

在对撞机上寻找新物理的研究

(申请清华大学理学博士学位论文)

培 养 单 位 : 工程物理系

学 科 : 物理学

研 究 生 : 许 方 舟

指 导 教 师 : 陈 少 敏 教 授

二〇二〇年五月

Search for New Physics at Colliders

Thesis Submitted to

Tsinghua University

in partial fulfillment of the requirement

for the degree of

Doctor of Philosophy

in

Physics

by

Xu Fangzhou

Thesis Supervisor: Professor Chen Shaomin

May, 2020

关于学位论文使用授权的说明

本人完全了解清华大学有关保留、使用学位论文的规定，即：

清华大学拥有在著作权法规定范围内学位论文的使用权，其中包括：（1）已获学位的研究生必须按学校规定提交学位论文，学校可以采用影印、缩印或其他复制手段保存研究生上交的学位论文；（2）为教学和科研目的，学校可以将公开的学位论文作为资料在图书馆、资料室等场所供校内师生阅读，或在校园网上供校内师生浏览部分内容。

本人保证遵守上述规定。

（保密的论文在解密后应遵守此规定）

作者签名： _____

导师签名： _____

日 期： _____

日 期： _____

摘要

Higgs 玻色子在大型强子对撞机 (LHC) 上的发现, 最终确认了标准模型 (SM) 的粒子谱。搜寻超出标准模型的新物理于是成为 LHC 的主要目标。在对撞能量和探测器分辨率均有限制的前提下, 设计算法以准确的重建事例与高效的提取信号始终是高能物理唯象研究的关键。

一种富有希望的新物理便是超对称 (SUSY)。由于 LHC 尚未观测到任何超对称粒子的信号, 超对称最简单自然的版本已被排除, 但以后在对撞机上发现超对称的希望仍在, 这需要对理论更精确细致的分析。

另一反面, 对撞机实验上持续观测到 B 介子衰变反常。能解释这一反常的新模型需包含一定的味道改变中性流, $b \rightarrow sll$ 。其中一种紫外 (UV) 完备模型, 是引入 Z' 玻色子和若干类矢量粒子的 $U(1)_X$ 模型。在对撞机上寻找这些新粒子可以检验此模型。

一般认为新物理的能标至少在 TeV 量级以上, 而在 LHC 上搜寻大质量粒子通常涉及喷注 (jet), 尤其是提速 (boosted) 喷注的重建与识别。使用深度学习方法可以提升识别效率及重建准确度。本论文主要工作如下:

- 在 split SUSY 中, 根据 Higgs 玻色子的质量和电弱真空稳定性, 研究超对称破缺能标的上限和未来对撞机上通过搜寻 gluino 发现超对称的前景。
- 分析了 $U(1)_X$ 模型参数空间, 指出在高亮度 LHC (HL-LHC) 上, 单纯借助双轻子道搜寻 Z' 玻色子可能由于其与夸克较弱的耦合而一无所获。此时可搜寻带电的左手类矢量轻子 E 。两个搜寻道可相互补充: 一是 $E^\pm \rightarrow \mu^\pm Z' (\rightarrow \mu^- \mu^+)$ 产生的 6 μ 子末态, 二是 $E^\pm \rightarrow \mu^\pm H (\rightarrow b\bar{b})$ 产生的双 μ 子加 4 b 夸克末态。HL-LHC 对 1 TeV 类矢量轻子寻找的灵敏度可达 2σ 。
- 考虑 R 宇称不守恒条件下, neutralino 的强子衰变。使用卷积神经网络 (CNN) 对 neutralino 喷注进行子结构分析, 改善 ATLAS 合作组给出 gluino 和 neutralino 质量限制的分析。
- 使用卷积神经网络优化 boosted Higgs 喷注的识别与重建。在单一场景实现了 50% 的探测效率提升。借助提升决策树 (BDT), 信号显著性提高超过 1 倍。而且新算法更不易受本底干扰。在多种复杂场景下, 重建的 Higgs 玻色子 4-动量可具有更高精度。

关键词: 超对称; 类矢量轻子; 喷注; 大型强子对撞机; 卷积神经网络

Abstract

The discovery of the Higgs boson at the Large Hadron Collider (LHC) has finally confirmed the particle spectrum of the Standard Model (SM). And then the searches for new physics beyond the SM have become the primary mission of the LHC. Under the premise that the collision energy and the detector resolution are finite, designing algorithms to reconstruct events accurately and extract signals efficiently is always the key to high energy phenomenology research.

