

لَهُ الْحُكْمُ وَإِلَيْهِ الْمُرْسَلُونَ



دانشگاه صنعتی اصفهان

دانشکده فیزیک

تخمین توزیع پایل آپ با استفاده از شبکه های عصبی گرافی در آزمایش CMS در برخورد دهنده بزرگ هادرولنی

پایان نامه کارشناسی ارشد فیزیک، گرایش ذرات بنیادی و نظریه میدان ها

مهرتاب جلال وندی

استاد راهنما

دکتر حامد بخشیان سهی



دانشگاه صنعتی اصفهان

دانشکده فیزیک

پایان نامه کارشناسی ارشد فیزیک، رشته گرایش ذرات بنیادی و نظریه میدان ها خانم مهتاب جلال وندی

تحت عنوان

تخمین توزیع پایل آپ با استفاده از شبکه های عصبی گرافی
در آزمایش CMS در برخورد دهنده بزرگ هادرولنی

در تاریخ ۱۴۰۳/۰۴/۱۶ توسط کمیته تخصصی زیر مورد بررسی و تصویب نهایی قرار گرفت:

دکتر حامد بخشیان سهی ۱- استاد راهنمای پایان نامه

دکتر صدیقه سجادیان ۲- استاد مشاور پایان نامه

دکتر عبیده جعفری ۳- استاد داور

دکتر سروش شاکری ۴- استاد داور

دکتر اسماعیل عبدالحسینی سارسری سرپرست تحصیلات تکمیلی دانشکده

تشکر و قدردانی

با قلبی سرشار از سپاس، از تک تک افرادی که در شکلگیری مسیر تحصیلی ام نقش آفرین بوده‌اند، قدردانی می‌کنم. عمیق‌ترین قدردانی را به راهنمای گرانقدرم، جناب آقای دکتر حامد بخشیان، ابراز می‌دارم. راهنمایی‌های ارزشمند ایشان نقشی اساسی در تکمیل موفقیت‌آمیز پروژه جذاب پایان‌نامه کارشناسی ارشد من ایفا نموده است. می‌خواهم عمیق‌ترین سپاسگزاری خود را به خانواده‌ام ابراز کنم. عشق و حمایتی که در طول زندگی به من داده‌اند، نقش اساسی در رسیدن به اهدافم داشته است. همچنین می‌خواهم از تمامی دوستانم، به خاطر تشویق‌های مداوم و ایمانشان به من تشکر کنم. حمایت شما برایم بسیار ارزشمند است و من قدردان تک تک شما هستم. بدون شما، من هرگز نمی‌توانستم به جایی که امروز هستم برسم. از صمیم قلب از شما متشرم.

کلیه حقوق مالکیت مادی و معنوی مربوط به این پایان نامه متعلق به دانشگاه صنعتی اصفهان و پدیدآورندگان است. این حقوق توسط دانشگاه صنعتی اصفهان و بر اساس خط مشی مالکیت فکری این دانشگاه، ارزشگذاری و سهم بندی خواهد شد.

هر گونه بهره برداری از محتوا، نتایج یا اقدام برای تجاری‌سازی دستاوردهای این پایان نامه تنها با مجوز کتبی دانشگاه صنعتی اصفهان امکان‌پذیر است.

تقدیم به

پدر و مادری که سایه‌شان پناهگاه و مهرشان بی‌دریغ.

فهرست مطالب

<u>صفحه</u>	<u>عنوان</u>
.....	فهرست مطالب
.....	فهرست شکل‌ها
.....	فهرست جداول‌ها
.....	چکیده
.....	مقدمه
.....	۱۰.۰ مروری بر فصل‌های پایان‌نامه
۵	۱: فیزیک ذرات نظری و تجربی
۶	۱.۱ مدل استاندارد فیزیک ذرات
۷	۱.۱.۱ ذرات بنیادی
۸	۲۰.۱ بوزون هیگر
۹	۳۰.۱ نیروهای بنیادی
۹	۲۰.۱ بروخورد دهنده بزرگ هادرولنی
۱۱	۳۰.۱ آزمایش CMS
۱۵	۱۰.۳.۱ انژری سنج الکترومغناطیسی
۱۶	۲۰.۳.۱ انژری سنج هادرولنی
۱۸	۳۰.۳.۱ سیستم میؤن
۱۹	۴۰.۱ بازسازی اشیاء
۱۹	۱۰.۴.۱ الگوریتم Particle Flow
۲۱	۲: درخشندگی و پایل‌آپ
۲۱	۱۰.۲ اندازه‌گیری درخشندگی در بروخورد دهنده ذرات
۲۳	۱۰.۱.۲ درخشندگی مطلق از پارامترهای ماشین
۲۴	۲۰.۱.۲ محاسبه درخشندگی با استفاده از سطح مقطع فیزیکی
۲۵	۳۰.۱.۲ محاسبه درخشندگی آنلاین
۲۶	۴۰.۱.۲ بروخورد دهنده بزرگ هادرولنی با درخشندگی بالا
۲۸	۲۰.۲ پایل‌آپ
۲۹	۱۰.۲.۲ چگالی پایل‌آپ مؤثر
۳۳	۳۰.۲ احتمالات پایل‌آپ و رویدادها در هر دسته عبوری
۳۵	۴۰.۲ احتمال پایل‌آپ برای کلاس‌های رویداد خاص
۳۵	۱۰.۴.۲ یک کلاس رویداد مشخص
۳۷	۲۰.۴.۲ دو کلاس رویداد مشخص
۳۹	۳۰.۴.۲ توزیع احتمال
۴۱	۵۰.۲ بروخورد
۴۳	۱۰.۵.۲ نرخ بروخورد

۴۴	۳ : شبیه‌سازی پایل آپ در CMS
۴۴	۱.۳ نحوه شبیه سازی پایل آپ
۴۷	۱.۱.۳ توزیع‌های پایل آپ اندازه‌گیری شده از داده
۴۷	۲۰.۱.۳ توزیع‌های پایل آپ اندازه‌گیری شده از شبیه‌سازی مونت کارلو
۴۹	۳۰.۱.۳ دسترسی به اطلاعات پایل آپ ذخیره شده در مجموعه داده‌ها
۴۹	۲.۰.۳ ذخیره خصوصیات ذرات
۴۹	۱۰.۲.۳ ابزارهای تحلیل پایل آپ
۵۰	۲۰.۲.۳ تجزیه و تحلیل اطلاعات پایل آپ در رویدادها با استفاده از EDAnalyzer
۵۳	۳۰.۲.۳ پردازش و ذخیره‌سازی متغیرهای پایل آپ
۵۸	۴۰.۲.۳ تعداد کاندیدای ذرات در یک رویداد
۷۴	۳.۰.۳ روش‌های تعیین شناسه نزدیک‌ترین رأس به هر ذره
۷۵	۱۰.۳.۳ جمع‌آوری داده‌های رأس
۷۶	۲۰.۳.۳ انواع روش‌های تعیین نزدیک‌ترین رأس
۷۹	۳۰.۳.۳ یافتن نزدیک‌ترین رأس
۸۲	۴۰.۳.۳ توزیع فاصله ذرات بر اساس نوع ذره
۸۷	۴ : شبکه‌های عصبی گرافی
۸۹	۱۰.۴ نمایش گرافی
۹۱	۲۰.۴ شبکه‌های عصبی
۹۶	۳۰.۴ انواع شبکه‌های عصبی گراف
۱۰۱	۵ : تخمین توزیع پایل آپ با استفاده از شبکه‌های عصبی گراف
۱۰۱	۱.۵ مرور داده‌ها و فرآیندهای آزمایشی
۱۰۲	۲.۰.۵ امکانات محاسباتی
۱۰۵	۳.۰.۵ در وهله اول اطلاعات رویدادها را می‌خوانیم
۱۰۹	۴.۰.۵ ساخت نمایش رویداد برخورد ذرات بصورت گراف
۱۱۲	۱۰.۴.۵ ارسال پیام
۱۱۳	۵.۰.۵ رویکردهای کاوش شده یادگیری ماشین
۱۱۳	۱۰.۵.۵ چارچوب مسئله وایازش و دسته‌بندی گراف
۱۱۳	۲۰.۵.۵ وظیفه وایازش و دسته‌بندی با MLP
۱۱۷	۶.۰.۵تابع هزینه
۱۱۷	۷.۰.۵ ارزیابی عملکرد
۱۱۸	۱۰.۷.۵ ارزیابی با L1Loss
۱۱۹	۲۰.۷.۵ ارزیابی عملکرد فیزیکی
۱۲۰	۸.۰.۵ فرایند آموزش مدل‌ها
۱۲۲	۹.۰.۵ توزیع خطاهای در پیش‌بینی‌های مدل
۱۲۴	۱۰.۰.۵ توزیع مقادیر پایل آپ، حقیقی و تخمینی
۱۲۹	۱۱.۰.۵ دقت مدل در تخمین تعداد پایل آپ در یک محدوده درخشنده
۱۳۲	۱۲.۰.۵ پراکندگی مقایسه خروجی مدل با داده‌های حقیقی

۱۳۲.....	۱۳۰.۵ نتیجه‌گیری
۱۳۶	پیوست‌ها
۱۳۶.....	۱. کد محاسبه یال‌ها در گراف‌ها
۱۳۸.....	۲. کد جمع‌آوری ویژگی‌های ذرات و ساخت گراف‌ها
۱۴۰.....	۳. دسته‌بندی وظایف
۱۴۳	مراجع

فهرست تصاویر

<u>صفحه</u>	<u>عنوان</u>
۸	شکل ۱.۱: مدل استاندارد فیزیک ذرات
۱۱	شکل ۲.۱: برخورد دهنده بزرگ هادرونی
۱۲	شکل ۳.۱: طرح آشکارساز CMS
۱۵	شکل ۴.۱: نمایی از قسمت مسیریاب
۱۶	شکل ۵.۱: انرژی سنج الکترومغناطیسی آشکارساز سیم لوله‌ی فشرده میؤنی
۱۹	شکل ۶.۱: نمایی از بخش‌های مختلف سیستم میؤن
۲۰	شکل ۷.۱: نمایی از سیستم میؤن
شکل ۱.۲: خلاصه‌ای از مجموع سطح مقطع تولید W^+ , W^- , W و Z به همراه کسرهای منشعب، نسبت‌های W به Z و W^+ به W^- و پیش‌بینی‌های نظری آن‌هاست.	
۲۳	شکل ۲.۲: تصویر عبور دسته‌ی پروتون‌ها با پارامترهای نسبی
۲۳	شکل ۳.۲: چپ: نمودار اسکن‌های VdM در صفحات x و y. راست: نمودار نرخ درخشنندگی سنج
۲۴	به عنوان تابع انحراف به دست آمده از اسکن VdM
۲۵	شکل ۴.۲: منطق پردازنده BCM1F سمت چپ: درخشنندگی آنلاین، عرض بازه ۶.۲۵ نانوثانیه (سومین بین داده‌های RHU). سمت راست: هیستوگرام‌های نرخ کانال چمغ شده، عرض بازه ۲۵ نانوثانیه (جمع ۴ بین RHU).
۲۸	شکل ۵.۲: طرح پایه LHC و HL-LHC برای دهه آینده و بعد از آن. عملکرد برخورد دهنده بزرگ هادرونی (LHC) یک دنباله از اجراهای عملیاتی ۴-۳ ساله و توقف‌های طولانی (LS) ۲ ساله است. پس از LS3 (۲۰۲۶-۲۰۲۴)، ماشین در حالت برخورد دهنده بزرگ هادرونی با درخشنندگی بالاخواهد بود. 4000 fb^{-1} هدف درخشنندگی یکپارچه نهایی است (مقدار طراحی HL-LHC 3000 fb^{-1} است). درخشنندگی طراحی شده، مقدار اصلی درخشنندگی است که LHC برای دستیابی به آن طراحی شده است.
۳۳	شکل ۶.۲: نمودار چگالی پایل آپ در IP5 یا IP1 در ابتدای پر شدن برای مبنای HL-LHC (خط قرمز) و تطابق گوسی آن (خط نقطه‌ای سیاه).
۴۲	شکل ۷.۲: $\langle N_t \rangle_{N_A \geq 1}$ برای ($\nu_t = 3.5\text{ cm}^{-2}\text{ s}^{-1}$) به عنوان یک تابع از ν_A برای حالت کلی و در تقریب برای ν_A کوچک.
۴۲	شکل ۸.۲: $\langle N_t \rangle_{N_A \geq 1}$ برای ($\nu_t = 0.188\text{ cm}^{-2}\text{ s}^{-1}$) به عنوان یک تابع از ν_A برای حالت کلی و در تقریب برای ν_A کوچک.
۴۷	شکل ۱.۳: توزیع تعداد برهmekنش‌ها در سال‌های مختلف
۴۸	شکل ۲.۳: شبیه سازی مونت کارلو برای داده‌های ۲۰۱۱
۴۸	شکل ۳.۳: شبیه سازی مونت کارلو برای داده‌های ۲۰۱۲
۵۴	شکل ۴.۳: توزیع تعداد رئوس بازسازی شده بر اساس معیارهای مختلف کیفیت رئوس.
۵۷	شکل ۵.۳: ارثبری کلاس recoCandidateReader از PackedCandidateReader

۵۸.....	شکل ۶.۳: توزیع تعداد کل ذرات.....
۵۹.....	شکل ۷.۳: توزیع تعداد هادرون های باردار.....
۵۹.....	شکل ۸.۳: توزیع تعداد هادرون های خنثی.....
۶۰.....	شکل ۹.۳: توزیع تعداد الکترون ها.....
۶۰.....	شکل ۱۰.۳: توزیع تعداد میون ها.....
۶۱.....	شکل ۱۱.۳: توزیع تعداد فوتون ها.....
۶۲.....	شکل ۱۲.۳: توزیع بار ذرات.....
۶۳.....	شکل ۱۳.۳: توزیع فی، تعداد ذرات شناسایی شده در هر مقدار فی را نشان می دهد.....
۶۴.....	شکل ۱۴.۳: توزیع تکانه در تصویر، تعداد ذرات شناسایی شده در هر مقدار تکانه را نشان می دهد.....
۶۵.....	شکل ۱۵.۳: نموداری از توزیع انرژی ذرات.....
۶۷.....	شکل ۱۶.۳: توزیع ذرات بر اساس شناسه گروه داده ذرات (PDGID).....
۶۹.....	شکل ۱۷.۳: Pt مولفه ای از تکانه ذره است که عمود بر جهت حرکت کلی آن است.....
۶۹.....	شکل ۱۸.۳: توریع پارامتر برخورد طولی ذرات.....
۷۰.....	شکل ۱۹.۳: توریع پارامتر برخورد عرضی ذرات.....
۷۳.....	شکل ۲۰.۳: توزیع ذرات را به عنوان تابعی از شبه تندی (η) در برخوردهای پروتون-پروتون (pp) در برخوردهای بزرگ هادرон (LHC) نشان می دهد. مقدار متوسط η برابر با صفر است.....
۷۳.....	شکل ۲۱.۳: در این شکل، مقادیر شبه تندی نمایش داده شده اند. در فیزیک ذرات، زاویه صفر معمولا هم جهت با محور پرتو است، بنابراین ذرات با مقادیر بالای شبه تندی معمولا از دست می روند و به همراه پرتو از فضای آشکارساز خارج می شوند. [۱۱].....
۷۳.....	شکل ۲۲.۳: شناسه رأس را برای هر ذره نشان می دهد. شناسه رأس خصوصیتی است که برای ردیابی محل پیدایش ذره در یک آزمایش برخورد ذرات استفاده می شود.....
۷۴.....	شکل ۲۳.۳: مقایسه توزیع های فاصله برای روش های مختلف تعیین نزدیکترین رأس.....
۷۸.....	شکل ۲۴.۳: نمودار نتایج به دست آمده از روش های مختلف را نشان می دهد. روش های مختلف شامل dzAsso-، NearVerPV، minVeryGoodVtx، minGoodVtx، minVtx و ciatedPV می باشد.....
۸۰.....	شکل ۲۵.۳: نمودارهای مربوط به توزیع شاخص رأس هایی که دارای بیشترین تکانه عرضی (pT) در هر رویداد هستند. در این نمودارها، سه روش مختلف شامل minVeryGoodVtxZIndex و minVtxZIndex مورد مقایسه قرار گرفته اند.....
۸۱.....	شکل ۲۶.۳: توزیع فاصله ذرات بر اساس نوع ذره.....
۸۲.....	شکل ۲۷.۳: توزیع فاصله انواع مختلف ذرات نسبت به رأس اولیه در راستای محور Z هر نمودار نشان دهنده نزدیکترین فاصله هر ذره نسبت به Vtx (رأس اولیه) است.....
۸۳.....	شکل ۲۸.۳: توزیع فاصله انواع مختلف ذرات نسبت به رأس GoodVtx هر نمودار نشان دهنده نزدیکترین فاصله هر ذره نسبت به GoodVtx است.....
۸۴.....	شکل ۲۹.۳: توزیع فاصله انواع مختلف ذرات نسبت به رأس VeryGoodVtx هر نمودار نشان دهنده نزدیکترین فاصله هر ذره نسبت به VeryGoodVtx است.....
۸۵.....	شکل ۳۰.۳: توزیع فاصله dz ذرات باردار مختلف نسبت به رأس اولیه در راستای محور Z محور افقی مقدار فاصله dz بر حسب سانتی متر و محور عمودی تعداد رویدادها را نشان می دهد.....
۸۶.....	

- شکل ۱.۴: شکل نشان می‌دهد که چگونه از شبیه‌سازی در آزمایش‌های فیزیک انرژی بالا برای یادگیری معکوس پیچیده استفاده می‌شود. در این روش، ابتدا فرایند فیزیکی با استفاده از شبیه‌سازی، بازسازی می‌شود تا یک «سوابق حقیقی» از رویداد فیزیکی که منجر به پاسخ خاصی در آشکارساز شده است، به دست آید. سپس از این «سوابق حقیقی» برای آموزش الگوریتم‌های یادگیری با نظارت استفاده می‌شود. این الگوریتم‌ها یاد می‌گیرند که شبیه‌سازی آشکارساز را معکوس کنند و از داده‌های اندازه‌گیری شده‌ی واقعی، ویژگی‌هایی در مورد فرایند فیزیکی زیربنایی استنتاج نمایند. در نهایت، این الگوریتم‌های آموزش‌دیده بر روی داده‌های واقعی که توسط آشکارساز اندازه‌گیری شده‌اند، به کار گرفته می‌شوند.....
۸۸
- شکل ۲.۴: (الف) خوشبندی برخوردهای آشکارساز ریاضی به مسیرها (ب) تقسیم سلول‌های کالری‌متر (پ) طبقه‌بندی رویدادها با انواع مختلف اشیاء فیزیکی (ت) طبقه‌بندی جت بر اساس ذرات مرتبط با جت
۸۹
- شکل ۳.۴: اجزای داخلی یک بلوک شبکه گراف توابع بروزرسانی و توابع تجمع هستند. شکل سمت چپ توابع بروزرسانی مجموعه‌ای از اشیاء با نمایش دارای اندازه‌ی ثابت را دریافت می‌کنند و همین تابع را روی هر یک از عناصر مجموعه اعمال می‌کنند که منجر به یک نمایش بروزرسانی شده (همچنین با اندازه‌ی ثابت) می‌شود. شکل سمت راست توابع تجمع مجموعه‌ای از اشیاء را دریافت می‌کنند و با استفاده از یک تابع ترتیب‌ناپذیر برای گروه‌بندی نمایش اشیاء (مانند مجموع روی عناصر)، یک نمایش با اندازه‌ی ثابت برای کل مجموعه ایجاد می‌کنند.....
۹۲
- شکل ۴.۴: یک بلوک شبکه گراف. یک گراف ورودی، $G = (u, V, E)$ ، پردازش می‌شود و یک گراف با همان ساختار لبه اما ویژگی‌های متفاوت، $G' = (u', V', E')$ ، به عنوان خروجی برگردانده شده‌اند.
۹۳
- شکل ۵.۴: بلوک‌های شبکه‌های گراف را می‌توان در معماری‌های محاسباتی پیچیده‌تر ترکیب کرد. ردیف بالا توالی‌ای از بلوک‌های شبکه‌های گراف مختلف را نشان می‌دهد که به صورت سری یا عمقی مرتب شده‌اند. ردیف پایین بلوک‌های شبکه‌های گراف متمایز را با یک پیکربندی مشترک و بازگشتی جایگزین می‌کند.
۹۴
- شکل ۶.۴: شبکه‌ی پیچشی گراف که نوعی شبکه‌ی عصبی با انتقال پیام است، می‌تواند به عنوان یک شبکه گراف، بدون یک ویژگی سراسری و یک تابع لبه‌ای خطی و غیرخطی، بیان شود.....
۹۵
- شکل ۷.۴: بازاریابی چشمگیرتر اجزای شبکه گرافی منجر به مدلی می‌شود که ویژگی‌های رأس (گره) را گردآوری می‌کند و آن‌ها را با یک ویژگی سراسری ترکیب می‌کند، سپس ویژگی‌های رأس را با استفاده از ویژگی ترکیبی به عنوان زمینه بروزرسانی می‌کند.....
۹۵
- شکل ۸.۴: شبکه گرافی پیچشی
۹۷
- شکل ۹.۴: شبکه گرافی توجه
۹۸
- شکل ۱۰.۴: شبکه گراف نمونه‌برداری و جمع‌آوری
۱۰۰
- شکل ۱.۵: اطلاعات ورتكس‌ها از جمله تعداد ورتكس‌ها و انواع مختلف آن‌ها
۱۰۵
- شکل ۲.۵: ویژگی‌های مختلف ذرات و نمودارهای مرتبط با توزیع هر ویژگی
۱۰۶
- شکل ۳.۵: ویژگی‌های مختلف ذرات و نمودارهای مرتبط با توزیع هر ویژگی
۱۰۷
- شکل ۴.۵: توزیع تعداد ذرات کاندیدا شامل فوتون‌ها، میون‌ها، الکترون‌ها، و هادرون‌های باردار و بدون بار در رویدادهای
۱۰۸

شکل ۵.۵: نمایش هایی از گراف تولید شده از داده‌ها. جایی که گره‌ها نماینده ذرات منفرد هستند و یال‌ها نشان‌دهنده ارتباطات بین آنها می‌باشند.	۱۱۱
شکل ۶.۵: معماری مدل PUGAT شکل ۷.۵: معماری مدل PUSage	۱۱۵ ۱۱۶
شکل ۸.۵: هر دو منحنی تابع هزینه PoissonNLLLOSS مجموعه آموزشی و اعتبارسنجی در تمام نمودارها به طور پیوسته کاهش می‌یابند، نشان می‌دهد که مدل به طور مؤثر از داده‌های آموزشی یاد می‌گیرد. اگر منحنی‌ها بعد از چندین دور به صورت افقی حرکت کنند یا افزایش یابند، ممکن است نشان‌دهنده بیشتر بازش یا ناکافی بودن داده‌های آموزشی باشد که آن قدری نیست وارد رژیم بیشتر بازش شویم.	۱۲۰
شکل ۹.۵: مقایسه تابع هزینه مدل‌ها	۱۲۱
شکل ۱۰.۵: ارزیابی اثربخشی مدل در برآورد پایل آپ: L1LOSS بر روی مجموعه داده‌ها به طور پیوسته کاهش می‌یابد، نشان می‌دهد که مدل در حال یادگیری انجام پیش‌بینی‌های دقیق‌تر بر روی داده‌های ناشناخته است. این یک نشانه مثبت است.	۱۲۲
شکل ۱۱.۵: مقایسه ی ماتریس عملکرد مدل‌ها	۱۲۳
شکل ۱۲.۵: شکل با قیماندها، بینش‌هایی درباره توزیع خطاهای مدل ارائه می‌دهد. واریانس خطاهای در سرتاسر محدوده پیش‌بینی‌ها ثابت نیست. با قیماندها تمایل دارند برای مقادیر بالاتر تعداد تعاملات بیشتر منفی باشند. این نشان می‌دهد که مدل تعداد تعاملات را در این موارد کم برآورد می‌کند.	۱۲۴
شکل ۱۳.۵: توزیع مقادیر پایل آپ، حقیقی و تخمینی توسط مدل را نمایش می‌دهند. نمودارها نشان می‌دهند که پیش‌بینی‌های مدل از نظر توزیع به مقادیر واقعی PU نزدیک هستند و این امر عملکرد کلی خوب مدل را در بدقت گرفتن توزیع زیربنایی نشان می‌دهد.	۱۲۵
شکل ۱۴.۵: هیستوگرام‌های متغیر پاسخ و مقادیر تخمینی. توزیع واقعی پایل آپ به صورت یکنواخت‌تر پخش شده است. توزیع تخمینی پایل آپ در مقادیر پایین‌تر پیک‌های بالاتری دارد و به سمت مقادیر بالاتر دم دارد. مدل تمایل به کم‌تخمین زدن مقادیر پایل آپ، بهویژه در مقادیر بالاتر دارد.	۱۲۶
شکل ۱۵.۵: نمودار تخمین چگالی کرنل KDE، همپوشانی بین توزیع‌های واقعی و تخمینی پایل آپ را به تصویر می‌کشد. دو توزیع باید با هم هماهنگ باشند، که نشان‌دهنده توافق خوب بین مقادیر واقعی و تخمینی است. مناطقی با چگالی بیشتر در نمودار نشان‌دهنده تعداد بیشتری از نقاط داده با مقادیر پایل آپ است.	۱۲۷
شکل ۱۶.۵: مقایسه توزیع خروجی مدل با توزیع پواسون . مورد انتظار. نمایانگر تعداد پایل آپ احتمال هر مقدار است. برای ارزیابی دقیقی از عملکرد مدل انجام می‌شود. منحنی آبی نمایانگر توزیع پواسون با میانگین 5° است، در حالی که نوارهای قرمز داده‌های خروجی مدل را نشان می‌دهند.	۱۲۸
شکل ۱۷.۵: مقایسه توزیع خروجی مدل‌ها با توزیع پواسون و تحلیل تفاوت میانگین و واریانس. نمودار سمت راست نمایانگر تعداد پایل آپ و احتمال هر مقدار است. در این نمودار، منحنی آبی نشان‌دهنده توزیع پواسون با میانگین 5° و نوارهای قرمز نشان‌دهنده داده‌های خروجی مدل‌ها است. نمودار سمت چپ تفاوت‌های میانگین و واریانس بین داده‌های مدل‌ها و پایل آپ را نشان می‌دهد که معیاری برای سنجش دقت مدل‌ها است.	

شکل ۱۸.۵ : توابع لگاریتم احتمال برای توزیع‌های پواسون مختلف. منحنی‌هایی با قله‌های بالاتر نشان‌دهنده تطابق بهتر بین توزیع تخمینی و توزیع پواسون واقعی هستند.....	۱۲۹
شکل ۱۹.۵ : بررسی دقت مدل در تخمین تعداد پایل آپ در یک محدوده درختنندگی.....	۱۳۰
شکل ۲۰.۵ : نتایج مقایسه‌ای مربوط به شب و نقاط برخورد برای چهار مدل مختلف را نمایش می‌دهد. این نمودار بهوضوح تفاوت عملکرد بین مدل‌ها را نشان می‌دهد.....	۱۳۱
شکل ۲۱.۵ : نمودار پراکندگی برای مقایسه خروجی مدل با داده‌های حقیقی. این نمودار‌ها مقایسه بصری از چگونگی پیش‌بینی تعداد پایل آپ توسط مدل‌های مختلف در مقایسه با مقادیر واقعی ارائه می‌دهند.....	۱۳۲
شکل ۱ : نمایی از دسته بندی وظایف.....	۱۴۱
شکل ۲ : کد PUSage	۱۴۲

فهرست جداول

صفحه	عنوان
۳۴.....	جدول ۱.۲ : $\sigma_{\text{elastic}} = ۳۰ \text{ mb}$ برای A کشسان با فرض $\mathcal{L}/k = ۳.۶$
۳۴.....	جدول ۲.۲ : $\nu_t = ۳.۵, \nu_A = ۰.۹۵, \frac{\nu_A}{\nu_t} = ۰.۲۷۳, ۱ - \frac{\nu_A}{\nu_t} = ۰.۷۲۷$ با $\nu_t = ۰.۱۸۸, \nu_A = ۰.۰۵۱, \frac{\nu_A}{\nu_t} = ۰.۲۷۳, ۱ - \frac{\nu_A}{\nu_t} = ۰.۷۲۷$
۳۶.....	$\mathcal{L}/k = ۳.۶ \times P(N_A, N_t)$ برای A کشسان با فرض $\sigma_{\text{elastic}} = ۳۰ \text{ mb}$ و $\nu_t = ۰.۱۸۸, \nu_A = ۰.۰۵۱, \frac{\nu_A}{\nu_t} = ۰.۲۷۳, ۱ - \frac{\nu_A}{\nu_t} = ۰.۷۲۷$ با $\nu_t = ۰.۲۲, \nu_A = ۰.۳۳, \frac{\nu_A}{\nu_t} = ۰.۹۵, ۱ - \frac{\nu_A}{\nu_t} = ۰.۰۵$
۳۶.....	جدول ۴.۲ : $\mathcal{L}/k = ۱.۹ \times P(N_A, N_t)$ برای A کشسان با فرض $\sigma_{\text{elastic}} = ۳۰ \text{ mb}$ و $\nu_t = ۰.۱۸۸, \nu_A = ۰.۰۵۱, \frac{\nu_A}{\nu_t} = ۰.۲۷۳, ۱ - \frac{\nu_A}{\nu_t} = ۰.۷۲۷$ با $\nu_t = ۰.۲۲, \nu_A = ۰.۳۳, \frac{\nu_A}{\nu_t} = ۰.۹۵, ۱ - \frac{\nu_A}{\nu_t} = ۰.۰۵$
۳۸.....	جدول ۵.۲ : $P(N_A, N_B, N_t)$ برای A کشسان، برای B غیرکشسان، (یعنی، $R = SD$) با فرض $\mathcal{L}/k = ۳.۶ \times ۱۰^{۲۹} \text{ cm}^{-۲} \text{ s}^{-۱}$ و $\sigma_{\text{elastic}} = ۳۰ \text{ mb}$ و $\sigma_{\text{NSD}} = ۷۰ \text{ mb}$ و $\nu_t = ۰.۲۲, \nu_A = ۰.۳۳, \frac{\nu_A}{\nu_t} = ۰.۹۵, ۱ - \frac{\nu_A}{\nu_t} = ۰.۰۵$
۳۸.....	جدول ۶.۲ : $P(N_A, N_B, N_t)$ برای A کشسان، برای B غیرکشسان، (یعنی، $R = SD$) با فرض $\mathcal{L}/k = ۱.۹ \times ۱۰^{۲۸} \text{ cm}^{-۲} \text{ s}^{-۱}$ و $\sigma_{\text{NSD}} = ۷۰ \text{ mb}$ و $\sigma_{\text{elastic}} = ۳۰ \text{ mb}$ و $\nu_t = ۰.۲۲, \nu_A = ۰.۳۳, \frac{\nu_A}{\nu_t} = ۰.۹۵, ۱ - \frac{\nu_A}{\nu_t} = ۰.۰۵$
۱۲۱.....	جدول ۱.۵ : PoissonNLLLoss
۱۲۲.....	جدول ۲.۵ : L1Loss

چکیده

آزمایش CMS در برخورده‌نده‌ی بزرگ هادرونی با افزایش سطح درخشندگی در بهروزرسانی‌های LHC، با چالش‌های زیادی در زمینه‌ی شبیه‌سازی و تحلیل برخوردهای همزمان پروتون-پروتون (پایل‌آپ) مواجه هستند. این پایان نامه رویکردی نوین برای تخمین پایل‌آپ را ارائه می‌دهد که از توان شبکه‌های عصبی گرافی استفاده می‌کند. همانطور که در کاربردهای اخیر برای بازسازی ذرات و کاهش پایل‌آپ نشان داده شده است، شبکه‌های عصبی گرافی به عنوان ابزاری قدرتمند برای نمایش داده‌های پیچیده‌ی رویدادها ظهرور کرده‌اند. ما روشی مبتنی بر شبکه‌های عصبی گرافی را برای تخمین تعداد برهم‌کنش‌های پایل‌آپ در هر رویداد پیشنهاد می‌کنیم. نتایج اولیه امیدوارکننده هستند و نشان می‌دهند که این روش را می‌توان برای بررسی پاسخ خطی لومینومترها، گامی کلیدی برای اندازه‌گیری مستقل درخشندگی، به کار برد. این پژوهش، پتانسیل متحولکننده‌ی شبکه‌های عصبی گرافی را برای تخمین پایل‌آپ و اندازه‌گیری درخشندگی در آزمایش‌های فیزیک انرژی بالا بررسی می‌کند. انتظار می‌رود این روش به عنوان یک ابزار ارزشمند به تکنیک‌های موجود اضافه شود و به پیشرفت‌های این حوزه کمک کند.

کلمات کلیدی

پایل‌آپ، درخشندگی، سطح مقطع، برخورد، شبکه‌های عصبی گرافی، توزیع پواسون، گراف

مقدمه

تحقیقات در زمینه فیزیک ذرات با انرژی بالا، به دلیل پیچیدگی و حجم عظیم داده‌های حاصل از آزمایش‌ها، در خط مقدم تجزیه و تحلیل داده‌ها قرار دارد. برای مثال، آزمایش CMS در LHC در سال‌های ۲۰۱۶ تا ۲۰۱۸ داده‌هایی با درخشندگی یکپارچه 150 fb^{-1} ثبت کرده است [۱]. این داده‌ها شامل سوابق برخورد پروتون-پروتون و اندازه‌گیری‌های میلیون‌ها کانال آشکارساز است که به اگرایت داده‌های قابل تجزیه و تحلیل برای هزاران فیزیکدان در سراسر جهان تبدیل می‌شوند. داده‌ها بر اساس الزامات انتخاب و مشاهدات فیزیکی خاص، مانند جرم ثابت یک سیستم فروپاشی، تجزیه و تحلیل می‌شوند. این روش‌ها، که به کشف‌های مهمی مانند بوزون‌های W و Z در دهه ۱۹۸۰ منجر شدند [۲]، همچنان نقش اساسی در توسعه مدل استاندارد فیزیک ذرات ایفا می‌کنند.

فیزیک ذرات به بررسی ویژگی‌های اصلی جهان ما در مقیاس زیراتومی می‌پردازد. با پیشرفت‌های صورت‌گرفته در سازمان تحقیقات هسته‌ای اروپا، به ویژه در برخورد دهنده بزرگ هادرونی، حجم بی‌سابقه‌ای از داده‌ها در طول برخورد ذرات تولید شده است. این داده‌ها با استفاده از تکنیک‌های یادگیری ماشین پردازش می‌شوند، و اکنون وارد عصری شده‌ایم که قدرت محاسباتی لازم برای تجزیه و تحلیل این حجم داده‌ها فراهم است. در دو دهه گذشته، علاوه بر پیچیده‌تر شدن سخت‌افزار آزمایش‌ها، تکنیک‌های تجزیه و تحلیل داده‌ها نیز به‌طور قابل توجهی تکامل یافته‌اند. روش‌های چندمتغیره به‌طور فرایندهای در تجزیه و تحلیل داده‌ها گنجانده شده‌اند، که از نتایج آن می‌توان به کشف بوزون هیگز در سال ۲۰۱۲ اشاره کرد. این کشف عمده‌تاً با استفاده از تکنیک‌های چندمتغیره و تقویت درختان تصمیم‌گیری انجام شد.

پیشرفت‌های اخیر در حوزه یادگیری ماشین این امکان را فراهم کرده است که ماشین‌ها بتوانند از داده‌ها بیاموزند و به‌طور مستقل تصمیم‌گیری کنند. تکنیک‌های یادگیری ماشین، که بسیاری از آن‌ها منبع باز هستند، اکنون می‌توانند به‌طور عملی در مسائل مربوط به فیزیک ذرات مورد استفاده قرار گیرند.

برخورد دهنده بزرگ هادرونی با دسته‌هایی از پروتون‌ها عمل می‌کند که هر کدام شامل $O(10^{11})$ پروتون است. این پروتون‌ها

به طور فشرده بسته شده‌اند تا درخشنندگی پرتو و در نتیجه نرخ برخورد افزایش یابد. با این حال، این افزایش درخشنندگی منجر به افزایش تعداد برخوردهای همزمان (پایل آپ) نیز می‌شود که اغلب شامل رویدادهای نرم با تکانه عرضی پایین است.

به دلیل شکل منطقه درخشنندگی، مختصات (x, y) برخوردهای پایل آپ در یک رویداد به برخوردهای مهم با تکانه عرضی بالا هم‌تراز شده‌اند که به آن‌ها سرنخ اصلی می‌گویند. از سوی دیگر، گرهای پایل آپ ممکن است تغییر مکان داده و تا چند سانتی‌متر در جهت \pm جابه‌جا شوند. اثرات پایل آپ با استفاده از الگوریتم‌های هدفمند که برای تخمین و کاهش این اثرات طراحی شده‌اند، کاهش می‌یابد.

حقیقان تکنیک‌های مختلفی را برای رسیدگی به پایل آپ توسعه داده‌اند. یکی از این تکنیک‌ها الگوریتم¹ PUPPI است که از سال ۲۰۱۵ به طور گسترده‌ای در تجزیه و تحلیل داده‌های LHC مورد استفاده قرار می‌گیرد. این الگوریتم به ذرات فردی بر اساس احتمال سیگنال یا انباشته‌شدنی آن‌ها وزن اختصاص می‌دهد. با این حال، هنوز جا برای بهبود عملکرد آن وجود دارد. شبکه‌های عصبی گرافی، یک کلاس از مدل‌های یادگیری عمیق هستند که برای پردازش داده‌های ساختاری گراف طراحی شده‌اند و در گرفتن روابط پیچیده بین گره‌ها (ذرات) در یک گراف برتقی دارند. در مقاله‌ای با عنوان "کاهش پایل آپ در برخورد دهنده هادرونی بزرگ با شبکه‌های عصبی گراف" یک رویکرد جدید معرفی می‌شود. نویسندهان در این مقاله، یک طبقه‌بندی‌کننده را بر اساس شبکه عصبی گرافی آموزش می‌دهند تا ذرات ناشی از برخوردهای با حرکت عرضی بالا را شناسایی و حفظ کنند، در حالی که ذرات ناشی از تعاملات پایل آپ را رد می‌کنند. آن‌ها با گسترش اطلاعات ورودی و استفاده از قابلیت‌های یادگیری شبکه عصبی گرافی، عملکرد بهتری را در مقایسه با راه حل‌های موجود مانند PUPPI نشان داده‌اند.

این تحقیق می‌تواند تأثیر مستقیمی بر دقت تجزیه و تحلیل داده‌های LHC داشته باشد. بهبود در تخمین پایل آپ نه تنها به اندازه‌گیری دقیق‌تر فرایندهای فیزیکی مانند کشف ذرات جدید یا مطالعه پدیده‌های نادر کمک می‌کند، بلکه به عنوان یک پارامتر ورودی در شبیه‌سازی‌ها نقش مهمی در اندازه‌گیری بازده آشکارساز نیز ایفا می‌کند. استفاده از شبکه‌های عصبی گرافی برای تخمین پایل آپ ما را قادر می‌سازد تا به اندازه‌گیری مستقل درخشنندگی برسیم. علاوه بر این، کاربرد شبکه‌های عصبی گرافی در فیزیک ذرات، امکان کشف راه حل‌های جدید در بازسازی و طبقه‌بندی رویدادها را فراهم می‌آورد. در نهایت، پیشرفت در برآورد پایل آپ به درک بهتر از ذرات بنیادی و قوانین حاکم بر جهان کمک خواهد کرد.

۱۰. مروری بر فصل‌های پایان‌نامه

این پایان‌نامه شامل پنج فصل اصلی است که به بررسی جامع روش‌های تخمین پایل آپ در فیزیک ذرات و کاربرد شبکه‌های عصبی گراف متمرکز در این زمینه می‌پردازد.

فصل ۱: به معرفی کلی فیزیک ذرات، اهمیت بررسی ذرات بنیادی و نیروهای بین آن‌ها، و تأثیر این مطالعات بر فهم بهتر ساختار و تکامل کیهان پرداخته می‌شود. همچنین، برخورده‌نده بزرگ هادرونی و نقش آزمایش CMS در این تحقیقات معرفی می‌شوند. در نهایت، مسئله پایل آپ به عنوان یکی از چالش‌های اصلی در تحلیل داده‌های حاصل از LHC بررسی می‌شود. این فصل مبانی نظری فیزیک ذرات از جمله مدل استاندارد ذرات بنیادی، نیروهای اساسی طبیعت، و نحوه تعامل ذرات را توضیح می‌دهد. همچنین، برخوردهای ذرات و شناسایی آن‌ها در آزمایش‌های فیزیک ذرات و ساختار و عملکرد آشکارسازهای CMS بررسی می‌شود.

فصل ۲: با تعریف درخشنندگی و پایل آپ آغاز می‌شود. در این فصل، روش‌های مختلف اندازه‌گیری این دو کمیت بررسی و مفاهیم مربوط به آن‌ها گنجانده می‌شود. این فصل اهمیت پایل آپ را به عنوان یک پارامتر کلیدی در شبیه‌سازی‌ها و تأثیر آن بر

¹ PileUp Per Particle Identification

اندازهگیری‌های دقیق فیزیکی بررسی می‌کند.

فصل ۳: به روش‌های شبیه‌سازی و تحلیل پدیده پایل‌آپ در آزمایش CMS اختصاص دارد. در این فصل، روش‌های سنتی و مدرن برای تخمین توزیع پایل‌آپ بررسی می‌شوند. همچنین، مزايا و معایب هر یک از این روش‌ها مورد بررسی قرار می‌گیرد. هدف این فصل بررسی دقیق تکنیک‌های مختلف برای بهبود تخمین پایل‌آپ و تأثیر آن بر تحلیل داده‌های آزمایش CMS است.

فصل ۴: به معرفی و بررسی استفاده از شبکه‌های عصبی گراف مرکز در تحلیل داده‌های حاصل از برخوردهای ذرات می‌پردازد. نحوه عملکرد شبکه‌های عصبی گراف، ساختار گرافی داده‌ها و چالش‌های مرتبط با آن‌ها در این فصل توضیح داده می‌شود. این فصل به بررسی عمیق کاربرد شبکه‌های عصبی گراف در فیزیک ذرات و پتانسیل آن‌ها برای تخمین دقیق‌تر پایل‌آپ می‌پردازد.

فصل ۵: عملکرد توزیع پایل‌آپ ارزیابی می‌شود و نتایج حاصل با روش‌های سنتی مقایسه می‌گردد. در این فصل، تأثیر استفاده از شبکه‌های عصبی گراف بر بهبود دقت و کارایی تحلیل داده‌ها مورد بررسی قرار می‌گیرد. همچنین، نتایج پژوهش و پیشنهادهایی برای کارهای آینده در این زمینه ارائه می‌شود. این فصل نگاهی جامع به کارایی شبکه‌های عصبی گراف در تخمین پایل‌آپ و پیشرفت‌های آینده در این حوزه دارد.

در مجموع، این پایان‌نامه به بررسی عمیق روش‌های موجود برای تخمین پایل‌آپ و بررسی استفاده از تکنیک‌های نوین مانند شبکه‌های عصبی گراف برای بهبود دقت و کارایی در فیزیک ذرات می‌پردازد و نشان می‌دهد که چگونه این رویکردها می‌توانند منجر به پیشرفت‌های قابل توجه در تحلیل داده‌های حاصل از برخوردهای ذرات شوند.

فصل اول

فیزیک ذرات نظری و تجربی

فیزیک انرژی بالا^۱ که به عنوان فیزیک ذرات نیز شناخته می‌شود، شاخه‌ای از علم فیزیک است که به مطالعه اجزای بنیادی ماده و نیروهایی که تعاملات آن‌ها را کنترل می‌کنند، می‌پردازد. این حوزه به بررسی ذرات بنیادی که سازنده‌های جهان هستند و همچنین نیروهای بنیادی که این ذرات از طریق آن‌ها با یکدیگر برهمنکش می‌کنند، می‌پردازد. مدل استاندارد فیزیک ذرات، چارچوب نظری فعلی است که این ذرات و نیروها را به جزگرانش توصیف می‌کند. این مدل شامل کوارک‌ها، لپتون‌ها و ذرات حامل نیرو مانند فوتون‌ها، بوزون‌های W و Z ، گلوئون‌ها و بوزون هیگز است.

فیزیکدانان در فیزیک انرژی بالا از روش‌های تجربی و نظری برای بررسی این ذرات و نیروها استفاده می‌کنند. شتاب‌دهنده‌های ذرات، مانند شتاب‌دهنده بزرگ هادرонی^۲ در سرن^۳ برای برخورد دادن ذرات با انرژی بالا استفاده می‌شوند، این شتاب‌دهنده شرایطی را بازسازی می‌کند که شبیه به شرایط لحظاتی پس از انفجار بزرگ است. و این امکان را به فیزیکدانان می‌دهد که خصوصیات ذرات را بررسی کرده و به دنبال پدیده‌های جدید باشد. فیزیکدانان نظری این آزمایش‌ها را با توسعه مدل‌ها برای توضیح داده‌های مشاهده شده و پیش‌بینی اثرات جدید تکمیل می‌کنند. آنها از چارچوب‌های ریاضی مانند نظریه میدان کوانتومی^۴ برای توصیف

¹High Energy Physics

²Large Hadron Collider

³ CERN

⁴Quantum Field Theory

تعامالت بین ذرات استفاده می‌کنند. با وجود موقعيت‌های مدل استاندارد، نواقصی هم در آن دیده می‌شود، مثلاً شامل گرانش نمی‌شود که توسط نظریه نسبیت عام توصیف می‌شود، و پدیده‌های ماده تاریک یا انرژی تاریک را نیز توضیح نمی‌دهد که بیشترین محتوای جرمی-انرژی جهان را تشکیل می‌دهند. نظریه پردازان انرژی بالا در حال کار بر روی نظریه‌هایی برای حل این اسرار هستند، مانند ابرتقارن و نظریه ریسمان. فیزیک انرژی بالا یک زمینه علمی بسیار همکاری محور است که شامل تیم‌های بین‌المللی بزرگی می‌شود که در آزمایش‌ها و توسعه‌های نظری مشارکت دارند. این حوزه تأثیرات عمیقی بر پیشرفت‌های علمی و فناوری دارد و الهام‌بخش نسل‌های آینده دانشمندان است. به طور خلاصه، فیزیک انرژی بالا در خط مقدم تحقیقاتی برای درک ساختار نهایی ماده و جهان قرار دارد. این حوزه حدود دانش بشری را گسترش داده و به دنبال کشف‌های جدید از ذرات زیراتمی تا حاشیه‌های دوردست فضا است.

در سرن، فیزیک ذرات محور اصلی تحقیقات است، اما برنامه‌های این مرکز شامل زمینه‌های گستردتری از جمله فیزیک هسته‌ای و مطالعات در زمینه‌های مختلف مانند ضدماده و پرتوهای کیهانی نیز می‌شود [۳]. از دهه‌ی ۱۹۷۰، فیزیک‌دانان در ذرات ساختار اساسی ماده را با استفاده از مجموعه‌ای شکیل از معادلات به نام مدل استاندارد شرح می‌دهند. فیزیک‌دانان در سرن از قوی‌ترین شتاب‌دهنده‌های ذرات و آشکارسازهای جهان استفاده می‌کنند تا پیش‌بینی‌ها و محدودیت‌های مدل استاندارد را آزمایش کنند. در طول سال‌ها این مدل بسیاری از نتایج آزمایشی را توضیح داده و پیش‌بینی دقیقی از طیفی از پدیده‌ها داده است، به طوری که امروزه به عنوان یک نظریه‌ی فیزیکی آزمایش شده محسوب می‌شود. اما این مدل تنها درباره‌ی ۴٪ از جهان شناخته‌شده صحبت می‌کند و سوالات باقی‌مانده‌اند. آیا ما در انرژی‌های بالای شتاب‌دهنده هادرون بزرگ، نیروها را یکپارچه می‌بینیم؟ چرا گرانش انقدر ضعیف است؟ چرا در جهان ماده بیشتر از پادماده است؟ آیا در انتظار کشف فیزیک جدید، شامل تئوری‌ها، پدیده‌ها یا ذراتی که غیرمعمول یا به طور معمول در تجارب روزمره مشاهده نمی‌شوند در انرژی‌های بالاتر است؟ آیا ما شواهدی برای نظریه‌ی ابرتقارن در شتاب‌دهنده هادرون بزرگ خواهیم یافت؟ یا فهمی از بوزون هیگر که به ذرات جرم می‌دهد؟ فیزیک‌دانان در سرن به دنبال پاسخ به این سوالات و بیشتر هستند [۳].

۱۰.۱ مدل استاندارد فیزیک ذرات

مدل استاندارد فیزیک ذرات یک چارچوب است که به ما اجازه می‌دهد تا ساختار اصلی و دینامیک‌های جهان را به صورت ذرات بنیادی درک کنیم، جایی که همه تعاملات بین ذرات بنیادی به وسیله برهمنش ذرات تسهیل می‌شود. به طور خلاصه، بر اساس درک فعلی ما، هنگامی که وارد حوزه زیراتمی می‌شویم کل جهان ما از یک آرایه بسیار پراکنده از ذرات بنیادی تشکیل شده است [۴]. ظهور انرژی پایین میدان الکترومغناطیس^۱ (تئوری میدان کوانتمی نیروی الکترومغناطیس) اتم‌ها را به عنوان حالت کوانتمی ذره‌ای که تحت یک پتانسیل قرار دارد با الکترون‌های با بار منفی (e^-) که دور هسته با بار مثبت (p) و نوترون‌های با

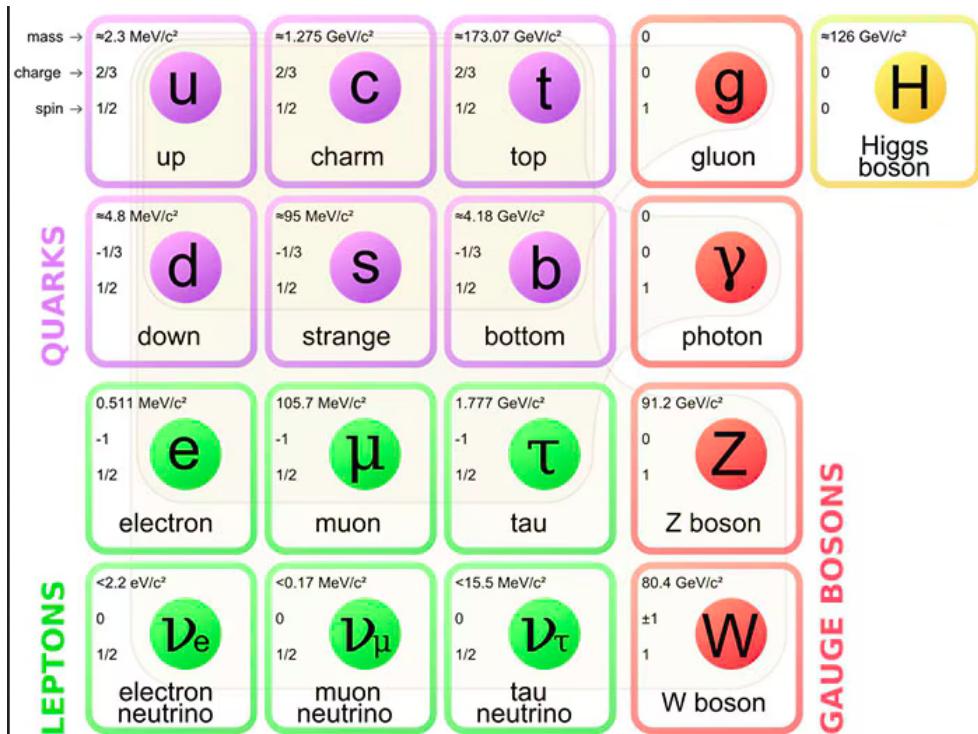
^۱QED (Quantum Electrodynamics)

بار الکتریکی خنثی (n) در نظر می‌گیرد و جاذبه الکترواستاتیکی این بارهای الکتریکی مخالف را درگیر میکند. کرومودینامیک کوانتمی^۱ تئوری اساسی برهم کش قوی است که پروتون‌ها و نوترون‌ها را در داخل هسته اتم به هم می‌چسباند. نوتريونی الکترونی تقریباً بی‌جرم در این فرآیندها تولید می‌شود. تقریباً همه پدیده‌های فیزیکی که در شرایط عادی رخ می‌دهند، توسط نیروهای الکترومناطقیس، قوی و ضعیف، گرانش (که بسیار ضعیف است اما ساختار بزرگ مقیاس جهان را توجیه می‌کند)، و فقط چهار ذره: الکترون، پروتون، نوترون و نوتريون توضیح داده می‌شود.

۱۰.۱ ذرات بنیادی

تمام ماده به جز ماده‌ی تاریک از مولکول‌ها ساخته شده است که خودشان از اتم‌ها تشکیل شده‌اند. درون اتم‌ها، الکترون‌ها به دور هسته می‌چرخند. هسته خود به طور کلی از پروتون‌ها و نوترون‌ها تشکیل شده است اما حتی این‌ها ذراتی هستند که از ذراتی دیگر ساخته شده‌اند. درون پروتون‌ها و نوترون‌ها، ما کوارک‌ها را پیدا می‌کنیم، اما به نظر می‌رسد این‌ها از ذرات کوچکتری ساخته نشده‌اند، به طور مشابه با الکترون‌ها. کوارک‌ها و الکترون‌ها برخی از ذرات بنیادی‌ای هستند که ما در سر و در دیگر آزمایشگاه‌ها مطالعه می‌کنیم. اما فیزیکدانان بیشتر این ذرات بنیادی را در آزمایش‌های مختلف پیدا کرده‌اند، به اندازه‌ای که پژوهشگران نیاز داشتند تا آنها را مرتب کنند، همانند کاری که مندلیف با جدول تناوبی خود انجام داد. که این در یک مدل نظری مختصر به نام مدل استاندارد خلاصه می‌شود. امروزه، ما ایده‌ی خیلی خوبی از آنچه ماده از چه چیزهایی تشکیل شده است، چگونه همه چیز با هم نگه داشته می‌شود و چگونه این ذرات با یکدیگر تعامل دارند داریم. در مقیاس‌های انرژی بالاتر، مانند آنچه که در آزمایش‌های انجام شده با استفاده از LHC (جایی که انرژی برخورد پروتون‌ها حدود ۱۳.۵ تراالکترون ولت است) بدست آمده است، پروتون‌ها و نوترون‌ها به عنوان حالت‌های محدودی از ذرات بنیادی به نام کوارک‌ها در نظر گرفته می‌شوند، به این صورت: پروتون‌ها از دو کوارک بالا و یک کوارک پایین (uud) تشکیل شده‌اند، در حالی که نوترون‌ها از دو کوارک پایین و یک کوارک بالا (ddu) تشکیل شده‌اند. شکل ۱.۱، مجموعه‌ای از ذرات (فرمیون‌ها) را نشان می‌دهد. مدل استاندارد فرمیون‌ها را به سه خانواده یا نسل تقسیم می‌کند. هر نسل شامل یک کوارک نوع بالا (با بار $-1/3$) و یک کوارک نوع پایین (با بار $+2/3$)، همچنین یک لپتون باردار (با بار -1) و یک لپتون بدون بار است. معادله دیراک وضعیت هر یک از دوازده فرمیون بنیادی را توصیف می‌کند و نشان می‌دهد که برای هر فرمیون، یک پادذره وجود دارد که دارای همان جرم اما بار مخالف است، که با یک خط افقی بالای نماد ذره نشان داده می‌شود.

^۱QCD (Quantum Chromodynamics)



شکل ۱۰.۱: مدل استاندارد فیزیک ذرات

هر چند که پذیرفته شده است که نوترینوها بی جرم نیستند، اما جرم آنها به گونه‌ای کوچک است که تاکنون به دقت مشخص نشده‌اند. تعاملات بین ذرات توسط چهار نیروی بنیادی تسهیل می‌شود، اما اثر گرانش در این مقیاس به اندازه‌ای کوچک است که می‌توان آن را بدون از دست دادن دقت نادیده گرفت. تمامی ذرات در تعاملات ضعیف شرکت می‌کنند و بنابراین به نیروی ضعیف وابسته هستند. نوترینوها همگی بدون بار الکتریکی هستند و به همین دلیل در تعاملات الکترومغناطیسی شرکت نمی‌کنند و می‌توان گفت که نسبت به این نیرو نامرئی هستند. کوارک‌ها با رنگی خوانده شده توسط کوانتم کرومودینامیک دارند و بنابراین تنها ذراتی هستند که احساس نیروی قوی می‌کنند. نیروی قوی کوارک‌ها را به وضعیت‌های محدود در داخل هادرон‌ها محدود می‌کند؛ بنابراین، کوارک‌ها به طور معمول زیر شرایط عادی به آزادی دیده نمی‌شوند.

۲۰.۱ بوزون هیگز

بوزون هیگز که وجود آن توسط همکاری‌های CMS و ATLAS در سرн در سال ۲۰۱۲ تأیید شد، با سایر ذرات مدل استاندارد تقاضت دارد، به دلیل اینکه یک ذره اسکالر است که سایر ذرات مدل استاندارد را با جرم مجهر می‌کند، یک خاصیت که بدون آن، تمام ذرات به طور مداوم با سرعت نور، c ، حرکت خواهند کرد. به تنهایی، تمام ذرات بی جرم هستند، اما با تعامل با میدان هیگز، که همیشه غیر صفر است، مکانیزم هیگز آنها را با جرم‌های متمایز مجهز می‌کند.

۳۰.۱ نیروهای بنیادی

از لحاظ تاریخی، نیوتن اظهار کرد که ماده می‌تواند با هر ماده دیگری بدون واسطه تماس مستقیم تعامل کند و به همین ترتیب، الکترومغناطیس کلاسیک، تعامل الکترواستاتیک بین ذرات را با استفاده از میدان‌ها توضیح می‌دهد. نظریه میدان کوانتوم از این توضیحات غیرمادی دوری می‌کند و شامل شرح هر یک از نیروهای بنیادی می‌شود. الکترومغناطیس توسط تئوری الکترودینامیک کوانتومی^۱ توضیح داده می‌شود، نیروی قوی توسط کرومودینامیک کوانتومی^۲، نیروی ضعیف توسط تئوری الکتروضعیف^۳. گرانش توسط مدل استاندارد توضیح داده نمی‌شود؛ بنابراین، نظریه عمومی نسبیت اینشتین هنوز بهترین توضیح برای این نیرو است. جستجو برای ادغام گرانش، مدل استاندارد، و همچنین پدیده‌های فیزیکی ناشناخته دیگر مانند طبیعت ماده تاریک و انرژی تاریک به یک نظریه وحدت بزرگ ادامه دارد و به پژوهش‌های جدید نظری مثل تئوری ریسمان و گرانش حلقه‌ای منجر شده است؛ این جریان‌های پژوهشی به طور کلی به عنوان فیزیک فراتر از مدل استاندارد^۴ شناخته می‌شوند. با نگاهی به الکترومغناطیس، تعامل بین ذرات باردار از طریق مبادله فوتون‌های مجازی رخ می‌دهد. به طور مشابه، برای هر دو نیروی قوی (۸ نوع گلوبن، که بی‌جرم هستند، همراه با کوارک‌ها) و نیروی ضعیف (بوزون‌های W^+ و W^- که حدود 80° برابر سنگین‌تر از پروتون هستند؛ و بوزون Z ، که تعامل جریان ضعیف نوتر دارد) ذرات مجازی وجود دارند. بوزون‌های مجازی همه دارای اسپین ۱ هستند، در مقابل فرمیون‌ها که همه اسپین $1/2$ دارند.

