

2022年度 修士論文

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

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

尾上 友紀

指導教員 末原 大幹 川越 清以

2023年1月12日



九州大学
KYUSHU UNIVERSITY

概要

2012 年に Large Hadron Collider (LHC) で発見されたヒッグス粒子は、素粒子の質量や宇宙の物質の起源の解明につながる粒子であり、そ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生成し、低バックグラウンドな環境で精密測定することで、ダークマターの候補になる新粒子探索などの標準理論を超える物理を切り開く足がかりとして期待されている。本研究で念頭に置いている国際リニアコライダー (International Linear Collider; ILC) 計画は次世代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期待されている。ヒッグスファクトリーの物理において重要な事象はジェットを含んでおり、ジェットを高い精度で測定・再構成することは、物理解析の性能に直結する。本研究ではジェット測定技術の研究として、検出器候補であるシリコンタングステン電磁カロリメータ (Sillicon Tangsten Electromagnetic Calorimeter; SiW-ECAL) の性能評価と、ジェット再構成におけるフレーバー識別の深層学習を用いたアルゴリズム開発を行なった。

ILC で重要となるイベントはジェットを生成するが、ILC の目標感度を達成するためにはおよそ $\sigma_E/E = 30\% \sqrt{E(\text{GeV})}$ の高いジェットエネルギー分解能が必要である。ILC では Particle Flow Algorithm (PFA) 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、これには非常に高精細な電磁カロリメータが求められる。SiW-ECAL はタングステンの吸収層とシリコンパッドセンサーの検出層が交互に積み重なったサンドウィッチ構造をしており、シリコンパッドセンサーは 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いる。また、シリコンパッドは ASIC を搭載した読み出し基板に接着されており、PFA に最適化したコンパクトな構造となっている。現在日本やフランスを中心とした国際協力によって技術プロトタイプの開発が行われており、本論文ではその性能評価について記述する。特に、欧州原子核研究機構 (CERN) の The Super Proton Synchrotron (SPS) で行なった、高エネルギーhadronビームを用いたテストビーム実験の結果から、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報告する。

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、深層学習技術を用いて行った。フレーバー識別は、実験データからジェットを再構成する過程において、ジェットの軌跡や崩壊点に関する情報から、親粒子のフレーバーを識別するプロセスである。従来技術である LCFIPlus では、フレーバー識別に機械学習技術である Boosted Decision Trees (BDTs) が採用されている。近年の粒子物理学の分野では、物理解析やシミュレーションに深層学習を導入することで精度や性能を改善する取り組みが行われており、本研究においても識別性能の向上を目的に深層学習を実装した。本研究では特にグラフ形式で記述されたデータを学習に扱うグラフニューラルネットワーク (GNN) を使用した。GNN では一般的なニューラルネットワークと比較して、物理現象の幾何学構造を活かし情報損失を減らすことができる。(本アルゴリズムの結果を LCFIPlus と比較したところ、現状ではより良い識別精度を得ることは出来なかったが、従来は別にプロセスを行なっていた崩壊点検出のアルゴリズム

を統合することができた。)

目次

図目次

表目次

第1章

序論

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 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ターである、国際リニアコライダー計画 (International Linear Collider, ILC) の概要に触れたのち、1.3 節で ILC が探索する物理、1.4 節で ILC のについて述べる。

1.1 素粒子物理学

1.1.1 標準模型

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、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??のような標準模型を掲げており、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。

標準模型は、主に次に挙げる 2 つの公理に沿って記述されている。1 つ目に、物質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互作用・弱い相互作用・強い相互作用の 3 つである。

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ンと呼び、同様に 6 種類存在する。クォーク・レプトンとともに 3 つの世代と 2 つの電荷タイプをもっており、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・レプトンへと崩壊する。

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

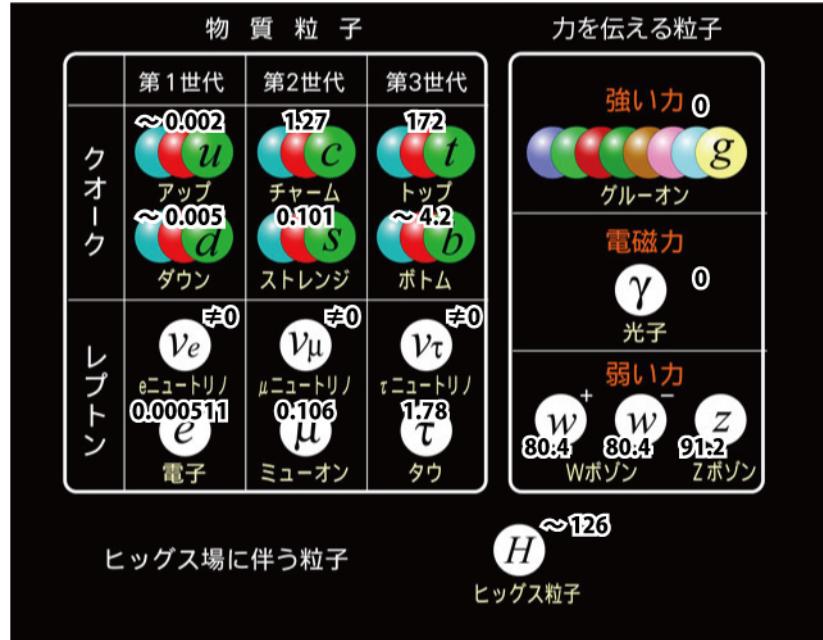


図 1.1: 素粒子の標準模型 (数値は質量 [GeV/c²])

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) 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。

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。このヒッグス機構では真空に複素 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は以下のように書ける。

$$V(\phi) = \mu^2 |\Phi|^2 + \lambda |\phi|^4 \quad (\text{ただし } \mu^2 < 0) \quad (1.1)$$

ポテンシャルは正であるため自己結合定数 $\lambda > 0$ であり、ポテンシャルは図??のような形状になる。

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

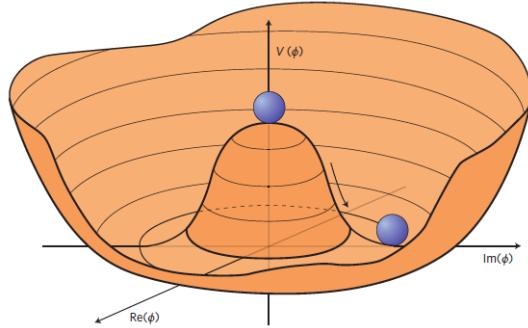


図 1.2: ヒッグスポテンシャル

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の有限の期待値として真空中に式??が現れる。

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

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、真空中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と、真空中のヒッグスの密度に比例すると考えることができ、質量は加速されにくさを表す量と考えることができる。

このヒッグス粒子は、2012 年 7 月に欧州原子核研究機構 (CERN) の大型ハドロン衝突型加速器 (LHC) 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ことを最大の目的としている。

1.2 国際リニアコライダー計画: ILC

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が計画されている電子陽電子衝突型線形加速器である。(図??) 全長 20km の線形加速器を用いて電子と陽電子を加速し、中央の 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。また ILC は重心系エネルギー $\sqrt{s} = 250\text{GeV}$ 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

