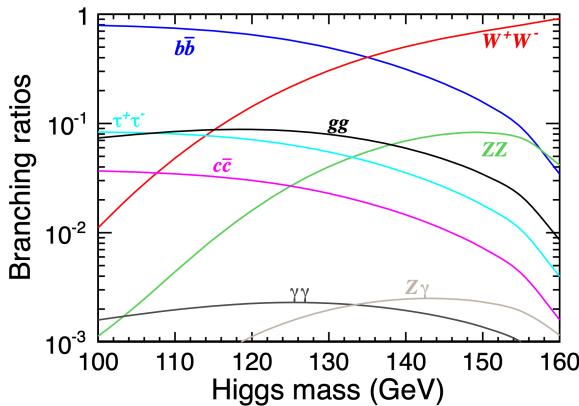


図 1.6: 標準模型におけるヒッグス粒子の質量と崩壊分岐比の関係

崩壊モード	崩壊分岐比
$b\bar{b}$	58.1 %
WW	21.5 %
gg	8.2 %
$\tau^+\tau^-$	6.3 %
$c\bar{c}$	2.9 %
ZZ	2.6 %
$\gamma\gamma$	0.2 %

表 1.1: $m_h = 125 \text{ GeV}$ における SM ヒッグス粒子の崩壊分岐比

361 以下では、結合定数の精密測定によって検証可能な新物理について述べる。図 1.7 は新物理
 362 の各シナリオにおける、標準模型ヒッグス粒子との結合定数のズレを表す。

363 ● 超対称性理論

364 超対称性理論 (Supersymmetry; SUSY) は、フェルミオンとボソンを交換する変換に
 365 対する不变性 (超対称性) を定義する理論である。この理論においては、標準模型にお
 366 けるすべての粒子に対してスピンが $1/2$ 異なる超対称性パートナーが導入される。超
 367 対称性が完全である場合、標準模型粒子と超対称性粒子の質量や相互作用が同じである
 368 必要性があるが、現段階では SUSY 粒子は発見に至っていない。SUSY では図 1.7 の
 369 ように主にヒッグス粒子に対して b クォーク・ τ レプトンとの結合が大きくなる。

370 ● 複合ヒッグス模型

371 ヒッグス粒子は TeV スケールにおいて複合粒子のように振る舞い、その内部により
 372 基本的な粒子を持っているとするモデルが複合ヒッグス模型である。この模型では
 Compositeness scale によって抑制される高次の作用素により、標準模型と比べて結合
 定数が大きくずれてしまう。Compositeness scale を f とすると、ヒッグス粒子のゲ
 ジボソンやフェルミオンへの結合は、

$$\frac{g_{hxx}}{g_{h_{SM}xx}} \simeq 1 \pm \mathcal{O}(v^2/f^2) \quad (1.4)$$

371 のオーダーで表される。図 1.7 において複合ヒッグス模型では、標準模型に比べて粒子
 372 との結合が小さくなっている。

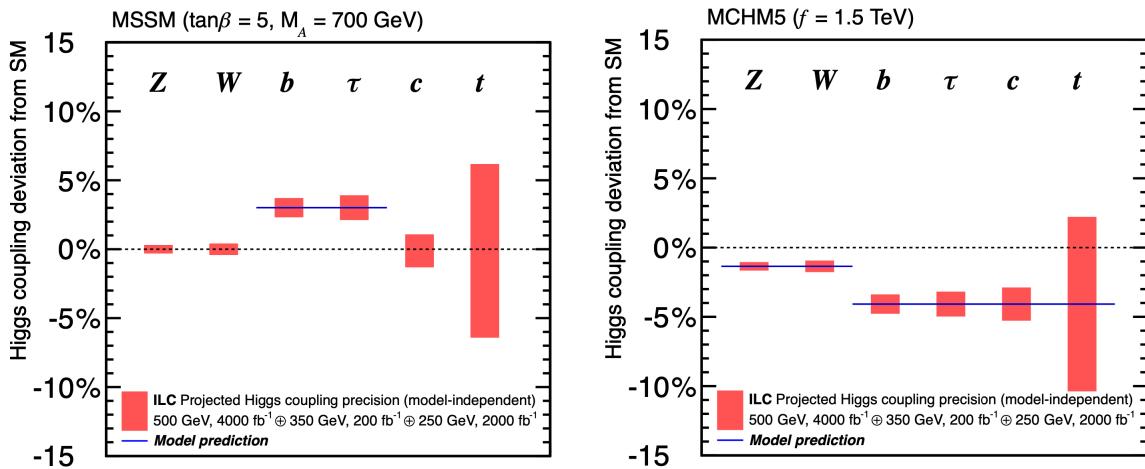


図 1.7: 2 つの新物理のモデルにおける、標準模型のヒッグス粒子結合定数とのずれ。 (左) 超対称性 (SUSY) モデル (右) 複合ヒッグスモデル。誤差棒は 1σ の範囲を表している。

373 ヒッグスの崩壊分岐比の測定精度は信号事象数を S 、背景事象数を N とすると $S/\sqrt{S+N}$
 374 となり、背景事象の影響を十分低減させることができた場合には不定性を 1% 以下まで下げる
 375 ことができる。図 1.8 では LHC と ILC におけるヒッグス粒子の結合定数の測定精度を表して
 376 おり、ILC では LHC に比べておよそ 1 桁優れた精度で測定することができる。これを達成す
 377 るために、高い検出器性能と精度の高い事象再構成・解析手法が求められている。

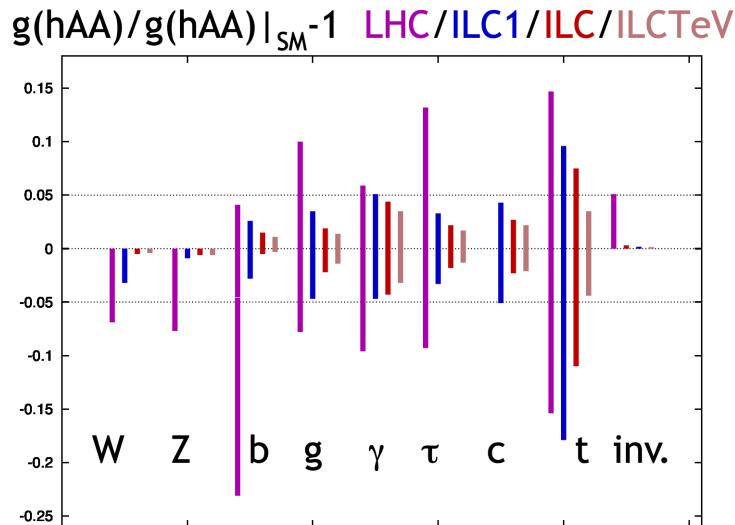


図 1.8: ILC と HL-LHC におけるヒッグス粒子の各粒子に対する結合定数の測定精度予測。誤差棒は 1σ の範囲を表している。(紫 : LHC (300fb^{-1})， 青 : ILC1 (250 GeV) , 膻脂 : ILC (500 GeV) , 薄橙 : ILCTeV (1 TeV)) [5]

378 1.3.3 ヒッグス自己結合

379 500 GeV 以上の ILC では図 1.9 のような、 $e^+e^- \rightarrow Zhh$ 反応の断面積を測定することで、
 380 自己結合について探索することができる。ゲージ不变性によりヒッグスの三点結合は、四点結
 381 合における真空凝縮としてのみ起こることができるため、三点結合を確認することでヒッグス
 382 場の真空凝縮について検証することができる。ILC500 GeV における図 1.9 の断面積は 0.2 fb
 383 と小さく高統計を得ることが難しいため、崩壊 $e^+e^- \rightarrow Zhh \rightarrow q\bar{q}bb\bar{b}$ における b フレ
 バーの識別精度が非常に重要になる。

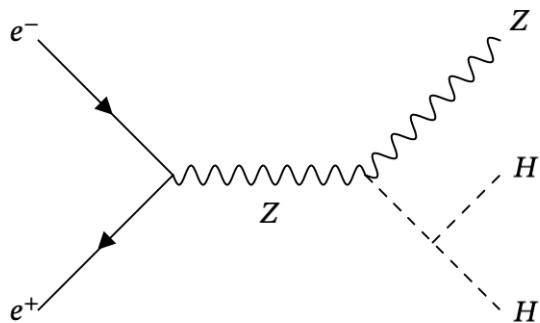


図 1.9: ヒッグス自己結合 $e^+e^- \rightarrow Zhh$

384

385 1.3.4 階層性問題

386 ヒッグス粒子の質量は LHC によって $125 \text{ GeV}/c^2$ と測定されている。しかしヒッグス粒子
 387 の質量は、繰り込みにおいて図 1.10 のような高次ダイアグラムから質量補正を受けることで
 388 発散してしまい、プランクスケール (10^{19} GeV) 程度の質量を持つてしまうことが分かっている。
 389 そのため標準模型を超える新物理による質量の量子補正をキャンセルする解決策がなけれ
 390 ば、量子補正を受ける前の質量がプランクスケール程度の質量であり、それに対して偶然プラ
 391 ンクスケール程度の負の量子補正がかかって $125 \text{ GeV}/c^2$ の質量を再現しているということに
 392 なってしまう（ファインチューニング）。これを回避するために、以下に挙げるような TeV ス
 393 ケールの超対称性理論や余剰次元理論など新物理によるシナリオが提案されている。これらの
 394 シナリオにおけるヒッグス粒子との結合定数は標準模型における予測からズレることとなるた
 395 め、ILC においてヒッグス粒子の精密測定を行うことで崩壊分岐比を決定することの意義は大
 396 きいと言える。

- 397 ● 超対称性理論

398 階層性問題においては、超対称性によりヒッグス粒子の質量補正に関する 2 次の発散
 399 をフェルミオンの寄与で打ち消す。これによって、対数による発散に落とすことができる。

400

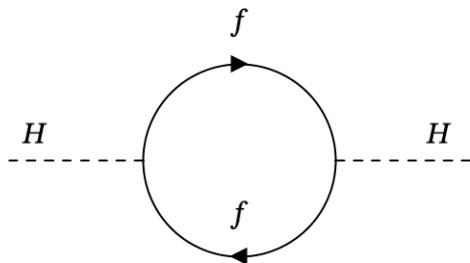


図 1.10: ヒッグス粒子の質量補正となるフェルミオンループ

401 ● 余剰次元理論

402 余剰次元理論とは、四次元時空以外にも次元があるとする理論である。この理論では空
403 間の次元数を増やすことで、増えた次元のゲージ場にヒッグス場の起源を求める。こ
404 の場合にはゲージ不变性により、繰り込みの発散が現れないため階層性問題に対応で
405 きる。

406 **1.3.5 その他の新物理**

407 上にあげた階層性問題に関する物理に加え、トップクォークの質量の精密測定、電弱相互作
408 用の精密検証が可能であり、ILC の実現やそのアップグレードを通して宇宙の謎に迫る大発見
409 を期待できる。

410 **1.4 ILC の検出器**

411 ILC の検出器(図 1.11)には、日本や欧米諸国が中心となって開発が進められている
412 International Large Detector (ILD) と、米国が中心となって開発が進められている Silicon
413 Detector (SiD) の二つのコンセプトが提案されており、ILC ではこれら 2 つの検出器が IP
414 を共有できるように push-pull 方式を採用している。また ILD、SiD ともに Particle Flow
415 Algorithm (PFA) という事象再構成アルゴリズムに沿って最適化されている。

416 **1.4.1 ILD で検出器可能な粒子群**

417 ILC の電子陽電子衝突で生じる粒子は、ヒッグス粒子などの重いボソンを介した後に、
418 クォークやグルーオン、レプトン、光子に崩壊する。この中でもクォークとグルーオンは
419 QCD における閉じ込めにより単体で存在することができず、クォーク対となって多数のハド
420 ロン（強い相互作用で結合した複合粒子）を発生させる。また、レプトンのうちニュートリノ
421 は、ILD の検出器と相互作用を起こさず通り抜けてしまうため、直接検出することは出来な
422 い。そのため実際に測定器で検出することができるのは、終状態としてハドロンが束になった

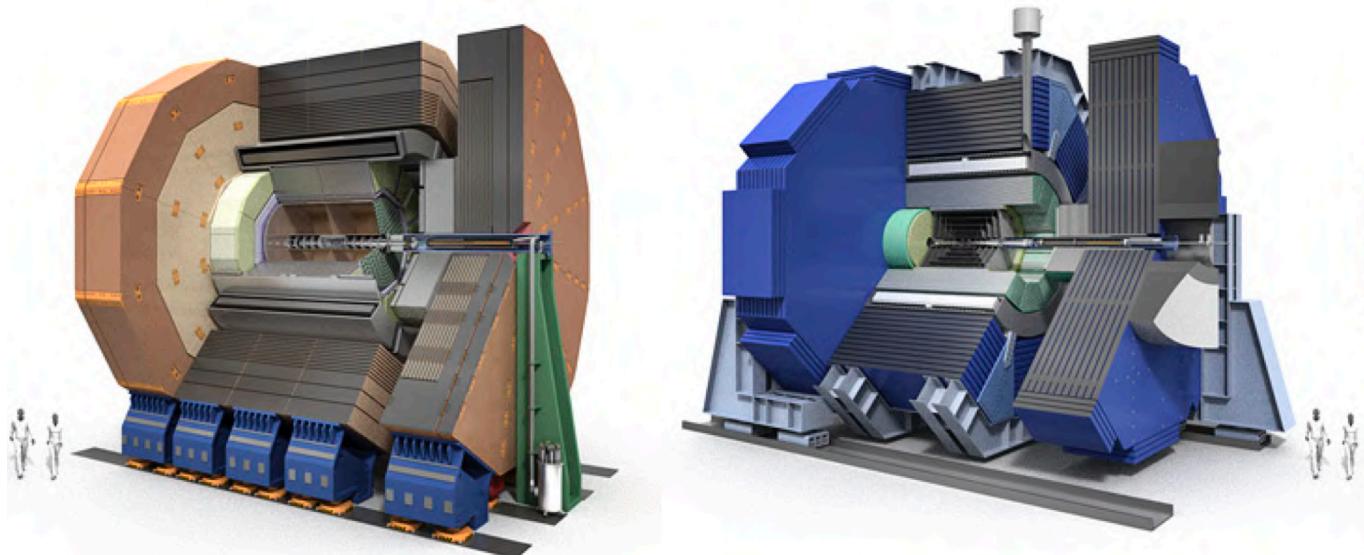


図 1.11: (左) ILD (右) SiD の全体図 [4]

₄₂₃ QCD ジェット、レプトン、光子となる。

₄₂₄ 1.4.2 Particle Flow Algorithm; PFA

前節の通り ILC で生成されるクォークやグルーオンは、ジェットの終状態として検出される。そのため ILC の物理を探求する上で必要となる粒子識別や事象再構成において、ジェットのエネルギーは非常に重要な情報であり、一般的にジェットのエネルギー分解能は

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \quad (1.5)$$

₄₂₅ として表される。ここで、 a, b, c を係数に E はエネルギーを表しており、第一項はカロリメータに由来する統計項を、第二項は補正項、第三項は定数項となっている。従来の素粒子実験におけるエネルギー測定では、およそ 7 割に相当する粒子がハドロンカロリメータでエネルギーを測定されるが、ハドロンカロリメータはそれ以外の検出器と比較してエネルギー分解能が低く、およそ $\sigma_E/E = 55\%/\sqrt{E(\text{GeV})}$ の分解能となっている。ILC ではジェットエネルギー分解能 $\sigma_E/E = 30\%/\sqrt{E(\text{GeV})}$ を目指しており、これを達成するために導入されているアルゴリズムが Particle Flow Algorithm である。PFA はジェット内の粒子をその種類ごとに最適な検出器でエネルギー測定を行うことでジェットエネルギー分解能を向上させる再構成手法であり、過去の実験結果から、ジェット中に含まれる主な粒子の種類の典型的な割合 ($Z \rightarrow q\bar{q}$ の場合) とそれに対応する検出器について以下のように分かっている（表 1.2）。

粒子のタイプ	検出器	ジェット中のエネルギー割合
荷電粒子	飛跡検出器	62%
光子	ECAL	27%
中性ハドロン	HCAL	10%
ニュートリノ	-	1%

表 1.2: ジェットを占める各粒子と対応する検出器

435 1.4.3 International Large Detector: ILD

436 ILD [4] は内側から順に崩壊点検出器、飛跡検出器、電磁カロリメータ、ハドロンカロリメー
437 タ、ミューオン検出器で構成されている。カロリメータとミューオン検出器の間には 3.5 T の
438 ソレノイドコイルが設置されている。図 1.12 に断面図を示す。

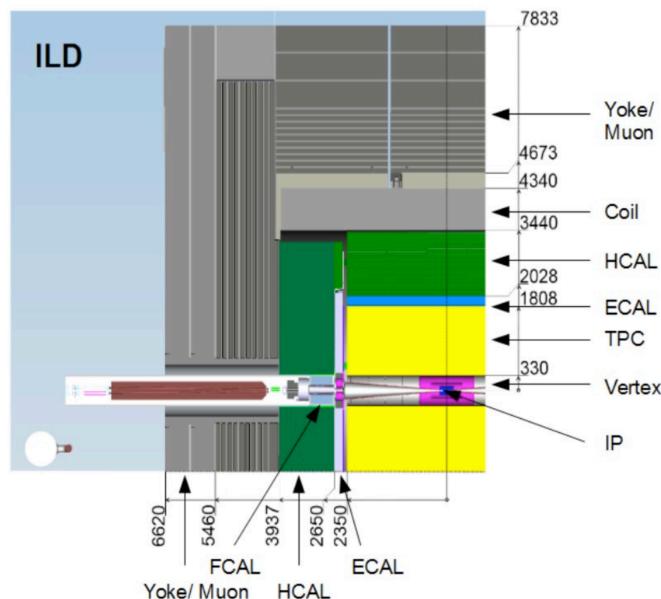


図 1.12: ILD の断面図

439 崩壊点検出器

崩壊点検出器は IP に最も近い場所に置かれる検出器であり、図 1.13 のようにシリコンピクセルセンサーが両面に貼られた層を ILD の半径方向に 3 層重ねた構造になっている。このシリコンピクセルセンサーで飛跡を高い位置分解能のもと検出することで、荷電粒子の生成点を高い精度で決定することができる。これによって短寿命粒子の崩壊点を高精度に再構成することができ、再構成された二次崩壊点はジェットのクォークフレーバーを識別する上で非常に重要な情報となる。ILC では飛跡検出における位置分解能 σ (式 1.6) を目標としており、そのた

めにCMOSセンサー、DEPFET、Fine Pixel CCD、SOIなど様々な技術候補が研究されている[4]。

$$\sigma[\mu m] = 5 \oplus \frac{10}{p[\text{GeV}/c] \sin^3 / 2\theta} \quad (1.6)$$

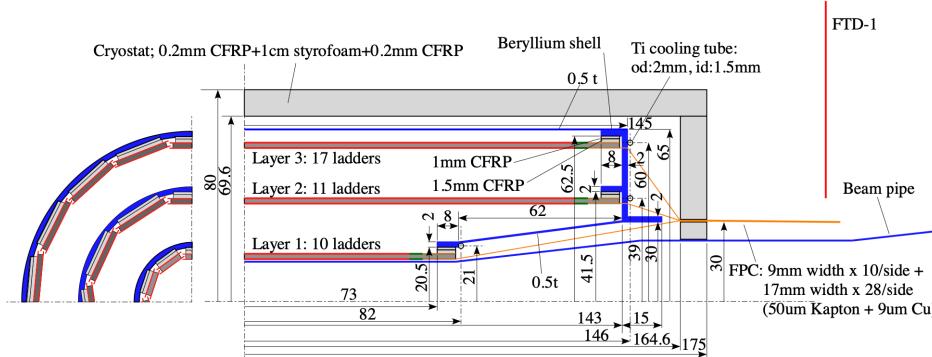


図 1.13: ILD 崩壊点検出器の構造

440 中央飛跡検出器

441 中央飛跡検出器は崩壊点検出器の外側に位置しており、Time Projection Chamber (TPC)
442 とその周囲に設置されるシリコン検出器のハイブリッドで構成されている。TPC は大型のガ
443 スチャンバーであり、荷電粒子の通過でガス内に生じる電離電子を電極間にかけられた電場に
444 よってドリフトし、ドリフト時間などの情報をもとに飛跡を 3 次元的に再構成する検出器であ
445 る。荷電粒子の飛跡を再構成することで運動量の測定が可能であり、粒子の検出数の多さから
446 高い運動量分解能を持つ。さらに信号の大きさからエネルギー損失も測定することが可能であ
447 り、これは粒子識別において重要な役割を果たす。

448 カロリメータ

449 カロリメータは入射粒子のエネルギーを測定するための検出器で、ILD では内側から電磁カ
450 ロリメータ (ECAL)、ハドロンカロリメータ (HCAL) によって構成されており、またビーム
451 軸方向に対して前方カロリメータ (FCAL) が設置される。これら ILD のカロリメータには
452 サンプリング型カロリメータが提案されており、シャワーを起こすための吸収層と生成された
453 シャワー内の粒子のエネルギーを測定する検出層が交互に組み合わさった構造となっている。

454 ECAL は主に電磁シャワー内の光子のエネルギーを測定するために利用される。ILD では
455 後述の PFA のためジェット内の粒子を分離できる高精細なカロリメータが必要とされてお
456 り、吸収層には物質量が大きいため放射長が短く、モリエール半径の小さいタンゲステンが検
457 討されている。また、検出層には読み出しセルが高精細なシリコン検出器を用いるシリコン
458 電磁カロリメータ (SiECAL) やシンチレータストリップを用いるシンチレータカロリメータ
459 (ScECAL) が提案されている。

460 HCAL は荷電ハドロンと中性ハドロンのエネルギー損失を分離し、中性ハドロンのエネル
461 ギーを測定するための検出器である。HCAL ではハドロンとの相互作用を起こすための吸収
462 層に鉄が用いられ、検出層には 3 cm 角のシンチレータタイルを用いてシンチレーション光を
463 検出するアナログカロリメータ (AHCAL) と、1 cm 角のガラス抵抗板チェンバー (RPC) を
464 用いて 2 bit の信号で読み出すセミデジタルカロリメータ (SDHCAL) の 2 つが提案されて
465 いる。

466 ミューオン検出器

467 ミューオン検出器はその名の通りミューオンを検出する検出器である。ミューオンは他の検
468 出器と相互作用を起こさないため IP から遠い検出器の最も外側に設置されており、RPC と
469 SiPM シンチレータストリップの両方が検討されている。

470 1.5 ILC のソフトウェア

471 1.5.1 イベントジェネレータと検出器シミュレーション

472 ILC をはじめとする線型加速器には「iLCSoft」というソフトウェアフレームワークが開発
473 されており、検出器シミュレーションから事象再構成までを実行することができる。iLCSoft
474 内では専用の LCIO フォーマットを使用し、C++ アプリケーションフレームワークである
475 Marlin によって運用され、検出器のジオメトリなど検出器記述には DD4hep というツール
476 キットを使用するという点で統一されている。

477 ILC は将来実験計画であるため、現在実験データは存在しないがシミュレーションによっ
478 て検出器応答や新物理探索など研究が可能であり、本論文における研究で扱うデータもモンテ
479 カルロ (Monte Carlo; MC) 法に基づくソフトウェア群を用いて生成したシミュレーション
480 データである。図 1.14 に ILC におけるソフトウェアの流れを示す。シミュレーションではま
481 ず、Whizard というイベントジェネレータを用いて、標準模型や様々な理論を背景とした物理
482 事象を生成する。Wizard では終状態で最大 8 粒子までの事象を生成し、Pythia によって粒
483 子の崩壊過程のシミュレーションを行う。更に生成されたイベントに対して、DDSim という
484 Geant4 をベースとした検出器シミュレーションを実行し、粒子から検出器ヒットデータが生
485 成される。