One promising new physics scenario is supersymmetry (SUSY). The null results of supersymmetric particle searches at the LHC excluded the most natural and straightforward version of supersymmetry. But there is still hope at future colliders, which requires a more accurate and careful analysis of the theory.

On the other hand, collider experiments have been observing anomalies in B -meson decays. New models addressing these anomalies need to accommodate specific amounts of flavor-changing neutral currents $b \rightarrow sll$. One ultraviolet (UV) complete model is the $U(1)_X$ model that introduces a Z' boson and several vector-like particles. Searching for these new particles at colliders could test this model.

It is generally recognized that new physics lies at the TeV scale or above. Searching for massive particles at the LHC involves typically reconstructing and identifying jets, especially boosted jets. Deep learning can help boost identification efficiency and reconstruction accuracy. The main works of this thesis are as follows:

- Study the upper bounds on supersymmetry breaking scales in split SUSY from the Higgs mass and electroweak vacuum stability, and analyse the prospect of discovering supersymmetry at future colliders by searching for gluinos.
- By analyzing the parameter space of the $U(1)_X$ model, we point out that it could be unfruitful to search for Z' boson in the dilepton channel at the High Luminosity LHC (HL-LHC) due to its weak couplings to quarks. We can search for the charged vector-like lepton E instead. Two channels are complementary: one is the 6-muon final state produced by $E^\pm \rightarrow \mu^\pm Z' (\rightarrow \mu^- \mu^+)$, the other is the dimuon plus four b -quarks final state produced by $E^\pm \rightarrow \mu^\pm H (\rightarrow b\bar{b})$. The HL-LHC is sensitive to the vector-like lepton of 1 TeV with 2σ significance.
- Study neutralino hadronic decays in R -parity violating supersymmetry. Use a Con-

volutional Neural Network (CNN) to analyse neutralino jet substructure, and improve the analysis by ATLAS Collaboration, which provided the bound of gluino mass and neutralino mass.

- Utilize a CNN to optimize the identification and reconstruction of boosted Higgs jets. The efficiency in a single scenario is improved by 50%. With the help of a boosted decision tree (BDT), the signal significance can be doubled. The algorithm is also less susceptible to background contamination. In various complex scenarios, the 4-momenta of reconstructed Higgs bosons are more accurate.

Key Words: Supersymmetry; Vector-Like Lepton; Jet; LHC; CNN

目 录

| | |
|------------------------------------|----|
| 第 1 章 引言 | 1 |
| 1.1 探索超出标准模型的新物理 | 1 |
| 1.1.1 标准模型 | 2 |
| 1.1.2 超对称 | 4 |
| 1.1.3 B 物理反常 | 6 |
| 1.2 对撞机物理 | 8 |
| 1.2.1 喷注算法及子结构 | 9 |
| 1.2.2 深度学习的运用 | 14 |
| 1.3 本论文组织架构 | 19 |
| 第 2 章 Split SUSY 的超对称破缺能标上限 | 20 |
| 2.1 介绍 | 20 |
| 2.2 sfermion 质量上限 | 21 |
| 2.3 在 LHC 上搜寻 gluino | 24 |
| 2.4 结论 | 26 |
| 第 3 章 类矢量轻子搜寻 | 28 |
| 3.1 $U(1)_X$ 模型 | 28 |
| 3.2 现有实验对参数空间的限制 | 31 |
| 3.3 在 LHC 上寻找类矢量轻子 | 33 |
| 3.3.1 衰变道 | 36 |
| 3.3.2 信号与本底模拟 | 42 |
| 3.3.3 事例重建与选取 | 42 |
| 3.4 结论 | 46 |
| 第 4 章 在 LHC 上搜寻违反 R 宇称的超对称 | 48 |
| 4.1 介绍 | 48 |
| 4.2 CNN 架构 | 49 |
| 4.3 CNN 的训练与测试 | 50 |
| 4.4 应用 CNN 在 LHC 上搜寻 gluino | 55 |
| 4.5 结论 | 61 |

| | |
|-----------------------------|----|
| 第 5 章 Higgs 喷注检测与重建 | 63 |
| 5.1 Mask R-CNN | 63 |
| 5.2 数据集 | 64 |
| 5.2.1 信号标定 | 65 |
| 5.3 算法表现 | 66 |
| 5.3.1 检测效率 | 68 |
| 5.3.2 重建精度 | 71 |
| 5.3.3 区分能力 | 79 |
| 5.4 结论 | 80 |
| 第 6 章 总结 | 83 |
| 参考文献 | 85 |
| 致 谢 | 96 |
| 声 明 | 97 |
| 个人简历、在学期间发表的学术论文与研究成果 | 98 |