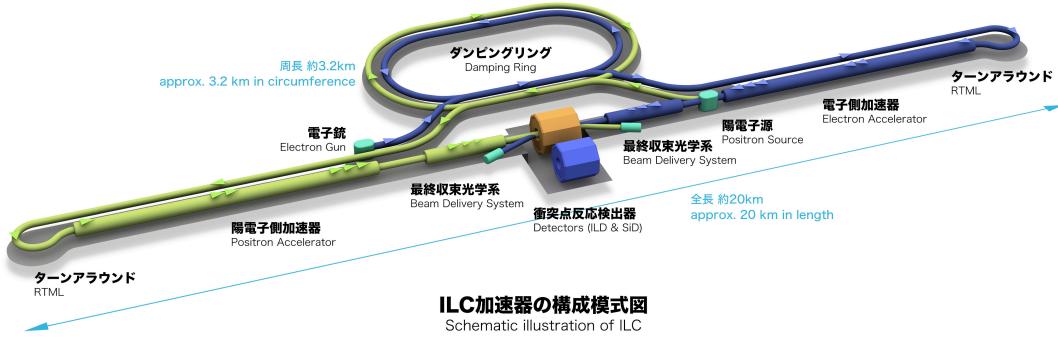


図 1.3: ILC の概略図

における物理プログラムのメインターゲットは次のようにになっている

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

ヒッグス粒子を発見した LHC と比較して ILC には以下の 5 つの利点が存在する。1 つ目は 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羅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。加えて全事象を記録しオフラインで事象選択を行うため、トリガーレスで運転することが可能である。3 つ目は ILC のビーム構造による点である。ILC のビームは 5Hz のパルス運転で、バンチ間隔が 554ns とレートが低く、トリガーレスの運転

で全データの取得が可能となる。4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 15mm 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 フレーバのみでなく c フレーバの識別も可能となる。5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。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 ヒッグス生成と結合定数の精密測定

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、ヒッグス粒子と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??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 ボソンを正確に測定することでヒッグス粒子の質量を高い精度で再構成することができる。

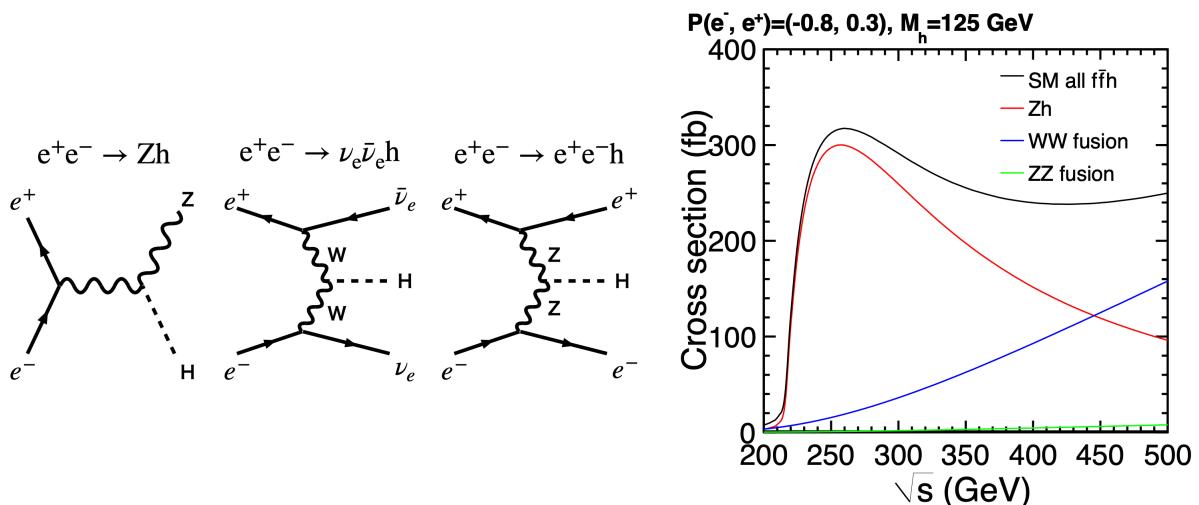


図 1.4: (左) ZH 随伴生成、WW 融合反応、ZZ 融合反応におけるファインマンダイアグラム。(右) ILC におけるヒッグス生成断面積。ヒッグス粒子の質量が $m_h = 125 \text{ GeV}$ であるとして、ZH 随伴生成、WW 融合反応、ZZ 融合反応をそれぞれ赤、青、緑線で示している。またこのとき電子・陽電子の偏極は、それぞれ電子が左巻き 90%、右巻き 10% の 80% であり、陽電子は左巻き 35%、右巻き 65% の 30% としている。

また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、質量の小さい 2つの粒子に崩壊する。標準模型におけるヒッグス粒子の崩壊分岐比は図??のようになっており、ILC 250 GeV における崩壊分岐比は表??のようになっている。この分岐比の測定精度は

信号事象を S 、背景事象を N とすると $S/\sqrt{S+N}$ 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% 以下まで下げることができる。そのため、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る。

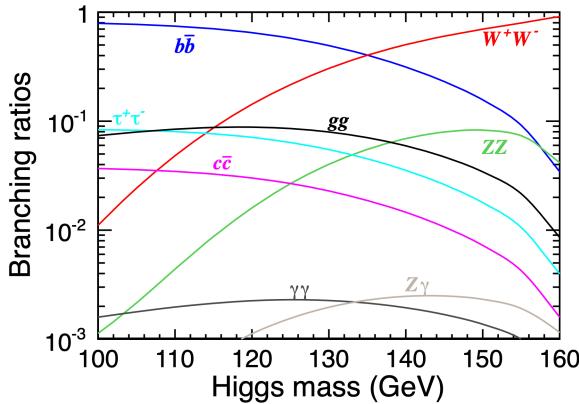


図 1.5: 標準模型におけるヒッグス粒子の質量と崩壊分岐比の関係

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%

図 1.6: ILC250GeV における SM ヒッグス粒子の崩壊分岐比

1.3.2 ヒッグス自己結合

500GeV 以上の ILC では図??のような、 $e^+e^- \rightarrow Zh h$ 反応の断面積を測定することで、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場の真空凝縮について検証することができる。ILC500GeV における図??の断面積は 0.2fb と小さく難しいため、崩壊 $e^+e^- \rightarrow Zh h \rightarrow q\bar{q}bb\bar{b}\bar{b}$ における b フレーバーの識別精度が非常に重要ななる。

1.3.3 階層性問題

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の質量は、繰り込みにおいて図??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発散してしまい、プランクスケール程度の質量を持つてしまうことが分かっている。そのため標準模型を超える新物理 (BSM) がないと仮定すると、質量の量子補正をキャンセルする解決策がなければ $125 \text{ GeV}/c^2$ 程度の質量を理論的に再現することができない。(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きいと言える。