486 1.5.2 事象再構成

487 前節までで生成された検出器ヒットをもとに、終状態粒子のエネルギーや飛跡を推定する事
488 象再構成が行われる。iLCSoft では、Marlin によって測定器出力のデジタル化や飛跡再構成、
489 PFA を行ったのちにジェットの再構成を実行し、物理解析へと繋げていく。iLCSoft におい
490 て、ジェットの再構成にあたる崩壊点検出からフレーバー識別のプロセスは、LCFIPplus [6] と
491 いうフレームワークで実行することができる。

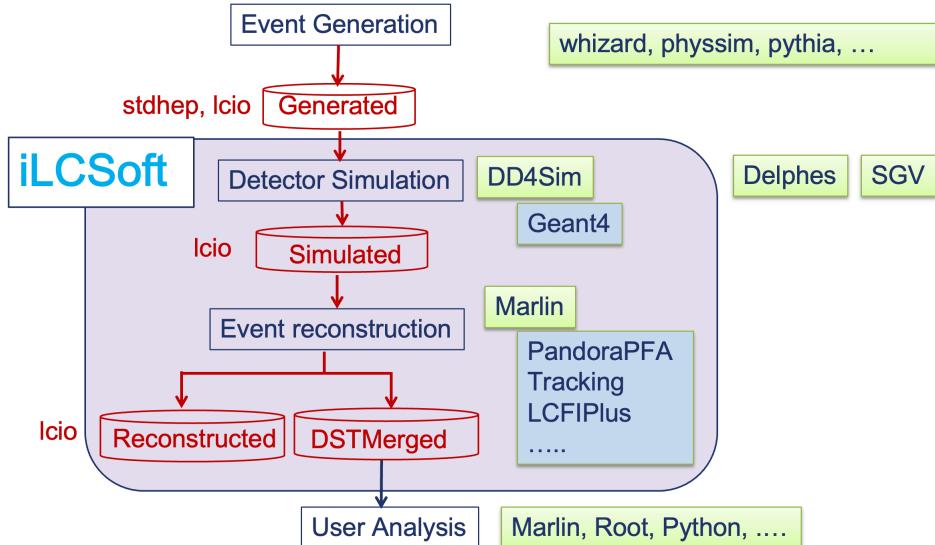


図 1.14: ILC におけるシミュレーションと事象再構成、物理解析の流れとソフトウェア

飛跡再構成

多数の飛跡を含むジェットは各検出器を通過するため、検出器ヒットをもとにフィッティングを行うことで飛跡は再構成することができる。特に崩壊点検出器では主に飛跡の方向情報を、中央飛跡検出器では運動量や時間情報を取得し、様々なパターン認識アルゴリズムを有する MarlinTrk によって再構成される。荷電粒子の飛跡が再構成されたのち、以降のプロセスでジェットの再構成を行う。

崩壊点検出

ジェットは IP で生成された粒子が崩壊を繰り返し、多くの飛跡を残すことで再構成される。この粒子が崩壊する点を崩壊点 (Vertex) と呼び、特に IP を primary vertex、そこで生成された粒子の二次崩壊点を secondary vertex と呼ぶ。LCFIPlus における崩壊点検出では、主に 2 本以上の飛跡の組み合わせから飛跡の発生源となる点を、 χ^2 値が最小値をとるフィッティングによって求める Vertex Fitter プロセスを用いる。この情報をもとに、primary vertex の再構成を tear-down アルゴリズムを用いて行う。tear-down アルゴリズムでは、全ての飛跡に対して IP を崩壊点とするフィッティングを Vertex Fitter によって行い、一定の閾値に到達するまで χ^2 値の大きい飛跡から排除していく。そして primary vertex に使用されていない全ての飛跡に対して再びフィッティングを行い、 χ^2 値や不变質量、運動量などを用いたカットや、飛跡同士の関係性の評価を行い、secondary vertex を再構成している。

509 ジェットクラスタリング

510 ILC の終状態の多くは 4 ジェット以上のマルチジェットであり、これらを幾つかのグルー
511 プに分ける（クラスタリング）することでフレーバー識別の精度を向上させることができる。
512 そのためジェットクラスタリングのプロセスでは、再構成された崩壊点やレプトンの情報を
513 ジェットのコアとし、Durham[7] のクラスタリングアルゴリズムを用いてクラスタリングを行
514 う。具体的には、崩壊点やレプトンの特徴的な物理量や飛跡同士の開き角とエネルギーを用
515 いて、全飛跡に対してクラスタリングを行う。

516 フレーバー識別

517 崩壊点の情報や飛跡の情報など 20 程度の物理量をもとに、多変量解析によってジェットの
518 親粒子のフレーバー識別が行われる。LCFIPlus では ROOT[8] の TMVA[9] パッケージを使
519 用して、従来の機械学習手法である Boosted Decision Trees (BDTs) を用いた識別が行われ
520 ている。LCFIPlus では b フレーバーのジェット、c フレーバーのジェット、u・d・s フレー
521 バーのジェットの 3 つを分類している。分類においては、b・c フレーバーのハドロンは衝突
522 点から離れたところで崩壊するという特徴がある。（図 1.15）ハドロンの寿命 τ は光速 c を用
523 いて、b フレーバーでおよそ $c\tau = 400 \sim 500 \text{ um}$ 、c フレーバーでおよそ $c\tau = 20 \sim 300 \text{ um}$
524 である。上記に加えて b フレーバーは派生して c フレーバーへ崩壊するため、1 つのジェット
525 中に崩壊点を 2 つもつ。これらの情報やジェットの物理量を用いてフレーバーを識別するこ
526 とができる。そのため崩壊点検出のプロセスにおいて、高精度に崩壊点を求めることが重要と
527 なる。

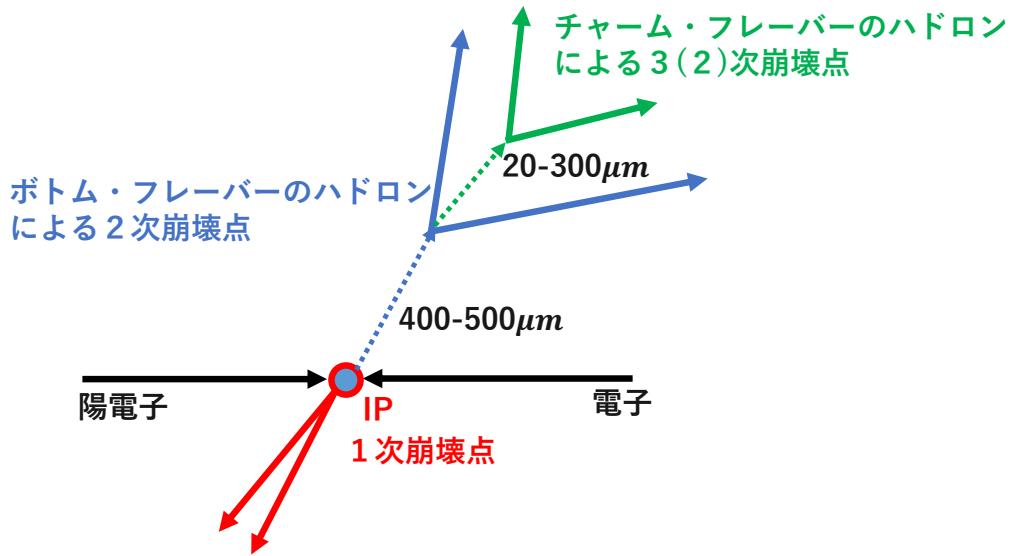


図 1.15: ジェットの崩壊の様子。赤線が 1 次崩壊点、青線が 2 次崩壊点、緑線が 3 次崩壊点を由来とする飛跡を表す。

1.6 本研究の目的

本研究の目的は、電子陽電子ヒッグスファクトリーである ILC におけるジェット測定技術の開発である。ヒッグス粒子や 350 GeV 以上の ILC で探索可能なトップクォークなど、ILC の物理において重要な事象の多くはジェットを含んでおり、ジェットを高い精度で測定することが物理解析の性能に直結する。そのジェット測定技術において、シリコンタングステン電磁カロリメータの性能評価、フレーバー識別アルゴリズムの開発の 2 つのテーマで研究を行った。

535

1.6.1 高エネルギービームによる SiW-ECAL の性能評価

ILC では粒子単位での再構成が求められ、そのためにはジェットのエネルギーと方向の分解能を高い精度で得ることが重要となる。高いジェットエネルギー分解能を達成するために提案されている PFA では、ジェット中の粒子を分離しクラスタリングが可能になるほどに高精細なカロリメータが求められており、日本やフランスのグループによって開発されたシリコンタングステン電磁カロリメータはその有望な候補である。本研究では、その技術プロトタイ

542 プに高エネルギービームを照射するビームテスト実験を行い、実験結果から更なる改善に向け
543 たフィードバックを得た。本論文では、2章で背景となる物理や SiW-ECAL の仕組みについ
544 て述べ、3章でビームテスト実験について報告する。

545

546 1.6.2 深層学習を用いたフレーバー識別アルゴリズムの開発

547 再構成において重要なプロセスとなるフレーバー識別アルゴリズムの開発を、深層学習技
548 術を用いて行った。本研究では、特にグラフ構造のデータを扱うグラフニューラルネットワー
549 クを実装し、従来技術である LCFIPlus と比較した識別精度の向上を目指した。グラフニュ
550 ラルネットワークでは、対象の特徴量に加え構造のトポロジー情報をデータに含んだ学習を行
551 うことができる。また同時に、全ての事象再構成アルゴリズムを深層学習に置き換えることを
552 最終目標に、フレーバー識別アルゴリズムと崩壊点検出アルゴリズムの統合を試みた。本論文
553 では、4章で深層学習に関する基本的事項について述べ、5章で実際に考案したアルゴリズム
554 について述べる。

555 第2章

556 シリコンタングステン電磁カロリ 557 メータ

558 本章では、ILD のシリコンタングステン電磁カロリメータについて説明する。まず検出器を
559 理解する上で必要な粒子と物質の相互作用について述べたのち、カロリメータの検出原理やシ
560 リコン検出器の検出原理について述べる。そして現在の ILD におけるシリコンタングステン
561 電磁カロリメータの読み出し方法、また ASIC の設計性能や読み出し方法、現在の技術プロト
562 タイプについて説明する。

563 2.1 入射粒子と物質の相互作用

564 素粒子実験で捉えたい素粒子やハドロンは、粒子と物質との相互作用によって捉えることが
565 できる。よって本節では入射粒子の種類ごとに物質との相互作用について述べる。

566 2.1.1 荷電粒子

567 荷電粒子のエネルギー損失の要因には、主に電離損失と制動放射が挙げられ、特にエネル
568 ギーの低いところでは電離損失の割合が高くなる。(電子ではエネルギーの高いところでは制
569 動放射の割合が高くなる。) 以下ではそれについて述べる。

570 電離損失

571 荷電粒子は物質を通過することで、物質中の原子を電離あるいは励起させ電離エネルギー損
572 失を生じる。この電離エネルギー損失は原子中の電子によるクーロン散乱によるものが支配的
573 であり、Bethe-Bloch の式に従う。

$$-\frac{dE}{dx} = 4\pi N_A r_e^2 m_e c^2 z^2 \frac{Z}{A} \frac{1}{\beta^2} \left[\frac{1}{2} \ln\left(\frac{2m_e c^2 \beta^2 \gamma^2 W_{\max}}{I^2}\right) - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right] \quad (2.1)$$

変数	値または単位
N_A : アボガドロ定数	$6.022 \times 10^{23} \text{ mol}^{-1}$
r_e : 古典電子半径	2.817 fm
m_e : 荷電粒子の質量	0.511 MeV
c : 光速	$2.998 \times 10^8 \text{ m/s}$
z : 入射粒子の電荷	-
v : 入射粒子の速度	m/s
Z : 物質の原子番号	-
A : 物質の相対原子質量	g mol^{-1}
β : 入射粒子の v/c	-
γ : $1/\sqrt{1 - \beta^2}$	-
W_{max} : 1回の衝突で物質に与える最大エネルギー	MeV
I : 物質の平均イオン化ポテンシャル	eV
$\delta(\beta\gamma)$: 密度効果による電離エネルギー損失の補正	$\sqrt{\rho \langle Z/A \rangle} \times 28.816 \text{ eV}$

574 Bethe-Bloch の式より、電離エネルギー損失 $-dE/dx$ は荷電粒子の入射速度に依存する。
 575 様々な物質に対する電離エネルギー損失と入射速度の関係を図 2.1 に示す。入射速度の小さい
 576 とき電離エネルギー損失は $1/\beta^2$ に比例しており、 $\beta\gamma \approx 3 \sim 4$ で電離エネルギー損失は最小
 577 値に達する。これを最小電離損失といい、この領域にある粒子を MIP (Minimum Ionization
 578 Particle) と呼ぶ。

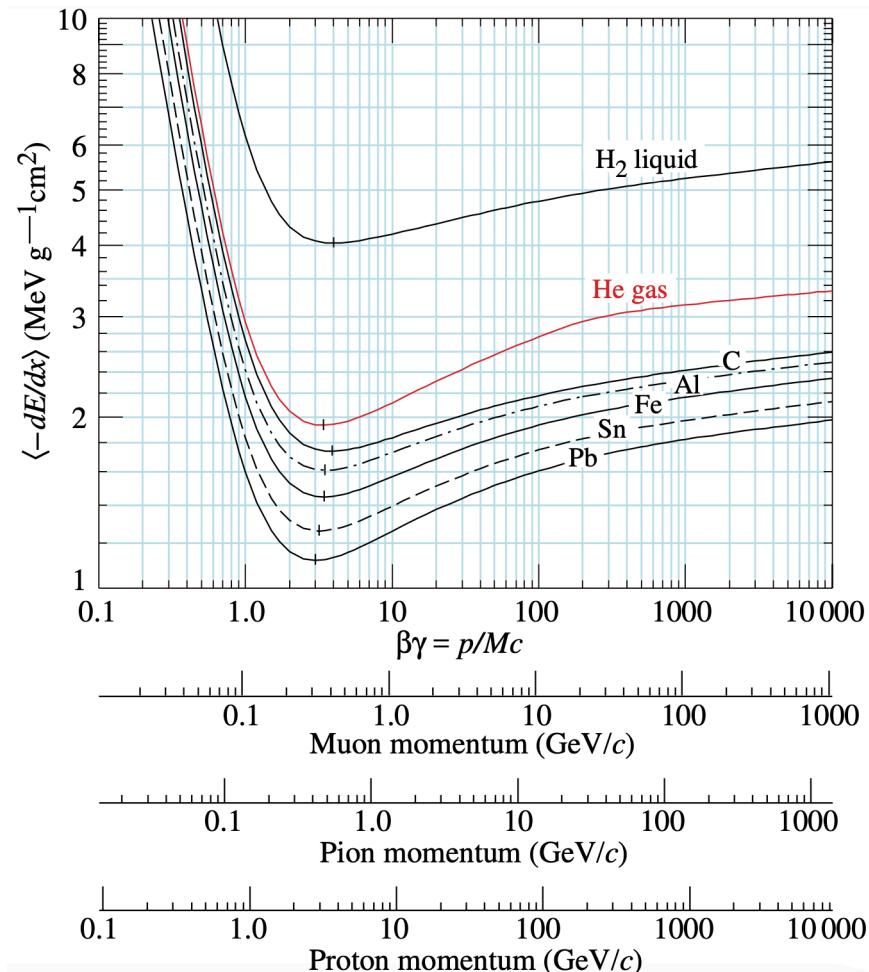


図 2.1: 水素(液体)、ヘリウム(気体)、炭素、アルミニウム、鉄、スズ、および鉛における平均エネルギー損失と、入射粒子(ミューオン、パイ中間子、陽子)の速度の関係。

579 制動放射

580 荷電粒子が物質を通過する際には電離の他に、原子核との衝突によって電磁波を放射しエネルギーを失うこともある。物質を構成する原子核はそれぞれ電場を持っており、電場によって
581 Rutherford 散乱を受けた荷電粒子は加速、減速をされ、光子を放射しエネルギーを失う。これ
582 を制動放射と呼ぶ。制動放射によって荷電粒子が失うエネルギー損失率は、
583

$$-\frac{dE}{dx} = \frac{E}{L_R} \quad (2.2)$$

L_R は放射長と呼ばれています。平均エネルギーが e の因子だけ小さくなる平均の長さを指す。
 L_R は以下のように与えられます。

$$\frac{1}{L_R} = 4 \left(\frac{\hbar}{mc} \right)^2 Z(Z+1) \alpha^3 n_\alpha \ln \left(\frac{183}{Z^{1/3}} \right) \quad (2.3)$$

式 (2.2) を積分することで、初期エネルギー E_0 を持った荷電粒子が物質を x だけ進むときのエネルギー損失は以下のようになる。

$$E = E_0 \exp(-x/L_R) \quad (2.4)$$

電離エネルギーと制動放射によって失うエネルギーの大きさが同じになる入射電子のエネルギーを臨界エネルギー E_c と呼び、この値よりもエネルギーが小さい場合はエネルギー損失が Bethe-Bloch の式に従い、大きい場合は制動放射によって主にエネルギーを失う。(臨界エネルギーは物質によって異なるがおおよそ $E_c \simeq 500 \text{ MeV}/Z$ と表される。) 荷電粒子の制動放射によって失われるエネルギーは質量の二乗に反比例するが、電離エネルギー損失は質量に強く依らないため、ほとんどの粒子に対しては制動放射よりも電離損失が支配的となる。一方で電子においては制動放射によるエネルギー損失が支配的であり、 E_c は電磁カロリメータなどの設計において重要なパラメータとなる。

2.1.2 光子

電荷を持たない光子は物質中で電離は起こさず、主に図 2.2 に示す光電効果、コンプトン散乱、電子陽電子対生成の 3 つの過程で相互作用する。

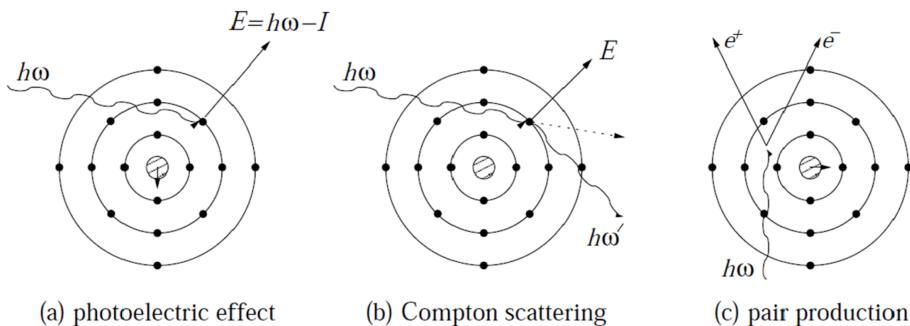


図 2.2: 光子と物質の相互作用

- 光電効果 (Photoelectric effect): 入射光子が物質に当たることで光子の持っていたエネルギー $h\nu$ が物質の電子に与えられる。これによって励起された電子が $h\nu - I$ の運動エネルギーで飛び出す現象。 $(I$ はイオン化エネルギー)
- コンプトン散乱 (Compton scattering): 入射光子と原子核に束縛されている 1 つの電子との弹性散乱。光電効果よりも光子のエネルギーが大きく、電子陽電子対生成反応よりも小さい時に支配的な反応である。
- 電子陽電子対生成 (Pair production): 入射光子が原子核のつくるクーロン場において消滅し、電子陽電子の対を生成する反応。この反応では、光子のエネルギーが電子陽電子の静止質量の和（およそ 1.02 MeV）よりも大きい必要がある。

光子のエネルギーによってこれらの反応確率は異なり、図2.3に光子のエネルギーに対する各反応の確率を示す。中でも数MeV以上の光子においては電子陽電子生成反応が主要なプロセスであり、ILCのような高エネルギーにおいては電子陽電子対生成が重要である。物質に入射した光子は電子陽電子を生成し、さらに制動放射によって光子を放出する。これを繰り返すことで電子陽電子と光子の数が指数関数的に増加していく、この現象が電磁シャワーと呼ばれている。この時、電子陽電子対生成過程の断面積は $E_\gamma \gg mc^2/\alpha Z^{1/3}$ において

$$\sigma_{pair} = \frac{7}{9} \frac{1}{n_a L_R} \quad (2.5)$$

と近似することができ、光子の飛程はおよそ $9/7 L_R$ となる。電磁シャワーは発展するにつれて光子の平均エネルギーが下がり、臨界エネルギー（ILCの検出器ではおよそ 10 MeV）に到達すると電子陽電子生成過程が起こらなくなり終息する。粒子のエネルギーを測定する場合には、シャワー内の荷電粒子を MIP とみなし、それらの粒子数が初めの光子のエネルギーに比例すると考えることで、検出器のセンサーに残したエネルギー損失の和をとることで測定する。

また電磁シャワーは進行方向だけでなく垂直方向にも広がり、モリエール半径 R_M によって広がりが測られる。モリエール半径とは、エネルギーの 90% が入るシャワーの半径を指し、以下の式で表される。

$$R_M \sim \frac{21(\text{MeV})L_R}{\text{臨界エネルギー } (\text{MeV})} (\text{g/cm}^2) \quad (2.6)$$

615

616 2.1.3 ハドロン

π 中間子や K 中間子などのハドロンは物質を構成する原子核と衝突し、非弾性散乱を繰り返すことでハドロンシャワーを生成する。ハドロンの相互作用長は典型的に放射長よりも大きく、ハドロンをカロリメータで測定する場合には非常に多くの物質を必要とする。相互作用長は以下のように表される。

$$\lambda = \frac{A}{N_a \rho} \sigma_{total} \quad (2.7)$$

617 ここで、 ρ は物質の密度を、 σ_{total} は反応断面積の総和を表す。ハドロンシャワーには $\pi^0 \rightarrow \gamma\gamma$
618 崩壊によって発生する電磁シャワーが混ざってしまっており、検出器のエネルギー応答が異なる
619 ことから、ハドロンシャワーのエネルギー分解能は電磁シャワーと比較して非常に悪くなってしまう。
620

621 2.2 粒子検出器の動作原理

622 素粒子実験では、前節の相互作用を用いて粒子を検出する。粒子の検出には、事象を区別するためには十分な時間分解能と位置分解能を持つ必要があり、また各粒子を識別するために、工

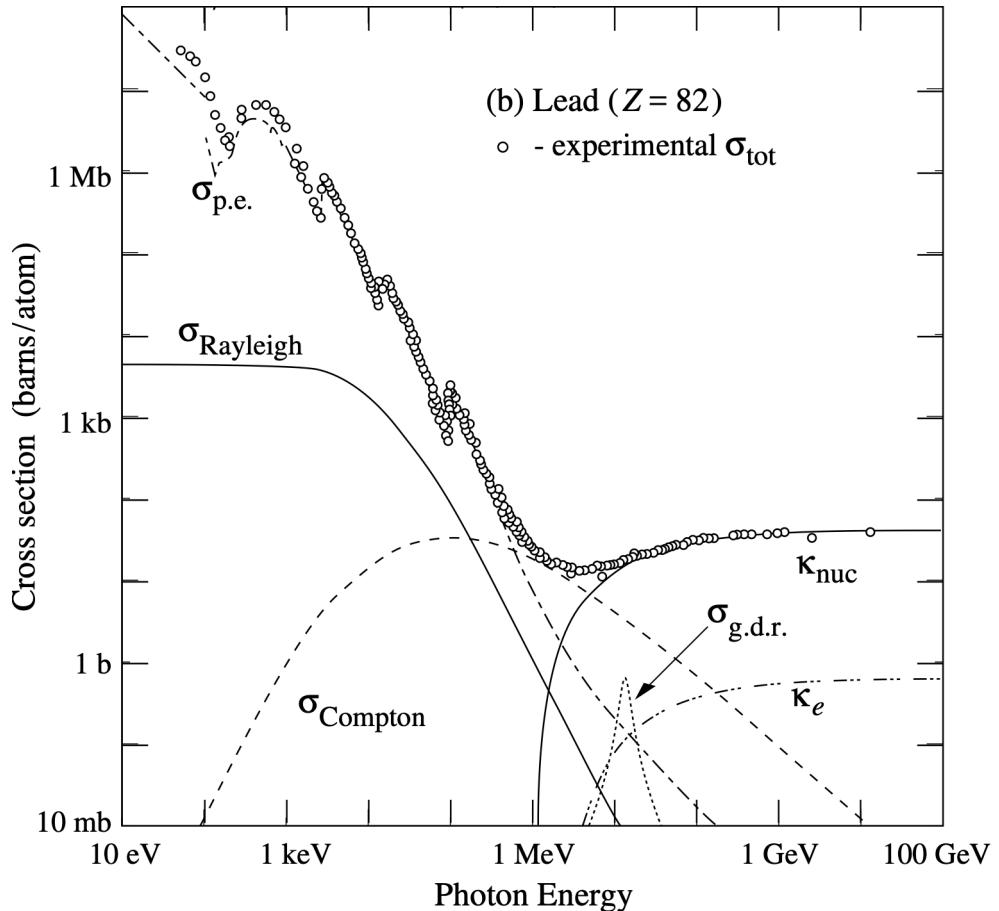


図 2.3: 光子と鉛の相互作用における断面積と光子のエネルギーの関係。 $\sigma_{p.e.}$ は光電効果を、 $\sigma_{Compton}$ はコンプトン散乱を、 κ_e, κ_{nuc} は電子陽電子対生成を指す。