۲۰.۱ برخورد دهنده‌ی بزرگ هادرولی

در پایان سال ۱۹۵۱، تصمیمی به تأسیس شورای اروپایی برای تحقیقات هسته‌ای سرن^۵ در یک جلسه بین‌المللی سازمان یونسکو در پاریس اتخاذ شد. پیش‌نویس نهایی کمیته سرن در سال ۱۹۵۳ توسط دوازده کشور امضا شد. اکنون، سرن یک سازمان واقعاً بین‌المللی است، با ۲۴ کشور عضو (بعضی از آن‌ها غیراروپایی)، که در تصمیم‌گیری‌های اصلی مشارکت دارند؛ چند کشور با وضعیت عضو همکار یا وضعیت ناظر؛ و کشورهای غیرعضو با توافقنامه‌های همکاری، از جمله آفریقای جنوبی مشارکت می‌کنند. مأموریت تحقیقاتی سرن در ارائه پاسخ به سوالات بنیادین درباره ساختار و تکامل کیهان، و همچنین منشأ آن، متمرکز است؛ این سازمان این اهداف را با ارائه دسترسی به تسهیلات شتاب‌دهنده ذرات و منابع محاسباتی خود به پژوهشگران بین‌المللی، که تحقیقاتی انجام می‌دهند که پیش روی دانش انسان را تسهیل می‌کند، به نفع بشریت تحقق می‌دهد. بنابراین، سرن به عنوان یک سازمان ای که جدای از سیاست عمل می‌کند، تأکید دارد که استدلال‌ها و تصمیم‌گیری‌ها باید بر پایه‌ی شواهد و اطلاعات موجود باشند. علاوه بر این، سرن بر تسهیل انتقال دانش از تحقیقات بنیادین به صنایع مختلف و توسعه نسل‌های

¹QED (Quantum Electrodynamics)

²QCD (Quantum Chromodynamics)

³EWT(Electroweak Theory)

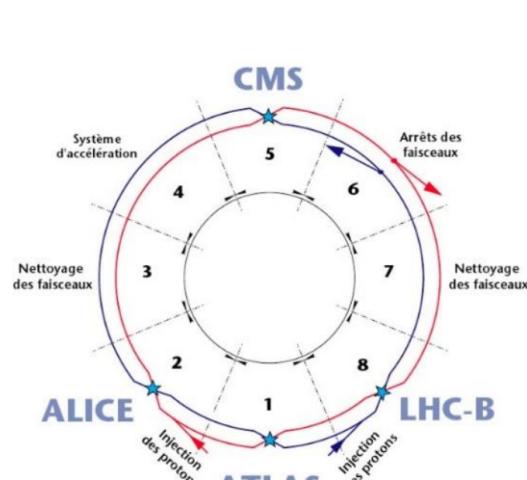
⁴BSM (Beyond the Standard Model)

⁵Conseil Européen pour la Recherche Nucléaire

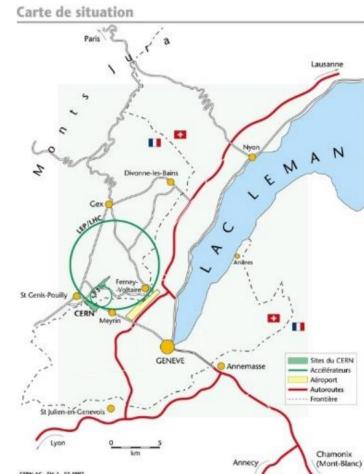
آینده‌ی دانشمندان و مهندسان تأکید دارد.

شتاب‌دهنده بزرگ هادرونی بزرگترین و قدرتمندترین شتاب‌دهنده ذرات در جهان است. این شتاب‌دهنده شامل یک حلقه ۲۷ کیلومتری از آهنرباهای ابرسانا است که دارای تعدادی ساختار شتاب‌دهنده برای افزایش انرژی ذرات در مسیر حرکتشان است. این شتاب‌دهنده برای اولین بار در ۱۵ سپتامبر ۲۰۰۸ راهاندازی شد و همچنان جدیدترین افروده به مجموعه شتاب‌دهنده‌های سرن است. LHC شامل یک حلقه ۲۷ کیلومتری از آهنرباهای ابرسانا است که دارای تعدادی ساختار شتاب‌دهنده برای افزایش انرژی ذرات در مسیر حرکتشان است. درون شتاب‌دهنده، دو پرتوی ذره با انرژی بالا تقریباً با سرعت نور حرکت می‌کنند قبل از این که به هم برخورد کنند. پرتوها در جهت‌های مخالف در لوله‌های پرتویی جداگانه حرکت می‌کنند - دو لوله که در خلاء بسیار بالا نگه داشته می‌شوند. آنها توسط میدان مغناطیسی قوی که توسط آهنرباهای الکترومغناطیسی ابرسانا حفظ می‌شود، در حلقه شتاب‌دهنده هدایت می‌شوند. آهنرباهای الکترومغناطیسی از سیم‌پیچ‌های مخصوصی ساخته شده‌اند که در حالت ابرسانا عمل می‌کنند و برق را بدون مقاومت یا از دست دادن انرژی به طور کارآمد هدایت می‌کنند. این نیاز به خنک کردن آهنرباهای به دمای ۳۰.۲۷۱ درجه سانتیگراد دارد - دمایی سردتر از فضای بیرونی. به همین دلیل، بخش زیادی از شتاب‌دهنده به سیستم توزیع هلیوم مایع که آهنرباهای را خنک می‌کند و همچنین به سایر خدمات تامین متصل است. هزاران آهنربای مختلف با اندازه‌ها و انواع مختلف برای هدایت پرتوها در اطراف شتاب‌دهنده استفاده می‌شوند. این‌ها شامل ۱۲۳۲ آهنربای دوقطبی به طول ۱۵ متر که پرتوها را خم می‌کنند و ۳۹۲ آهنربای چهارقطبی هر کدام به طول ۵ تا ۷ متر که پرتوها را متمرکز می‌کنند، هستند. درست قبل از برخورد، نوع دیگری از آهنربا برای "فسردن" ذرات به هم نزدیکتر برای افزایش احتمال برخورد استفاده می‌شود. ذرات آنقدر کوچک هستند که وظیفه برخورد دادن آنها شبیه شلیک دو سوزن از فاصله ۱۰ کیلومتری با دقیقی که در میانه راه به هم برستند، است.

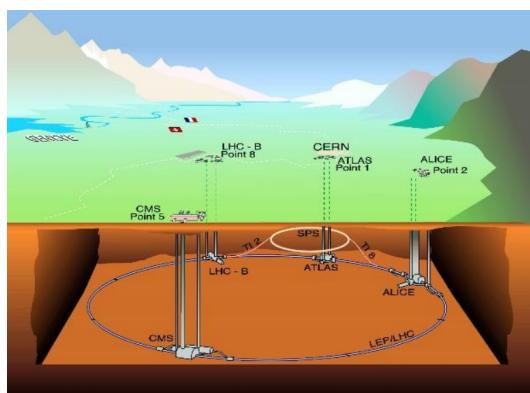
تمام کنترل‌های شتاب‌دهنده، خدمات آن و زیرساخت‌های فنی آن در یک ساختمان در مرکز کنترل سرن قرار دارد. از اینجا، پرتوها در داخل LHC در چهار نقطه اطراف حلقه شتاب‌دهنده به هم برخورد می‌کنند که با موقعیت‌های چهار آشکارساز ذرات - ATLAS، CMS، ALICE و LHCb - مطابقت دارند.



(ب) ۴ نقطه برخورد و آزمایش های سرن



(الف) بروتونهای بزرگ هادرونی در مرز فرانسه و سوئیس



(ج) بُرخورد دهنده بزرگ هادرولوئی در عمق $5-75$ متری زیر زمین

شکل ۲.۱: بی خورد دهنده بزرگ هادرone

۳.۱ آزمایش CMS

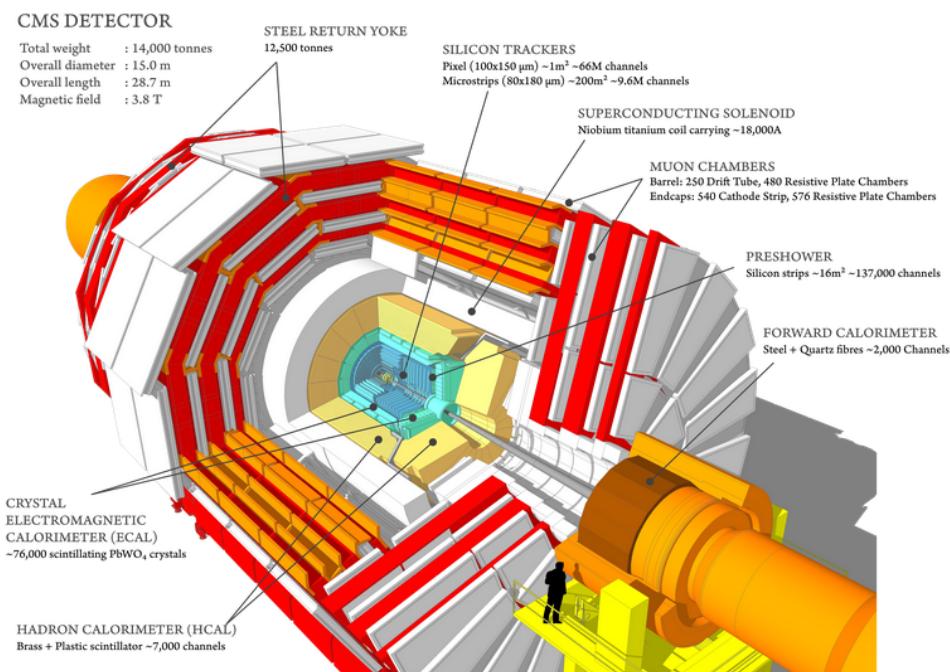
شتابدهنده بزرگ هادرونی پروتون‌ها را تقریباً تا سرعت نور شتاب می‌دهد در جهت‌های ساعت‌گرد و پادساعت‌گرد و سپس آنها را در چهار نقطه اطراف حلقه خود به هم برخورد می‌دهد. در این نقاط، انرژی برخورد ذرات به جرم تبدیل می‌شود و ذرات به همه جهات پراکنده می‌شوند. آشکارساز سیم لوله فشرده میؤنی (CMS) در یکی از این چهار نقطه برخورد قرار دارد. این یک آشکارساز چندمنظوره است؛ به این معنی که برای مشاهده هر پدیده جدید فیزیکی که ممکن است LHC آشکارکند، طراحی شده است. CMS مانند یک دوربین غولپیکر و پرسرعت عمل می‌کند و تا ۴۰ میلیون بار در هر ثانیه از برخورد ذرات از همه جهات عکس‌های سه‌بعدی می‌گیرد. اگرچه اکثر ذرات تولید شده در برخوردها "ناپایدار" هستند، اما به سرعت به ذرات پایدار تبدیل می‌شوند که توسط CMS قابل شناسایی هستند. با شناسایی (تقریباً) همه ذرات پایدار تولید شده در هر برخورد، اندازه‌گیری تکانه و انرژی آنها و سپس کثار هم قرار دادن اطلاعات همه این ذرات مانند قرار دادن قطعات یک پازل، آشکارساز می‌تواند یک

”تصویر“ از برخورد برای تحلیل بیشتر بازسازی کند.

چگونه CMS کار می‌کند؟

آشکارساز ۱۴۰۰۰ تنی نام خود را از این موارد گرفته است:

با ۱۵ متر ارتفاع و ۲۱ متر طول، واقعاً برای تمام مواد آشکارسازی که در خود دارد کاملاً فشرده^۱ است؛ برای شناسایی ذراتی به نام میئن‌ها^۲ با دقت بسیار بالا طراحی شده است؛ و قوی‌ترین آهنربای سیم‌لوله‌ای^۳ که تاکنون ساخته شده را دارد.



شکل ۳.۱: طرح آشکارساز CMS

آشکارساز CMS به شکل یک پیاز استوانه‌ای است که دارای چندین لایه هم‌مرکز از اجزاء مختلف است. این اجزاء به تهیه ”عکس‌ها“ی هر رویداد برخورد با تعیین ویژگی‌های ذرات تولید شده در آن برخورد خاص کمک می‌کنند. این کار با انجام موارد زیر صورت می‌گیرد:

خم کردن ذرات

یک آهنربای قوی برای خم کردن ذرات باردار زمانی که از نقطه برخورد به بیرون پرتاپ می‌شوند، لازم است. خم کردن مسیر ذرات دو هدف را دنبال می‌کند:

- کمک به شناسایی بار ذره: ذرات باردار مثبت و منفی در یک میدان مغناطیسی مشابه در جهت‌های مخالف خم می‌شوند.

¹Compact

²Muon

³Solenoid

- امکان اندازهگیری تکانه ذره: در یک میدان مغناطیسی یکسان، ذرات با تکانه بالا کمتر از ذرات با تکانه پایین خم می‌شوند. آهنربای سیم‌لوله‌ای که نام CMS از آن گرفته شده است، از یک سیم‌پیچ استوانه‌ای از فیبرهای ابررسانا تشکیل شده است. زمانی که برق (۱۸۵۰۰ آمپر!) در این سیم‌پیچ‌ها جریان پیدا می‌کند، هیچ مقاومتی تجربه نمی‌کنند – جادوی ابررسانا! – و می‌توانند میدان مغناطیسی حدود ۴ تسلا تولید کنند که حدود ۱۰۰۰۰۰ برابر قوی‌تر از میدان مغناطیسی زمین است. این میدان مغناطیسی قوی باید به حجم آشکارساز محدود شود که با "یوغ" فولادی که بخش عمدۀ جرم آشکارساز را تشکیل می‌دهد، انجام می‌شود. سیم‌پیچ‌های آهنربای و یوغ بازگشتی آن وزنی معادل ۱۲۵۰۰ تن دارند که سنگین‌ترین جزء CMS است! این سیم‌لوله بزرگ‌ترین آهنربای نوع خود است که تاکنون ساخته شده و به ردیاب و انرژی‌سنجه اجازه می‌دهد تا داخل سیم‌پیچ قرار گیرند و نتیجه آن یک آشکارساز "فسرده" در مقایسه با آشکارسازهای با وزن مشابه است.

شناسایی مسیرها

فقط خم کردن ذرات کافی نیست – CMS باید مسیرهایی را که این ذرات باردار خم شده طی می‌کنند با دقت بسیار بالا شناسایی کند. این کار توسط یک ردیاب سیلیکونی انجام می‌شود که از حدود ۷۵ میلیون حسگر الکترونیکی منفرد تشکیل شده است که در لایه‌های هم‌مرکز قرار گرفته‌اند. هنگامی که یک ذره باردار از لایه ردیاب عبور می‌کند، به صورت الکترومغناطیسی با سیلیکون تعامل می‌کند و یک ضربه تولید می‌کند – این ضربه‌های منفرد می‌توانند به هم متصل شوند تا مسیر ذره عبوری شناسایی شود.

اندازهگیری تکانه ذرات برای کمک به ما در ساخت تصویری از آنچه در قلب برخورد اتفاق می‌افتد، بسیار مهم است. یک روش برای محاسبه تکانه یک ذره، دنبال کردن مسیر آن از طریق یک میدان مغناطیسی است؛ هرچه مسیر اhhh بیشتری داشته باشد، تکانه ذره کمتر است. ردیاب CMS مسیرهای گرفته شده توسط ذرات باردار را با اندازهگیری موقعیت‌های آنها در چندین نقطه ثبت می‌کند.

ردیاب می‌تواند مسیرهای میؤن‌ها، الکترون‌ها و هادرон‌های با انرژی بالا (ذراتی که از کوارک‌ها تشکیل شده‌اند) را بازسازی کند و همچنین می‌تواند مسیرهای ناشی از فروپاشی ذرات بسیار کوتاه‌عمر مانند کوارک‌های بیوتی یا "ب‌کوارک‌ها" را ببیند که برای مثال برای مطالعه تفاوت‌های بین ماده و پادماده استفاده می‌شوند.

ردیاب نیاز دارد تا مسیرهای ذرات را با دقت ثبت کند ولی در عین حال سبک باشد تا کمترین اختلال را در ذره ایجاد کند. این کار را با اندازهگیری موقعیت‌ها به قدری دقیق انجام می‌دهد که مسیرها می‌توانند به طور قابل اعتمادی با استفاده از چند نقطه اندازهگیری بازسازی شوند. هر اندازهگیری دقتی برابر با ۱۰ میکرومتر دارد، که بخشی از عرض موی انسان است. این بخش همچنین داخلی‌ترین قسمت آشکارساز CMS است و بنابراین بیشترین حجم ذرات را دریافت می‌کند؛ بنابراین مواد ساخت به دقت برای مقاومت در برابر تابش انتخاب شده‌اند.

طرح نهایی شامل ردیابی است که کاملاً از عناصر تشخیص سیلیکونی ساخته شده است: پیکسل‌ها، در هسته آشکارساز و

مواجه با بالاترین شدت ذرات، و ردیاب‌های میکرو نوار سیلیکونی که آن را احاطه کردند. همانطور که ذرات از ردیاب عبور می‌کنند، پیکسل‌ها و نوارهای میکرو سیلیکونی سیگنال‌های الکتریکی کوچکی تولید می‌کنند که تقویت و شناسایی می‌شوند. سیگنال‌ها به مدت چند میکروثانیه در حافظه تراشه‌ها ذخیره شده و سپس پردازش می‌شوند قبل از اینکه به لیزری فرستاده شوند تا به پالس‌های مادون قرمز تبدیل شوند. این پالس‌ها سپس از طریق یک کابل فیبر نوری ۱۰۰ متری برای تجزیه و تحلیل در محیط بدون تابش منتقل می‌شوند. ردیاب از ۴۰۰۰۰ لینک فیبر نوری استفاده می‌کند که روشی کم‌صرف و سبک برای انتقال سیگنال فراهم می‌کند. بسیاری از فناوری‌های پشت الکترونیک ردیاب از توآوری در همکاری با صنعت به دست آمدند.

ردیاب از حسگرهایی استفاده می‌کند که مساحتی به اندازه یک زمین تنیس را پوشش می‌دهند، با ۱۳۵ میلیون کانال خوانش الکترونیکی مجزا: در ردیاب پیکسل حدود ۶۶۰۰۰ اتصال در هر سانتی‌متر مربع وجود دارد. حسگرهای سیلیکونی به دلیل پاسخ سریع و وضوح مکانی خوب، بسیار مناسب برای دریافت تعداد زیادی ذره در فضای کوچک هستند.

ردیاب نوار CMS در اواخر سال ۲۰۰۷ در CMS نصب شد و از سال ۲۰۰۸ به طور مداوم در حال کار بوده است. ردیاب پیکسل اصلی CMS در سال‌های ۲۰۰۹ تا ۲۰۱۶ کار کرد و در یک توقف طولانی در پایان سال ۲۰۱۷/۲۰۱۶ با ردیاب پیکسل فاز-۱ CMS جایگزین شد.

اندازه‌گیری انرژی

اطلاعات در مورد انرژی‌های مختلف ذرات تولید شده در هر برخورد برای درک آنچه در نقطه برخورد رخ داده است، حیاتی است. این اطلاعات از دو نوع "انرژی‌سنج" در CMS جمع‌آوری می‌شود. انرژی‌سنج الکترومغناطیسی (ECAL) لایه داخلی است و انرژی الکترون‌ها و فوتون‌ها را با توقف کامل آنها اندازه‌گیری می‌کند. هادرون‌ها، که ذرات ترکیبی تشکیل شده از کوارک‌ها و گلوبون‌ها هستند، از ECAL عبور کرده و توسط لایه بیرونی به نام انرژی‌سنج هادرونی (HCAL) متوقف می‌شوند.

شناسایی میؤن‌ها

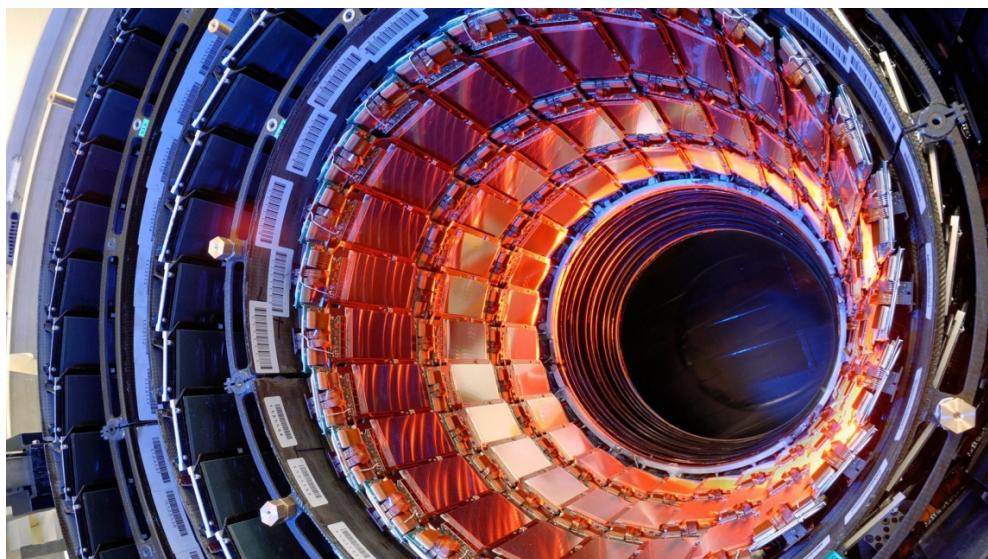
آخرین ذره‌ای که CMS مستقیماً مشاهده می‌کند، میؤن است. میؤن‌ها به همان خانواده ذرات الکترون تعلق دارند، اگرچه حدود ۲۰۰ برابر سنگین‌تر هستند. آنها توسط انرژی‌سنج‌ها متوقف نمی‌شوند، بنابراین زیرآشکارسازهای ویژه‌ای باید برای شناسایی آنها در حین عبور از CMS ساخته شوند. این زیرآشکارسازها با یوغ بازگشته سولنوئید تداخل دارند. آهنربای بزرگ CMS همچنین اجازه می‌دهد تا تکانه هر میؤن هم در داخل سیم‌پیچ ابررسانا (توسط دستگاه‌های ردیابی) و هم در خارج از آن (توسط اتاق‌های میؤن) اندازه‌گیری شود.

یکی از ویژگی‌های منحصر به فرد آشکارساز CMS این است که به جای ساختن در محل مانند سایر آشکارسازهای عظیم آزمایش‌های LHC، در سطح زمین ساخته شد و سپس به یک غار زیرزمینی نزدیک سسی در فرانسه پایین آورده شده و دوباره مونتاژ شد. مهندسان دریافتند که ساخت بخش‌ها در سطح زمین به جای ساخت آنها در غار با تمام مسائل دسترسی و ایمنی، زمان ارزشمندی را صرفه‌جویی کرد. نتیجه‌گیری مهم دیگر این بود که زیرآشکارسازها باید به گونه‌ای ساخته شوند که

دسترسی آسان‌تری داشته باشند تا نگهداری و تعمیرات سریع‌تر و آسان‌تر انجام شود.

بنابراین CMS به صورت پانزده بخش جداگانه یا "برش" طراحی شد که در سطح زمین ساخته شده و به صورت آماده به غار پایین آورده شدند. توانایی کار موازی در حفاری غار و ساخت آشکارساز زمان ارزشمندی را صرفه‌جویی کرد. این برش‌ها، همراه با طراحی دقیق کابل‌کشی و لوله‌کشی، همچنین تضمین می‌کنند که بخش‌ها می‌توانند با حداقل اختلال به‌طور کامل باز و بسته شوند و هر قطعه در داخل غار قابل دسترسی باقی بماند.

این ملاحظات، همراه با شرایط منحصر به فرد، LHC بر طراحی هر لایه از آشکارساز تأثیر گذاشت[۵].



شکل ۴.۱: نمایی از قسمت مسیریاب

۱۰.۳.۱ انرژی سنج الکترومغناطیسی

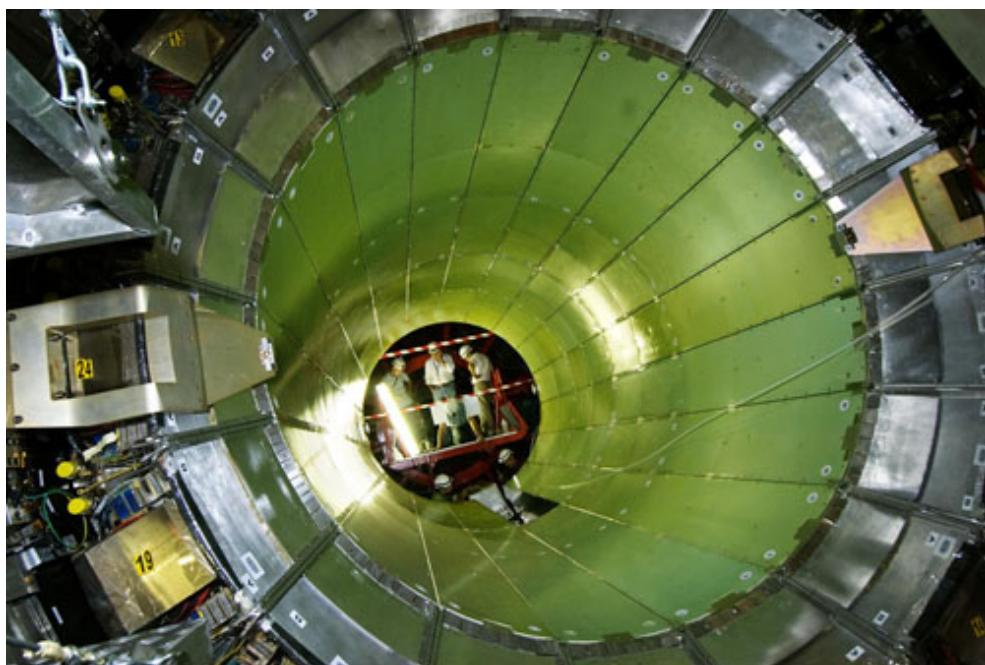
برای ساخت تصویری از رویدادهایی که در LHC رخ می‌دهد، CMS باید انرژی ذراتی که به بیرون می‌آیند را اندازه‌گیری کند. الکترون‌ها و فوتون‌ها به ویژه مورد توجه هستند، به دلیل نقشی که در کشف بوزون هیگر و فیزیک جدید دیگر دارند.

انرژی‌های الکترون‌ها و فوتون‌ها با استفاده از کالری‌متر الکترومغناطیسی CMS (ECAL) اندازه‌گیری می‌شود. اندازه‌گیری انرژی‌های آنها با دقت لازم در شرایط بسیار سخت LHC - میدان مغناطیسی قوی، سطوح بالای تابش و فقط ۲۵ نانوثانیه بین برخوردها - نیاز به مواد آشکارساز خاص دارد.

کریستال تنگستات سرب عمدتاً از فلز ساخته شده و سنگین‌تر از فولاد ضد زنگ است، اما با یک لمسه اکسیژن در این فرم کریستالی، بسیار شفاف بوده و هنگامی که الکترون‌ها و فوتون‌ها از آن عبور می‌کنند "سوسو می‌زنند". این بدان معناست که متناسب با انرژی ذره ورودی نور تولید می‌کند. این کریستال‌های با چگالی بالا نور را در نوسانات کوتاه و سریع و مشخص تولید می‌کنند که اجازه می‌دهد یک آشکارساز دقیق، سریع و نسبتاً فشرده باشد.

آشکارسازهای نوری که بهویژه برای کار در میدان مغناطیسی قوی طراحی شده‌اند، به پشت هر یک از کریستال‌ها چسبانده شده‌اند تا نور سوسوزن را شناسایی کرده و آن را به سیگنال الکتریکی تبدیل کنند که تقویت شده و برای تحلیل فرستاده می‌شود. ECAL، که از یک بخش "استوانه" و دو "کلاهک انتهایی" تشکیل شده است، لایه‌ای بین ردیاب و HCAL را تشکیل می‌دهد. استوانه استوانه‌ای شامل ۶۲۰۰ کریستال است که در ۳۶ "ابر مازول" تشکیل شده‌اند، هر کدام حدود سه تن وزن دارند و ۱۷۰۰ کریستال را در خود جای داده‌اند. کلاهک‌های انتهایی تخت استوانه را در هر انتهای می‌بندند و شامل تقریباً ۱۵۰۰۰ کریستال بیشتر هستند.

برای دقت مکانی بیشتر، ECAL همچنین شامل یک آشکارساز پیش‌دوش است که در جلوی کلاهک‌های انتهایی قرار دارد. این‌ها به CMS اجازه می‌دهند تا بین فوتون‌های تک با انرژی بالا (که اغلب نشانه فیزیک هیجان‌انگیز هستند) و جفت‌های نزدیک فوتون‌های با انرژی پایین که کمتر جالب هستند، تمایز قائل شود^[۵].



شکل ۵.۱: انرژی سنج الکترومغناطیسی آشکارساز سیم لوله‌ی فشرده میؤنی

کریستال‌ها هر کدام ۵۰.۱ کیلوگرم وزن دارند اما با حجمی تقریباً معادل یک فنجان کوچک قهوه، شامل نزدیک به ۸۰۰۰۰ چنین کریستالی است که هر کدام دو روز طول می‌کشد تا رشد کنند.

۲۰.۳.۱ انرژی سنج هادرونی

انرژی سنج هادرونی (HCAL) انرژی "هادرون‌ها"، ذراتی که از کوارک‌ها و گلوبون‌ها تشکیل شده‌اند (مانند پروتون‌ها، نوترون‌ها، پایون‌ها و کائون‌ها)، را اندازه‌گیری می‌کند. علاوه بر این، اندازه‌گیری غیرمستقیم حضور ذرات بدون بار و غیرقابل

تعامل مانند نوتريونها را نيز فراهم می‌کند.

اندازه‌گيري اين ذرات مهم است زира می‌توانند به ما بگويند که آيا ذرات جديدي مانند بوزون هيگر يا ذرات ابرمتقارن (نسخه‌های بسيار سنگين‌تر از ذرات استانداردي که می‌شناسيم) تشکيل شده‌اند يا خير.

زمانی که اين ذرات فروپاشی می‌کنند ممکن است ذرات جديدي توليد کنند که هيچ اثری از حضورشان در هيچ بخشی از آشكارساز CMS به جا نمی‌گذارند. برای شناسايي اين ذرات، HCAL باید "هرمتيك" باشد، به اين معني که باید تا حد امکان هر ذره‌اي که از برخوردها بیرون می‌آيد را ثبت کند. به اين ترتيب اگر ببینيم که ذرات از يك طرف آشكارساز بیرون می‌روند ولی از طرف ديگر نه، با عدم تعادل در تکانه و انژي (که در جهت "عرضي" نسبت به خط باريکه اندازه‌گيري می‌شود)، می‌توانيم نتیجه بگيريم که ما در حال توليد ذرات "نامائي" هستيم.

برای اطميان از اينکه چيزی جديد را مشاهده می‌کنيم، نه فقط ذرات آشنايي که به طور نامشخص فرار کرده‌اند، لايه‌های HCAL به صورت پله‌اي ساخته شده‌اند تا هيچ شكافي در خطوط مستقيم وجود نداشته باشد که ذرات آشنا از آن فرار کنند. HCAL يك كالري متر نمونه برداري است که به اين معنى است که موقعيت، انژي و زمان رسيدن ذره را با استفاده از لايه‌های متناوب "جادب" و مواد فلورسانس "سوسوزن" که در هنگام عبور ذره نور پالس سريعي توليد می‌کنند، پيدا می‌کند. فيبرهای نوري ويزه اين نور را جمع کرده و به جعبه‌های خوانش می‌فرستند که در آنجا فوتواشكارسازها سيگنال را تقويت می‌کنند. زمانی که مقدار نور در يك منطقه مشخص در عمق لايه‌های متعدد کاشی‌ها، به نام "برج"، جمع‌آوري شود، اين مقدار کل نور يك اندازه‌گيري از انژي ذره است.

با توجه به اينکه HCAL بزرگ و ضخيم است، جاي دادن آن در CMS "فسرده" يك چالش بود، زيرا ريزش‌های ذرات توليد شده وقتی يك هادرон به مواد جاذب چگال برخورد می‌کند (که به عنوان "شاور" شناخته می‌شود) بزرگ هستند و حداقل مقدار ماده مورد نياز برای محصور کردن و اندازه‌گيري آنها حدود يك متر است.

برای دستيابي به اين هدف، HCAL به بخش‌های استوانه‌اي (HO و HB)، انتهائي (HE) و جلو (HF) سازماندهی شده است. ۳۶ "گوه" استوانه‌اي، هر کدام با وزن ۲۶ تن، تشکيل‌دهنده آخرین لايه آشكارساز در داخل سيم‌پيچ آهنربا هستند در حالی که چند لايه اضافي، استوانه بيروني، (HO) در خارج از سيم‌پيچ قرار دارند تا اطميان حاصل شود که هيچ انژي از پشت HB به طور ناشناس نشست نمی‌کند. به طور مشابه، ۳۶ گوه انتهائي انژي ذرات را زمانی که از انتهائي آهنرباي سولنوئيدي بیرون می‌آيند، اندازه‌گيري می‌کنند.

در نهاييت، دو كالري متر هادروني جلو (HF) در هر دو انتهائي CMS قرار دارند تا ذرات مختلفي که از منطقه برخورد با زويای کم نسبت به خط باريکه خارج می‌شوند را ثبت کنند. اين‌ها بيشتر انژي ذرات موجود در برخورد را دريافت می‌کنند، بنابراین باید بسيار مقاوم به تابش باشند و از مواد مختلفي نسبت به ساير قسمت‌های HCAL استفاده می‌کنند.

از بيش از يك ميليون پوسته برنجي جنگ جهاني دوم از نيروي دريایي روسيه در ساخت برخى از اجزاي آشكارساز استفاده

کرده است؛ از ۳۶ گوه تشکیل شده است که هر کدام وزنی معادل ۶ فیل آفریقایی دارند؛ شامل بیش از ۴۰۰ واحد "رمزگشای نوری" است که همه آنها توسط دانشآموزان دبیرستانی آمریکایی از طریق برنامه QuarkNet ساخته شده‌اند^[۵].

۳.۳.۱ سیستم میؤن

یکی از مهم‌ترین وظایف CMS تشخیص میؤن‌ها است. میؤن‌ها ذرات باردار هستند که مانند الکترون‌ها و پوزیترون‌ها عمل می‌کنند، اما ۲۰۰ برابر سنگین‌تر هستند. انتظار داریم که آنها در تجزیه چندین ذره جدید محتمل تولید شوند؛ به عنوان مثال، یکی از وضوح‌های واضح بوسون هیگز، تجزیه آن به چهار میؤن است.

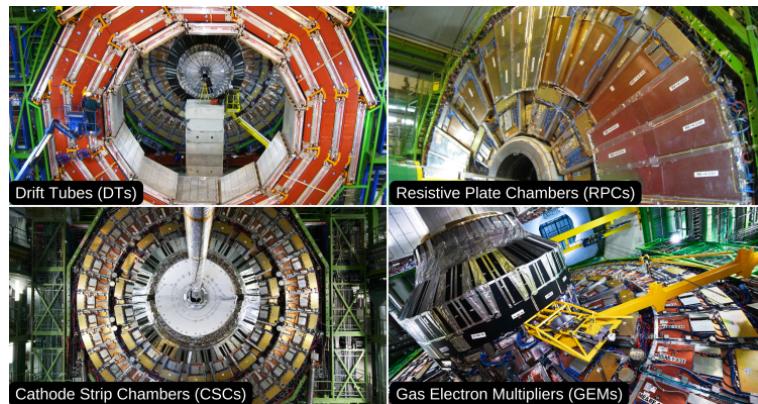
زیرا میؤن‌ها می‌توانند چندین متری از مواد را با انرژی کم از دست بدهنند، برخلاف اکثر ذرات، توسط هیچ‌کدام از کالری‌مترهای CMS متوقف نمی‌شوند. بنابراین، اتاق‌هایی برای شناسایی میؤن‌ها در بخش خارجی آزمایشگاه قرار داده شده‌اند که احتمالاً تنها ذراتی هستند که یک سیگنال روش تولید می‌کنند.

یک ذره به وسیله ساخت یک منحنی با استفاده از ضربات ثبت شده در چهار ایستگاه میؤن، که در بخش خارجی پلاک‌های آهن "یوک بازگشت" جای گذاشته شده‌اند، اندازه‌گیری می‌شود. مسیر ذره توسط پیگیری موقعیت آن از طریق لایه‌های فعال چندگانه هر ایستگاه اندازه‌گیری می‌شود؛ برای دقت بهتر، این اطلاعات با اندازه‌گیری‌های ردیاب سیلیکونی CMS ترکیب می‌شود. اندازه‌گیری مسیر، اندازه‌گیری تکانه ذره را فراهم می‌کند. در واقع، میدان مغناطیسی قوی تولید شده توسط سولنوید CMS مسیر ذره را انحنا می‌دهد، با شعاع انحنایی که به تکانه آن بستگی دارد؛ هر چه مسیر مستقیم‌تر باشد، تکانه بالاتری دارد.

کلیه میؤن‌ها شامل ۱۴۰۰ اتاق میؤن هستند: ۲۵۰ لوله دریفتی (DTs) و ۵۴۰ اتاق نوار کاتد (CSCs) موقعیت‌های ذرات را پیگیری کرده و تریگر را فراهم می‌کنند، در حالی که ۶۱۰ اتاق پلاک مقاوم (RPCs) و ۷۲ اتاق گس مولتیپلایر گازی (GEMs) سیستم تریگر تکراری را تشکیل می‌دهند که به سرعت تصمیم می‌گیرد که داده‌های میؤن اکتساب شده را نگه دارد یا دور بریزد. به دلیل بسیاری از لایه‌های آزمایشگاه و تخصص‌های مختلف هر نوع، سیستم به طور طبیعی قدرتمند و قادر به فیلتر کردن نویز پس‌زمینه است.

یک میؤن، در صفحه متعامد بر شعاع‌های LHC، یک مسیر منحنی را در چهار لایه از آشکارسازهای میؤن ("ایستگاه‌ها") ترک می‌کند.

RPCs و DTs به شکل مخروطی متمرکز در اطراف خط بیمار (ناحیه بشکه) ترتیب داده شده‌اند، در حالی که، CSCs چهارچوبی و GEMs دیسک‌های "پایانه" را تشکیل می‌دهند که انتهای بشکه را "بسته" می‌کنند. شامل ۲ میلیون سیم اتاق نوار کاتد است که هر کدام به اندازه موی انسان نازک هستند؛ با ردیاب مرکزی به حدود یک ششم میلی‌متر تطبیق داده شده است تا آشکارسازها بتوانند با هم در باسازی مسیرها کار کنند؛ از مؤلفه‌های ساخته شده در ۱۵ کشور تشکیل شده است. تا به امروز، گروه سیستم میؤن از ۶۳ مؤسسه از ۲۲ کشور تشکیل شده است. ساخت اتاق‌های میؤن سخت



شکل ۶.۱: نمایی از بخش‌های مختلف سیستم میون

بود

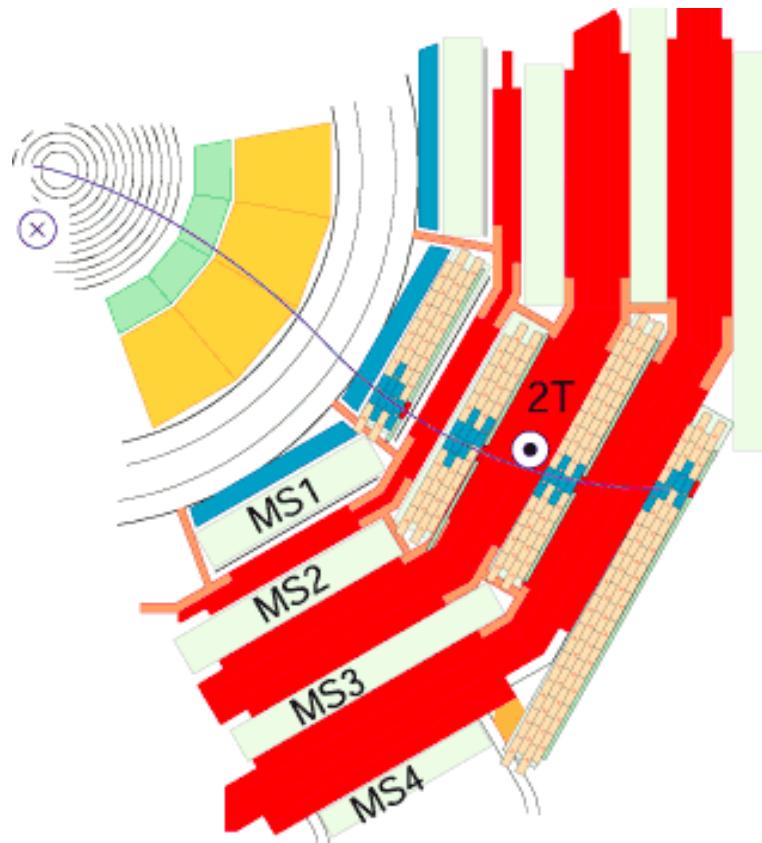
بیشتر اتاق‌های میون در آزمایشگاه‌های مختلف در سراسر جهان ساخته شدند پیش از اینکه به سراسر سرن ارسال شوند، بنابراین اتاق‌ها باید مقاومت مکانیکی آنها و عملکرد آنها را در چندین مرحله از پروتوتاپ، طراحی، پرتوهای آزمایشی، مونتاژ و بهره‌برداری در آزمایشگاه‌های مربوطه و در سرن آزمایش کردند[۵].

۴.۱ بازسازی اشیاء

برای بررسی ذرات در آشکارساز نیاز است مسیر حرکت ذرات بازسازی شود. این به این معنی است که با توجه به سیگنال‌های که در قسمت‌ها مختلف که در آشکارساز ثبت شده است مسیری برای تشخیص ساخته می‌شود که می‌توانیم آن‌ها را بررسی کنیم. برای این کار از الگوریتم‌های مختلفی استفاده می‌شود هدف اصلی این الگوریتم‌ها بازسازی اشیاء با کاربرد زیاد و حذف اشیائی است که به اشتباه تشخیص داده شده‌اند. فرایند بازسازی رویداد در آزمایش CMS شامل چندین مرحله است، هر کدام طراحی شده‌اند تا از قطعات مختلفی از آشکارساز را استخراج شوند. در ادامه روش بازسازی اشیاء را توضیح می‌دهیم و برای جت و فتون را که دو شیء اصلی در این پایان نامه هستند یخش‌ها اساسی‌تر در نظر گرفته‌ایم.

۱۰.۴.۱ الگوریتم Particle Flow

بازسازی سنتی رویدادها در برخوردهای هادرونی به این صورت بود که جت‌ها متشکل از هادرون‌ها و فوتون‌ها در نظر گرفته می‌شدند و انرژی آن‌ها به‌طور انحصاری توسط انرژی سنج‌ها اندازه‌گیری می‌شد، بدون شناسایی ذرات فردی درون جت‌ها و بدون استفاده از اطلاعات ردیاب‌ها و آشکارسازهای میونی. اندازه‌گیری تکانه عرضی از دست رفته (



شکل ۷.۱: نمایی از سیستم میؤن

فصل دوم

درخشدگی و پایل آپ

۱۰.۲ اندازه‌گیری درخشدگی در برخورد دهنده ذرات

درخشدگی یکی از پارامترهای کلیدی آزمایش‌ها در برخورد دهنده ذرات است که برای تعیین پتانسیل فیزیکی از لحاظ نرمال کردن آمار رویدادها بکار می‌رود. اندازه‌گیری دقیق درخشدگی از اهمیت حیاتی برخوردار است، زیرا عدم قطعیت درخشدگی به طور مستقیم به عدم قطعیت اندازه‌گیری سطح مقطع برخورد ذرات منتقل می‌شود. درخشدگی بالا برای اندازه‌گیری سطح مقطع پایین مورد نیاز است، مثل اندازه‌گیری هیگز یا جستجوهای ابرتقارن. برای فرآیندهای با سطح مقطع بالا، مانند تولید بوزون‌های W و Z ، عدم قطعیت درخشدگی اغلب عدم قطعیت اصلی است. این در شکل ۱۰.۲ نشان داده شده است که نسبت‌های سطح مقطع تجربی به نظری و نسبت‌های سطح مقطع برای بوزون‌های W و Z را نشان می‌دهد. جعبه سایه‌دار، عدم قطعیت در اندازه‌گیری درخشدگی را نشان می‌دهد. نوارهای خطای داخلی عدم قطعیت‌های تجربی را نشان می‌دهند، نوارهای خطای خارجی همچنین عدم قطعیت‌های پیش‌بینی‌های نظری را شامل می‌شوند.

علاوه بر این، اندازه‌گیری درخشدگی به ازای هر دسته عبوری برای تشخیص و بهینه‌سازی شتاب دهنده اهمیت دارد. برای ارائه یک اندازه‌گیری پیوسته از درخشدگی و امکان بررسی سطح مقطع دائمی برای ردیابی در زمان واقعی احتمال عیب‌یابی‌های آشکارسازها، استفاده از بیش از یک درخشدگی سنج برای آزمایش مفید است. عملکرد در بالاترین درخشدگی ممکن برای کاوش

فرآیندهای سطح مقطع پایین مورد نیاز است. پس از اولین دوره خاموشی طولانی مدت^۱ انتظار می‌رفت که LHC با فاصله دسته ۲۵ نانوثانیه و انرژی ۶.۵ تراکترون ولت برای هر شاره عمل کند. برای عملکرد در بالاترین درخشندگی ۱۰^{۳۴} سانتیمتر مکعب بر ثانیه با دسته هایی هایی با شدت (O¹¹) پروتون، درخشندگی سنج CMS ارتقا یافته یا بازطراحی شدند. پس از LS1، آزمایش CMS با سه درخشندگی سنج آنلاین مستقل در حال اجرا بود: انرژی سنج هادرон جلو، تلسکوپ درخشندگی پیکسل^۲ و نظارت شرایط تابش سریع بار.

the Beam Condition Monitor “Leakage” (BCML) •

the Beam Halo Monitor (BHM) •

the Fast Beam Conditions Monitor (BCM1F) •

از سال ۲۰۱۸، CMS قادر است اندازه‌گیری درخشندگی را از چندین منبع مختلف ارائه دهد. این اندازه‌گیری‌ها هم به صورت آنلاین (برخط) و در زمان واقعی به LHC و CMS و هم به صورت آفلاین (ناخط) با تحلیل‌های بعدی برای ارائه بهترین اندازه‌گیری‌های درخشندگی برای تحلیل‌های فیزیکی انجام می‌شود. اندازه‌گیری‌های آنلاین درخشندگی:

اندازه‌گیری‌های درخشندگی آنلاین توسط تلسکوپ درخشندگی پیکسلی، (PLT) کالریومتر پیش‌رونده هادرونی، (HF) مانیتور شرایط پرتو سریع (BCM1F) و لوله‌های لقی (DT) انجام می‌شود. اندازه‌گیری‌های آفلاین درخشندگی:

علاوه بر موارد آنلاین، شمارش خوش‌های پیکسلی (PCC) با آشکارگر اصلی CMS و همچنین شمارش رأس^۳ نیز به صورت آفلاین قابل استفاده است.

الگوریتم‌ها:

PLT از روش شمارش صفر (PLTZ) و HF از هر دو روش جمع (HFET) ET و روش مبتنی بر اشغال (HFOC) استفاده می‌کند.

اندازه‌گیری درخشندگی در Run 1:

در Run 1، اندازه‌گیری درخشندگی در CMS بر اساس اندازه‌گیری اشغال HF (HFOC) بود. جزئیات عدم قطعیت‌های سیستماتیک برای Run 1 مربوط به ثابت عادی‌سازی را می‌توان در [۶] یافت.

تایید درخشندگی:

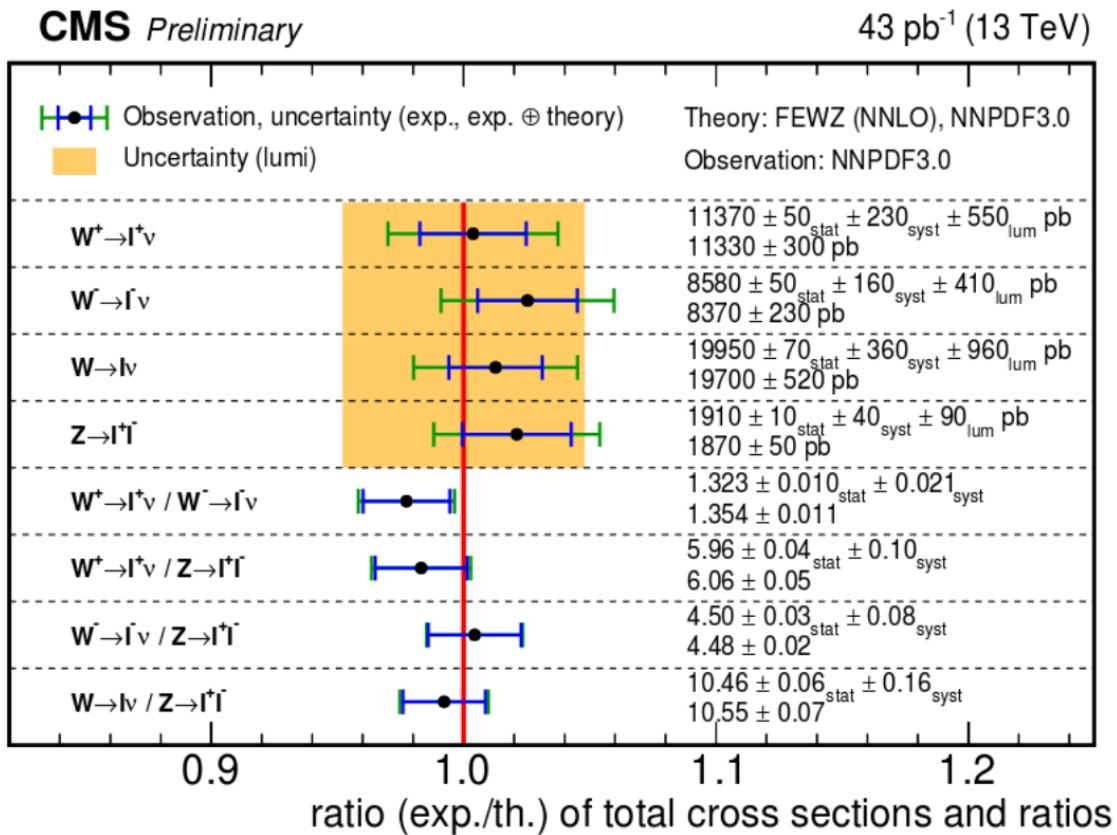
درخشندگی با استفاده از اسکن‌های جداسازی که اولین بار توسط سایمون ون در میر در ISR پیشگام شد، کالیبره می‌شود.

¹LS1(Long Shutdown 1)

²Pixel Luminosity Telescope

³vertex counting

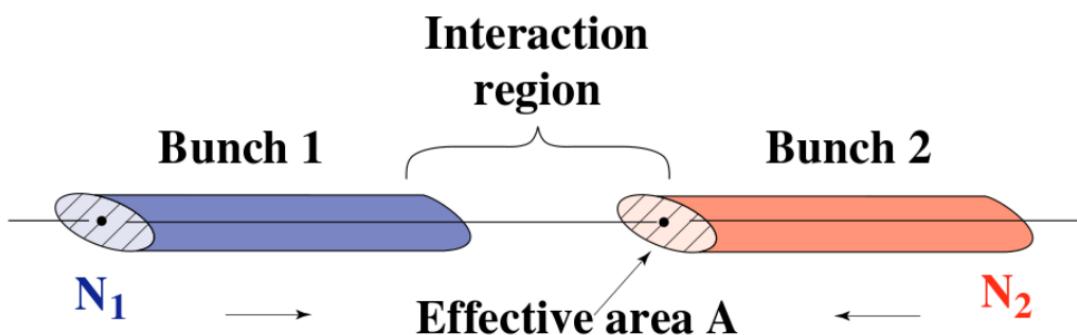
این "اسکن‌های ون در میر" امکان تعیین درخشندگی مطلق را به عنوان تابعی از پارامترهای پرتو فراهم می‌کند و بنابراین می‌توان از آن‌ها برای توسعه کالیبراسیون برای درخشندگی سنج‌ها استفاده کرد.



شکل ۱۰.۲: خلاصه‌ای از مجموع سطح مقطع تولید W^+ ، W^- ، W و Z به همراه کسرهای منشعب، نسبت‌های W به Z و W^+ به W^- و پیش‌بینی‌های نظری آن‌هاست.

۱۰.۲ درخشندگی مطلق از پارامترهای ماشین

درخشندگی مطلق در LHC از پارامترهای ماشین تعیین می‌شود، که نیازی به دانش سطح مقطع ذرات ندارد.



شکل ۲۰.۲: تصویر عبور دسته‌ی پروتون‌ها با پارامترهای نسبی

برای تلاقی دسته‌هایی با تعداد ذرات N_1 و N_2 ، همانطور که در شکل ۲.۲ نشان داده شده است، درخشنندگی به شکل زیر

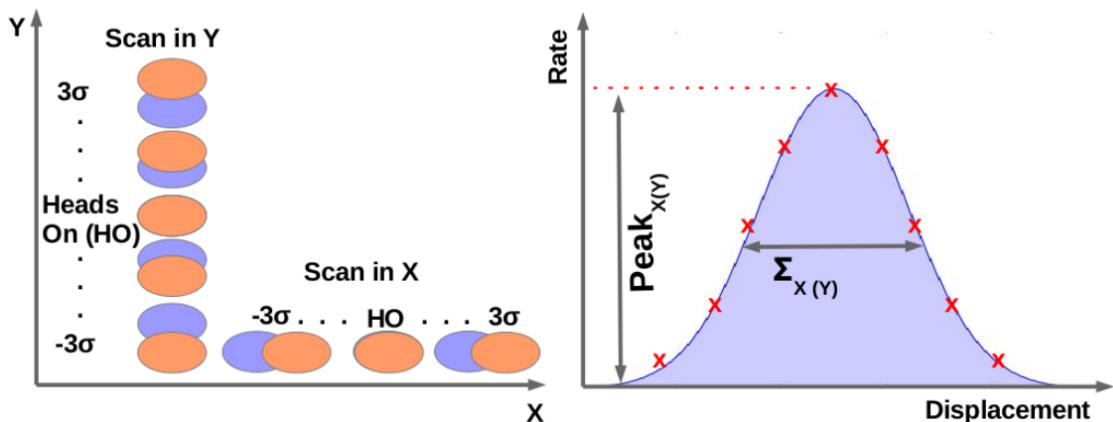
تعریف می‌شود:

$$L = \frac{N_1 \cdot N_2 \cdot f \cdot n_b}{A_{\text{eff}}} \quad (2.2)$$

فرکانس چرخش (f)، ناحیه درخشنان (A_{eff}) و تعداد دسته‌های برخورد (n_b) از اندازه‌گیری‌های دستگاه LHC به دست می‌آید. ناحیه درخشنان، با فرض توزیع گوسی چگالی ذرات در دسته‌های برخورد، به شکل زیر محاسبه می‌شود:

$$A_{\text{eff}} = 4\pi \cdot \sigma_x \cdot \sigma_y \quad (2.2)$$

که σ_x و σ_y عرض‌های گوسی توزیع چگالی ذرات در کناره دسته‌ها هستند. برای دستیابی به درخشنندگی بالا، باید پرتوها را در نقطه تعامل به اندازه‌گیری ممکن‌ترین ناحیه ممکن تمرکز دهیم. عرض ناحیه درخشنان برای کالیبراسیون درخشنندگی سنج‌ها لازم است و از اسکن‌های VdM¹ (شکل ۳.۲) به دست می‌آید.



شکل ۳.۲: چپ: نمودار اسکن‌های VdM در صفحات x و y. راست: نمودار نرخ درخشنندگی سنج به عنوان تابع انحراف به دست آمده از اسکن VdM.

۲۰.۲ محاسبه درخشنندگی با استفاده از سطح مقطع فیزیکی

نرخ برخورد N_{col} برای یک فرآیند با سطح مقطع معلوم σ_{crop} که توسط یک دستگاه با درخشنندگی L تولید شده است:

$$N_{\text{col}} = L \times \sigma_{\text{proc}} \quad (3.2)$$

¹van der Meer scans

اگر سطح مقطع یک فرآیند معلوم باشد، درخشندگی می‌تواند از این رابطه از نرخ رویداد مشاهده شده اندازه‌گیری شود. در برخورد دهنده‌های الکترون-پوزیترون، پراکندگی $e^+e^- \rightarrow e^+e^-$ یا فرآیند بابا^۱ معروف برای این منظور استفاده می‌شود. برای برخورد pp در LHC با انرژی ۱۴ ترا الکترون‌ولت در مرکز جریان، نرخ تولید W و Z بالاست و اعتقاد بر این است که برای محاسبه درخشندگی مناسب خواهد بود، اگر دقت محاسبه سطح مقطع برای تولید W و Z بهبود یافته باشد. وقتی برهمنهی دسته‌ها را در نظر می‌گیریم، درخشندگی آنی ناحیه به نسبت تعداد برخوردها در هر دسته عبوری (μ) است. تعداد برهمنکش‌ها بر اساس سطح مقطع کمینه تصادفی pp (σ_0) به درخشندگی مرتبط است:

$$L = \frac{f \cdot \mu}{\sigma_0} \quad (4.2)$$

که f فرکانس چرخش LHC است. محاسبه درخشندگی به این روش تنها توسط پارامترهای دستگاه انجام می‌شود، که محدودیت‌هایی دارد که ناشی از دقت سطح مقطع پیش‌بینی شده σ_0 است.

۳.۱۰.۲ محاسبه درخشندگی آنلاین

با توجه به این که تعداد رویدادها در هر دسته عبوری از یک توزیع پواسون پیروی می‌کند، احتمال وجود تعداد معین از رویدادها n هم، از توزیع پواسون پیروی می‌کند:

$$P(n) = \frac{\mu^n e^{-\mu}}{n!} \quad (5.2)$$

که در آن μ میانگین تعداد رویدادها در هر دسته عبوری است.

احتمال اینکه هیچ رویدادی وجود نداشته باشد، (p^0 ، برای اندازه‌گیری مقدار میانگین μ استفاده می‌شود:

$$\mu = -\ln[p^0] = -\ln[1 - p(n \neq 0)] \quad (6.2)$$

ماژول‌های RHU^۲ برای اندازه‌گیری نرخ آنلاین BCM1F استفاده می‌شوند. نرم‌افزار استفاده شده برای خواندن داده‌ها از RHU در چارچوب نرم‌افزاری گرفتن داده آنلاین CMS با نام XDAQ پیاده‌سازی شده است و این سیستم ماژول‌لار است، بنابراین داده‌ها از عنصر به عنصر از طریق یک صفحه رویدادی گذر می‌کند. ماژولی که با سخت‌افزار RHU ارتباط برقرار می‌کند، به نام منبع نامیده می‌شود.

¹Bhabha process

²Real-time Histogramming Unit

منبع هیستوگرام‌های RHU را جمع‌آوری می‌کند و آن‌ها را از طریق صفحه رویدادی به پردازنده منتقل می‌کند، که مازولی است که از الگوریتم zero-counting برای محاسبه درخشندگی و همچنین اندازه‌ی پس‌زمینه، همراه با اطلاعات نظرارتی و هیستوگرام‌های نرخ کانال جمع شده برای تجزیه و تحلیل آفلاین، مطابق با شکل ۴.۲، استفاده می‌کند. مقادیر محاسبه‌شده سپس از پردازنده از طریق صفحه رویدادی به جمع‌آوری داده ارسال می‌شود که مقادیر را آنلاین منتشر می‌کند و برای استفاده آفلاین ذخیره می‌کند. نرخ شمارش RHU در هنگام اسکن VdM^۱ برای کالیبره کردن BCM1F به عنوان یک درخشندگی‌سنجد استفاده می‌شود.

هیستوگرام‌های RHU که در طول چهار بار درخشندگی جمع‌آوری می‌شوند، شامل ۴ باکس در هر دسته عبوری هستند. محتوای باکس سوم برای محاسبه درخشندگی استفاده می‌شود. ابتدا بر تعداد مدارها که ۱۶۳۸۴ است تقسیم می‌شود و سپس از ۱ کم می‌شود، که نسبت شمارش‌های صفری (p^0) است. سپس معادله ۶.۲ برای تعیین میانگین مقادیر ضربه برای هر دسته عبوری در هر چرخش استفاده می‌شود.