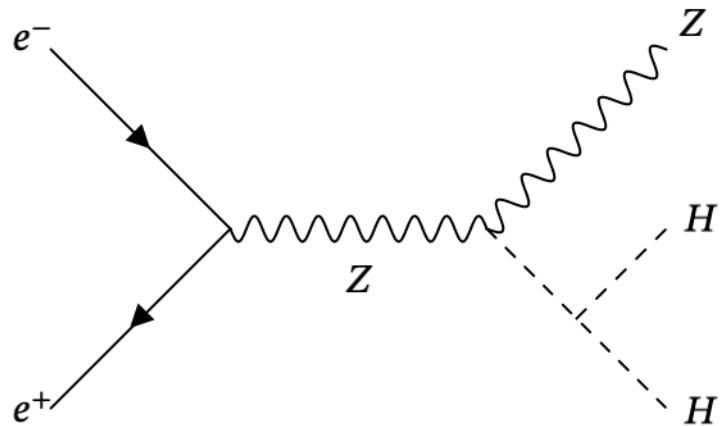
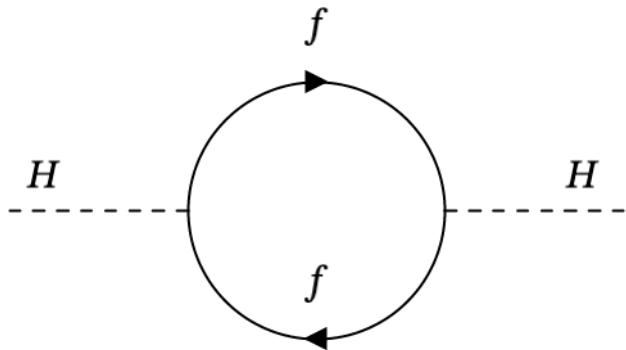
図 1.7: ヒッグス自己結合 $e^+e^- \rightarrow Zhh$ 

図 1.8: ヒッグス粒子の質量補正となるフェルミオンループ

- 超対称性理論超対称性理論 (Supersymmetry, SUSY) は、フェルミオンとボソンを交換する変換に対する不变性（超対称性）を定義する理論である。またこの理論においては、標準模型にお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。超対称性が完全である場合、標準模型粒子と質量や相互作用が同じである必要性があるが、現段階では SUSY 粒子は発見に至っていない。しかし階層性問題においては、超対称性によりヒッグスボソンの質量補正に関する 2 次の発散をフェルミオンの寄与で打ち消し、対数による発散に落とすことができる。
- 余剰次元理論余剰次元理論とは、四次元時空以外にも次元があるとする理論である。この理論では時空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きる。

1.3.4 その他の新物理

上にあげた階層性問題に関する物理に加え、WIMP (Weakly Interacting Massive Particle) のような暗黒物質の探索や、トップクォークの質量の精密測定、電弱相互作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を期待することができる。

1.4 ILC の検出器

ILC の検出器 (図??) には、日本の国々が中心となって開発が進められている 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 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 を共有できるように push-pull 方式を採用している。また ILD、SiD ともに後述の Particle Flow 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。

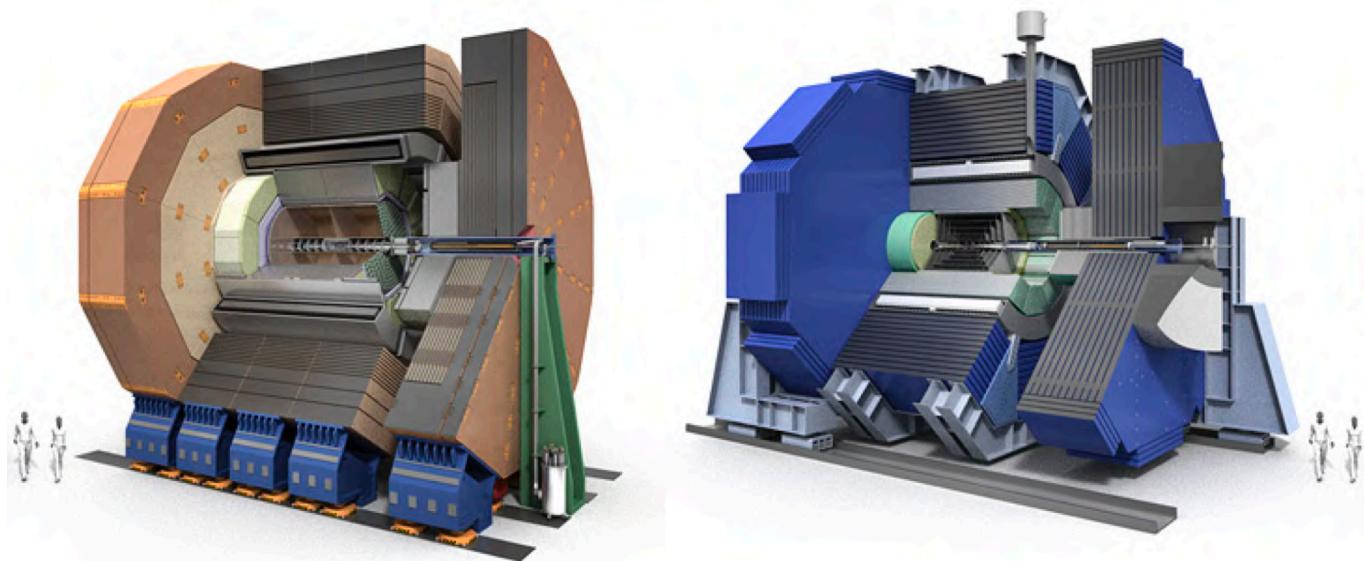


図 1.9: (左) ILD (右) SiD の全体図

1.4.1 International Large Detector: ILD

ILD 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 3.5T のソレノイドコイルが設置されている。図??に断面図を記す。

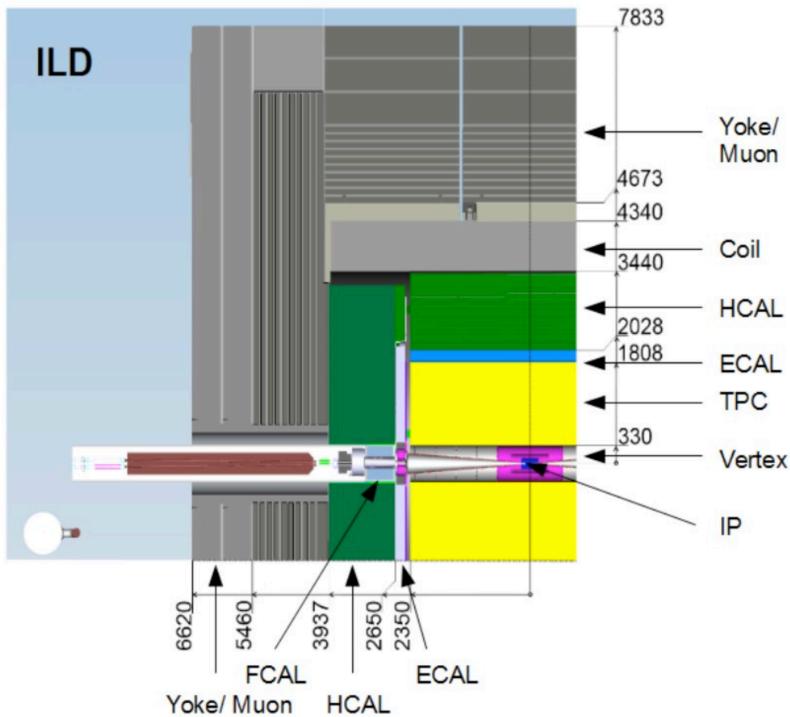


図 1.10: ILD の断面図

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、シリコンピクセルセンサーによって高い分解能で荷電粒子の生成点を測定する。飛跡の高精度な測定によって、短寿命粒子の崩壊点を高精度に再構成することができ、ILC では小型化や高速読み出しに向けて CMOS センサー、DEPFET、Fine Pitch CCD、SOI など様々な技術候補が研究されている。