624 ネルギーと運動量を十分な精度で測定する必要がある。以下では、測定器を構成する検出器の
625 うち特に重要なものを取り上げる。

626 2.2.1 ガス検出器

627 ガス検出器は、主にアルゴンのような活性の低いガスを検出器内に充填した検出器である。
628 荷電粒子がガス中を通過することで電離反応を起こし、生成される電子と陽イオンを電極に集
629 める、あるいは電離の軌跡を可視化することで荷電粒子を検出することができる。主な検出器
630 としては、電極への印加電圧が小さい領域では電離箱が、大きい領域ではワイヤーチェンバー
631 や RPC などが挙げられる。

632 2.2.2 半導体検出器

633 半導体検出器とは半導体材料を使用した検出器を指し、光や電子をはじめとして様々な粒子
634 を検出するものがある。半導体材料には主にシリコンやゲルマニウムが用いられ、接合ダイ

635 オードの原理を使用して製造される。以下では、シリコン検出器の構造と動作原理について説
636 明する。まず基本構造としては、検出器の一方に正孔の多い p 型半導体、もう一方の面に自由
637 電子が多い n 型半導体の p-n 接合が作られてある。それぞれに対して逆バイアス電圧（p 型に
638 負、n 型に正）を印加することで、p 型と n 型との間で正孔と自由電子の結合が進み、空乏層
639 と呼ばれる安定化した領域が検出器の接合面を中心に広がる。この空乏層を通過した荷電粒子
640 は、検出器内のシリコン原子を励起し、電子正孔対を生成する。飛跡に沿って生成された電子
641 正孔対は、空乏層内の電場によって両印加極板までドリフトされ、パルス電流として測定され
642 る。また、シリコンなど半導体のバンドギャップは 1eV 程度であり、電子正孔対を生成する
643 ために必要なエネルギーはおよそ 3,4eV となっている。この信号の小ささから、半導体検出
644 器では読み出しにおいて増幅を必要とするため、増幅回路が近傍（あるいは半導体内部）に存
645 在している。半導体検出器は、電極が平面構造のピクセル検出器や帯状のストリップ検出器、
646 検出器基板上に増幅回路を形成するモノリシック検出器など、電極や増幅回路の実装によって
647 様々な構造が存在する。

648 2.2.3 シンチレーション検出器

649 励起エネルギーの一部が、より低いエネルギー準位へ遷移する際に可視光として表れる物質
650 をシンチレータという。シンチレータ検出器は、荷電粒子の通過によって発生した蛍光（シン
651 チレーション光）を光検出器によって測定することで動作する検出器である。シンチレーション
652 光は非常に弱い光信号であるため、光電子増倍管や半導体光センサーなどを用いて信号を検
653 出する。

654 2.3 シリコンタングステン電磁カロリメータ SiW-ECAL

655 2.3.1 SiW-ECAL の全体構造

656 ILD の SiW-ECAL は、図 2.5 のようにタングステンの吸収層とシリコンパッドセンサーの
657 検出層が、30 層サンドウィッチ状に交互に重なったサンプリング型カロリメータである。1 つ
658 のモジュールが 10 ほどのサブモジュールに分かれており、サブモジュールにはそれぞれ 4 枚
659 のシリコン半導体センサーが貼り付けられる。

660 2.3.2 シリコン半導体検出器

661 SiW-ECAL では検出層にシリコン半導体検出器を用いる。センサーの大きさは 1 枚あたり
662 $9 \times 9 \text{ cm}^2$ で、1 枚に $5.5 \times 5.5 \text{ mm}^2$ のピクセルが 16×16 個並んでおり、1 つのサブモジュー
663 ルあたり 1024 チャンネル読み出しが可能となっている。またシリコンセンサーには、電極と
664 してアルミニウム (Al)、絶縁層に二酸化ケイ素 (SiO_2) が使用される。厚さ $320\mu\text{m}$ のセン
665 サーでは、1MIP あたり 86.8 keV のエネルギー損失が起こり、臨界エネルギーに達するまで

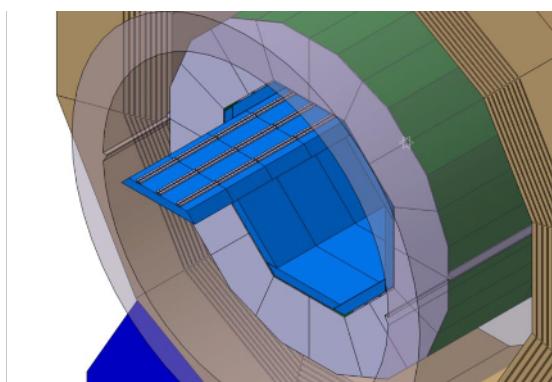


図 2.4: ILD および ECAL の全体図

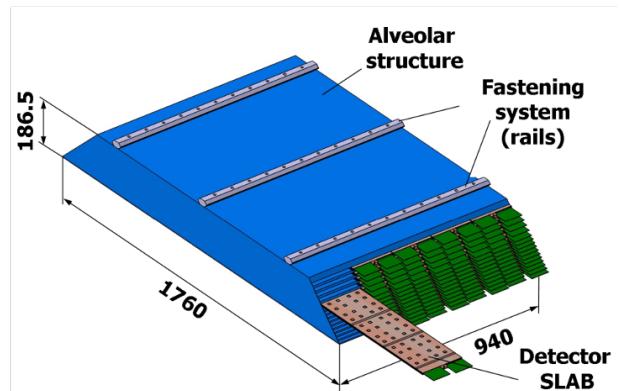


図 2.5: SiW-ECAL の構造

666 に生成される電子正孔対はおよそ 24,000、電荷にして 4 fC となる。シリコンセンサーは、常
667 溫硬化型導電性接着剤によって回路基板 (PCB) と接着されており、PCB を通して信号の読
668み出しが行われる。以下にセンサーの仕様と 1 枚のシリコンセンサーパッドを示す。

表 2.1: シリコンセンサーの仕様

制作会社	浜松ホトニクス株式会社
サイズ	89.7 × 89.7 mm ²
セルサイズ	5 × 5 mm ²
セル数	16 × 16 = 256
厚さ	320/500/650 μm
完全空乏化電圧	40/70/110 V

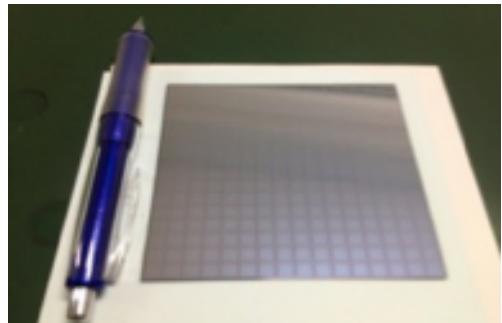


図 2.6: シリコンパッドセンサー

669 2.3.3 読み出しシステム

670 SiW-ECAL は高精細であることから読み出しチャンネル数が非常に多くなっており、30 層
671 の ECAL 全体ではおよそ 1 億にもおよぶ。そのため、シリコンセンサーからの信号読み出
672 しをコンパクトにする必要があり、読み出し専用の Application Specific Integrated Circuit
673 (ASIC) が開発された。現在の技術プロトタイプに実装されている ASIC には、フランスの
674 Omega グループが開発した Silicon Kalorimeter Integrated ReadOut Chip (SKIROC) シ
675 リーズの第二バージョンである SKIROC2A を用いている。

676 まず、読み出しに用いる ASIC に求められる性能には次に挙げる項目が求められる。

- 677 • 自動トリガー：
678 信号に対して ASIC 自身でトリガーをかける。
- 679 • 完全デジタル出力：

シリコン半導体検出器からのアナログ情報をすべてデジタル変換し、データ量を圧縮して Data Acquisition (DAQ) へ送信する。

● 発熱量の抑制 :

電力消費によって発生するジュール熱を、1 チャンネルあたり $25\mu\text{W}$ 以下に抑える。

● 1 MIP 相当の信号を識別できる高い Signal to Noise 比 (S/N 比) :

PFA においてジェットエネルギー分解能を向上させるために、高い精度でノイズからシグナルを分離する必要がある。

続いて、SKIROC2A の基本的な仕様を以下に示し、図 2.7 に SKIROC2A のアナログ部の回路図を示す。

● Austria Micro Systems 社製 0.35 um SiGe

● $7.5 \times 8.5\text{ mm}^2/1$ チップ

● 1 チップあたり 64 チャンネル読み出し可能

● 2 種類のダイナミックレンジをもつ Analog to Digital Converter (ADC) mode :

電磁シャワー内の粒子が 1 つのチャンネルに大量に入射したときに全エネルギーを測定出来るよう、幅広いゲインのレンジを持つ。

– High gain ... $0.5 \sim 150$ MIP 相当の信号に対応

– Low gain ... $150 \sim 2500$ MIP 相当の信号に対応

● TDC (Time to Digital Converter) mode : 1 ns 程の時間分解能で時間情報を保存

● 1 チャンネルあたり 15 イベント保持可能な Analog memory cell

ビームバンチ構造に対応するため、200 ns の間イベントを保持することが可能。1 つの Memory cell では、High gain ADC、Low gain ADC を、時間情報である Bunch crossing ID (BCID) と紐づけて保存。

● 数珠つなぎ型読み出し : 順番に読み出すことで読み出していないチップの電源を必要とせず、消費電力を減らす

● 0.5 MIP での自動トリガー

● 全 64 チャンネルの閾値を個別で同時設定可能な 10 bit Digital Analog Converter (DAC) threshold

● Power pulsing mode

データ収集の手順は、以下のとおりである。まずシリコンセンサーからのアナログ信号が各チャンネルに入力され、前置増幅器によって前段増幅を行う。ここでの増幅率は、前置増幅器の feedback capacitance によって変更可能となっており、増幅率が最大となる 0 pF から 6.0 pF まで 0.4 pF 刻みで決定することができる。前段増幅を経た信号は、Fast shaper と 2 つの Slow shaper (Low/High gain) の 3 つに分割される。Fast shaper に入った信号は、CRRC shaper によって信号をさらに増幅したのち、discriminator にて閾値を超えた場合のみトリガー信号を出し、同時に入ってきた Slow shaper の信号電圧を Memory cell に保持する。こ

こで、CRRC shaper とは微分回路と積分回路を組み合わせた回路であり、信号増幅率と立ち上がり時間を調整している。Memory cell に保持された電荷は、読み出し時間が経過、あるいは Memory cell が満たされたタイミングでマルチプレクサー (MUX) に送られる。一方、Slow shaper においても信号が CRRC shaper によって増幅率 1、10 倍に増幅され、トリガー信号によって MUX に送られる。MUX では、Slow shaper (Low or High) と TDC の組み合わせを決定し、Wilkinson 型 12 bit ADC によってアナログ信号をデジタル化し、メモリに保存されたのち外部へ転送される。

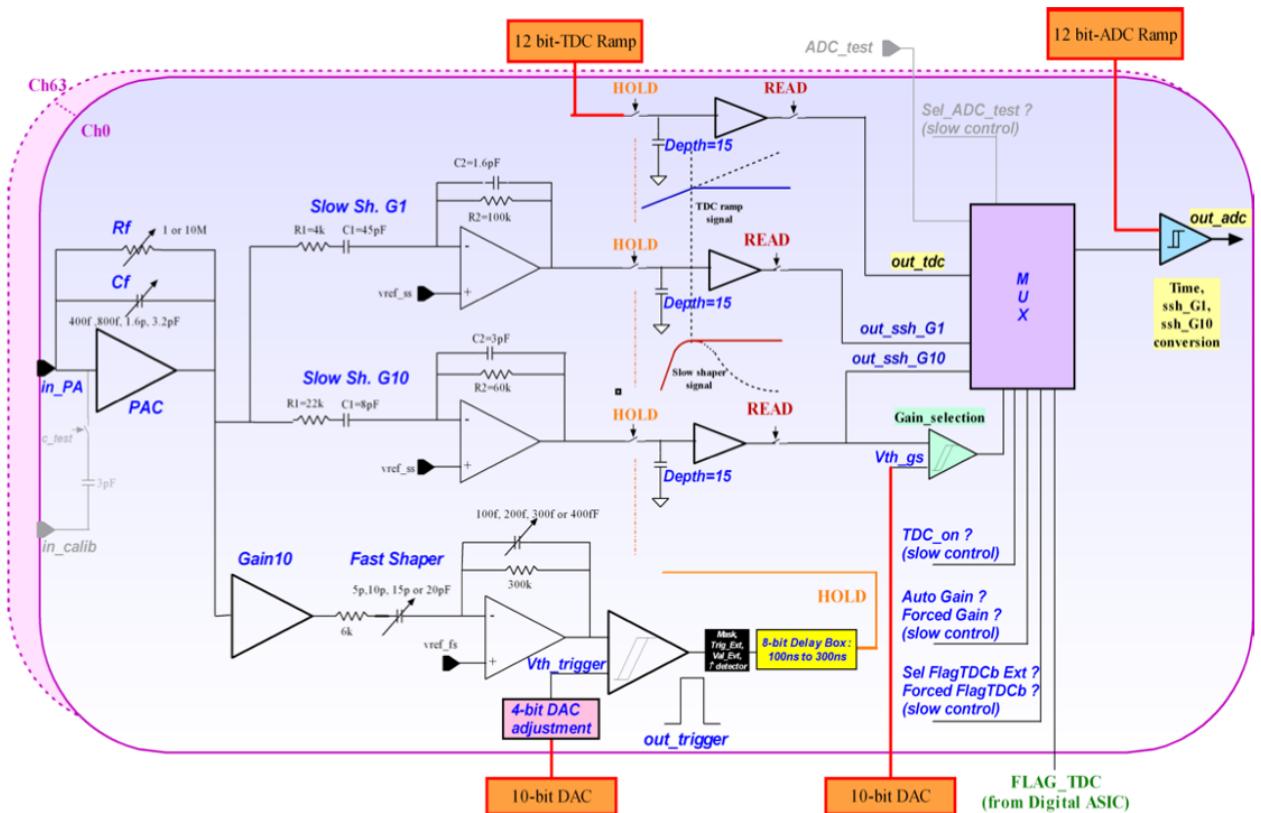


図 2.7: SKIROC2A のアナログ部の回路図

2.3.4 タングステン吸収層

ILD の SiW-ECAL の吸収層では PFA の要件を満たすために、相互作用長が長く放射長の短い物質でシャワーの広がりを小さく抑える物質を採用する必要がある。そのため、モリエール半径が小さく、放射長に比べ相互作用長が大きいタングステンが採用されている。表 2.2 に物質量が大きく吸収層の候補となる物質の性質について示す。

物質	λ/cm	L_R/cm	R_M/cm
鉄	16.8	1.76	1.69
銅	15.1	1.43	1.52
タングステン	9.6	0.35	0.93
鉛	17.1	0.56	1.00

表 2.2: 物質量の大きい吸収層の候補物質 (λ は相互作用長、 L_R は放射長、 R_M はモリエール半径を示す。)

727 2.3.5 技術プロトタイプ

728 SiW-ECAL の技術プロトタイプとして、図 2.8 のような構造が考えられている。図 2.8 は、
 729 1 層の検出器 (Slab) の一部として作製された Short slab である。図の上から構造体として炭
 730 素繊維強化プラスチック (CFRP) の板があり、その下に読み出し基板 (FEV) や ASIC の制
 731 御基板があり、基板下にはセンサーに接触し電荷を印加する導電性シートがあり、さらにカー
 732 ポンの板で挟まれている。1 層あたり、FEV には PCB 上に SKIROC2A が 16 チップ実装
 733 されており、裏面にはシリコンパッドセンサー 4 枚が導電性接着剤で接着されている。また、
 734 ASIC は FPGA を通して制御されており、さらに図 2.10 のに示すモジュールを通して、複数
 735 の Slab からの信号を同時に読み出している。

736 読み出し基板には、大きく分けて FEV と COB の 2 つの案がある。(図 2.9) FEV は PCB
 737 の裏面に 4 枚のシリコンセンサーを導電性接着剤で接着しており、その内側にフレキシブル
 738 基板を接着することでシリコンセンサーへの電圧供給を行う。また PCB 表面には ASIC が
 739 16chip 実装されており、センサーからの信号をデジタル変換している。FEV には FEV7 から
 740 13 までのバージョンがあり、内部配線や外部接続の面でアップデートが行われている。一方
 741 で COB (Chip On Board) では、ASIC が溝みにワイヤーボンディングされており、これに
 742 よって読み出し基板が FEV よりも薄いコンパクトな構造になっている点が特徴である。しか
 743 し同時に、FEV と比較してノイズの影響を受けやすいなど欠点も抱えている。

744 これら技術プロトタイプの開発は、フランスと日本を中心に CALICE グループによって国
 745 際協力で行われている。CALICE (Calorimeter for Linear Collider Experiment) は ILC に
 746 向け、カロリメータの共同開発を目的として結成された国際コラボレーションであり、九州大
 747 学のグループもこの一員である。また、モジュールの生産・組み立ては国内においても可能と
 748 なっている。

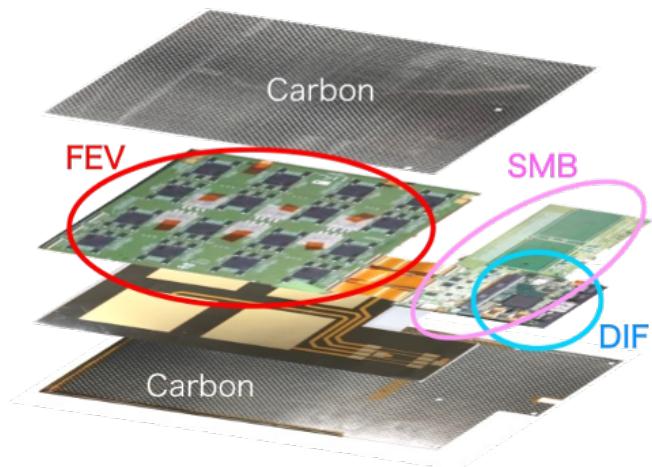


図 2.8: short slab の構造



図 2.9: (左) FEV13 (右) COB



図 2.10: 多層読み出しのための CORE モジュール

749 第3章

750 ビームテストによる評価試験

751 これまでに作成された SiW-ECAL の技術プロトタイプ (FEV および COB) の性能評価実
752 験を、2023 年 6 月 7 日から 2023 年 6 月 22 日の期間に CERN SPS 加速器のビームラインに
753 て行った。本実験の主な目的は、電磁カロリメータとハドロンカロリメータの技術プロトタイ
754 プを同じビーム軸上に設置し、同時に運転を行いデータを取得すること、またこれまでの評価
755 実験の中でも最高エネルギーのハドロンビームを用いて 15 層の SiW-ECAL の評価を行うこ
756 との 2 点であった。また本実験におけるハドロンカロリメータは、同じく CALICE グループ
757 においてドイツやチェコが中心となって開発を進めている AHCAL を用いた。以下では実験
758 の詳細と、結果について述べる。

759 3.1 ビームライン

760 ビームテストは、フランスとスイスの国境付近に位置する CERN の SPS 加速器のビーム
761 ラインを用いて行った。SPS は、現在 LHC の前段加速器として利用されており、陽子シンク
762 ロトロンから来た 26 GeV の陽子を、周長 7km の加速器によって 400 ~ 450 GeV まで加速
763 している。SPS のビームラインでは、陽子ビームを固定ターゲットに入射することで、電子、
764 ミューオン/パイ中間子の 2 次ビームを運動量 10 ~ 400 GeV で得ることができる。本実験で
765 はこれらのビームを用いて、North エリアにある H2A ビームラインで実験を行った。実験に
766 用いたビームパラメータを表 3.1 に示す。

767

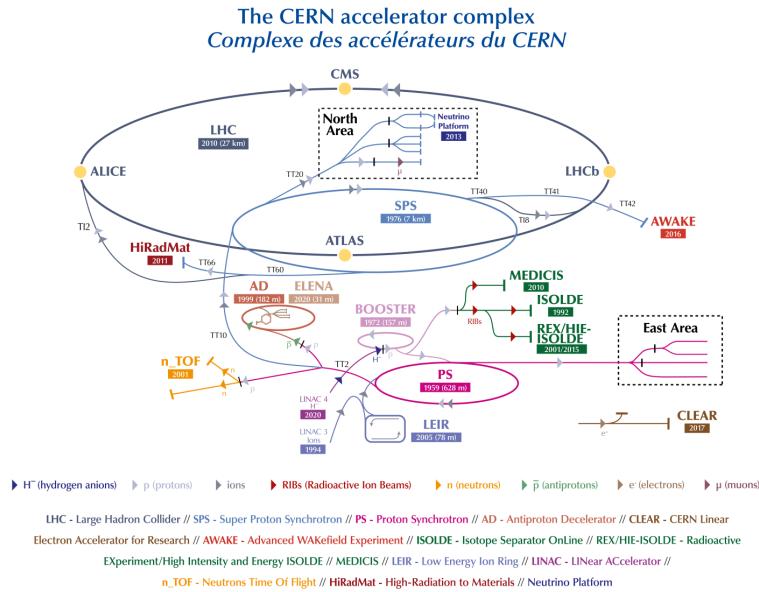


図 3.1: CERN の全体図

運動量	10-200[GeV/c]
電子の purity	10-99.5%
Max $\delta p/p$	2%
ビームの高さ	2460[mm]

表 3.1: H2A ビームラインにおけるビームパラメータ

3.2 実験セットアップ

3.2.1 測定機器のセットアップ

本実験でのセットアップの概観を図 3.2 に示す。図中右手からビームが照射され、ILD の構成と同様に上流側に SiW-ECAL が、下流側に AHCAL を設置した。SiW-ECAL のプロトタイプは、FEV が 13 層（うち FEV10 が 1 層、FEV11 が 3 層、FEV12 が 2 層、FEV13 が 7 層）COB が 2 層を組み合わせた計 15 層からなる検出層と鉛板 15 層の吸収層からなるモジュールを組み立てた。今回のビームテストの検出層は、最も良い性能が期待できる FEV13 を前方に設置し、性能比較のために COB を加えた構成となった。15 層の検出層において、センサーと読み出しボードに供給する電源は、15 層に対して図 3.3 左のように並列で印加した。また、信号はカプトンケーブルを通して 15 層分の信号を一括して CORE モジュールに送り、PC で読み出しを行った。また、AHCAL もサンプリング型ハドロンカロリメータであり、鉄の吸収層とシンチレータ検出層によって構成されていて、読み出しがシンチレータに統合され