主要符号对照表

| | |
|---------------|-----------|
| SM | 标准模型 |
| SUSY | 超对称 |
| MSSM | 最小超对称标准模型 |
| LHC | 大型强子对撞机 |
| QCD | 量子色动力学 |
| LSP | 最轻超对称粒子 |
| RPV | R 宇称破坏 |
| BR | 分支比 |
| CNN | 卷积神经网络 |
| BDT | 提升决策树 |
| H, h | Higgs 玻色子 |
| b | 底夸克 |
| g | 胶子 |
| j | 胶子和轻夸克喷注 |
| ℓ | 轻子 |
| y | 快度 |
| η | 赝快度 |
| ϕ | 方位角 |
| p_T | 横向动量 |
| E_T | 缺失的横向能量 |
| v | 真空期望值 |
| m | 质量 |
| σ | 散射截面 |
| θ | 混合角 |
| Γ | 衰变宽度 |
| ε | 效率 |
| \mathcal{L} | 拉式量 |
| \mathcal{P} | 概率 |

第 1 章 引言

1.1 探索超出标准模型的新物理

粒子物理标准模型 (Standard Model, SM) 无疑是目前描述物质基本相互作用最成功的理论。已观测到的基本粒子, 全部归属于同一规范群的不同表示, 其相互作用形式和强度则由规范对称性和 19 个实参数完全确定。数十年来, 标准模型的预言经受住了无数实验检验, 各个自由参数的测量也愈发精确^[1]。2012 年 Higgs 玻色子的发现^[2-3] 更是为标准模型拼上了最后一块拼图。然而, 即使不考虑与引力相关的问题, 仍有许多实验现象无法纳入标准模型的框架, 比如天文学和宇宙学观测中证据确凿的暗物质^[4], 宇宙中正反物质的不对称性^[5], 强作用电荷共轭-宇称对称性问题 (strong *CP* problem)^[6], 和中微子静质量非零导致的中微子振荡等^[7-8]。这些现象表明标准模型只是一个低能有效理论, 在更高能标 ($10^4 - 10^{19}$ GeV) 应存在一个更加完备的理论。一方面, 理论模型需要扩展以解释这些未解之谜; 另一方面, 物理学家们希望在精确测量实验中能够看到与标准模型预言相违背的结果, 为筛选新模型提供线索。遗憾的是, 目前仍没有十分可靠的迹象直接指明新物理的方向。扩展标准模型既有理论动机主导的大统一理论 (GUT), 包括 SU(5), SO(10), E_6 模型等^[9]; 超对称 (SUSY), 包括最小超对称标准模型 (MSSM)^[10], 次最小超对称标准模型 (NMSSM)^[11] 等; 额外维, 如 Kaluza-Klein 理论^[12], Randall-Sundrum 模型^[13-14], 弦论^[15-16] 等。也有唯象模型, 如双 Higgs 二重态模型 (2HDM)^[17], 左右对称模型^[18], 预言了轴子的 Peccei-Quinn 理论^[19], 以及其他各类提供暗物质候选者^[20] 或解释中微子质量产生机制^[21] 的模型等。上述模型预言的质子衰变, 轴子与电磁场相互作用, 暗物质与原子核碰撞等在相应实验, 如日本的超级神冈探测器 (Super-Kamiokande)^[22], 美国的轴子暗物质试验 (ADMX)^[23], 清华大学主导的中国暗物质实验 (CDEX)^[24] 和上海交通大学主导的粒子和天体物理氦探测器 (PandaX)^[25] 等目前都呈现零结果。而对撞机作为验证粒子物理理论最有说服力的武器之一, 也未收集到对标准模型预言有明显超出的数据。在大型强子对撞机 (LHC) 开启之前, 高能物理学界就普遍认为 Higgs 玻色子的发现几乎是必然的, 人们本来寄望于超对称的发现, 然而除了将超对称粒子的质量下限一再推高之外, LHC 在这方面可以说一无所获^[26]。物理学家只好更加仔细分析海量数据, 从中寻找新物理的蛛丝马迹。如 2015 年 LHC 观测到的双光子超出^[27-28], 一度吸引了广大粒子物理学家的注意, 可惜后来随着数据量的累积被证明只是统计涨落^[29-30]。自 2013 年以来, *B* 物理反常^[31] 也逐渐成为一个热门领域, 因为这是目前超出标准模