درخشندگی در هر دسته عبوری به صورت زیر بدست می‌آید:

$$L(\text{BCID}) = \frac{\mu}{\sigma_{\text{vis}}} \times f \quad (7.2)$$

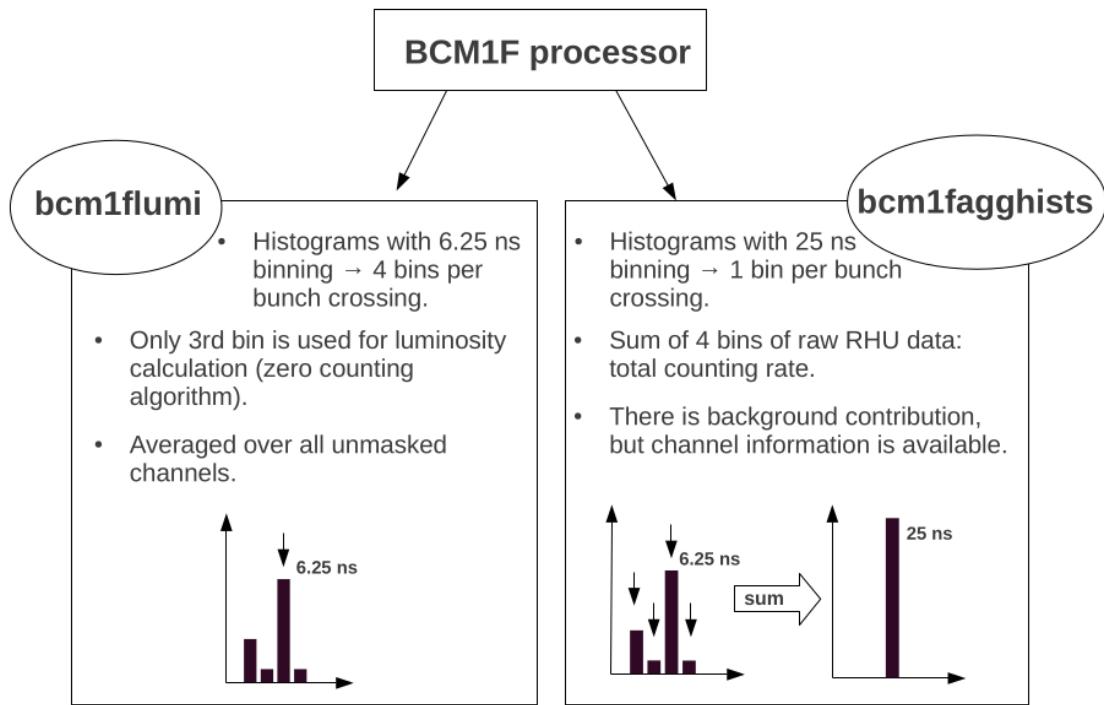
که BCID شماره شناسایی دسته‌عبوری است. برای هر چهار نوبت درخشندگی میانگین $L(\text{BCID})$ برای تمام کانال‌های غیرماسک شده محاسبه، ذخیره و به CMS و LHC به عنوان درخشندگی آنلاین ارسال می‌شود. مقادیر ذخیره شده در تجزیه و تحلیل اسکن VdM استفاده می‌شوند.

۴.۱۰.۲ برخورد دهنده بزرگ هادرونی با درخشندگی بالا

پروژه برخورد دهنده بزرگ هادرونی با درخشندگی بالا^۱ به هدف افزایش عملکرد شتاب دهنده LHC با هدف افزایش امکان کشف‌ها پس از سال ۲۰۲۹ می‌پردازد. هدف، افزایش درخشندگی مجموع به ۱۰ برابر از مقدار طراحی شده LHC است. درخشندگی یک نشانگر مهم از عملکرد یک شتاب دهنده است: این متناسب با تعداد برخوردهایی است که در یک مدت زمان مشخصی رخ می‌دهد. هرچه درخشندگی بیشتر باشد، داده‌های بیشتری که آزمایش‌ها می‌توانند جمع‌آوری کنند تا به آن‌ها اجازه دهد فرایندهای نادر را مشاهده کنند.

برخورد دهنده بزرگ هادرونی با درخشندگی بالا، که در ابتدای سال ۲۰۲۹ بهره‌برداری خواهد شد، به فیزیکدان‌ها امکان مطالعه مکانیزم‌های شناخته شده را با جزئیات بیشتری می‌دهد، مانند بوزون هیگز، و مشاهده پدیده‌های نو و نادری که ممکن است فاش شوند. به عنوان مثال، LHC با درخشندگی بالا حداقل ۱۵ میلیون بوزون هیگز را در سال تولید می‌کند، در مقایسه با حدود سه میلیون بوزون هیگز که از LHC در سال ۲۰۱۷ تولید شد.

^۱High-Luminosity LHC



شکل ۴.۲: منطق پردازندۀ BCM1F

سمت چپ: درخشندگی آنلاین، عرض بازه ۶.۲۵ نانوثانیه (سومین بین داده‌های RHU).

سمت راست: هیستوگرام‌های نرخ کanal چمۀ شده، عرض بازه ۲۵ نانوثانیه (جمع ۴ بین RHU).

پروژه برخورد دهنده بزرگ هادرونی با درخشندگی بالا به عنوان اولین استراتژی اروپایی برای فیزیک ذرات در سال

۲۰۱۳ اعلام شد.

توسعۀ آن وابسته به چند نوآوری فناورانه است. فاز اول پروژه در سال ۲۰۱۱ آغاز شد و تا حدی توسط کمیته اروپایی FP7

^۱ پشتیبانی مالی شد. این فاز آزمایشگاه‌های بسیاری از کشورهای عضو CERN همچنین آمریکا، ژاپن و روسیه را گرد هم

آورد. مطالعه طراحی با انتشار یک گزارش طراحی مقدماتی^۲ در اکتبر ۲۰۱۵ به پایان رسید و عملیات ساخت برای پروژه در

CERN و در صنعت آغاز شد. گزارش طراحی فنی LHC با درخشندگی بالا^۳ در سال ۲۰۱۷ منتشر شد - بازتاب کننده طراحی

قبل از تغییرات بررسی مجدد در سال ۲۰۱۶ - و نسخه نهایی آن در سال ۲۰۲۰. کارهای عمرانی از آوریل ۲۰۱۸ آغاز شد.

این پروژه تحت رهبری سرن با حمایت از همکاری بین‌المللی ۴۴ موسسه در ۲۰ کشور انجام می‌شود - کشورهای مختلف

اروپا از جمله ایتالیا، اسپانیا، سوئد و انگلستان - و شامل تعدادی از کشورهای غیرعضو CERN مانند ایالات متحده، ژاپن و

کانادا. پروژه HL-LHC بهبود عملکرد ستارگان LHC را با افزایش پتانسیل کشف‌ها پس از سال ۲۰۲۹ هدف گرفته است.

¹ European Commission's seventh framework programme

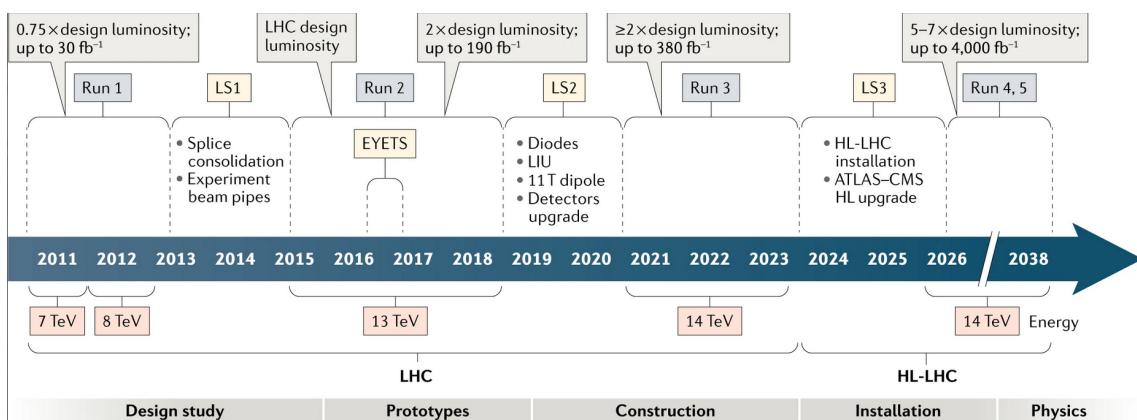
² Preliminary Design Report (PDR)

³ Technical Design Report (TDR)

افزایش درخشندگی در مرتبه ۱۰ برابر مقدار طراحی LHC هدف این پروژه است.

درخشندگی یک شاخص مهم از عملکرد یک شتابگر است: هر چه درخشندگی بالاتر باشد، داده‌های بیشتری آزمایش‌ها

می‌توانند جمع‌آوری کنند تا به آن‌ها امکان مشاهده پدیده‌های نادر را بدهد.^[۱۲]



شکل ۵.۲: طرح پایه LHC و HL-LHC برای دهه آینده و بعد از آن.

عملکرد برخورد دهنده بزرگ هادرونی (LHC) یک دنباله از اجراهای عملیاتی ۴-۳ ساله و توقف‌های طولانی (LS) ۲ ساله است. پس از LS3 (۲۰۲۶-۲۰۲۴)، ماشین در حالت برخورد دهنده بزرگ هادرونی با درخشندگی بالا خواهد بود. 4000 fb^{-1} هدف درخشندگی یکپارچه نهایی است (مقدار طراحی LHC درخشندگی 300 fb^{-1} است). درخشندگی طراحی شده، مقدار اصلی درخشندگی است که LHC برای دستیابی به آن طراحی شده است.

۲۰۲ پایل آپ

اندازه‌گیری دقیق سطح مقطع هادرونیک، ورودی مدل‌های پدیدارشناسی^۱ را فراهم می‌کند و برای تنظیم تولیدکننده‌های رویداد مونت کارلو مورد نیاز است. مقدار سطح مقطع غیرکشسان و در تخمین تعداد برهمنکنش‌های همزمان pp که در دسته پرتو عبوری رخ می‌دهند که از این پس به عنوان "پایل آپ" نامیده می‌شود، استفاده می‌شود.

پایل آپ اثربخش است که ناشی از چندین برخورد در هر برخورد دسته و یا برهمنکش رویدادهای همپوشانی است. این رویدادهای همپوشانی که با نام رویدادهای "بایاس حداقل"^۲ یا "پایل آپ" شناخته می‌شوند، باعث ایجاد نویز پس‌زمینه قابل توجهی در پاسخ سیستم آشکارساز می‌شوند. در درخشندگی‌های بالا، احتمال وقوع چندین رویداد مجزا در یک برخورد دسته وجود دارد.

هدف اصلی از ارتقای HL-LHC^۳ افزایش درخشندگی یکپارچه به حدود 250 fb^{-1} در سال است. درخشندگی یکپارچه اندازه‌ای از تعداد کلیه برخوردهای ذرات در طول زمان است که برای کشف و تحقیق در فیزیک اطلاعات بیشتری فراهم می‌کند. به علت سطوح بالای درخشندگی، آشکارسازها با چالش عملياتی مواجه خواهند بود که در سطوح بسیار بالای پایل آپ عمل

¹phenomenological

²minimum bias

³High-Luminosity Large Hadron Collider

کنند (میانگین تعداد رویدادها در هر دسته عبوری). مشخص شده است که نه تنها پایل آپ تأثیری بر کارایی آشکارساز دارد، بلکه توزیع طولی آن در اطراف نقطه برخورد نیز تأثیرگذار است.

۱۰.۲ چگالی پایل آپ مؤثر

پایل آپ μ میانگین تعداد رویدادها در هر عبور از دسته است و مقدار آن در طول زمان تغییر می‌کند، چرا که درخشندگی، تغییر می‌کند،

$$\mu(t) = \frac{\sigma \cdot L(t)}{f \cdot n_b} \quad (8.2)$$

در اینجا، σ نشان‌دهنده سطح مقطع برخورد غیرکشسان پروتون-پروتون است، $L(t)$ درخشندگی آنی در زمان t است، n_b تعداد دسته‌های برخورده‌شونده، و f فرکانس چرخش LHC است. یک توصیف دقیق و بهروز از سطح پایل آپ در حالت‌های عملیاتی مختلف HL-LHC برای آزمایش‌ها بسیار مهم است: بازسازی^۱ و شناسایی^۲ نقاط اولیه^۳ رویدادها در یک برخورد، نقطه عطف بیشتر آنالیز هاست، و نرخ آن‌ها که به تعداد نقاط شبیه‌سازی شده بستگی دارد به پایل آپ وابسته است. یک شکل ساده شده از فرمول درخشندگی لحظه‌ای با فرض توزیع‌های گوسی برای دسته‌های عرضی، به شرح زیر است:

$$L = \frac{\gamma f n_b N \gamma}{4\pi \sigma^* \times \sigma_{||}^* \cos \theta} \iint \rho_s(s-z) \rho_s(s+z) K(s, z) ds dz \quad (9.2)$$

¹Reconstruction

²Identification

³Primary Vertices

$$K(s, z) = \frac{1}{\sqrt{1 + (s/\beta_{\times}^*)^2} \sqrt{1 + (s/\beta_{\parallel}^*)^2}}$$

$$\times \exp \left\{ -\frac{\left[(\theta/2)s - (1/k_{\times}) (\theta_{\times}/2) \sin(k_{\times}s) \cos(k_{\times}z) \right]^2}{\sigma_{\times}^{*2} \left[1 + (s/\beta_{\times}^*)^2 \right]} \right.$$

$$\left. - \frac{\left[(1/k_{\parallel}) (\theta_{\parallel}/2) \cos(k_{\parallel}s) \sin(k_{\parallel}z) \right]^2}{\sigma_{\parallel}^{*2} \left[1 + (s/\beta_{\parallel}^*)^2 \right]} \right\}$$

صفحه تلاقي \times و صفحه جدایي موازي \parallel هستند. و N تعداد پروتون ها در هر دسته است (فرض می شود برای تمام دسته ها در هر دو پرتو، یکسان است); θ زاويه کامل دو پرتو عبوری است؛ ρ_s چگالی طولي ذرات دو دسته بربور دكنته است. β_{\parallel}^* و β_{\times}^* توابع بتاترون و اندازه پرتو در نقطه تلاقي هستند، به ترتيب؛ و $\varepsilon_{\times, \parallel}$ انتشار ها در صفحه های تلاقي و جدایي هستند. آنتن های RF CC با زاويه ها و موج های $k_{\times}, k_{\parallel}$ مشخص می شوند، که $f_{\times, \parallel}$ فرکانس آنهاست. انتگرال ها بر روی مختصات طولي s و فاصله تا ذره مرجع z ، در دامنه های مربوطه شان گسترش می یابند؛ نقطه تلاقي در $s = 0$ قرار دارد.

میزان چگالی پایل آپ (s, t) ، توزيع محلی رویدادها در امتداد s را توصیف می کند:

$$\mu(t) = \int \rho(s, t) ds \quad (10.2)$$

توجه کنید که از روابط (۵) و (۶) مشخص می شود که $\frac{dL}{ds}(s, t) = \frac{fn_b\rho(s, t)}{\sigma}$ است. آشکارساز های ردیابی آزمایشگاهی حساس به پایل آپ خطی هستند. مقادير مورد توجه دیگري وجود دارد مثل پایل آپ بيشينه که معمولاً در نزديکی نقطه بربور داده (IP) واقع هستند.

$$\hat{\rho}(t) = \rho(s = 0, t) \quad (11.2)$$

و منطقه درخشش ميانگين مربعی $(rms)_{s, lum}(t)$ به صورت زير تعریف می شود:

$$\sigma_{s, lum}(t) = \sqrt{\frac{1}{\mu(t)} \int_0^{\infty} s^2 \rho(s, t) ds} \quad (12.2)$$

هر سناریوی عملیاتی یک چگالی محلی (و منطقه درخشش) با ویژگی‌های خاص خود را ارائه می‌دهد. مانند پایل‌آپ، مطالعه‌ی دقیقی از این مقادیر برای ارائه‌ی ورودی برای شبیه‌سازی‌های دقیق عملکرد آشکارساز برای تنظیمات مختلف آشکارساز لازم است.

وضوح طولی ردیاب (حدود ۰.۱ میلیمتر) به طور قابل توجهی کوچکتر از میانگین فاصله بین رأس‌های پایل‌آپ است، بنابراین منطقی است که تأثیر چگالی پایل‌آپ بر کارایی آشکارساز را در محدوده‌ی فعلی برخورد، خطی فرض کنیم. چندین مقدار به عنوان تابعی از چگالی پایل‌آپ به طور اولیه مورد مطالعه قرار گرفته است؛ مشخص شده است که بازده محلی رویدادهای بازسازی شده یک رابطه تقریباً خطی با چگالی پایل‌آپ ارائه می‌دهد، این فرضیه ساده‌اما نوین را تأیید می‌کند. فرض خطی تأثیر چگالی پایل‌آپ بر بازده آشکارساز امکان تعریف یک چگالی پایل‌آپ موثر را فراهم می‌کند - همانطور که در زیر تعریف می‌شود - برای اندازه‌گیری تأثیر یکپارچه‌ای که به مراتب جذاب‌تر از قله است: این شاخص عملکرد درخشنده‌ی را تکمیل می‌کند، به جای قله درخشنده‌ی، به عنوان یک پارامتری که برای آزمایش‌ها مورد توجه است.

بر اساس نتایج توصیف شده، پیشنهاد شده که استفاده از مقدار داده‌های مفید برای تجزیه و تحلیل، که توسط $L_{\text{int}}^{\text{fill}}$ نمایان می‌شود، به عنوان یک بخش از درخشنده‌ی یکپارچه تحویل داده شده $L_{\text{int}} = \int_0^t L(t) dt$ برای مدت زمان پر شدن t_{fill} ، از طریق یک تابع f از μ و چگالی محلی ρ پارامتری‌سازی شود.

$$\mathcal{L}'_{\text{int}} = \int_0^{t_{\text{fill}}} \int f(\mu, \rho) \mathcal{L}(s, t) ds dt \quad (13.2)$$

بسط f حول \circ ، $\rho = \circ$

$$f(\mu, \rho) \approx f(\mu, \circ) + \frac{\partial f(\mu, \rho)}{\partial \rho} \Bigg|_{\rho=\circ} \rho(s, t) \quad (14.2)$$

در ثابت \circ ، که در HL-LHC برای بیشتر مدت زمان پر شدن اتفاق می‌افتد، انتگرال بر روی اولین عبارت می‌تواند به شکل تقریبی تخمین زده شود.

$$\int_0^{t_{\text{fill}}} \int f(\mu(t), \circ) \mathcal{L}(s, t) ds dt \approx f(\mu \circ) \mathcal{L}_{\text{int}} \quad (15.2)$$

و

$$\int_0^{t_{\text{fill}}} \int \frac{\partial f(\mu \circ, \rho)}{\partial \rho} \Bigg|_{\rho=\circ} \rho(s, t) \mathcal{L}(s, t) ds dt \approx -\varepsilon \int_0^{t_{\text{fill}}} \int \rho(s, t) \mathcal{L}(s, t) ds dt \equiv -\varepsilon \bar{\rho} \mathcal{L}_{\text{int}} \quad (16.2)$$

حایی که

$$\varepsilon \equiv - \left. \frac{\partial f(\mu_0, \rho)}{\partial \rho} \right|_{\rho=0} \quad (17.2)$$

می‌توانیم درخشنده‌گی یکپارچه را در معادله ۱۳.۲ بصورت زیر بازنویسی کنیم:

$$\mathcal{L}'_{int} \approx [f(\mu_0) - \varepsilon(\mu_0) \bar{\rho}] \mathcal{L}_{int} \quad 0 \leq f(\mu_0) \leq 1 \quad \varepsilon > 0 \quad (18.2)$$

و از این رو، این امکان را به ما می‌دهد تا راه‌های مختلف را از نظر درخشنده‌گی یکپارچه و $\bar{\rho}$ برای پایه‌گذاری ثابت μ_0 مقایسه کنیم. معادله ۱۶.۲ چگالی پایل آپ (خطی) موثر $\bar{\rho}$ میانگین چگالی پایل آپ در طول پر شدن، که توسط درخشنده‌گی یکپارچه (یا به طور معادل، پایل آپ کل در طول پر شدن) وزن داده شده است را تعریف می‌کند.

$$\bar{\rho} = \frac{\int_0^t \text{fill} \int \rho^2(s, t) ds dt}{\int_0^t \text{fill} \mu(t) dt} \quad (19.2)$$

واحدهای $\bar{\rho}$ میلی‌متر هستند، به معنای رویدادها در هر میلی‌متر. هر چه تراکم موثر پایین‌تر باشد، کارایی آشکارساز برای بازسازی راس رویداد (به عبارت دیگر، کیفیت داده بالاتر) افزایش می‌یابد. باید توجه داشت که برای دیگر مقادیر آزمایشی، مانند آن‌هایی که به ردیاب وابسته نیستند، فقط سطح پایل آپ مربوط است.

به طور کلی، چگالی پایل آپ یکتابع پیچیده از پارامترهای پرتو و اپتیکی در نقطه برخورد است. با این حال، توصیف گوسی

با نقطه اوج $\hat{\rho}(t)$ و مقدار $\sigma_{s, \text{lum}}(t)$ که توسط $\sigma_{s, \text{lum}}$ داده شده است،

$$\rho(s, t) = \hat{\rho}(t) \exp \left(-\frac{s^2}{2\sigma_{s, \text{lum}}^2(t)} \right) \quad (20.2)$$

اغلب حتی زمانی که شبیه‌سازی‌های واقع‌گرایانه را در نظر بگیریم، تقریب خوبی است، همانطور که در شکل ۶.۲ دیده می‌شود.

در این صورت، معادله ۱۹.۲ می‌تواند ساده‌سازی شود به

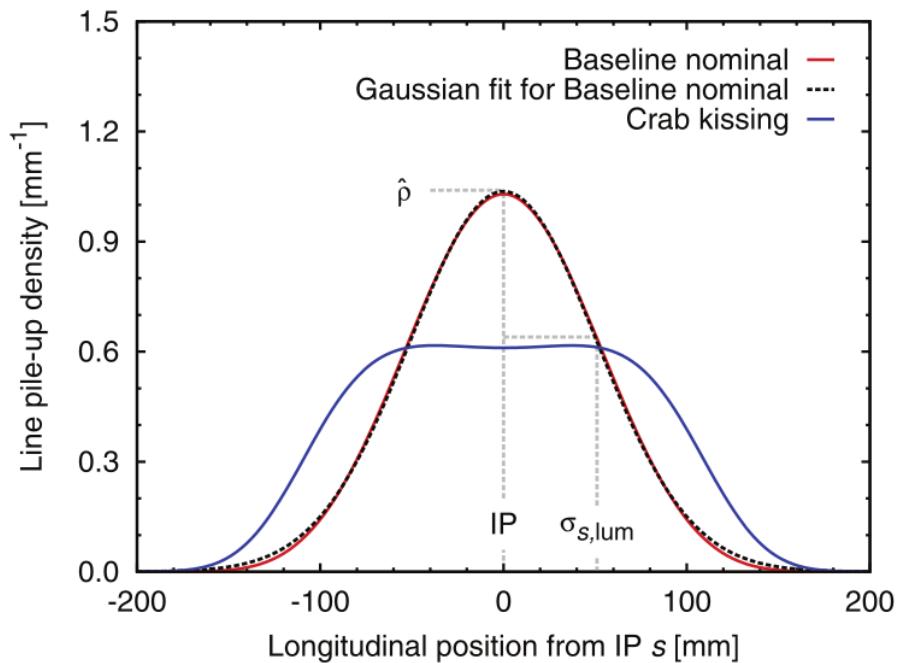
$$\bar{\rho}_{\text{Gauss}} = \frac{1}{\sqrt{2}} \frac{\int_0^t \text{fill} \hat{\rho}^2(t) dt}{\int_0^t \text{fill} \hat{\rho}(t) dt} \quad (21.2)$$

یک توزیع رویداد مسطح، به عنوان مثال در شکل ۶.۲ درآغاز پرشدن مشاهده می‌شود.

برای مشتق تراکم موثر به طور تحلیلی، ما یکتابع مستطیلی (جعبه) با عرض کامل $\Delta(t)$ را در نظر می‌گیریم،

$$\rho(s, t) = \hat{\rho}(t) \text{box}[s, \Delta(t)] \quad (22.2)$$

سپس چگالی موثر پایل آپ توسط معادله زیر داده می‌شود،



شکل ۲۴.۲: نمودار چگالی پایل آپ در ابتدای پر شدن برای مبنای HL-LHC (خط قرمز) و تطابق گوسی آن (خط نقطه‌ای سیاه).

$$\bar{\rho}_{\text{flat}} = \frac{\int_0^t \text{fill} \hat{\rho}(t) dt}{\int_0^t \text{fill} \hat{\rho}(t) dt} \quad (24.2)$$

از مقایسه‌ی ۲۱.۲ و ۲۳.۲، مشاهده می‌شود که $\hat{\rho}$ یک شاخص قوی نیست و برای مواردی که توزیع رویدادها به طور تقریزی گوسی است با فاکتور $\sqrt{2}$ بدینانه است، مانند پایه‌ی HL-LHC. به همین دلیل، $\bar{\rho}$ معرفی می‌شود تا چگونگی عملکرد آشکارساز را در ثابت پایل آپ ارزیابی کنند.

۳.۲ احتمالات پایل آپ و رویدادها در هر دسته عبوری

تعداد کل رویدادها در هر دسته عبوری از یک توزیع پواسون پیروی می‌کند.

$$P(N_t) = \frac{\nu_t^{N_t}}{N_t!} e^{-\nu_t} \quad (24.2)$$

با میانگین

$$\nu_t = \langle N_t \rangle = \sigma_t \mathcal{L} \langle \Delta t_{\text{bunch}} \rangle, \quad (25.2)$$

که فاصله میانگین گروه‌ها توسط

$$\langle \Delta t_{\text{bunch}} \rangle = \frac{1}{f_{LHC} k}. \quad (26.2)$$

تعیین می‌شود. فرکانس چرخش LHC برابر است با $f_{LHC} = 11.245 \text{ kHz}$. برای سطح مقطع کلی ما فرض می‌کنیم $\sigma_t = 110 \text{ mb}$. درخشنده‌گی \mathcal{L} و تعداد دسته‌ها k به شرایط اجرایی وابسته‌اند. ما دو حالت را در نظر می‌گیریم:

• یک سناریو معمولی $\mathcal{L}/k = 3.6 \times 10^{29} \text{ cm}^{-2} \text{ s}^{-1}$ و $\mathcal{L} = 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$: $\beta^* = 0.5 \text{ m}$

در این حالت، $\langle \Delta t_{\text{bunch}} \rangle = 31.67 \text{ ns}$ (نه حداقل فاصله دسته ۲۵ ns!). این باعث می‌شود $\nu_t = 3.5$ رویدادها

در هر دسته عبوری شود.

برخی از مقادیر توزیع برای $P(N_t)$ در جدول ۱۰.۲ آورده شده است. ما به این نتیجه می‌رسیم که احتمال مشاهده حداقل

۱ رویداد برابر با $1 - 0.03 = 97\%$ است، و احتمال پایل آپ $P(N_t \geq 2) = 86.4\%$ است.

جدول ۱۰.۲ مقادیر توزیع برای $P(N_t)$ را نشان می‌دهد.

جدول ۱۰.۲

n	$P(N_t = n)$	$P(N_t > n)$
۰	۰.۱۰۶	۰.۸۹۳
۱	۰.۱۸۵	۰.۸۱۵
۲	۰.۲۱۶	۰.۷۸۳
۳	۰.۱۸۹	۰.۷۱۱
۴	۰.۱۳۲	۰.۸۶۸
۵	۰.۱۴۲	۰.۸۵۶

• یک سناریو معمولی $\mathcal{L}/k = 1.9 \times 10^{28} \text{ cm}^{-2} \text{ s}^{-1}$ و $\mathcal{L} = 3 \times 10^{30} \text{ cm}^{-2} \text{ s}^{-1}$: $\beta^* = 90 \text{ m}$

در این حالت، $\langle \Delta t_{\text{bunch}} \rangle = 570 \text{ ns}$ و در هر برخورد دسته، میانگین $\nu_t = 188$ رویداد مشاهده می‌شود.

جدول ۲۰.۲ توزیع پواسون برای تعداد کل رویدادها در هر دسته عبوری را برای $\mathcal{L}/k = 1.9 \times 10^{28} \text{ cm}^{-2} \text{ s}^{-1}$

نشان می‌دهد.

جدول ۲۰.۲

n	$P(N_t = n)$	$P(N_t > n)$
۰	۰.۸۲۸۵	۰.۱۷۱۵
۱	۰.۱۵۵۸	۰.۸۴۴۲
۲	۰.۱۴۷	۰.۸۵۳۱
۳	۰.۰۰۰۹	۰.۹۹۹۱

۴.۲ احتمال پایل آپ برای کلاس‌های رویداد خاص

سوال: احتمال مشاهده در یک دسته عبوری با تعداد کل N_t عضو از کلاس A، N_B از کلاس B و ...

هستند چیست؟ سطح مقطع کل را به شکل زیر تجزیه می‌کنیم:

$$\sigma_t = \sigma_A + \sigma_B + \dots + \sigma_R \quad (27.2)$$

یا، از لحاظ تعداد رویدادها،

$$N_t = N_A + N_B + \dots + N_R \quad (28.2)$$

جایی که کلاس رویداد R برای بقیه نا مشخص است. در ادامه ما دو مورد ساده‌تر با ۱ و ۲ کلاس رویداد مشخص را در نظر

می‌گیریم.

۱۰.۲ یک کلاس رویداد مشخص

سوال ساده شده به این شکل است: احتمال مشاهده در یک دسته عبوری تعداد کل N_t رویداد است که N_A از نوع

هستند، به عنوان مثال رویدادهای پراکندگی کشسان؟ تجزیه تعداد رویدادها اکنون به شکل زیر است:

$$N_t = N_A + N_R \quad (29.2)$$

مقدار مورد نیاز احتمال دو بعدی $P(N_A, N_t)$ است. از مستقل بودن کلاس‌های A و R به صورت مستقیم نتیجه می‌شود:

$$P(N_A, N_t) = \frac{\nu_A^{N_A}}{N_A!} e^{-\nu_A} \frac{(\nu_t - \nu_A)^{N_t - N_A}}{(N_t - N_A)!} e^{-(\nu_t - \nu_A)} \quad (30.2)$$

رویکرد دیگری - که به طور رسمی مستقل بودن A و R را تأیید می‌کند - این است که $P(N_A, N_t)$ را به عنوان زیر بیان

کنیم:

$$P(N_A, N_t) = P(N_A|N_t)P(N_t) \quad (31.2)$$

جایی که $P(N_t)$ توسط معادله ۲۴.۲ داده شده است، و احتمال شرطی $P(N_A|N_t)$ توزیع دو جمله‌ای را دنبال می‌کند:

$$P(N_A|N_t) = \binom{N_t}{N_A} \left(\frac{\nu_A}{\nu_t} \right)^{N_A} \left[1 - \frac{\nu_A}{\nu_t} \right]^{N_t - N_A} \quad (32.2)$$

توجه کنید که در اینجا پیش‌نیازهای تقریب پواسون برآورده نمی‌شوند: N_t بسیار بزرگ نیست و $\frac{\nu_A}{\nu_t}$ به طور ضروری بسیار کوچک نیست (برای مثال اگر A برای پراکندگی کشسان باشد). با ترکیب معادله‌های ۲۴.۲ و ۳۲.۲ در معادله ۳۱.۲ به دست می‌آوریم:

$$P(N_A, N_t) = \binom{N_t}{N_A} \left(\frac{\nu_A}{\nu_t} \right)^{N_A} \left[1 - \frac{\nu_A}{\nu_t} \right]^{N_t - N_A} \frac{\nu_t^{N_t}}{N_t!} e^{-\nu_t} \quad (33.2)$$

که معادل با ۳۰.۲ است.

به عنوان مثال، جداول ۳ و ۴ برای سناریوهای معمولی $m^{\circ.5}$ و $m^{\circ.9}$ نمایش داده شده‌اند، در صورتی که A برای پراکندگی کشسان باشد.

$P(N_A)$	N_t						N_A
	۵	۴	۳	۲	۱	۰	
۰.۳۸۷	۰.۰۲۷	۰.۰۵۳	۰.۰۸۳	۰.۰۹۸	۰.۰۷۷	۰.۰۳۰	۰
۰.۳۶۸	۰.۰۵۱	۰.۰۸۰	۰.۰۹۳	۰.۰۷۳	۰.۰۲۹	۰	۱
۰.۱۷۵	۰.۰۳۸	۰.۰۴۵	۰.۰۳۵	۰.۰۱۴	۰	۰	۲
۰.۰۵۵	۰.۰۱۴	۰.۰۱۱	۰.۰۰۴۴	۰	۰	۰	۳
۰.۰۱۳	۰.۰۰۳	۰.۰۰۱	۰	۰	۰	۰	۴

جدول ۳.۲ برای A کشسان با فرض $P(N_A, N_t) : ۳.۶ \times \sigma_{\text{elastic}} = ۳.۰ \text{ mb}$ ، برای $\mathcal{L}/k = ۳.۰$ cm $^{-2}$ s $^{-1}$. $\left(\nu_t = ۳.۰, \nu_A = ۰.۹۵, \frac{\nu_A}{\nu_t} = ۰.۲۷۳, 1 - \frac{\nu_A}{\nu_t} = ۰.۷۲۷ \right)$

$P(N_A)^{*2}$	N_t				N_A^{*2}
	۳	۲	۱	۰	
۹۴۹۹.۰	۰۰۰۳۵.۰	۰۰۷۷.۰	۱۱۳۳.۰	۸۲۸۰.۰	۰
۰۴۸۸.۰	۰۰۰۴۰.۰	۰۰۵۸.۰	۰۴۲۶.۰	۰	۱
۰۰۱۳.۰	۰۰۰۱۵.۰	۰۰۱۱.۰	۰	۰	۲
۲.۱×۱۰^{-۵}	۱.۸×۱۰^{-۵}	۰	۰	۰	۳

جدول ۴.۲ برای A کشسان با فرض $P(N_A, N_t) : ۴.۲ \times \sigma_{\text{elastic}} = ۳.۰ \text{ mb}$ ، برای $\mathcal{L}/k = ۱.۹$ cm $^{-2}$ s $^{-1}$. $\left(\nu_t = ۰.۱۸۸, \nu_A = ۰.۰۵۱, \frac{\nu_A}{\nu_t} = ۰.۲۷۳, 1 - \frac{\nu_A}{\nu_t} = ۰.۷۲۷ \right)$

حد خاص:

چه احتمالی برای مشاهده N_t رویداد وجود دارد که یکی از آن‌ها از نوع A بسیار کمیاب است، یعنی $N_A = ۱$ و $0 \rightarrow ? \frac{\nu_A}{\nu_t}$

معادله ۳۲.۲ حالا به صورت زیر کاوش می‌یابد:

$$P(N_A = 1 | N_t) = N_t \frac{\nu_A}{\nu_t} \left[1 - \frac{\nu_A}{\nu_t} \right]^{N_t-1} \rightarrow N_t \frac{\nu_A}{\nu_t},$$

و نتیجه ما است:

$$P(N_A = 1, N_t) \rightarrow N_t \frac{\nu_A}{\nu_t} \frac{\nu_t^{N_t}}{N_t!} e^{-\nu_t} = \nu_A \frac{\nu_t^{N_t-1}}{(N_t-1)!} e^{-\nu_t}$$

۲۰.۴.۲ دو کلاس رویداد مشخص

احتمال مشاهده کلی در یک دسته عبوری از N_t رویداد که N_A تعلق به کلاس A و N_B به کلاس B دارد چقدر است؟

مشابه قسمت قبل، $P(N_A, N_B, N_t)$ می‌تواند به دو صورت نوشته شود:

$$P(N_A, N_B, N_t) = \frac{\nu_A^{N_A}}{N_A!} e^{-\nu_A} \frac{\nu_B^{N_B}}{N_B!} e^{-\nu_B} \frac{(\nu_t - \nu_A - \nu_B)^{N_t - N_A - N_B}}{(N_t - N_A - N_B)!} e^{-(\nu_t - \nu_A - \nu_B)}$$

یا به صورت

$$\begin{aligned} P(N_A, N_B, N_t) &= \binom{N_A + N_B}{N_A} \left(\frac{\nu_A}{\nu_A + \nu_B} \right)^{N_A} \left[1 - \frac{\nu_A}{\nu_A + \nu_B} \right]^{N_B} \\ &\times \binom{N_t}{N_A + N_B} \left(\frac{\nu_A + \nu_B}{\nu_t} \right)^{N_A + N_B} \left[1 - \frac{\nu_A + \nu_B}{\nu_t} \right]^{N_t - N_A - N_B} \\ &\times \frac{\nu_t^{N_t}}{N_t!} e^{-\nu_t} \end{aligned}$$

جداول ۵ و ۶ مورد A کشسان و B غیرکشسان برای دو سناریو مدنظر نشان می‌دهند.

جدول ۲ برای A کشسان، برای B غیرکشسان، (یعنی، $R = SD$) با فرض $\mathcal{L}/k = ۳.۶ \times ۱۰^{۲۹} \text{ cm}^{-۲} \text{ s}^{-۱}$ ، برای غیرکشسان $\sigma_{NSD} = ۷^{\circ} \text{ mb}$ و $\sigma_{\text{elastic}} = ۳^{\circ} \text{ mb}$ ، $\nu_t = ۳.۵$ ، $\nu_{SD} = ۰.۳۳$ ، $\nu_{NSD} = ۲.۲۲$ ، $\nu_{\text{elastic}} = ۰.۹۵$.

۵	۴	۳	۲	۱	۰	N_A, N_B
۸×۱۰^{-۷}	۲×۱۰^{-۵}	۰.۰۰۰۲	۰.۰۰۲	۰.۰۱۰	۰.۰۳۰	۰,۰
۴×۱۰^{-۵}	۰.۰۰۰۴	۰.۰۰۴	۰.۰۲۲	۰.۰۶۷	۰	۱,۰
۲×۱۰^{-۵}	۰.۰۰۰۲	۰.۰۰۲	۰.۰۱۰	۰.۰۲۹	۰	۰,۱
۰.۰۰۰۴	۰.۰۰۴	۰.۰۲۵	۰.۰۷۵	۰	۰	۲,۰
۰.۰۰۰۴	۰.۰۰۳	۰.۰۲۱	۰.۰۶۴	۰	۰	۱,۱
۸×۱۰^{-۵}	۰.۰۰۰۷	۰.۰۰۵	۰.۰۱۴	۰	۰	۰,۲
۰.۰۰۳	۰.۰۱۸	۰.۰۵۵	۰	۰	۰	۳,۰
۰.۰۰۴	۰.۰۲۳	۰.۰۷۱	۰	۰	۰	۲,۱
۰.۰۰۲	۰.۰۱۰	۰.۰۳۰	۰	۰	۰	۱,۲
۰.۰۰۰۲	۰.۰۰۱	۰.۰۰۴	۰	۰	۰	۰,۳
۰.۰۱۰	۰.۰۳۱	۰	۰	۰	۰	۴,۰
۰.۰۱۷	۰.۰۵۳	۰	۰	۰	۰	۳,۱
۰.۰۱۱	۰.۰۳۴	۰	۰	۰	۰	۲,۲
۰.۰۰۳	۰.۰۱۰	۰	۰	۰	۰	۱,۳
۰.۰۰۰۳	۰.۰۰۱	۰	۰	۰	۰	۰,۴

۳	۲	۱	۰	N_A, N_B
۷×۱۰^{-۷}	۰.۰۰۰۱۲	۰.۰۱۴۰	۰.۸۲۸۵	۰,۰
۱×۱۰^{-۵}	۰.۰۰۱۷	۰.۰۹۹۳	۰	۱,۰
۶×۱۰^{-۶}	۰.۰۰۰۷۲	۰.۰۴۲۵	۰	۰,۱
۱×۱۰^{-۴}	۰.۰۰۵۹	۰	۰	۲,۰
۹×۱۰^{-۵}	۰.۰۰۵۱	۰	۰	۱,۱
۲×۱۰^{-۵}	۰.۰۰۱۱	۰	۰	۰,۲
۲.۴×۱۰^{-۴}	۰	۰	۰	۳,۰
۳.۱×۱۰^{-۴}	۰	۰	۰	۲,۱
۱.۳×۱۰^{-۴}	۰	۰	۰	۱,۲
۲×۱۰^{-۵}	۰	۰	۰	۰,۳

جدول ۳ برای A کشسان، برای B غیرکشسان، (یعنی، $R = SD$) با فرض $\nu_{\text{elastic}} = \nu_t = ۰.۱۸۸$ ، $\mathcal{L}/k = ۱.۹ \times ۱۰^{۲۸} \text{ cm}^{-۲} \text{ s}^{-۱}$ ، $\sigma_{NSD} = ۷^{\circ} \text{ mb}$ و $\sigma_{\text{elastic}} = ۳^{\circ} \text{ mb}$ ، $\nu_{SD} = ۰.۰۱۷$ ، $\nu_{NSD} = ۰.۱۲۰$ ، $\nu_{\text{elastic}} = ۰.۰۵۱$

فرض کنید که ما رویدادهای کلاس A را (که حتی می‌تواند یک مقدار مشخص‌کننده حداقل بیاس را به عنوان یک حالت خاص نشان دهد) فعال می‌کنیم. در یک تلاقی می‌تواند حداکثر یک مقدار مشخص‌کننده وجود داشته باشد، حتی اگر چند رویداد از نوع A رخ داده باشد. بنابراین، شرط مقدار مشخص‌کننده را می‌توان به صورت $1 \geq N_A \geq 1$ نوشت. توزیع احتمال و میانگین تعداد کل رویدادها در یک تلاقی تحت شرط دادن یک مقدار مشخص‌کننده چیست؟

۳۴.۲ توزیع احتمال

توزیع احتمال شرطی که به دنبال آن هستیم به صورت رسمی به شکل زیر بیان می‌شود:

$$P(N_t | N_A \geq 1) = \frac{P(N_t, N_A \geq 1)}{P(N_A \geq 1)} \quad (34.2)$$

از آنجا که $P(N_A)$ یک توزیع پواسون را دنبال می‌کند، مخرج به شکل زیر می‌شود:

$$P(N_A \geq 1) = 1 - e^{-\nu_A} \quad (35.2)$$

که در صورت واجب با استفاده از معادلات ۳۱.۲ و ۳۲.۲ حل می‌شود:

$$\begin{aligned} P(N_t, N_A \geq 1) &= P(N_t) \sum_{N_A=1}^{N_t} \binom{N_t}{N_A} \left(\frac{\nu_A}{\nu_t}\right)^{N_A} \left[1 - \frac{\nu_A}{\nu_t}\right]^{N_t-N_A} \\ &= P(N_t) \left[\sum_{N_A=0}^{N_t} \binom{N_t}{N_A} \left(\frac{\nu_A}{\nu_t}\right)^{N_A} \left(1 - \frac{\nu_A}{\nu_t}\right)^{N_t-N_A} - \left(1 - \frac{\nu_A}{\nu_t}\right)^{N_t} \right] \\ &= P(N_t) \left[1 - \left(1 - \frac{\nu_A}{\nu_t}\right)^{N_t} \right] \end{aligned} \quad (36.2)$$

در نهایت

$$\begin{aligned} P(N_t | N_A \geq 1) &= P(N_t) \frac{1 - \left(1 - \frac{\nu_A}{\nu_t}\right)^{N_t}}{1 - e^{-\nu_A}} \\ &= \frac{\nu_t^{N_t}}{N_t!} e^{-\nu_t} \frac{1 - \left(1 - \frac{\nu_A}{\nu_t}\right)^{N_t}}{1 - e^{-\nu_A}} \end{aligned} \quad (37.2)$$

باید حالا دو حالت خاص را مورد بررسی قرار دهیم:

- تریگر روی کمترین حالت : در این حالت، $\nu_t = \nu_A$. بنابراین

$$P(N_t | N_t \geq 1) = \frac{\nu_t^{N_t}}{N_t!} \frac{e^{-\nu_t}}{1 - e^{-\nu_t}} \quad (38.2)$$

- تریگر روی یک کلاس رویداد بسیار کم؛ در این حد، $1 \ll \nu_A$.

$$\begin{aligned} P(N_t | N_A \geq 1) &\rightarrow \frac{\nu_t^{N_t}}{N_t!} e^{-\nu_t} \frac{1 - \left(1 - N_t \frac{\nu_A}{\nu_t}\right)}{1 - 1 + \nu_A} \\ &= \frac{\nu_t^{N_t-1}}{(N_t-1)!} e^{-\nu_t} \\ &= P(N_t - 1) \end{aligned} \quad (39.2)$$

میانگین تعداد رویدادها

میانگین تعداد رویدادها در یک تلاقی گروهی که یک تریگر مشاهده شده است، توسط

$$\langle N_t \rangle_{N_A \geq 1} = \sum_{N_t=0}^{\infty} N_t P(N_t | N_A \geq 1) \quad (40.2)$$

استفاده از ۳۷.۲ برای حالت عمومی می‌دهد

$$\begin{aligned} \langle N_t \rangle_{N_A \geq 1} &= \frac{\sum_{N_t=0}^{\infty} N_t \frac{\nu_t^{N_t}}{N_t!} e^{-\nu_t} \left[1 - \left(1 - \frac{\nu_A}{\nu_t}\right)^{N_t} \right]}{1 - e^{-\nu_A}} \\ &= \nu_t \frac{\sum_{N_t=1}^{\infty} \frac{\nu_t^{N_t-1}}{(N_t-1)!} e^{-\nu_t} \left[1 - \left(1 - \frac{\nu_A}{\nu_t}\right)^{N_t-1} \left(1 - \frac{\nu_A}{\nu_t}\right) \right]}{1 - e^{-\nu_A}} \\ &= \nu_t \frac{\sum_{N_t=0}^{\infty} \frac{\nu_t^{N_t}}{N_t!} e^{-\nu_t} - \left(1 - \frac{\nu_A}{\nu_t}\right) \sum_{N_t=0}^{\infty} \frac{\nu_t^{N_t}}{N_t!} e^{-\nu_t} \left(1 - \frac{\nu_A}{\nu_t}\right)^{N_t}}{1 - e^{-\nu_A}} \\ &= \frac{1 - \left(1 - \frac{\nu_A}{\nu_t}\right) e^{-\nu_A} \sum_{N_t=0}^{\infty} \frac{\left[\nu_t \left(1 - \frac{\nu_A}{\nu_t}\right)\right]^{N_t}}{N_t!} e^{-\nu_t \left(1 - \frac{\nu_A}{\nu_t}\right)}}{1 - e^{-\nu_A}} \\ &= \nu_t \frac{1 - \left(1 - \frac{\nu_A}{\nu_t}\right) e^{-\nu_A}}{1 - e^{-\nu_A}} \\ &= \nu_t \left[1 + \frac{\nu_A}{\nu_t} \frac{e^{-\nu_A}}{1 - e^{-\nu_A}} \right] \end{aligned} \quad (41.2)$$

دوباره، ما این فرمول را با دو حالت خاص مانند قسمت قبل توضیح می‌دهیم.

- تریگر حداقل بار؛ در این حالت، $\nu_t = \nu_A$. بنابراین

$$\langle N_t \rangle_{N_t \geq 1} = \nu_t \left[1 + \frac{e^{-\nu_t}}{1 - e^{-\nu_t}} \right] \quad (42.2)$$

برای $\nu_t = 3.5$ ، ما داریم

$$\langle N_t \rangle_{N_t \geq 1} = 3.5 \left(1 + \frac{0.030}{0.970} \right) = 3.6, \quad (43.2)$$

و برای $\nu_t = 0.188$

$$\langle N_t \rangle_{N_t \geq 1} = 0.188 \left(1 + \frac{0.829}{0.171} \right) = 1.097. \quad (44.2)$$

- تریگر کردن روی یک کلاس حالت بسیار کم: در این حد بسیار کوچک است، $1 \ll \nu_A$.

$$\begin{aligned} \langle N_t \rangle_{N_A \geq 1} &\rightarrow \nu_t \left[1 + \frac{\nu_A}{\nu_t} \frac{1 - \nu_A}{1 - 1 + \nu_A} \right] \\ &= \nu_t + 1 - \nu_A \end{aligned} \quad (45.2)$$

$$\rightarrow \nu_t + 1$$

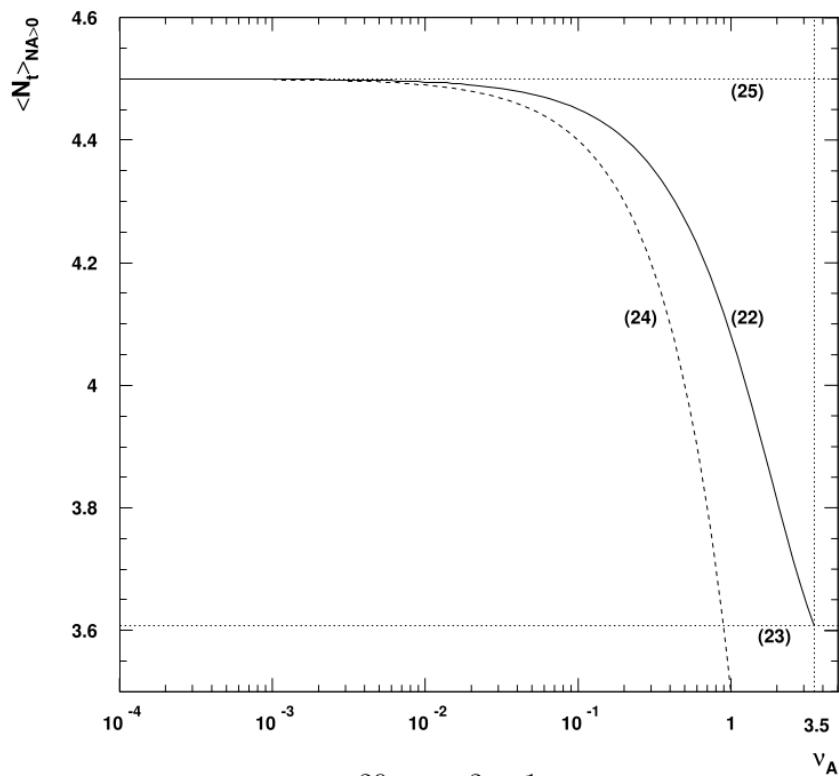
شکل ۷.۲ را به عنوان تابعی از ν_A برای $\nu_t = 3.5$ در حالت کلی و در تقریب برای ν_A کوچک نشان می‌دهد. توجه کنید که در موارد پراش کشسان ($\nu_A = 0.33$) و پراش تک مسیره ($\nu_A = 0.95$) باید از فرمول کلی استفاده شود در حالی که برای تبادل دوگانه پومرون ($\sigma_{DPE} = 1 \text{ mb}$) با فرض $\nu_A = 0.03$ هر دو تقریب ۲۴.۲ و ۲۵.۲ در $\pm 33\%$ از مقدار واقعی است.

شکل ۸.۲ نشان می‌دهد همان کار را برای $\nu_t = 0.188$ انجام می‌دهد.

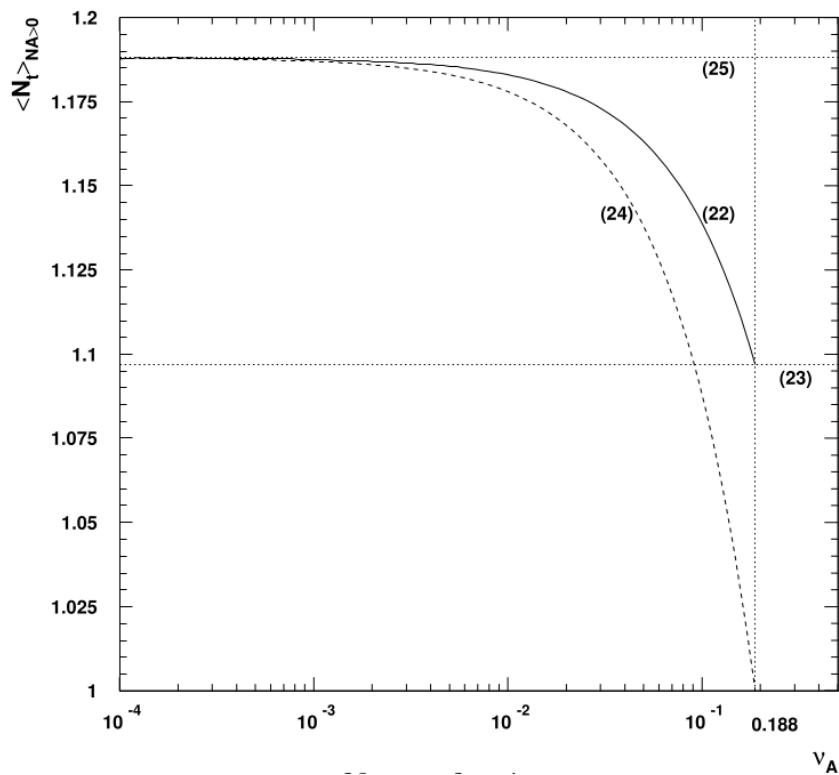
۵.۲ برخورد

- هدف این است که اندازه شعاع را در نقطه تلاقی به حداقل اندازه فشرده کنیم تا احتمال برخورد بیشتر شود.
- با این حال پروتون‌ها بسیار کوچکی هستند.
- بنابراین، حتی اگر ۱۰۰۰۰۰ میلیون پروتون در هر دسته را به ۶۴ میکرون (حدود عرض موی انسان) در نقطه برخورد فشرده کنیم، تنها حدود ۲۰ برخورد در هر تلاقی با جریان‌های پرتو^۱ داریم.
- دسته‌ها (هر ۲۵ نانوثانیه) عبور می‌کنند، بنابراین در ابتدای پر کردن با جریان مورد انتظار، حدود ۶۰۰ میلیون برخورد در ثانیه داریم.

¹Nominal Beam Currents



شکل ۷.۲: $\mathcal{L}/k = 3.6 \times 10^{24} \text{ cm}^{-2} \text{ s}^{-1}$ برای $(\nu_t = 3.5)$ به عنوان یک تابع از v_A برای حالت کلی و در تقریب برای v_A کوچک.



شکل ۸.۲: $\mathcal{L}/k = 1.9 \times 10^{25} \text{ cm}^{-2} \text{ s}^{-1}$ برای $(\nu_t = 0.188)$ به عنوان یک تابع از v_A برای حالت کلی و در تقریب برای v_A کوچک.

- اکثر پروتون‌ها از یکدیگر بدون برخورد عبور می‌کنند و دور حلقه ساعت‌ها در حال چرخیدن هستند.

۱۰.۲ نرخ برخورد

- به اصطلاح، نرخ رویداد = درخشندگی × سطح مقطع.
- سطح مقطع کلی پروتون-پروتون در $7 \text{ تراالکترون ولت حدود } 11^{\circ} \text{ میکروبarn}$ است. این مقدار کلی می‌تواند از دستاوردها مانند:

- غیرکشسان ($\sigma_{\text{in}} = 60 \text{ میکروبarn}$)

- تک انحرافی ($\sigma_{\text{sd}} = 12 \text{ میکروبarn}$)

- کشسان ($\sigma_{\text{el}} = 40 \text{ میکروبarn}$)

- عرض مقطع حاصل از پراکنش کشسان پروتون‌ها و رویدادهای تک انحرافی توسط آشکارسازها مشاهده نخواهد شد زیرا فقط پراکنش‌های غیرکشسان باعث تولید ذرات با زاویه‌های کافی بالا نسبت به محور شعاع می‌شود.
- به اصطلاح، نرخ رویداد غیرکشسان در درخشندگی نامی بنابراین $= 60^{\circ} \times 10^{-24} \times 10^{34} \times 60 \times 10^3 \text{ میلیون در ثانیه برای هر آزمایش با درخشندگی بالا است - حدود } 19 \text{ رویداد غیرکشسان در هر تلاقی.}$
- فاصله دسته‌ها در 25 نانوثانیه است، با این حال، فاصله‌های بزرگتری وجود دارد (برای اجازه دادن به زمان دهنده دامپ و غیره).
- یک شعاع 25 نانوثانیه از ما نرخ تلاقی اوجی 40 مگاهرتز می‌دهد. به دلیل وجود فاصله‌ها، میانگین نرخ تلاقی = تعداد دسته‌ها × فرکانس چرخش $= 31.6 = 11245 \times 11245 \times 2808 \times 10^8 \text{ مگاهرتز است. ضرب در } 19 \text{ رویداد در هر تلاقی در درخشندگی نامی به ما می‌دهد } 600 \text{ میلیون رویداد غیرکشسان در ثانیه.}$

فصل سوم

شبیه‌سازی پایل‌آپ در CMS

۱.۳ نحوه شبیه‌سازی پایل‌آپ

در LHC، برخوردهای زیادی به طور همزمان در یک برخورد دسته پروتون-پروتون رخ می‌دهد. این برهمکنش‌های اضافی، فراتر از برهمکنش "اصلی"^۱ از برخورد دسته فعلی، "پایل‌آپ همزمان"^۲ نامیده می‌شوند. پاسخ آشکارساز ممکن است تحت تأثیر برهمکنش‌های همزمان از برخورد دسته فعلی و همچنین برخوردهای دسته‌های مجاور قرار گیرد. برهمکنش‌های ناشی از برخوردهای دسته‌های مجاور، "پایل‌آپ ناهمزمان"^۳ نامیده می‌شوند. برخوردهای پایل‌آپ در شبیه‌سازی CMS مدل‌سازی می‌شوند.^[۶]

شبیه‌سازی مونت کارلو پایل‌آپ^۴ با استفاده از کد و فایل‌های پیکربندی در مأذول SimGeneral/MixingModule نرم‌افزار CMSSW کنترل می‌شود. شبیه‌سازی به شرح زیر پیش می‌رود:

۱. توزیعی از میانگین تعداد برخوردها به ازای هر برخورد دسته برای یک سال معین از دریافت داده، ایجاد می‌شود. مقداری که از این توزیع انتخاب می‌شود، به عنوان میانگین تعداد برخوردهایی که باید در یک رویداد برای همه برخوردهای دسته

¹"primary" interaction

²in-time pileup

³out-of-time pileup

⁴ Monte Carlo simulation of pile-up

شبیه سازی شود، اختصاص می یابد. این مقدار به عنوان "تعداد واقعی برخوردها" برای رویداد ثبت می شود. این مقدار،

درخشندگی لحظه‌ای را که برای همه برخوردهای دسته در آن رویداد شبیه‌سازی می‌شود، تعیین می‌کند.^[۶]

۲. تعداد برخوردها به ازای هر برخورد دسته (همزمان و ناهمزمان) در رویداد به طور تصادفی از یک توزیع پواسون با میانگین

برابر با مقداری که در (۱) انتخاب شده است، نمونه برداری می‌شود. این تعداد تعیین می‌کند که چه تعداد پایل‌آپ به

رویداد شبیه سازی شده اضافه شود. این مقدار به عنوان "تعداد پایل‌آپ" در رویداد ثبت می‌شود.^[۷]

توجه داشته باشید که اعمال همان میانگین تعداد برخوردها به همه برخوردهای دسته در یک رویداد، الزام می‌کند که همه

دسته‌ها دارای درخشندگی لحظه‌ای یکسانی باشند. نادیده گرفتن تغییرات درخشندگی دسته به دسته، تأثیر ناچیزی دارد: در عمل،

پراکنده‌گری درخشندگی دسته به دسته بسیار کمتر از پراکنده‌گری توزیع پواسون است و دسته‌های بلند استفاده شده در LHC باعث

می‌شوند که تأثیر نادیده گرفتن هرگونه تغییری از ساختار دسته نیز بسیار ناچیز باشد.^[۸]

ماژول مخلوط‌سازی^۱، رویدادهای پایل‌آپ را بر روی یک رویداد سیگنال (اصلی) قرار می‌دهد. این کار شامل پایل‌آپ

غیرهمزمان نیز می‌شود، بنابراین ماژول مخلوط‌سازی باید با توالی برخوردهای دسته کار کند تا به درستی سهم‌هایی که روی

برخورد دسته همزمان اثر می‌گذارند را شبیه‌سازی کند. این قابلیت به سایر جریان‌های ثانویه که رویدادهای آن‌ها قابل افزودن

به رویدادهای سیگنال هستند، مانند رویدادهای پرتوهای کیهانی و هاله پرتو^۲، نیز تعیین داده شده است. در سطح برخورد^۳

ترکیب انجام می‌شود. در این سطح، اطلاعات آشکارگر به همراه داده‌های مونت‌کارلو از جریان‌های مختلف با هم ترکیب شده و در

نتیجه محصول جدیدی به نام CrossingFrame ایجاد می‌شود. این CrossingFrame شامل اطلاعات مربوط به برخوردهای

مختلف (Crossing) است. اشیایی که با جریان‌های ورودی ثانویه مخلوط می‌شوند عبارتند از: PCaloHits، PSimHits

HepMCProduct و SimVertices، SimTracks، SimTracks آشکارگر CMS را فراهم می‌کند.

عملکردهای کلیدی

برهم نهی سطح ضربه: این ماژول، اطلاعات آشکارساز (برخوردها) و داده‌های مونت‌کارلو را از رویداد سیگنال و رویدادهای

پایل‌آپ در سطح برخورد ترکیب می‌کند.