たエレクトロニクスによって行われる。

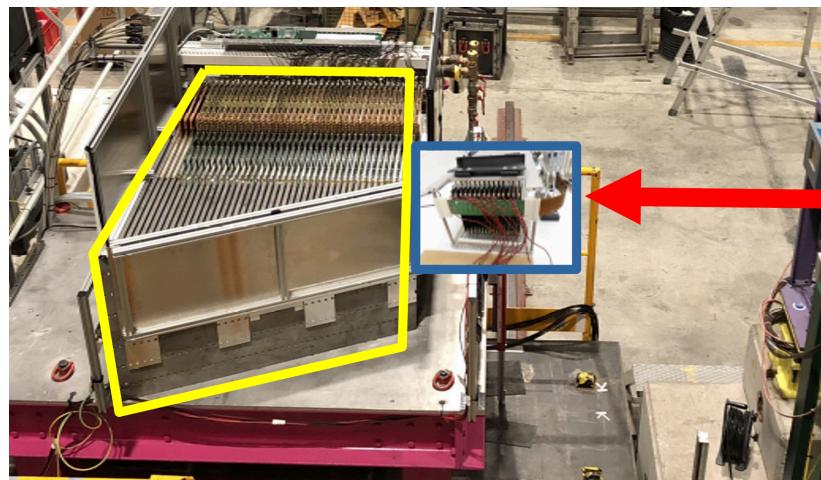


図 3.2: セットアップの全体図。黄枠 : AHCAL、青枠 : SiW-ECAL、赤 : ビーム位置

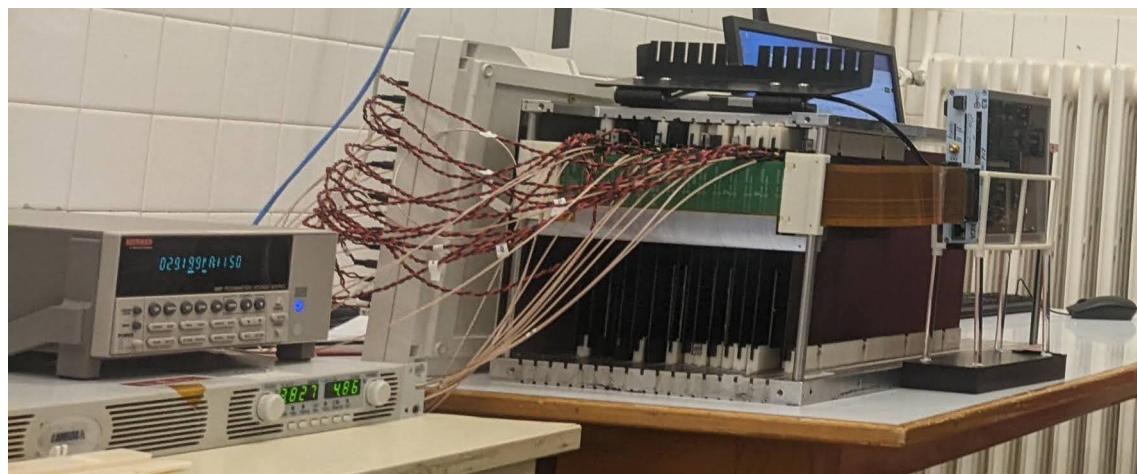


図 3.3: SiW-ECAL のプロトタイプモジュールと電源供給、読み出しシステム

層数	ボードの種類	ウェハー厚み [μm]	タンクスティン厚み [mm]
0	FEV13	650	4.2
1	FEV13	650	4.2
2	FEV13	650	4.2
3	FEV13	650	4.2
4	FEV13	500	4.2
5	FEV13	500	4.2
6	COB	500	4.2
7	FEV12	500	4.2
8	COB	500	5.6
9	FEV12	500	5.6
10	FEV11	320	5.6
11	FEV11	320	5.6
12	FEV10	320	5.6
13	FEV13	320	5.6
14	FEV11	320	5.6

表 3.2: SiW-ECAL のレイヤー構成 (0 層目がビーム上流側)

3.2.2 信号読み出し

ECAL と HCAL は独立に DAQ を行うが、本実験では EUDAQ という読み出しフレームワークによって HCAL と同期した読み出しを行った。EUDAQ では、ECAL と HCAL の間で Clock and Control Card (CCC) を同期させており、CCC ではクロック、スタート、ストップ信号を PC へ送っているため、PC 上で時間情報である Bunch Crossing ID (BCID) が同じイベントを収集することができる。

また、ECAL では専用の DAQ ソフトウェア [10] が稼働しており、DAQ のみでなくチャンネル単位での閾値の設定やイベントのモニター (図 3.7) が可能となっている。データは raw 形式で保存され、解析用のソフトウェアを通してツリー形式にデータを保存した root ファイルへ変換し解析に利用することができる。

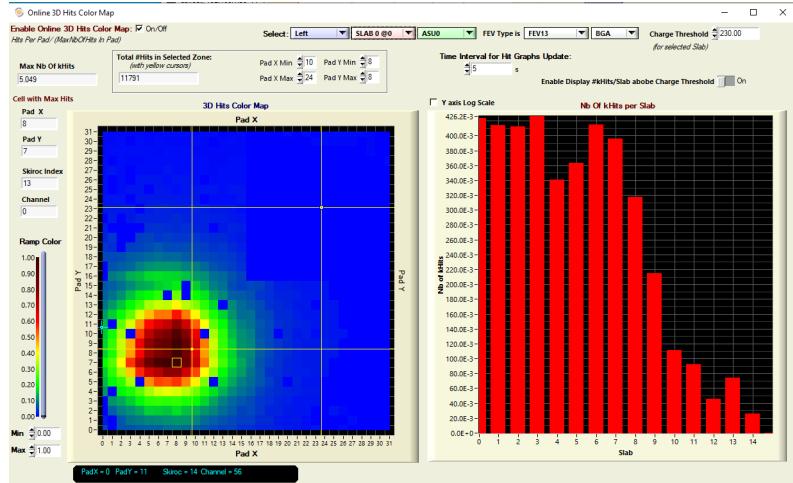


図 3.4: SiW-ECAL 読み出しソフトウェアにおけるイベントモニター (左) ヒットマップ (右) 各層あたりのヒット数

3.3 実験結果

3.3.1 検出器応答

各 run に対して層ごとのヒットマップを作成し、各層の応答を確認した。作成したヒットマップが図 4.3 である。ビームはヒットマップ左下のセンサーを中心に照射を行った。電子ビームは相互作用を起こし電磁シャワーを形成するため、ヒット領域が大きくなっていることが確認できる。またミューオンビームは、電子ビームと比較して相互作用を起こさず、ビームサイズが小さい。また白く抜けている箇所はノイズが多いためマスクしている、あるいは信号がないチャンネルであるが、ビームの種類によらず全体を通して 4 つほど四角く抜けている部分が確認できる。これは 1 枚のセンサーの場所と対応しており、実験後に確認したところセンサーと PCB の間の導電性接着剤が剥がれていたことが発覚した。さらにビームのエネルギーを上げていくと、80 GeV 以上のエネルギーで図 4.3 (b) 第 6 層のような、1 センサー全体にヒットが集中してしまう現象が確認された。エネルギーが大きくなった場合にはセンサーに入ってくる信号が多くなることから、センサー周辺での放電等が原因として考えられるが、現在調査を行なっている。

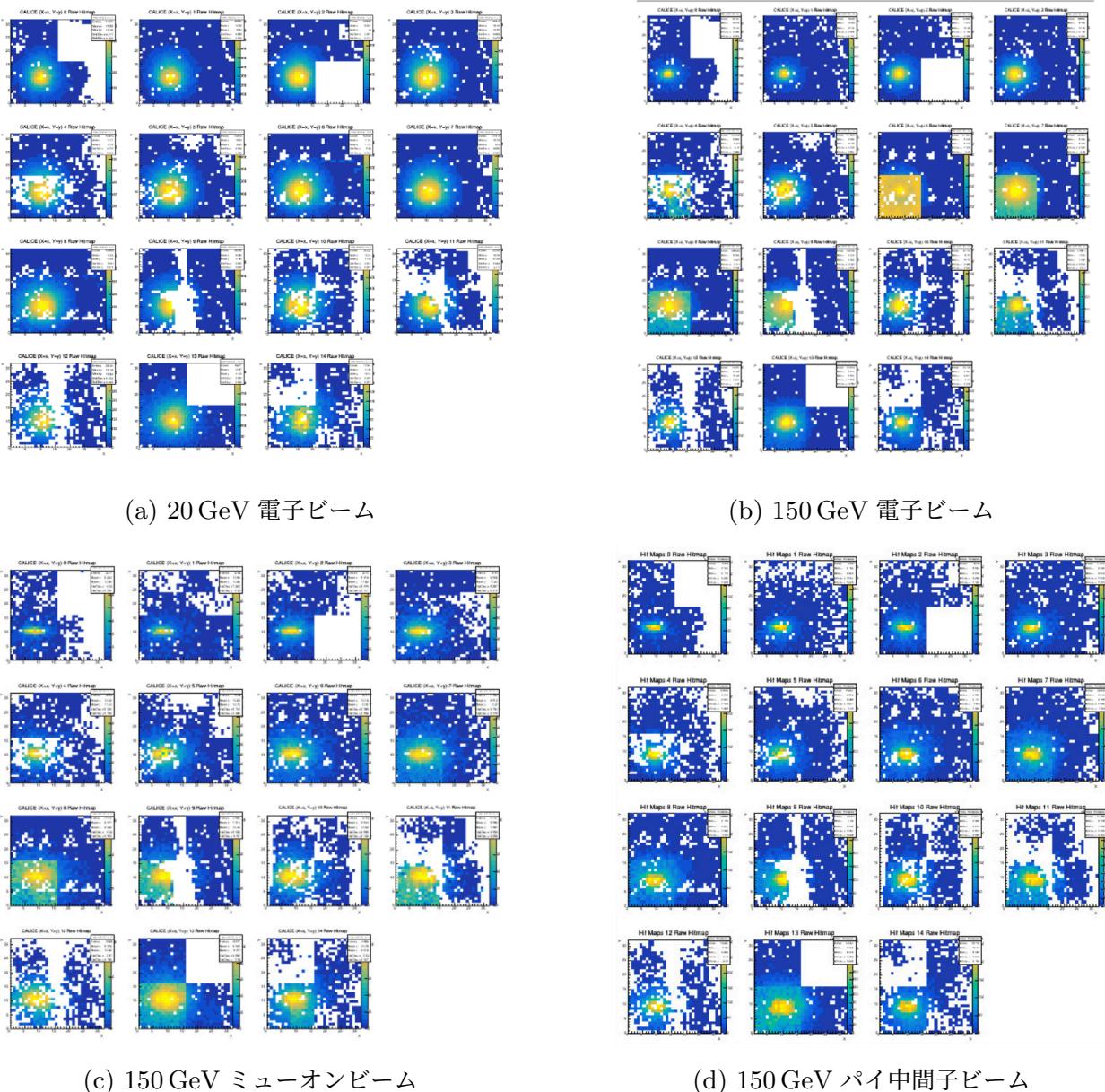


図 3.5: 各ビーム種類におけるヒットマップ。左上から 0 層右上が 3 層、右下が 14 層と順に並んでいる。

3.3.2 ペデスタル

続いて、ペデスタルの解析を行った。ペデスタルとはトリガーが入っていない時の信号の大きさで、エレクトロニクス由来のノイズ等によってその値の幅は変化する。実際のヒットによる信号の大きさは、得られたすべての値からペデスタルの値を引いた値になるため、ペデスタルについて十分に解析を行うことは重要となる。一般的にペデスタルはガウス関数でフィッティングすることができ、その中央値はペデスタル信号の大きさを、標準偏差はその揺らぎの大きさ(ノイズ)に相当する。図 3.6 に各エネルギーの電子ビームの run における同じチャン

812 ネルのペデスタルを示した。

813

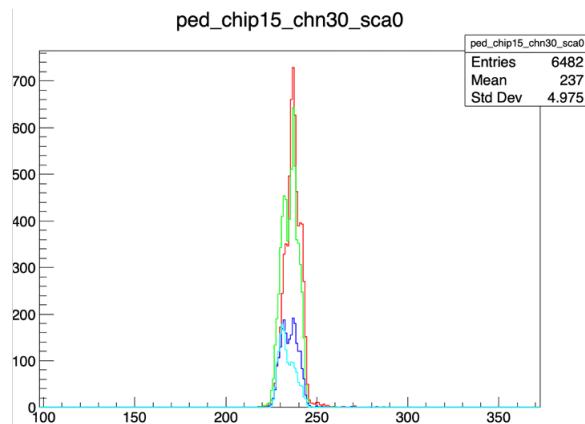


図 3.6: 40/60/100/150GeV の電子ビームにおけるペデスタル

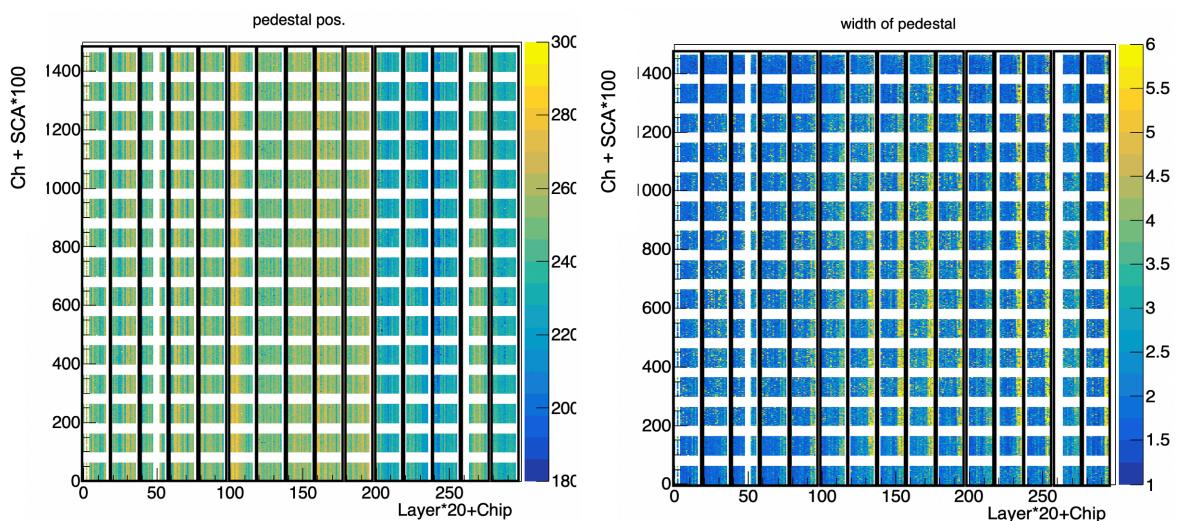


図 3.7: ペデスタルをガウス関数フィットした際のパラメータ。横軸がチップを、縦軸がチャネル番号を表しており、それぞれ全層、全 SCA での可視化のため、層数や SCA をかけた値となっている。(左) ガウス関数の中央値 (右) ガウス関数の幅の大きさ

814 また、チャネルによって図 3.8 に示すような 2 つのピークを持つペデスタルが確認された。このペデスタルピークが 2 つ見える現象をダブルペデスタルと呼んで調査を行なった。

815

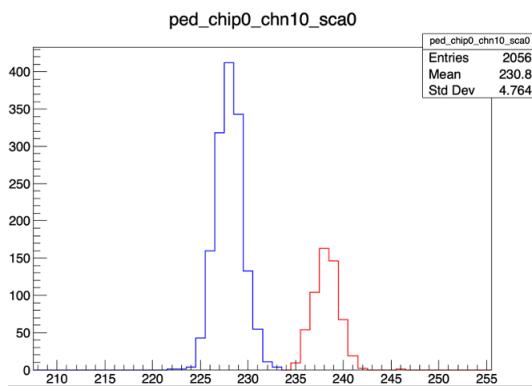


図 3.8: ダブルペデスタルの例

3.4 まとめと考察

記入予定（メモ：各 Slab は実験前にモジュールへのインストールと動作確認を行い、緩衝材で梱包して輸送した。）

819 第 4 章

820 深層学習

821 本章では、本研究で提案する手法である深層学習の理論を述べる。初めに、深層学習の基礎
822 技術であるパーセプトロンについて説明する。そしてパーセプトロンを多層にしたニューラル
823 ネットの構造と計算技術について説明する。最後に深層学習のネットワークについて、特にグ
824 ラフ構造のデータを扱うグラフニューラルネットワークについて紹介する。

825 4.1 ニューラルネットワーク

826 4.1.1 パーセプトロン（単層ニューラルネットワーク）

827 ニューラルネットワークの基礎となるパーセプトロンは、ローゼンブラットにより 1957 年
828 に考案された [11]。パーセプトロンの基本構造は、信号を入力として受け取り論理回路を通し
829 て出力信号を出すものである。図 4.1 に最も基本的なパーセプトロンの例 [12] を示す。

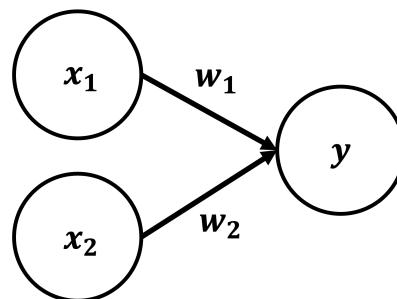


図 4.1: パーセプトロン

x_1, x_2 は入力信号、 y が output 信号であり、 w_1, w_2 がそれぞれの入力信号にかかる重みを表す。また、図中における ○ はノードと呼ぶ。入力信号はノードに送られる前に重みが掛けられ、出力ノードにてそれらの総和をとる。出力ノードでの演算（活性化関数）をステップ関数（階段関数）とすると、その総和が閾値 θ を超えている場合のみ出力信号は 1 を出力すること

になる。数式で示すと以下のようになる。

$$y = \begin{cases} 0 & (w_1x_1 + w_2x_2) \leq \theta \\ 1 & (w_1x_1 + w_2x_2) > \theta \end{cases} \quad (4.1)$$

830 パーセプトロンにおいて重要なのは入力信号に対する固有の重みであり、重みは各信号
 831 の重要性を操作する要素として働く。すなわち重みが大きいほど、対応する信号の全体における重要性が高くなる。この重みを更新する操作を学習と呼び、ニューラルネットワークでは学
 832 833 習を繰り返すことによって重みパラメータを理想とする値に近づけていく。

また、入力信号が 3 つ以上の場合についても考えることができ、以下のような式で表される。入力信号 $x = \{x_1, x_2, \dots, x_n\}$ 、重みパラメータ $w = \{w_1, w_2, \dots, w_n\}$ 、活性化関数（ここではステップ関数）を $h(x)$ とすると、出力ベクトル y は以下のようになる。

$$y = h(w^T x) = \begin{cases} 0 & (w_1x_1 + w_2x_2 + \dots + w_nx_n) \leq \theta \\ 1 & (w_1x_1 + w_2x_2 + \dots + w_nx_n) > \theta \end{cases} \quad (4.2)$$

834 4.1.2 多層パーセプトロン（多層ニューラルネットワーク）

835 パーセプトロンの演算では線形領域のみしか表現できず、非線形領域においても扱えるよう
 836 入力層と出力層の間に中間層（隠れ層）を加えるニューラルネットワークに改良された。この
 837 ような中間層を複数重ねたパーセプトロンを多層パーセプトロン（Multi Layer Perceptron,
 838 MLP）と呼ぶ。多層パーセプトロンの簡単な例を図 4.2 に示す。

839

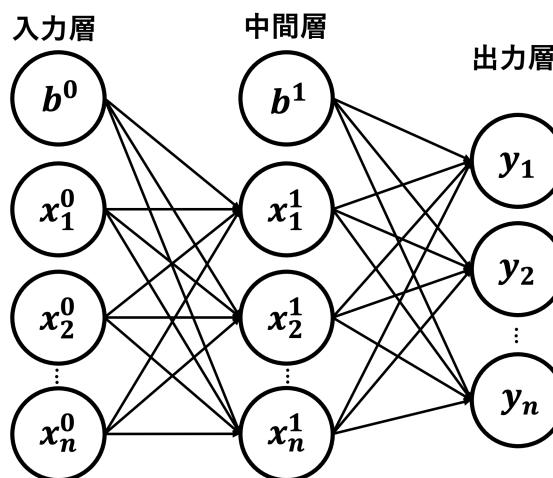


図 4.2: 多層パーセプトロン（ニューラルネットワーク）

最も左のノード列を入力層、真ん中のノード列を中間層、一番右のノード列を出力層とすると、以下のような数式で表される。入力信号 $x^0 = \{x_1^0, x_2^0, \dots, x_n^0\}$ 、中間層の各ノードに入ってくる信号 $x^1 = \{x_1^1, x_2^1, \dots, x_n^1\}$ 、入力層と中間層の信号にかかる重みパラメータがそれぞ

れ $w^0 = \{w_1^0, w_2^0, \dots, w_n^0\}$, $w^1 = \{w_1^1, w_2^1, \dots, w_n^1\}$ 、活性化関数を $h(x)$ とすると、出力ベクトル y_n は

$$x_1^1 = h(w_1^0 x_1^0 + w_2^0 x_2^0 + w_3^0 x_3^0 + \dots + w_n^0 x_n^0 + b^0) \quad (4.3)$$

$$y_n = h(w_1^1 x_1^1 + w_2^1 x_2^1 + w_3^1 x_3^1 + \dots + w_n^1 x_n^1 + b^1) \quad (4.4)$$

となる。ここで、 b^n としてより学習にパラメータを加えるため、各層に実数値のバイアスを導入した。重みと信号の積の和を a として、上式に行列を用いると簡略に表現できる。

$$\mathbf{y} = h(\mathbf{a}) \quad (4.5)$$

$$\mathbf{a} = \mathbf{W}\mathbf{x} + \mathbf{b} \quad (4.6)$$

以下ではニューラルネットワークの学習における、学習の仕組みや重要な技術について取り上げる。

活性化関数

活性化関数はニューラルネットワークにおける入力の重み線形和から、出力を決定するための関数である。活性化関数には、非線形演算によって表現力を高めるために非線形関数が用いられることが多く、以下に主なものについて示す。中でも ReLU 関数は、勾配の最大値が 1 であることから勾配消失を起こしにくく、計算の安定性の理由から多くのモデルにおいて用いられている。(sigmoid 関数の場合 $x = 0$ で最大となり、それ以外で急速に小さくなってしまう)

- ステップ（階段）関数

$$h(a) = \begin{cases} 0 & (a \leq \theta) \\ 1 & (a > \theta) \end{cases} \quad (4.7)$$

- sigmoid 関数

$$h(a) = \frac{1}{1 + \exp(-a)} \quad (4.8)$$

- tanh 関数

$$h(a) = \tanh(a) \quad (4.9)$$

- ReLU 関数（ランプ関数）

$$h(a) = \begin{cases} 0 & (a \leq \theta) \\ a & (a > \theta) \end{cases} \quad (4.10)$$

- LeakyReLU 関数 ($s=0.01$ が多い)

$$h(a) = \begin{cases} sa & (a \leq \theta) \\ a & (a > \theta) \end{cases} \quad (4.11)$$