中央飛跡検出器

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) 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 は大型のガスチェンバーであり、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る。また荷電粒子の飛跡を再構成することで運動量の測定や、信号の大きさからエネルギー損失を測定することができ、粒子識別において重要な役割を果たす。

カロリメータ

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

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。

電磁カロリメータ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グステンが検討されている。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(ScECAL)が提案されている。

ハドロンカロリメータは中性ハドロンのエネルギーを測定するための検出器である。HCALはECALに比べ大型であるため吸収層には鉄が用いられ、検出層には3cm角のSiPMタイルを用いてシンチレーション光を検出するアナログカロリメータ(AHCAL)と、1cm角のセルをRPCを用いてバイナリ信号で読み出すデジタルカロリメータ(SDHCAL)の2つが提案されている。

ミューオン検出器

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チャンバーとSiPMシンチレータストリップの両方が検討されている。

1.5 ILCのソフトウェア

Particle Flow Algorithm: PFA

ILCの電子陽電子衝突で生じる粒子は、ジェットの終状態で検出される。このジェットのエネルギー情報は粒子識別や事象再構成において重要であり、一般的にジェットのエネルギー分解能は次のように書ける。

$$\frac{\sigma_E}{E} = \frac{\alpha}{\sqrt{E}} \oplus \beta(E) \quad (1.3)$$

ここで、第一項はカロリメータに由来するであり、第二項は補正項である。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 $\sigma_E/E = 55\%/\sqrt{E(GeV)}$ 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 $\sigma_E/E = 30\%/\sqrt{E(GeV)}$ 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(PFA)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割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??)

崩壊モード	崩壊分岐比	ジェット内のエネルギー割合
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 1.1: ジェットを占める各粒子と対応する検出器

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

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 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 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 というツールキットを使用するという点で統一されている。

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使用するデータもシミュレーションデータである。シミュレーションにおいてはまず、モンテカルロ (Monte Carlo, MC) 法に基づく Wizard 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 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成される。図??に ILC におけるソフトウェアの流れを示す。

1.5.2 事象再構成

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と飛跡を推定する事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 PFA による再構成を行っている。さらに LCFIPlus を用いて事象を選別し、物理解析を行う。

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。ここで再構成された飛跡をもとに、さらに高次の再構成が行われる。

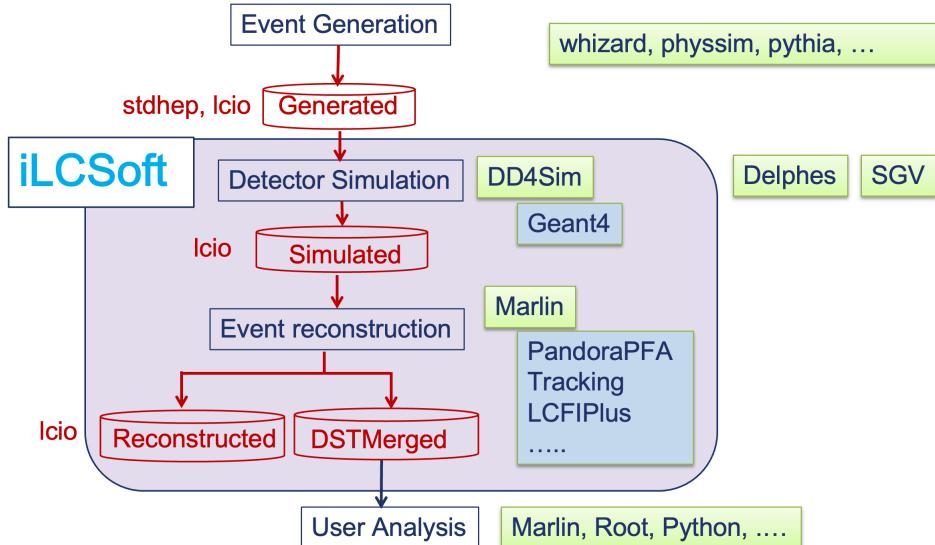


図 1.11: ILC におけるシミュレーションと事象再構成、物理解析の流れとソフトウェア

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一次崩壊点を secondary vertex と呼ぶ。一つのハドロンから発生する飛跡が 2 本以上ある場合、崩壊点は直接再構成することができ、LCFIPlus の崩壊点検出アルゴリズムでは 2 本以上の飛跡の交点をフィッティングすることで求める。iLCSoft においては崩壊点検出からフレーバー識別までを LCFIPlus というフレームワークで実行することができる。

ジェットクラスタリング

ILC の終状態の多くは 4 ジェット以上のジェットを持っており、これらを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。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の情報やレプトンの情報の特徴的な物理量や飛跡同士の開き角などを用いて、全飛跡に対してクラスタリングを行う。

フレーバー識別

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親粒子のフレーバー識別が行われる。LCFIPlus では ROOT の TMVA パッケージを使用して、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それ以外のフレーバーに比べて寿命が長いことから、衝突点から離れたところで崩壊するという特徴があ

る。(図??) ハドロンの寿命 τ は高速 c を用いて、 b フレーバーでおよそ $c\tau = 400 \sim 500 \mu m$ 、 c フレーバーでおよそ $c\tau = 20 \sim 300 \mu m$ 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なる。

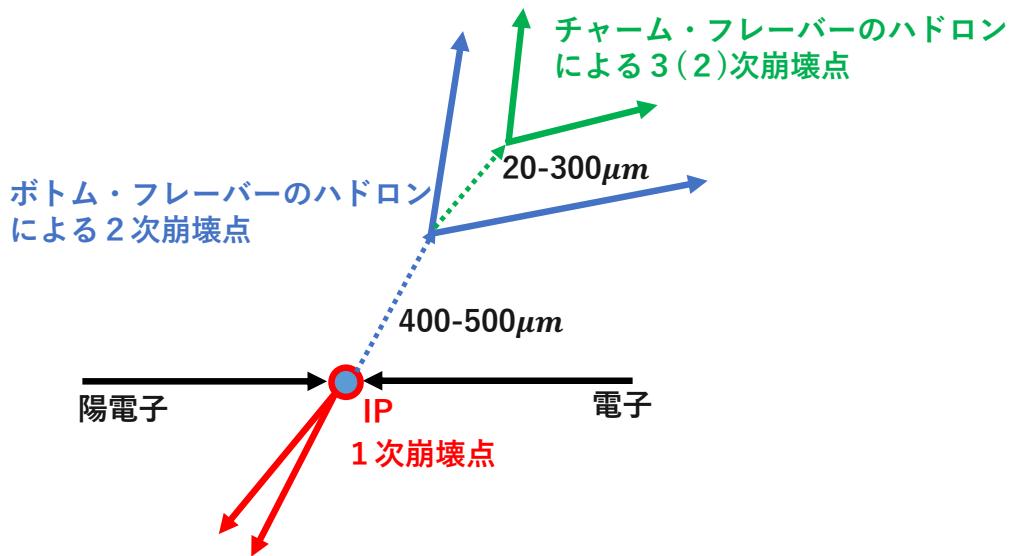


図 1.12: b フレーバージェットの崩壊の様子。赤線が1次崩壊点、青線が2次崩壊点、緑線が3次崩壊点を由来とする飛跡を表す。

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 350GeV 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。