اشیاء پشتیبانی شده: ماژول مخلوط‌سازی با انواع مختلفی از اشیاء مرتبط با شبیه‌سازی آشکارساز کار می‌کند، از جمله:

PSimHits (برخوردهای شبیه‌سازی شده در آشکارساز)

PCaloHits (برخوردهای انرژی‌سنجد)

SimTracks (مسیرهای شبیه‌سازی شده ذرات)

¹MixingModule

²Beam Halo

³Hit

(راس‌های اولیه برخورد شبیه‌سازی شده) SimVertices
 (اطلاعات حاصل از ژنراتور رویداد) HepMCProduct
 CrossingFrame: اطلاعات اضافه شده در یک ساختار داده جدید به نام CrossingFrame ذخیره می‌شود. هر شیء CrossingFrame اختصاصی خود را دارد. کد MixingModule و CrossingFrame به ترتیب در بسته‌های زیر قرار دارند:

SimDataFormats/CrossingFrame	Mixing/Base	SimGeneral/MixingModule
------------------------------	-------------	-------------------------

ماژول مخلوطسازی: شبیه‌سازی پاسخ کامل آشکارساز
 ماژول مخلوطسازی پاسخ کامل آشکارساز را شبیه‌سازی می‌کند. کاری که برای هر رویداد انجام می‌دهد به شرح زیر است:
 کپی کردن داده سیگنال به CrossingFrame در برخورد دسته^۰: این ماژول ابتدا داده‌های سیگنال را در CrossingFrame مربوط به برخورد دسته^۰ کپی می‌کند.
 حلقه زدن روی تمام برخوردهای دسته: ماژول سپس روی تمام برخوردهای دسته تعریف شده در پیکربندی حلقه می‌زند.
 تصمیم‌گیری برای هر برخورد دسته: برای هر برخورد دسته، بر اساس پیکربندی، تعداد رویدادهایی که باید از هر منبع اضافه شوند، تعیین می‌شود. این رویدادها توسط سرویس‌های ورودی ثانویه تأمین می‌گردند.
 افروzen اشیاء: برای هر برخورد دسته و برای هر نوع داده (SimVertices, SimTracks, PCaloHits, PSimHits, HepMCProducts)، اشیاء مربوطه از رویدادهای خوانده شده از جریان‌های ثانویه اضافه می‌شوند. همزمان، اطلاعات زمان بر اساس مقدار برخورد دسته جابه‌جا شده و شناسه‌های رویداد کدگذاری شده^۱ پر می‌شوند. از ماژول مخلوطسازی را می‌توان در دو حالت مختلف اجرا کرد:

حالت استاندارد (A): جریان‌های رویداد ثانویه روی جریان رویداد سیگنال سوپرنشانده می‌شوند.
 حالت چند مرحله‌ای (B): جریان‌های رویداد ثانویه ابتدا با هم ادغام می‌شوند، سپس نتیجه با جریان رویداد سیگنال ادغام می‌شود:

B1 ۱ مرحله

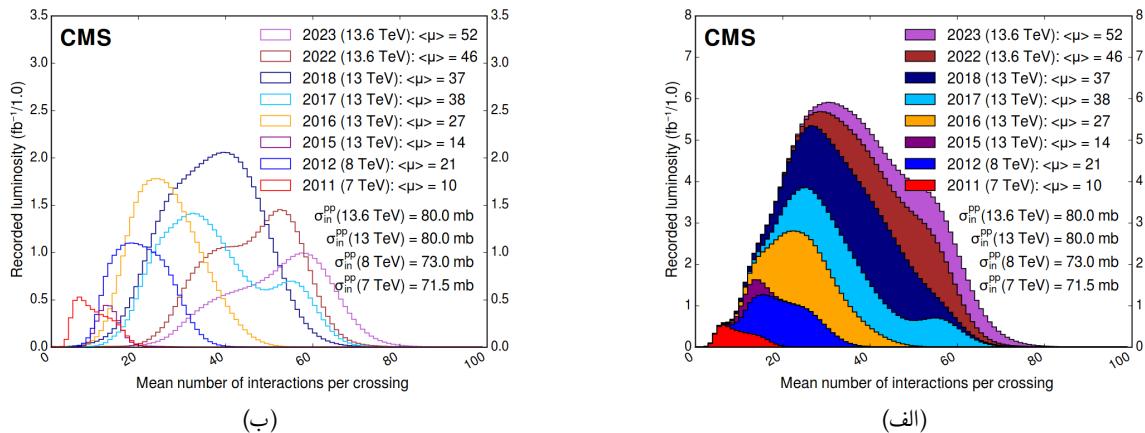
تولید داده‌های ترکیبی: نتایج ماژول سازی، یعنی CrossingFrame‌ها، به عنوان اشیاء با نام PCrossingFrames در یک فایل ذخیره می‌شوند. در این حالت، داده‌های سیگنال به CrossingFrame‌ها اضافه نمی‌شوند.

¹EncodedEventIds

مرحله ۲: B2

مخلوط کردن با استفاده از داده‌های منبع ترکیبی از مرحله ۱: CrossingFrame‌ها روی جریان رویداد سیگنال اصلی قرار داده می‌شوند. [۸]

۱۰.۳ توزیع‌های پایل‌آپ اندازه‌گیری شده از داده برهمنکش در هر برخورد (پایل‌آپ) برای سال‌های ۲۰۱۱–۲۰۲۳ و ۲۰۱۵ و ۲۰۲۲ و ۲۰۲۳ و ۲۰۲۰ و ۲۰۱۸–۲۰۱۲ و ۲۰۱۷ و ۲۰۱۶ و ۲۰۱۵ (آبی) برای بروتون-پروتون در سال‌های ۲۰۱۱ (قرمز)، ۲۰۱۲ (آبی)، ۲۰۱۵ (بنفش)، ۲۰۱۶ (نارنجی)، ۲۰۱۷ (آبی روش)، ۲۰۱۸ (آبی تیره) و ۲۰۲۲ (قهوه‌ای) نشان داده شده است. همچنین مقادیر میانگین کلی و سطح مقطع‌های حداقل سوگیری نمایش داده شده‌اند. این نمودارها فقط از داده‌هایی استفاده می‌کنند که تأییدیه طلایی (یعنی همه زیر آشکارگرهای CMS) برای هر نوع استفاده‌ای در تحلیل فیزیک تأیید شده‌اند) و مقادیر مقاطع عرضی حداقل سوگیری را شامل می‌شوند که از پیش‌بینی‌های نظری از Pythia گرفته شده‌اند و باید برای مقایسه با سایر آزمایش‌های LHC استفاده شوند. [۹]



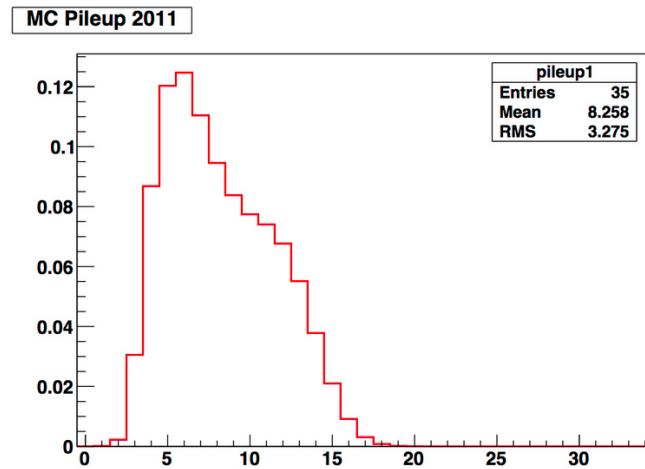
شکل ۱۰.۳: توزیع تعداد برهمنکش‌ها در سال‌های مختلف

۲۰.۳ توزیع‌های پایل‌آپ اندازه‌گیری شده از شبیه‌سازی مونت کارلو

این توزیع برای هر سال متفاوت است. نمودارهای زیر از مقادیر عددی (مطابق با سناریوی پایل‌آپ هر سال) ارائه شده در آن صفحه ساخته شده‌اند. محور x هیستوگرام‌های زیر "میانگین تعداد برخورد در هر برخورد دسته" است.

۲۰۱۱

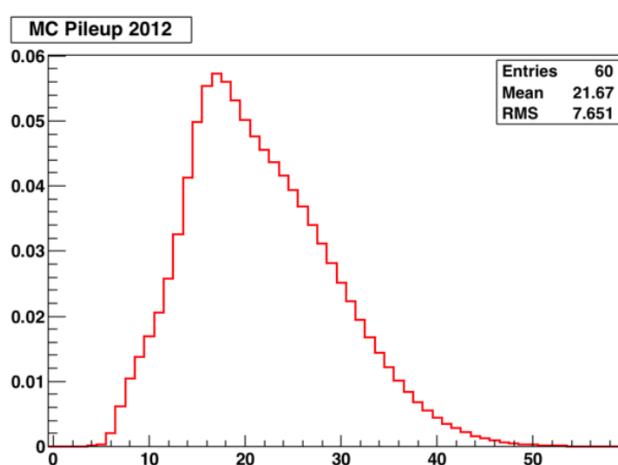
برای داده‌های سال ۲۰۱۱، توزیعی که در اینجا نشان داده شده است برای تولید رویدادهای مونت کارلو استفاده شده است. این توزیع مستقیماً با توزیع مشاهده شده در داده‌های اجرای ۲۰۱۱ مطابقت دارد، بنابراین برای دستیابی به توافق خوب بین داده و شبیه‌سازی از نظر برهمنکش همزمان، نیاز به بازنگارش کمی یا اصلاً وجود ندارد. [۹]



شکل ۲۰.۳: شبیه سازی مونت کارلو برای داده های ۲۰۱۱

۲۰۱۲

برای داده های سال ۲۰۱۲، توزیعی که در اینجا نشان داده است برای تولید رویدادهای مونت کارلو استفاده شده است. برخلاف سال ۲۰۱۱، این توزیع تطابق دقیقی با داده ها ندارد، بنابراین برای دستیابی به توافق خوب، بازنگارش نمونه های مونت کارلو ضروری است.^[۶]



شکل ۲۰.۴: شبیه سازی مونت کارلو برای داده های ۲۰۱۲

۳۰.۱.۳ دسترسی به اطلاعات پایل آپ ذخیره شده در مجموعه داده‌ها

MiniAOD و AOD

اطلاعات مونت کارلو در مورد پایل آپ به صورت برداری از اشیاء PileupSummaryInfo برای برخوردهای دسته مختلف در دسترس است. در AOD، مجموعه addPileupInfo است که دارای اطلاعات دقیق برای هر دو برخورد دسته در زمان^۱ و خارج از زمان^۲ است. در MiniAOD، این مجموعه slimmedAddPileupInfo نامیده می‌شود و با ذخیره نکردن اطلاعات دقیق برای برخوردهای دسته خارج از زمان، خلاصه شده است.[\[۶\]](#)

به این اطلاعات با استفاده از دوتابع تعریف شده در CMSSW نرم‌افزار SimDataFormats/PileupSummaryInfo دسترسی می‌توان داشت:

getTrueNumInteractions() •

این تابع میانگین واقعی توزیع پواسون را بازیابی می‌کند که تعداد برهمنکنش‌ها در هر برخورد دسته رویداد از آن نمونه برداری شده است. به عبارت دیگر، این تابع مقدار مورد انتظار تعداد برهمنکنش‌ها را بر اساس توزیع پواسون زمینه که برای شبیه سازی مونت کارلو استفاده می‌شود، برمی‌گرداند.[\[۷\]](#)

getPUNumInteractions() •

این تابع تعداد واقعی برهمنکنش‌هایی را که اضافه شده‌اند بازیابی می‌کند. برخلاف *getTrueNumInteractions* که میانگین توزیع زمینه را نشان می‌دهد، این تابع تعداد واقعی برهمنکنش‌هایی را که برای یک رویداد خاص شبیه سازی شده است، برمی‌گرداند.[\[۷\]](#)

۲۰.۳ ذخیره خصوصیات ذرات

۱۰.۲.۳ ابزارهای تحلیل پایل آپ

سلسله مراتب کلاس ابزاری ارزشمند برای محققان جهت دسترسی و تحلیل اطلاعات پایل آپ ذخیره شده در مجموعه داده‌های MiniAOD فراهم می‌کند. در اینجا نحوه تسهیل تحلیل پایل آپ توسط این کلاس‌ها شرح داده می‌شود:

سلسله مراتب کلاس

HaNaBaseMiniAnalyzer •

این کلاس پایه قابلیت‌های هسته‌ای را برای برنامه EDAnalyzer ما فراهم می‌کند. این کلاس راهاندازی اولیه را مدیریت می‌کند، اشیاء خواننده را برای انواع مختلف داده (رئوس، ذرات و غیره) مدیریت می‌کند و به طور بالقوه ابزارهای تحلیل

¹in-time

²out-of-time

اساسی ارائه می‌دهد.

SimplePUAnalyzer •

این کلاس از `HaNaBaseMiniAnalyzer` به ارث می‌برد و به طور خاص برای تجزیه و تحلیل اطلاعات پایل آپ و ذرات درون یک رویداد تخصص دارد. این کلاس از اشیاء خواننده از کلاس پایه برای دسترسی به داده‌های رئوس و کاندیدای ذرات استفاده می‌کند.

جریان پردازش رویداد

برنامه ما با ساخت نمونه‌ای از کلاس (`SimplePUAnalyzer`) شروع می‌شود. در طول راهنمایی (روشن سازنده یا تابع `beginJob`، تحلیلگر پارامترهای پیکربندی را بازیابی می‌کند و اشیاء خواننده را بر اساس آن پارامترها راهنمایی می‌کند. این اشیاء خواننده به احتمال یک کلاس پایه مشترک (`BaseEventReader`) مشتق شده‌اند و انواع خاصی از داده‌ها (به عنوان مثال `VertexReader` برای رئوس، `PackedCandidateReader` برای کاندیداهای ذرات) را مدیریت می‌کنند.

خواندن و استخراج اطلاعات پایل آپ

تحلیلگر ما از قابلیت‌های موجود در `CMSSW` برای خواندن و استخراج اطلاعات مرتبط با پایل آپ از مجموعه داده‌های `MiniAOD` استفاده می‌کند.

این امر شامل دسترسی به اشیاء `PileupSummaryInfo` از طریق توابعی مانند

`getTrueNumInteractions()` •

`getPUNumInteractions()` •

است.

این توابع امکان دسترسی به تعداد «واقعی» برهم‌کنش‌ها برای شبیه‌سازی (`getTrueNumInteractions()`) و تعداد «واقعی» برهم‌کنش‌های پایل آپ اضافه شده به رویداد (`getPUNumInteractions()`) را فراهم می‌کنند.

۲۰.۲.۳ تجزیه و تحلیل اطلاعات پایل آپ در رویدادها با استفاده از `EDAnalyzer`

این بخش به بررسی تحلیل اطلاعات پایل آپ درون رویدادها با استفاده از کلاس‌های تعریف شده ما در چارچوب نرم‌افزاری `CMSSW` (نسخه ۶.۴.۱۲) می‌پردازد. درک و تحلیل پایل آپ برای تفسیر صحیح داده‌های جمع‌آوری شده توسط آزمایش ضروری است.

ساختار کلی `EDAnalyzer` بصورت زیر خلاصه شده است:

۱. فایل های header برای چارچوب EDM (مدیریت رویداد) ، هیستوگرامها و اطلاعات پایل آپ
 ۲. کلاس پایه ما : HaNaBaseMiniAnalyzer برای خواندن ، آنالیز داده ها و سازنده کلاس متغیرهای عضو ^۱ مختلف را بر اساس فایل پیکربندی (مجموعه پارامتر) ، مقداردهی اولیه می کند. این متغیرها شامل پرچم هایی برای شناسایی نمونه های داده یا شبیه سازی، دایرکتوری ورودی برای فایل های اضافی و خواننده هایی برای اشیاء فیزیکی خاص مانند مسیرها^۲ ، رئوس^۳ ، میؤن ها و غیره می باشد.
 ۳. تعريف کلاس همچنین شامل روش هایی مانند beginJob ، endJob و analyze است که در مراحل مختلف اجرای حلقه رویداد ^۴ در چارچوب پردازش داده آزمایش (CMSSW) فراخوانده می شوند. از متدهای fillDescriptions برای تعریف پارامترهای پیکربندی که می توان آنها را در فایل پیکربندی برای سفارشی سازی مشخص کرد، استفاده می شود.
 ۴. تعريف کلاس SimplePUAnalyzer که از SimplePUAnalyzer ارث می برد.
- از قابلیت های کلاس پایه استفاده مجدد می کند و مراحل آنالیز خاصی را برای مطالعه پایل آپ اضافه می کند.

متغیرهای عضو در SimplePUAnalyzer

- به یک شیء هیستوگرام اشاره می کند که برای ذخیره تعداد واقعی برهمکنش ها (pileup) برای رویدادهای شبیه سازی شده (در صورت وجود) استفاده می شود.
- این اشاره گر به یک شیء TTree اشاره می کند، که یک ساختار داده است که در ROOT استفاده می شود.
۴. یک پارامتر است ^۵ را برای پیکربندی آنالیزگر دریافت می کند. کلاس را مقداردهی اولیه می کند. پیغامی Constructor چاپ می کند که نشان دهنده ایجاد شیء است.
- زمانی که شیء حذف می شود فراخوانده می شود. پیغامی چاپ می کند که نشان دهنده حذف شیء است.
- Destructor : fillDescriptions

- پیاده سازی را از کلاس پایه به ارث می برد، که احتمالاً برای تعریف پارامترهای پیکربندی برای تحلیلگر استفاده می شود.
- : beginJob
- (الف) یک بار در ابتدای کار فراخوانده می شود.
- (ب) یک سرویس TFile برای نوشتن داده ها به یک فایل ROOT ایجاد می کند.
- (ج) زیرمجموعه ای به نام "Trees" را درون فایل خروجی ROOT ایجاد می کند.

¹member variable

²Tracks

³Vertices

⁴event loop

⁵ParameterSet

(د) شاخه‌هایی را در شیء theTree تعریف می‌کند تا مقادیر مختلف مربوط به اشیاء بازسازی شده فیزیکی مانند رئوس

با کیفیت خوب، الکترون‌ها، میون‌ها، ترک‌ها و غیره را ذخیره کند.

(ه) شاخه‌هایی را برای ذخیره سینماتیک (تکانه، انرژی، موقعیت و غیره) تک تک ذرات بازسازی شده در رویداد تعریف

می‌کند.

(و) شاخه‌هایی را برای ذخیره تعداد ذرات بازسازی شده از انواع مختلف تعریف می‌کند.

(ز) شاخه‌هایی را برای ذخیره اطلاعات مربوط به رأس اولیه بازسازی شده که برای اتصال ترک‌ها استفاده می‌شود، تعریف

می‌کند.

:Analyze

(الف) برای هر رویداد در طول اجرای حلقه رویداد فراخوانده می‌شود.

(ب) اطلاعات مربوط به رئوس و ذرات بازسازی شده را از اشیاء vertexReader و packedReader می‌خواند.

(ج) اگر خواندن رئوس با مشکل مواجه شود، تابع بدون پردازش بیشتر رویداد خارج می‌شود.

(د) شاخه‌های شیء theTree را با داده‌های رویداد جاری پر می‌کند.

۵. ثبت پلاگین SimplePUAnalyzer را در چارچوب EDM ثبت می‌کند.

۶. و در آخر کامپایل و اجرای این قطعه کد پایه در CMSSW.

پر کردن یک درخت روت (ROOT TTree)

این تحلیل‌گر، اطلاعات استخراج شده و پردازش شده مربوط به پایل آپ را درون یک شیء درخت روت پر می‌کند. این درخت به عنوان فرمت داده‌ای ساختاریافته برای تحلیل و تجسم بیشتر عمل می‌کند. هر شاخه درون درخت به یک متغیر خاص پایل آپ مربوط می‌شود و امکان دسترسی و دستکاری کارآمد را فراهم می‌سازد.

مباحث موجود در مورد شبیه‌سازی و اندازه‌گیری پایل آپ در CMS را تکمیل می‌کند. این ابزار با ارائه‌ی یک راه حل عملی برای تحلیل مستقیم ویژگی‌های پایل آپ درون تک‌تک رویدادها، فراتر از مباحث نظری می‌رود. این تحلیل می‌تواند برای موارد زیر مورد استفاده قرار گیرد:

- مطالعه توزیع پایل آپ در سراسر رویدادها

- بررسی تأثیر پایل آپ بر کارایی و کیفیت بازسازی رئوس

- ارزیابی تأثیر بالقوه‌ی پایل آپ بر نتایج سایر تحلیل‌ها (به عنوان مثال، اندازه‌گیری تکانه عرضی گمشده)

اگرچه SimplePUAnalyzer در درجه اول برای تحلیل اطلاعات پایل آپ درون داده‌های مونت‌کارلو^۱ طراحی شده است،

اما این پتانسیل را دارد که برای کار با داده‌های واقعی در مطالعات پایل آپ نیز سازگار شود. این سازگاری ممکن است نیازمند

¹ MC

اصلاحاتی برای در نظر گرفتن تفاوت‌های احتمالی در نحوه ذخیره‌سازی اطلاعات پایل آپ در داده‌های واقعی نسبت به داده‌های مونت‌کارلو باشد.

برنامه از چارچوب ROOT برای ذخیره و تجزیه و تحلیل داده استفاده می‌کند. شیء TTree برای ذخیره اطلاعات در یک فرمت ساختاریافته برای پردازش بیشتر استفاده می‌شود.

تجزیه و تحلیل پایل آپ به نظر می‌رسد تمرکز اصلی کلاس SimplePUAnalyzer باشد، که از تأکید بر اطلاعات رئوس و متغیر npv (تعداد رئوس اولیه) به عنوان معیاری برای پایل آپ مشهود است.

در حالی که SimplePUAnalyzer بر روی اطلاعات پایل آپ و ذرات تمرکز دارد، کلاس پایه HaNaBaseMiniAnalyzer احتمالاً قابلیت‌هایی را برای خواندن و تجزیه و تحلیل سایر انواع داده‌های مرتبط با اهداف تحلیل فیزیک خاص شما (به عنوان مثال، جت‌ها، تکانه عرضی گشده و غیره) ارائه می‌دهد.

در نتیجه کلاس SimplePUAnalyzer با ارائه یک ابزار کاربرپسند برای تحلیل اطلاعات پایل آپ در مجموعه داده‌های MiniAOD درون چارچوب نرم‌افزاری CMSSW (نسخه ۱۲.۴.۶)، درک عمیقت‌تری از تأثیر پایل آپ بر داده‌های آزمایش CMS را تسهیل می‌کند. این درک برای بهبود کیفیت داده، اصلاح تکنیک‌های تحلیل و در نهایت دستیابی به نتایج فیزیکی قابل اعتمادتر، ضروری است.

۳۰.۲.۳ پردازش و ذخیره‌سازی متغیرهای پایل آپ

معیارهای برش برای شناسایی رئوس با کیفیت در VertexReader

در کد ارائه شده، دو نوع معیار برش^۱ برای شناسایی رئوس با کیفیت در کلاس VertexReader استفاده می‌شود:

معیارهای برش عمومی

موقعیت رأس (z) مقدار مطلق z رأس باید کمتر از ۲۴.۰° باشد. این آستانه، رأس را در راستای محور برخورددهنده (محوری که پروتون‌ها در آن برخورد می‌کنند) محدود می‌کند. درجات آزادی^۲: ndof عددی است که نشان‌دهنده تعداد محدودیت‌های مستقل در برآذش رئوس است. مقدار ndof رأس باید از ۰.۴° بیشتر باشد تا به عنوان یک رأس با کیفیت خوب در نظر گرفته شود. فاصله عرضی از خط برخورد^۳: rho فاصله رأس از خط برخورددهنده در صفحه عمود بر محور برخورد است. مقدار rho باید کمتر از ۰.۲° سانتی‌متر باشد.

معیارهای برش سخت‌گیرانه‌تر

موقعیت رأس (z) مقدار مطلق z رأس باید کمتر از ۰.۱۰° باشد. درجات آزادی^۴: ndof رأس باید از ۰.۸° بیشتر باشد.

¹cuts

²ndof

³rho

⁴ndof

انتخاب رأس اولیه: رأس اولیه، اولین رأس موجود در مجموعه رئوس بازسازی شده در رویداد در نظر گرفته می‌شود. این فرض بر این است که اولین رأس، رأس اولیه‌ای است که در فرآیند پراکندگی سخت تولید شده است.

ذخیره اطلاعات پایل آپ تحلیل‌گر VertexReader متغيرهای ضروری مرتبط با پایل آپ را پردازش و ذخیره می‌کند. این

متغيرها شامل موارد زیر هستند:

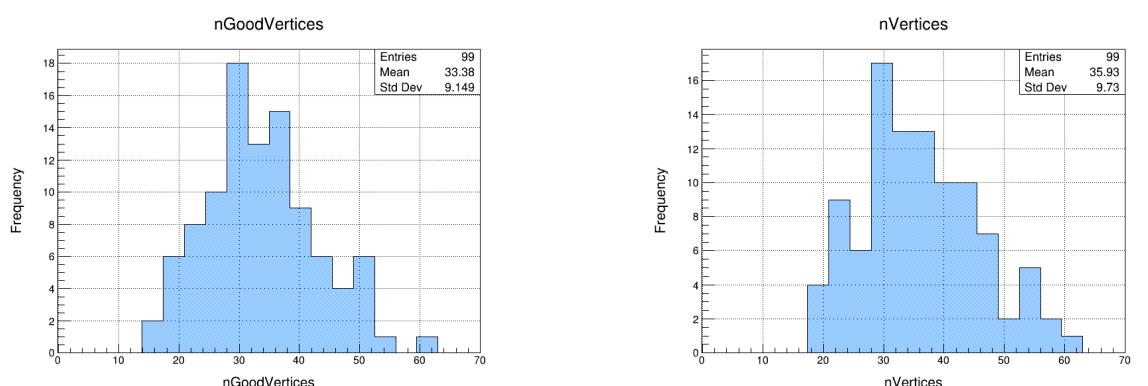
تعداد کل رئوس بازسازی شده^۱: این متغير تعداد کل رئوس بازسازی شده در رویداد را نشان می‌دهد. (شکل ۴.۳-الف)

تعداد رئوس با کیفیت خوب^۲: این متغير تعداد رئوس بازسازی شده که از معیارهای برش عمومی عبور می‌کنند را نشان می‌دهد.

(شکل ۴.۳-ب) تعداد رئوس مطابق با معیارهای کیفی سخت‌گیرانه‌تر (nVGoodVertices، nVGGoodVertices): این

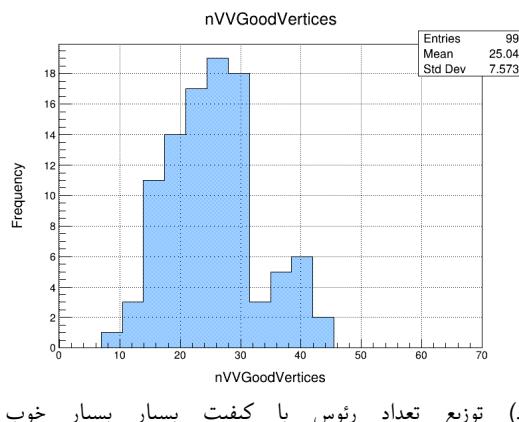
متغيرها تعداد رئوس بازسازی شده که از معیارهای برش سخت‌گیرانه‌تر عبور می‌کنند را نشان می‌دهند. (شکل ۴.۳-ج و شکل

۴.۳-د)

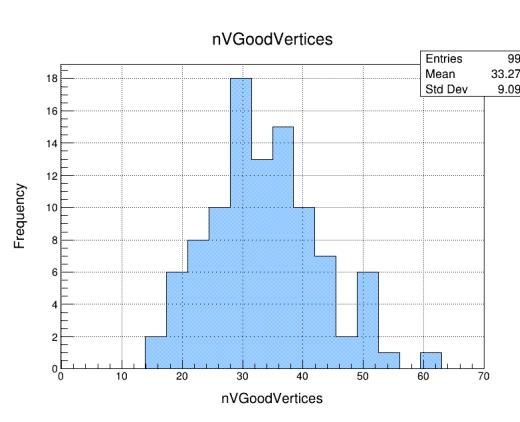


.(ب) توزیع تعداد رئوس با کیفیت خوب .nGoodVertices

.(الف) توزیع تعداد کل رئوس بازسازی شده .nVertices



.(د) توزیع تعداد رئوس با کیفیت بسیار خوب .nVGGoodVertices



.(ج) توزیع تعداد رئوس با کیفیت بسیار خوب .nVGoodVertices

شکل ۴.۳: توزیع تعداد رئوس بازسازی شده بر اساس معیارهای مختلف کیفیت رئوس.

نحوه خواندن و ذخیره‌سازی اطلاعات مربوط به اشیاء PackedCandidate

این تحلیل‌گر به معرفی قابلیت‌های پایه‌ای و توضیح انواع کاندیدای ذره‌ای در فیزیک ذرات می‌پردازد.

¹nVertices

²nGoodVertices

کاندیداهای ذرات چیست؟ کاندیداهای ذرات نشان دهنده ذرات فیزیکی بازسازی شده هستند. این به این معنی است که دانشمندان از اطلاعات جمع آوری شده توسط آشکارسازهای ذرات برای استنتاج وجود و ویژگی های ذرات استفاده می کنند. انواع مختلفی از کاندیداهای ذرهای وجود دارد که هر کدام نشان دهنده ذره خاصی است.

کلاس پایه مشترک برای همه این کاندیداهای، کلاسی به نام `Candidate` است. این کلاس سه قابلیت اساسی را برای همه انواع کاندیداهای فراهم می کند:

۱. دسترسی به اطلاعات سینماتیک (حرکتی)

پیاده سازی شامل بردار لورنتس تکانه، راس، بار الکتریکی و شناسه `PDG` ، و تعدادی دیگر است. نمونه هایی از این کاندیداهای عبارتند از:

```
double pt = cand.pt(), eta = cand.eta(), phi = cand.phi();
int q = cand.charge(), id = cand.pdgId(), st = cand.energy();
int qx3 = cand.dz();
```

۲. دسترسی به اجزای زیربنایی که می توانند از هر نوعی باشند. به عنوان مثال، پیوندهایی به یک مسیر^۱، ابر خوشة^۲ و غیره. نمونه:

```
TrackRef trk = cand.get<TrackRef>();
SuperClusterRef cluster = cand.get<SuperClusterRef>();
```

برای دسترسی به اجزای زیرین یک کاندیدای ذرهای، از روش `الگوی <...> get` استفاده می شود. تعریف این روش `الگو در فایل هدر DataFormats/RecoCandidate/interface/RecoCandidate.h` قرار دارد که باید آن را در برنامه خود وارد کنیم.^۳.

در صورتی که یک کاندیدای ذرهای چندین جزو از یک نوع داشته باشد، مانند یک میؤن که سه تناسب مسیر^۴ مختلف دارد، می توانید یک آرگومان اضافی با نام «تگ»^۵ ارسال کرد. برای دسترسی به تناسبهای مسیر میؤن، می توانید به صورت زیر عمل کنید:

```
CaloTowerRef ct = jetConstituent.get<CaloTowerRef>();
```

¹track

²super-cluster

³include

⁴track fit

⁵tag

کاندیدهای ذرات کجا استفاده می‌شوند؟

الکترون‌ها، میئون‌ها، فوتون‌ها: کاندیداهای ذرات به عنوان بلوک‌های سازنده برای شناسایی و تحلیل ذرات شناخته شده‌ای مانند الکترون‌ها، میئون‌ها و فوتون‌ها عمل می‌کنند.

Particle Flow: این تکنیک برای بازسازی دقیق ذرات خروجی از برخوردها استفاده می‌شود. این تکنیک به شدت بر روی کاندیداهای ذرات برای شناسایی و ترکیب اطلاعات از زیرسیستم‌های مختلف آشکارگر تکیه می‌کند.

مجموعه ابزار تحلیل فیزیکی^۱: چارچوب PAT از کاندیداهای ذرات به عنوان ورودی برای ایجاد اشیاء فیزیکی با کیفیت بالاتر استفاده می‌کند. این اشیاء برای تحلیل‌های بعدی بهینه‌سازی شده‌اند.

هادرон‌های باردار^۲: کاندیداهای ذرهای در شناسایی و مطالعه هادرон‌های باردار، که ذرات مرکب حاوی کوارک‌ها هستند، نقش دارند.

هادرون‌های خنثی^۳: شناسایی و تحلیل هادرون‌های خنثی مانند نوترون‌ها و مزون‌های خنثی نیز به کاندیداهای ذرات وابسته است.

BaseEventReader< pat::PackedCandidateCollection > از کلاس پایه PackedCandidateReader

به ارث می‌برد.

این امر نشان می‌دهد که PackedCandidateReader از قابلیت‌های کلاس پایه برای خواندن کلی داده رویداد استفاده می‌کند و روی مدیریت مجموعه‌هایی از اشیاء pat::PackedCandidate تمرکز دارد.

انواع کاندیدای ذرهای

کاندیداهای ذرهای در دو دسته اصلی قرار می‌گیرند:

۱. کاندیداهای برگی^۴: این نوع کاندیداهای مستقیماً از اطلاعات جمع‌آوری شده توسط زیرآشکارگرهای CMS بازسازی می‌شوند. کاندیداهای برگی هیچ کاندیدای فرزند ندارند و باید از کلاس پایه LeafCandidate به ارث برسند.

۲. کاندیداهای مرکب^۵: این نوع کاندیداهای از ترکیب اشیاء کاندیدای دیگر بازسازی می‌شوند. کاندیداهای مرکب حاوی ارجاع داخلی یا کلون‌هایی از کاندیداهای فرزند استفاده شده در بازسازی خود هستند.

کاندیداهای RECO

کاندیداهایی که هیچ فرزندی ندارند و دارای اجزایی از اشیاء دیگر RECO/AOD هستند، باید از کلاس recoCandidate

¹ Physics Analysis Toolkit (PAT)

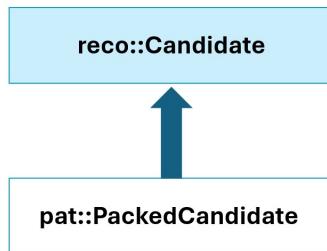
² Charged Hadrons

³ Neutral Hadrons

⁴ Leaf Candidate Type

⁵ Composite Candidate Types

به ارث برستند. این رابط interface روشهایی را برای پیمایش در مجموعه اجزای استاندارد RECO/AOD ارائه می‌دهد. برای اشیائی که از recoCandidate به ارث می‌برند، بررسی تداخل «خودکار»^۱ برای یافتن احتمالی اجزای مشترک RECO پیاده‌سازی شده است. کلاس (PackedCandidateReader) ما از recoCandidate و در آنالیز ما نقش مهمی ایفا خواهد کرد که در ادامه نحوه دسترسی به آن‌ها و سلسله مراتب کلاس را توضیح میدهیم.



شکل ۵.۳: ارث بری کلاس recoCandidate از PackedCandidateReader

سازنده PackedCandidateReader، اشیاء PackedCandidateReader را مقداردهی اولیه می‌کند. متغیرهای عضو را برای ذخیره اطلاعات مربوط به ذرات مختلف (مانند تعداد الکترون‌ها، میؤن‌ها، هادرون‌های باردار و خنثی و فوتون‌ها) مقداردهی اولیه می‌کند.

آرایه‌های بزرگ را برای ذخیره مقادیر دینامیکی (مانند تکانه، انرژی، وغیره) برای حداکثر ۴۰۰۰۰ کاندیدا اختصاص می‌دهد. تمام عناصر این آرایه‌ها را با صفر مقداردهی اولیه می‌کند تا از داده‌های باقی‌مانده از رویدادهای قبلی جلوگیری کند. سازنده حلقه‌ای روی تمام کاندیداهای موجود در مجموعه pat::PackedCandidateCollection اجرا می‌کند.

برای هر کاندیدا:

نوع ذره را با گرفتن مقدار مطلق شناسه ذره‌ای^۲ آن تعیین می‌کند. شمارنده‌های مربوط به انواع مختلف ذرات را بر اساس نوع ذره به روز رسانی می‌کند. مقادیر دینامیکی (مانند تکانه، انرژی، وغیره) را در آرایه‌های مربوطه ذخیره می‌کند. اطلاعات اضافی مانند بار ذره، displacement از خط برخورد (dxy, dz)، کلید رأس مرتبط، وغیره را نیز ذخیره می‌کند. در نهایت، تعداد کل کاندیداهای خوانده شده را برمی‌گردداند.

این کد از کلاس pat::PackedCandidate برای دسترسی به اطلاعات مربوط به ذرات بازسازی شده استفاده می‌کند.

آرایه‌های بزرگ برای ذخیره اطلاعات دینامیکی کاندیداهای اختصاص داده شده‌اند تا از ایجاد آرایه‌های جدید در هر فراخوانی تابع

¹Overlap Checking

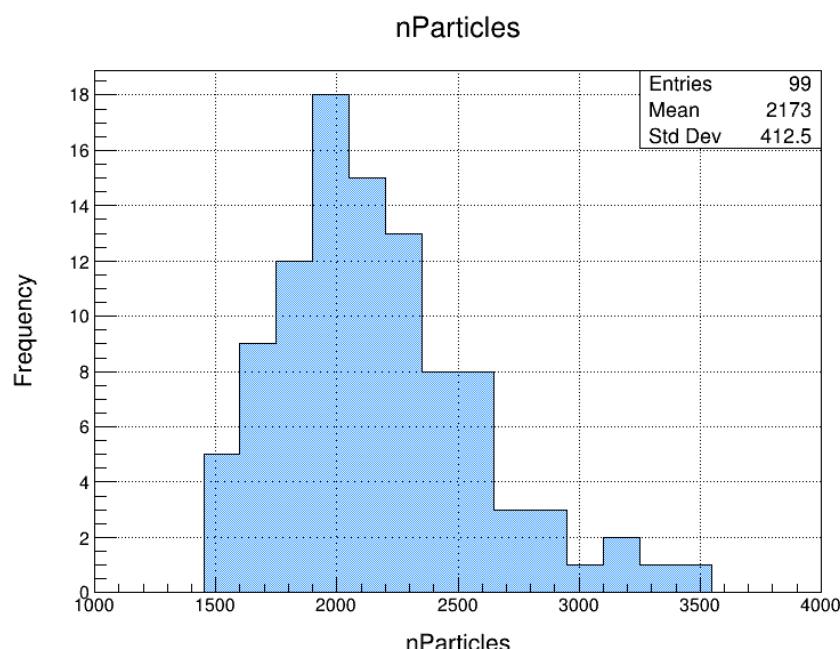
²pdgId

جلوگیری شود.

در نتیجه کلاس PackedCandidateReader ابزاری مفید برای خواندن و پردازش آبجکت‌های در یک تحلیلگر رویداد فیزیک انرژی بالا است. این کلاس می‌تواند انواع مختلفی از ذرات را شناسایی و شمارش کند و همچنین اطلاعات دینامیکی و اطلاعات اضافی مربوط به هر کاندیدا را ذخیره کند.

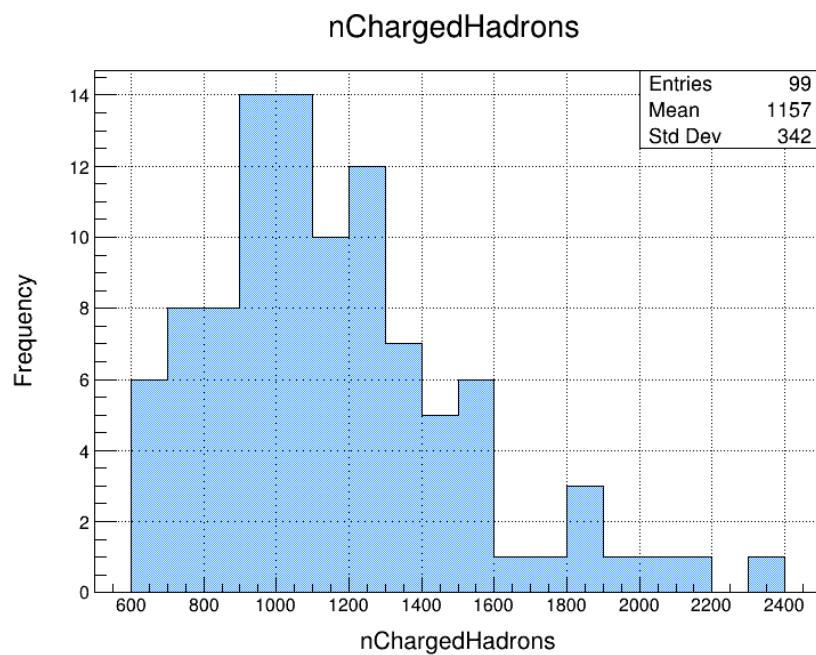
۴.۲.۳ تعداد کاندیدای ذرات در یک رویداد

توزیع تعداد کل ذرات



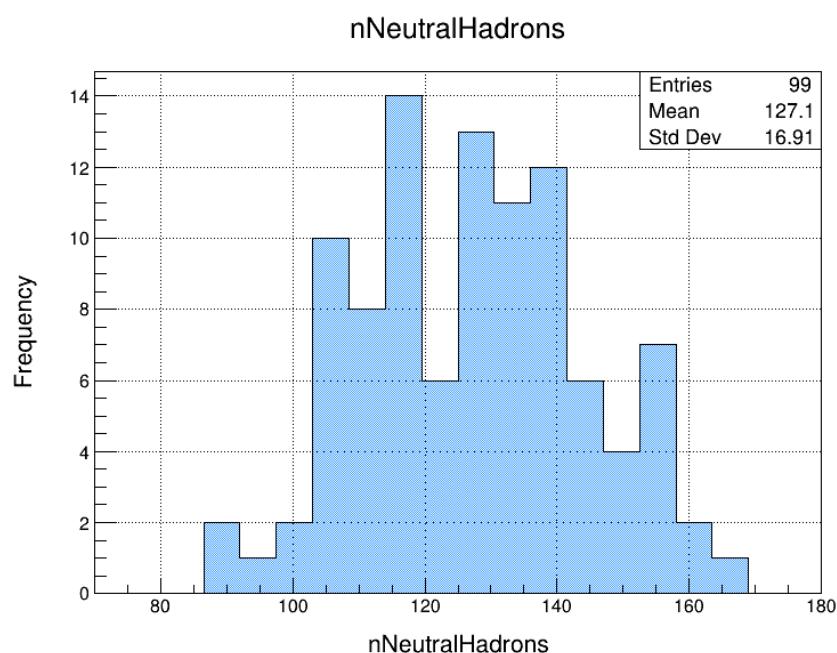
شکل ۶.۳: توزیع تعداد کل ذرات

توزیع تعداد هادرون های باردار



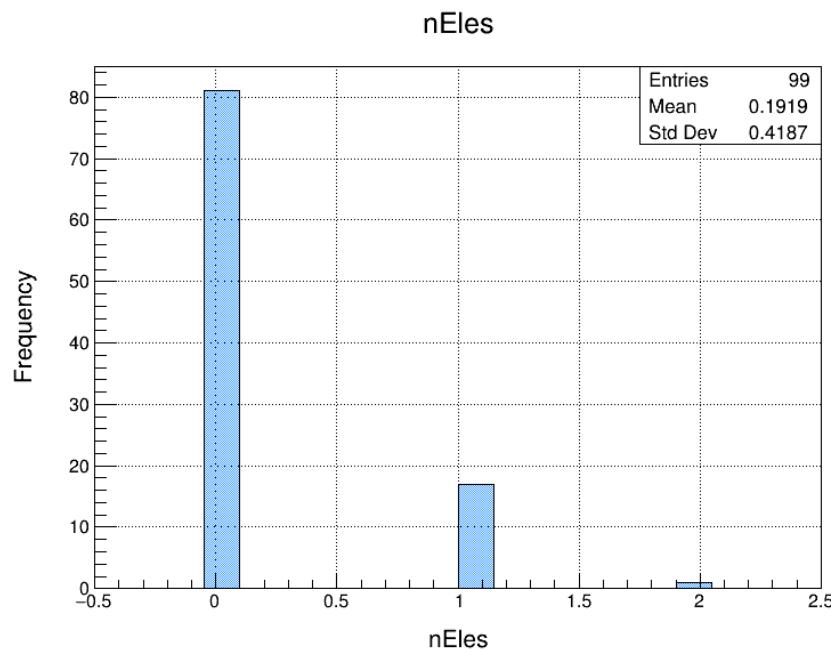
شکل ۷.۳: توزیع تعداد هادرون های باردار

توزیع تعداد هادرون های خنثی



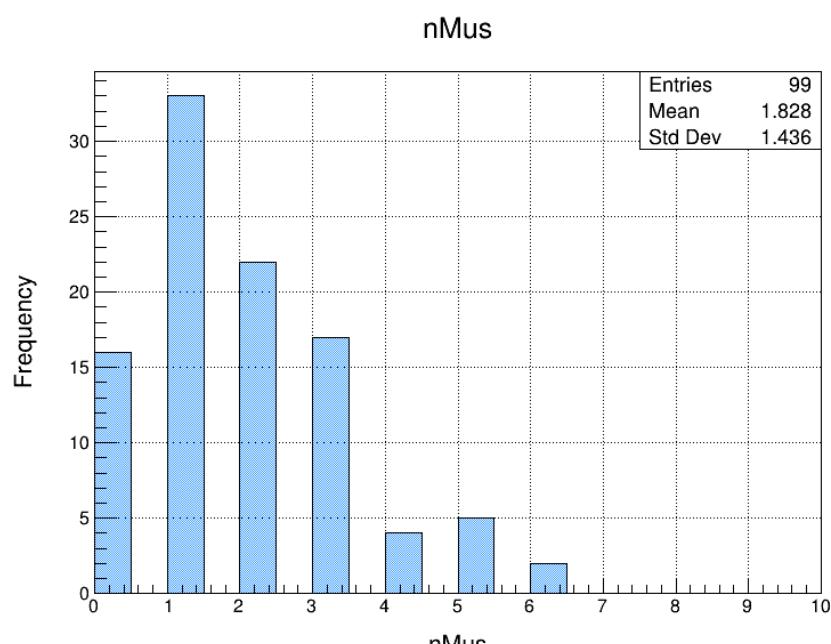
شکل ۸.۳: توزیع تعداد هادرون های خنثی

توزيع تعداد الكترون ها



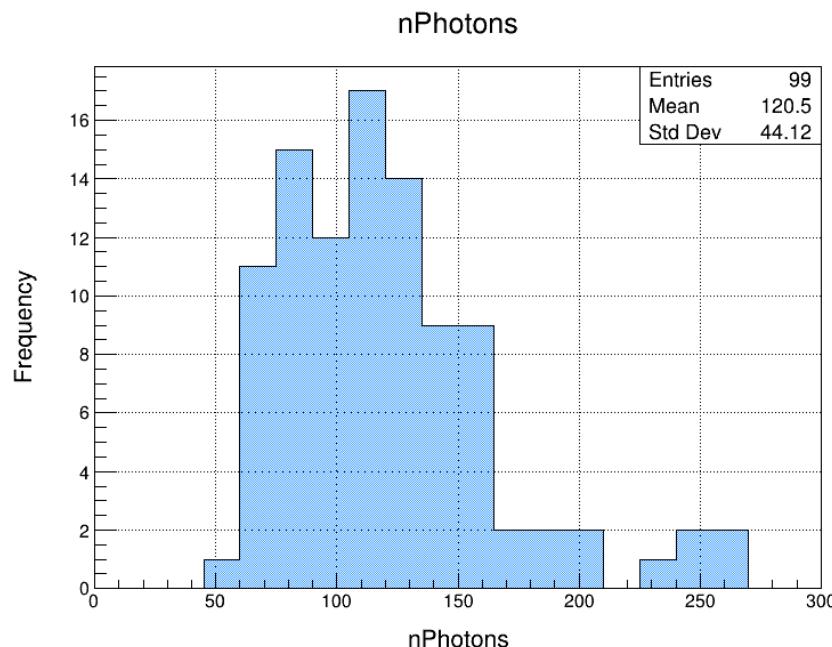
شكل ٩.٣: توزيع تعداد الكترون ها

توزيع تعداد ميؤن ها



شكل ١٠.٣: توزيع تعداد ميؤن ها

توزیع تعداد فوتون‌ها



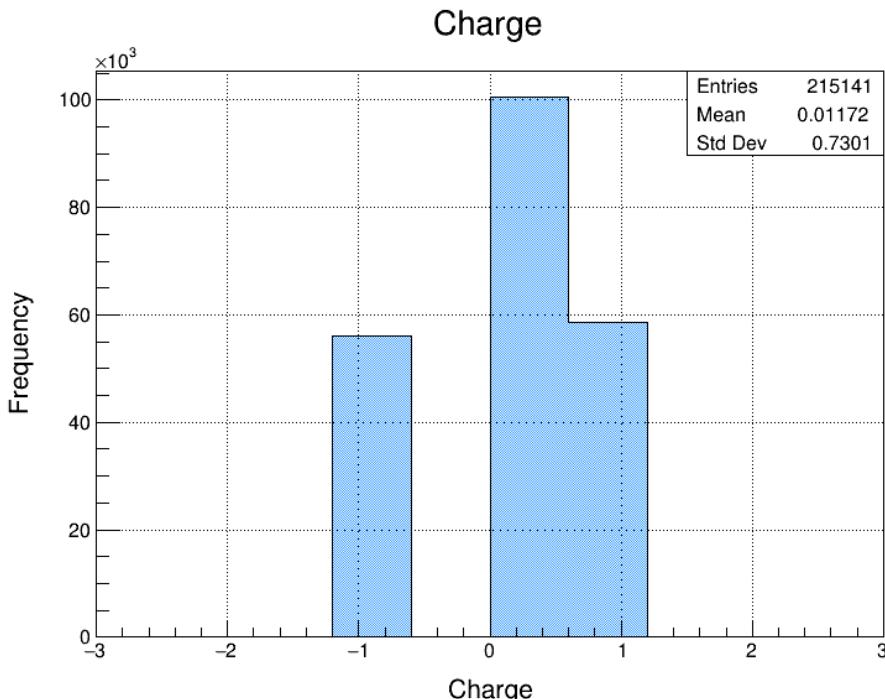
شکل ۱۱.۳: توزیع تعداد فوتون‌ها

خصوصیات ذرات

خصوصیات ذرات برای توصیف ویژگی‌های ذرات تولید شده در برخوردها استفاده می‌شوند، اما اطلاعات متفاوتی را ارائه می‌دهند

توزیع بار ذرات

شکل ۱۲.۳ نمودار توزیع بار ذرات را نشان می‌دهد. در اینجا اطلاعات موجود در تصویر را شرح می‌دهیم:



شکل ۱۲.۳: توزیع بار ذرات

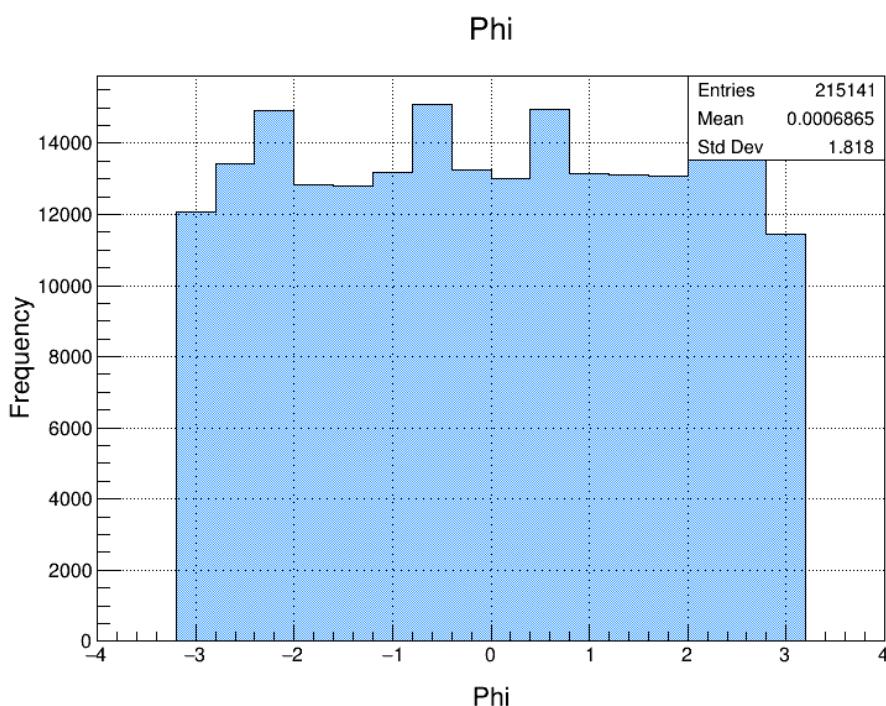
تفسیر توزیع بار

ذرات با بار e : این دسته شامل الکترون‌ها و میئون‌ها با بار $+e$ و همچنین پوزیترون‌ها (ضد الکترون‌ها) با بار $-e$ می‌شود.
 پیک در سمت چپ نمودار (نزدیک به $-1e$) مربوط به پوزیترون‌ها است، در حالی که پیک در سمت راست (نزدیک به $+1e$) مربوط به الکترون‌ها و میئون‌ها است.
 ذرات با بار 0 : این دسته شامل نوترون‌ها و فوتون‌ها می‌شود. نوترون‌ها ذرات خنثی هستند، در حالی که فوتون‌ها حاملان نیروی الکترومغناطیسی بدون جرم هستند و هیچ باری ندارند. پیک در مرکز نمودار (نزدیک به $0e$) مربوط به نوترون‌ها و فوتون‌ها است.

- **نسبت ذرات:** توزیع بار در نمودار نشان می‌دهد که تعداد الکترون‌ها و میئون‌ها به طور قابل توجهی بیشتر از تعداد پوزیترون‌ها و پادذره‌ها است. این امر با فراوانی نسبی ذرات در طبیعت مطابقت دارد.
- **انحراف معیار:** انحراف معیار 1.730° نشان می‌دهد که توزیع بار نسبتاً گسترشده است. این به این معنی است که بار ذرات در این داده‌ها از مقادیر منفی کوچک تا مقادیر مثبت بزرگ متغیر است.

توزیع فی Φ

یک توزیع فی یک هیستوگرام صاف خواهد داشت، که نشان می‌دهد ذرات با احتمال برابر در تمام جهات اطراف محور برخورد منتشر می‌شوند. با این حال، توزیع مشاهده شده در تصویر زیر یکنواخت نیست. در شکل ۱۳.۳، نموداری از توزیع فی (Φ) را در یک شبیه‌سازی داریم. در اینجا اطلاعات موجود در تصویر را شرح می‌دهیم:



شکل ۱۳.۳: توزیع فی، تعداد ذرات شناسایی شده در هر مقدار فی را نشان می‌دهد.

- Φ (فی): این برچسب محور عمودی است که نشان‌دهنده فی است.
- تعداد کل ورودی^۱: ۲۱۵۱۴۱:
- میانگین^۲: $6.865e^{-4}$ ؛ این نشان می‌دهد که مقدار متوسط فی برابر با (6.865×10^{-4}) است.
- انحراف معیار^۳: این نشان می‌دهد که انحراف معیار فی برابر با ۰.۸۱۸ است.

توزیع فی (Φ):

- نشان‌دهنده زاویه سمتی ذره در صفحه برخورد است.
- برخورد را به عنوان یک سطح صاف (صفحه) تصور کنید. فی زاویه حول این صفحه را که ذره از آن خارج می‌شود تعریف می‌کند.

¹Entries Phi

²Mean

³Std Dev

- درجه مطابق با محور x مثبت، 90° درجه مطابق با محور y مثبت و غیره است.
- توزیع فی یکنواخت (ذرات به طور مساوی در تمام جهات ظاهر می‌شوند) نشان می‌دهد که جهت ترجیحی برای انتشار ذرات وجود ندارد.
- انحراف از توزیع یکنواخت ممکن است نشان دهنده فرآیندهای خاص یا برهمنکنش‌هایی در داخل برخورد باشد.

کاربردها

از توزیع فی در مطالعات مختلف فیزیک ذرات استفاده می‌شود، از جمله:

- درک دینامیک برخورد: تجزیه و تحلیل چگونگی توزیع فی می‌تواند اطلاعاتی در مورد نیروهای اساسی و برهمنکنش‌های درون برخورد را آشکار کند.
- شناسایی کانال‌های فروپاشی خاص: فروپاشی برخی ذرات ممکن است مقادیر فی خاصی را برای ذرات حاصل ترجیح دهد.
- مطالعات پس زمینه: در برخی موارد، نویز پس زمینه در آشکارسازها ممکن است توزیع فی مشخصی داشته باشد و به تشخیص آن از سیگنال مورد نظر کمک کند.

غیر یکنواخت بودن توزیع فی

- غیر یکنواخت بودن توزیع فی نشان می‌دهد که فرآیندها یا برهمنکنش‌های زمینه‌ای در برخورد وجود دارند که مقادیر خاصی از فی را نسبت به مقادیر دیگر ترجیح می‌دهند. این می‌تواند به دلایل زیر باشد:
- ناهمگنی در سیستم برخورد: اگر ذرات برخورد کننده یا برهمنکنش زمینه‌ای دارای عدم تقارن ذاتی باشند، می‌تواند منجر به توزیع فی غیر یکنواخت شود.
 - کانال‌های فروپاشی خاص: اگر ذرات شناسایی شده محصولات فرآیندهای فروپاشی خاصی باشند، این فرآیندهای فروپاشی ممکن است ترجیحی برای مقادیر خاص فی داشته باشند.
 - اثرات آشکارساز: در برخی موارد، خود آشکارساز ممکن است حساسیت یا سوگیری داشته باشد که بر توزیع فی مشاهده شده تأثیر می‌گذارد.

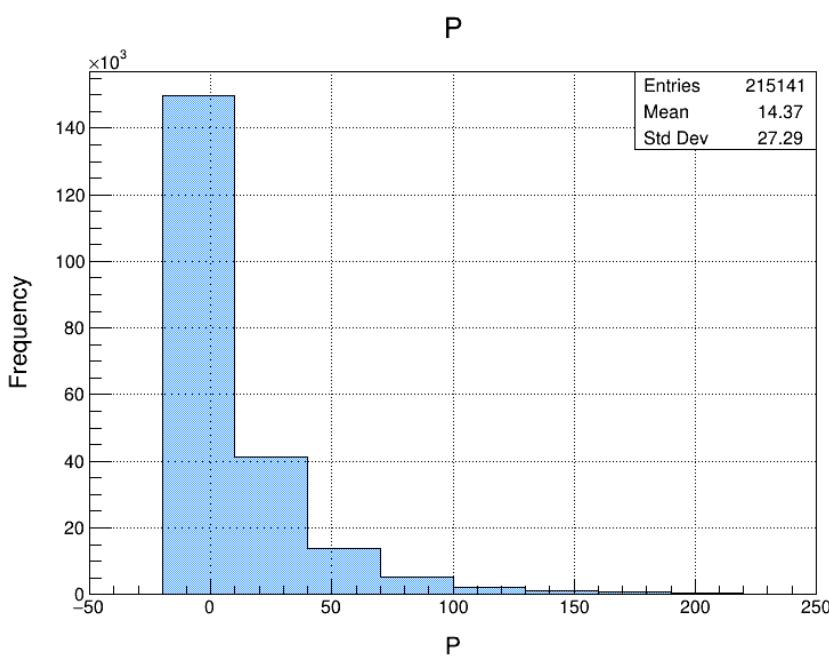
نتیجه‌گیری

توزیع فی ابزاری ارزشمند برای مطالعه دینامیک برخوردهای ذرات است. با تجزیه و تحلیل توزیع فی، محققان می‌توانند به درک فرآیندها و برهمنکنش‌های زمینه‌ای که ذرات مشاهده شده را تولید می‌کنند، دست یابند.

- تفسیر توزیع فی به آزمایش یا شبیه‌سازی خاص بستگی دارد.
- توزیع فی غیر یکنواخت اغلب نشان‌دهنده پدیده‌های فیزیک جالب است.

توزیع تکانه

یک توزیع تکانه یک هیستوگرام صاف خواهد داشت، که نشان می‌دهد ذرات با احتمال برابر در تمام جهات حرکت می‌کنند. با این حال، توزیع مشاهده شده در تصویر یکنواخت نیست.



شکل ۱۴.۳: توزیع تکانه در تصویر، تعداد ذرات شناسایی شده در هر مقدار تکانه را نشان می‌دهد.