型新物理最显著的信号之一。

本论文将从对撞机直接探测的角度，试图给新物理提供一些理论限制，搜寻方向和实用算法。内容包括超对称破缺能标的限制和未来的探测前景，类矢量轻子的质量限制与搜索策略，使用神经网络帮助提高对撞机对新物理的灵敏度等。下面首先介绍本论文的理论 and 实验基础，包括标准模型和超对称模型的框架，实验观测到的 B 物理反常，和对撞机唯象学的研究方法。

1.1.1 标准模型

粒子物理标准模型由描述电弱相互作用的 Glashow-Weinberg-Salam 模型和描述强相互作用的量子色动力学 (quantum chromodynamics, QCD) 整合而成^[32]。由于新物理模型普遍在一定条件下，如较低能标，退化为标准模型，且新粒子的发现也是通过其与标准模型粒子的耦合来设计探测方案，故深入理解和精确测量标准模型对于探索新物理至关重要。标准模型是定义在 $SU(3)_C \times SU(2)_L \times U(1)_Y$ 规范群上的规范场论。其中下标 C, Y 分别代表色荷，超荷， L 表示 $SU(2)$ 只作用于左手场。标准模型粒子在规范群上的表示如表1.1所示。第二列和第三列为群表示的维度，最后一列为超荷。

表 1.1 标准模型粒子的表示

| 场 | $SU(3)_C$ 表示 | $SU(2)_L$ 表示 | $U(1)_Y$ 荷 |
|-------|--------------|--------------|----------------|
| q | 3 | 2 | $\frac{1}{6}$ |
| u_R | 3 | 1 | $\frac{2}{3}$ |
| d_R | 3 | 1 | $-\frac{1}{3}$ |
| l | 1 | 2 | $-\frac{1}{2}$ |
| e_R | 1 | 1 | -1 |
| H | 1 | 2 | $\frac{1}{2}$ |
| g | 8 | 1 | 0 |
| W | 1 | 3 | 0 |
| B | 1 | 1 | 0 |

表中，三代费米子由外尔 (Weyl) 旋量场构成。其中，左手夸克 q 与轻子 l 二重态分别为，

$$q = \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \quad l = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}, \quad (1-1)$$

而 u_R, d_R 和 e_R 为右手费米子。传递强相互作用的胶子，传递弱相互作用的 W, Z 玻色子，和传递电磁相互作用的光子由矢量场构成，它们的规范本征态 g, W 和

B 处于规范群的伴随表示。唯一的标量场 Higgs 场 H 是一个复标量二重态，其中获得了真空期望值的分量构成 Higgs 玻色子。

根据规范对称性，Lorentz 对称性和可重整性的要求，标准模型拉式量具有以下形式：

1) 规范作用项，分别由 $SU(3)_C$ ， $SU(2)_L$ ， $U(1)_Y$ 规范场的场强张量构成。

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4}G_{\mu\nu}^a G^{a\mu\nu} - \frac{1}{4}W_{\mu\nu}^a W^{a\mu\nu} - \frac{1}{4}B_{\mu\nu} B^{\mu\nu} \quad (1-2)$$

2) 费米子动能项，由费米子场及其协变微分构成。

$$\mathcal{L}_{\text{fermion}} = i\bar{q}\not{D}q + i\bar{u}_R\not{D}u_R + i\bar{d}_R\not{D}d_R + i\bar{l}\not{D}l + i\bar{e}_R\not{D}e_R + \text{h.c.} \quad (1-3)$$

其中 h.c. 代表厄米共轭。

3) Yukawa 项，由费米子场及 Higgs 场构成，这里 i, j 指标表示不同代费米子。

$$\mathcal{L}_{\text{Yukawa}} = -y_{ij}^u \bar{q}^i \tilde{H} u_R^j - y_{ij}^d \bar{q}^i H d_R^j - y_{ij}^e \bar{l}^i H e_R^j + \text{h.c.} \quad (1-4)$$