1.6.1 高エネルギーhadron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、非常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プに高エネルギーなhadron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たフィードバックを得た。

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

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。グラフニュー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。

第2章

シリコンタングステン電磁カロリメータ

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電磁カロリメータの読み出し方法、また ASIC の設計性能やの読み出し方法、現在の技術プロトタイプについて説明する。

2.1 入射粒子と物質の相互作用

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。

2.1.1 荷電粒子

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、特にエネルギーの低いところでは電離損失の割合が、エネルギーの高いところでは制動放射の割合が高くなる。以下ではそれについて述べる。

電離損失

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	-
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	$\sqrt{\rho \langle Z/A \rangle} \times 28.816 \text{ eV}$

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。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??に示す。入射速度の小さい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 Particle) と呼ぶ。

制動放射

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 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。これを制動放射 (Bremsstrahlung) 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、

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

L_R は放射長 (Radiation Length) と呼ばれており、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る。

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

式 (??) 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

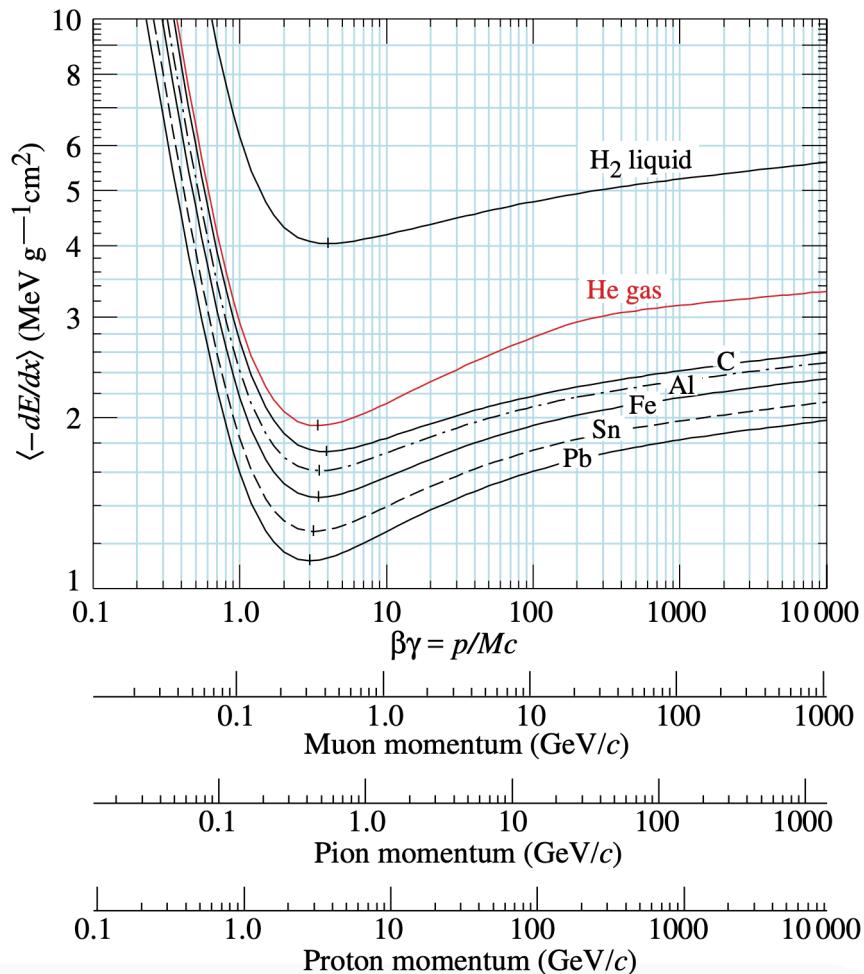


図 2.1: 水素 (液体)、ヘリウム (気体)、炭素、アルミニウム、鉄、スズ、および鉛における平均エネルギー損失と、入射粒子 (ミューオン、パイ中間子、陽子) の速度の関係。

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り、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??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。

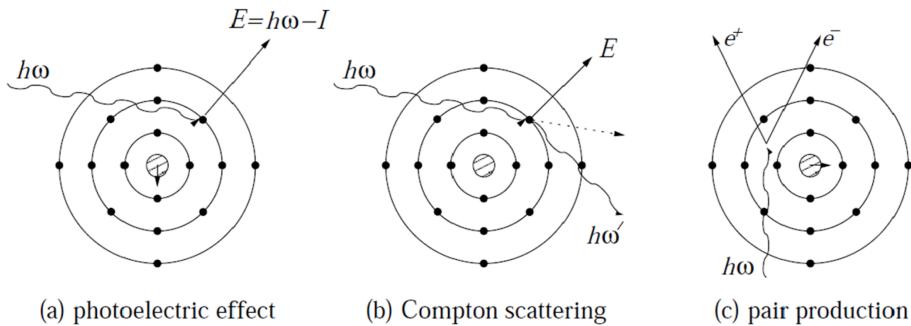


図 2.2: 光子と物質の相互作用

- 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 つの電子との弾性散乱。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 1.02MeV）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??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。電磁シャワーは発展するにつれて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収束する。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。