غیر یکنواختی توزیع تکانه

غیر یکنواختی توزیع تکانه نشان می‌دهد که فرآیندها یا برهمنکنش‌های زمینه‌ای در برخورد ذرات وجود دارند که مقادیر خاصی از تکانه را نسبت به مقادیر دیگر ترجیح می‌دهند. این می‌تواند به دلایل زیر باشد:

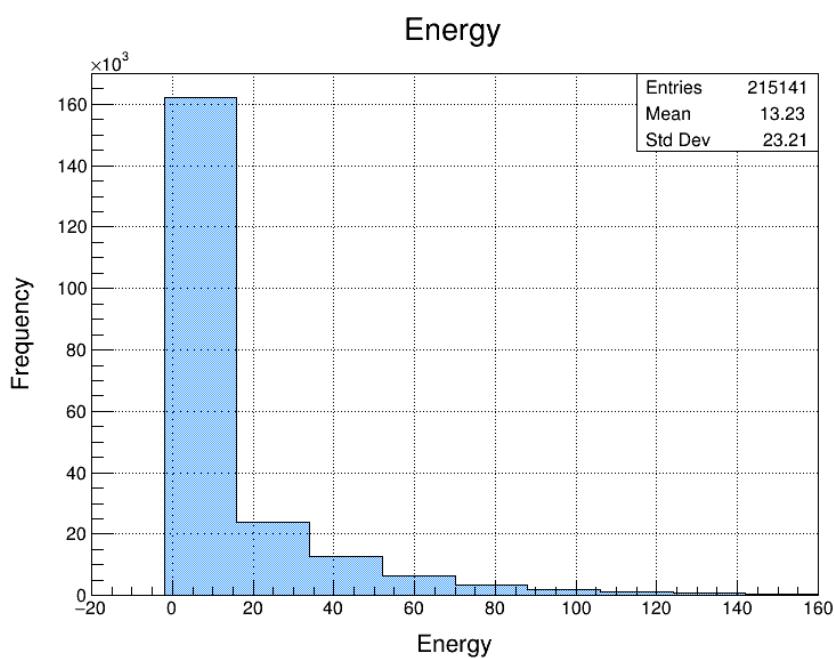
- ناهمگنی در سیستم برخورد: اگر ذرات برخورد کننده یا برهمنکنش زمینه‌ای دارای عدم تقارن ذاتی باشند، می‌تواند منجر به توزیع تکانه غیر یکنواخت شود.
- فروپاشی ذرات خاص: اگر ذرات شناسایی شده محصولات فرآیندهای فروپاشی خاصی باشند، این فرآیندهای فروپاشی ممکن است ترجیحی برای مقادیر خاص تکانه داشته باشند.
- اثرات آشکارساز: در برخی موارد، خود آشکارساز ممکن است حساسیت یا سوگیری داشته باشد که بر توزیع تکانه

مشاهده شده تأثیر می‌گذارد.

توزیع تکانه ابزاری ارزشمند برای مطالعه دینامیک برخوردهای ذرات است. با تجزیه و تحلیل توزیع تکانه، محققان می‌توانند

به درک فرآیندها و برهمنکنش‌های زمینه‌ای که ذرات مشاهده شده را تولید می‌کنند بپردازنند.

توزیع انرژی ذرات



شکل ۱۵.۳: نموداری از توزیع انرژی ذرات

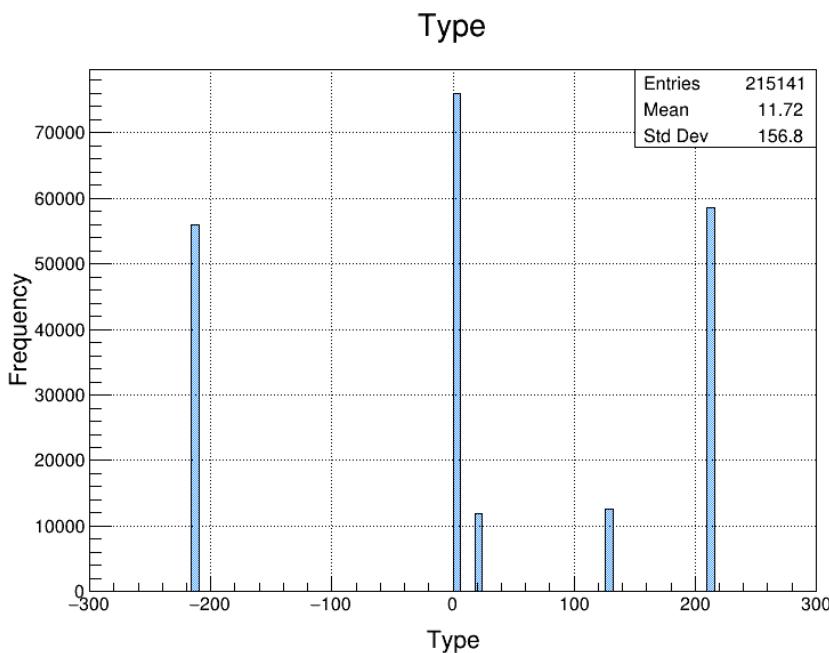
محور عمودی تعداد ذراتی را نشان می‌دهد که دارای مقدار انرژی خاصی در محور افقی هستند. نمودار نشان می‌دهد که اکثر ذرات دارای انرژی نزدیک به مقدار متوسط (14.95 GeV) هستند و تعداد ذرات با دور شدن انرژی از میانگین کاهش می‌یابد.

توزیع نوع ذرات در برخوردها شناسه گروه داده ذرات (PDGID)

شناسه گروه داده ذرات (PDGID) یک استاندارد بین‌المللی برای شناسایی منحصر به فرد انواع مختلف ذرات زیر اتمی است. این شناسه به طور گسترده در فیزیک ذرات برای طبقه‌بندی و تجزیه و تحلیل ذرات شناسایی شده در برخورد دهنده‌های ذرات و سایر آزمایش‌ها استفاده می‌شود.

توزيع ذرات بر اساس PDGID

شکل ۳.۵.ب توزیع تعداد ذرات شناسایی شده در برخوردها را بر اساس PDGID آنها نشان می‌دهد. محور افقی (x) نشان‌دهنده PDGID ذره و محور عمودی (y) تعداد ذراتی است که با هر PDGID خاص شناسایی شده‌اند.^[۹]



شکل ۱۶.۳: توزیع ذرات بر اساس شناسه گروه داده ذرات (PDGID)

همانطور که در این شکل مشاهده می‌شود، ذرات با PDGID‌های خاص، فراوانی بیشتری نسبت به سایرین دارند. برخی از نمونه‌های برجسته عبارتند از:

• PDGID 11 (e^-): الکترون

• PDGID 13 (μ^-): میون

• PDGID 211 (π^-): هادرон باردار (به طور خاص، پیون منفی)

• PDGID 130 (K^0): هادرон خنثی (به طور خاص، کائون صفر) یا هادرон خنثی (بسطه به زمینه)

• PDGID 22 (γ): فوتون

در اینجا توضیحاتی در مورد هر ذره ارائه شده است:

- الکترون (e^-): یک ذره بنیادی با بار الکتریکی منفی، الکترون یکی از رایج‌ترین ذرات در جهان است.^[۹]
 - میئون (μ^-): یک ذره بنیادی با بار الکتریکی منفی و جرمی تقریباً ۲۰۷ برابر جرم الکترون.^[۹]
 - هادرон باردار (π^-): یک هادرон باردار که به خانواده مزون‌ها تعلق دارد. جرم آن حدود ۱۳۹ برابر جرم الکترون است و در نیروی هسته‌ای قوی شرکت می‌کند.^[۹]
 - هادرон خنثی (K^0): یک هادرون خنثی که به خانواده مزون‌ها تعلق دارد. جرم آن کمی بیشتر از جرم پیون باردار است و در نیروی هسته‌ای ضعیف (بسته به ایزوتوپ خاص کائون صفر) نقش دارد. توجه: بسته به زمینه، "id == 130" می‌تواند به طور کلی نشان‌دهنده تمام هادرون‌های خنثی باشد، نه فقط کائون‌ها.^[۹]
 - فوتون (γ): ذره‌ای بدون جرم که نیروی الکترومغناطیسی را منتقل می‌کند. خود فوتون هیچ بار الکتریکی ندارد.^[۹]
- فراوانی بالای این ذرات نشان می‌دهد که آنها برخی از رایج‌ترین ذرات تولید شده در برخوردهای خاص هستند. توزیع PDGID ذرات در یک آزمایش به عوامل مختلفی از جمله نوع برخورد، انرژی برخورد و کارایی آشکارساز بستگی دارد.

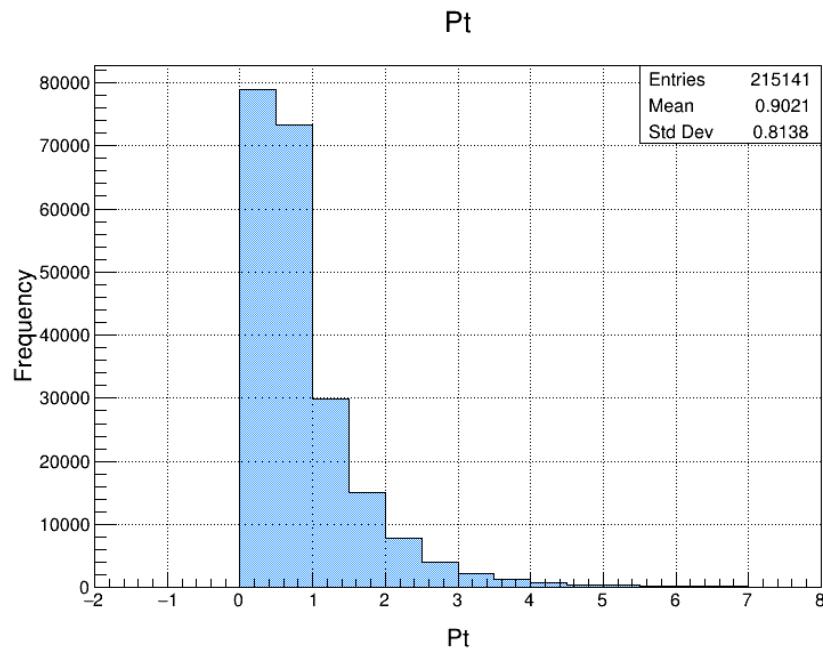
اهمیت توزیع ذرات بر اساس PDGID

تحلیل توزیع ذرات بر اساس PDGID اطلاعات ارزشمندی در مورد دینامیک برخوردهای ذرات ارائه می‌دهد. با درک انواع ذرات تولید شده، دانشمندان می‌توانند مکانیسم‌های بنیادی حاکم بر فرآیندهای برخورد را بهتر درک کنند. این اطلاعات برای توسعه مدل‌های نظری برخورد ذرات و پیش‌بینی انواع ذرات تولید شده در برخوردهای مختلف بسیار مهم است.^[۱۰]

توزیع تکانه عرضی

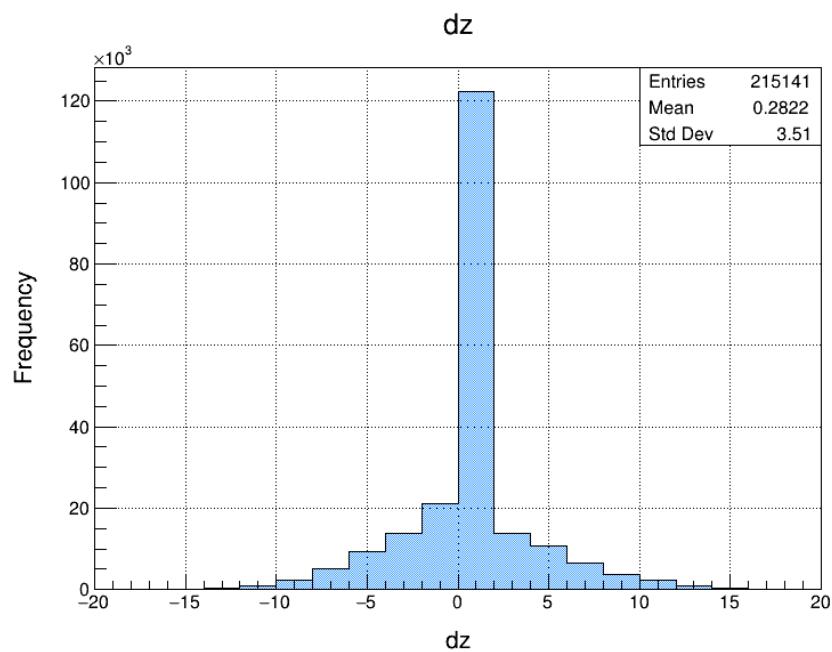
شکل ۳.۵ نموداری از توزیع تکانه عرضی^۱ ذرات را نشان می‌دهد. تکانه عرضی، اندازه حرکت ذره را عمود بر جهت پرتو (جهت برخورد ذرات) اندازه‌گیری می‌کند.

^۱Pt distribution



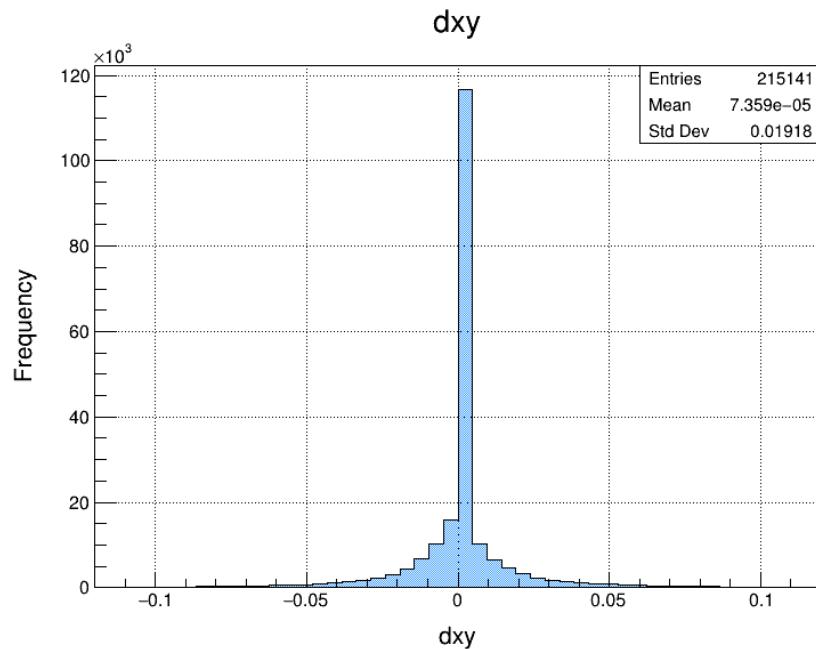
شکل ۱۷.۳: مولفه‌ای از تکانه ذره است که عمود بر جهت حرکت کلی آن است.

توزيع dz



شکل ۱۸.۳: توزیع پارامتر برخورد طولی ذرات

توزيع dxy



شکل ۱۹.۳: توزیع پارامتر برخورد عرضی ذرات

توزیع شبه‌تندی (η)

شبه‌تندی^۱ (η) یک مختصات فضایی رایج است که زاویه یک ذره را نسبت به محور پرتو در فیزیک تجربی ذرات توصیف می‌کند. این کمیت به صورت زیر تعریف می‌شود:

$$\eta = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right] \quad (1.3)$$

که در آن θ زاویه بین اندازه حرکت سه بعدی ذره (p) و جهت مثبت محور پرتو است:

$$\theta = 2 \arctan(e^{-\eta}) \quad (2.3)$$

شبه‌تندی را می‌توان به عنوان تابعی از اندازه حرکت سه بعدی (p) به صورت زیر نوشت:

^۱Pseudorapidity

$$\eta = \frac{1}{2} \ln \left(\frac{|p| + p_L}{|p| - p_L} \right) = \left(\frac{p_L}{|p|} \right) \quad (3.3)$$

که در آن p_L مولفه اندازه حرکت در راستای محور پرتو (یعنی اندازه حرکت طولی - با استفاده از سیستم مختصات رایج در فیزیک برخوردهای هادرон، این کمیت اغلب با p_z نشان داده می‌شود) است.^[۱۱] در حد جرم ناچیز ذره (زمانی که سرعت ذره به سرعت نور نزدیک می‌شود)، شبه تندی به صورت زیر درمی‌آید:

$$\eta \approx y = (\beta) \quad (4.3)$$

که در آن y تندی ذره و β سرعت ذره به عنوان کسری از سرعت نور است.

ارتباط با تندی (y)

تندی معیاری از میزان تقویت سرعت ذره در جهت طولی است. به صورت زیر تعریف می‌شود:

$$y = \frac{1}{2} \ln \left(\frac{E + p_L}{E - p_L} \right) \quad (5.3)$$

که در آن E انرژی ذره و p_L اندازه حرکت طولی آن است.

در حد جرم ناچیز ذره، تندی به صورت زیر درمی‌آید:

$$y \approx \eta \quad (6.3)$$

این بدان معنی است که برای ذرات بدون جرم، شبه تندی و تندی اساساً یکسان هستند.

مزایای شبه تندی

شبه تندی اغلب به زاویه قطبی (θ) ترجیح داده می‌شود زیرا تولید ذره تمایل دارد به عنوان تابعی از η ثابت بماند. دلیل این امر آن است که تبدیل لورنتس برای η صرفاً یک تغییر مکان در η است، در حالی که تبدیل لورنتس برای θ پیچیده‌تر است.^[۱۱] همچنین، اختلافات در η تحت تقویت‌های (boost) هم جهت با محور پرتو، لورنتس-ناورد هستند. این بدان معناست که

اختلاف η بین دو ذره، بدون توجه به میزان تقویت آنها در امتداد محور پرتو، یکسان باقی می‌ماند. این ویژگی برای تحلیل داده‌های برخورددهنده‌های بزرگ هادرتون که در آن ذرات برخوردکننده دارای اندازه حرکت‌های متفاوتی هستند، مفید است.^[۱۱]

معیار ΔR

معیار ΔR اندازه‌ای از جدایی زاویه‌ای بین ذرات است که با استفاده از η (یا y) و زاویه سمتی (ϕ) تعریف می‌شود:

$$\Delta R = \sqrt{(\eta_1 - \eta_2)^2 + (\phi_1 - \phi_2)^2} \quad (7.3)$$

که در آن η_1 و η_2 شبه تندی‌های دو ذره و ϕ_1 و ϕ_2 زوایای سمتی آنها هستند.

معیار ΔR نیز تحت تقویت‌های هم جهت با محور پرتو، لورنتس-ناورده است. این ویژگی آن را به ابزاری ارزشمند برای تحلیل فیزیک ذرات تبدیل می‌کند.

کاربردهای شبه تندی

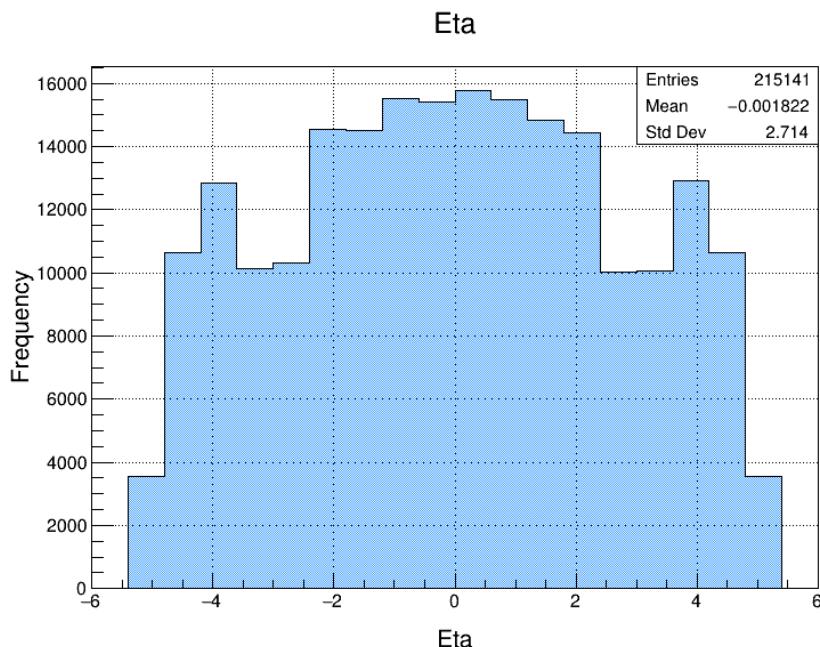
شبه تندی در کاربردهای مختلفی در فیزیک ذرات استفاده می‌شود، از جمله:

- شناسایی جت‌ها: جت‌ها پاشش‌های جهت‌دار ذرات هستند که در برهمکنش قوی تولید می‌شوند. از شبه تندی می‌توان برای شناسایی جت‌ها و اندازه‌گیری ویژگی‌های آنها مانند اندازه حرکت عرضی و انرژی آنها استفاده کرد.
- مطالعه تولید ذره: از شبه تندی می‌توان برای مطالعه تولید ذرات در مناطق مختلف فضای فاز استفاده کرد. این امر می‌تواند برای آزمایش مدل‌های تولید ذره و جستجوی ذرات جدید به کار رود.
- اندازه‌گیری همبستگی بین ذرات: از شبه تندی می‌توان برای اندازه‌گیری همبستگی بین ذرات استفاده کرد. این امر می‌تواند برای مطالعه ساختار هادرتون‌ها و جستجوی پدیده‌های جدید به کار رود.

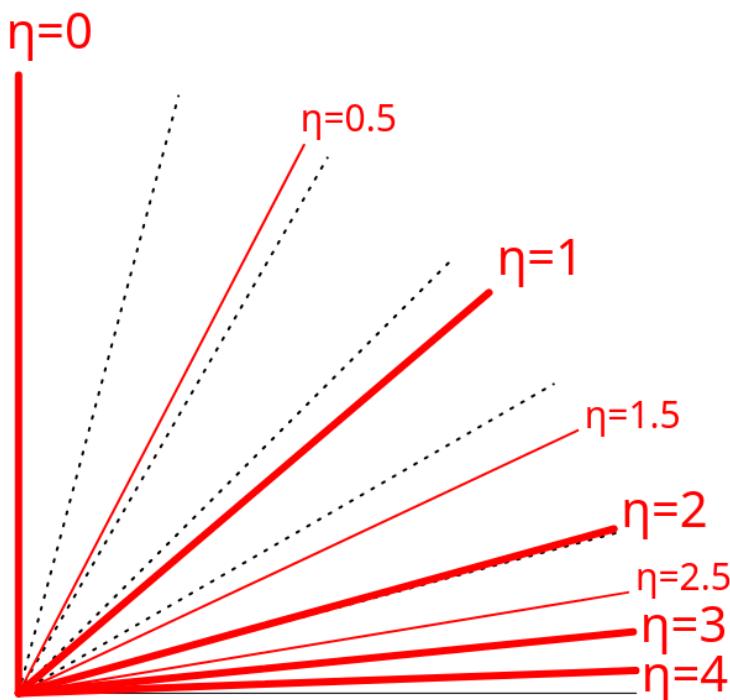
توزیع شناسه رأس ذرات

نقش شناسه رأس در ردبایی ذرات همانطور که در محور عمودی (y) شکل ۲۲.۳ نشان داده شده، شناسه رأس به نزدیکترین محل ایجاد یا برهمکنش یک ذره در رویداد برخورد اشاره دارد. شناسایی این رأس، به فیزیکدانان کمک می‌کند تا محل دقیق منشأ یک ذره را مشخص کنند.

سیستم کلیدگذاری برای ردبایی آسان‌تر به هر رأس یک شناسه منحصر به فرد، یعنی همان شناسه رأس، اختصاص داده می‌شود. این سیستم کلیدگذاری که در محور افقی (x) شکل قابل مشاهده است، به محققان امکان می‌دهد تا ذرات را به طور

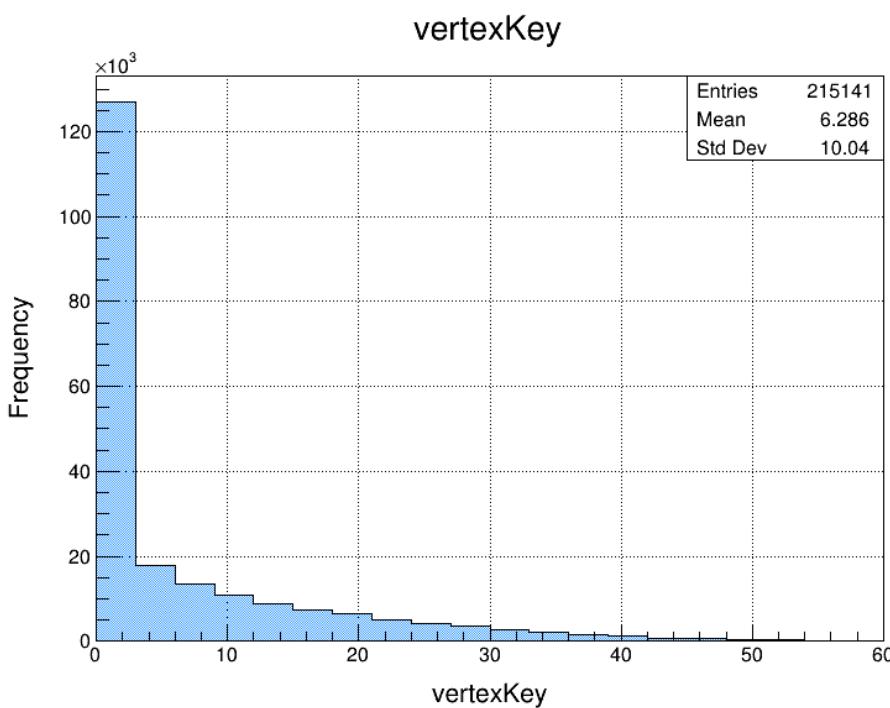


شکل ۲۰.۳: توزیع ذرات را به عنوان تابعی از شبه تندی (η) در برخوردهای پروتون-پروتون (pp) در برخورددهنده بزرگ هادرон (LHC) نشان می‌دهد. مقدار متوسط η برابر با صفر است.



شکل ۲۱.۳: در این شکل، مقادیر شبه تندی نمایش داده شده‌اند. در فیزیک ذرات، زاویه صفر معمولاً هم جهت با محور پرتو است، بنابراین ذرات با مقادیر بالای شبه تندی معمولاً از دست می‌روند و به همراه پرتو از فضای آشکارساز خارج می‌شوند. [۱۱]

مؤثر و دقیق به نقاط اولیه‌ی آن‌ها ردیابی کنند. یافتن نزدیکترین رأس و اختصاص شناسه رأس مربوطه را به عنوان یک روش بنیادی در مازول‌های EDAnalyzer معرفی می‌کند. این مازول‌های نرم‌افزاری به بخشی از یک چارچوب تحلیل بزرگ‌تر در آزمایش‌های فیزیک ذرات هستند، با ردیابی ذرات به نقاط اولیه‌ی آن‌ها، شناسایی نزدیکترین رأس نقشی اساسی در کشف پیچیدگی‌های برهمنکنش‌های ذرات ایفا می‌کند. درک این برهمنکنش‌ها برای رمزگشایی از مکانیزم‌های بنیادی حاکم بر دنیای زیر اتمی ضروری است. نمودار موجود در شکل ۲۴.۳ توزیع شناسه‌های رأس را برای مجموعه‌ای از ذرات را نشان می‌دهد. میانگین شناسه رأس 6.294×10^3 و انحراف معیار آن 10.06 است. این مقادیر، بینش آماری در مورد فراوانی شناسه‌های رأس خاص در مجموعه داده‌ها ارائه می‌دهند.



شکل ۲۴.۳: شناسه رأس را برای هر ذره نشان می‌دهد. شناسه رأس خصوصیتی است که برای ردیابی محل پیدایش ذره در یک آزمایش برخورد ذرات استفاده می‌شود.

۳.۳ روش‌های تعیین شناسه نزدیکترین رأس به هر ذره

ما برای تعیین شناسه نزدیکترین رأس به هر ذره از پنج روش مختلف استفاده کردیم. این روش‌ها بر اساس اصول و الگوریتم‌های متفاوتی عمل می‌کنند و هر کدام مزایا و معایب خاص خود را دارند. با توجه به اهمیت دقت در تعیین شناسه رأس، به ویژه در مواردی که تعداد زیادی ذره به رأس اول نزدیک هستند، روش انتخاب نهایی از بین این پنج روش، بر اساس نتایج آزمون PT انجام می‌شود. آزمون pT ، به عنوان یک معیار مهم در ارزیابی عملکرد روش‌های مختلف، به ما کمک می‌کند تا روشی را انتخاب کنیم که قادر به تشخیص دقیق‌تر و قابل اعتمادتر رأس‌های اولیه باشد. در واقع، روشهایی که در نهایت انتخاب می‌شود،

روشی است که در آزمون pT بهترین عملکرد را داشته باشد و بتواند تعداد بیشتری از ذرات را به رأس اول صحیح مرتبط کند. به منظور تعیین دقیق‌ترین رأس برای هر ذره، ابتدا رأس‌های موجود با توجه به کیفیت آن‌ها به گروه‌های مختلفی تقسیم‌بندی می‌شوند. این گروه‌ها با متغیرهایی مانند Vtx، GoodVtx و ... نشان داده می‌شوند. هر یک از این متغیرها تعداد رأس‌هایی را که دارای کیفیت مشخصی هستند، مشخص می‌کند. برای هر ذره، فاصله dz تا تمام رأس‌های موجود محاسبه می‌شود. فاصله dz به عنوان کوتاه‌ترین فاصله مستقیم بین دو نقطه در فضای dz تعریف می‌شود. از بین رأس‌هایی که فاصله آن‌ها تا ذره محاسبه شده است، رأسی که کمترین فاصله را دارد به عنوان نزدیک‌ترین رأس به آن ذره انتخاب می‌شود. استفاده از روش‌های VeryGoodVtx، GoodVtx و ... به ما این امکان را می‌دهد تا تأثیر کیفیت رأس بر دقت نتایج را بررسی کنیم. انتظار می‌رود که با افزایش کیفیت روش‌ها، دقت تعیین نزدیک‌ترین رأس نیز افزایش یابد.

۱۰.۳.۳ جمع‌آوری داده‌های رأس

- **مجموعه رأس‌ها:** جمع‌آوری تمامی رأس‌های موجود که معیارهای تعریف‌شده را برآورده می‌کنند. هر رأس دارای مختصات (x, y, z) در فضا خواهد بود.

أنواع رأس‌ها

بازسازی رأس جزء حیاتی از بازسازی رویدادها است. انواع مختلفی از رأس‌ها شناسایی و بازسازی می‌شوند که هر کدام هدف خاصی را دنبال می‌کند:

رأس در کلیت، رأس^۱ در آزمایش CMS به نقطه‌ای اشاره دارد که مسیرهای ذره‌ای با هم برخورد می‌کنند و نشان‌دهنده تعامل یا برخورد ذرات است. بازسازی رأس شامل شناسایی این نقاط از مسیرهای ذرات باردار با استفاده از Tracker سیلیکونی می‌باشد.^[۱۲]

رأس اولیه رأس اولیه^۲ نقطه‌ای است که برخورد ابتدایی پروتون-پروتون در آن رخ می‌دهد. این رأس، اصلی‌ترین رأس در یک رویداد است، زیرا برای اختصاص مسیرهایها به برخوردها و تعیین سینماتیک رویداد استفاده می‌شود. بازسازی دقیق رأس اولیه برای درک ساختار کلی رویداد و اهداف تریگر ضروری است.^[۱۲]

رأس ثانویه رأس‌های ثانویه^۳ از تجزیه ذرات با عمر طولانی، مانند هادرون‌های سنگین و لیپتون‌های^۴، که از رأس اولیه دور می‌شوند، تشکیل می‌شوند. شناسایی این رأس‌ها برای مطالعه خواص این ذرات و برای برچسب‌گذاری طعم، که به تمایز بین

¹ Vertex

² Primary Vertex

³ Secondary Vertex

انواع مختلف تعاملات ذره‌ای کمک می‌کند، اهمیت دارد^[۱۲].

رأس تجمعی رأس‌های تجمعی^۱ نتیجه تعاملات متعدد پروتون-پروتون هستند که در همان عبور دسته‌ای اتفاق می‌افتد. این تعاملات می‌توانند بازسازی رأس اولیه را پیچیده کنند و نیاز است که برای تحلیل دقیق رویدادها به آن‌ها توجه شود. آزمایش CMS از الگوریتم‌های پیشرفته برای تمایز بین رأس اولیه و رأس‌های تجمعی و کاهش اثرات آن‌ها استفاده می‌کند^[۱۲]. در مجموع، بازسازی رأس در آزمایش CMS شامل الگوریتم‌ها و تکنیک‌های پیشرفته‌ای است که برای شناسایی و تفکیک دقیق این انواع رأس‌ها، که برای تحلیل دقیق رویدادها و اندازه‌گیری‌های فیزیک ضروری است، به کار می‌رود^[۱۲].

۲۰.۳.۳ انواع روش‌های تعیین نزدیکترین رأس

در این بخش، روش‌های مختلف برای تعیین نزدیکترین رأس به هر ذره توضیح داده می‌شود. این روش‌ها شامل محاسبه فاصله، جمع‌آوری داده‌ها و استفاده از معیارهای مختلف برای انتخاب رأس نزدیکترین هستند:

- محاسبه فاصله: برای هر ذره، فاصله تا هر رأس محاسبه می‌شود. این فاصله معمولاً در بعد dz محاسبه می‌شود که کوتاه‌ترین فاصله در جهت z بین موقعیت ذره و رأس است.

$$dz = |\tilde{z}_{\text{ذره}} - \tilde{z}_{\text{رأس}}| \quad (8.3)$$

- جمع‌آوری داده‌ها: مقادیر فاصله را برای هر جفت ذره-رأس ذخیره می‌کنیم.

روش‌های مختلف تعیین نزدیکترین رأس عبارتند از:

- **minVtx**: این روش ساده‌ترین روش برای تعیین نزدیکترین رأس است. در این روش، فاصله ذره تا تمامی رأس‌های موجود محاسبه می‌شود و رأسی که کمترین فاصله را دارد به عنوان نزدیکترین رأس انتخاب می‌شود.

- **minGoodVtx**: در این روش، تنها رأس‌هایی که شرایط کیفی خاصی را برآورده می‌کنند، در نظر گرفته می‌شوند. این شرایط کیفی معمولاً بر اساس پارامترهایی مانند تعداد ترک‌های مرتبط با رأس، کیفیت ردیابی و سایر معیارها تعیین می‌شود. سپس، نزدیکترین رأس از میان رأس‌های با کیفیت خوب انتخاب می‌شود.

- **minVeryGoodVtx**: مشابه روش minGoodVtx ، اما با شرایط کیفی سختگیرانه‌تر برای انتخاب رأس‌ها استفاده می‌شود. این روش برای انتخاب رأس‌هایی با کیفیت بسیار بالا مناسب است.

- **particleDz**: این روش نیز مبتنی بر فاصله dz است اما به شکل متفاوتی محاسبات خود را انجام می‌دهد. در روش particleDz ، فاصله dz هر ذره نسبت به تمام رأس‌ها محاسبه و ثبت می‌شود. سپس این مقادیر مورد استفاده قرار می‌گیرند تا نزدیکترین رأس با دقت بیشتری انتخاب شود. در این روش، علاوه بر محاسبه دقیق فاصله، سایر

^۱Pile-Up Vertex

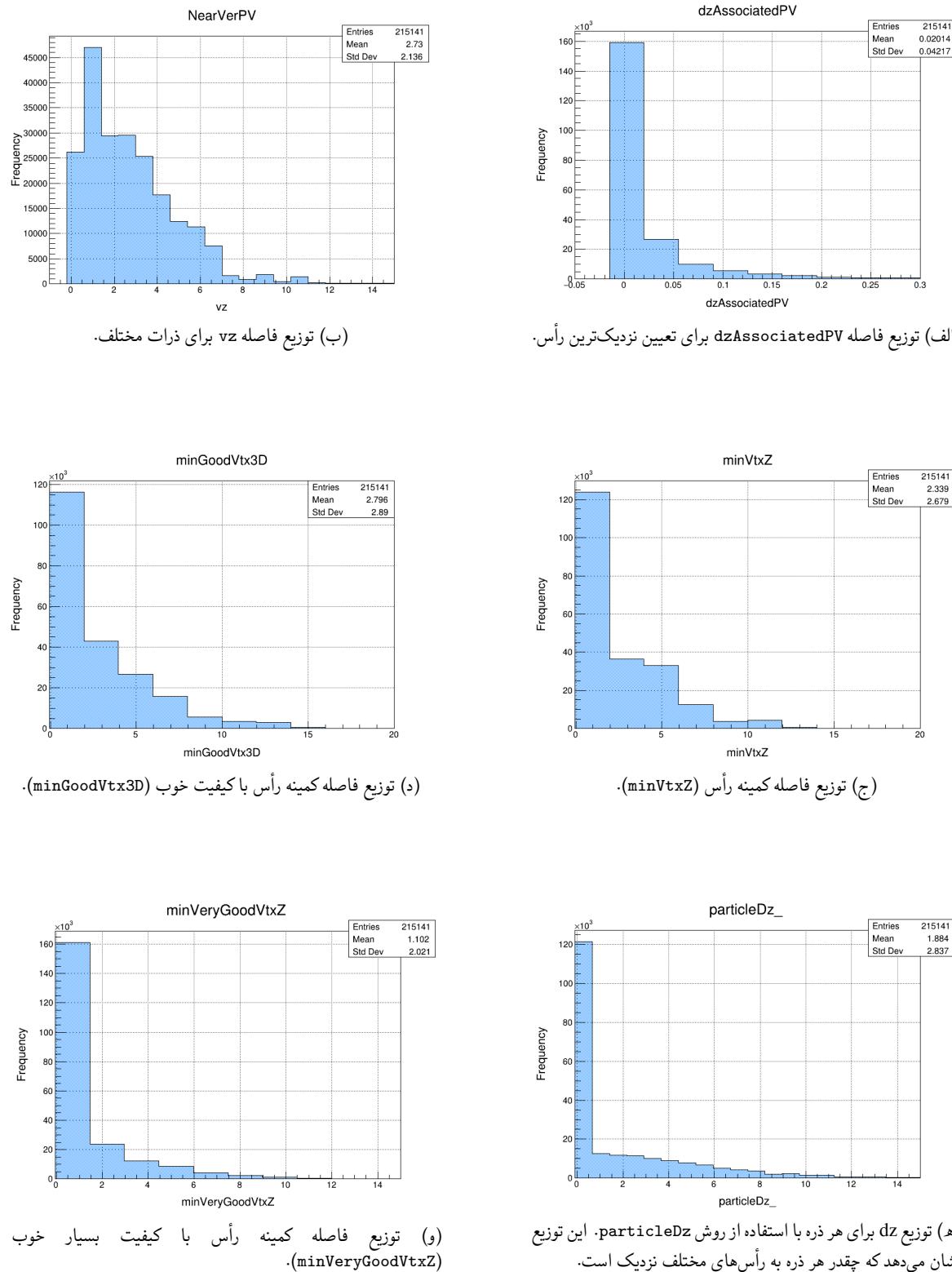
ویژگی‌های ذره نیز در نظر گرفته می‌شوند تا به انتخاب نزدیک‌ترین رأس کمک کنند.

• **NearVerPV**: در این روش، از اطلاعات ارتباط ذره با رأس اولیه استفاده می‌شود. رأس اولیه به عنوان یک نقطه مرجع

در نظر گرفته شده و نزدیک‌ترین رأس به این نقطه انتخاب می‌شود.

• **dzAssociatedPV**: این روش بر اساس فاصله ذره تا رأس اولیه محاسبه می‌شود. رأسی که دارای کمترین مقدار

است، به عنوان نزدیک‌ترین رأس در نظر گرفته می‌شود.



شکل ۲۳.۳: مقایسه توزیع‌های فاصله برای روش‌های مختلف تعیین نزدیکترین رأس.

۳.۳.۳ یافتن نزدیکترین رأس

هدف اصلی این بخش مقایسه‌ی عملکرد روش‌های مختلف در تعیین نزدیکترین رأس به ذرات است. برای این منظور، به نکات زیر توجه می‌کنیم:

- **تعیین کمترین فاصله:** برای هر ذره، رأس با کوچکترین فاصله (dz) را شناسایی می‌کنیم. توجه به توزیع ذرات در فاصله از نزدیکترین رأس با استفاده از روش‌های مختلف، بسیار مهم است. نمودارهایی که این توزیع را نشان می‌دهند، به ما کمک می‌کنند تا عملکرد هر روش را بررسی کنیم. هرچه قله‌ی نمودار به سمت چپ متمایل‌تر باشد، به این معناست که تعداد بیشتری از ذرات به رأس نزدیک‌تر هستند و بنابراین روش مورد نظر در پیدا کردن نزدیکترین رأس عملکرد بهتری دارد.

- **انتخاب روش با بیشترین تعداد ذرات در فاصله‌ی کم:** ابتدا نمودارهایی را که بیشترین تعداد ذرات در فاصله‌ی کم از رأس اولیه دارند، انتخاب می‌کنیم. روش‌هایی که در این دسته قرار دارند، در پیدا کردن نزدیکترین رأس برای تعداد بیشتری از ذرات موفق‌تر هستند و بنابراین انتخاب بهتری خواهند بود.

- **توجه به پهنانی توزیع:** علاوه بر تعداد ذرات، به پهنانی توزیع نیز توجه می‌کنیم. روش‌هایی که توزیع باریک‌تری دارند، دقت بیشتری در تعیین فاصله‌ی دقیق از رأس دارند. این توجه به پهنانی توزیع به ما کمک می‌کند تا روش‌هایی را شناسایی کنیم که از نظر دقت عملکرد بهتری دارند.

روش **dzAssociatedPV** در نمودار ۲۴.۳ نشان داده شده است که برای بیشتر ذرات فاصله‌ای نزدیک به رأس اولیه دارد و قله توزیع این روش به سمت چپ نمودار متمایل است. این به این معناست که **dzAssociatedPV** به تعداد بیشتری از ذرات نزدیکترین رأس را به درستی تشخیص می‌دهد. همچنین، پهنانی توزیع این روش نسبت به روش‌های دیگر کمتر است که نشان‌دهنده دقت بالای آن در تعیین فاصله دقیق از رأس است.

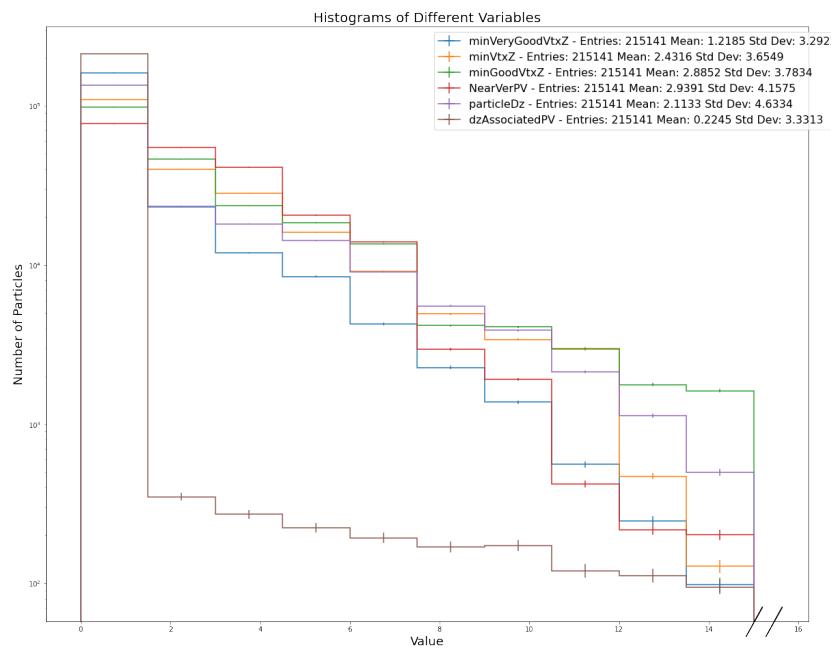
با توجه به نتایج به دست آمده از نمودار ۲۴.۳، روش **dzAssociatedPV** به عنوان بهترین روش انتخاب می‌شود. این روش بیشترین تعداد ذرات را در فاصله کم از رأس اولیه شناسایی کرده و همچنین پهنانی توزیع باریک‌تری نسبت به سایر روش‌ها دارد. این ویژگی‌ها نشان‌دهنده دقت و کارایی بالای این روش در تعیین نزدیکترین رأس به ذرات است.

- **آزمون pT :** هدف از آزمون pT این است که در هر روش بررسی کنیم که آیا رأس اول در هر رویداد دارای بیشترین تکانه عرضی است یا خیر. روشی که در آن رأس اول در هر رویداد بیشترین تکانه عرضی را داشته باشد، به عنوان روش بهینه انتخاب می‌شود.

سه روش مختلف که در آن‌ها شناسه‌های رأس و تکانه عرضی آن‌ها مقایسه می‌شوند، عبارتند از:

minVtxZIndex – در این روش شناسه‌ی رأس minVtxZ که به عنوان نزدیکترین رأس انتخاب شده است.

minVeryGoodVtxZIndex – این روش با استفاده از شناسه‌ی رأس minVeryGoodVtxZ که با دقت

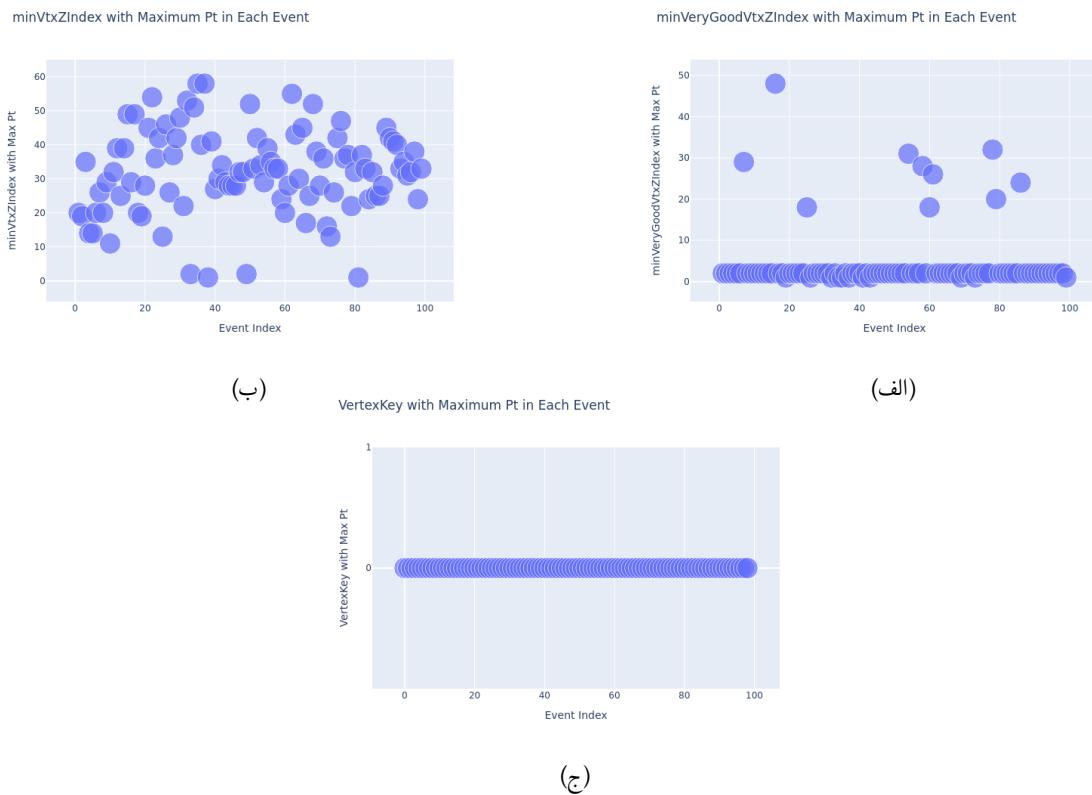


شکل ۲۴.۳: نمودار نتایج به دست آمده از روش‌های مختلف را نشان می‌دهد. روش‌های مختلف شامل `dzAssociatedPV`, `NearVerPV`, `minVeryGoodVtx`, `minGoodVtx`, `minVtx`

بیشتر شناسایی شده است.

شناخته شده است، `VertexKey - dzAssociatedPV` است، شناسه‌ی رأس مربوط به ذرهای که

به عنوان نزدیک‌ترین رأس شناخته شده است،



شکل ۲۵.۳: نمودارهای مربوط به توزیع شاخص رأس‌هایی که دارای بیشترین تکانه عرضی (pT) در هر رویداد هستند. در این نمودارها، سه روش مختلف شامل `minVtxZIndex`, `minVeryGoodVtxZIndex` و `VertexKey` مورد مقایسه قرار گرفته‌اند.

با توجه به نمودارها:

- در روش `minVtxZIndex` (شکل (ب)) مشاهده می‌شود که تکانه عرضی در اکثر رویدادها متعلق به رأس اول نیست، بلکه توزیع آن به صورت پراکنده است.
- در روش `minVeryGoodVtxZIndex` (شکل (الف)) نیز مشاهده می‌شود که تکانه عرضی متعلق به رأس اول نیست و توزیع مشابهی با روش قبلی دارد.
- در مقابل، در روش `dzAssociatedPV` با استفاده از شناسه `VertexKey` (شکل (ج)), تمامی رویدادها نشان می‌دهند که رأس اول همیشه دارای بیشترین تکانه عرضی است.

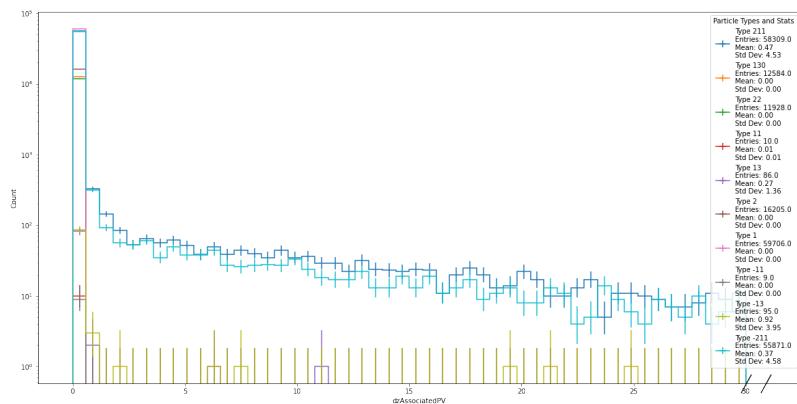
این نتایج نشان می‌دهد که روش‌های `minVeryGoodVtxZIndex` و `minVtxZIndex` در آزمون PT عملکرد مناسبی ندارند، زیرا تکانه عرضی در این روش‌ها متعلق به رأس اول نیست. در مقابل، روش `dzAssociatedPV` با استفاده از شناسه `VertexKey` بهترین عملکرد را دارد و در تمامی رویدادها، رأس اول بیشترین تکانه عرضی را دارد.

- **انتخاب نهایی:** با توجه به نتایج آزمون PT و تحلیل نمودارها، روش `dzAssociatedPV` با استفاده از شناسه `VertexKey` به عنوان روش نهایی و بهینه انتخاب می‌شود، چرا که در تمامی رویدادها، رأس اول دارای بیشترین

تکانه عرضی است. این روش بهترین عملکرد را برای تعیین نزدیکترین رأس با بیشترین تکانه عرضی نشان داده است.

۴.۳.۳ توزیع فاصله ذرات بر اساس نوع ذره

شکل ۲۶.۳ توزیع انواع ذرات را بر اساس فاصله آنها (dz) از نزدیکترین رأس که با استفاده از روش $dz_{AssociatedPV}$ تعیین شده است، نشان می‌دهد. هر خط نشان‌دهنده یک نوع ذره است و تعداد نمونه‌ها، میانگین و انحراف استاندارد مقادیر dz برای هر نوع ذره ارائه شده است.

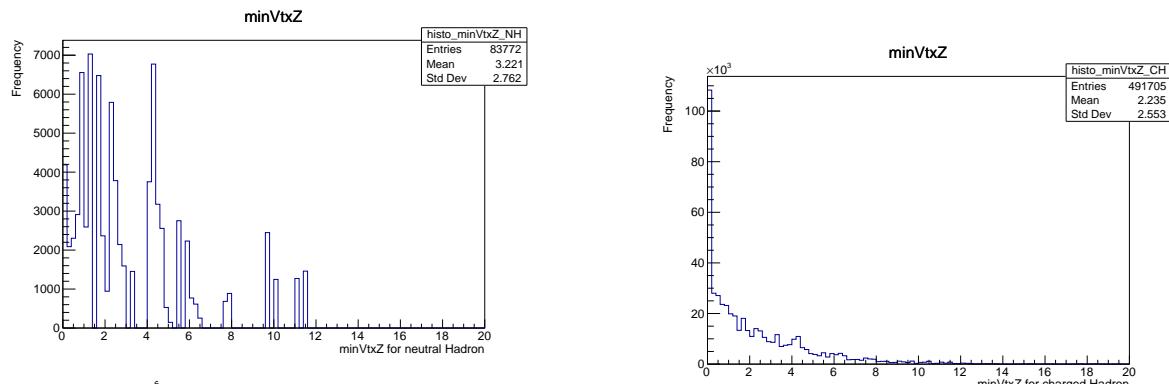


شکل ۲۶.۳: توزیع فاصله ذرات بر اساس نوع ذره

در شکل های بعدی، توزیع فاصله ذرات باردار مختلف (هادرون‌های باردار میون‌ها و الکترون‌ها) نسبت به رأس اولیه^۱ در راستای محور Z نمایش داده شده است. محور افقی مقدار فاصله ذره تا رأس اولیه بر حسب سانتی‌متر و محور عمودی تعداد رویدادها را نشان می‌دهد. قله هر توزیع نشان‌دهنده میانگین فاصله ذرات از رأس اولیه است.

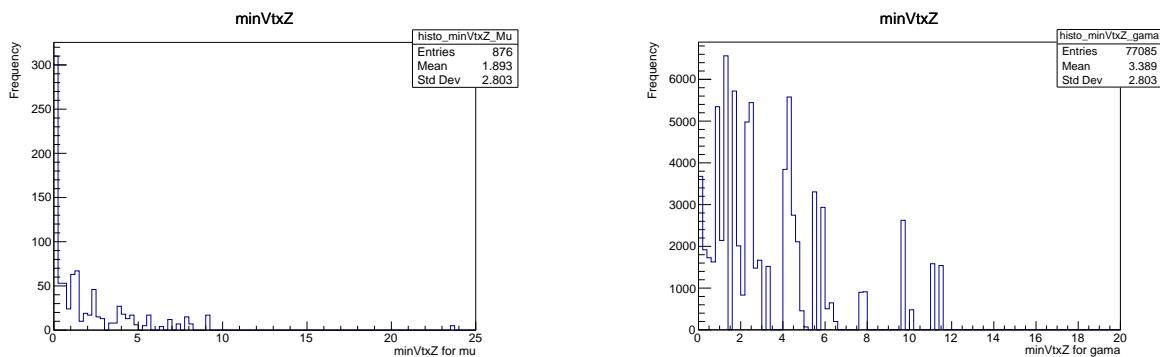
همانطور که مشاهده می‌شود، توزیع فاصله برای ذرات مختلف متفاوت است. این تفاوت به دلیل تفاوت در فرایند تولید، برهم‌کنش‌ها و عمر ذرات مختلف است. به عنوان مثال، الکترون‌ها به دلیل عمر طولانی‌تر، تمایل دارند فاصله بیشتری از رأس اولیه داشته باشند.

¹Primary Vertex



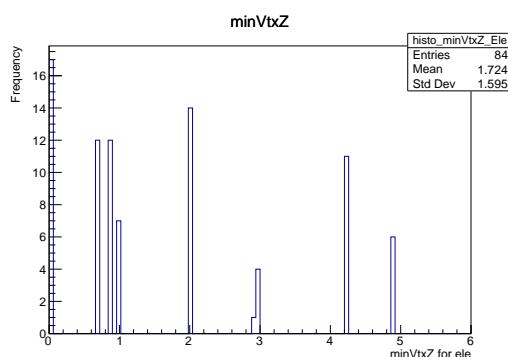
(ب) توزیع فاصله هادرон‌های خنثی نسبت به رأس اولیه. به دلیل عدم برهمنکش مستقیم با میدان‌های الکتریکی و مغناطیسی، روش‌های تشخیص و ردیابی آنها متفاوت است.

(الف) توزیع فاصله هادرون‌های باردار نسبت به رأس اولیه. اکثر هادرون‌های باردار در فاصله‌ای حدود ۲ از محور Z ایجاد می‌شوند.



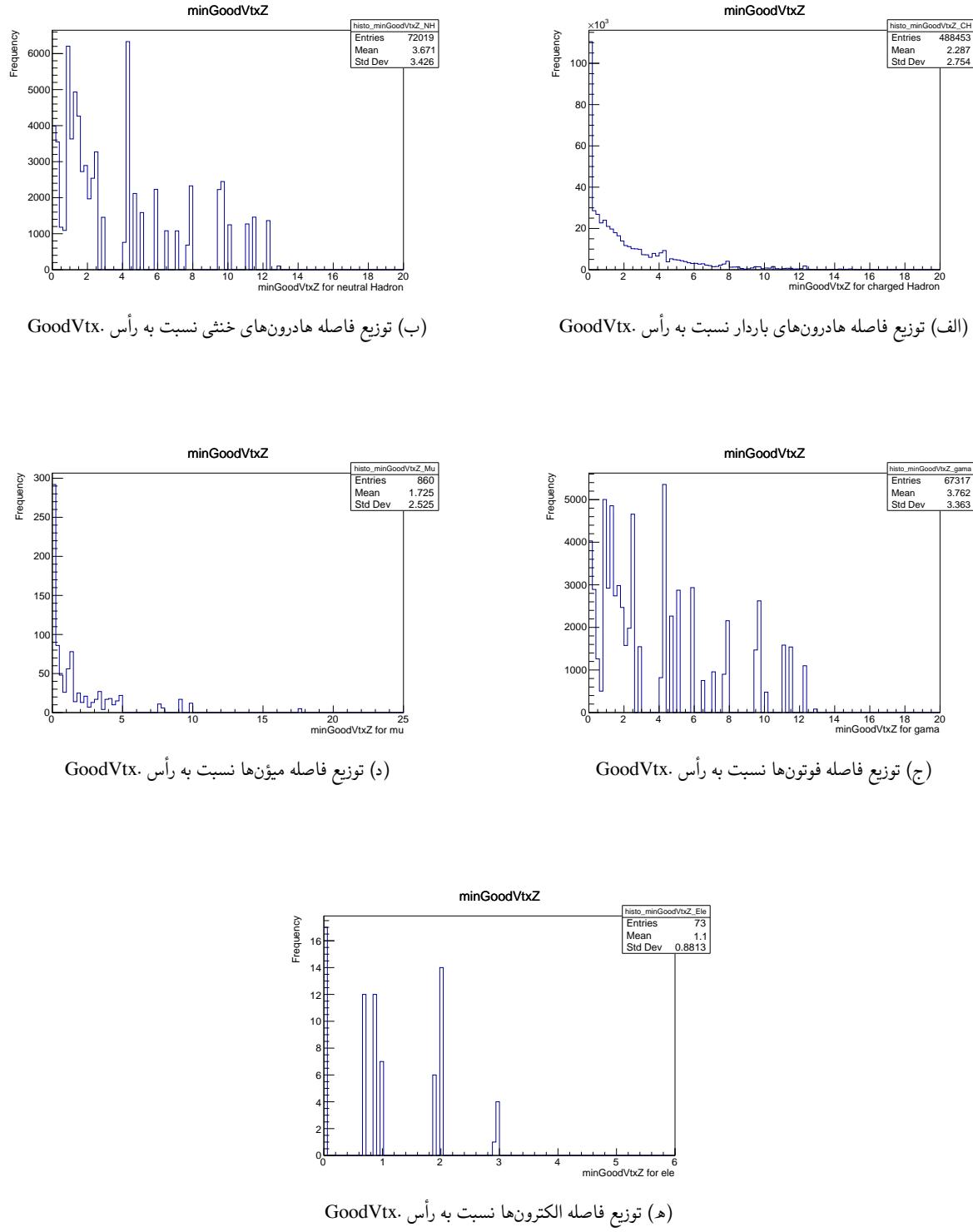
(د) توزیع فاصله میون‌ها نسبت به رأس اولیه. میون‌ها بیشتر در نزدیکی محور Z ایجاد می‌شوند.

(ج) توزیع فاصله فوتون‌ها نسبت به رأس اولیه. فوتون‌ها در فاصله‌ای متوسط بین میون‌ها و هادرون‌های باردار ایجاد می‌شوند.

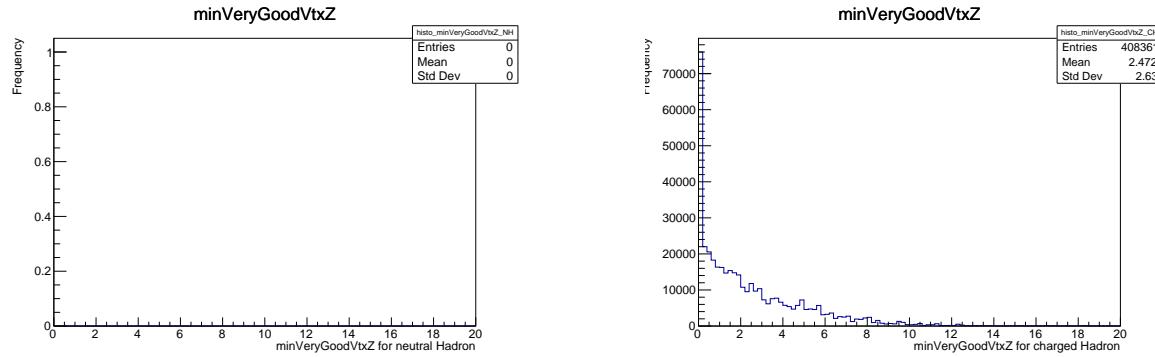


(ه) توزیع فاصله الکترون‌ها نسبت به رأس اولیه. اکثر الکترون‌ها بسیار نزدیک به محور Z ایجاد می‌شوند.