其中 $\tilde{H} = i\sigma_2 H^*$ 。

4) Higgs 场的动能及势能项。

$$\mathcal{L}_{\text{Higgs}} = D^\mu H^\dagger D_\mu H + m^2 H^\dagger H - \lambda(H^\dagger H)^2 \quad (1-5)$$

其中 m^2 和 λ 分别表示 Higgs 场的二次和四次耦合系数。

电弱能标下，Higgs 场获得真空期望值 v ，

$$\langle H \rangle = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{m}{\sqrt{2\lambda}} \end{pmatrix} \quad (1-6)$$

以及质量 $m_H = \sqrt{2}m$ 。

费米子及规范玻色子通过与 Higgs 玻色子耦合而获得质量，其中费米子质量由 Yukawa 耦合大小决定。后面我们将要研究的新粒子也通过相同机制获得质量。Higgs 真空期望值使 $SU(2)_L \times U(1)_Y$ 对称性自发破缺至 $U(1)_{\text{EM}}$ ，即电磁场的规范对称性。 $SU(2)_L$ 规范场中带电的部分构成了 W 玻色子，而不带电部分和 $U(1)_Y$ 的规范场重新组合，有质量的本征态成为 Z 玻色子，无质量的本征态即我们熟知的光子场 A ，对应于没有破缺的 $U(1)_{\text{EM}}$ 对称性。用 g_L ， g_Y 分别表示 $SU(2)_L$ ， $U(1)_Y$ 的耦合系数， e 表示电磁耦合系数， θ_w 表示 Weinberg 角，则有

$$\tan \theta_w = \frac{g_Y}{g_L}, \quad e = g_L \sin \theta_w, \quad m_W = \frac{1}{2}g_L v, \quad m_Z = \frac{m_W}{\cos \theta_w} \quad (1-7)$$

$$\begin{pmatrix} W_\mu^+ \\ W_\mu^- \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{i}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} W_\mu^1 \\ W_\mu^2 \end{pmatrix}, \quad \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos \theta_w & -\sin \theta_w \\ \sin \theta_w & \cos \theta_w \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix} \quad (1-8)$$

由于 Yukawa 耦合系数是 3×3 的矩阵，将三代夸克和三代轻子质量对角化的转换矩阵分别构成了 Cabibbo-Kobayashi-Maskawa (CKM) 矩阵^[33-34] 与 Pontecorvo-Maki-Nakagawa-Sakata (PMNS) 矩阵^[35-36]，它们分别出现在夸克和轻子与 W 玻色子的弱相互作用项中，

$$\mathcal{L} \supset -\frac{g_L}{\sqrt{2}} V_{ij}^{\text{CKM}} \left(\bar{u}_L^i W^+ d_L^j + \bar{d}_L^i W^- u_L^j \right) - \frac{g_L}{\sqrt{2}} U_{ij}^{\text{PMNS}} \left(\bar{\nu}_L^i W^+ e_L^j + \bar{e}_L^i W^- \nu_L^j \right) \quad (1-9)$$

这里费米子场表示其质量本征态，区别于前面的味本征态。

Yukawa 耦合影响了费米子参与弱作用的形式，且使得不同代的夸克和轻子质量差异极大，标准模型的物理因此变得复杂多样，味物理学也成为粒子物理的主要分支之一，研究方向包括介子的混合，CKM 矩阵中电荷共轭-宇称 (CP) 相角产生的 CP 对称性破坏等。其中 B 物理即是专门研究 B 强子（至少包含一个 bottom 夸克的强子）的领域。

1.1.2 超对称

超对称 (supersymmetry, SUSY) 是一种时空对称性，众所周知，超对称变换下，玻色子变为费米子，反之亦然。加入超对称以后，描述狭义相对论时空对称性的庞加莱代数被扩展，成为超庞加莱代数，额外引进的生成元代有 $1/2$ 自旋。高能唯象学研究，特别是寻找新粒子时，仅关注只有一组超对称生成元的理论，即 $\mathcal{N} = 1$ 的超对称。将全局的超对称性（不考虑引力）引入标准模型，进行最小程度的扩展，我们就得到了最小超对称标准模型 (MSSM)。相比标准模型，MSSM 有几项被广泛认同的优势^[10]：

1) 由于费米子和玻色子之间存在对称性，它们对 Higgs 玻色子质量的二次发散项贡献被抵消了。

2) 与标准模型不同， $SU(3)_C$ ， $SU(2)_L$ ， $U(1)_Y$ 的规范耦合常数，在 $SU(5)$ 大统一理论的假设下^[37]，可以在高能标以很高的精度统一。