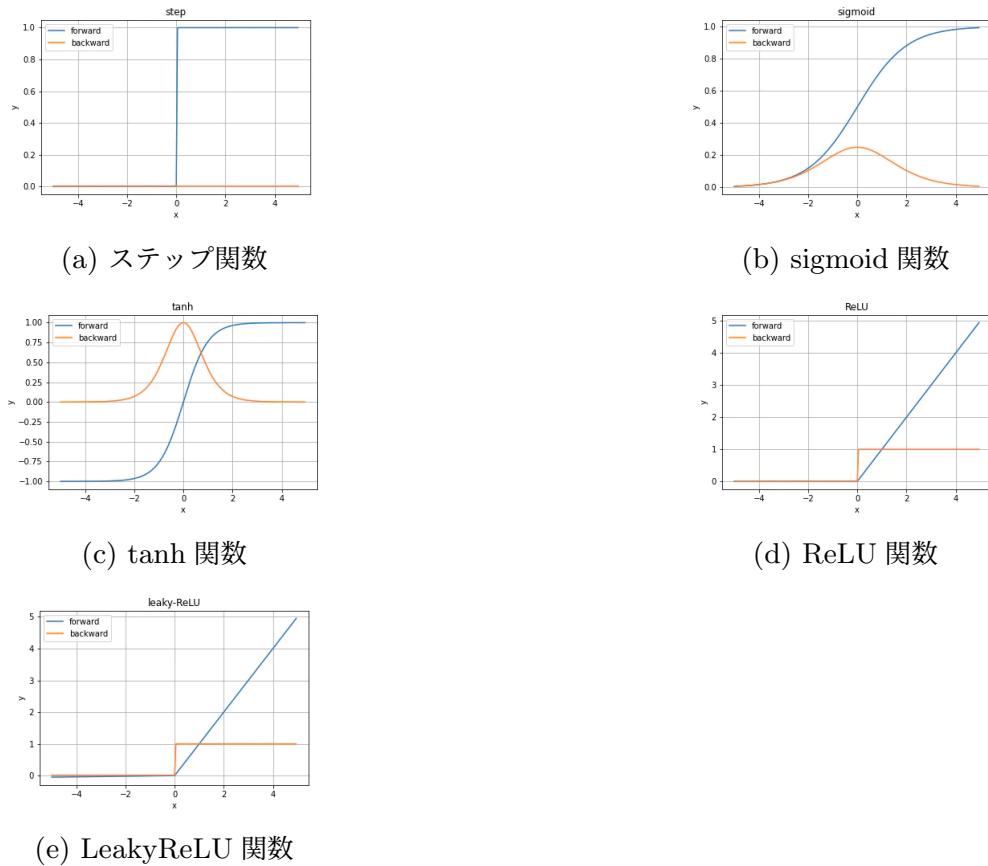


図 4.3: 活性化関数のグラフ。Forward は順方向、backward は逆方向（後に誤差逆伝播法で説明）の際の演算を表す。

850 出力層の設計

ニューラルネットワークで扱える問題は、主に回帰問題と分類問題に分けられる。それぞれの問題によって出力層の設計が異なり、回帰問題においては恒等関数が、分類問題においてはソフトマックス関数が用いられる。恒等関数では入力された値をそのまま出力する。ソフトマックス関数は以下の式 4.12 のように、0 から 1 までの値を出力する関数であり、それぞれのカテゴリに分類される確率を表す。ここで y_k はニューラルネットワークの出力、 x_k は出力層へ入ってくる信号を表す。また、出力層のノード数は問題に合わせて適宜調整する必要があり、分類問題であればカテゴリ数だけノードを設計する必要がある。

$$y_k = \frac{\exp(x_k)}{\sum_{i=1}^n \exp(x_i)} \quad (4.12)$$

851 損失関数

852 先述の通り、ニューラルネットワークでは学習によって重みを更新するが、その際に学習結果を正しい答えと照らし合わせて評価し、その評価値を最小にするように重みを更新する。この評価関数を損失関数 (loss function) と呼ぶ。損失関数には主に以下の 2 つが用いられる。

- 二乗和誤差関数 (Mean Squared Error)

二乗和誤差関数は、以下の式 4.13 で定義される関数である。ここで、 y_k はニューラルネットワークの出力、 t_k は正解ラベルを表し、 k はデータの次元数を表す。二乗和誤差関数の微分値は y の一次関数となっていることから、出力・正解ラベルが共に連続値であり恒等関数を出力層を持つ回帰問題で採用される。

$$L = \frac{1}{2} \sum_k (y_k - t_k)^2 \quad (4.13)$$

- 交差エントロピー誤差 (Cross Entropy Error)

交差エントロピー誤差は、以下の式 4.14 で定義される関数である。ここで、 y_k はニューラルネットワークの出力で、 t_k は 0 か 1 の one-hot 表現の正解ラベルを表す。交差エントロピー誤差は主にソフトマックス関数を出力層に用いる分類問題において採用される。

$$L = - \sum_k t_k \log(y_k) \quad (4.14)$$

855 **誤差逆伝播法**

856 これまでではニューラルネットワークの順方向の伝播 (forward propagation) について見て
857 きたが、出力層において学習結果と正解ラベルを比較し、逆方向に信号を伝播させ重みを更新
858 するアルゴリズムを誤差逆伝播法 (back propagation) という。誤差逆伝播法では、次のような
859 処理を行う。

- 860 1. ニューラルネットワークにおいて順方向に学習を行い、出力層で損失関数によって正解
861 ラベルとの誤差を求める。
862 2. 誤差から各出力層ノードについて期待される出力と重要度、誤差を計算する。(局所
863 誤差)
864 3. 特に重要度の高い前層の入力が、局所誤差に影響を及ぼしているとして重みを調整する。
865 4. 合成関数の微分によって、さらに前層へと処理を繰り返す。

866 これによって重みを修正していく学習が可能になっており、また勾配の計算には活性化関数の
867 偏微分が積をとって含まれているため、ネットワークの演算は全体を通して微分可能となっ
868 いる。

869 **前処理**

870 ニューラルネットワークでは、前処理を行うことで識別性能の向上や学習の高速化を見込む
871 ことができる。前処理の手法には以下にあげるようなものがあり、データ全体の分布を考慮し
872 て適応する必要がある。

- 正規化：最小値を 0、最大値を 1 とするスケーリング

$$x'_i = \frac{x_i - \min(x)}{\max(x) - \min(x)} \quad (4.15)$$

- 標準化：平均を 0、分散を 1 とするスケーリング

$$x'_i = \frac{x_i - \bar{x}}{\sigma} \quad (4.16)$$

- 対数変換：外れ値による分散を小さくし、0 付近の値を区別しやすくする。

$$x'_i = \ln(x_i) \quad (4.17)$$

873 ミニバッチ処理

874 ニューラルネットワークを学習させるにあたって、データを 1つ1つ学習させるわけではな
875 い。実際にはミニバッチと呼ばれる、学習データをいくつかまとめて束としたものを一度に学
876 習させる。また、このミニバッチのサイズ、つまりいくつのデータをまとめて束にするかとい
877 う値のことをバッチサイズという。数値計算を扱うライブラリの多くは、大きな配列の計算を
878 効率よく処理できるよう最適化がなされており、ミニバッチによる学習を行うことで、処理時
879 間を短縮することができる。一方で、ミニバッチのデータは誤差逆伝播においてバッチ内で損
880 失関数の和を用いているため、バッチサイズによって学習結果が異なり、サイズが大きい場合
881 には学習精度が悪くなってしまう可能性もある。

882 最適化アルゴリズム

883 ニューラルネットワークの学習では、損失関数の値が最小となるような最適なパラメータを
884 探索する。しかし損失関数のパラメータ空間は非常に複雑であることから、最適化は難しい。
885 以下では、勾配降下法をはじめとする最適化手法について述べる。また、一度の学習で更新す
886 るパラメータの度合いを学習率 (learning rate) と呼んでおり、ネットワークの重みなどのパ
887 ラメータとは異なり、学習率のような人の手で設定する必要のあるパラメータをハイパーパラ
888 メータと呼ぶ。

- 勾配降下法

現在のネットワークのパラメータの微分 (勾配) を計算し、その微分の値を手がかりにパラメータの値を徐々に更新する方法を、勾配降下法 (gradient descent method) という。勾配降下法は以下の式 4.18 のように表される。ここで、 \mathbf{W} は更新する重みパラメータを、 L は損失関数を、 η は学習率を表す。

$$\mathbf{W} \leftarrow \mathbf{W} - \eta \frac{\partial L}{\partial \mathbf{W}} \quad (4.18)$$

889 また、ミニバッチ学習を用いた勾配降下法は特に、確率的勾配降下法 (stochastic
890 gradient descent, SGD) と呼ばれており、現在のニューラルネットワークの最適化法

は主に SGD に基づいて設計されている。しかし、SDG には関数の形状が等方的でない場合、勾配の方向が最終的な最小値と異なるため探索が非効率になるという欠点があり、単純に勾配方向へ進む以外の方法としてさまざまな最適化手法が考案されている。

● モーメンタム

モーメンタム (Momentum) は、それまでの学習における損失関数上で更新ステップの動きを考慮することで SGD の振動を抑えるアルゴリズムである。モーメンタムにおける更新方法は、物理学の速度にあたる変数 v を加え、以下の式のように表される。

$$v \leftarrow \alpha v - \eta \frac{\partial L}{\partial W} \quad (4.19)$$

$$W \leftarrow W + v \quad (4.20)$$

上式における αv が、U の字の斜傾を転がるボールが徐々に減速する運動のような役割を果たし、振動を抑えている。

● AdaGrad

AdaGrad では、モーメンタムと同様に SGD の振動を抑えるが、学習率を減衰させることによってこれを達成するアルゴリズムである。AdaGrad の更新方法は次のような式で表される。

$$h \leftarrow h + \frac{\partial L}{\partial W} \odot \frac{\partial L}{\partial W} \quad (4.21)$$

$$W \leftarrow W - \eta \frac{1}{\sqrt{h}} \frac{\partial L}{\partial W} \quad (4.22)$$

ここで、 \odot は要素ごとの積を行うアダマール積を表し、 h はこれまでの勾配の値を二乗和として保持する役割を持つ。そして $\eta \frac{1}{\sqrt{h}}$ によって学習率のスケールを調整することができる。これによって動いた大きさに合わせてパラメータ毎に学習率の減衰を行うことができる。

● Adam

Adam はモーメンタムの考え方と AdaGrad の考え方を融合させた手法であり、モーメンタムの変数 2 つと前のステップまでの学習係数を表す変数 1 つの計 3 つをハイパープラメータにもつ。これによって効率的にパラメータ空間を探索することができる。

● RAdam

RAdam[13] は Adam に Warmup による改良を行った最適化手法である。これまで挙げた手法では、学習の初期段階ではサンプル数が少ないために、適応学習率の分散が極めて大きくなり粗悪な局所的最適解に陥ってしまうという課題があった。それに対して Warmup 手法によって学習初段階の学習率を下げ、学習全体の学習率の分散を抑えるようなパラメータ空間の探索が可能となっている。

910 過学習

911 深層学習には、ネットワークモデルが学習用データに過度に適合し過ぎてしまい、新しい
912 データに対する性能が低下してしまう、過学習 (Overfitting) という状態が発生する場合があ
913 る。これはデータがパラメータを大量に持ち、表現力が極めて高いモデルである場合や、学習
914 用データの数が少ない場合に多く見られ、モデルがトレーニングデータに含まれるノイズ、ま
915 たは特異的な特徴に過度に適合するために起きてしまうものとされている。そのため、深層学
916 習の実装においては過学習を防止するいくつかの方法を用いることが多く、以下に代表的なも
917 のを挙げる。

918 ● ドロップアウト (Dropout)

919 ニューラルネットワークでは全てのノード同士が繋がっており演算を行っていたが、
920 ニューロンをランダムに消去することで表現力が高すぎることによる過学習を抑制する
921 というアプローチがあり、ドロップアウトという。これは機械学習におけるアンサンブル
922 学習に近く、各エポック（学習の回数）の学習でそれぞれ違うモデルを学習させてい
923 ると解釈することができる。

922 ● 荷重減衰 (Weight Decay)

923 適合しすぎる学習では、重みパラメータが極端に大きい値をとってしまっている場合が
924 多く存在する。そのため重みに制限をかけることで抑制する荷重減衰 (Weight Decay)
925 という手法が存在する。具体例として、重みの二乗ノルム (L2) を損失関数 L に加算す
926 る場合には、以下のような制限がかかる。

$$L = E + \frac{1}{2} \lambda \sum_k (w_k)^2 \quad (4.23)$$

927 ここで、 E は通常の誤差関数（損失関数）、 λ は正則化の強さを表す正則化パラメータ、
928 w が重みパラメータを表す。また L2 正則化の他にも、重みの絶対値の和をとる L1 正
929 則化や絶対値最大の成分の絶対値に λ をかける L_∞ などが存在する。これらによって
930 重みパラメータに強いペナルティが課せられた上で損失関数の値を最小にする学習を行
931 うことができる。

930 ● 重み初期化

931 荷重減衰と同様に重みを大きくしないための手法として、重みの初期値を定める手法も
932 存在する。重みは小さければ良いというものではなく、0 の場合には誤差逆伝播法にお
933 いて全ての重みの値が同じように更新されてしまうため、正しくは重みを対称的な構造
934 を持たないものにすることが重要となる。その手法にはランダムな重みを振る手法など
935 様々なものがあり、中でも Xavier Glorot による手法 [14] が一般的に用いられる。こ
936 の手法では、ニューラルネットにおいて入力層と出力層の重みが、ノード数を n とした
937 時 $\frac{1}{\sqrt{n}}$ の標準偏差を持つガウス関数になるという初期化を行う。

936 ● 正規化 (Batch Normalization)

ミニバッチ毎の入力特徴量をスケール 1、平均 0 の分布に正規化を行う手法が存在する。これによって学習の高速化と各層で活性化関数の値が正規化されることで、活性化関数の値のスケールが統一され、勾配がほぼ 0 になってしまい学習が上手くいかなくなる勾配消失問題を抑止することができる。

4.1.3 ディープニューラルネットワーク

多層ニューラルネットワークにおいて、特に図 4.4 のように層の数を多数持つモデルに対しては、構造が深いことからディープニューラルネットワーク (Deep Neural Network; DNN; 深層学習) と呼ぶ。ディープニューラルネットワークの演算には過去に勾配消失のような技術的課題が存在していたが、計算機性能の向上に加え以下に挙げる計算技術の工夫などによって学習が可能となった。

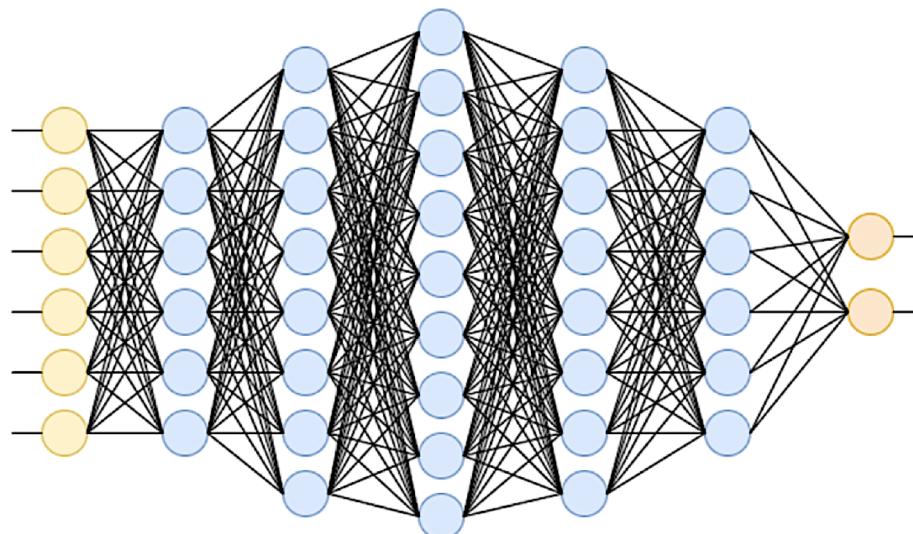


図 4.4: ディープニューラルネットワーク

4.2 グラフニューラルネットワーク

深層学習では数値データのみでなく、画像認識や自然言語処理など様々なデータに対して目覚ましい成果を挙げている。その中で特に近年、グラフで表される構造データに対する研究が非常に盛んになっており、本研究において取り上げるグラフニューラルネットワーク (Graph Neural Network; GNN) [15] もその一つである。グラフ構造データとは、オブジェクトの集合（ノード）を関係（エッジ）で結んだデータのことと、ノードのみが特徴量を持つ場合と、ノードとエッジの両方が特徴量を持つ場合がある。代表的には友人関係や論文の引用関係、化合物などをグラフデータとして構築することは有用であるとされており、データ間の相互関係を用いたモデリング・学習が可能であるため、高い表現力を持つことを強みとしている。

957 グラフの種類は大きく分けてノードの次元が同じ同種グラフと、異なる次元のデータを扱う異
 958 種グラフの 2 種類に分けられる。更に、エッジが方向性を持っている有向グラフと無向グラ
 959 フ、データが時系列で変化する動的グラフと変化しない静的グラフに分けられる。そしてこの
 960 ようなグラフデータを扱うニューラルネットをグラフニューラルネットワークという。グラフ
 961 構造や演算技術、目的とする課題によって GNN のネットワークモデルの種類は多岐にわたっ
 962 ており、以下では基本的なグラフでの演算や本論文に関連した幾つかの種類のモデルを取り上
 963 げる。

964

965 4.2.1 メッセージパッシング

GNN のアーキテクチャにおいて、グラフは集合 $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ で構成されており、 \mathcal{V} はノードの集合、 \mathcal{E} はエッジの集合を表す。ニューラルネットワークの場合と同様にノードは特徴量を持っており、エッジも特徴量を持つことができる。図 4.5 に GNN における一般的な演算処理の流れ（メッセージパッシング）を示す。各ノードを \mathbf{h}_i とすると、隣接しているノードおよびその間のエッジの特徴量を集約し、ノードを更新 \mathbf{h}'_1 する。これによってエッジによって関連づけられたノード間で特徴量を更新していくことができる。エッジが特徴量を持つ場合は、エッジについても更新を行う。これらを数式にまとめると次のようになる。

$$\mathbf{h}_{(i,j)} = f_{\text{edge}}(\mathbf{h}_i, \mathbf{h}_j, x_{(i,j)}) \quad (4.24)$$

$$\mathbf{h}'_i = f_{\text{node}}(\mathbf{h}_i, \sum_{j \in \mathcal{N}_i} \mathbf{h}_{(j,i)}, x_i) \quad (4.25)$$

966 ここで、 $\mathbf{h}_{(i,j)}$ はエッジを、 x_i はノードの特徴量を、 \mathcal{N}_i は隣接しているノードの集合を表す。

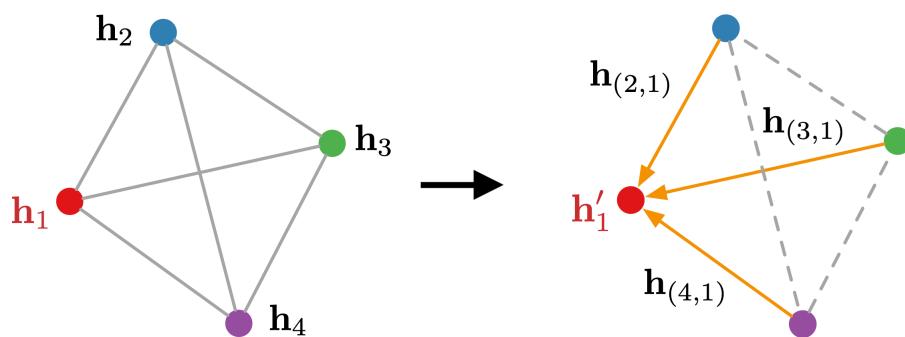


図 4.5: (左) 4 つのノードを持つ全結合グラフニューラルネットワーク。 h_i はノード表現を表す。 (右) メッセージパッシングの処理。

4.2.2 Graph Convolution Network (GCN)

GCN とは、メッセージパッシングにおいて畳み込み（Convolution）を用いる手法である。一般的に機械学習における畳み込みは、あるフィルターを用いて対象と掛け合わせたものの和をとることで、周辺の情報を含ませることのできる処理であり、画像処理などにおいて広く用いられている。グラフの畳み込みにおいてはスペクトルによるアプローチと空間的なアプローチの 2 つがあり、以下ではそれについて説明する。

Spectral Graph Convolution

スペクトルによる畳み込み [16] は信号処理の考えに基づいたアプローチである。音声などの信号処理においては、関数をフーリエ変換によって周波数成分に変換し、ノイズ除去をして逆変換する。これをグラフデータに置き換えると、グラフラプラシアンの固有ベクトルが張る空間への変換・逆変換となる。グラフ信号を \mathbf{x} とすると、グラフフーリエ変換 $\mathcal{F}(\mathbf{x})$ 、逆グラフフーリエ変換 $\mathcal{F}^{-1}(\mathbf{x})$ は次のように表される。

$$\mathcal{F}(\mathbf{x}) = \mathbf{U}^T \mathbf{x} \quad (4.26)$$

$$\mathcal{F}^{-1}(\mathbf{x}) = \mathbf{U} \mathbf{x} \quad (4.27)$$

ここで、 \mathbf{U} は正規化グラフラプラシアン \mathbf{L}' の固有ベクトル行列を表す。グラフラプラシアン \mathbf{L} は、グラフの隣接を表す隣接行列 \mathbf{A} とグラフの各ノードに接続したノード数を対角成分にもつ次数行列 \mathbf{D} を用いて、 $\mathbf{L} = \mathbf{D} - \mathbf{A}$ で求められる実対称正方行列であり、正規化グラフラプラシアンは $\mathbf{L}' = \mathbf{I} - \mathbf{D}^{\frac{1}{2}} \mathbf{A} \mathbf{D}^{\frac{1}{2}}$ とかける。

これらを用いて、グラフにおける Spectral な畳み込み演算は次のように定義される。

$$\mathbf{g} \star \mathbf{x} = \mathcal{F}^{-1}(\mathcal{F}(\mathbf{g}) \odot \mathcal{F}(\mathbf{x})) \quad (4.28)$$

$$= \mathbf{U}(\mathbf{U}^T \mathbf{g} \odot \mathbf{U}^T \mathbf{x}) \quad (4.29)$$

$\mathbf{g} \star \mathbf{x}$ は畳み込み演算を、 $\mathbf{U}^T \mathbf{g}$ はスペクトルにおける畳み込みのフィルターを表す。学習により更新される集合である \mathbf{g} に焦点を当ててより単純化すると、畳み込みは次のように書くことができる。

$$\mathbf{g} \star \mathbf{x} = \mathbf{U} \mathbf{g} \mathbf{U}^T \mathbf{x} \quad (4.30)$$

上記のように、スペクトルによる畳み込みは一度の更新の中でグラフ構造の一部が全体に影響を与えるような演算であるという特徴を持っている。さらに、行列の次元が固定されてしまうことから異なる構造のグラフ間でパラメータを共有できないことや、行列の固有値分解など計算量が多くなってしまうという問題点がある。

978 **Spatial Graph Convolution**

グラフデータにおける畠み込みは空間的なメッセージパッシングの考えに基づいており、グラフ内の1つのノードが持っている特徴量に、隣接関係にあるノードの特徴量に重みをかけたものを加えていく。これによりノード自体の特徴量に加え、隣接関係や隣接ノードの特徴量の情報などを含んだ演算を行うことができる。メッセージパッシングにおける工夫に応じて様々な畠み込みのモデルが提案されているが、以下では最も一般的なモデルである GraphSAGE (Graph SAmples and aggreGatE) [17] について述べる。

GraphSAGE では、隣接から特徴量をサンプリングして集約することで自身のノードを更新する。GraphSAGE における畠み込み演算は式 4.25 と同様に、次の式のように表される。

$$\mathbf{h}_{\mathcal{N}(v)}^k \leftarrow \text{AGGREGATE}_k(\{\mathbf{h}_u^{k-1}, \forall u \in \mathcal{N}(v)\}) \quad (4.31)$$

$$\mathbf{h}_v^k \leftarrow \sigma(\mathbf{W}^k \cdot \text{CONCAT}(\mathbf{h}_v^{k-1}, \mathbf{h}_{\mathcal{N}(v)}^k)) \quad (4.32)$$

$$(\forall k \in \{1, \dots, K\}, \forall v \in \mathcal{V})$$

979 ここで、 K は集約における隣接の深さを、 \mathbf{W}^k は重みパラメータ行列を、 σ は非線形関数
980 を表しており、1 式目 AGGREGATE で集約を、2 式目 CONCAT で更新を行なっている。
981 GraphSAGE はあらかじめサンプリングによって、隣接するノード数が異なる場合やグラフ
982 構造が変わった場合にも畠み込み演算を行うことができ、さらに大規模なグラフ演算に対して
983 は計算コストを改善することができる。