$$R_M \sim \frac{21(MeV)L_R}{\text{臨界エネルギー } (MeV)} (g/cm^2) \quad (2.5)$$

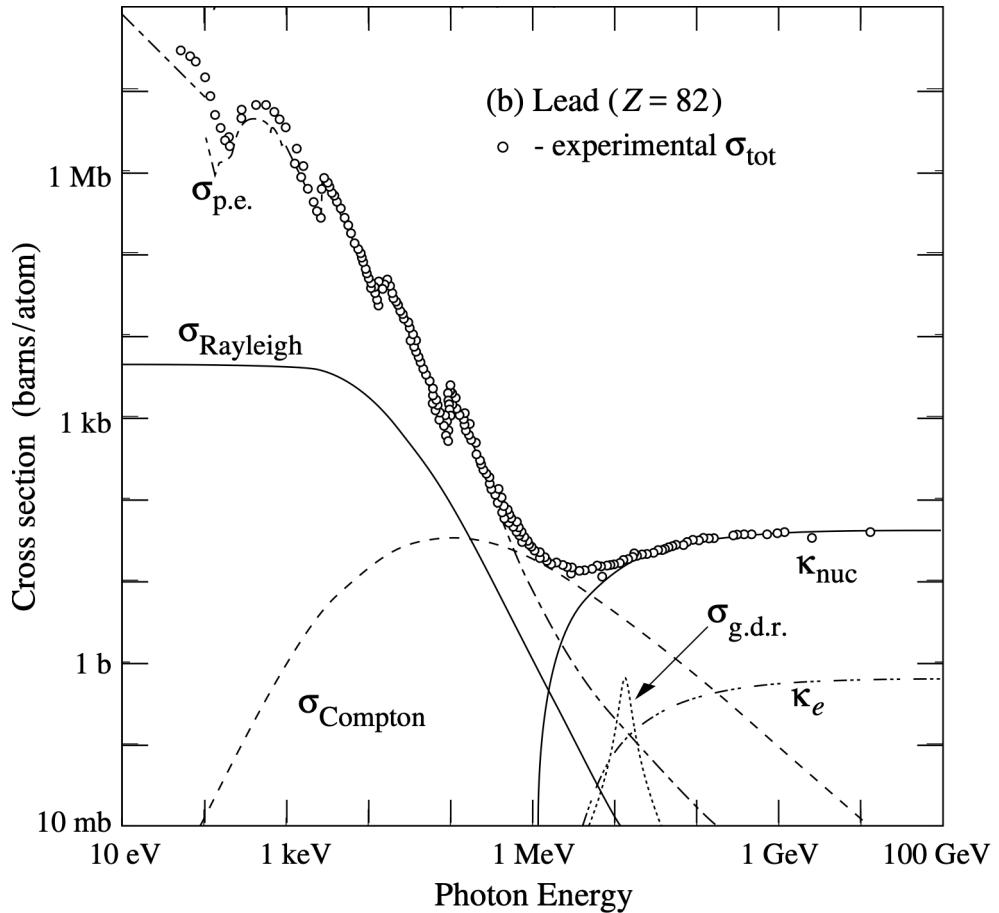


図 2.3: 光子と鉛の相互作用における断面積と光子のエネルギーの関係。 $\sigma_{p.e.}$ は光電効果を、 $\sigma_{Compton}$ はコンプトン散乱を、 κ_e, κ_{nuc} は電子陽電子対生成を指す。

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。また、ハドロンシャワー内には $\pi^0 \rightarrow \gamma\gamma$ 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。

2.2 粒子検出器の動作原理

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、エ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

うち特に重要なものを取り上げる。

2.2.1 ガス検出器

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。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や抵抗版チェンバー（RPC）などが挙げられる。

2.2.2 半導体検出器

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、半導体検出器では読み出しにおいて增幅を必要とするため、增幅回路が近傍（あるいは半導体内部）に存在している。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、検出器基盤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様々な構造が存在する。

2.2.3 シンチレーション検出器

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をシンチーレタという。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。シンチレーション光は非常に弱い光信号であるため、検出においては光電子増倍管を用いて信号を增幅し検出する。

2.3 シリコンタングステン電磁カロリメータ SiW-ECAL

2.3.1 SiW-ECAL の全体構造

ILD の SiW-ECAL は、図??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 枚のシリコン半導体センサーが貼り付けられる。

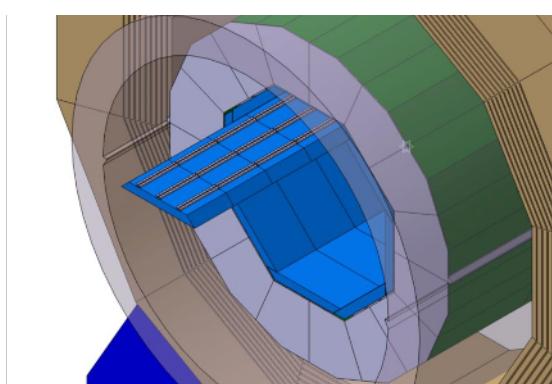


図 2.4: ILD および ECAL の全体図

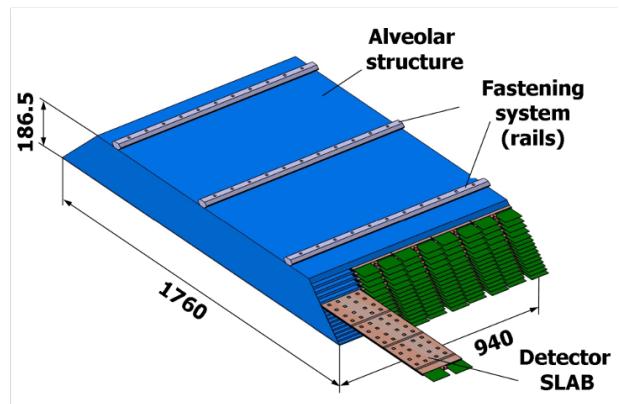


図 2.5: SiW-ECAL の構造

2.3.2 シリコン半導体検出器

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 $9 \times 9\text{cm}^2$ 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 つサブモジュールあたり 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 $320\mu\text{m}$ のセンサーでは、1MIP あたり 86.8keV 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に生成される電子正孔対はおよそ 24,000、電荷にして 4fC となる。シリコンセンサーは、常温硬化型導電性接着剤によって回路基盤 (PCB) と接着されており、PCB を通して信号の読み出しが行われる。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。

図 2.6: 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	$89.7 \times 89.7 \text{ cm}^2$
セルサイズ	$5 \times 5 \text{ cm}^2$
セル数	$16 \times 16 = 256$
厚さ	$320/500/650\mu\text{m}$
完全空乏化電圧	40/70/110 V

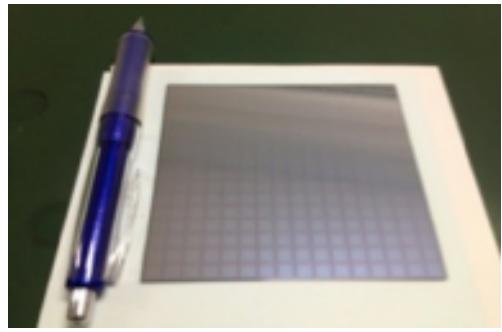


図 2.7: シリコンパッドセンサー

2.3.3 読み出しシステム

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しをコンパクトにする必要があり、読み出し専用の ASIC (Application Specific Integrated Circuit) 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 Omega グループが開発した SKIROC (Sillicon Kalorimeter Integrated ReadOut Chip) シリーズの第二バージョンである SKIROC2A を用いている。

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。

- 自動トリガー：
信号に対して ASIC 自身でトリガーをかける。
- 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、DAQ へ送信する。
データ量を圧縮し、またデータの劣化を防ぐことができる。
- 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、高い精度で粒子を識別する必要がある。

続いて、SKIROC2A の基本的な仕様を以下に示す。

- Austria Micro Systems 社製 $0.35 \mu\text{m}$ 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 ADC (Analog to Digital Converter) 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 ~ 150 MIP 相当の信号に対応
- Low gain … 150 ~ 2500 MIP 相当の信号に対応
- TDC (Time to Digital Converter) mode : 1ns 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、200ns 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 BCID (Bunch crossing ID) 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 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 0pF から 6.0pF まで 0.4pF 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、Memory cell にイベントを保持する。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 に保持された信号は、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 12bit 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??に SKIROC2A のアナログ部の回路図を示す。