3) 如果 R 宇称守恒，最轻的超对称粒子 (lightest supersymmetric particle, LSP) 自动成为暗物质候选者。

MSSM 的粒子谱，相比 SM 有如下扩展：费米子的超对称伙伴， 0 自旋，称为 sfermion，包括 squark 和 slepton。规范玻色子的超对称伙伴， $1/2$ 自旋，称为 gaugino，包括 gluino, wino, 和 bino。Higgs 玻色子的超对称伙伴， $1/2$ 自旋，称

为 Higgsino。

MSSM 中有两个 Higgs 二重态，分别是 H_u 和 H_d ，因为反常消除要求作为费米子的 Higgsino 必须成对出现。另一方面，由于拉式量中的超势（superpotential）是全纯函数，上型（up-type）和下型（down-type）夸克（或带电轻子）的质量必须来自与不同 Higgs 场的 Yukawa 耦合。MSSM 电弱破缺的模式也与 SM 略有不同。两个 Higgs 二重态同时获得真空期望值，其比值是唯象研究中一个重要参数。

$$\tan \beta \equiv \frac{v_u}{v_d} \quad (1-10)$$

粗略估计 $\tan \beta$ 的值在 1.2 至 65 之间^[10]。 H_u 与 H_d 的混合角则通常用 α 表示。

上面所述均为规范本征态，唯象上更有意义的是质量本征态。中性的 Higgsino，中性的 wino，以及 bino 混合形成 4 个 neutralino。带电的 Higgsino 则与带电的 wino 混合形成 chargino。而 sfermion 中，只有第三代由于较大的 Yukawa 耦合和软破缺项耦合，使其左右手场之间具有较大的混合角。前两代 sfermion 的混合可忽略，gluino 则不存在混合问题。

由于并没有观察到与标准模型粒子质量相同的超对称伙伴，显然超对称在低能区是破缺的。尽管破缺机制尚不明确，但我们仍然可以通过已知的参数对超对称破缺能标进行一些估计。在 MSSM 中，超对称破缺由拉式量中的软破缺项来描述，其中包括 gaugino 质量项，sfermion 质量项，与 Yukawa 耦合一一对应的 Higgs-sfermion-sfermion 耦合项以及 Higgs 部分的二次耦合项。

为了避免可重整阶的快速质子衰变，MSSM 的要素之一是 R 宇称守恒。 R 宇称定义为：

$$R = (-1)^{3(B-L)+2S} \quad (1-11)$$

其中 B ， L ，和 S 分别表示粒子的重子数，轻子数和自旋。由于 SM 粒子 $R = +1$ ，而超对称粒子 $R = -1$ ，所以 R 宇称守恒直接导致：

- 1) LSP 绝对稳定，这是其作为暗物质候选的前提。
- 2) 对撞机上的超对称粒子必须同时产生偶数个，一般是成对产生。
- 3) 超对称粒子经级联衰变后的最终产物中至少有一个 LSP，这在对撞机上表现为缺失的横向能量 \cancel{E}_T 。

R 宇称守恒情况下，MSSM 只能具有如下超势：

$$W = y_u Q H_u U^c + y_d Q H_d D^c + y_e L H_d E^c + \mu H_u H_d \quad (1-12)$$

上述超势中隐去了代指标。放宽这一要求，只保留规范不变性和可重整性，则超

势中会出现违反轻子数和重子数守恒的项^[38],

$$W_{\Delta L=1} = \frac{1}{2}\lambda_{ijk}L_iL_jE_k^c + \lambda'_{ijk}L_iQ_jD_k^c + \mu_iL_iH_u \quad (1-13)$$

$$W_{\Delta B=1} = \frac{1}{2}\lambda''_{ijk}U_i^cD_j^cD_k^c \quad (1-14)$$

其中 Q, L, U^c, D^c, E^c, H_u , 和 H_d 为手征超场, 分别代表左手夸克二重态, 左手轻子二重态, 右手上型夸克, 右手下型夸克, 右手轻子, 上型 Higgs, 和下型 Higgs。根据 $SU(2)_L$ 和 $SU(3)_C$ 规范不变性可知, λ_{ijk} 关于指标 ij 反对称, λ''_{ijk} 关于指标 jk 反对称。为了规避质子衰变, λ 和 λ'' 中, 至少有一个的数值非常小。当 R 宇称不再守恒, LSP 也不再稳定。通常假定 LSP 是一个 neutralino, 那么它将通过上面的算符衰变成三个标准模型粒子。后面我们将研究 R 宇称破坏 (RPV) 情况下 neutralino 的衰变信号。