شکل ۲۷.۳: توزیع فاصله انواع مختلف ذرات نسبت به رأس اولیه در راستای محور Z. هر نمودار نشان‌دهنده نزدیکترین فاصله هر ذره نسبت به Vtx (رأس اولیه) است.

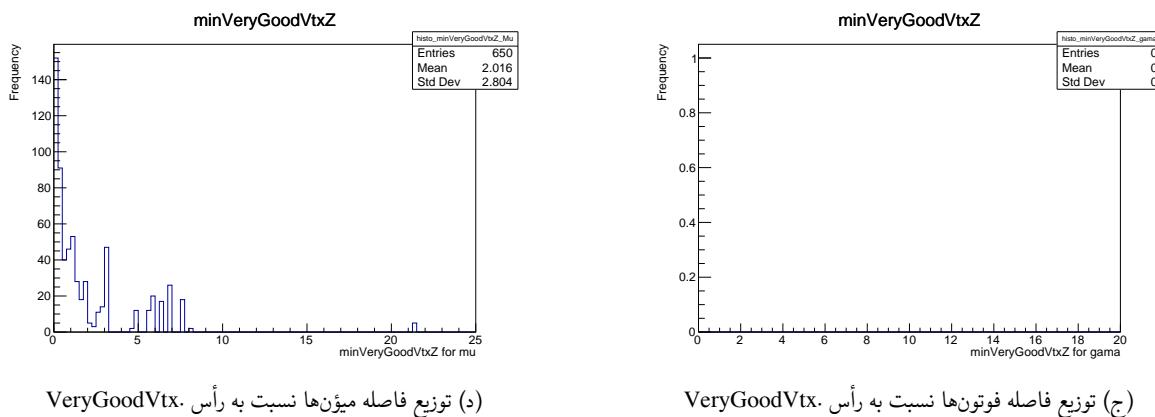


شکل ۲۰.۳: توزیع فاصله انواع مختلف ذرات نسبت به رأس. GoodVtx. هر نمودار نشان‌دهنده تزدیک‌ترین فاصله هر ذره نسبت به GoodVtx است.



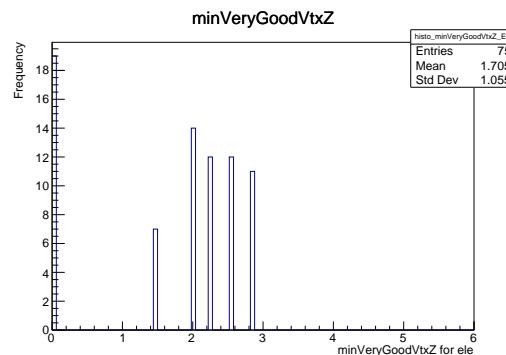
(ب) توزیع فاصله هادرون‌های خنثی نسبت به رأس. VeryGoodVtx.

(الف) توزیع فاصله هادرون‌های باردار نسبت به رأس. VeryGoodVtx.



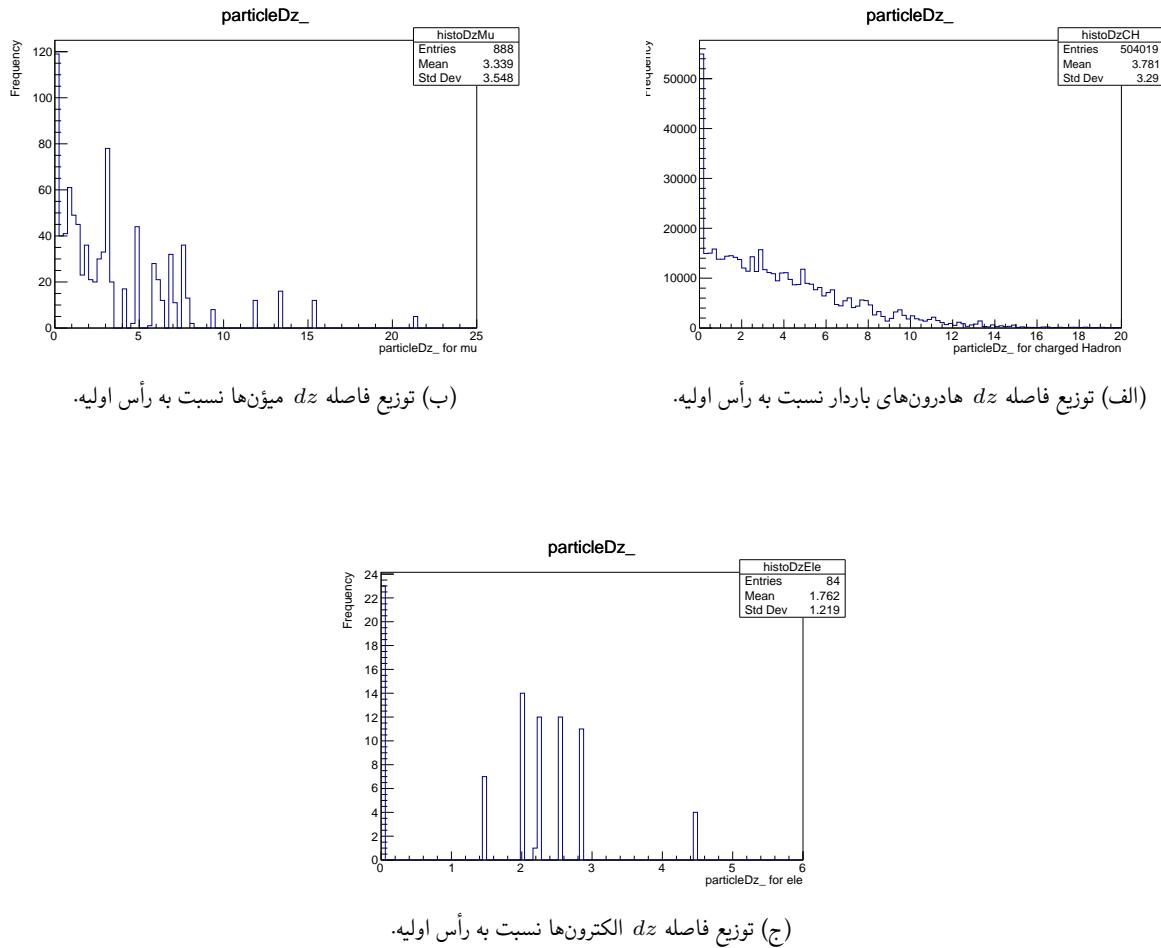
(د) توزیع فاصله میؤن‌ها نسبت به رأس. VeryGoodVtx.

(ج) توزیع فاصله فوتون‌ها نسبت به رأس. VeryGoodVtx.



(ه) توزیع فاصله الکترون‌ها نسبت به رأس. VeryGoodVtx.

شکل ۲۹.۳: توزیع فاصله انواع مختلف ذرات نسبت به رأس. VeryGoodVtx. هر نمودار نشان‌دهنده نزدیکترین فاصله هر ذره نسبت به VeryGoodVtx است.



شکل ۳۰.۳: توزیع فاصله dz ذرات باردار مختلف نسبت به رأس اولیه در راستای محور Z . محور افقی مقدار فاصله dz بر حسب سانتی‌متر و محور عمودی تعداد رویدادها را نشان می‌دهد.

فصل چهارم

شبکه‌های عصبی گرافی

فیزیکدانان برای اندازه‌گیری ذرات بنیادی، آزمایش‌هایی طراحی می‌کنند و از روش‌های آماری برای بررسی صحت مدل‌های مختلف استفاده می‌کنند. داده‌های حاصل از این آزمایش‌ها اغلب نمونه‌گیری پراکنده‌ای از یک فرایند فیزیکی در زمان و مکان هستند. یادگیری ماشین به طور سنتی نقش مهمی در فیزیک ذرات داشته است، با کاربردهای دسته‌بندی^۱ و وايانش^۲ با استفاده از تکنیک‌های کلاسیک مانند درختان تصمیم‌گیری^۳ تقویت شده، ماشین بردار پشتیبان^۴، پرسپترون‌های چندلایه ساده^۵ و غیره. با الهام از موفقیت یادگیری عمیق در دستیابی به عملکرد فراتر از انسان در کارهای مختلف، حوزه‌های متنوعی در علوم فیزیکی، از جمله فیزیک ذرات، شروع به کاوش یادگیری عمیق به عنوان ابزاری منحصر به فرد برای رسیدگی به مشکلات علمی دشوار کرده‌اند که فراتر از دسته‌بندی ساده است. این ابزار به سازماندهی و درک منابع عظیم داده، استنتاج درباره‌ی عوامل علیت مشاهده‌نشده و حتی کشف اصول فیزیکی زیربنای پدیده‌های پیچیده کمک می‌کند.

آزمایش‌های فیزیک انژی بالا از ماشین برای یادگیری توابع معکوس پیچیده استفاده می‌کنند. در این روش، هدف استنتاج ویژگی‌های فرایند فیزیکی زیربنایی از اطلاعاتی است که آشکارساز اندازه‌گیری می‌کند. شکل ۱۰.۴ این طرح را به تصویر می‌کشد.

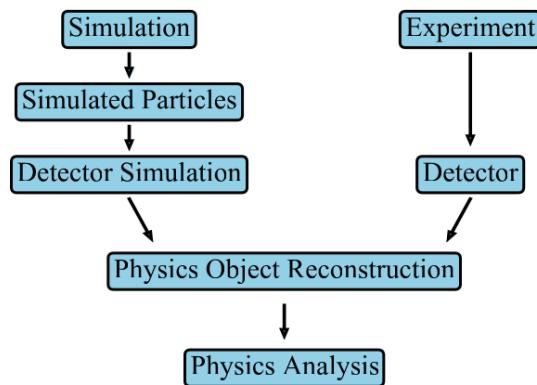
¹classification

²regression

³boosted decision trees

⁴support vector machine

⁵multi-layer perceptrons



شکل ۱.۴: شکل نشان می‌دهد که چگونه از شبیه‌سازی در آزمایش‌های فیزیک انرژی بالا برای یادگیری معکوس پیچیده استفاده می‌شود. در این روش، ابتدا فرایند فیزیکی با استفاده از شبیه‌سازی، بازسازی می‌شود تا یک «سوابق حقیقی» از رویداد فیزیکی که منجر به پاسخ خاصی در آشکارساز شده است، به دست آید. سپس از این «سوابق حقیقی» برای آموزش الگوریتم‌های یادگیری با نظرارت استفاده می‌شود. این الگوریتم‌ها یاد می‌گیرند که شبیه‌سازی آشکارساز را معکوس کنند و از داده‌های اندازه‌گیری شده واقعی، ویژگی‌هایی در مورد فرایند فیزیکی زیربنایی استنتاج نمایند. در نهایت، این الگوریتم‌های آموزش دیده بر روی داده‌های واقعی که توسط آشکارساز اندازه‌گیری شده‌اند، به کار گرفته می‌شوند.

ظهور شبکه‌های عصبی گرافی در فیزیک انرژی بالا

در حالی که سه نوع رایج از اجزای سازنده‌ی یادگیری عمیق - شبکه‌های کاملاً متصل^۱، شبکه عصبی پیچشی^۲ و شبکه‌های عصبی بازگشتی^۳ - در بسیاری از حوزه‌های علمی مفید واقع شده‌اند، تمرکز این آنالیز بر روی کلاسی از معماری‌ها به نام شبکه‌های عصبی گرافی^۴ است، همانطور که در ادامه توضیح داده می‌شود، ما توجه خود را به عنوان یک معماری مبتنی بر گراف معطوف می‌کنیم. شبکه‌های عصبی گرافی را می‌توان با استفاده از داده‌ها آموزش داد تا توابعی را روی گراف‌ها بیاموزند. بسیاری از مسائل شامل داده‌هایی هستند که به صورت مجموعه‌های بدون ترتیب از عناصر با روابط و تعاملات غنی با یکدیگر نشان داده می‌شوند و می‌توان آنها را به طور طبیعی به صورت گراف بیان کرد.

ما از داده‌های آزمایش CMS استفاده کردیم. این آشکارساز در هر ثانیه میلیون‌ها اندازه‌گیری با ابعاد بالا را ثبت می‌کنند و از زیرآشکارسازهای مختلفی استفاده می‌کند. همانطور که در فصول قبل گفته شد داده‌های خام به دست آمده از آشکارساز حجمی و پیچیده هستند. هر برخورد ذره‌ای منجر به مجموعه‌ای از سیگنال‌های الکترونیکی می‌شود که در زیرآشکارسازهای مختلف ثبت می‌گردد. این سیگنال‌ها نیاز به پردازش و بازسازی اولیه دارند تا بتوانند ویژگی‌های فیزیکی ذرات برهمنش‌کننده را استخراج

¹Fully Connected (FC)

²Convolutional neural network (CNN)

³Recurrent neural network (RNN)

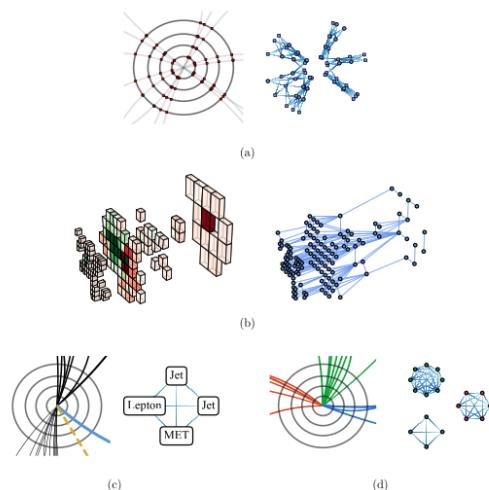
⁴Graph neural network (GNN)

کنند. داده‌های بازسازی شده شامل اطلاعاتی مانند انرژی، تکانه، جهت و نوع ذرات است.

برای تجزیه و تحلیل دقیق‌تر، داده‌های بازسازی شده به صورت گراف نشان داده می‌شوند. در این گراف، گره‌ها نشان‌دهنده‌ی ذرات شناسایی شده و یال‌ها نشان‌دهنده‌ی روابط بین آن‌ها (مانند فاصله) هستند. این نمایش گرافی به الگوریتم‌های یادگیری ماشین امکان می‌دهد تا روابط پیچیده بین ذرات را بیاموزند و ویژگی‌های فیزیکی برخورد را با دقت بیشتری استخراج کنند. برخلاف آشکارسازهای فیزیک انرژی بالا، آشکارسازهای اختوفیزیک معمولاً بزرگ‌تر (تا کیلومترها) بوده و با یک تکنولوژی اندازه‌گیری واحد ساخته می‌شوند (مانند IceCube و Antares). به همین دلیل، داده‌های اختوفیزیک همگن هستند. در هر دو مورد، اندازه‌گیری‌ها به دلیل طراحی هندسی حسگرها، ذاتاً در فضا پراکنده‌اند. بنابراین، ماهیت داده‌ها از ابتدا با ساختارهای داده‌ی همگن و شبکه‌ای همخوانی ندارد.

۱۰.۴ نمایش گرافی

شکل ۲۰.۴ نشان می‌دهد که داده‌های LHC را می‌توان برای بسیاری از کاربردها به صورت گراف نشان داد:



شکل ۲۰.۴: (الف) خوشبندی برخوردهای آشکارساز ردیابی به مسیرها (ب) تقسیم سلول‌های کالری‌متر (پ) طبقه‌بندی رویدادها با انواع مختلف اشیاء فیزیکی (ت) طبقه‌بندی جت بر اساس ذرات مرتبط با جت

(الف) خوشبندی برخوردهای آشکارساز ردیابی به مسیرها: در این کاربرد، گراف برخوردهای ثبت شده توسط آشکارساز ردیابی را به عنوان گره نشان می‌دهد. یال‌ها گره‌هایی را به هم وصل می‌کنند که احتمالاً از مسیر حرکت یک ذره خاص نشأت گرفته‌اند. با تحلیل گراف، محققان می‌توانند مسیر حرکت ذرات را در آشکارساز شناسایی و بازسازی کنند.

(ب) تقسیم سلول‌های کالری‌متر: کالری‌مترها انرژی ذرات را اندازه‌گیری می‌کنند. در این کاربرد، گراف سلول‌های کالری‌متر را به عنوان گره نشان می‌دهد. یال‌ها سلول‌های همسایه را که احتمالاً از یک بارش ذرات مشترک نشأت گرفته‌اند، به هم وصل

می‌کنند. با تحلیل گراف، محققان می‌توانند سلول‌های کالری‌متر را با هم گروه کرده و خوش‌های انژی یا «ابرخ خوش‌ها» را تشکیل دهند.

(پ) دسته‌بندی رویدادها با انواع مختلف اشیاء فیزیکی: رویدادهای فیزیکی می‌توانند شامل انواع مختلفی از ذرات باشند. در این کاربرد، گراف اشیاء فیزیکی بازسازی شده (مانند مسیرها، الکترون‌ها، میون‌ها) را به عنوان گره نشان می‌دهد. یال‌ها اشیائی را به هم وصل می‌کنند که احتمالاً از یک فرایند فیزیکی پایه‌ای مشترک نشأت گرفته‌اند. با تحلیل گراف، محققان می‌توانند رویدادها را بر اساس نوع و روابط بین اشیاء بازسازی شده طبقه‌بندی کنند.

(ت) دسته‌بندی جت بر اساس ذرات مرتبط با جت: جت‌ها افشاره‌های باریکه‌ای از ذرات هستند که در برخوردهای ذرات با انژی بالا تولید می‌شوند. در این کاربرد، گراف ذرات مرتبط با یک جت را به عنوان گره نشان می‌دهد. یال‌ها ذرات درون جت را به هم وصل می‌کنند. با تحلیل گراف، محققان می‌توانند در مورد خواص جت، مانند انژی و ترکیب آن، اطلاعات کسب کنند. داده‌های فیزیک انژی بالا را می‌توان تا حدودی به عنوان تصویر تفسیر کرد و بنابراین، تکنیک‌های بینایی رایانه (شبکه‌های عصبی پیچشی) با کارایی بهبود یافته در حال اعمال هستند. با این حال، نمایش‌های تصویری با برخی محدودیت‌ها مانند هندسه نامنظم آشکارسازها یا پراکندگی داده‌ها مواجه هستند. به دلیل از دست رفتن ذاتی اطلاعات، نمایش‌های تصویری ممکن است میزان اطلاعات قابل استخراج از داده‌ها را محدود کنند.

اندازه‌گیری‌ها و اشیاء بازسازی شده را می‌توان به عنوان توالی در نظر گرفت که ترتیبی بر اساس درک نظری یا تجربی داده‌ها بر آنها اعمال می‌شود. بنابراین، روش‌هایی که در غیر این صورت برای پردازش زبان طبیعی به کار می‌روند، مورد بررسی قرار گرفته‌اند. در حالی که ترتیبی که استفاده می‌شود معمولاً از نظر تجربی قابل توجیه است، اما اغلب تحمیل شده و بنابراین نحوه ارائه داده‌ها به مدل‌ها را محدود می‌کند. همچنین در برخی موارد می‌توان این ترتیب را با استفاده از دانش تجربی قبلی در مورد فرایند فیزیکی مورد نظر آموخت. با این حال، این همیشه صادق نیست و ممکن است انتظار رود که ترتیب تحمیل شده، عملکرد یادگیری را کاهش دهد - ترتیبی که همانطور که در ادامه خواهیم دید، ضروری نیست. به عنوان مثال، شواهدی را نشان می‌دهد که یک شبکه‌ی بدون تغییر در برابر جایگشت، عملکرد بهتری نسبت به یک الگوریتم مبتنی بر توالی که از دقیقاً همان ویژگی‌های ورودی برای همان کار دسته‌بندی استفاده می‌کند، دارد.

داده‌ها در بسیاری از سطوح، ذاتاً مجموعه‌ای (مجموعه‌ای بدون ترتیب) از آیتم‌ها هستند. اگر روابط بین آیتم‌ها (هندسی یا فیزیکی) را در نظر بگیریم، یک مجموعه با افزودن ماتریس مجاورت به یک گراف تبدیل می‌شود. محدودیت کمتری برای اعمال یادگیری عمیق بر روی این نمایش ذاتی داده‌ها نسبت به سایر سطوح ذکر شده در بالا وجود دارد. نمونه‌هایی از داده‌های مختلف فیزیک انژی بالا و فرمول‌بندی آن‌ها به صورت گراف در شکل ۲.۴ نشان داده شده است.

۲۰.۴ شبکه‌های عصبی

شبکه‌های عصبی عمیق^۱ در دهه گذشته نقشی اساسی در پیشرفت‌های یادگیری ماشین و هوش مصنوعی داشته‌اند، و می‌توان آن‌ها را حاصل همگرایی چند عامل کلیدی دانست. اولًا، شبکه‌های عصبی بزرگ می‌توانند توابع بسیار پیچیده‌ای را نشان دهند. دوم، اطلاعات ارزشمند موجود در داده‌های حجیم را می‌توان از طریق روش‌های آموزشی مبتنی بر گرادیان، در پارامترهای شبکه‌های عصبی بزرگ کدگذاری کرد. سوم، سخت‌افزار رایانه‌ای موازی می‌تواند چنین آموزشی را در مدت زمان چند ساعت یا چند روز انجام دهد که برای بسیاری از موارد استفاده مهم، به اندازه‌ی کافی کارآمد است. چهارم، چارچوب‌های نرم‌افزاری به خوبی طراحی شده، مانند TensorFlow و PyTorch موانع فنی توسعه و توزیع کاربردهای یادگیری عمیق را کاهش می‌دهند و ابزارهای قدرتمند یادگیری ماشین را برای افراد متخصص به طور گسترده در دسترس قرار می‌دهند.

داده‌های با ساختار گراف در سراسر علوم، مهندسی و بسیاری از حوزه‌های دیگر مسائل فراگیر هستند. یک گراف، حداقل، به عنوان مجموعه‌ای از گره‌ها و همچنین مجموعه‌ای از یال‌های مجاور به زوج‌های گره تعریف می‌شود. انواع غنی‌تر و موارد خاص‌تر شامل: درخت‌ها، که در آن‌ها دقیقاً یک توالی از یال‌ها هر دو گره را به هم وصل می‌کند؛ گراف‌های جهت‌دار، که در آن‌ها دو گرهی مرتبط با یک یال مرتب هستند؛ گراف‌های ویژگی‌دار، که شامل ویژگی‌های سطح گره، سطح یال یا سطح گراف هستند؛ چندگراف‌ها، که در آن‌ها ممکن است بیش از یک یال بین یک یال جفت گردد و وجود داشته باشد؛ ابر گراف‌ها، که در آن‌ها بیش از دو گره با یک یال مرتبط هستند و غیره. نکته‌ی اساسی این است که گراف‌ها روشی طبیعی و قدرتمند برای نمایش بسیاری از سیستم‌های پیچیده هستند، به عنوان مثال درخت‌ها برای نمایش تکامل گونه‌ها، یا ساختار سلسله‌مراتبی جملات؛ شبکه‌ها و توری‌ها به ترتیب برای نمایش گسترش‌سازی‌های منظم و نامنظم فضا؛ شبکه‌های پویا برای نمایش ترافیک جاده‌ها و روابط اجتماعی در طول زمان.^[۱۲]

شبکه‌های عصبی گراف کلاسی از معماری‌های یادگیری عمیق هستند که سوگیری‌های القایی رابطه‌ای قوی را برای یادگیری توابعی که روی گراف‌ها عمل می‌کنند، پیاده‌سازی می‌کنند. آن‌ها شکلی از ارسال پیام پارامتری را اجرا می‌کنند که از طریق آن اطلاعات در سراسر گراف منتشر می‌شود و اجازه می‌دهد تا خروجی‌های پیچیده‌ی سطح لبه، گره و گراف محاسبه شوند. در داخل یک شبکه عصبی گرافی، یک یا چند بلوك سازنده‌ی استاندارد شبکه‌ی عصبی، به‌طور معمول لایه‌های کاملاً متصل، وجود دارند که محاسبات پیام و توابع انتشار را پیاده‌سازی می‌کنند.^[۱۳]

فرمول‌بندی شبکه گراف

در اینجا بر فرمول‌بندی شبکه گراف^۲ تمرکز می‌کنیم، که شبکه‌های عصبی گراف مختلف و همچنین روش‌های دیگر را تعمیم می‌دهد. شبکه گراف توابعی از گراف به گراف هستند که گراف‌های خروجی آن‌ها همان ساختار گره و یال ورودی را دارند. با

¹Deep Neural network

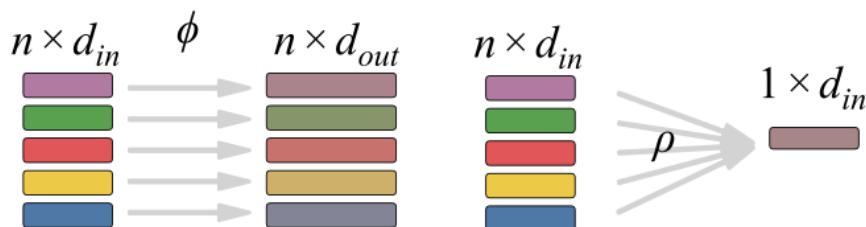
²Graph Network

پذیرش فرمول‌بندی، یک گراف را می‌توان به صورت $G = (\mathbf{u}, V, E)$ نشان داد، با N_v رأس و N_e یال. \mathbf{u} نشان‌دهنده ویژگی‌های سطح گراف است. مجموعه گره‌ها (یا رأس‌ها) $V = \{\mathbf{v}_i\}_{i=1:N_v}$ هستند، جایی که \mathbf{v}_i نشان‌دهنده ویژگی‌های گره i است. مجموعه یال‌ها $E = \{(\mathbf{e}_k, r_k, s_k)\}_{k=1:N_e}$ هستند، جایی که \mathbf{e}_k نشان‌دهنده ویژگی‌های یال k است و r_k و s_k شاخص‌های دو گره (به ترتیب گیرنده و فرستنده) متصل به یال k است [۱۲].

مراحل پردازش در یک بلوک شبکه گراف

$$\begin{array}{ll} \mathbf{e}'_k = \phi^e \left(\mathbf{e}_k, \mathbf{v}_{l_k}, \mathbf{v}_{s_k}, \mathbf{u} \right) & \bar{\mathbf{e}}'_i = \rho^{e \rightarrow v} \left(E'_i \right) \\ \mathbf{v}'_i = \phi' \left(\bar{\mathbf{e}}'_i, \mathbf{v}_i, \mathbf{u} \right) & \bar{\mathbf{e}}' = \rho^{e \rightarrow u} \left(E' \right) \\ \mathbf{u}' = \phi^u \left(\bar{\mathbf{e}}', \bar{\mathbf{v}}, \mathbf{u} \right) & \bar{\mathbf{v}}' = \rho^{v \rightarrow u} \left(V' \right) \end{array} \quad \begin{array}{l} \text{بلوک یال} \\ \text{بلوک گره} \\ \text{بلوک سراسری} \end{array}$$

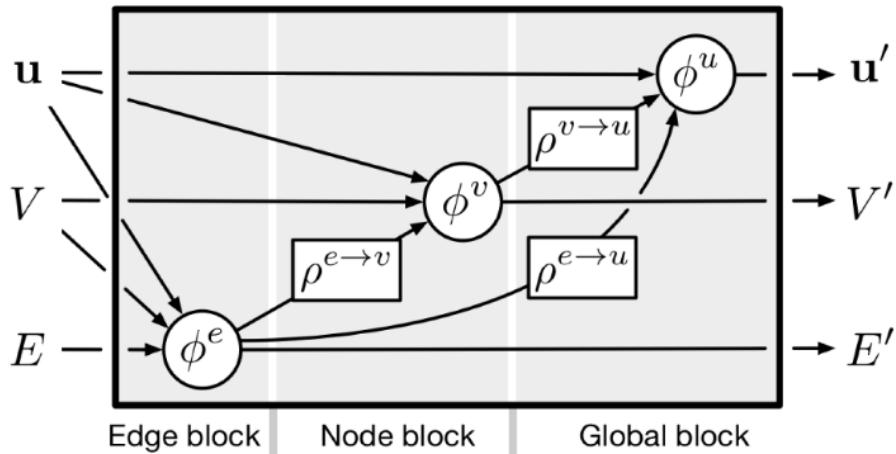
یک بلوک شبکه گراف شامل ۶ تابع داخلی است: ۳ تابع بهروزرسانی (ϕ^e, ϕ^v و ϕ^u) و ۳ تابع جمع‌آوری ($\rho^{e \rightarrow v}, \rho^{e \rightarrow u}$ و $\rho^{u \rightarrow v}$). فرمول‌بندی شبکه گراف یک معماری مدل خاص نیست، تعیین نمی‌کند که این توابع دقیقاً چیستند. توابع بهروزرسانی توابعی با ورودی و خروجی با اندازه ثابت هستند، و توابع جمع‌آوری یک مجموعه ورودی با اندازه متغیر (مانند مجموعه‌ای از یال‌های متصل به یک گره خاص) را به یک نمایش با اندازه ثابت تبدیل می‌کنند. این در شکل ۳.۴ نشان داده شده است [۱۲]. بلوک یال یک خروجی برای هر یال، \mathbf{e}'_k ، محاسبه می‌کند و آن‌ها را بر اساس گره دریافت‌کننده متناظر، $\bar{\mathbf{e}}'_i$ ، جمع‌آوری می‌کند.



شکل ۳.۴: اجزای داخلی یک بلوک شبکه گراف تابع بهروزرسانی و توابع تجمع هستند. شکل سمت چپ توابع بهروزرسانی مجموعه‌ای از اشیاء با نمایش دارای اندازه‌ی ثابت را دریافت می‌کنند و همین تابع را روی هر یک از عناصر مجموعه اعمال می‌کنند که منجر به یک نمایش بهروزرسانی شده (همچنین با اندازه‌ی ثابت) می‌شود. شکل سمت راست توابع تجمع مجموعه‌ای از اشیاء را دریافت می‌کنند و با استفاده از یک تابع ترتیب‌ناپذیر برای گروه‌بندی نمایش اشیاء (مانند مجموع روى عناصر)، یک نمایش با اندازه‌ی ثابت برای کل مجموعه ایجاد می‌کنند.

جایی که E'_i مجموعه‌ی یال‌های واقع بر روی گره i است. بلوک گره یک خروجی برای هر گره، \mathbf{v}'_i ، محاسبه می‌کند. خروجی‌های سطح یال و گره همه به هم جمع‌آوری می‌شوند تا بلوک سراسری را محاسبه کنند. خروجی شبکه گراف مجموعه‌ای از

تمامی خروجی‌های سطح یال، گره و گراف است، $G' = (\mathbf{u}', V', E')$. به شکل ۱۰.۴ مراجعه کنید.



شکل ۱۰.۴: یک بلوک شبکه گراف. یک گراف ورودی، $G = (u, V, E)$ ، پردازش می‌شود و یک گراف با همان ساختار لبه اما ویژگی‌های متفاوت، $G' = (u', V', E')$ ، به عنوان خروجی برگردانده شده‌اند.

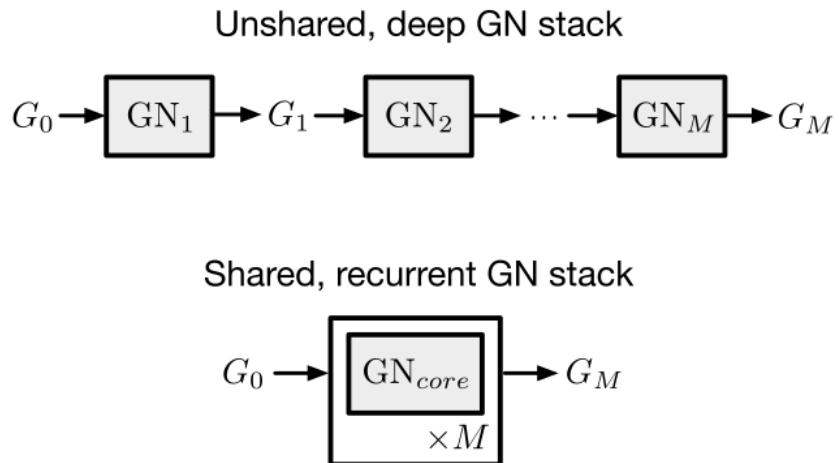
در عمل، توابع ϕ^e, ϕ^v و ϕ^u به عنوان یک شبکه عصبی آموزش‌پذیر ساده پیاده‌سازی می‌شوند، مثل یک شبکه کاملاً متصل. توابع ρ^{e-v}, ρ^{e-u} و ρ^{v-u} به عنوان عملگرهای کاهش تبدیل ناپذیر پیاده‌سازی می‌شوند^۱، مانند جمع‌های عنصر به عنصر، میانگین‌ها یا بیشینه‌ها. اگر بلوک شبکه گرافی بخواهد خاصیت هم‌ارزی تبدیل‌پذیری را حفظ کند، تابع ρ باید تبدیل ناپذیر باشد. [۱۳]

مزایای کلیدی شبکه‌های گراف

برخی از مزایای کلیدی شبکه‌های گراف عمومی بودن آن‌هاست: اگر مسئله‌ای نیاز به نگاشت گراف به گراف دیگر یا برخی خروجی‌های خلاصه داشته باشد، شبکه‌های گراف اغلب مناسب هستند. همچنین، آن‌ها تمایل دارند به خوبی به گراف‌هایی که در طول آموزش، دیده نشده‌اند، تعمیم دهند، زیرا یادگیری بر روی سطح یال و گره متمرکز است - در واقع، اگر بلوک سراسری حذف شود، شبکه‌های گراف حتی از کل گراف در هیچ یک از محاسباتش آگاه نیست، زیرا بلوک‌های یال و گره فقط مناطق مربوطه خود را به عنوان ورودی می‌گیرند. با این حال، وقتی چندین بلوک شبکه‌های گراف به صورت عمیق یا بازگشتی تنظیم شوند، همانطور که در شکل ۵.۴ نشان داده شده است، اطلاعات می‌توانند در سراسر ساختار گراف پردازش و انتشار یابند تا محاسبات پیچیده‌تر و بلند برد انجام شوند. [۱۳]

فمول‌بندی شبکه‌های گراف یک چارچوب کلی است که می‌تواند انواع مختلفی از معماری‌های شبکه‌های عصبی گرافی دیگر را در بر بگیرد. این معماری‌ها می‌توانند با حذف یا بازآرایی اجزای داخلی بلوک شبکه‌های گراف عمومی در شکل‌های ۵.۴ و ۵.۵

^۱permutation invariant reduction operators



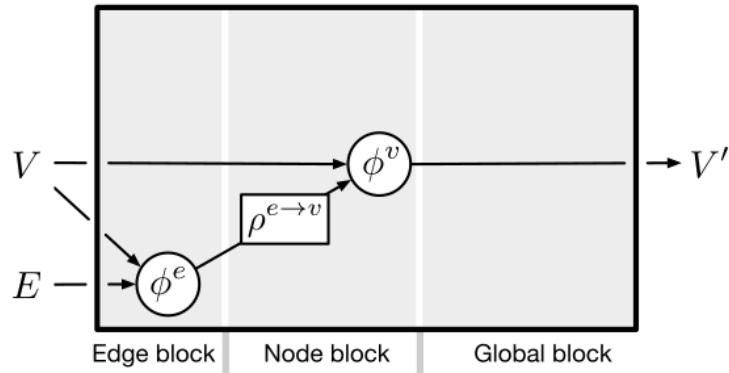
شکل ۵.۴: بلوک‌های شبکه‌های گراف را می‌توان در معماری‌های محاسباتی پیچیده‌تر ترکیب کرد. ردیف بالا توالی‌ای از بلوک‌های شبکه‌های گراف مختلف را نشان می‌دهد که به صورت سری یا عمقی مرتب شده‌اند. ردیف پایین بلوک‌های شبکه‌های گراف متمایز را با یک پیکربندی مشترک و بازگشتی جایگزین می‌کند.

و پیاده‌سازی توابع مختلف ϕ و ρ با استفاده از فرم‌های عملکردی خاص، بیان شوند. برای مثال، یکی از معماری‌های بسیار محبوب شبکه‌های عصبی گرافی، شبکه پیچشی گراف^۱ است. با استفاده از فرمول‌بندی شبکه گرافی، یک شبکه پیچشی گراف می‌تواند به صورت زیر بیان شود:

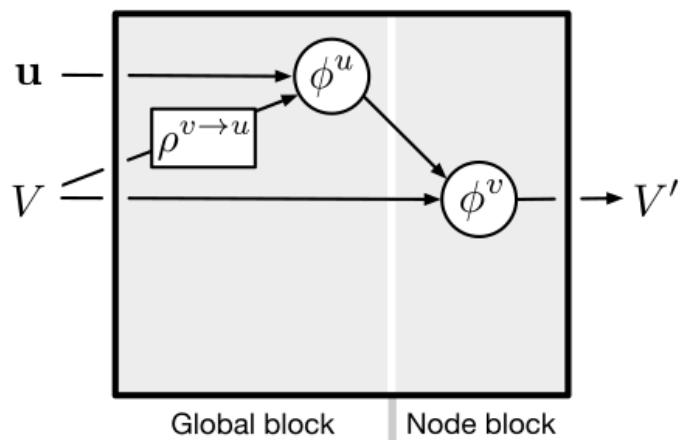
$$\begin{aligned} \mathbf{e}'_k &= \phi^e \left(\mathbf{e}_k, \mathbf{v}_{s_k} \right) = \mathbf{e}_k \mathbf{v}_{s_k}, \quad \text{که در آن } \mathbf{e}_k = \frac{1}{\sqrt{\binom{r_k}{\text{درجه}} \binom{s_k}{\text{درجه}}}} \\ \bar{\mathbf{e}}'_i &= \rho^{e \rightarrow v} \left(E'_i \right) = \sum_{\{k | r_k = i\}} \mathbf{e}'_k \\ \mathbf{v}'_i &= \phi^v \left(\bar{\mathbf{e}}'_i \right) = \sigma \left(\bar{\mathbf{e}}'_i W \right) \end{aligned}$$

شکل ۶.۴ تطابق بین شبکه پیچشی گراف و شبکه گراف نشان داده شده در شکل ۵.۴ و ۴.۴ را نشان می‌دهد. ملاحظاتی را که هنگام تصمیم‌گیری در مورد انتخاب پیاده‌سازی واقعی توابع داخلی شبکه گرافی در نظر گرفته می‌شود، مورد بحث قرار خواهیم داد. انتخاب معماری خاص با انگیزه روابطی صورت می‌گیرد که بین عناصر داده ورودی و وظیفه‌ای که فرد می‌خواهد با مدل حل کند، وجود دارد. [۱۳]

¹Graph Convolutional Networks



شکل ۶.۴: شبکه‌ی پیچشی گراف که نوعی شبکه‌ی عصبی با انتقال پیام است، می‌تواند به عنوان یک شبکه گراف، بدون یک ویژگی سراسری و یک تابع لبه‌ای خطی و غیرخطی، بیان شود.



شکل ۷.۴: بازآرایی چشمگیرتر اجزای شبکه گرافی منجر به مدلی می‌شود که ویژگی‌های رأس (گره) را گردآوری می‌کند و آن‌ها را با یک ویژگی سراسری ترکیب می‌کند، سپس ویژگی‌های رأس را با استفاده از ویژگی ترکیبی به عنوان زمینه بروزرسانی می‌کند.

۳.۴ انواع شبکه‌های عصبی گراف

معماری‌های شبکه‌های عصبی گراف به چهار شاخه اصلی طبقه‌بندی می‌شوند: مکرر^۱، کانولوشنال^۲، مکانی-زمانی^۳ و رمزگذار خودکار گراف^۴. در RecGNNs، گره‌ها به صورت تکراری پیام‌ها را با همسایگانشان تبادل می‌کنند تا به تعادل پایداری برسند. در مقابل، GAEs^۵ گره‌ها یا کل گراف را به یک فضای نهان کد می‌کنند و به صورت بدون نظارت اطلاعات گراف را از اطلاعات کد شده بازیابی می‌کنند. در همین حال، STGNNs^۶ به دنبال یادگیری الگوهای پنهان از گراف‌های فضا-زمانی هستند. ConvGNNs^۷ از میان تنوع GNNs بیشترین استفاده را دارند، که ایده کانولوشن را از داده‌های Grid به گراف‌ها گسترش می‌دهند و با جمع‌آوری ویژگی‌های خود گره‌ها با ویژگی‌های همسایگانشان نمایندگی‌های گره‌ها را ایجاد می‌کنند. بر خلاف ConvGNNs، RecGNNs^۸ از چندین لایه کانولوشن گرافی برای استخراج نمایندگی‌های پیچیده‌تر گره‌ها استفاده می‌کنند. به طور مشخص، مدل‌های ConvGNN^۹ مانند ParticleNet نشان داده‌اند که در مسائل مربوط به فیزیک ذرات در حوزه انرژی بالا عملکرد عالی دارند. به همین دلیل، این بخش یک مرور دقیق از عملکرد ConvGNNs^{۱۰} ارائه می‌دهد. سه نوع برجسته از شبکه‌های عصبی گراف عبارتند از GCN^{۱۱}، GAT^{۱۲} و GraphSAGE^{۱۳}. هر یک از این مدل‌ها مکانیزم‌های منحصر به‌فردی برای گدازه‌سازی اطلاعات از همسایگان یک گره برای یادگیری نمایش گره‌ها دارند.

شبکه پیچشی گراف (GCN)

برای یادگیری نیمه‌نظرارت شده روی داده‌های ساختاریافته گراف طراحی شده‌اند. یادگیری را از داده‌های سنتی ساختاریافته شبکه‌ای (مانند تصاویر) به داده‌های گراف گسترش می‌دهند.

- **عملیات پیچشی:** شبکه‌های پیچشی گراف (GCNs) عملیات پیچشی را مستقیماً بر روی گراف انجام می‌دهند. این با جمع‌آوری اطلاعات ویژگی از محیط محلی یک گره (رأس) انجام می‌شود.
- **تبديل‌های گرافی طیفی:** عملکرد پیچش از طریق تقریب محلی اولین نظر درباره تبدیل‌های گرافی طیفی است. این شامل استفاده از لایپلاسین گراف برای تعریف پیچش می‌شود.
- **نرم‌السازی:** ویژگی‌های گره‌ها توسط درجه گره‌ها نرم‌السازی می‌شوند تا پایداری را تضمین کرده و مشکلات عددی را جلوگیری کنند.

¹RecGNN

²ConvGNNs

³STGNN

⁴GAEs

⁵Graph Convolutional Network

⁶Graph Attention Network

⁷GraphSAGE (SAmple and aggreGatE)

فرمول بندی

تعبیرگذاری گره $h_i^{(l+1)}$ در لایه $l+1$ به شکل زیر محاسبه می‌شود:

$$h_i^{(l+1)} = \sigma \left(\sum_{j \in \mathcal{N}(i)} \frac{1}{\sqrt{|\mathcal{N}(i)|} \sqrt{|\mathcal{N}(j)|}} W^{(l)} h_j^{(l)} \right)$$

که:

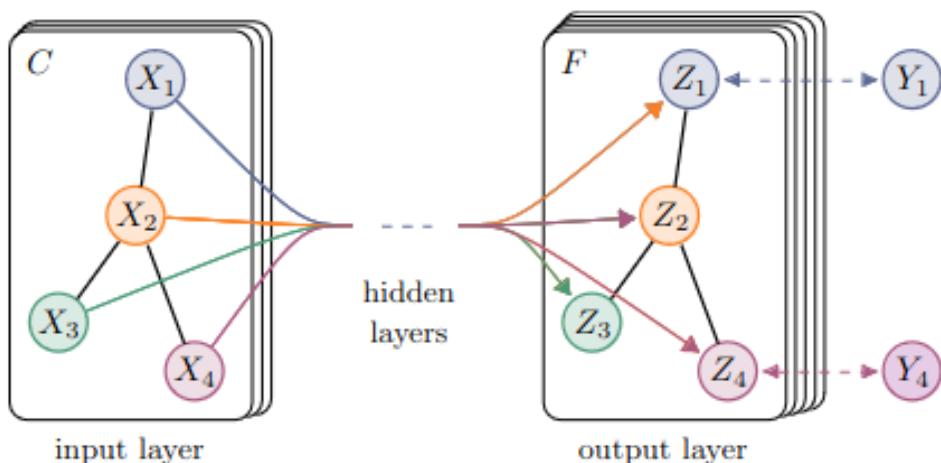
- $\mathcal{N}(i)$ مجموعه‌ای از همسایگان گره i است.

- $W^{(l)}$ ماتریس وزن برای لایه l است.

- σ یک تابع فعال‌ساز^۱ است (مانند ReLU).

شبکه‌های پیچشی گراف در وظایف مختلفی مانند دسته‌بندی گره، دسته‌بندی گراف و سیستم‌های پیشنهادی استفاده می‌شوند.

[۱۴]



شکل ۸.۴: شبکه گرافی پیچشی

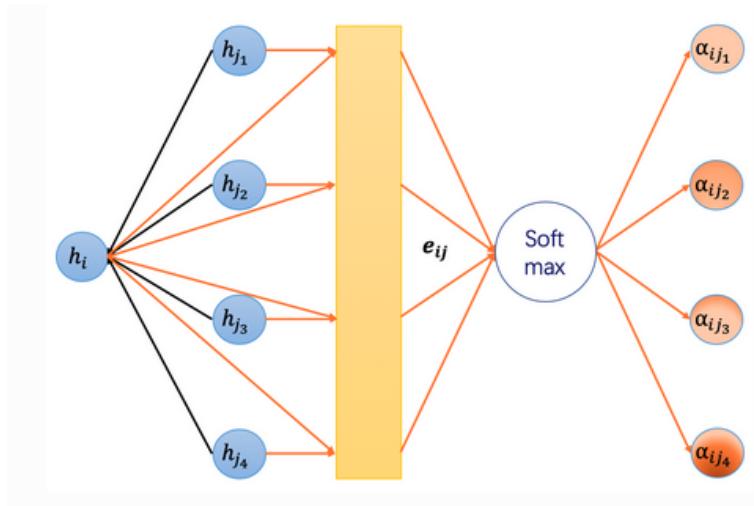
شبکه توجه گراف (GAT)

شبکه‌های توجه گرافی یک مکانیزم توجه را به شبکه‌های عصبی گرافی معرفی می‌کنند که به مدل اجازه می‌دهد تا اهمیت‌های مختلفی را به گره‌های مختلف در یک همسایگی اختصاص دهد.

- مکانیزم توجه: از لایه‌های خودتوجه ماسک شده برای محاسبه ضرایب توجه برای هر لبه در گراف استفاده می‌کنند. این امکان را به مدل می‌دهد که اهمیت گره‌های همسایه را به طریق متفاوتی می‌سنجد.

^۱Activation Function

- موازی‌سازی: محاسبه ضرایب توجه می‌تواند در تمامی لبه‌ها به صورت موازی صورت گیرد و جمع‌آوری می‌تواند در تمامی گره‌ها به صورت موازی انجام شود.
- یادگیری عمیمی: نیازی به ساختار کل گراف از پیش ندارند که آن‌ها را برای وظایف یادگیری عمیمی مناسب می‌کند.



شکل ۹.۴: شبکه گرافی توجه

فرمول‌بندی

۱. تبدیل خطی

$$z_i^{(l)} = W^{(l)} h_i^{(l)}$$

۲. امتیاز توجه:

$$e_{ij}^{(l)} = \text{LeakyReLU} \left(\vec{a}^{(l)} \cdot [z_i^{(l)} \| z_j^{(l)}] \right)$$

۳. نرمال‌سازی:

$$\alpha_{ij}^{(l)} = \frac{\exp(e_{ij}^{(l)})}{\sum_{k \in \mathcal{N}(i)} \exp(e_{ik}^{(l)})}$$

۴. جمع‌آوری:

$$h_i^{(l+1)} = \sigma \left(\sum_{j \in \mathcal{N}(i)} \alpha_{ij}^{(l)} z_j^{(l)} \right)$$

شبکه‌های توجه گرافی عملکردی برتر در وظایفی مانند دسته‌بندی گره، پیش‌بینی پیوند و دسته‌بندی گرافی نشان داده‌اند [۱۵].

شبکه گراف نمونه‌برداری و جمع‌آوری (GraphSAGE)

نمونه‌برداری و جمع‌آوری گراف^۱ برای تولید نمایش‌های گره برای داده‌هایی که قبلاً دیده نشده‌اند (یادگیری تعمیمی) با نمونه‌برداری و جمع‌آوری ویژگی‌ها از محیط محلی یک گره طراحی شده است.

- نمونه‌برداری: به جای استفاده از همه همسایگان، GraphSAGE یک مجموعه با اندازه ثابت از همسایگان برای هر گره نمونه‌برداری می‌کند.

- جمع‌آوری: از توابع جمع‌آوری مختلف (میانگین، LSTM، استخراج) برای جمع‌آوری ویژگی‌های همسایگان نمونه‌برداری شده استفاده می‌کند.

- قابلیت تعمیم: با یادگیری توابع جمع‌آوری به جای نمایش‌های خاص گره، GraphSAGE می‌تواند به گره‌هایی که دیده نشده‌اند تعمیم دهد. [۱۶]

فرمول‌بندی

تعابیر نمایش گره $h_i^{(l+1)}$ در لایه $l+1$ به شکل زیر محاسبه می‌شود:

- نمونه‌برداری:

$$\mathcal{N}_S(i) \subseteq \mathcal{N}(i)$$

- جمع‌آوری:

$$h_{\mathcal{N}_S(i)}^{(l)} = \text{AGGREGATE}^{(l)} \left(\{h_j^{(l)} : j \in \mathcal{N}_S(i)\} \right)$$

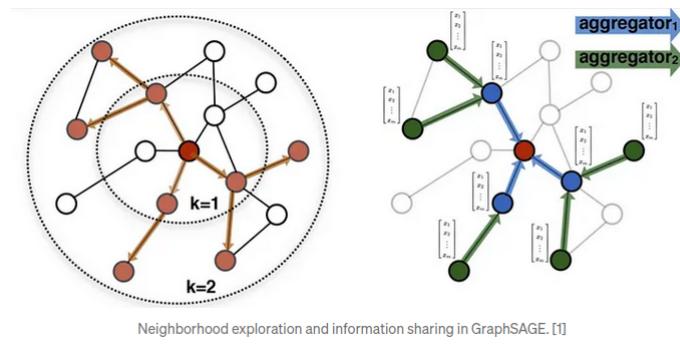
- بروزرسانی:

$$h_i^{(l+1)} = \sigma \left(W^{(l)} \cdot [h_i^{(l)} || h_{\mathcal{N}_S(i)}^{(l)}] \right)$$

- شبکه‌های کانولوشن گرافی (GCNs): در وظایفی مانند دسته‌بندی گره، دسته‌بندی گراف و سیستم‌های پیشنهادی استفاده می‌شود.

- شبکه‌های توجه گرافی (GATs): در وظایفی مانند دسته‌بندی گره، پیش‌بینی پیوند و دسته‌بندی گرافی به عنوان بهترین

^۱GraphSAGE (SAmple and aggreGatE)



شکل ۱۰.۴: شبکه گراف نمونه‌برداری و جمع‌آوری

عملکرد شناخته شده است.

- شبکه گرافی نمونه‌برداری و جمع‌آوری (GraphSAGE): برای داده‌های گراف بزرگ و برنامه‌هایی که نیازمند یادگیری تعمیمی هستند، مانند تحلیل شبکه‌های اجتماعی و سیستم‌های پیشنهادی، مفید است.

یادگیری ماشین می‌تواند در بسیاری از مراحل رویداد استفاده شود: از جمله فعال‌سازی، بازسازی و شبیه‌سازی. شبکه‌های عصبی گرافی به سه روش مختلف برای پیش‌بینی استفاده می‌شوند: در سطح گراف، گره یا لبه، به تبع وظیفه موجود. ما در فصل بعد به اختصار چالش‌ها و روش‌هایی که اعمال شده‌اند تا پایل آپ را تخمین بزنند، اشاره می‌کنیم و به جزئیات بیشتر باز می‌گردیم. تمام روش‌های ارائه شده بر روی رویدادهای شبیه‌سازی شده توسعه داده شده‌اند و هیچ عملکردی بر روی داده‌های واقعی گزارش نشده است. در هر خط کاری که در زیر توضیح داده شده است، ابتدا تصمیمی در مورد این گرفته شده است که داده‌ها چگونه می‌توانند به عنوان یک گراف بیان شوند: چه موجودات و ارتباطاتی که به ترتیب به عنوان گره‌ها و لبه‌ها نمایان می‌شوند؟ چه خروجی مورد نیاز است، یعنی پیش‌بینی‌های سطح لبه، گره یا گراف؟ از آنجا، تصمیماتی در مورد معماری خاص GNN گرفته شده است تا محاسبات مورد نظر را منعکس کنند: آیا شبکه خروجی جهانی برای تولید خروجی‌های سطح گراف لازم است؟ آیا تعاملات جفتی بین گره‌ها باید محاسبه شود، یا بیشتر مانند جمع‌آوری و تبدیل غیرخطی؟ GCN چند مرحله گذار پیام باید استفاده شود، به منظور انتقال اطلاعات بین گره‌های دور از هم در گراف؟

فصل پنجم

تخمین توزیع پایل آپ با استفاده از شبکه های عصبی گراف

این فصل به بررسی جامع آزمایش های انجام شده و نتایج آنها پرداخته، محدودیت های شناسایی شده در رویکرد آزمایشی را مورد بحث قرار می دهد و راه های احتمالی برای تحقیقات آینده را بر جسته می کند. در اینجا، عملکرد روش های بررسی شده برای تخمین پایل آپ گزارش شده و پیاده سازی یک گراف از رویدادهای برخورد ذرات بهره برداری می کند، مورد بحث قرار می گیرد. الگوریتم های مبتنی بر یادگیری ماشین به نظر می رسد برای هندسه نامنظم CMS مناسب تر باشند. به عنوان نتیجه ای از ملاحظات دقیق بر اساس مطالعات منتشر شده قبلی، این کار تمرکز اصلی خود را بر پیش بینی پایل آپ با استفاده از شبکه های عصبی گراف در گراف های مصنوعی ساخته شده از رویدادهای برخورد قرار می دهد. این فصل مفاهیم شبکه عصبی گراف مرتبط با این کار را مرور می کند، معماری پیشنهادی را معرفی می کند و توضیحاتی درباره انتخاب های طراحی آن ارائه می دهد. در نهایت، راهبردهایی برای ارزیابی عملکرد راه حل های پیشنهادی تخمین پایل آپ توضیح داده می شود.

۱.۵ مرور داده ها و فرآیندهای آزمایشی

داده های مدل از داده های MiniAOD تولید شده اند، از اطلاعات ذرات برای ساخت نمایش گرافی استفاده کردیم که در فصل ۳ مورد بحث قرار گرفت. سپس، آزمایش به سمت سناریوهای پایه گذاری مدل پیش می رود. رویکرد همسایگی های گره با استفاده از مکانیزم های پیام رسانی که در فصول ۴ و ۵ توضیح داده شده، گسترش می یابد. برای شناسایی مدل های با بهترین عملکرد برای

هر مجموعه داده، از معیارهای L1Loss و فیزیکی استفاده می‌شود.

۲.۵ امکانات محاسباتی

پردازش مسیر بازسازی ذرات برای کل آشکارساز CMS در چارچوب CMSSW پیاده‌سازی شده است که عمدتاً به زبان C++ نوشته شده است. سیستم هایی برای استخراج ویژگی‌ها، پیش‌پردازش ورودی، تفسیر نتایج مدل و پس‌پردازش نیز پیاده‌سازی شوند. برای ساده‌سازی فرآیند کار، تمامی آزمایش‌ها در پایتون انجام می‌شوند. برای انجام تحلیل داده‌ها و آموزش شبکه به صورت تعاملی، از سرویس ابری (Service for Web-based ANalysis) SWAN در CERN استفاده می‌کنم. SWAN مورد استفاده مجهر به پردازنده Intel(R) Xeon(R) Silver 4216 CPU @ 2.10GHz با ۱۶ گیگابایت RAM و کارت گرافیک NVIDIA Tesla T4 با ۳۲ گیگابایت حافظه است. آزمایش‌ها با استفاده از مجموعه‌ای از Jupyter Notebookها در SWAN و کدهای همراه پایتون انجام می‌شوند. داده‌های مورد استفاده در آزمایش‌ها در CERNBox ذخیره شده‌اند و به صورت مستقیم از محیط SWAN قابل دسترسی هستند. در این کار از چندین کتابخانه پایتون برای دستکاری داده‌های فیزیک، توسعه مدل‌ها و پیاده‌سازی معیارهای ارزیابی استفاده می‌شود. داده‌های رویداد برخوردها با استفاده از ماژول NTuplizer در ROOT به فرمت CMSSW مبتنی بر آرایه سریال‌سازی می‌شوند که با کمک کتابخانه Uproot در پایتون خوانده می‌شوند. برای مدیریت آرایه‌های تو در توی با اندازه متغیر به شیوه tensor، از کتابخانه Torch استفاده می‌شود. علاوه بر این، از مجموعه‌ای از ابزارها برای تحلیل گراف، مصورسازی و معیارهای ارزیابی بهره می‌برم که شامل کتابخانه networkx برای تحلیل گراف، matplotlib و plotly برای مصورسازی و scikit-learn و scipy برای معیارهای ارزیابی از پیش آماده می‌شود. در زمینه یادگیری ماشین، مدل‌های MLP و GNN را با استفاده از PyTorch پیاده‌سازی کرده و شبکه‌های آموزش‌دیده را برای استنتاج به فرمت pt ذخیره می‌کنم. به طور خاص، این کار از کتابخانه PyTorch Geometric استفاده می‌کند.

ذخیره اطلاعات رویداد با استفاده از کتابخانه روت اطلاعات ذخیره شده رویدادهای برخورد ذرات با استفاده از کتابخانه ROOT نشان داده می‌شود. ویژگی‌های مختلف ذرات و ورتكس‌ها به صورت نمودارها و توضیحات مرتبط نمایش داده می‌شوند. اطلاعات ذخیره شده شامل:

- خصوصیات ذرات
- اطلاعات ورتكس‌ها
- تعداد ذرات

پیش‌پردازش مراحل پیش‌پردازش داده‌های رویدادها قبل از اعمال مدل‌های یادگیری ماشین شرح داده می‌شوند. این شامل آماده‌سازی داده‌ها برای ورود به مدل‌ها است. فایل‌های h5 را در دایرکتوری مشخصی پردازش می‌کنیم، وجود فایل‌های مرتبط

.json را بررسی کرده و متادیتای آن‌ها را خلاصه می‌کنیم. تابعی بررسی می‌کند که آیا یک فایل دارای پسوند خاصی است یا خیر. در حالی که تابعی دیگر فایل‌های h5 را می‌خواند، متادیتای مربوط به فایل‌های json آن‌ها را استخراج کرده و این داده‌ها را به یک دیکشنری خلاصه‌سازی می‌کند و آن را در فایلی ذخیره می‌کند. خطاهایی که در این فرآیند رخ می‌دهند ثبت می‌شوند. تابع یک ایندکس جهانی را بر اساس داده‌های فراوانی کلاس‌ها در یک دیکشنری معین به یک کلاس خاص و ایندکس محلی نگاشت می‌کند. بهبودهای پیشنهادی برای این مرحله شامل بهبود مدیریت خطا، بهینه‌سازی پردازش فایل‌ها برای کارایی بیشتر، بهبود لگ‌نویسی با اطلاعات جزئی‌تر و افزایش خوانایی کد با مستندسازی بهتر و نام‌گذاری متغیرهای مناسب‌تر است.

پردازش پردازش و تبدیل داده‌های مرتبط با گراف به فرمتهای سازگار با کتابخانه PyTorch Geometric. ابتدا، فایل‌های داده (شامل فرمتهای pt و h5) با استفاده از مکانیزم‌های بازخوانی بارگذاری می‌شوند و در صورت عدم دسترسی به فایل‌ها، خطاهای ثبت می‌گردند. در ادامه، ساختارهای گراف ناهمگن از طریق ایجاد ماتریس‌های مجاورت و افزودن خود-حلقه‌ها ساخته می‌شوند. ویژگی‌های مختلف گره‌ها و لبه‌ها را مدیریت کرده و از صحت قالب‌بندی و وزن‌دهی آنها اطمینان حاصل می‌کند. داده‌های گراف با تبدیل ویژگی‌های گره‌ها، ویژگی‌های یال‌ها و شاخص‌ها به شیء Data در PyTorch Geometric آماده می‌شوند که این عملیات شامل افزودن خود-حلقه‌ها و تغییرات ویژگی‌ها نیز می‌باشد.