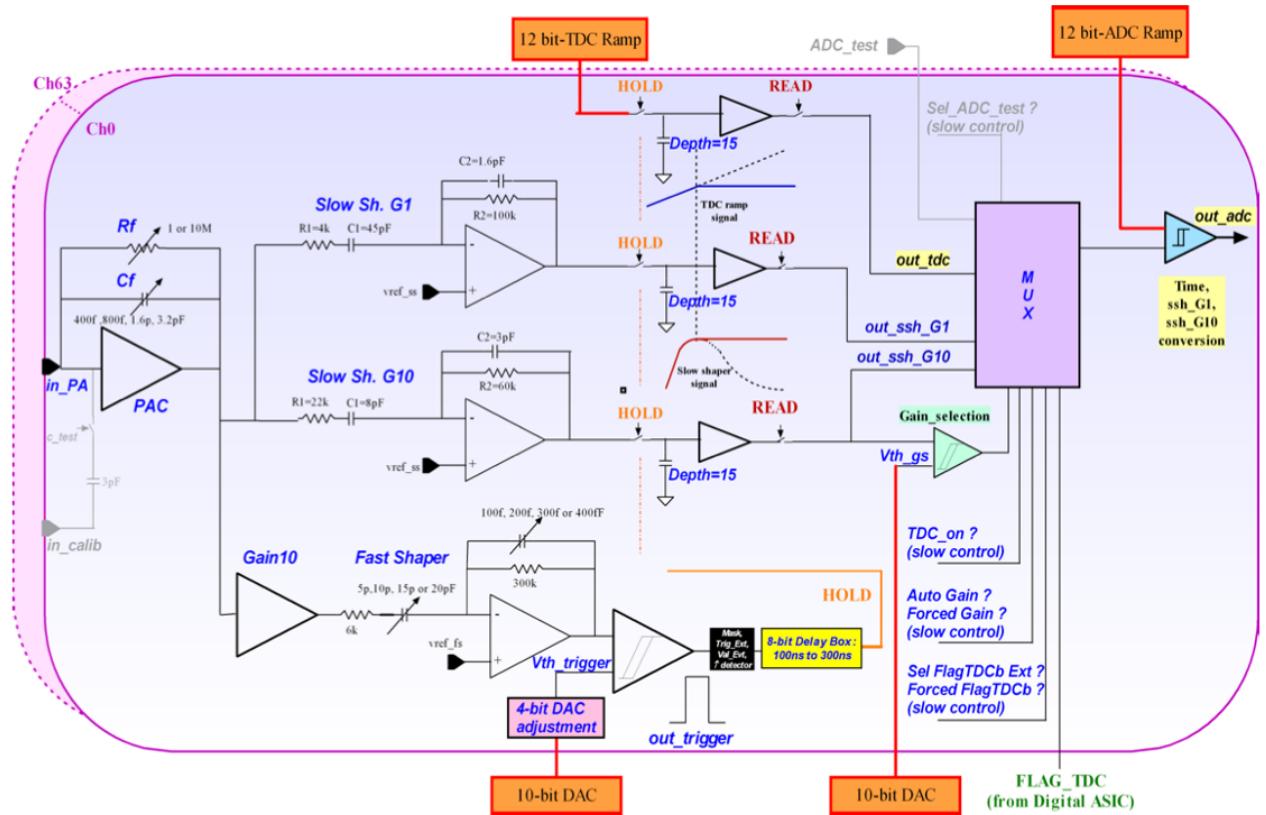


図 2.8: SKIROC2A のアナログ部の回路図

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??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 2.1: 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。)

2.3.5 技術プロトタイプ

SiW-ECAL の技術プロトタイプとして、図??のような構造が考えられている。図??は、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カーボンの板があり、その下に読み出し基板 (FEV) や ASIC の制御基板があり、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、さらにカーボ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、ASIC は FPGA を通して制御されており、さらに図??のようなモジュールを通して、複数の Slab からの信号を同時に読み出している。この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際協力で行われており、国内においてもモジュールの生産・組み立ては可能となっている。

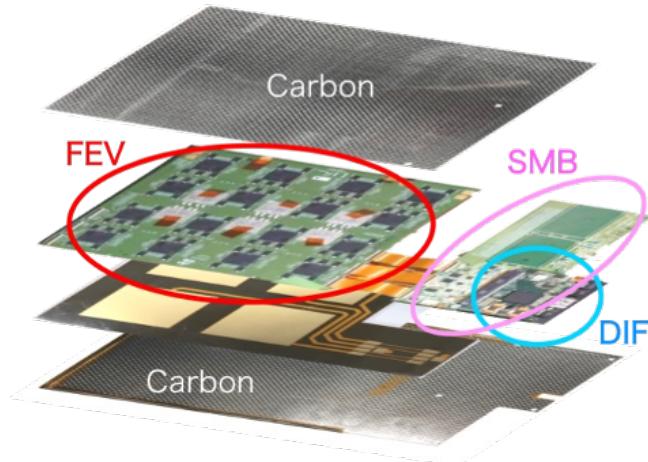


図 2.9: short slab の構造



図 2.10: 多層読み出しのための CORE モジュール

第3章

ビームテストによる評価実験

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。以下では実験の詳細と、結果について述べる。

3.1 CERN SPS

3.2 実験セットアップ

3.2.1 測定機器のセットアップ

3.2.2 EUDAQ による信号読み出し

3.3 実験結果

3.3.1 検出器応答

3.3.2 ペデスタル

3.3.3 スクエアイベント

3.4 まとめと考察

第4章

深層学習

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。

4.1 ニューラルネットワーク

4.1.1 パーセプトロン（単層ニューラルネットワーク）

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 年に考案された。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て出力信号を出すものである。図??に最も基本的なパーセプトロンの例を示す。

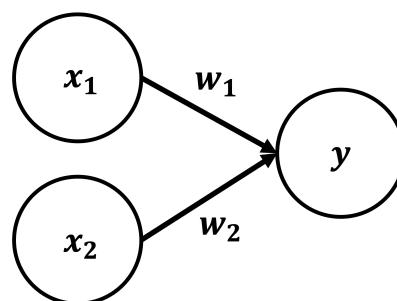


図 4.1: パーセプトロン

x_1, x_2 は入力信号、 y が出力信号であり、 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 \circ 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。

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

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習を繰り返すことで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2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。

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

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）

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, MLP）と呼ぶ。多層パーセプトロンの簡単な例を図??に示す。

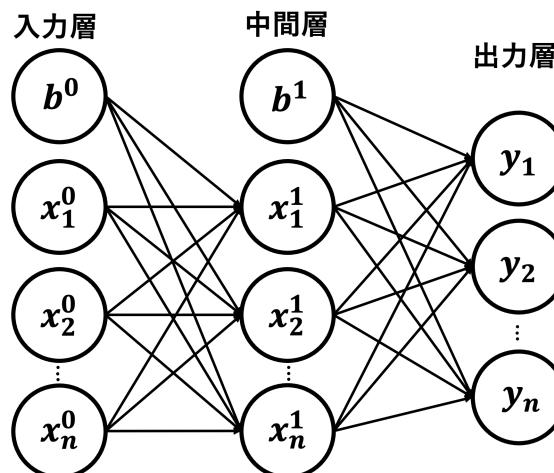


図 4.2: 多層パーセプトロン（ニューラルネットワーク）

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。活性化関数には主に非線形関数が用いられ、以下に主なものについて示す。