由于 LHC 在多年运行以后并未发现任何 SUSY 的踪迹, 相关的参数空间已经受到了极大限制, 人们不得不放弃 SUSY 模型中最简单的版本, 转而考虑其他可能。比如, 放弃理论自然性要求的 split SUSY^[39-41]。在这种情形下, 一个很轻的 Higgs 玻色子需要的精细调节是可以接受的, sfermion 的质量极重, 而 gaugino 和 Higgsino 质量仍然在 TeV 量级左右。由于 sparticle 的质量谱包含很重的标量波色子和相对很轻的费米子, 它被叫作 split SUSY。Higgs 玻色子的精细调节之所以被接受, 是因为它被放在和宇宙学常数的精细调节问题相同的地位上, 即认为在某种机制下, 低能标中对紫外敏感的参数显得是被精细调节过的^[40]。通过提高 sfermion 的质量, 额外的 CP 破坏效应, 过多的味道改变过程, 和过快的质子衰变等都被压低。另外, 由 R 对称性 (或称手征对称性) 保护, gaugino 和 Higgsino 等费米子质量仍可以保持在 TeV 量级左右。文献 [39-41] 表明, 超对称粒子谱在包含极重的标量玻色子和 TeV 能标的费米子时, 规范耦合系数的统一依然可以实现。另外最轻的 neutralino 仍然可以作为暗物质候选。于是, SUSY 两个重要的优点都得以保留。后面, 我们会通过已知的实验参数, 在 split SUSY 的情形下, 根据模型参数重整化群能标的跑动, 探究 SUSY 破缺的能标限制, 以及未来对撞机上搜寻相关信号的前景。

1.1.3 B 物理反常

B 物理研究 B 强子, 事实上主要是 B 介子的衰变, 振荡 (或混合) 等。由于过程涉及 b 夸克味道的转变, 故而提供了 CKM 矩阵第三行的信息, 对于检验 CKM 的么正性至关重要。为专门研究 B 物理, 粒子物理学家设计建造了能大量产生和探测 B 介子的 B 工厂, 包括日本的 Belle 实验和美国的 BaBar 实验。此外, LHC 上的

LHCb 探测器也专做 B 强子探测。2013 年, LHCb 合作组在 $B \rightarrow K^*(\rightarrow K\pi)\mu^+\mu^-$ 这一罕见衰变的产物角度分布中观测到了与标准模型预言的差异^[31], 从此 B 反常开始在高能界引起持续关注。图1.1是标准模型中贡献这一过程的简明费曼图。

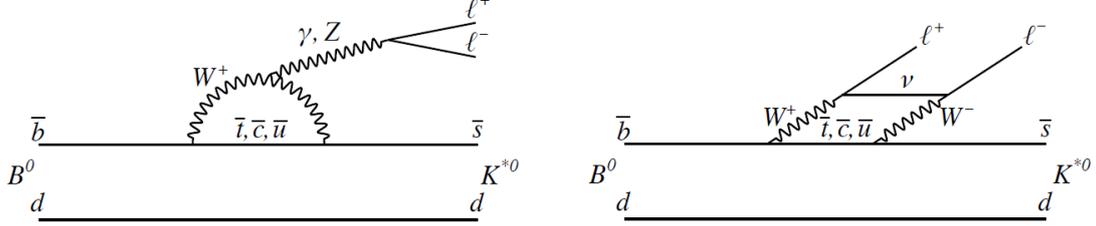


图 1.1 标准模型中 $B \rightarrow K^*ll$ 费曼图, 主要是两个一圈图。

得益于较小的强子效应不确定性, B 介子半轻子罕见衰变被认为是探索超出标准模型新物理的有力工具。过去几年, Belle, BaBar, LHCb 都报告了 B 介子和 B 重子的半轻子罕见衰变分支比^[42-44] 和产物角度分布^[45-46] 与标准模型有一致的偏离, 这些过程都是由 $b \rightarrow sll$ 诱导的。考虑到轻子味道普适性, B 介子衰变到不同轻子的分支比应该大体相同, 它们的比值是非常干净的可观测量,

$$R_{K^{(*)}} \equiv \frac{\text{BR}(B \rightarrow K^{(*)}\mu^+\mu^-)}{\text{BR}(B \rightarrow K^{(*)}e^+e^-)} \quad (1-15)$$