عملیات پردازش شامل گسترش ویژگی‌ها از طریق پردازش انواع خاصی از گره‌ها و محاسبه اتصالات مبتنی بر تکانه می‌پردازد. همچنین، برای اطمینان از یکنواختی داده‌ها، وجود مقادیر NaN یا inf در تنسورها بررسی می‌شود. در صورت وجود مقادیر نامعتبر، گره‌های دارای NaN یا inf حذف و لیست‌های یال‌ها تنظیم می‌گردد. علاوه بر این، لبه‌هایی که ویژگی‌های آنها بالاتر از یک آستانه مشخص باشد، فیلتر می‌شوند. این رویکرد جامع و سیستماتیک، امکان مدیریت قوی و انعطاف‌پذیر داده‌های گراف را فراهم آورده و آنها را برای پردازش‌های بیشتر یا آموزش مدل‌های یادگیری ماشین آماده می‌کند. فرایند

فرایند آموزش و ارزیابی مدل‌ها از خواندن داده‌ها^۱، مدیریت دیتاست‌ها^۲، بارگذاری داده‌ها^۳، آموزش مدل‌ها^۴، و تحلیل نتایج^۵ تشکیل شده است. از ویژگی‌های کلیدی می‌توان به استفاده از PyTorch Geometric برای مدیریت گراف‌ها، ابزارهای پردازش و پیش‌پردازش داده، و همچنین قابلیت ایجاد و ذخیره نمودارهای پیچیده با استفاده از کتابخانه Plotly اشاره کرد. نمودارهای آموزش و اعتبارسنجی برای مدل‌های مختلف شامل PUGAT و PUSage ارائه می‌شوند.

در مدل یادگیری ماشین، پارامترها و هایپرپارامترها موجود است:

¹DataReader

²BoostedDataset

³DataLoader

⁴Trainer

⁵Analyzer

پارامترها

- وزن و بایاس در لایه : این پارامترها در فرآیند آموزش شبکه عصبی گرافی یادگرفته می‌شوند.

ابرپارامترها

- تعداد دوره‌های آموزشی

RAdam

- الگوریتم بهینه‌سازی

• نرخ یادگیری برای بهینه‌ساز

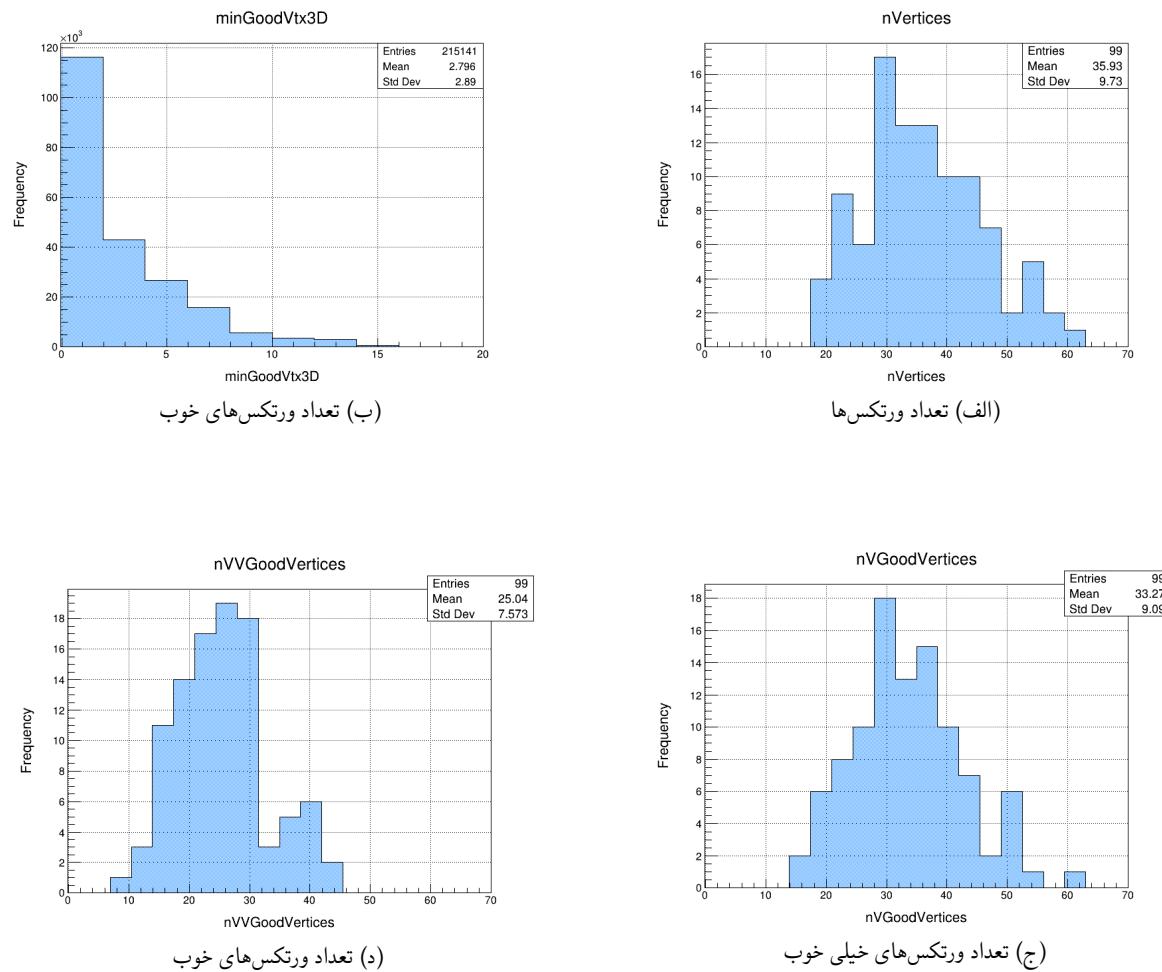
- تابع هزینه

داده‌ها

مجموعه داده‌های استفاده شده به صورت زیر تعریف شده است: به ازای پایل آپ‌های ۲۰ تا ۷۰۰ گراف به عنوان ورودی به مدل داده شد که ۲۰٪ از آنها برای آزمون و ۱۰٪ برای اعتبار سنجی و باقی برای آموزش استفاده شدند.

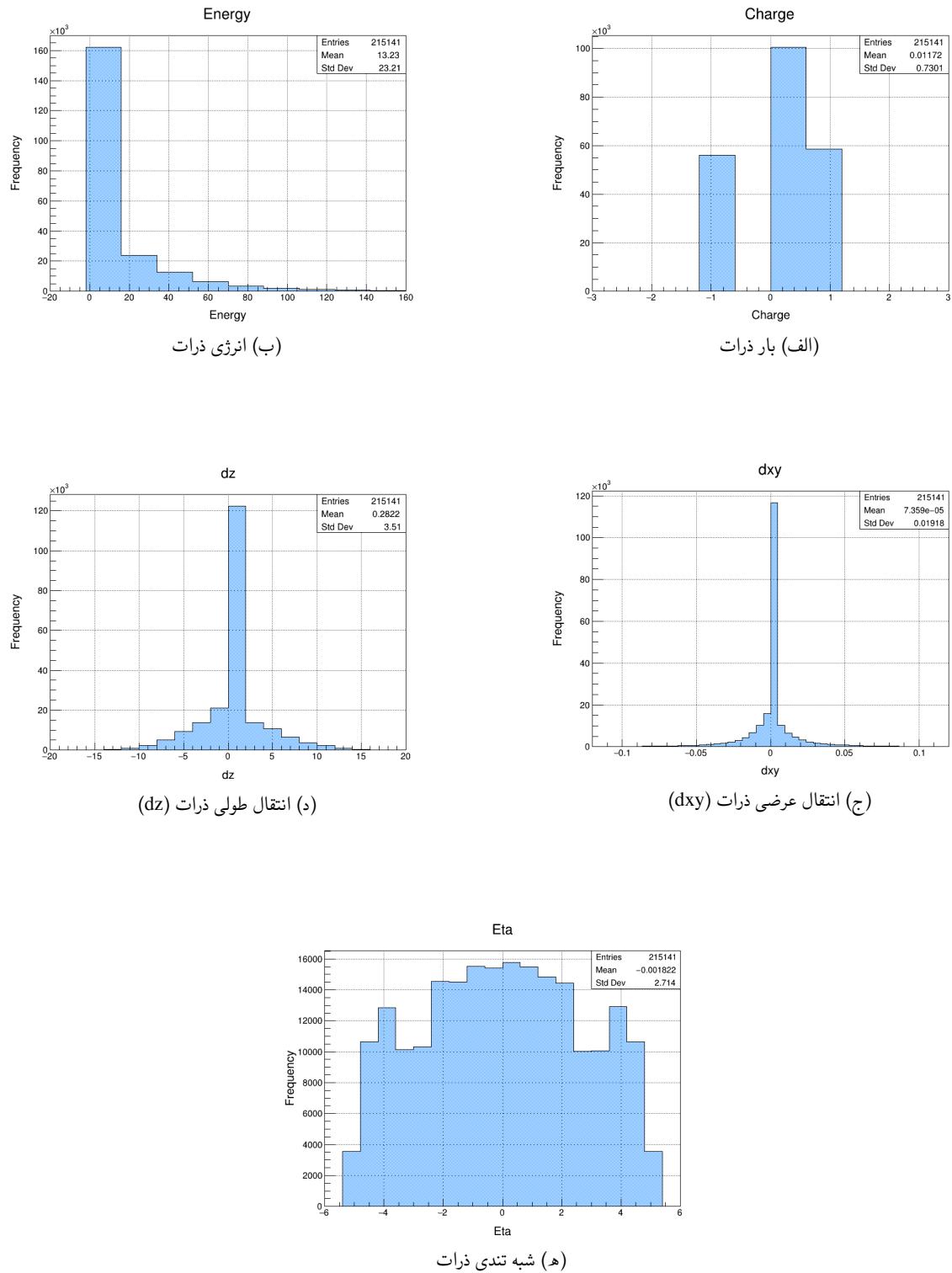
۳.۵ در وهله اول اطلاعات رویداد ها را می خوانیم

اطلاعات رأس ها

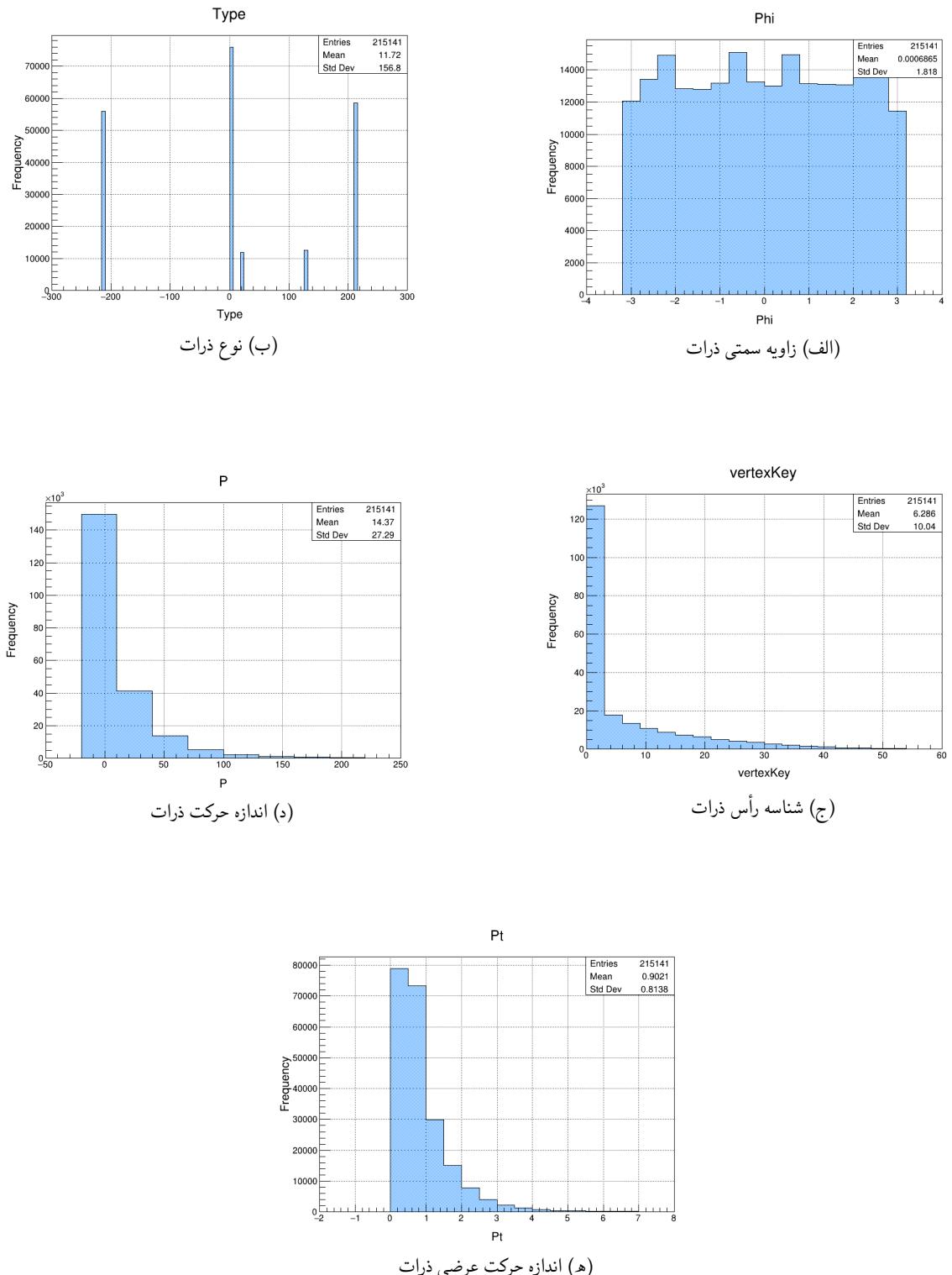


شکل ۱.۵: اطلاعات ورتكس ها از جمله تعداد ورتكس ها و انواع مختلف آنها

اطلاعات کاندیدای ذرات

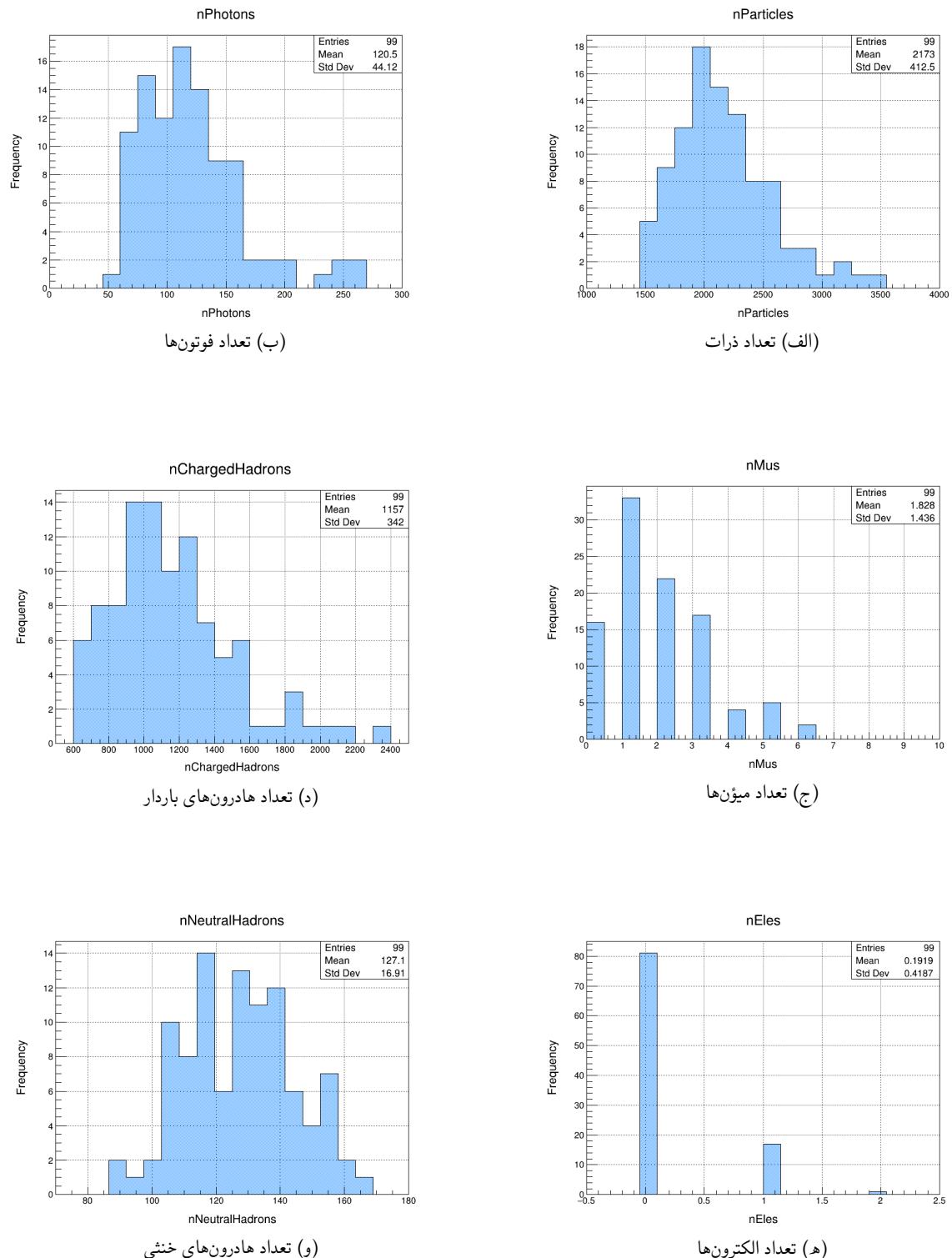


شکل ۲.۵: ویژگی‌های مختلف ذرات و نمودارهای مرتبط با توزیع هر ویژگی



شکل ۳.۵: ویژگی‌های مختلف ذرات و نمودارهای مرتبط با توزیع هر ویژگی

اطلاعات تعداد ذرات



شکل ۴.۵: توزیع تعداد ذرات کاندیدا شامل فوتون‌ها، میئن‌ها، الکترون‌ها، و هادرون‌های باردار و بدون بار در رویدادهای

۴.۵ ساخت نمایش رویداد برخورد ذرات بصورت گراف

داده‌های خام آشکارساز در اصل مجموعه‌ای نامرتب از N شیء هستند. با این حال، با در نظر گرفتن Ne پیوند هندسی یا فیزیکی بین اشیاء، می‌توان مجموعه را به یک گراف تکمیل کرد. در اصطلاح مسئله پیوند، گره‌ها در گراف نمایانگر ذرات هستند و یال‌ها رابطه دو به دوی آن‌ها را نشان می‌دهند. این نمایش گراف توسط $G = (V, E)$ ارائه می‌شود، که در آن V مجموعه‌ای از N گره است و $E \subseteq V \times V$ مجموعه‌ای از Ne یال‌ها است. هر گره در این نقطه $V = \{x_1, \dots, x_N\} \subseteq \mathbb{R}^F$ شامل خصوصیت گره است، و هر یال می‌تواند دارای ویژگی‌های F' خاص خود باشد.

رابطه بین گره‌ها ممکن است به عنوان جهت‌دار یا بدون جهت دسته‌بندی شود. این می‌تواند توسط یک ماتریس مجاورت نمایانگر شود که یک ماتریس دودویی $N \times N$ (معمولًاً پراکنده) است و عناصر آن نشان می‌دهند که آیا یک گره به یک گره دیگر متصل است یا خیر. یک شاخص یال در فرمت فهرست مختصات، که به عنوان یک ماتریس $Ne \times 2$ تعریف می‌شود و هر ستون شامل شاخص‌های گره‌های هر یال است، جایگزین کدگذاری ماتریس مجاورت گراف است. اگرچه این نمایش فشرده برای تولید ماتریس افزایشی و نیازهای حافظه کوچکتر سودمند است، تبدیل به فرمت فشرده یا متراتکم اغلب برای عملیات حسابی مورد نیاز است.

رویکردهای ساخت گراف شبکه عصبی گراف در مرحله پیش پردازش داده‌ها ویژگی‌های مختلف ذرات مانند شبه سریع^۱، زاویه سمتی^۲، مؤلفه‌های تکانه^۳، اطلاعات رأس^۴، نوع ذره^۵، انرژی^۶، بار ذره^۷ و فواصل فضایی^۸ را استخراج کردیم. علاوه بر این، ویژگی‌های سطح رویداد مانند تعداد رئوس^۹ بازیابی شدند.

ویژگی‌های گره: ویژگی‌های گره^{۱۰} را با انباسته کردن چندین تنسور یک بعدی حاوی خواص ذرات در یک تنسور دو بعدی واحد تعریف می‌کند. این تنسور اساساً نقطه داده هر ذره را در گراف نشان می‌دهد.

ویژگی‌های مختلف ذرات از داده‌ها استخراج شدند برای جزئیات بیشتر در مورد کد و روش استفاده شده، به پیوست ۲۰ مراجعه کنید.

تعریف یال: تابع ای که برای تعریف یال‌ها نوشته‌یم^{۱۱} نقش اساسی در تعریف یال‌ها در گراف ایفا می‌کند. فاصله بین هر

¹eta

²phi

³pt, p

⁴vertexKey

⁵Type

⁶Energy

⁷Charge

⁸dz, dxy

⁹nVertices

¹⁰node_features

¹¹GetEdgeInfo

جفت ذره را بر اساس مقادیر eta و phi آنها محاسبه می‌کند. از یک آستانه فاصله^۱ برای تعیین اینکه آیا ذرات متصل در نظر گرفته می‌شوند یا خیر (یعنی یال‌ای بین آنها وجود دارد) استفاده می‌شود. جزئیات کامل این تابع و نحوه عملکرد آن در پیوست آورده شده است (نگاه کنید به پیوست ۱). این تابع دو تنسور را برمی‌گرداند: edge_index که حاوی شاخص‌های ذرات متصل است و edge_attr که حاوی فاصله بین آن ذرات متصل است. از شیء tg.data.Data از کتابخانه torch_geometric است و برای ساخت گراف استفاده شد. این شیء تمام مؤلفه‌های ضروری را در بر می‌گیرد:

- x: تنسور ویژگی‌های گره

- edge_index: تنسور شاخص‌های یال

- edge_attr: تنسور ویژگی‌های یال

- graph_attr: تنسوری حاوی ویژگی‌های سطح رویداد مانند تعداد رئوس

کد ما یک مقدار "nInt" را از داده رویداد بازیابی می‌کند که به عنوان برچسب گراف عمل می‌کند. این مربوط به تعداد پایل‌آپ در رویداد برخورد خاص است. داده‌های پردازش شده برای هر رویداد (ویژگی‌های گره، اطلاعات یال، ویژگی‌های گراف و برچسب) با استفاده از کتابخانه h5py در یک فایل HDF5 ذخیره می‌شود. داده‌ها برای بازیابی کارآمد بر اساس نوع تعامل (پایل‌آپ) در فایل HDF5 سازماندهی می‌شوند.

در اصل، ما گراف‌هایی را ایجاد می‌کنیم که در آن:

- گره‌ها نماینده ذرات منفرد هستند.

- یال‌ها ذرات را در یک آستانه فاصله فضایی خاص به هم متصل می‌کنند.

- ویژگی‌های گره خواص مختلف هر ذره را ضبط می‌کنند.

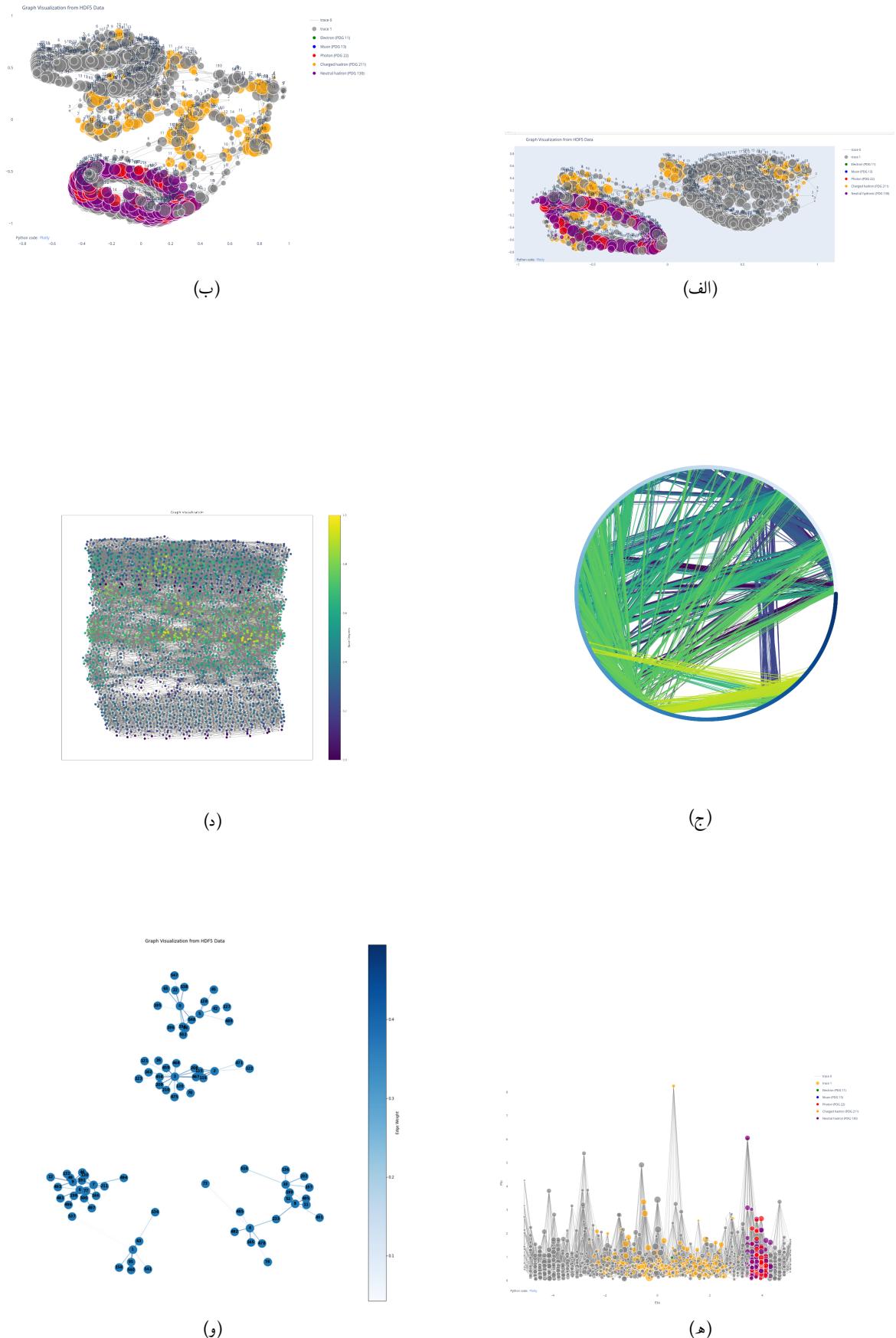
- ویژگی‌های یال نشان‌دهنده فاصله بین ذرات متصل است.

- برچسب‌های گراف به طور بالقوه تعداد پایل‌آپ را در رویداد نشان می‌دهد.

این نمایش گراف به ما امکان می‌دهد از تکنیک‌های قدرتمند شبکه‌های عصبی گراف برای تحلیل بیشتر تعاملات پیچیده ذرات مشاهده شده در داده‌های LHC استفاده کنیم.

شکل ۵.۵ نمایشی از گرافی است که تولید کرده ایم:

^۱max_distance



شکل ۵.۵: نمایش هایی از گراف تولید شده از داده‌ها. جایی که گره‌ها نماینده ذرات منفرد هستند و یا ها نشان‌دهنده ارتباطات بین آنها می‌باشند.

۱۰.۴.۵ ارسال پیام

ما از دو مکانیسم خاص ارسال پیام، در این کار استفاده کردیم، انتقال پیام از طریق لایه‌های GATv2Conv و SAGEConv می‌شود، هرکدام از این لایه‌ها از مکانیسم‌های متفاوتی برای بهروزرسانی نمایش‌های گره در داده‌های ساختارمند گراف استفاده می‌کنند. لایه GATv2Conv یا شبکه توجه گراف نسخه ۲، انتقال پیام را با یکپارچگی مکانیزم‌های توجه ارتقا می‌دهد. این امکان را به گره‌ها می‌دهد که بر اساس ضرایب توجه یادگرفته شده، اهمیت همسایگان خود را انتخابی کنند و در حین جمع‌آوری، به اطلاعات مربوطتر تمرکز کنند. به عبارت دیگر، GATv2Conv به گره‌ها این امکان را می‌دهد که به صورت انتخابی وزن همسایگان خود را مشخص کنند. از سوی دیگر، SAGEConv از یک رویکرد ساده‌تر استفاده می‌کند که هر گره اطلاعات را به صورت یکنواخت از همسایگان خود جمع‌آوری می‌کند، معمولاً با استفاده از استراتژی‌های جمع‌آوری مانند میانگین یا جمع که توسط edge_index تعریف شده‌اند. این مکانیسم اطمینان می‌دهد که گره‌ها اطلاعات محلی را به طور موثری در بر می‌گیرند در حالی که کارایی محاسباتی را حفظ می‌کنند.

با هم، این مکانیسم‌های انتقال پیام، به مدل‌های ما این امکان را می‌دهند که روابط پیچیده درون گراف را گرفتار کنند، که برای وظایفی مانند دسته‌بندی گره یا پیش‌بینی سطح گراف که در تخمین پایل آپ و اهمیت گره‌ها اساسی است، مناسب هستند. شبکه‌های عصی گراف با ارسال پیام‌ها در سراسر یک گراف عمل می‌کنند که، از دو مرحله اصلی تشکیل شده است: جمع‌آوری و بهروزرسانی. ابتدا، تابع بهروزرسانی یال مقادیر یال بهروزرسانی شده، که به عنوان پیام‌ها نیز شناخته می‌شوند، بر اساس یال

$$\mathbf{e}_{ij}^{t-1} \text{ و خصوصیات گره } \mathbf{h}_i^{t-1} \text{ در سطح انتقال پیام } (1-t), \text{ با مجموعه ویژگی‌های قابل یادگیری } \Theta \text{ محاسبه می‌شود:}$$

$$\mathbf{e}_{ij}^t = \phi_{t,\Theta}^e \left(\mathbf{e}_{ij}^{t-1}, \mathbf{h}_i^{t-1}, \mathbf{h}_j^{t-1} \right).$$

سپس، پیام‌های محاسبه شده به صورت گروهی برای هر گره جمع‌آوری می‌شوند با در نظر گرفتن همه یال‌ها یا زیرمجموعه‌ای از یال‌ها که به این گره i متصل هستند که توسط $N(i) = j_1, \dots, j_{n_i}$ نشان داده می‌شود:

$$\bar{\mathbf{e}}_i^t = \rho_{t,j:N(i)} \left(\mathbf{e}_{ij_1}^t, \dots, \mathbf{e}_{ij_{n_i}}^t, \mathbf{h}_i^{t-1} \right).$$

در نهایت، تابع بهروزرسانی گره

$$\mathbf{h}_i^t = \phi_{t,\Theta'}^v \left(\bar{\mathbf{e}}_i^t, \mathbf{h}_i^{t-1} \right)$$

استفاده می‌شود تا ویژگی‌های گره تجدید شده $\mathbf{h}_i^t \in \mathbb{R}^{F_t}$ را محاسبه کند. توابع خاص بهروزرسانی و جمع‌آوری به معماری

GNN بستگی دارد. توابع به روزرسانی $\phi_{t,\Theta}^e$ و $\phi_{t,\Theta'}^v$ به عنوان شبکه‌های عصبی قابل آموزش پیاده‌سازی می‌شوند، در حالی که توابع جمع‌آوری $\rho_{t,j:N(i)}$ ثابت‌به-جای-ترتیب هستند و تعداد متغیری از آرگومان‌ها را می‌پذیرند (به عنوان مثال، توابعی مانند میانگین، حداقل، حداقل، جمع، و غیره).

۵.۰.۵ رویکردهای کاوش شده یادگیری ماشین

در آزمایش‌های مربوط به مدل‌های GNN، یک رویکرد تدریجی اتخاذ شد که با معماری‌های ساده‌تر شروع و به تدریج پیچیدگی آن‌ها افزایش یافت. به طور مشخص، مجموعه اول آزمایش‌ها شامل مقادیر آستانه یال^۱ برابر با ۰.۱ و ۰.۲ است، با و بدون استفاده از شناسه رأس ذرات^۲ در ویژگی‌های گره. سپس از شبکه عصبی گراف استفاده شده و در نهایت با افزودن مکانیزم‌های توجه و نمونه‌برداری و جمع‌آوری تکمیل می‌شود. جزئیات این آزمایش‌ها در بخش‌های بعدی توصیف شده‌اند. هر دو شبکه می‌توانند بر روی کل گراف رویداد اعمال شوند.

۱۰.۵.۵ چارچوب مسئله وایازش و دسته‌بندی گراف

چارچوب مسئله وایازش و دسته‌بندی گراف رویداد را به عنوان یک وظیفه وایازش و دسته‌بندی سطح گراف تحت نظرارت با استفاده از شبکه‌های عصبی در گراف‌های ساخته‌شده از رویدادهای برخورده تعریف می‌کند. به طور مشخص، روش پیشنهادی شامل آموزش یک مدل وایازش و یک مدل دسته‌بندی برای پیش‌بینی برچسب یک گراف بر اساس تعییه‌های مربوطه‌اش است، با استفاده از مکانیزم‌های پیشرفت‌هه انتقال پیام مدل‌های PUSage و PUGAT در حالی که استفاده از پیش‌بینی برچسب گراف برای وایازش و دسته‌بندی ممکن است در ابتدا ضد و نقیض به نظر بررسد، روش‌های جمع‌آوری همسایگی که توسط شبکه‌های عصبی گراف به کار گرفته می‌شود، این فرایند را به طور موثری تسهیل می‌کند.

در مدل PUGAT، لایه‌های GATv2Conv از مکانیزم‌های توجه برای وزن‌دهی به اهمیت گره‌های همسایه استفاده می‌کنند، بنابراین در هنگام جمع‌آوری، بر اطلاعات مربوطتر تمرکز می‌کنند. از سوی دیگر، PUSage از لایه‌های SAGEConv استفاده می‌کند که اطلاعات را از همسایگان با استفاده از استراتژی‌های میانگین یا مجموع بر اساس اتصالات آن‌ها جمع‌آوری می‌کند.

۲۰.۵.۵ وظیفه وایازش و دسته‌بندی با MLP

در وظایف وایازش، هدف پیش‌بینی یک خروجی مقدار پیوسته بر اساس ویژگی‌های ورودی است. یک MLP برای وایازش معمولاً شامل چندین لایه از نورون‌ها است که هر نورون در یک لایه به تمام نورون‌های لایه بعدی وصل است. به این صورت یک MLP برای وایازش ساختاردهی می‌شود:

¹edges_threshold

²vertexKey

لایه ورودی: نورون‌هایی که متناظر با ویژگی‌های ورودی هستند. لایه‌های مخفی: لایه‌های میانی از نورون‌ها با تابع فعال‌سازی (مانند ReLU). لایه خروجی: یک نورون بدون تابع فعال‌سازی (برای واپاش) یا با تابع فعال‌سازی خطی (برای پیش‌بینی مقادیر پیوسته). فرآیند آموزش شامل تنظیم وزن‌ها و بایاس‌ها با استفاده از تکنیک‌های بهینه‌سازی مانند نزول گرادیان برای کمینه کردن اختلاف بین مقادیر پیش‌بینی شده و واقعی (مانند خطای میانگین مربعات) است.

در وظایف دسته‌بندی، هدف اختصاص ورودی‌ها به دسته‌های پیش‌ تعیین شده است. یک MLP برای دسته‌بندی مشابه واپاش است اما به طور معمول دارای لایه خروجی متفاوت و از تابع هزینه متفاوت است:

لایه ورودی: نورون‌هایی که متناظر با ویژگی‌های ورودی هستند. لایه‌های مخفی: لایه‌های میانی با تابع فعال‌سازی (مانند ReLU). لایه خروجی: نورون‌هایی که متناظر با تعداد کلاس‌ها هستند، با تابع فعال‌سازی softmax (برای دسته‌بندی چندکلاسه) یا تابع فعال‌سازی sigmoid (برای دسته‌بندی دودویی). در طول آموزش، MLP یاد می‌گیرد که احتمالات خروجی را برای هر کلاس تولید کند و تابع هزینه برای بهروزرسانی وزن‌ها و بایاس‌ها استفاده می‌شود.

پایه‌گذاری معماری مدل

PUGAT با توجه به شکل ۶.۵ می‌توان این مدل را به این صورت توضیح داد، داده ورودی: ترکیبی از ویژگی‌های گره‌ها^۱، اطلاعات مجاورت^۲، ویژگی‌های گراف^۳ و اطلاعات دسته‌بندی^۴. لایه‌ها^۵: لایه یک مکانیزم توجه گرافی را به داده‌های ورودی اعمال می‌کند تا روابط بین گره‌ها را درک کند. نرمال‌سازی لایه: خروجی لایه قبلی را نرمال می‌کند تا فرآیند آموزش را تثبیت و تسريع کند. شبکه پرسپترون چندلایه: ویژگی‌های گره‌های بهصورت جهانی جمع می‌کند و ویژگی‌های گراف را به عنوان ورودی دریافت می‌کند. خروجی با ابعاد $2 * \text{hidden_channels}$ تولید می‌کند. لایه: یک لایه توجه گرافی دیگر برای پردازش بیشتر ویژگی‌های گره‌ها. نرمال‌سازی لایه: خروجی لایه قبلی را نرمال می‌کند. شبکه پرسپترون چندلایه: ویژگی‌های گره‌های بهصورت جهانی تجمع یافته، ویژگی‌های اضافی و یک متغیر g را به عنوان ورودی دریافت می‌کند. خروجی با ابعاد $2 // \text{hidden_channels}$ تولید می‌کند. نتیجه: خروجی نهایی مدل.

PUSage با توجه به شکل ۷.۵ داده ورودی: ترکیبی از ویژگی‌های گره‌ها^۶، x شاخص یال‌هاء^۷، ویژگی‌های یال‌ها^۸، برچسب‌ها

و اطلاعات دسته‌بندی^۹. لایه **SAGEConv1**: یک کانولوشن GraphSAGE اعمال می‌کند تا اطلاعات را از گره‌های

¹ x

²adj

³features

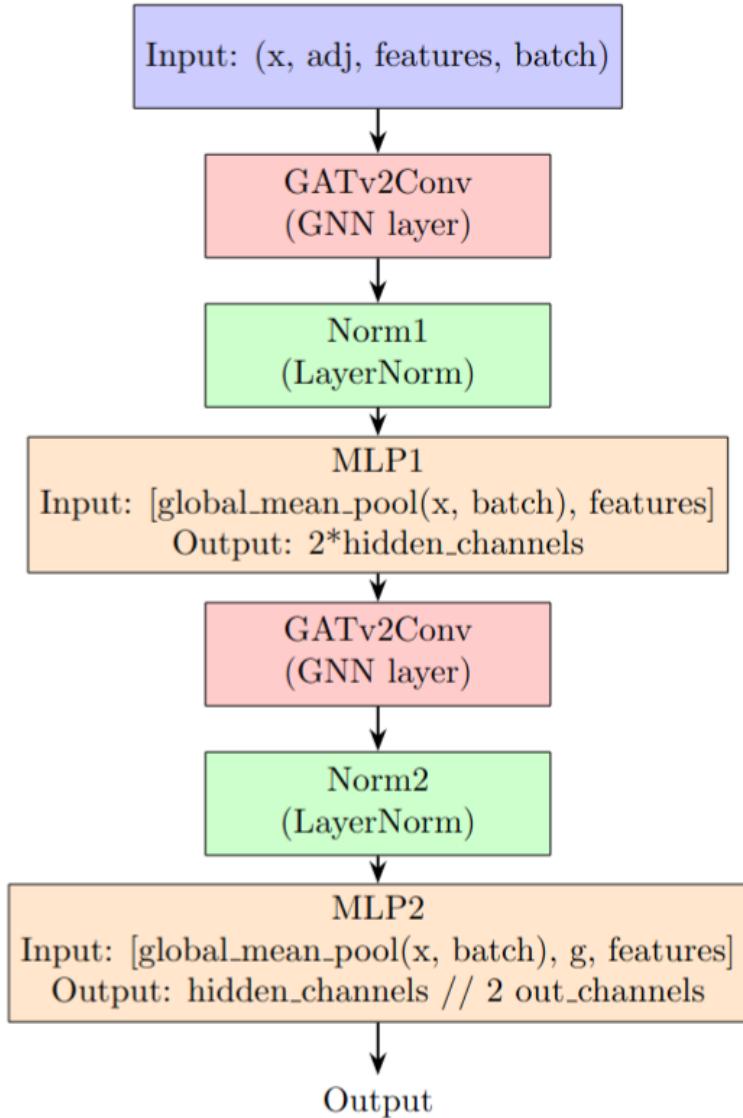
⁴batch

⁵GATv2Conv1 (GNN layer)

⁶edge_index

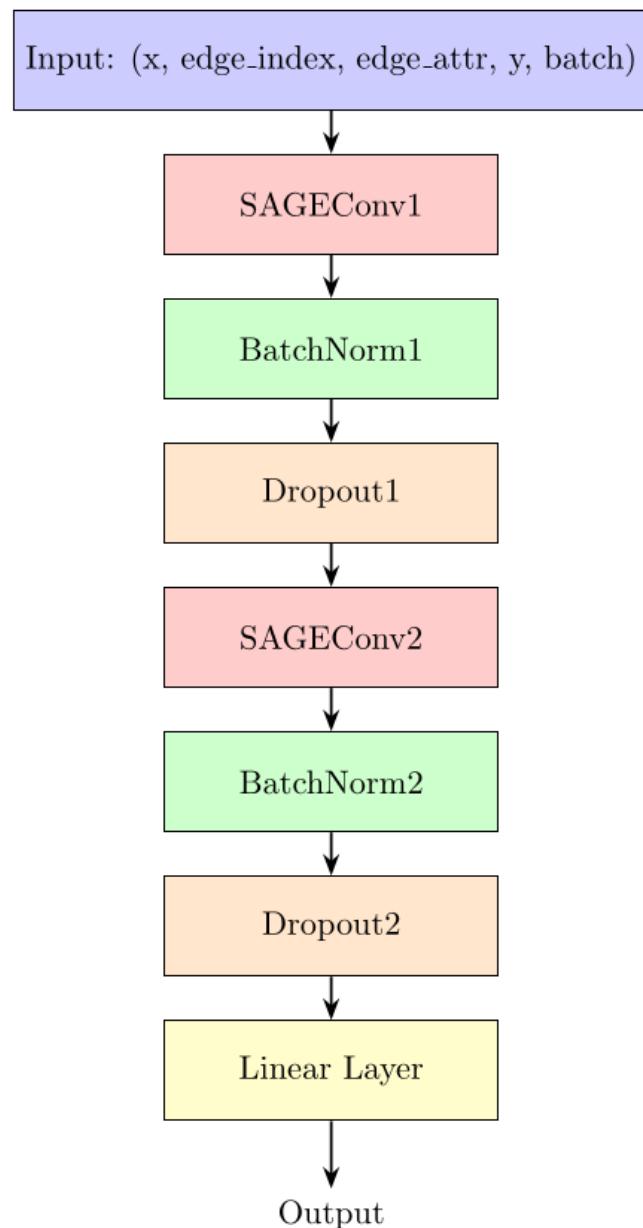
⁷edge_attr

⁸batch



شکل ۶.۵: معماری مدل PUGAT

همسایه جمع‌آوری کند. **BatchNorm1**: نرمال‌سازی دسته‌ای را برای ثبیت فرآیند آموزش با نرمال کردن خروجی لایه قبلی اعمال می‌کند. **Dropout1**: به صورت تصادفی یک بخش از واحدهای ورودی را به صفر می‌رساند تا از بیش‌برازش جلوگیری کند. لایه **SAGEConv2**: یک لایه کانولوشن GraphSAGE دیگر برای پردازش بیشتر ویژگی‌های گره‌ها. **BatchNorm2**: نرمال‌سازی دسته‌ای را برای خروجی لایه قبلی اعمال می‌کند. **Dropout2**: یک لایه درآپاوت دیگر برای جلوگیری از بیش‌برازش. لایه **خطی**: یک لایه کاملاً متصل که ورودی را به ابعاد خروجی نهایی تبدیل می‌کند. نتیجه: خروجی نهایی مدل.



شکل ۷.۵: معماری مدل PUSage

۶.۵ تابع هزینه

تابع هزینه یک مؤلفه کلیدی در فرآیند آموزش مدل است. اندازهگیری اختلاف بین خروجی‌های پیش‌بینی شده مدل و مقادیر هدف واقعی را انجام می‌دهد. این تابع اندازهگیری کمی از اینکه چقدر پیش‌بینی‌های مدل با نتایج مورد انتظار همخوانی دارند. در طول آموزش، الگوریتم بهینه‌سازی (مانند نزول گرادیان) از تابع خسارت برای تنظیم پارامترهای مدل استفاده می‌کند. هدف از این عمل این است که خطای را کمینه کند و از این طریق عملکرد مدل را بهبود بخشد. الگوریتم بهینه‌سازی، گرادیان تابع هزینه نسبت به پارامترهای مدل را محاسبه کرده و پارامترها را در جهت کاهش خسارت بهروزرسانی می‌کند. تابع هزینه منفی لگاریتم احتمال پواسون یا PoissonNLLLoss از کتابخانه PyTorch برای یک پیش‌بینی واحد λ (ورودی) و یک تعداد مشاهده شده y (هدف) به صورت زیر است:

$$\text{loss}(\lambda, y) = \lambda - y \log(\lambda) + \log(y!)$$

با دادن 'ورودی'^۱ (نرخ پیش‌بینی λ) و 'هدف'^۲ (تعداد مشاهده شده y) به عنوان بردارها:

$$\text{input} = [\lambda_1, \lambda_2, \dots, \lambda_n]$$

$$\text{target} = [y_1, y_2, \dots, y_n]$$

تابع هزینه برای دسته‌ای از داده‌های صورت زیر محاسبه می‌شود:

$$\text{loss}(\text{input}, \text{target}) = \sum_{i=1}^n (\lambda_i - y_i \log(\lambda_i) + \log(y_i!))$$

۷.۵ ارزیابی عملکرد

با توجه به پیچیدگی و تنوع مسئله‌ی تخمین پایل آپ، تکیه بر یک معیار برای ارزیابی اثربخشی یک الگوریتم تخمین عملی نیست. برای ارزیابی دقیق عملکرد، ما توصیه می‌کنیم از مجموعه‌ای از معیارها استفاده شود. نتیجه‌ی کلی را می‌توان از نظر عملکرد وایازش و دسته‌بندی ارزیابی کرد، بدون توجه به فرمول‌بندی مسئله یا روش به کار رفته. از آنجا که توزیع پواسونی موجود است، می‌توانیم از معیارهایی که در بخش بعد معرفی شده‌اند، استفاده کنیم. یک مدل با دقت بالا ممکن است برای پردازش رویدادها زمان زیادی بگیرد که آن را برای سناریوهای واقعی غیرعملی می‌کند. بنابراین، باید بین پیچیدگی مدل و عملکرد آن تعادل برقرار شود.

¹Input

²target

۱۰.۷.۵ ارزیابی با L1Loss

روشی برای اندازه‌گیری عملکرد مدل بر روی داده‌های اعتبارسنجی یا آزمون ارائه می‌دهند. کمک می‌کنند تا بفهمیم مدل تا چه اندازه در زمینه دقت، صحت، بازآوری، نمره F1 و غیره، بسته به وظیفه خاص، عملکرد دارد. برای مقایسه مدل‌های مختلف یا نسخه‌های مختلف یک مدل استفاده می‌شوند. کمک می‌کنند تا بر اساس عملکرد در مجموعه اعتبارسنجی بهترین مدل انتخاب شود. برخلاف توابع هزینه که ممکن است در زمینه اهمیت واقعی قابل تفسیر نباشند، ماتریس‌های ارزیابی اغلب به گونه‌ای انتخاب می‌شوند که قابل تفسیر و مرتبط با مسئله خاص باشند.

تعريف ریاضی

- یک تانسور ورودی x با المان‌های x_i
 - یک تانسور هدف y با المان‌های y_i
 - هر دو تانسور x و y دارای N المان هستند (یعنی $\{1, 2, \dots, N\}$)
- تفاوت مطلق^۱ $L1$ برای هر المان i به صورت زیر: تعریف می‌شود

$$l_i = |x_i - y_i|$$

پارامتر کاهش^۲ مشخص می‌کند که اتلاف‌های هر المان چگونه جمع می‌شوند:

۱. بدون کاهش^۳

$$\ell(x, y) = \{l_1, l_2, \dots, l_N\}$$

در اینجا، خروجی یک تانسور شامل تفاوت‌های مطلق هر المان است

۲. کاهش به صورت میانگین^۴

$$\ell(x, y) = \frac{1}{N} \sum_{i=1}^N |x_i - y_i|$$

این روش میانگین تفاوت‌های مطلق را در تمام المان‌ها محاسبه می‌کند.

^۱Element-wise Loss

^۲reduction

^۳reduction='none'

^۴reduction='mean'

۳. کاهش به صورت مجموع^۱

$$\ell(x, y) = \sum_{i=1}^N |x_i - y_i|$$

این روش مجموع تفاوت‌های مطلق را در تمام المان‌ها محاسبه می‌کند.

این فرمول‌ها درک ریاضی واضحی از نحوه محاسبه خطرا در PyTorch ارائه می‌دهند.

۲۰۷.۵ ارزیابی عملکرد فیزیکی

در این روش ضمن ارزیابی اولیه توسط L1Loss با بازسازی توزیع پواسونی از داده‌های حقیقی به ازای هر پایل آپ، می‌توان عملکرد مدل را در تخمین پایل آپ ارزیابی کرد.

ارزیابی اربیبی بازسازی توزیع پواسونی و برآورد میانگین توزیع توسط مدل یکی از اهداف بهکارگیری مدل شبکه عصبی گرافی بر روی داده‌های واقعی، این است که بتوانیم با تخمین میانگین تعداد وقایع، میانگین توزیع خروجی را به عنوان برآورده‌گر پارامتر توزیع پواسونی اصلی گزارش کنیم:

$$\begin{aligned}\hat{\lambda} &= \text{avg} \left(\left\{ \text{GNN} \left(\text{event}_i \right) \forall i \in \{ \text{luminosity given a at collection data} \} \right\} \right) \\ &= L \times \sigma\end{aligned}$$

ه منظور فراهم کردن شرایط مشابه، کاری که انجام می‌دهیم این است که از مجموعه داده‌های شبیه‌سازی شده یک توزیع را رسم کنیم که یک توزیع پواسونی داشته باشد. پس از آن، این مجموعه نمونه‌گیری شده را به عنوان ورودی به شبکه عصبی آموزش دیده می‌دهیم و سپس توزیع خروجی را بررسی می‌کنیم. به منظور برآورده‌گر پارامتر توزیع پواسونی از روش بیشینه درست‌نمایی استفاده می‌کنیم که برای توزیع پواسونی، برآورده‌گر همان میانگین داده‌ها خواهد بود:

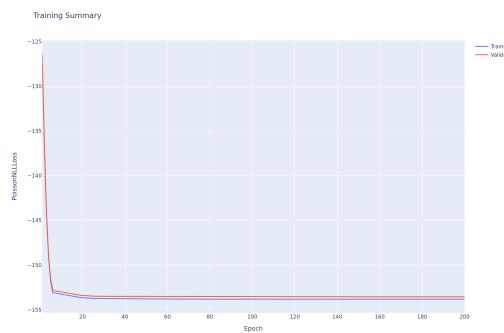
$$\hat{\lambda}_{MLE} = \text{mean}(\{\text{GNN}(\text{data}), \forall \text{data} \in \text{collection}\})$$

توقع داریم اگر همین کار را برای نمونه‌های متفاوت با توزیع‌های مختلف تکرار کنیم، آنگاه نمودار میانگین برآورده شده بر حسب میانگین اصلی، خط راستی باشد با شیب ۱ و عرض از مبدأ صفر. این تحلیل نشان می‌دهد که مدل شبکه عصبی گرافی ما باید به خوبی پارامتر توزیع پواسونی را برآورد کند و نتایج دقیق و قابل اعتمادی ارائه دهد. این مدل با استفاده از تخمین میانگین وقایع و ارزیابی بر اساس توزیع‌های مختلف، باید قادر باشد بازسازی دقیقترا از توزیع داده‌های واقعی ارائه دهد و به پژوهشگران

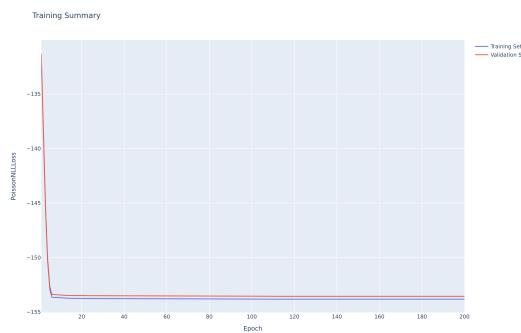
^۱reduction='sum'

امکان می‌دهد تا تحلیل‌های آماری دقیق‌تری انجام دهد.

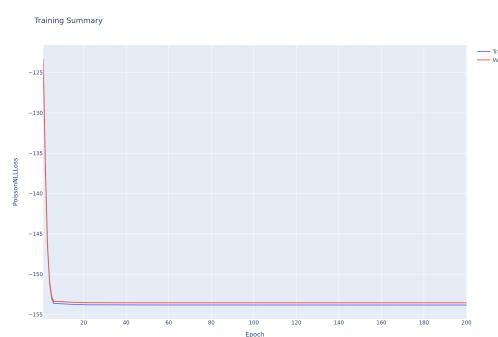
۸.۵ فرایند آموزش مدل‌ها



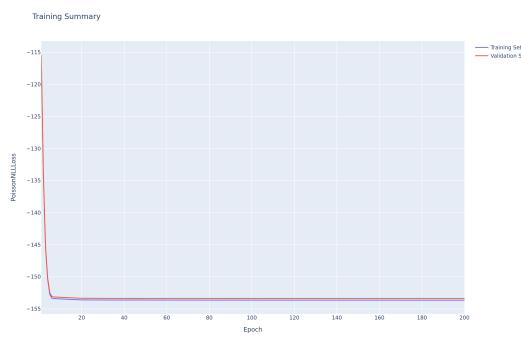
(ب) فرایند آموزش مدل PUGAT با وجود شناسه ذرات



(الف) فرایند آموزش مدل PUGAT بدون وجود شناسه ذرات

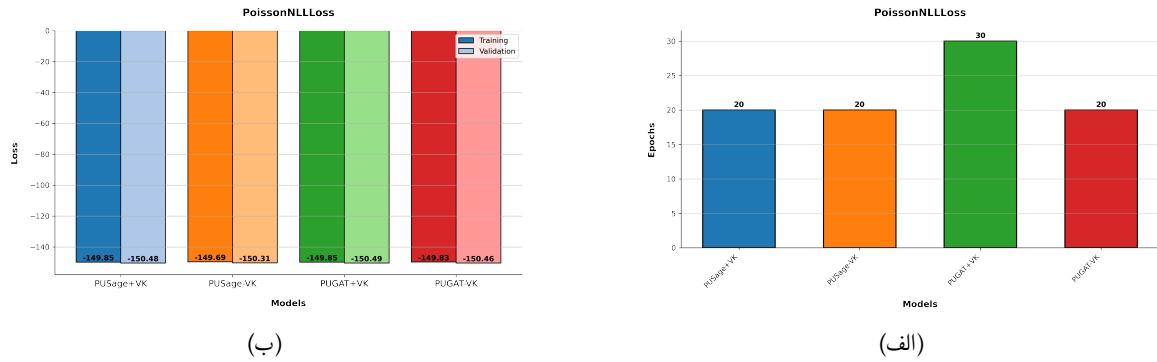


(د) فرایند آموزش مدل PUSage با وجود شناسه ذرات



(ج) فرایند آموزش مدل PUSage بدون وجود شناسه ذرات

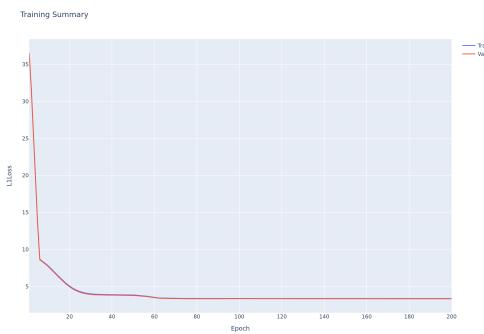
شکل ۸.۵: هر دو منحنی تابع هزینه PoissonNLLLoss مجموعه آموزشی و اعتبارسنجی در تمام نمودارها به طور پیوسته کاهش می‌یابند، نشان می‌دهد که مدل به طور مؤثر از داده‌های آموزشی یاد می‌گیرد. اگر منحنی‌ها بعد از چندین دور به صورت افقی حرکت کنند یا افزایش یابند، ممکن است نشان‌دهنده بیش‌برازش یا ناکافی بودن داده‌های آموزشی باشد که آن قدری نیست وارد رژیم بیش‌برازش شویم:



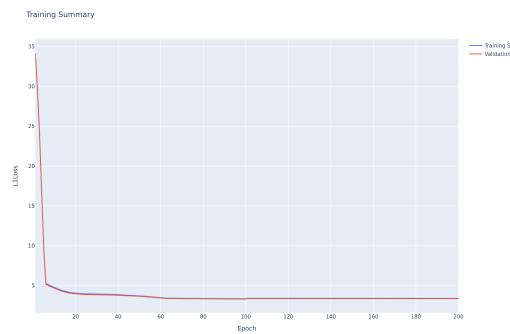
شکل ۹.۵: مقایسه تابع هزینه مدل ها

جدول ۱.۵: PoissonNLLLoss

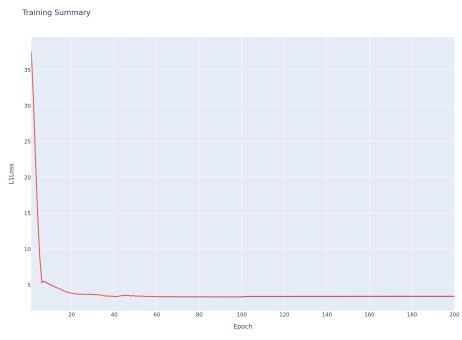
اعتبارسنجی	آموزش	
-۱۵۰.۴۸	-۱۴۹.۸۵	PUSage+VK
-۱۵۰.۳۱	-۱۴۹.۶۹	PUSage-VK
-۱۵۰.۴۹	-۱۴۹.۸۵	PUGAT+VK
-۱۵۰.۴۶	-۱۴۹.۸۳	PUGAT-VK



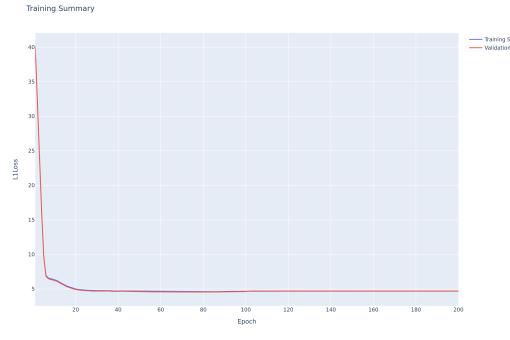
(ب) ارزیابی با خطای مطلق برای PUGAT با وجود شناسه ذرات.
خطای مجموعه آموزشی و اعتبارسنجی حدود ۳.۲ است.



(الف) ارزیابی با خطای مطلق برای PUGAT بدون وجود شناسه ذرات.
خطای مجموعه آموزشی و اعتبارسنجی حدود ۳.۵ است.



(د) ارزیابی با خطای مطلق برای PUSage با وجود شناسه ذرات.
خطای مجموعه آموزشی و اعتبارسنجی حدود ۳.۳ است.