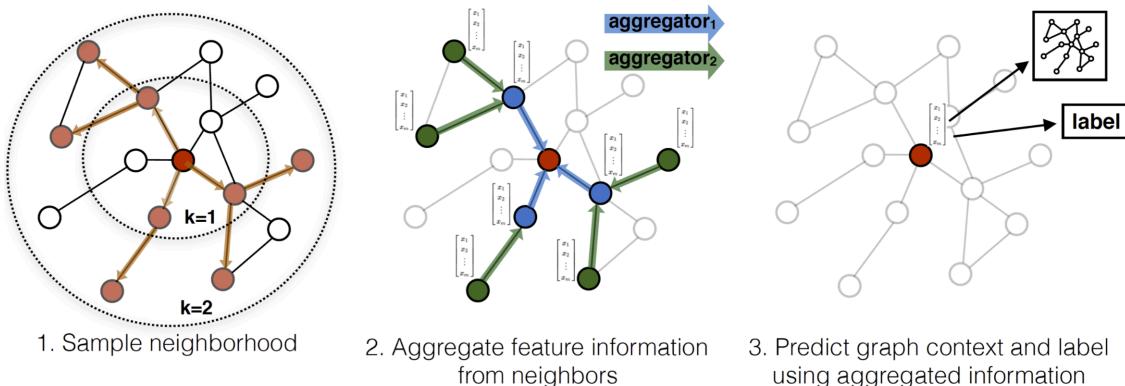


図 4.6: GraphSAGE における処理 (1. サンプリング, 2. 隣接からの集約, 3. 学習結果による推論)

984 **4.2.3 Graph Attention Network (GAT)**

深層学習では Attention という、データの中でも特に有益な場所に重み付けを行う手法があり、自然言語処理の分野などにおいてよく用いられている。Attention 機構では各ノード

ドに対して重要度を表す重みを導入し、それらを掛けた和をとることで必要な場所に注目（attention）を向けることができる。Attention をグラフ学習に適用したものが、Graph Attention Networks (GAT) [18] である。GAT では、重みパラメータ \mathbf{W} に加えて隣接したノードの重要度を表す attention 係数 α_{ij} を導入して、ノードの特徴量の更新を以下のように行う。

$$\mathbf{h} = \sigma \left(\sum_{j \in \mathcal{N}_i} \alpha_{ij} \mathbf{W} \mathbf{h}_i \right) \quad (4.33)$$

ここで、 α_{ij} は attention 処理 \mathbf{a} を用いて、次のように表される。

$$\alpha_{ij} = \mathbf{a}(\mathbf{W} \mathbf{h}_i, \mathbf{W} \mathbf{h}_j) \quad (4.34)$$

$$= \frac{e^{\text{LeakyReLU}(\mathbf{a}^T [\mathbf{W} \mathbf{h}_i \| \mathbf{W} \mathbf{h}_j])}}{\sum_{\mathcal{N}_i} e^{\text{LeakyReLU}(\mathbf{a}^T [\mathbf{W} \mathbf{h}_i \| \mathbf{W} \mathbf{h}_j])}} \quad (4.35)$$

985 GAT では attention 処理に 1 層のニューラルネットワークを用いており、 \mathbf{a} として学習の
 986 重みベクトルを用いている。式 4.35においては全ノード間で正規化して確率値を出力するた
 987 めの softmax 関数の適用と、活性化関数 (Leaky ReLU) の適用を行っている。（ $\|$ はテンソル
 988 の concatenate を表す。）GAT は、隣接に任意の重みを割り当てることから次数の異なるノー
 989 ドにも適応可能であり、未知のグラフ構造にも一般化することができる。

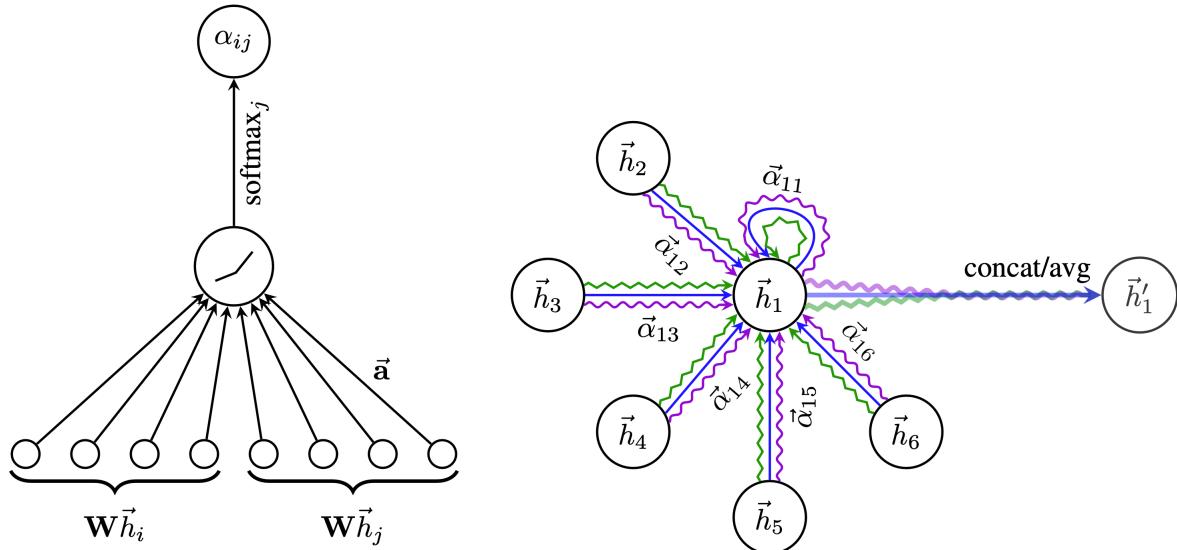


図 4.7: (左) 重みベクトル \mathbf{a} を用いた Attention 処理。 (右) 1 つのノード \mathbf{h}_1 に対する隣接ノード $\mathbf{h}_{\neq 1}$ の Attention と、ノード特徴量の更新 \mathbf{h}'_1

990 第 5 章

991 深層学習を用いたジェットフレー
992 バー識別

993 本章では、深層学習を用いて開発したジェットフレーバー識別アルゴリズムについて述べ
994 る。まず 5.1 節では、フレーバー識別に関する事象の詳細についてや、現在ジェットフレー
995 バー識別に用いられている LCFIPlus について述べる。次に 3.2 節で、今回のアルゴリズムの
996 評価において使用した MC シミュレーションデータについて述べる。またシミュレーション
997 データには深層学習における精度を上げるために前処理を行なっており、これについても述
998 べる。そして 5.3 節、5.4 節それぞれでディープニューラルネットワーク、グラフニューラル
999 ネットワークによる実装について述べ、5.5 節で LCFIPlus と比較を行って学習の性能につ
1000 いて議論する。

1001 5.1 イベントサンプルと先行研究

1002 本研究では、MC イベントジェネレーターである whizard による事象生成を行い、ILD 検
1003 出器のフルシミュレーションを使用したデータを用いてフレーバー識別を行った。シミュレー
1004 ションデータについての詳細を表 5.1 に示す。

イベントジェネレーター	whizard
シミュレーション	ILD Full simulation
検出器モデル	ILD_15_250GeV_v02-02
反応過程	$e^+ e^- \rightarrow \nu \bar{\nu} h \rightarrow \nu \bar{\nu} b\bar{b}/c\bar{c}/q\bar{q}$ ($q = u, d, s$)
重心系エネルギー	250 GeV
Initial State Radiation	あり
ビーム偏極	なし

表 5.1: シミュレーションデータのパラメータ

1005 **5.1.1 LCFIPlus におけるフレーバー識別**

1006 LCFIPlus のフレームワークでは、フレーバー識別は崩壊点検出、ジェットクラスタリング
1007 の後に行われる。フレーバー識別は、飛跡や崩壊点の特徴量をもとに ROOT の TMVA パッ
1008 ケージにおける Boosted Decision Trees (BDTs) を用いて識別が行われる。この Boosted
1009 Decision Trees は、複数のモデルを組み合わせて全体的な性能を向上させるアンサンブル学習
1010 (Boosting) と、入力された特徴量に基づいてデータをクラスに分割していく決定木 (Decision
1011 Tree) の手法を組み合わせたものであり、複数の弱い決定木モデルを組み合わせて決定木の精
1012 度を向上させる機械学習手法である。データは、初めに崩壊点の数で場合分けが行われ、次
1013 ページに示す、表 5.2 にある入力変数をもとに、最終的に各ジェットを b/c/uds の 3 つに分類
1014 する学習を行う。

変数名	説明
trk1d0sig	d_0 の significance が最も高い飛跡の d_0 (d_0 : xy 平面射影における IP と飛跡の距離, Appendix を参照)
trk2d0sig	d_0 の significance が 2 番目に高い飛跡の d_0
trk1z0sig	z_0 の significance が最も高い飛跡の d_0 (z_0 : sz 平面射影における IP と飛跡の距離, Appendix を参照)
trk2z0sig	z_0 の significance が 2 番目に高い飛跡の d_0
trk1pt	d_0 が最も大きい飛跡の横方向運動量
trk2pt	d_0 が 2 番目に大きい飛跡の横方向運動量
jprobr	全飛跡を用いた $r\text{-}\phi$ 平面での結合確率
jprobr5sigma	5σ 以上のパラメータを持つ全飛跡を用いた $r\text{-}\phi$ 平面での結合確率
jprobz	全軌跡を用いた z 軸射影での結合確率
jprobz5sigma	5σ 以上のパラメータを持つ全飛跡を用いた z 軸射影での結合確率
d0bprob	全飛跡に対して b,c,uds フレーバーの d_0 分布を用いた d_0 の b-クォーク確率の積
d0cprob	全飛跡に対して b,c,uds フレーバーの d_0 分布を用いた d_0 の c-クォーク確率の積
d0qprob	全飛跡に対して b,c,uds フレーバーの d_0 分布を用いた d_0 の uds-クォーク確率の積
z0bprob	全飛跡に対して b,c,uds フレーバーの z_0 分布を用いた d_0 の b-クォーク確率の積
z0cprob	全飛跡に対して b,c,uds フレーバーの z_0 分布を用いた d_0 の c-クォーク確率の積
z0qprob	全飛跡に対して b,c,uds フレーバーの z_0 分布を用いた d_0 の uds-クォーク確率の積
nmuon	ミューオンの数
nelectron	電子の数
trkmass	d_0/z_0 が 5σ を超える全飛跡の質量
vtxprob	崩壊点に関連した全飛跡を結合した時の崩壊点確率
vtxlen1	ジェット内の 2 番目の崩壊点の崩壊長
vtxlen2	ジェット内の 3 番目の崩壊点の崩壊長
vtxlen12	ジェット内の 2 番目の崩壊点と 3 番目の崩壊点の距離
vtxsig1	vtxlen1 の significance
vtxsig2	vtxlen2 の significance
vtxsig12	vtxlen12 を 2 番目の崩壊点と 3 番目の崩壊点の共分散行列の和の誤差で割った値
vtxdirang1	運動量と 2 番目の崩壊点の変位との間の開き角
vtxdirang2	運動量と 3 番目の崩壊点の変位との間の開き角
vtxmult1	2 番目の崩壊点に含まれる飛跡数
vtxmult2	3 番目の崩壊点に含まれる飛跡数
vtxmult	secondary vertex を構成するために用いられる飛跡の数
vtxmom1	2 番目の頂点に結合された全飛跡の運動量のベクトル和
vtxmom2	3 番目の頂点に結合された全飛跡の運動量のベクトル和
vtxmass1	飛跡の四元運動量の和から計算される 2 番目の崩壊点質量
vtxmass2	飛跡の四元運動量の和から計算される 3 番目の崩壊点質量
vtxmass	secondary vertex を構成する全 飛跡の四元運動量の和から計算される崩壊点質量
vtxmasspc	primary/secondary vertex の誤差行列で許容可能な pt 補正を行なった崩壊点質量
vtxprob	崩壊点が形成される確率

表 5.2: LCFIPlus におけるフレーバー識別の入力変数

1015 5.2 ディープニューラルネットワークによる実装

1016 本節では、ディープニューラルネットワーク（多層ニューラルネットワーク）によるフレー
1017 バー識別の実装について述べる。

1018 5.2.1 実装目的

1019 フレーバー識別では、シグナル効率に対して排除できる背景事象の割合が多くなるほど、
1020 ヒッグスなどの物理解析において有利であるため、その性能が非常に重要である。LCFIPlus
1021 において実装されている BDTs は、崩壊点やジェットの特徴量の場合分けによって分類を行
1022 うため、ツリー構造が視覚化されており解釈が容易である。一方で、BDTs の場合分けにおけ
1023 る閾値は人為的に決める必要があり、ILC のフレーバー識別のようなデータが高次元で複雑な
1024 タスクに対しては、より複雑な表現学習の方が適している可能性がある。さらに BDTs には
1025 ノイズに敏感である点や、ツリー数が多くなると過学習を起こしやすくなる点など問題点もあ
1026 る。そこで本研究では、より最適化したモデルの獲得によるフレーバー識別の更なる性能向上
1027 を目的に、ディープニューラルネットワークによる実装を行った。

1028 5.2.2 入力変数とネットワークアーキテクチャ

1029 学習には ILD フルシミュレーションにおける $e + e^- \rightarrow \nu\nu h$ 事象を 200 万イベント用い
1030 た。そして入力変数には、LCFIPlus において用いた変数（表 5.2）と同様の変数を使用した。
1031 また、学習値に対しては前処理を行い、0 付近の値が多い変数や値幅が大きい変数が存在した
1032 ため、各変数に対して対数変換や正規化を実行した。また出力も LCFIPlus と同様に b フレー
1033 バー、c フレーバー、uds フレーバーとした。

1034 ネットワークには図 5.1 のように、レイヤー間で全てのノード同士が結合した全結合層を用い
1035 たニューラルネットワークを採用した。さらに、過学習を抑制するためにバッチ正規化を各全
1036 結合層の後に行い、活性化関数には勾配消失への対策として全て LeakyReLU 関数を使用し
1037 た。損失関数には分類問題に適した交差エントロピーを用い、最適化アルゴリズムには Adam
1038 を学習率 0.01 で使用した。また学習率は 25 エポックごとに 0.1 倍させ、学習の安定化を目的
1039 に学習率を減衰させた。表 5.3 に最適化によって求めた学習におけるハイパーパラメータをま
1040 とめる。また、上記のネットワークは深層学習ライブラリである PyTorch を用いて実装を行
1041 なった。PyTorch は Python で使用可能なライブラリであり、高速な数値計算を行うための
1042 GPU サポートや、動的な計算グラフを導入しているため実行時コンパイル（Define-by-Run）
1043 という特徴を持っている。



図 5.1: フレーバー識別のためのディープニューラルネットワークの概略図

ノード数	(124, 124, 124)
活性化関数	LeakyReLU 関数 (0.001)
損失関数	交差エントロピー
最適化アルゴリズム	Adam
学習率	0.01 (25 エポックあたり 0.1 倍)
エポック数 (全データでの学習回数)	100
バッチサイズ	1024

表 5.3: ディープニューラルネットワークにおけるハイパーパラメータ

5.2.3 ハイパーパラメータの最適化

1044 学習におけるハイパーパラメータ (表 5.3) の一部にはチューニングを行った。ハイパーパ
 1045 ラメータのチューニングにおいて最適化する対象はブラックボックスであることが多く、ベイ
 1046 ズ最適化 [19] のような目的関数を推定しつつ大域的最適解を探索する手法がよく用いられる。
 1047 ベイズ最適化法は、ガウス過程回帰によって未知の関数をデータから学習し、可能な限り少な
 1048 い試行数で最適解の推定を行う。ガウス過程回帰とは、入力変数 x と推定精度を表す y との関
 1049 係性関数 $y = f(x)$ を推定する手法であり、入力の集合に対して出力の集合が多変量ガウス分
 1050 佈である。

1051 布に従うとして推定の幅を算出し、次の候補点を決定している。本アルゴリズムのハイパーパ
 1052 ラメータのチューニングでは、Preferred Networks 社のフレームワーク Optuna[20] を使用し
 1053 た。
 1054 チューニングを行なったパラメータとその値の範囲を表 5.4 に示す。パラメータの最適化試行
 1055 回数は 10 回で、図 5.2 に評価関数の値の経過を示す。今回の最適化では損失関数を評価関数
 1056 に用い、値が最小になるような最適化を行った。図 5.3 は 2 つのパラメータに対する評価関数
 1057 の 2 次元ヒストограмであり、中間層のノード数が 250 程度で極大値を取っている。また、図
 1058 5.4 に各パラメータの評価関数に与える重要度を示す。

変数名	最適化範囲
中間層のノード数	32 ~ 1024
ドロップアウト	0.1 ~ 0.9
学習率	0.00001 ~ 0.1
weight decay	0.000001 ~ 0.1

表 5.4: 最適化を行ったハイパーパラメータとその値の範囲。

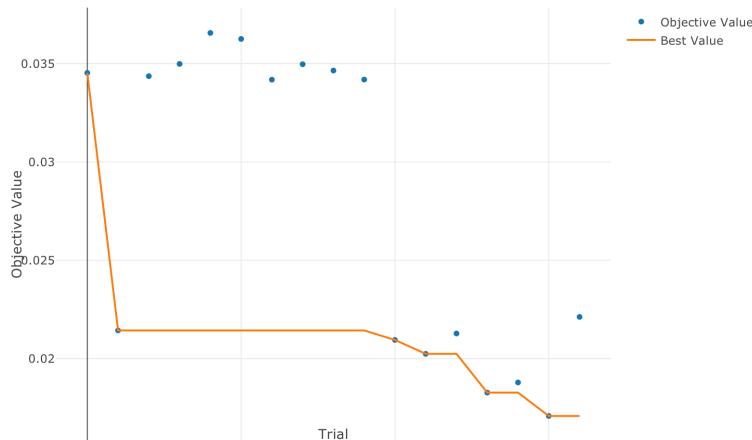


図 5.2: 各試行における評価関数の値。値が最小になるようなアルゴリズムを実行している。

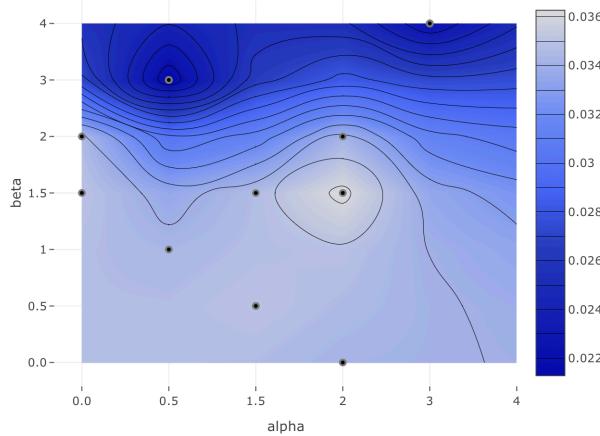
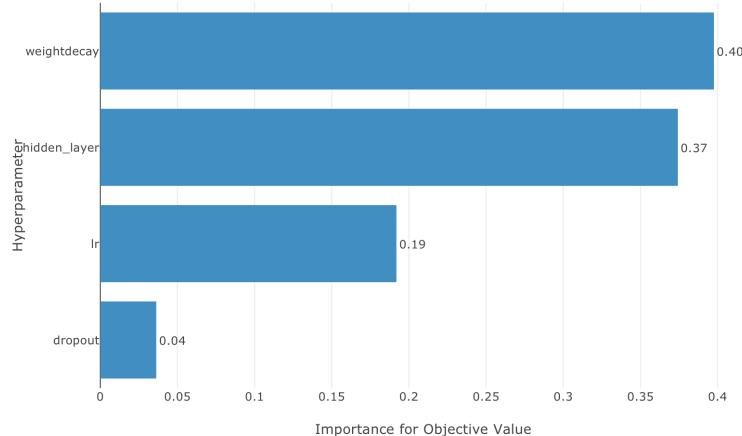
図 5.3: w_{node} (図中の α), w_{edge} (図中の β) に対する評価関数の 2 次元ヒストグラム。

図 5.4: 各パラメータの重要度

1059 5.2.4 学習評価

1060 学習における損失関数の値と識別精度を図 5.5 に示す。100 エポック の学習によって、損失
 1061 関数の値は降下したのちおおよそ安定しており、学習精度はおよそ 82.5% となっている。ま
 1062 た、図 5.6 に混合行列を示す。混合行列とは、分類問題で出力された結果をまとめた表であり、
 1063 縦軸(行)が実際の答えを横軸(列)が学習結果を表していて、実際の答えごと(行)に正規化
 1064 されている。そのため、左上から右下の対角成分が実際の答えと学習結果が一致している割合
 1065 を表しており、図 5.5 における学習精度は $(\text{学習精度}) = (\text{対角成分の和}) / (\text{全成分の和})$ とし
 1066 て求められる。図 5.6 より、ディープニューラルネットワークでは c ジェットについて uds
 1067 ジェットと識別しにくいことがわかる。

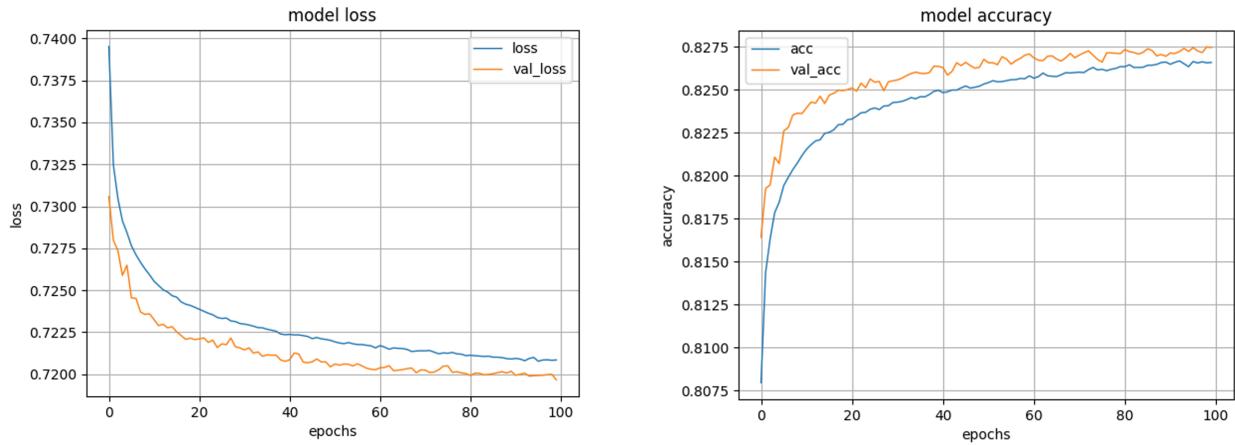


図 5.5: (左) 学習の経過における損失関数。(右) 学習の経過における学習精度

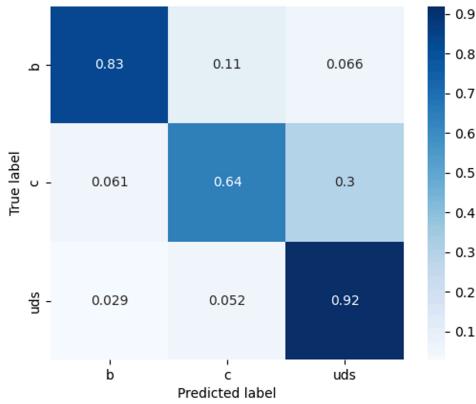


図 5.6: ディープニューラルネットワークの学習における混合行列。縦軸が実際の答えを、横軸が学習結果を表している。

1069 次に、LCFIPlus との比較を示す。本アルゴリズムのモデルの出力値は 1 つのジェットに
 1070 対してそれぞれのフレーバーである確率値が得られる。そして、物理解析においてはフレー
 1071 バー識別の効率が重要であり、識別効率 (= フレーバーがある確率値で決定した場合の正し
 1072 い認識率) に対して背景事象と誤認識してしまう割合が少なくなることが望ましい。図 5.7,
 1073 5.8 は b/c ジェットの識別効率を示している。b ジェットの識別効率は LCFIPlus と比較して
 1074 劣っており、特に識別効率 80% における識別割合は、LCFIPlus が c ジェットの識別割合が
 1075 7.3%、uds ジェットの識別効率が 0.74% であるのに対して、ディープニューラルネットの c
 1076 ジェットの識別割合は 20% 程度、uds ジェットの識別効率は 2% 程度と、uds ジェットの識
 1077 別割合は 30 倍近く高くなっている。一方で c ジェットの識別効率は識別効率 80% において、
 1078 LCFIPlus が b ジェットの識別割合が 22%、uds ジェットの識別効率が 24% であるのに対し
 1079 て、ディープニューラルネットの b ジェットの識別割合は 数 % 程度、uds ジェットの識別効
 1080 率は 24% 程度と、b ジェットの背景割合は半分以下になっている。本来フレーバー識別は b、
 1081 c クォークの寿命が長いという特徴を活かして識別するという発想にあるため、崩壊点に関する