- 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）

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

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??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。

$$y_k = \frac{\exp(x_k)}{\sum_{i=1}^n \exp(x_i)} \quad (4.11)$$

損失関数

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、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??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。

$$E = \frac{1}{2} \sum_k (y_k - t_k)^2 \quad (4.12)$$

- 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??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 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。

$$E = - \sum_k t_k \log(y_k) \quad (4.13)$$

誤差逆伝播法

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、次のような処理を行う。

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ラベルとの誤差を求める。
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。これを局所誤差といふ。
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。
4. さらに前層へと処理を繰り返す。

ミニバッチ処理

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習させる。この束をミニバッチという。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、処理時間を短縮することができる。

最適化アルゴリズム

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。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??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。

$$\mathbf{W} \leftarrow \mathbf{W} - \eta \frac{\partial L}{\partial \mathbf{W}} \quad (4.14)$$

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 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 \mathbf{v} を加え、以下の式のように表される。

$$\mathbf{v} \leftarrow \alpha \mathbf{v} - \eta \frac{\partial L}{\partial \mathbf{W}} \quad (4.15)$$

$$\mathbf{W} \leftarrow \mathbf{W} + \mathbf{v} \quad (4.16)$$

上式における $\alpha \mathbf{v}$ 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。

$$\mathbf{h} \leftarrow \mathbf{h} + \frac{\partial L}{\partial \mathbf{W}} \odot \frac{\partial L}{\partial \mathbf{W}} \quad (4.17)$$

$$\mathbf{W} \leftarrow \mathbf{W} - \eta \frac{1}{\sqrt{h}} \frac{\partial L}{\partial \mathbf{W}} \quad (4.18)$$

ここで、 \mathbf{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 はモーメントの考え方と AdaGrad の考え方を融合させた手法であり、モーメントの変数 2つと前のステップまでの学習係数を表す変数 1つの 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。

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??のように層の数を多数持つモデルに対しては深い（ディープな）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。

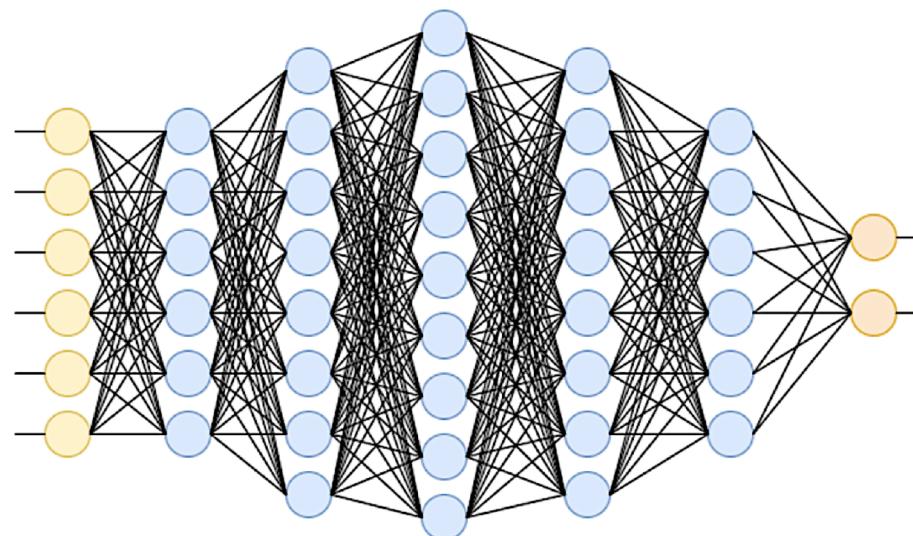


図 4.3: ディープニューラルネットワーク

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、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、化学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、高い表現力を持つことを強みとしている。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（Graph Neural Network, GNN）という。グラフ構造や演算技術、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げる。

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??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 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 h'_i 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。

$$h_{(i,j)} = f_{edge}(h_i, h_j, x_{(i,j)}) \quad (4.19)$$

$$h'_i = f_{node}(h_i, \sum_{j \in \mathcal{N}_i}, h_{(j,i)}, x_i) \quad (4.20)$$

ここで、 $h_{(i,j)}$ 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。

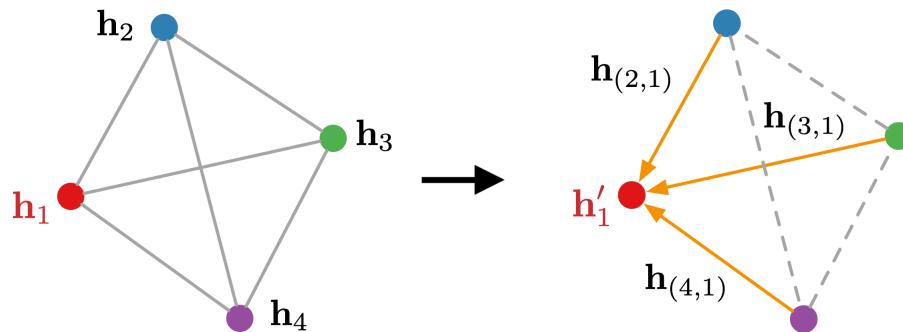


図 4.4: (左) 4 つのノードを持つ全結合グラフニューラルネットワーク。 h_i はノード表現を表す。 (右) メッセージパッシングの処理。

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性成分に変換し、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。

$$\mathcal{F}(\mathbf{x}) = \mathbf{U}^T \mathbf{x} \quad (4.21)$$

$$\mathcal{F}^{-1}(\mathbf{x}) = \mathbf{U} \mathbf{x} \quad (4.22)$$

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。

$$\mathbf{g} * \mathbf{x} = \mathcal{F}^{-1}(\mathcal{F}(\mathbf{g}) \odot \mathcal{F}(\mathbf{x})) \quad (4.23)$$

$$= \mathbf{U}(\mathbf{U}^T \mathbf{g} \odot \mathbf{U}^T \mathbf{x}) \quad (4.24)$$

$\mathbf{g} \star \mathbf{x}$ は畳み込み演算を、 $\mathbf{U}^T \mathbf{g}$ はスペクトル領域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、次のようにかくことができる。

$$\mathbf{g} \star \mathbf{x} = \mathbf{U}\mathbf{g}\mathbf{U}^T\mathbf{x} \quad (4.25)$$

上記のように、スペクトルによる畳み込みは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おり、さらに行列の固有値分解など計算量が多くなってしまう欠点がある。

Spatial Graph Convolution

グラフデータにおける畳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。

4.2.3 Graph Attention Network (GAT)

4.2.4 グラフニューラルネットワークの応用

第 5 章

深層学習を用いたジェットフレーバー識別

5.1 ジェットフレーバー識別アルゴリズム

5.1.1 事象再構成

5.2 イベントサンプル

5.2.1 前処理

5.3 ディープニューラルネットワークによる実装

5.3.1 ネットワークの構造

5.3.2 ハイパーパラメータの最適化

5.3.3 学習結果

5.4 グラフニューラルネットワークによる実装

5.4.1 飛跡によるグラフデータセット

5.4.2 ネットワークの構造

5.4.3 ハイパーパラメータの最適化

5.4.4 学習結果

5.5 LCFIPPlus との比較

5.6 まとめと今後の展望

第6章

まとめと今後の展望

こんな感じで作ってください。

付録 A

付録 A

謝辞