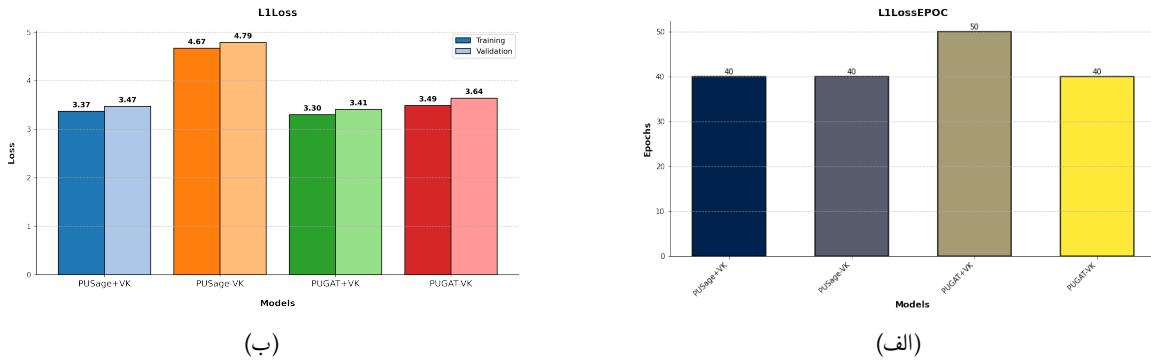


(ج) ارزیابی بدون خطای مطلق برای PUSage با وجود شناسه ذرات.
خطای مجموعه آموزشی و اعتبارسنجی حدود ۴.۶ است.

شکل ۱۰.۵: ارزیابی اثربخشی مدل در برآورد پایل آپ: L1Loss بر روی مجموعه داده‌ها به طور پیوسته کاهش می‌یابد، نشان می‌دهد که مدل در حال یادگیری انجام پیش‌بینی‌های دقیق‌تر بر روی داده‌های ناشناخته است. این یک نشانه مثبت است.

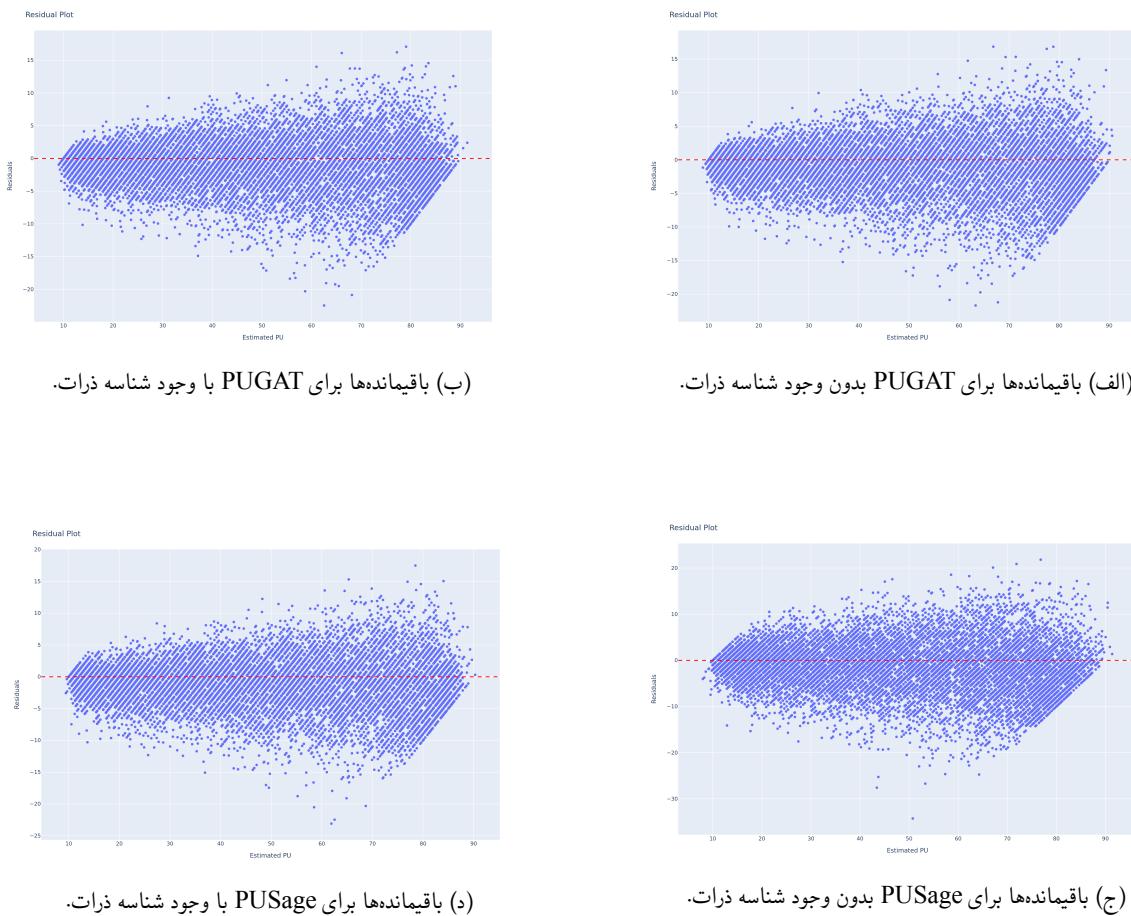
جدول ۲.۵

اعتبارسنجی	آموزش	
۳.۴۷	۳.۳۷	PUSage+VK
۴.۷۹	۴.۶۷	PUSage-VK
۳.۴۱	۳.۳۰	PUGAT+VK
۳.۶۴	۳.۴۹	PUGAT-VK



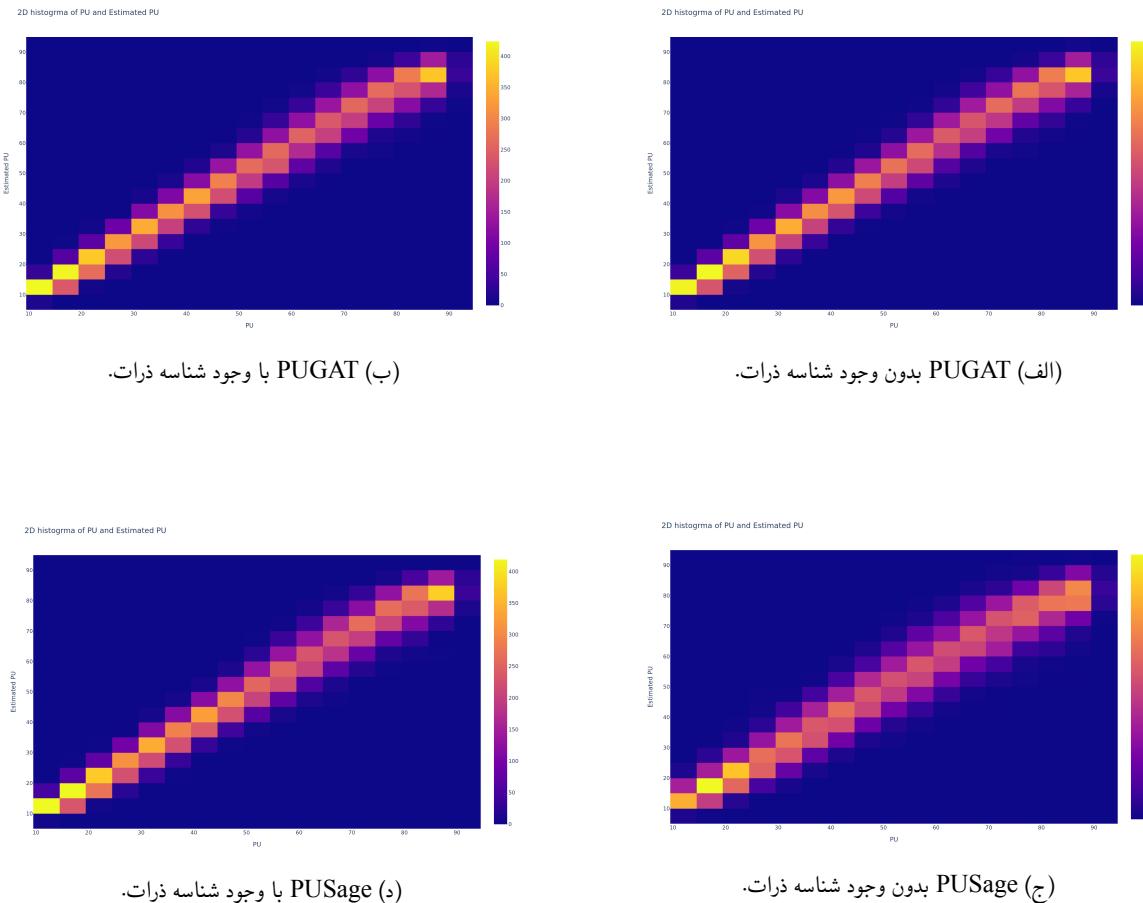
شکل ۱۱.۵: مقایسه ماتریس عملکرد مدل ها

۹.۰۵ توزیع خطاهای در پیش‌بینی‌های مدل



شکل ۱۲.۵: شکل باقیماندهای بینش‌هایی درباره توزیع خطاهای در پیش‌بینی‌های مدل ارائه می‌دهد. واریانس خطاهای در سرتاسر محدوده پیش‌بینی‌ها ثابت نیست. باقیماندهای تمایل دارند برای مقادیر بالاتر تعداد تعاملات بیشتر منفی باشند. این نشان می‌دهد که مدل تعداد تعاملات را در این موارد کم برآورد می‌کند.

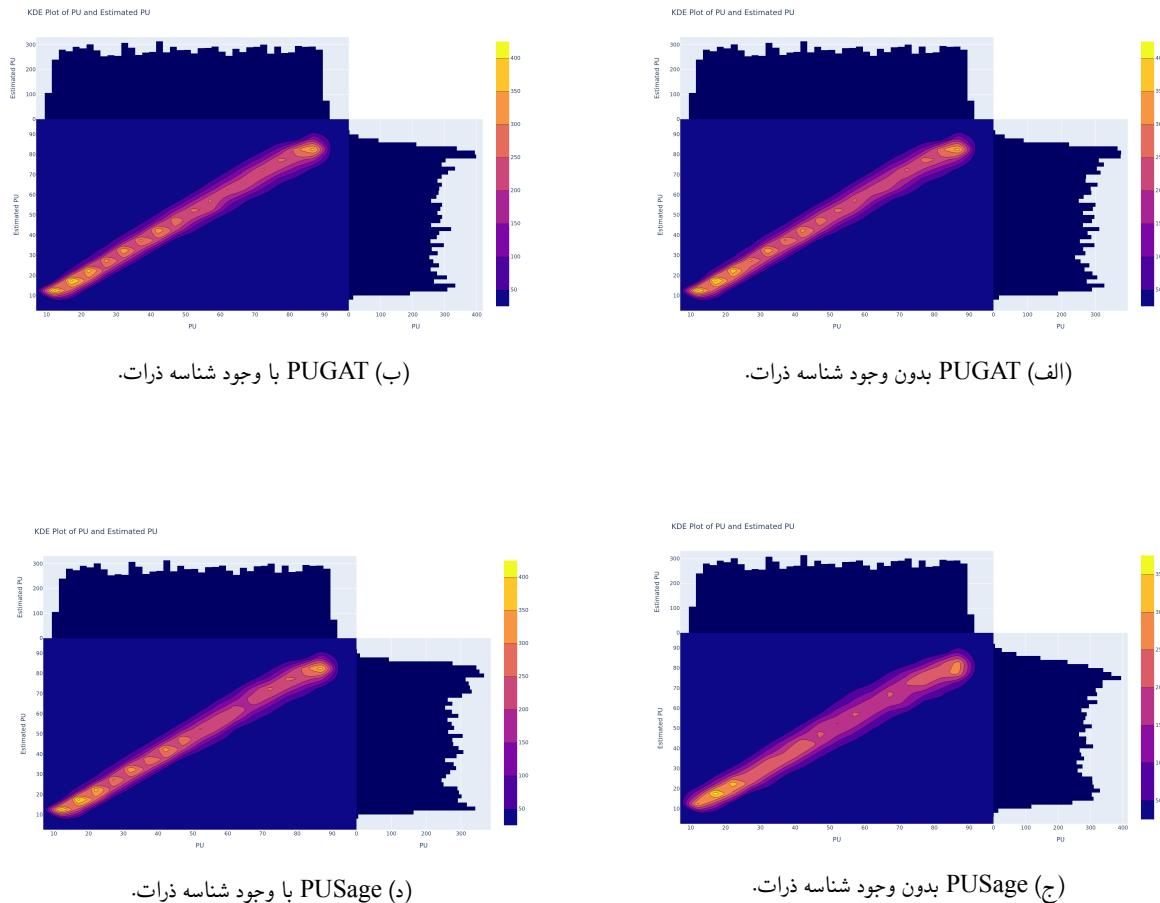
۱۰.۵ توزیع مقادیر پایل آپ، حقیقی و تخمینی



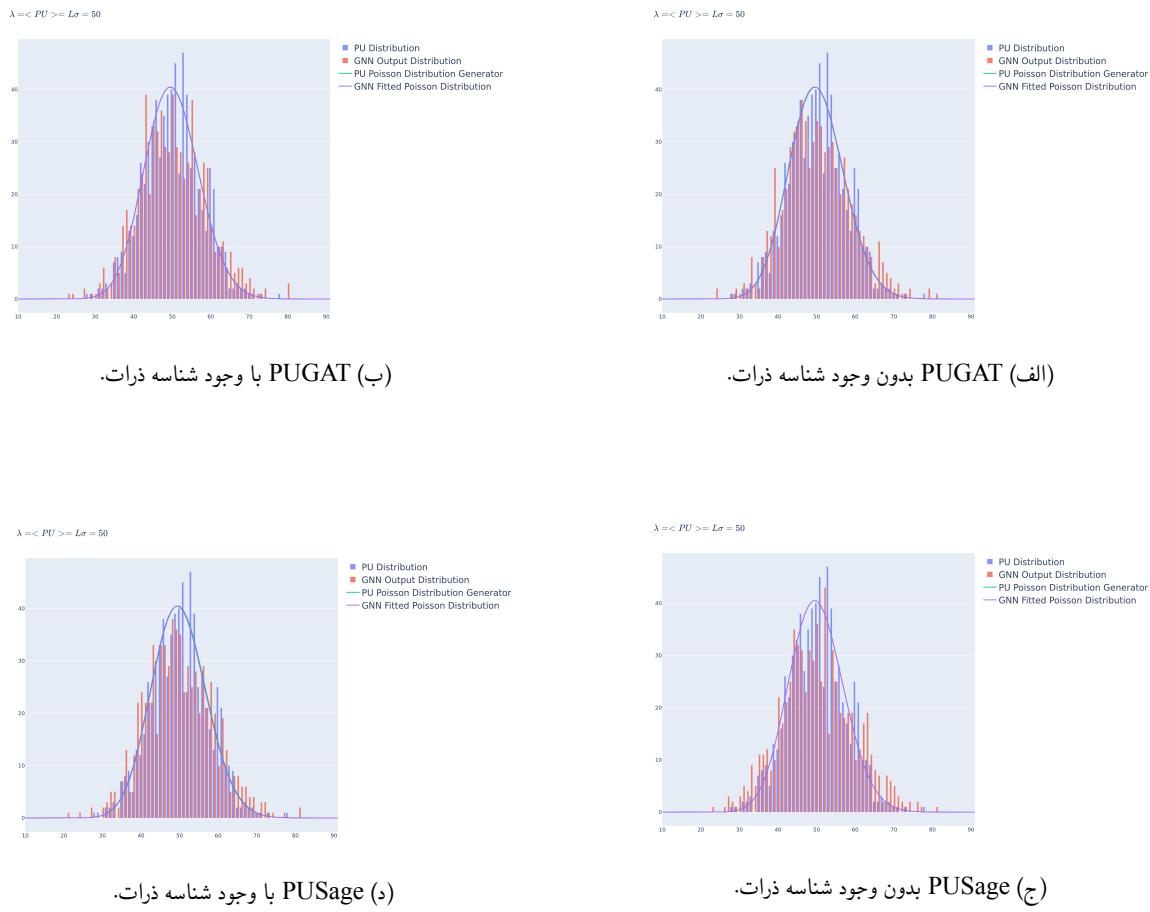
شکل ۱۳.۵: توزیع مقادیر پایل آپ، حقیقی و تخمینی توسط مدل را نمایش می‌دهند. نمودارها نشان می‌دهند که پیش‌بینی‌های مدل از نظر توزیع به مقادیر واقعی PU نزدیک هستند و این امر عملکرد کلی خوب مدل را در بدقت گرفتن توزیع زیربنایی نشان می‌دهد.



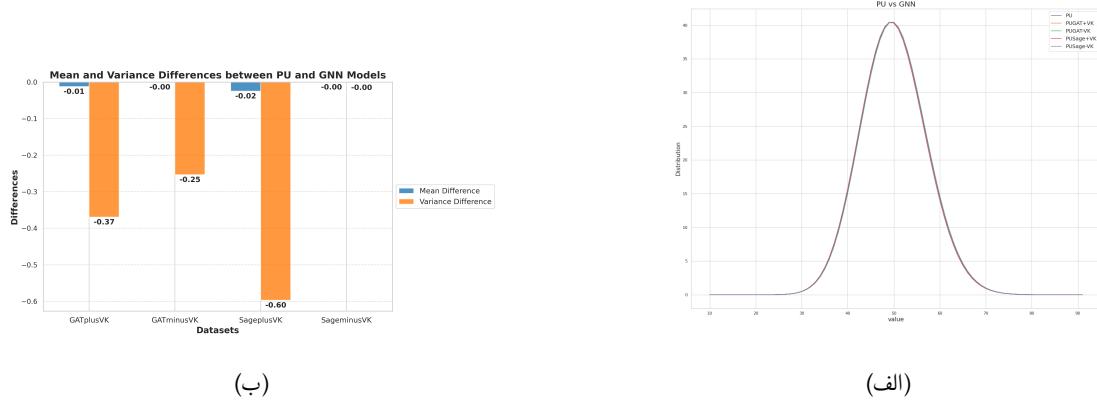
شکل ۱۴.۵: هیستوگرام‌های متغیر پاسخ و مقادیر تخمینی. توزیع واقعی پایل آپ به صورت یکنواخت‌تر پخش شده است. توزیع تخمینی پایل آپ در مقادیر پایین‌تر پیک‌های بالاتری دارد و به سمت مقادیر بالاتر دم دارد. مدل تمایل به کم‌تخمین زدن مقادیر پایل آپ، به ویژه در مقادیر بالاتر دارد.



شکل ۱۵.۵: نمودار تخمین چگالی کرنل (KDE)، همپوشانی بین توزیع‌های واقعی و تخمینی پایل آپ را به تصویر می‌کشد. دو توزیع باید با هم هماهنگ باشند، که نشان‌دهنده توافق خوب بین مقادیر واقعی و تخمینی است. مناطقی با چگالی بیشتر در نمودار نشان‌دهنده تعداد بیشتری از نقاط داده با مقادیر پایل آپ است.



شکل ۱۶.۵: مقایسه توزیع خروجی مدل با توزیع پواسون . مورد انتظار. نمایانگر تعداد پایل آپ احتمال هر مقدار است. برای ارزیابی دقیقی از عملکرد مدل انجام می شود. منحنی آبی نمایانگر توزیع پواسون با میانگین ۵۰ است، در حالی که نوارهای قرمز داده های خروجی مدل را نشان می دهد.



شکل ۱۷.۵: مقایسه توزیع خروجی مدل‌ها با توزیع پواسون و تحلیل تفاوت میانگین و واریانس. نمودار سمت راست نمایانگر تعداد پایل آپ و احتمال هر مقدار است. در این نمودار، منحنی آبی نشان‌دهنده توزیع پواسون با میانگین 5° و نوارهای قرمز نشان‌دهنده داده‌های خروجی مدل‌ها است. نمودار سمت چپ تفاوت‌های میانگین و واریانس بین داده‌های مدل‌ها و پایل آپ را نشان می‌دهد که معیاری برای سنجش دقت مدل‌ها است.



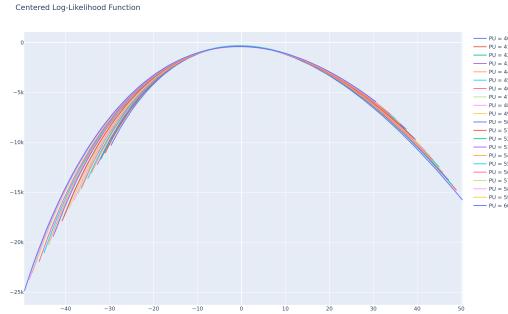
(ب) PUGAT با وجود شناسه ذرات.



(الف) PUGAT بدون وجود شناسه ذرات.



(د) PUSage با وجود شناسه ذرات.



(ج) PUSage بدون وجود شناسه ذرات.

شکل ۱۸.۵: توابع لگاریتم احتمال برای توزیع‌های پواسون مختلف. منحنی‌هایی با قله‌های بالاتر نشان‌دهنده تطابق بهتر بین توزیع تخمینی و توزیع پواسون واقعی هستند.

۱۱.۵ دقیق مدل در تخمین تعداد پایل آپ در یک محدوده درخشنندگی

در این بخش، دقیق مدل‌های مختلف در تخمین تعداد پایل آپ‌ها در یک محدوده درخشنندگی با استفاده از توزیع پواسون بررسی می‌شود. نتایج به دست آمده برای چهار مدل مختلف در شرایط مختلف (با و بدون ورود شناسه ذرات) مقایسه می‌شود. هر مدل با دقیق برآذش داده‌های شبیه‌سازی شده در برابر توزیع نظری پواسون ارزیابی شده است. شکل‌های ۲۱.۵-الف، ۲۱.۵-ب، ۲۱.۵-ج و ۲۱.۵-د نشان‌دهنده اعتبار مدل‌ها در تخمین تعداد پایل آپ با استفاده از توزیع پواسون می‌باشند. خطوط قرمز در این نمودارها نمایانگر خط ایده‌آل است که مدل باید به آن نزدیک شود تا نشان‌دهنده صحیح برآذش باشد.

شکل الف (PUGAT بدون ورود شناسه ذرات): این نمودار نشان می‌دهد که مدل PUGAT بدون ورود شناسه ذرات عملکرد نسبتاً خوبی دارد، به طوری که نقاط داده‌ها به خط ایده‌آل نزدیک هستند. این امر حاکی از آن است که این مدل در تخمین تعداد پایل آپ در این شرایط مناسب عمل کرده است.

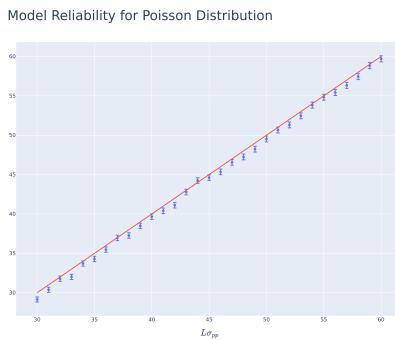
شکل ب (PUGAT با ورود شناسه ذرات): در حضور شناسه ذرات، دقیق مدل PUGAT کاهش می‌یابد و انحراف از خط

ایدهآل افزایش می‌یابد. این ممکن است به دلیل اثرات پیچیده شناسه ذرات باشد که باعث افزایش عدم قطعیت در تخمین‌ها می‌شود.

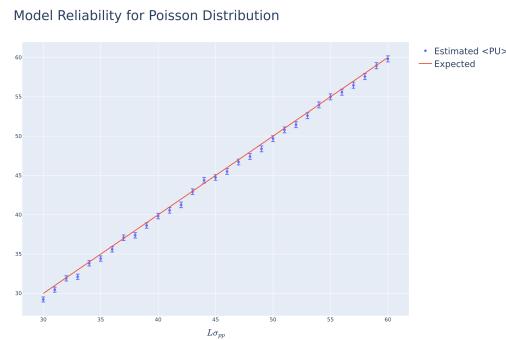
شکل ج (PUsage بدون ورود شناسه ذرات): در این حالت نیز مدل PUsage بدون ورود شناسه ذرات عملکرد خوبی را نشان می‌دهد و نقاط داده‌ها به خط ایدهآل نزدیک هستند. این نشان‌دهنده سازگاری مدل با داده‌ها شده است.

شکل د (PUsage با ورود شناسه ذرات): مشابه مدل PUGAT، در اینجا نیز مشاهده می‌شود که ورود شناسه ذرات باعث کاهش دقت مدل می‌شود. انحراف از خط ایدهآل در حضور شناسه ذرات بیشتر است.

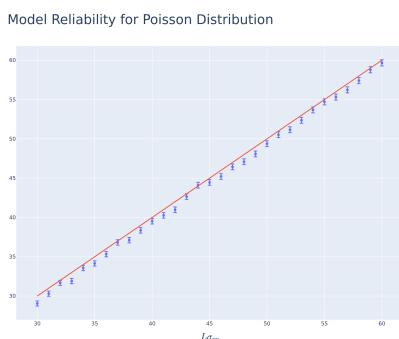
نتایج به دست آمده نشان می‌دهند که هر دو مدل PUGAT و PUsage در غیاب شناسه ذرات عملکرد بسیار خوبی دارند و به خوبی با توزیع پواسون برازش پیدا می‌کنند. اما در حضور شامت ذرات، دقت هر دو مدل کاهش می‌یابد. این نکته نشان می‌دهد که ورود شناسه ذرات پیچیدگی‌های جدیدی به تحلیل اضافه می‌کند که باید در نظر گرفته شود.



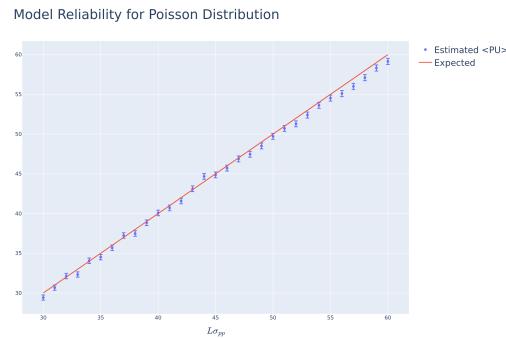
(ب) PUGAT با وجود شناسه ذرات.



(الف) PUGAT بدون وجود شناسه ذرات.

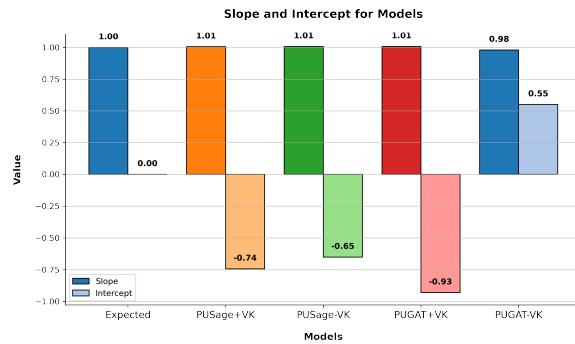


(د) PUSage با وجود شناسه ذرات.



(ج) PUSage بدون وجود شناسه ذرات.

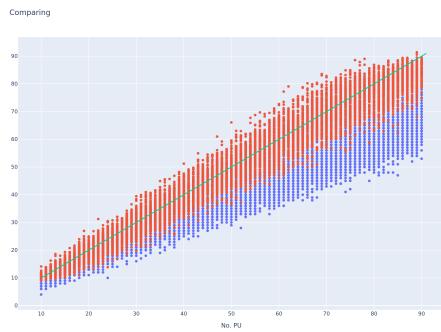
شکل ۱۹.۵: بررسی دقت مدل در تخمین تعداد پایل آپ در یک محدوده درخشنندگی.



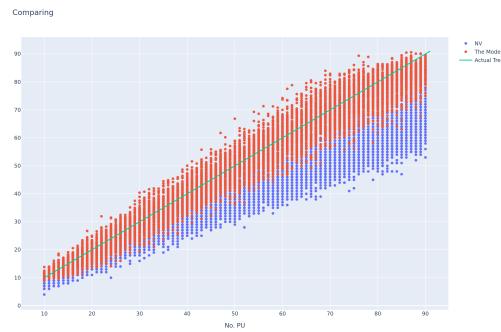
شکل ۲۰.۵: نتایج مقایسه‌ای مربوط به شیب و نقاط برخورد برای چهار مدل مختلف را نمایش می‌دهد. این نمودار به وضوح تفاوت عملکرد بین مدل‌ها را نشان می‌دهد.

نرده‌یک به ° هستند، به خط ایده‌آل برازش نزدیک‌تر بوده و به عنوان مدل‌های دقیق‌تر در تخمین پایل آپ شناخته می‌شوند. اختلاف مدل‌ها: این اختلاف نشان‌دهنده چالش‌های مرتبط با پیش‌بینی تعداد پایل آپ به ویژه در حضور اثرات شناسه ذرات است

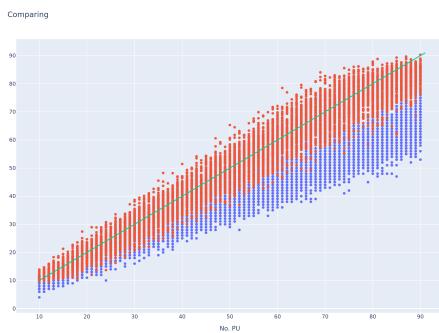
۱۲.۵ پراکندگی مقایسه خروجی مدل با داده های حقیقی



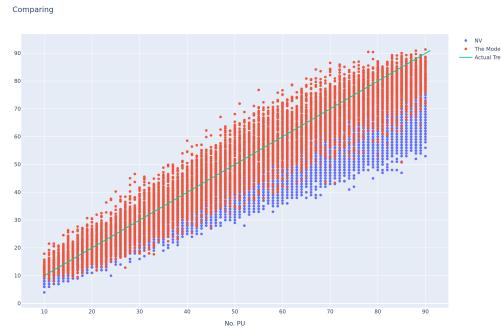
(ب) PUGAT با وجود شناسه ذرات.



(الف) PUGAT بدون وجود شناسه ذرات.



(د) PUSage با وجود شناسه ذرات.



(ج) PUSage بدون وجود شناسه ذرات.

شکل ۲۱.۵: نمودار پراکندگی برای مقایسه خروجی مدل با داده های حقیقی. این نمودارها مقایسه بصری از چگونگی پیش بینی تعداد پایل آپ توسط مدل های مختلف در مقایسه با مقادیر واقعی ارائه می دهند.

۱۳.۵ نتیجه‌گیری

جمع‌بندی نتایج اصلی در این تحقیق، مدل شبکه عصبی گراف برای تخمین تعداد پایل آپ در محدوده‌های مختلف درخشنگی مورد بررسی قرار گرفت. نتایج اصلی به شرح زیر است:

- مدل GNN توانست تعداد میانگین برخوردهای چندگانه را با دقت مناسبی تخمین بزند.
- مقادیر تخمین زده مدل با مقادیر واقعی همخوانی خوبی داشتند، که در نمودارهای ارائه شده به خوبی نمایان است.
- نمودارهای لاغ-احتمال نشان‌دهنده توانایی مدل در یافتن بیشینه مقدار لاغ-احتمال برای توزیع‌های مختلف پوآسون بودند.

تفسیر و تحلیل نتایج نتایج به دست آمده نشان‌دهنده این است که مدل GNN توانسته است با دقت خوبی پایل آپ را تخمین بزند. این نتایج اهمیت بالایی در مطالعات برخوردهای پروتون-پروتون و تخمین مقادیر مربوط به آن‌ها دارد. تطابق خوب بین مقادیر تخمین زده شده و مقادیر واقعی، نشان‌دهنده کارایی مدل در شرایط شبیه‌سازی شده است. این تحقیق دارای محدودیت‌هایی است که می‌تواند بر نتایج تأثیرگذار باشد: داده‌های استفاده شده برای آموزش و ارزیابی مدل، شبیه‌سازی شده هستند و ممکن است تفاوت‌هایی با داده‌های واقعی داشته باشند. مدل GNN به دلیل پیچیدگی‌های محاسباتی ممکن است نیاز به منابع محاسباتی بالایی داشته باشد که در شرایط عملی محدودیت ایجاد کند. تعداد داده‌های استفاده شده برای آموزش و ارزیابی ممکن است کافی نباشد و نیاز به داده‌های بیشتری برای افزایش دقت مدل باشد.

در این پژوهش، برای اولین بار در آزمایش‌های برخوردهای بزرگ هادرونی، موفق به استخراج توزیع پایل آپ به صورت مستقل از درخشندگی شدیم. این دستاورده نه تنها نشان‌دهنده پیشرفت در دقت محاسبات پایل آپ است، بلکه می‌تواند به طور مستقیم به بهبود تحلیل‌های مربوط به برخوردهای ذرات منجر شود. پایل آپ یکی از چالش‌های اصلی در آزمایش‌های LHC است، چرا که می‌تواند داده‌ها را تحت تأثیر قرار داده و باعث افزایش خطأ در اندازه‌گیری‌ها شود. این روش جدید می‌تواند به طور موثری این مسئله را کاهش دهد و نتایج دقیق‌تری را ارائه دهد.

برای مدل‌سازی این برخوردها، از یک رویکرد نوین استفاده کردیم که در آن رویدادهای برخورد ذرات به صورت یک گراف نمایش داده شدند. در این مدل، ذرات به عنوان گره‌ها و ارتباطات میان آن‌ها به عنوان یال‌ها تعریف شدند. این روش مدل‌سازی گرافی امکان تحلیل دقیق‌تر و جامع‌تری از ساختار و دینامیک برخوردهای ذرات را فراهم کرد. یکی از نکات کلیدی این روش، توانایی آن در بررسی روابط پیچیده بین ذرات و اثرات متقابل آن‌ها است که در مدل‌های سنتی کمتر به آن توجه شده است. در ادامه، مطالعه‌ای جامع در مورد اهمیت نزدیکی ذرات به رأس انجام شد. این بررسی نشان داد که فاصله ذرات از رأس برخورد تأثیر مستقیمی بر کیفیت نتایج تحلیل دارد. ذراتی که به رأس نزدیک‌تر هستند، معمولاً به دلیل ارتباط قوی‌تر با رویداد اصلی، داده‌های دقیق‌تری ارائه می‌دهند. در مقابل، ذرات دورتر ممکن است تحت تأثیر عوامل جانبی قرار گیرند که می‌تواند دقت تحلیل را کاهش دهد. این یافته‌ها اهمیت ویژه‌ای در انتخاب داده‌های مناسب برای تحلیل‌های دقیق دارد.

برای بررسی تأثیر این ویژگی در مدل‌سازی، دو مدل شبکه عصبی گرافی، یعنی شبکه توجه گرافی و شبکه نمونه برداری و جمع آوری را توسعه و آموزش دادیم. در این مدل‌ها، شناسه نزدیک‌ترین رأس به عنوان یک ویژگی کلیدی به ماتریس ویژگی‌های هر گره اضافه شد. این کار باعث شد تا مدل‌ها بتوانند با دقت بیشتری توزیع پایل آپ را شناسایی کنند.

به طور خاص، افزودن اطلاعات مربوط به نزدیک‌ترین رأس به ماتریس ویژگی‌ها، باعث بهبود عملکرد مدل‌ها در شناسایی و جدا کردن رویدادهای پایل آپ از رویدادهای اصلی شد. این بهبود عملکرد، نشان‌دهنده اهمیت این ویژگی در تحلیل داده‌های پیچیده برخوردهای ذرات است.

در نهایت، نتایج به دست آمده نشان داد که روش پیشنهادی ما می‌تواند به طور موثری پایل آپ را شناسایی و درخشندگی را تعیین

کند. این موفقیت نه تنها باعث افزایش دقت در تحلیل داده‌های برخورده‌تنه بزرگ هادرونی خواهد شد، بلکه می‌تواند به عنوان یک روش مرجع برای سایر آزمایش‌های برخورد ذرات نیز به کار گرفته شود. این روش می‌تواند در آینده به بهبود رویکردهای تحلیل داده در آزمایش‌های فیزیک ذرات و سایر تحقیقات مرتبط کمک شایانی کند. همچنین، این دستاوردهای می‌تواند به توسعه مدل‌های پیشرفته‌تر برای شناسایی دقیق‌تر پایل آپ و بهبود کیفیت داده‌ها در مطالعات آتی منجر شود.

پیشنهادات برای بهبود و گسترش تحقیق

۱. تبدیل گراف به گراف ناهمگون: با تبدیل گراف به گراف ناهمگون و پیچیده‌تر کردن ساختار گراف، شبکه عصبی گراف قادر خواهد بود روابط پیچیده‌تری را بین داده‌ها شناسایی و مدل‌سازی کند. این می‌تواند دقت تخمین‌ها را بهبود بخشد.
۲. اعمال روش‌های تحلیل شبکه‌های پیچیده: استفاده از روش‌هایی مانند شبکه‌های معنایی، تجزیه و تحلیل شبکه‌های تطبیقی، و تحلیل شبکه‌های دینامیکی برای بهبود تفاوت‌های دینامیکی در گراف‌های رویداد.
۳. مقایسه با رویکردهای دیگر: اعمال روش‌های تحلیل شبکه‌های پیچیده مقایسه گراف رویداد با سایر روش‌ها و مدل‌های استفاده شده در فیزیک سامانه‌های پیچیده برای بررسی و تحلیل فرآیندهای پیچیده.
۴. تعمیق مفاهیم نظریه گراف بررسی بهتر و تعمیق در مفاهیم نظریه گراف مورد نیاز برای مدل‌سازی و تحلیل دقیق‌تر گراف‌های رویداد.
۵. افزایش پیچیدگی مدل: استفاده از مدل‌های پیچیده‌تر و عمیق‌تر مانند شبکه‌های عصبی چندلایه یا مدل‌های ترکیبی می‌تواند به بهبود عملکرد مدل کمک کند.
۶. افزایش داده‌های آموزشی: جمع‌آوری و استفاده از داده‌های بیشتر و متنوع‌تر می‌تواند به مدل کمک کند تا الگوهای پیچیده‌تری را یاد بگیرد و عملکرد بهتری داشته باشد.
۷. استفاده از تکنیک‌های پیشرفته داده‌افزایی: استفاده از تکنیک‌هایی مانند تولید داده‌های مصنوعی یا oversampling برای افزایش تعداد نمونه‌های با PU بالا می‌تواند به بهبود تخمین‌های مدل کمک کند.
۸. استفاده از توابع خطای وزنی: استفاده از توابع خطای وزنی که به خطاهای در محدوده‌های خاص وزن بیشتری می‌دهند می‌تواند مدل را به بهبود تخمین‌ها در این محدوده‌ها هدایت کند.
۹. ارزیابی مستمر عملکرد مدل: استفاده از معیارهای ارزیابی حساس به عدم همگنی و خطای تخمین مانند میانگین خطای مطلق یا ریشه میانگین مربع خطای برای ارزیابی مستمر عملکرد مدل می‌تواند به شناسایی نقاط ضعف و بهبود آن‌ها کمک کند.
۱۰. گسترش تحقیق به شاخه‌های دیگر CMSSW: با کامل‌تر کردن تحقیق، نتایج و روش‌های ارائه شده می‌توانند به عنوان افزونه‌ای در شاخه‌های دیگر CMSSW مورد استفاده قرار گیرند و به بهبود عملکرد کل سیستم کمک کنند.

این پیشنهادات می‌توانند به بهبود دقت تخمین‌ها و کارایی مدل کمک کرده و راه را برای تحقیقات آینده هموارتر کنند.

کاربردها و پیامدهای عملی در تحقیقات فیزیک ذرات و مطالعات برخوردهای پروتون-پروتون برای تخمین دقیق‌تر پایل آپ. در طراحی و بهبود دستگاه‌های آشکارسازی برخوردها و بهینه‌سازی عملکرد آن‌ها. در تحلیل داده‌های تجربی و بهبود روش‌های تحلیل داده‌های برخوردها.

نتیجه‌گیری نهایی این تحقیق نشان‌دهنده کارایی و دقت مدل شبکه عصبی گراف در تخمین تعداد میانگین برخوردهای چندگانه در شرایط شبیه‌سازی شده است. با استفاده از این مدل می‌توان به تخمین‌های دقیقی از برخوردهای پروتون-پروتون دست یافت که اهمیت بالایی در مطالعات فیزیک ذرات دارد. آینده این پژوهش می‌تواند با استفاده از داده‌های واقعی و بررسی کاربردهای مختلف مدل، بهبود یابد و نتایج ارزشمندی را به همراه داشته باشد.

پیامدها و چشم‌اندازهای آینده بهبود تخمین پایل آپ می‌تواند بازسازی ذرات، انتخاب رویدادها و ردیابی پس‌زمینه را بهبود بخشد و در نهایت حساسیت جستجوهایی برای پدیده‌های فیزیکی جدید را افزایش دهد. همچنین، در گسترش مرزهای درخشندگی، روش‌های نوآورانه مانند این رویکردها برای حفظ دقت و کیفیت بالا در تحلیل داده‌ها اساسی خواهند بود.

پیوست‌ها

۱۰. کد محاسبه یال‌ها در گراف‌ها

تابع برای محاسبه یال‌ها:

تابع `GetEdgeInfo` برای تعریف یال‌ها بین گره‌ها در گراف استفاده شده است. این تابع فاصله‌های بین هر جفت ذره را بر اساس مقادیر شبه سریعی (η) و زاویه آزموتال (ϕ) محاسبه می‌کند. یک آستانه به نام $max_distance$ تعیین می‌کند

که آیا ذرات متصل در نظر گرفته شوند یا خیر (یعنی آیا یال‌ای بین آنها وجود دارد). مراحل اجرای تابع به شرح زیر است:

- محاسبه تفاوت‌های ϕ بین هر جفت ذره و تنظیم آنها برای مقادیر بیشتر از π .
- محاسبه تفاوت‌های η بین هر جفت ذره.
- محاسبه فاصله اقلیدسی در فضای $\eta\text{-}\phi$.
- تنظیم مقادیر قطر اصلی ماتریس فاصله به مقدار بالا برای جلوگیری از یال‌های خودی (self-loop).
- تنظیم مقادیر مثلث پایینی ماتریس فاصله به مقدار بالا برای جلوگیری از یال‌های تکراری.
- اعمال آستانه فاصله و استخراج یال‌های معتبر.
- استخراج ویژگی‌های یال برای یال‌های معتبر.

کد مربوط به تابع `GetEdgeInfo` در زیر آورده شده است:

```
def GetEdgeInfo(eta , phi ,max_distance):
    pi = ROOT.TMath.Pi()
    phi_diff = abs( phi.unsqueeze(1) - phi.unsqueeze(0) )
    #phi_diff = phi_diff.remainder(2 * pi)
    mask = phi_diff > pi
    delta_phi = phi_diff - mask*2*pi
    eta_diff = eta.unsqueeze(1) - eta.unsqueeze(0)
    distances = torch.sqrt(eta_diff ** 2 + delta_phi ** 2)

    torch.diagonal(distances).fill_(1000.0)
    distances[ torch.tril(torch.ones(distances.shape, dtype=torch.bool)) ] = 1000.0

    mask = distances < max_distance
    indices = mask.nonzero().t()
    edge_index = indices.flip(0)

    edge_attr = distances[mask]

    return edge_index , edge_attr
```

توضیح

محاسبه تفاوت‌های `:phi`

تفاوت مطلق بین مقادیر `phi` هر جفت ذره $\text{phi.unsqueeze}(0) - \text{abs}(\text{phi.unsqueeze}(1)) = \text{phi_diff}$ • را محاسبه می‌کند.

مسکنی ایجاد می‌کند که تفاوت‌هایی که از π بیشتر هستند را مشخص می‌کند.

$\text{pi} > \text{phi_diff} = \text{mask}$ • تفاوت‌های بیشتر از π را تنظیم می‌کند تا تناوب زاویه آزمیوتال را در نظر بگیرد.

محاسبه تفاوت‌های `:eta`

تفاوت در `eta` بین هر جفت ذره را محاسبه می‌کند. $\text{eta.unsqueeze}(0) - \text{eta.unsqueeze}(1) = \text{eta_diff}$ • محاسبه فاصله‌ها:

فاصله اقلیدسی در فضای `eta-phi` را $\text{distances} = \text{torch.sqrt}(\text{eta_diff} ** 2 + \text{delta_phi} ** 2)$ می‌کند.

تنظیم مقادیر قطر و مثلث پایینی به مقدار بالا:

- مقدار بالا تنظیم می‌کند تا از یال‌های خودی جلوگیری شود.
- $0.1000 = \text{dtype}=\text{torch.bool}))]$ $\text{distances}[\text{torch.tril}(\text{torch.ones}(\text{distances.shape},$
- مثلث پایینی ماتریس فاصله را به مقدار بالا تنظیم می‌کند تا از یال‌های تکراری جلوگیری شود.
- اعمال آستانه فاصله:

ماسکی ایجاد می‌کند برای فاصله‌هایی که کمتر از آستانه هستند.

$\text{max_distance} < \text{distances} = \text{mask}$

شاخص‌های یال‌های معتبر را استخراج می‌کند.

$\text{mask.nonzero()}.t() = \text{indices}$

شاخص‌ها را برای ترتیب صحیح معکوس می‌کند.

$\text{indices.flip}(0) = \text{edge_index}$

استخراج ویژگی‌های یال:

$\text{distances}[\text{mask}] = \text{edge_attr}$

بازگرداندن نتایج:

شاخص‌ها و ویژگی‌های یال‌ها را بازمی‌گرداند.

این تابع در ساخت گراف‌هایی که نماینده تعاملات ذرات هستند بسیار اساسی است. در این گراف‌ها، گره‌ها نماینده ذرات منفرد هستند و یال‌ها نماینده اتصالات فضایی بین آنها بر اساس یک آستانه فاصله مشخص هستند.

۲۰. کد جمع‌آوری ویژگی‌های ذرات و ساخت گراف‌ها

در این بخش، کدی که برای جمع‌آوری ویژگی‌های ذرات و ساخت گراف‌ها استفاده شده است، ارائه می‌شود. این کد برای هر رویداد، ویژگی‌های مختلف ذرات را استخراج کرده و از آنها برای ساخت گراف استفاده می‌کند. مراحل اجرای این کد به شرح زیر است:

```
file = ROOT.TFile(opt.input)
tree = file.Get("PUAnalyzer/Trees/Events")
```

توضیح

در این کد:

```

num_events = tree.GetEntries()
for e in tqdm( range(num_events) ):
    tree.GetEntry(e)
    # Collect features for particles
    phi = torch.tensor(tree.Phi, dtype=torch.float32)
    energy = torch.tensor(tree.Energy, dtype=torch.float32)
    p = torch.tensor(tree.P, dtype=torch.float32)
    pt = torch.tensor(tree.Pt, dtype=torch.float32)
    dz = torch.tensor(tree.dz, dtype=torch.float32)
    dxy = torch.tensor(tree.dxy, dtype=torch.float32)
    eta = torch.tensor(tree.Eta, dtype=torch.float32)
    Type = torch.tensor(tree.Type, dtype=torch.int32)
    charge = torch.tensor(tree.Charge, dtype=torch.int32)
    vertexKey = torch.tensor(tree.vertexKey, dtype=torch.int32)

    # Collect attributes for events
    nVertices = torch.tensor([tree.nVertices], dtype=torch.int32)
    nVGoodVertices = torch.tensor([tree.nVGoodVertices], dtype=torch.int32)

    # node_features = torch.stack((phi, charge, energy, p, pt, dz, dxy, eta, Type), dim=1)
    # graph_attr = torch.stack((nVertices, nVGoodVertices), dim=0)

    node_features = torch.stack((phi, pt, dz, dxy, eta, Type, charge, energy, p, vertexKey), dim=1)
    graph_attr = torch.stack((nVertices,), dim=0)

    edge_index, edge_attr = GetEdgeInfo(eta, phi, opt.maxDR)

```

- تعداد رویدادها را از درخت داده‌ها دریافت می‌کند.
- حلقه `for e in tqdm(range(num_events))` برای هر رویداد تکرار می‌شود.
- دستور `tree.GetEntry(e)` رویداد فعلی را از درخت داده‌ها بارگذاری می‌کند.
- ویژگی‌های مختلف ذرات مانند `phi`, `charge`, `Type`, `eta`, `dxy`, `dz`, `p`, `pt`, `energy` استخراج می‌شوند و به صورت تنسورهای `torch` ذخیره می‌شوند.
- ویژگی‌های سطح رویداد مانند `nVGoodVertices` و `nVertices` نیز به صورت تنسورهای `torch` ذخیره می‌شوند.
- ویژگی‌های گره‌ها (`node_features`) با استفاده از دستور `torch.stack` انباشته می‌شوند.
- ویژگی‌های گراف (`graph_attr`) نیز با استفاده از دستور `torch.stack` انباشته می‌شوند.
- یال‌ها و ویژگی‌های یال با استفاده از تابع `GetEdgeInfo` محاسبه می‌شوند.
- شیء `tg.data.Data` برای ساخت گراف نهایی استفاده می‌شود که شامل ویژگی‌های گره‌ها، شاخص‌های یال، ویژگی‌های یال و ویژگی‌های گراف است.

```

data = tg.data.Data(
    x=node_features,
    edge_index=edge_index,
    edge_attr=edge_attr,
    graph_attr=graph_attr
)

```

۳. دسته بندی و ظایف

شبکه‌های عصبی گرافی می‌توانند برای سه نوع اصلی از ظایف استفاده شوند: ظایف سطح گراف، سطح گره و سطح پال. این ظایف ممکن است مشمول مسائل دسته‌بندی یا رگرسیون باشند. تفاوت‌های کلیدی بین الگوریتم‌های رگرسیون و دسته‌بندی در یادگیری ماشین عبارتند از نوع خروجی:

- رگرسیون برای مسائلی استفاده می‌شود که هدف پیش‌بینی یک مقدار یا مقدار عددی بر اساس متغیرهای مستقل ورودی است.
- دسته‌بندی برای مسائلی استفاده می‌شود که هدف اختصاص نمونه‌های داده به یکی از چندین دسته یا کلاس پیش‌ تعیین شده بر اساس متغیرهای ورودی است.
- بعضی از تفاوت‌های مهم دیگر شامل موارد زیر است:
 - معیارهای ارزیابی: رگرسیون از معیارهایی مانند خطای میانگین مربعات استفاده می‌کند، در حالی که دسته‌بندی از دقت، دقت، بازیابی و غیره استفاده می‌کند.
 - الگوریتم‌ها: الگوریتم‌های معمول رگرسیون شامل رگرسیون خطی، درخت‌های تصمیم، جنگل‌های تصادفی و غیره است. الگوریتم‌های معمول دسته‌بندی شامل رگرسیون لجستیک، نویس بیز ساده، ماشین‌های بردار پشتیبان و غیره است.
 - انواع داده: رگرسیون با متغیرهای هدف پیوسته سروکار دارد، در حالی که دسته‌بندی با متغیرهای هدف گسسته سروکار دارد.

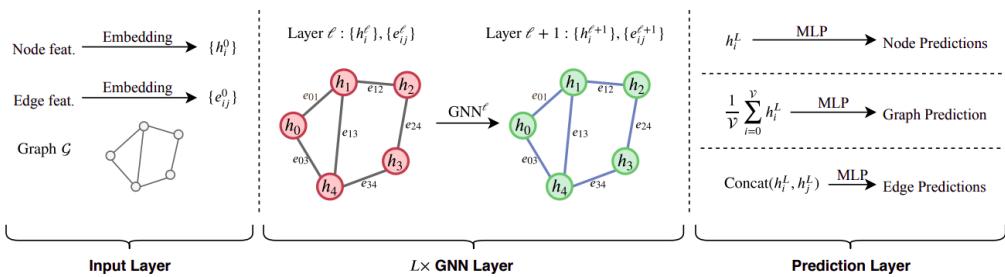
بنابراین، به طور خلاصه، تمایز اساسی این است که رگرسیون یک خروجی کمی را پیش‌بینی می‌کند، در حالی که دسته‌بندی برچسب‌های کласیفیکه یا دسته‌بندی را به عنوان خروجی پیش‌بینی می‌کند.

• وظایف سطح گراف:

- دسته‌بندی^۱: پیش‌بینی برچسب یا دسته‌ای برای کل گراف (برای مثال، دسته‌بندی رویدادها به عنوان یک پایل آپ

[۱۷] مشخص).

^۱Classification



شکل ۱: نمایی از دسته بندی وظایف

- رگرسیون^۱: پیش‌بینی یک مقدار عددی برای کل گراف (برای مثال، پیش‌بینی پایل آپ برای یک رویداد[۱۸]).
- وظایف سطح گره:
- دسته‌بندی: پیش‌بینی کلاس یا برچسب گره‌ها در یک گراف (برای مثال، دسته‌بندی هر ذره که در هر رویداد به پایل آپ مشخصی تعلق دارد[۱۹][۲۰]).
- رگرسیون: پیش‌بینی یک مقدار عددی برای گره‌ها در یک گراف (برای مثال، پیش‌بینی یک عدد برای هر ذره در یک گراف)[۱۹][۲۰].
- وظایف سطح یال:
- دسته‌بندی: پیش‌بینی اینکه آیا یک یال بین دو گره وجود دارد یا خیر (برای مثال، پیش‌بینی نزدیکترین ذره)[۱۷].
- رگرسیون: پیش‌بینی یک مقدار عددی برای یال‌ها در یک گراف (برای مثال، پیش‌بینی قدرت تعاملات بین ذره‌ها).

معماری PUSage

معماری گرافی PUSage است. این معماری با استفاده از لایه‌های گراف کانولوشن نوع SAGE²، نرمال‌سازی دسته‌ای^۳ و مکانیزم Dropout ساخته شده است. هدف از این مدل استخراج ویژگی‌های گره‌ها از گراف‌ها به صورت مؤثر و با کارایی بالا است.

ساختار معماری

مدل شامل سه بخش اصلی است:

- لایه ورودی: این لایه داده‌های اولیه مربوط به گراف شامل ویژگی‌های گره‌ها و لبه‌ها را دریافت می‌کند.
- لایه‌های پنهان: چندین لایه از نوع SAGEConv همراه با نرمال‌سازی و Dropout برای استخراج ویژگی‌های گره‌ها در

¹Regression

²Sample and Aggregate

³Batch Normalization

گراف به کار می‌روند.

- لایه خروجی: لایه خطی (Fully Connected) برای تولید خروجی نهایی از مدل استفاده می‌شود.

فرمول‌های ریاضی

لایه کانولوشن گراف SAGEConv به شکل زیر تعریف می‌شود:

$$h_v^{(k)} = \sigma \left(W^{(k)} \cdot \text{AGGREGATE} \left(h_v^{(k-1)}, \{h_u^{(k-1)} | u \in \mathcal{N}(v)\} \right) \right)$$

که در آن:

- $h_v^{(k)}$ نمایش ویژگی‌های گره v در لایه k است.

- همسایگان گره v است. $\mathcal{N}(v)$

- وزن‌های لایه k است. $W^{(k)}$

- σ تابع فعال‌سازی (مانند ReLU) است.

کد مربوط به معماری PUSage

```

1 import torch
2 import torch.nn.functional as F
3 from torch_geometric.nn import SAGEConv, global_mean_pool
4 from torch_geometric.utils import add_self_loops
5
6 class PUSage(torch.nn.Module):
7     def __init__(self, input_dim, hidden_dim, output_dim, num_layers, dropout_rate):
8         super(PUSage, self).__init__()
9
10        self.convs = torch.nn.ModuleList()
11        self.bns = torch.nn.ModuleList()
12        self.dropouts = torch.nn.ModuleList()
13
14        # Input layer
15        self.convs.append(SAGEConv(input_dim, hidden_dim))
16        self.bns.append(torch.nn.BatchNorm1d(hidden_dim))
17        self.dropouts.append(torch.nn.Dropout(dropout_rate))
18
19        # Hidden layers
20        for _ in range(1, num_layers):
21            self.convs.append(SAGEConv(hidden_dim, hidden_dim))
22            self.bns.append(torch.nn.BatchNorm1d(hidden_dim))
23            self.dropouts.append(torch.nn.Dropout(dropout_rate))
24
25        # Output layer
26        self.linear = torch.nn.Linear(hidden_dim, output_dim)
27
28    def forward(self, data):
29        x, edge_index, edge_attr, y = data['x'], data['edge_index'], data.get('edge_attr'), data['y'].float()
30        edge_index, _ = add_self_loops(edge_index, num_nodes=x.size(0))
31
32        for conv, bn, dropout in zip(self.convs, self.bns, self.dropouts):
33            x = F.relu(bn(conv(x, edge_index)))
34
35        x = global_mean_pool(x, data['batch'])
36
37        x = self.linear(x)
38
39        return x
40
41 # Hyperparameters
42 input_dim = 15
43 hidden_dim = 64
44 output_dim = 1
45 num_layers = 5
46 dropout_rate = 0.5
47

```

شکل ۲: کد PUSage

مراجع

- [1] “CMS Luminosity Measurements for the 2016 Data Taking Period”, 2017.
- [2] Friedman, J. H., “Data Analysis Techniques for High-Energy Particle Physics”, in *3rd CERN School of Computing*, p. 271, 10 1974.
- [3] CERN, “Cern physics”, Accessed on March 22, 2024.
- [4] Thomson, M., *Modern Particle Physics*, Cambridge, UK: Cambridge University Press, 2013.
- [5] CMS Collaboration, “Cms detector”, Accessed: 2024-06-15, Accessed on: 2024-06-15.
- [6] “CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update”, tech. rep., CERN, Geneva, 2013.
- [7] CMS Collaboration, “CMS Guide to Pileup Simulation”, 2011.
- [8] CMS Collaboration, “Mixing Module”, 2024.
- [9] Particle Data Group, “Particle data group”, 2024.
- [10] Particle Data Group, “Review of particle physics: Monte carlo techniques”, 2007.
- [11] contributors, W., “Pseudorapidity”, 2024.
- [12] Adam, W., “Track and vertex reconstruction in CMS”, *Nucl. Instrum. Meth. A*, Vol. 582, pp. 781–784, 2007.
- [13] Shlomi, J., Battaglia, P., and Vlimant, J.-R., “Graph neural networks in particle physics”, *Machine Learning: Science and Technology*, Vol. 2, p. 021001, dec 2020.
- [14] Kipf, T. N. and Welling, M., “Semi-supervised classification with graph convolutional networks”, 2017.
- [15] Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., and Bengio, Y., “Graph attention networks”, 2018.
- [16] Hamilton, W. L., Ying, R., and Leskovec, J., “Inductive representation learning on large graphs”, 2018.
- [17] of Amsterdam Deep Learning Course, U., “Graph neural networks overview”, 2024, Accessed: 2024-06-01.
- [18] Zhu, X., Yan, Z., Xu, Z., Wang, Z., and Jha, N. K., “Unifying graph convolutional neural networks and graph neural tangent kernels: Towards understanding and improving graph representations”, *arXiv preprint arXiv:2307.07840*, 2023.
- [19] Deep Graph Library, “Training a gnn for node classification”, 2023.
- [20] Exchange, A. S., “Are there explainable gnn methods for node regression tasks?”, 2023, Accessed: 2024-06-01.

Pile up distribution estimation with Graph Neural Network (GNN) in CMS experiment for Large Hadron Collider

Mahtab Jalavandi

June, 30, 2024

Department of physics

Isfahan University of Technology, Isfahan 84156-83111, Iran

Degree: Master of Science in physics, Elementary Particle physics and field theory orientation

Language: Farsi

Supervisor: Dr. Hamed Bakhshian sohi, .

Abstract

High-energy physics experiments, like the CMS experiment at the Large Hadron Collider (LHC), face growing challenges in simulating and analyzing the pile-up of simultaneous proton-proton collisions as luminosity levels increase with LHC upgrades. This paper presents a novel approach for pile-up estimation that leverages the power of Graph Neural Networks (GNNs). GNNs have emerged as a powerful tool for representing complex event data, as shown in recent applications for particle reconstruction and pile-up mitigation. We propose a GNN-based method to estimate the number of pile-up interactions within each event. Preliminary results are promising, indicating that the method can be used to study the linearity response of the luminometers, a critical step for independent luminosity measurement. This research explores the transformative potential of GNNs for pile-up estimation and luminosity measurement in high-energy physics experiments. We expect this method to become a valuable addition to existing techniques, contributing to advancements in the field.

Keywords

pile up, luminosity, cross section, collision, Graph neural network, GNN, Graph, poisson distribution



Isfahan University of Technology

Department of physics

Pile up distribution estimation with Graph Neural Network (GNN) in CMS experiment for Large Hadron Collider

A Thesis

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in physics, Elementary Particle
physics and field theory orientation

By

Mahtab Jalavandi

Evaluated and Approved by the Thesis Committee, on June, 30, 2024

- 1- Dr. Hamed Bakhshian sohi (Supervisor)
- 2- Dr. Sedighe Sajadian (Advisor)
- 3- Dr. Abideh Jafari (Examiner)
- 4- Dr. Soroosh Shakeri (Examiner)

Department Graduate Coordinator: Dr. Ismaeil Abdolhosseini Sarsari