る変数で線形にモデリングした BDTs では、c ジェットよりも b ジェットの方が識別効率が高くなっている。一方でディープニューラルネットワークでは、関係性が逆転している。機械学習アルゴリズムの性能はデータの特性やタスクの種類に大きく依存するため、同じデータを用いていてもアルゴリズムが異なる場合、同じタスクに対する学習特性は異なることがある。今回の場合は c ジェットの識別において、ディープニューラルネットワークでデータ内の複雑な関係を処理できたため、このような結果になったと考えることができる。

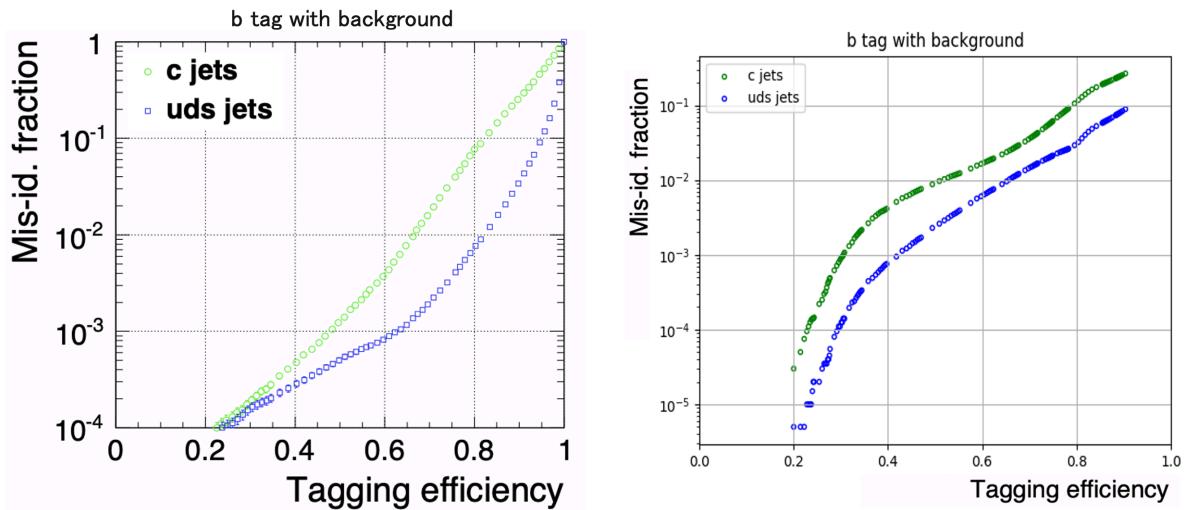


図 5.7: LCFIPlus(左)とディープニューラルネットワーク(右)による b フレーバージェットの識別効率の比較。緑: b ジェットに対する c ジェットの識別効率、青: b ジェットに対する uds ジェットの識別効率

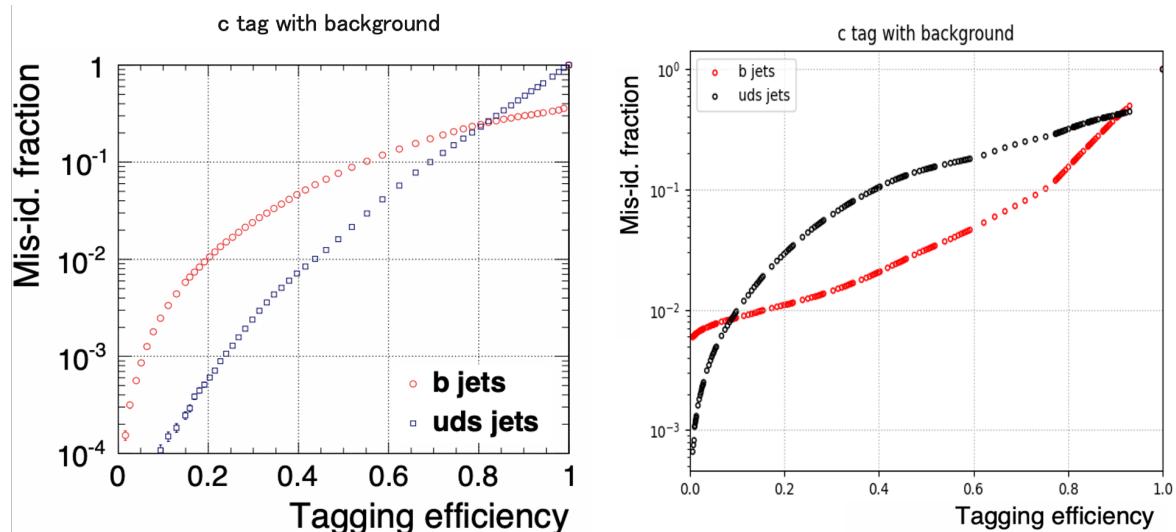


図 5.8: LCFIPlus(左)とディープニューラルネットワーク(右)による c フレーバージェットの識別効率の比較。赤: c ジェットに対する b ジェットの識別効率、黒: c ジェットに対する uds ジェットの識別効率

識別効率=0.8	背景ジェット	誤認率	
		LCFIPplus	DNN
b ジェット	c ジェット	0.073	~ 0.2
	uds ジェット	0.0074	~ 0.02
c ジェット	b ジェット	0.22	~ 0.7
	uds ジェット	0.24	~ 0.24

表 5.5: 識別効率 (Tagging Efficiency) に対するジェット誤認率 (Mis-id fraction)

5.3 グラフニューラルネットワークによる実装

本節では、グラフニューラルネットワークによるフレーバー識別の実装について述べる。

5.3.1 実装目的

5.2 節におけるディープニューラルネットワークの実装では、一部で精度の向上が見られたものの顕著な識別精度の改善はなかった。そこで、ILC のフレーバー識別により最適化した学習を行うために、より高い表現力を持つグラフニューラルネットワークによる実装を考えた。ILC におけるジェットの物理現象の振る舞いはトポロジカルな構造を持っており、またジェット内の飛跡同士は崩壊点を共有するものが存在している。そのため、グラフデータによってトポロジカルな構造を再現し、データ間の相互関係を活かした学習を行うことで、高い表現力によるアルゴリズムモデルの最適化を目指した。

さらにグラフニューラルネットワークの実装では、データをグラフ構造化する上でジェットの内部構造を理解するようなモデルを構築する必要があり、これを補助学習としたアルゴリズムの統合が可能であると考えた。つまり、これまでのようなプロセスを分離した高度なアルゴリズムではなく、1つのアルゴリズム内で飛跡に関する入力変数のみから、飛跡の分類や崩壊点の予想と同時にジェットフレーバーの識別を行うことを可能とするものである。これによつて、情報損失を少なくすることで性能の向上を目指すとともに、アルゴリズム調整の単純化や、飛跡の分類結果など将来的に他のアプリケーションに利用可能な情報の生成までもを期待することができる。

5.3.2 飛跡によるグラフデータセット

ILC におけるジェットの振る舞いを再現するため、本アルゴリズムでは 1 つのジェットに対して、飛跡をノードとし全ノードがエッジによって結ばれる 1 つの同種グラフを構築した。(図 5.9) 各ノードは表 5.5 に挙げる特徴量を入力変数として持ち、エッジは特徴量を持たないものとした。ノードの特徴量には、飛跡に関するパラメータ (Appendix を参照) を用いた。入

1111 力データの生成のため、まず 250 GeV の ILD フルシミュレーションにおける $e^+e^- \rightarrow \nu\nu H$
 1112 事象を 24 万イベント使用し、飛跡再構成を行った。続いて LCFIPlus の Vertex Fitter プロ
 1113 セスによって、飛跡対が結合する確率値を得た。またシミュレーションにおける答えをもとに
 1114 次の 3 つの学習に対して次のような答えラベルを準備した。それぞれ、ノードの答えラベルは
 1115 飛跡がどのフレーバージェットのどの vertex から来たものであるのか、エッジの答えラベルは
 1116 飛跡対が同じ崩壊点を共有するかのバイナリ形式、グラフの答えラベルはどのフレーバーで
 1117 あるのかを答えとして設定した。各学習の答えラベルについて表 5.7～表 5.9 にまとめる。ま
 1118 た、入力変数には前処理を行った。さらにデータ数は答えラベル間で偏りがあり、ノードに關
 1119 するラベルあたりのデータ数 5.10 は Primary Vertex を由来とする飛跡が突出して多かった
 1120 ため、比の逆数を重みとして損失関数で考慮することで対応した。また、本アルゴリズムの実
 1121 装は PyTorch とグラフニューラルネットワーク向けのグラフ構造を扱うための PyTorch ライ
 1122 ブラリである PyTorch Geometric を用いて行なった。

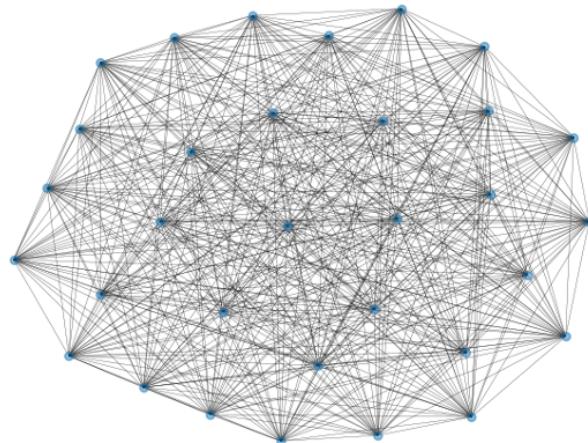


図 5.9: 今回の学習に用いたグラフデータの一例。ノードは飛跡を、グラフ全体が 1 つのジェットに対応する。

変数名	説明
d_0	xy 平面射影における IP と飛跡の距離
ϕ	xy 平面射影における飛跡軌道の方位角
ω	xy 平面射影における飛跡の曲率
z_0	sz 平面射影における IP と飛跡の距離
$\tan \lambda$	sz 平面射影における dz/ds
$\sigma(d_0)$	d_0 のフィッティングにおける誤差
$\sigma(z_0)$	d_0 のフィッティングにおける誤差

表 5.6: 各ノード (飛跡) が持つ特徴量。詳細については Appendix を参照。

ラベル	説明
PV	primary vertex 由来の飛跡
SVBB	b フレーバージェットの secondary vertex 由来の飛跡
TVCC	b フレーバージェットの tertiary vertex 由来の飛跡
SVCC	c フレーバージェットの secondary vertex 由来の飛跡
Others	上記の崩壊点を持たない飛跡

表 5.7: ノード分類の答えラベル

ラベル	説明
Connected	エッジの結ぶ飛跡対同士が崩壊点を構成する
Not-Connected	エッジの結ぶ飛跡対同士が崩壊点を構成しない

表 5.8: リンク予測の答えラベル

ラベル	説明
$b\bar{b}$	b フレーバージェット
$c\bar{c}$	c フレーバージェット
$q\bar{q}$	uds フレーバージェット

表 5.9: グラフ分類の答えラベル

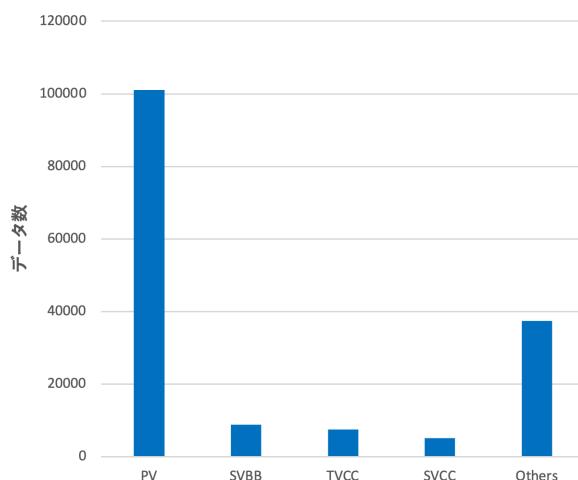


図 5.10: ノードラベル毎のデータ数

1123 5.3.3 ネットワークアーキテクチャ

1124 上述の通り、本アルゴリズムではフレーバー識別に加えて、補助学習を導入する。具体的に
1125 はノード分類とリンク予測、グラフ分類の 3 つのタスクを同時に使うネットワークで、図 5.11
1126 のようなモデルを構築した。

1127 はじめに、飛跡の変数をより多次元ベクトル化し潜在表現を得るために全結合層を設置する。
1128 その後、Graph Attention Network (GAT) を 3 層設置し、各 GAT 層の後にはバッチ正規化
1129 と LeakyReLU 関数の活性化関数を置く。そして学習をノード分類、リンク予測、グラフ分類
1130 の 3 つに分岐させる。ノード分類では、飛跡の由来となる崩壊点の種類によって 5 つの出力が
1131 全結合層によって設計されている。リンク予測では、エッジの隣接行列を取得してノード間に
1132 隣接があるかないかを判断する 2 分類問題を行っており、これはエッジが崩壊点となりうるの
1133 かの判断に等しい。グラフ分類では、pooling によってグラフ全体の特徴量を 1 つの次元に置
1134 き換え、各フレーバーらしさを出力する設計になっている。

1135 GAT の学習は 4.2.3 節で述べた通りであり、本アルゴリズムのグラフデータは全結合のグラ
1136 フニューラルネットワークを構築しており、ノードに対するアテンション α はエッジの隣接
1137 関係を用いて式 4.35 のように学習される。このようなアテンション機構によってグラフ全体、
1138 すなわちジェットフレーバーの識別を補助することを目的に、ノードの識別の補助学習を導入
1139 した。また、実際に飛跡対が崩壊点を共有している場合、その隣接関係は物理的に重要なもの
1140 となるため、その情報を学習に活かすことを目的にリンク予測を補助学習に加えた。また、3
1141 つの学習の出力層に置いた全結合層はそれぞれ独立に実行される。

1142

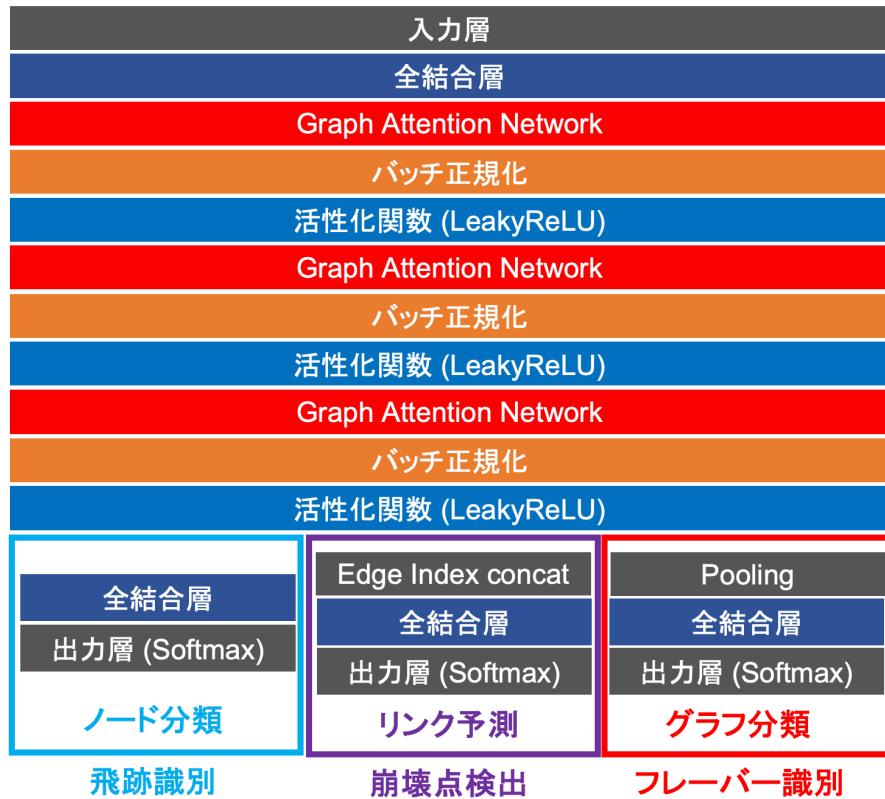


図 5.11: フレーバー識別のためのグラフデータを用いたネットワークの概略図

また本アルゴリズムでは 3 つの異なる学習を同時に行う必要があり、目的によって相対的な学習のしやすさが異なるため、次のような損失関数をデザインした。

$$L_{\text{total}} = w_{\text{node}} L_{\text{node}} + w_{\text{edge}} L_{\text{edge}} + L_{\text{graph}} \quad (5.1)$$

1143 L はそれぞれの損失関数を、 w_{node} と w_{edge} はそれぞれノード、エッジの損失関数にかける重
1144 みを表す。グラフ分類と同程度の損失関数に収束するために、ノード識別とリンク予測における
1145 損失関数に重み付けを行った。

1146 また、学習におけるハイパーパラメータにはベイズ最適化を行った上で、表 5.9 に挙げるも
1147 のを用いた。最適化アルゴリズムには、Adam に徐々に学習率を上げる Warmup 機構を加え
1148 た RAdam による学習が最も精度が高くなった。

ノード数	(512, 512, 512)
アテンションヘッド数	16
活性化関数	LeakyReLU 関数 (0.001)
損失関数	交差エントロピー
最適化アルゴリズム	RAdam ($\beta_1 = 0.8, \beta_2 = 0.9, \text{eps} = 10^{-8}$)
学習率	0.01
Weight decay	0.0001
w_{node}	3.0
w_{edge}	1.0
エポック数	100
バッチサイズ	128

表 5.10: グラフニューラルネットワークにおけるハイパーパラメータ

¹¹⁴⁹ 5.3.4 ハイパーパラメータの最適化

¹¹⁵⁰ ネットワークのハイパーパラメータは、ベイズ最適化法によってチューニングを行った。
¹¹⁵¹ チューニングしたパラメータは表 5.9 に加え、RAdam における学習率のハイパーパラメータ
¹¹⁵² $\beta_1, \beta_2, \text{eps}$ などについて実行した。チューニングにはディープニューラルネットワークの実装
¹¹⁵³ におけるものと同じ Optuna ライブラリを用いたが、評価関数として検証用データの損失関数
¹¹⁵⁴ のうち L_{graph} のみを用いて、最小になるような最適化(図 5.12)を行った。今回の学習において
¹¹⁵⁵ 重要になる損失関数の割合 w_{node} (図中の α)、 w_{edge} (図中の β)に関する 2 次元ヒストグラム
¹¹⁵⁶ の値は、図 5.13 のように α が 0.5、 β が 3 のあたりで評価関数の値が極小値をとった。また、
¹¹⁵⁷ 今回の最適化に使用したパラメータの重要度は GAT 層におけるドロップアウトと、RAdam
¹¹⁵⁸ におけるパラメータ β_2 が最大となった(図 5.14)。

変数名	最適化範囲
中間層のノード数	32 ~ 1024
アテンションヘッドの数	1 ~ 20
ドロップアウト	0.1 ~ 0.9
LeakyReLU	0.00001 ~ 0.1
β_1	0.5 ~ 0.9
β_2	0.5 ~ 0.9
eps	$10^{-9} \sim 10^{-7}$
学習率	0.00001 ~ 0.1
weight decay	0.000001 ~ 0.1
$w_{\text{node}}(\alpha)$	0.0 ~ 5.0
$w_{\text{edge}}(\beta)$	0.0 ~ 5.0

表 5.11: 最適化を行ったハイパーパラメータとその値の範囲。

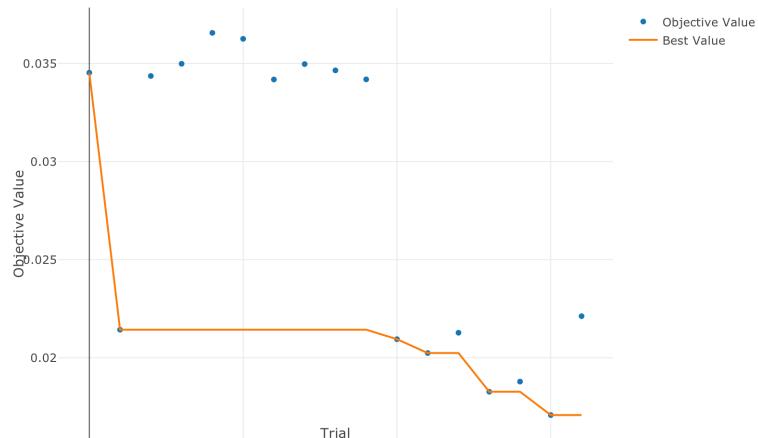


図 5.12: 各試行における評価関数の値。値が最小になるようなアルゴリズムを実行している。

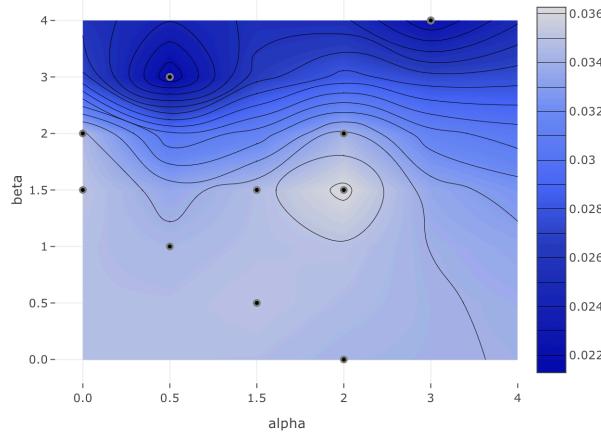


図 5.13: w_{node} (図中の α), w_{edge} (図中の β) に対する評価関数の 2 次元ヒストグラム。

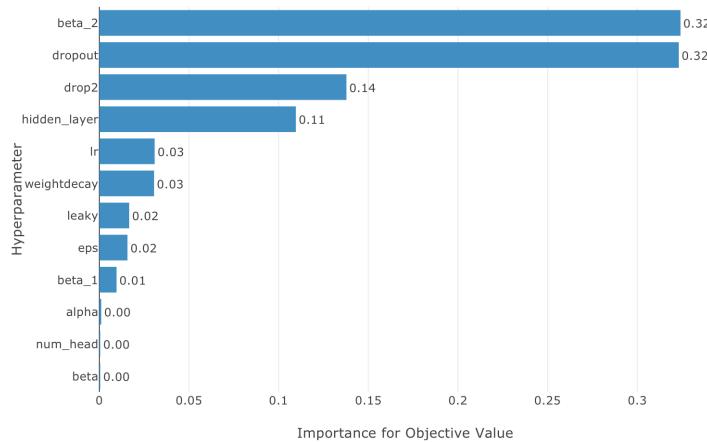


図 5.14: 各パラメータの重要度。 β_2 は RAdam における学習の進行度のパラメータ、ドロップアウトは学習におけるアテンション係数の隣接ノードへの参照率を表す。

1159 5.3.5 学習評価

1160 はじめに、学習全体の経過を図 5.15 に示す。学習の経過を通して損失関数は減少しており、
 1161 特に 80 エポック付近で急激な降下を見せた。しかしそれ以降は過学習が進んでしまい、損失
 1162 関数の値は安定化させることが難しかった。また、ノード分類、リンク予測、グラフ分類それ
 1163 ぞれの精度においても、学習の経過を通して向上が見られた。

1164

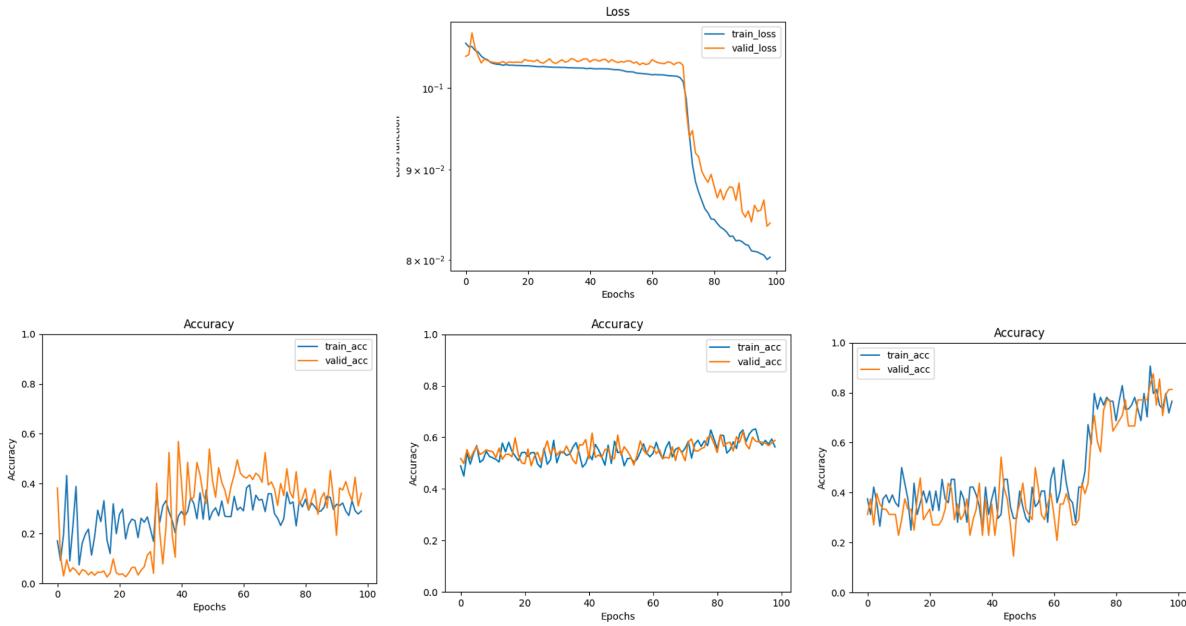


図 5.15: (上) 学習の経過における損失関数。(左下) 学習の経過におけるノードの学習精度、(中央下) リンクの学習精度、(右下) グラフの学習精度。

1165 続いて、ノード、リンク、グラフ分類における混合行列を図 5.16 に示す。ノード、リンクの
 1166 学習精度は十分良いと言える制度ではなく、一方でグラフの識別においてはディープニューラ
 1167 ルネットワークでの実装に匹敵する精度が得られた。ノードの混合行列では PV と SVBB と
 1168 識別する割合が高く、一方で TVCC、Others について分類が出来ず PV に分類されてしまう
 1169 傾向にあった。

1170

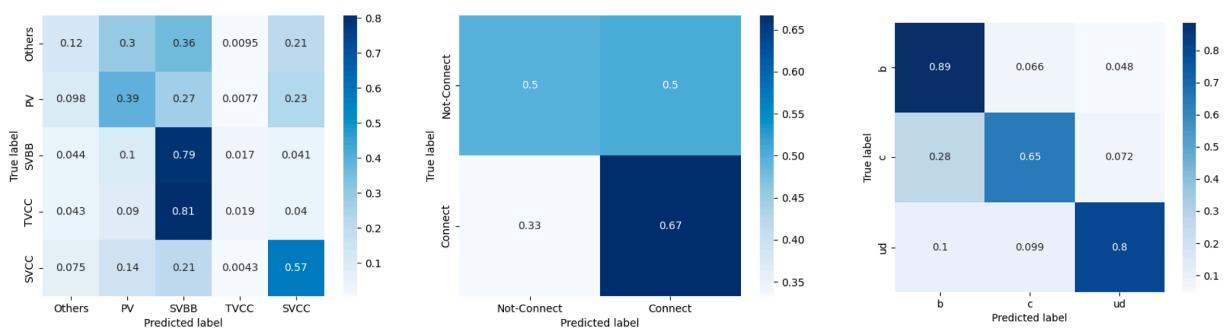


図 5.16: (左) ノード分類の混合行列、(中央) リンク予測の混合行列、(右) グラフ分類の混合行列。

1171 最後に、LCFIPlus との比較を示す。ディープニューラルネットワークの比較と同様に、図
 1172 5.17, 5.18 は b/c ジェットの識別効率を示している。b ジェットの識別効率は c ジェットにつ
 1173 いて向上しており、識別効率 80% における識別割合は、LCFIPlus が c ジェットの識別割合

が 7.3%、uds ジェットの識別効率が 0.74% であるのに対して、グラフニューラルネットワークの c ジェットの識別割合はおよそ 1.5%、uds ジェットの識別効率もおよそ 1% ほどという結果になり、c ジェットの誤認率が 5 分の 1 程度になっているものの、uds ジェットに関しては若干の精度低下となった。また、c ジェットの識別効率は非常に悪い結果となってしまい、識別効率 80% における識別割合は、LCFIPlus が b ジェットの識別割合が 22%、uds ジェットの識別効率が 24% であるのに対して、グラフニューラルネットワークの b ジェットの識別割合はおよそ 40%、uds ジェットの識別効率はおよそ 15% ほどという結果になり、b ジェットの誤認率が 2 倍近く上がってしまった。ノード分類結果では、PV と SVBB と識別する割合が高いことから、b ジェットに対する識別効率が高くなっていることが考えられる。また、c ジェットの b ジェットとの分離が悪い点に関しては、主に SVBB, TVCC に対する SVCC の分類精度が課題となっていると推測される。

1185

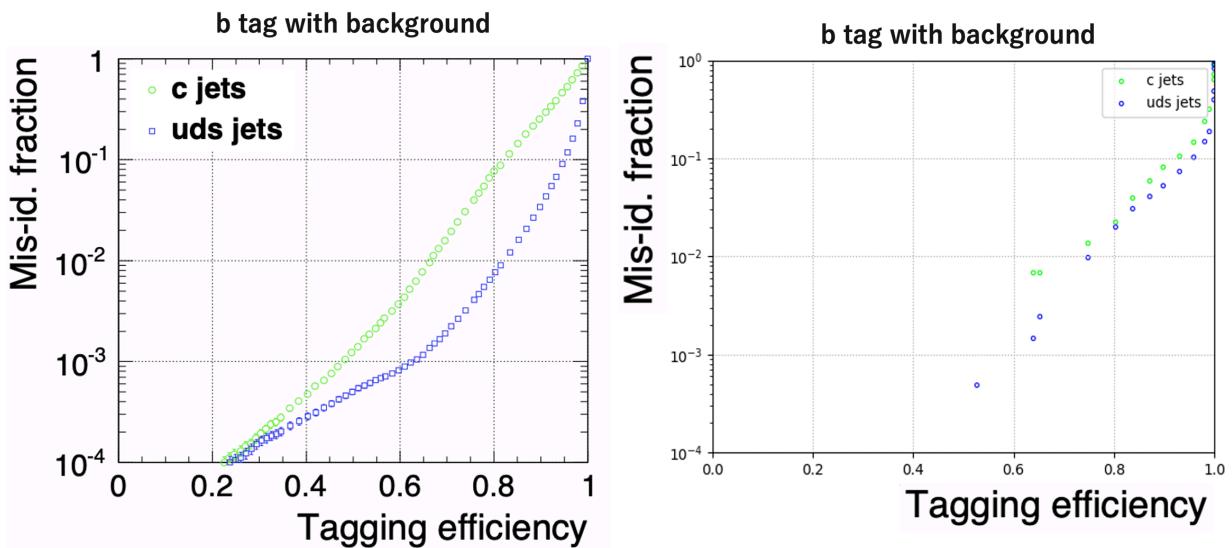


図 5.17: LCFIPlus(左) とグラフニューラルネットワーク(右)による b フレーバージェットの識別効率の比較。緑: b ジェットに対する c ジェットの識別効率、青: b ジェットに対する uds ジェットの識別効率。

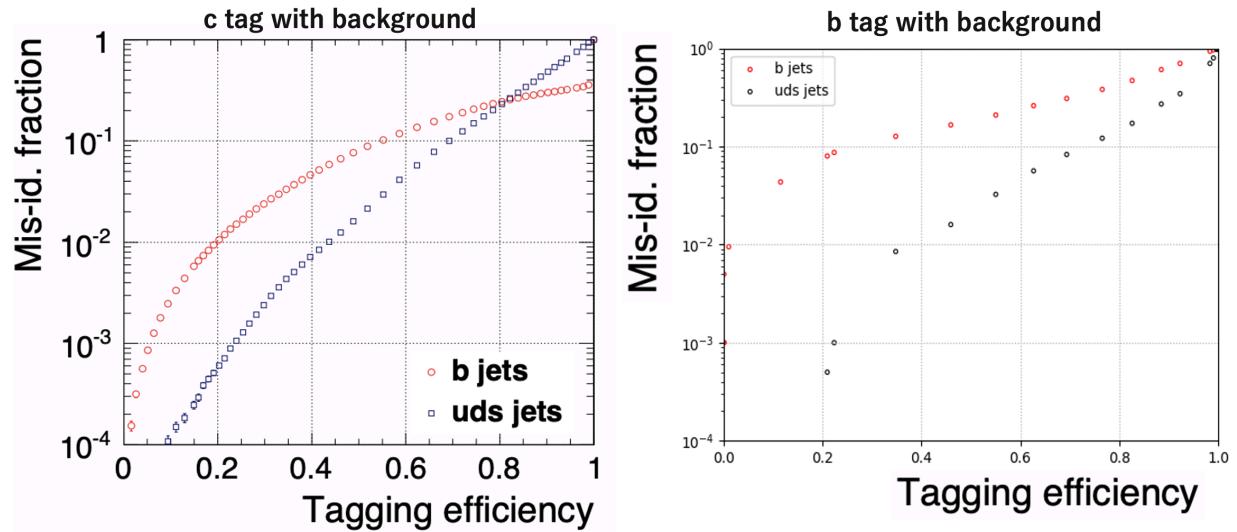


図 5.18: LCFIPlus(左) とグラフニューラルネットワーク(右)による c フレーバージェットの識別効率の比較。赤:c ジェットに対する b ジェットの識別効率、黒:c ジェットに対する uds ジェットの識別効率。

識別効率=0.8	背景ジェット	誤認率	
		LCFIPlus	GNN
b ジェット	c ジェット	0.073	~ 0.015
	uds ジェット	0.0074	~ 0.01
c ジェット	b ジェット	0.22	~ 0.40
	uds ジェット	0.24	~ 0.15

表 5.12: 識別効率 (Tagging Efficiency) に対するジェット誤認率 (Mis-id fraction)

結果として、グラフニューラルネットワークによる実装では LCFIPlus と比較して顕著な性能向上があったとは言えない結果になった。しかしながらエッジリンク予測の実行によって、これまで別のプロセスで実行していた崩壊点検出を、1つのネットワーク内で内包することが出来た。これによって同程度の性能ではあるが、低レベルの入力に対して高レベルの出力をを行う、中間層の情報損失の少ない学習を行うことが可能になった。

第 6 章

まとめと今後の展望

1193 本論文では、電子陽電子ヒッグスファクトリーのためのジェット測定技術の研究として、特
 1194 に ILC を念頭に置いた 2 つのテーマの研究を行った。1 つ目のテーマでは、ILC の検出器案
 1195 である ILD の電磁カロリメータのうち、SiW-ECAL の技術プロトタイプをハドロンビーム試
 1196 験によって性能評価した。2 つ目のテーマでは、深層学習を用いたフレーバー識別アルゴリズ
 1197 ムの開発を行った。本章ではそれぞれの研究のまとめと今後の展望について述べる。

1198 ILD Si-W ECAL プロトタイプのハドロンビーム試験による性能評価

1199 解析が完了次第記入

1200 深層学習を用いたフレーバー識別アルゴリズム

1201 本研究では、深層学習を用いたフレーバー識別アルゴリズムの開発を行った。ヒッグスファ
 1202 クトリーの目的であるヒッグス粒子は、クォークやグルーオンなどに崩壊し、多数のハドロン
 1203 の束であるジェットとして検出器に到達する。そのためジェットの再構成でクォークのフレ
 1204 バーを識別することは、ヒッグスの物理を知るための解析において非常に重要な情報となる。
 1205 フレーバー識別はジェットの構成粒子の種類や運動量、崩壊点に関する情報から、ジェットの
 1206 元となるクォークのフレーバーを識別するプロセスであり、現在ジェットの再構成に用いられ
 1207 ている LCFIPlus では従来の機械学習手法である BDTs が用いられている。これに対して深
 1208 層学習を導入することで、識別性能の向上や崩壊点検出アルゴリズムとの統合などを目指し
 1209 た。

1210 はじめに、最も単純な構造のネットワークモデルであるディープニューラルネットワークに
 1211 よる実装を行った。アルゴリズムには過学習対策や学習の効率化に向けた手法を用いて、全結
 1212 合層を中心としたディープニューラルネットワークのモデルを、PyTorch を用いて構築した。
 1213 学習の入力データとして、250 GeV の ILD フルシミュレーションのイベント 400 万イベント
 1214 を、変数には BDTs の学習に用いた変数と同じものを用いた。学習を重ねるにつれて損失は
 1215 減少し、学習全体の精度はおよそ 82.5% で得られた。また LCFIPlus との比較にあたっては、

1216 識別効率あたりの誤認率によって比較した。結果として、c ジェットの識別など一部で改善が
1217 見られたものの、LCFIPplus と大差ない、あるいはやや劣る結果となった。

1218 続いて、グラフデータを用いたフレーバー識別アルゴリズムの実装を行った。動機として
1219 は、フレーバー識別において重要な IP 付近の物理現象をグラフ構造のデータとして構築する
1220 ことで、物理量を羅列する 1 次元の数値データと比較して表現力が上がり、性能が向上するの
1221 ではないかという狙いがあった。またグラフデータを構築する際の副産物として、以前は別プロ
1222 プロセスであった崩壊点検出を 1 回の識別で同時に処理できるという長所もあった。グラフデ
1223 タは独自のデータセットクラスを作成し、飛跡をノードとする全結合のグラフを、1 ジェット
1224 あたり 1 グラフ構築した。また、このグラフではディープニューラルネットワークと異なり、
1225 ノードのみ飛跡再構成におけるフィッティングパラメータを特徴量を持つとした。イベントは
1226 250 GeV の ILD フルシミュレーションのイベント 240 万イベントを用いた。ネットワークアル
1227 ゴリズムは GAT を 3 層使用し、損失関数の工夫やデータ次元変更などの手法を用いてノ
1228 ノード識別/リンク予測/グラフ分類の 3 つのタスクを同時に実行し、重みを共有するようなネット
1229 ワークモデルを PyTorch Geometric を用いて実装した。学習は不安定であったものの損失は
1230 減少し、ディープニューラルネットワークに匹敵するグラフ分類の精度が得られた。また識別
1231 効率あたりの誤認率のプロットによる LCFIPplus との比較では、b ジェット識別において改善
1232 が見られたが、c ジェット識別では非常に悪い結果となった。またノード識別、リンク予測に
1233 対しては十分に良い精度を得ることができなかった。しかし、低い精度ながら崩壊点検出アル
1234 ゴリズムの統合を達成することができた。

1235 今後の展望として、フレーバー識別アルゴリズムには大きく 2 つの方向性で改善が可能である
1236 と考えている。

1237 1 つ目は深層学習の理論・技術面での方向性である。GAT に挙げられるネットワークの処理
1238 に関してより理解を深め、フレーバー識別により最適なネットワーク構造を提案することや、
1239 物理的な性質をより組み込むなどフレーバー識別に最適化された損失関数を設計するなど、演
1240 算手法によって更なる改善の見込みがあると考えている。また、グラフデータの設計について
1241 も改善ができると考えており、具体例としては今回のデータセットにおいて崩壊点となるエッジに
1242 特徴量を持たせることや、グラフデータの構成要素を変更することなどが挙げられる。現在
1243 エッジは特徴量を持っておらず、接続したノードの情報のみで更新を行っているが、エッジ
1244 自身が特徴量を持つことでデータの学習パラメータが増え、精度が向上すると考えることができる。
1245 また、本研究では飛跡をノードとするグラフを構成したが、本来飛跡は点ではなく曲線
1246 の形状を取るため、やはり崩壊点のような物理的に点となる量をノードとしてすることで、実際の
1247 物理現象に近いグラフの構築が期待できると考えている。

1248 2 つ目は、iLCSoft としての実装に向けた方向性である。例えば出力情報の設計や C++ 環境
1249 への移行があると考えている。出力設計に関して、今回のグラフデータを用いたフレーバー識
1250 別では崩壊点検出を包含するアルゴリズムを開発した。しかし、現段階では解析において崩壊
1251 点の情報が必要になった際、物理量として出力することは出来ない。そのため方針としては、
1252 中間層で得たネットワークの潜在的なパラメータを回帰問題として設計することで、新しく崩

1253 壊点など数値データの出力を得ることなどが考えられる。また別の例としては、開発したフ
1254 レーバー識別アルゴリズムの C++ 環境への移植である。現在ジェット再構成に用いられている
1255 LCFIPlus は Marlin プロセッサーの 1 部であり、C++ 環境で動作している一方で、本研
1256 究で開発したアルゴリズムは python によって記述されている。そのため PyTorch/PyTorch
1257 Geometric C++ API を用いる、あるいは C++ によって書き換えを行うことによる iLCSoft
1258 への実装は、今後の課題の一つである。

1259

1260 付録 A

1261 付録 A

1262 A.1 LCIO parameter

1263 記入予定です。

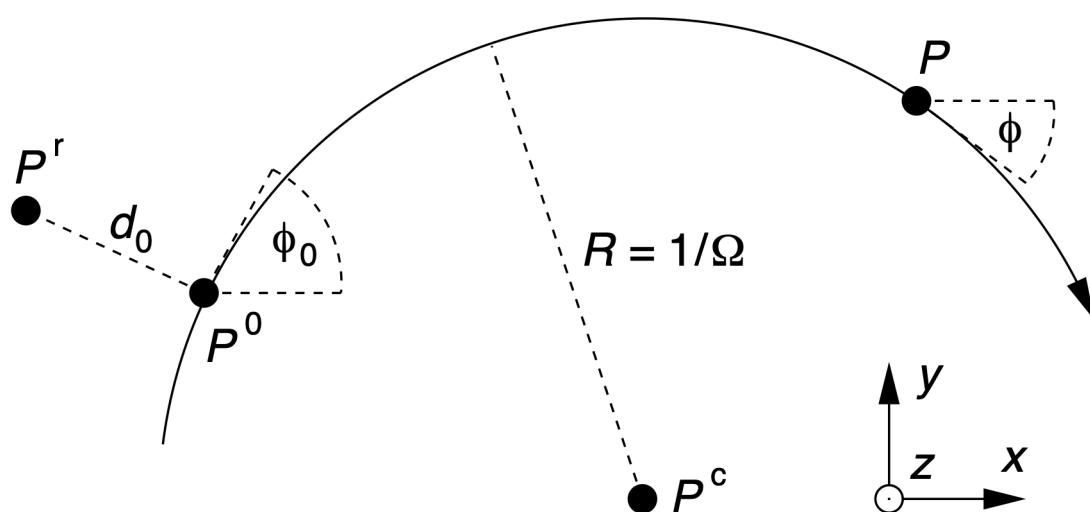
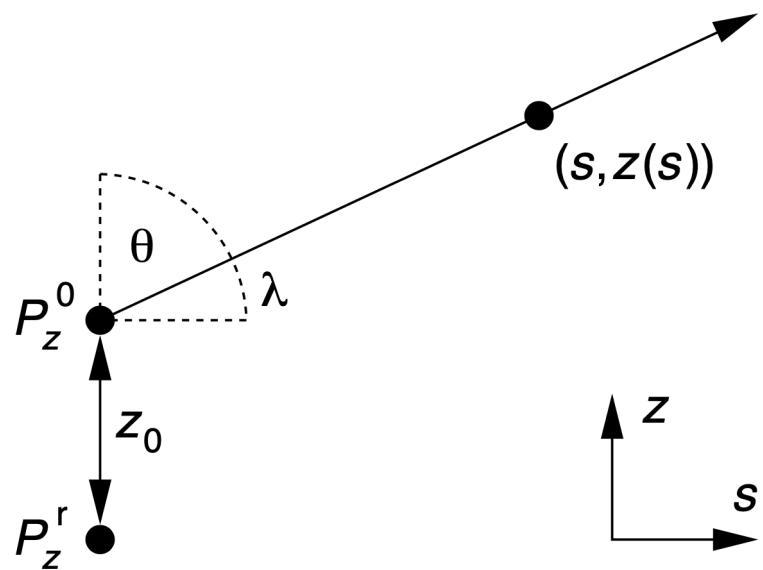


図 A.1: xy 平面

図 A.2: sz 平面

謝辞

1264

参考文献

- [1] S. L. Glashow, Partial-symmetries of weak interactions, Nuclear Physics (1960).
- [2] 藤井恵介, 高エネルギー加速器セミナー OHO: ILC の物理.
- [3] S. D. Bass, A. D. Roeck, and M. Kado, The higgs boson implications and prospects for future discoveries, Nature Review Physics (2021).
- [4] H. Baer *et al.*, The international linear collider: Technical design report, (2013).
- [5] K. Fujii *et al.*, Physics case for the international linear collider, (2015).
- [6] T. Suehara and T. Tanabe, Lcfiplus: A framework for jet analysis in linear collider studies, Nucl.Instrum.Meth. A808 , 12 (2016).
- [7] S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock, and B. Webber, New clustering algorithm for multijet cross sections in e+e – annihilation, Physics Letters B **269**, 432 (1991).
- [8] R. Brun and F. Rademakers, ROOT - An Object Oriented Data Analysis Framework, 1996.
- [9] J. Therhaag, TMVA Toolkit for multivariate data analysis in ROOT, PoS ICHEP2010 (2010).
- [10] A. Irles, Siwecal-tb-analysis, (2019).
- [11] R. Frank, The perceptron: A probabilistic model for information storage and organization in the brain, Psychological Review (1958).
- [12] 斎藤康毅, ゼロから作る DeepLearning volume 1 (オンラインジャパン, 2016).
- [13] L. Liu *et al.*, On the variance of the adaptive learning rate and beyond, ICLR 2020 (2020).
- [14] X. Glorot and Y. Bengio, Understanding the difficulty of training deep feedforward neural networks, Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics (2010).
- [15] J. Zhou *et al.*, Graph neural networks: A review of methods and application, KeAi (2020).
- [16] 村田剛志, グラフニューラルネットワーク Pytorch による実装 (オーム社, 2022).
- [17] W. L. Hamilton, R. Ying, and J. Leskovec, Inductive representation learning on large graphs, NIPS (2017).

- 1295 [18] P. Veličković *et al.*, Graph attention networks, ICLR 2018 , 12 (2017).
- 1296 [19] P. I. Frazier, A tutorial on bayesian optimization, arXiv:1807.02811 (2018).
- 1297 [20] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, Optuna: A next-generation
1298 hyperparameter optimization framework, KDD 2019: Proceedings of the 25th ACM
1299 SIGKDD International Conference on Knowledge Discovery Data Mining (2019).