由于强子形状因子和潜在的系统误差在很大程度上抵消, $R_{K^{(*)}}$ 具有极小的理论和实验不确定度。然而目前实验数据显示它们明显小于标准模型的计算值。表1.2列出了在特定双轻子不变质量区间内的理论与测量值。

表 1.2 $R_{K^{(*)}}$ 的标准模型理论计算值与实验测量值

| 可观测量 | 理论计算 ^[47] | 实验测量 ^[48-49] |
|---|----------------------|-------------------------------------|
| $R_K(q^2 = [1, 6] \text{ GeV}^2)$ | 1.00 ± 0.01 | $0.745^{+0.090}_{-0.074} \pm 0.036$ |
| $R_{K^*}(q^2 = [0.045, 1.1] \text{ GeV}^2)$ | 0.92 ± 0.02 | $0.660^{+0.110}_{-0.070} \pm 0.024$ |
| $R_{K^*}(q^2 = [1.1, 6] \text{ GeV}^2)$ | 1.00 ± 0.01 | $0.685^{+0.113}_{-0.069} \pm 0.047$ |

根据标准模型, 分支比差异只来源于 μ 子, 电子的质量差异, 这在过程能标较低时更为明显。 R_K, R_{K^*} 测量值超出标准模型预言的显著性分别只有 $2.6\sigma, 2.5\sigma$ 。若假设 R_K, R_{K^*} 相关, 则显著性提升至 4σ 。若假设所有 B 反常源于同一因素, 则可达 5σ ^[50]。

$b \rightarrow sll$ 是典型的味道改变中性流 (FCNC) 过程。标准模型在树图阶不存在 FCNC, 且由于 Glashow-Iliopoulos-Maiani (GIM)^[51] 机制, 即使在圈图阶也被高度压低。新物理能标远高于 B 反常过程, 对于实验数据, 通常采用有效算符进行

模型无关的分析。 $b \rightarrow s$ 转换按照惯例由以下有效哈密顿量描述^[52],

$$\mathcal{H}_{\text{有效}} = -\frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \frac{e^2}{16\pi^2} \sum_i^{10} (C_i \mathcal{O}_i + C'_i \mathcal{O}'_i) + \text{h.c.} \quad (1-16)$$

$\mathcal{O}_i^{(\prime)}$ 为 5 维或 6 维算符, $C_i^{(\prime)}$ 称作 Wilson 系数。6 维算符 $\mathcal{O}_{1-6}^{(\prime)}$ 描述 4 夸克相互作用, 5 维算符 $\mathcal{O}_7^{(\prime)}$, $\mathcal{O}_8^{(\prime)}$ 分别描述夸克-光子和夸克-胶子相互作用, 只有 6 维算符 $\mathcal{O}_9^{(\prime)}$, $\mathcal{O}_{10}^{(\prime)}$ 描述 $b \rightarrow s \ell \ell$ 所属的夸克-轻子相互作用,

$$\mathcal{O}_9 = (\bar{s} \gamma_\mu P_L b) (\bar{\ell} \gamma^\mu \ell), \quad \mathcal{O}_{10} = (\bar{s} \gamma_\mu P_L b) (\bar{\ell} \gamma^\mu \gamma_5 \ell) \quad (1-17)$$

$$\mathcal{O}'_9 = (\bar{s} \gamma_\mu P_R b) (\bar{\ell} \gamma^\mu \ell), \quad \mathcal{O}'_{10} = (\bar{s} \gamma_\mu P_R b) (\bar{\ell} \gamma^\mu \gamma_5 \ell) \quad (1-18)$$

有数个小组对 B 反常相关实验做了全局拟合, 结论大体一致^[53-59]。有三种情形具有较高显著性, 下面列出的 Wilson 系数已减去标准模型的贡献。

$$C_9^\mu \simeq -1.1 \quad (1-19)$$

$$C_9^\mu = -C_{10}^\mu \simeq -0.6 \quad (1-20)$$

$$C_9^\mu = -C_9^{\prime\mu} \simeq -1.0 \quad (1-21)$$

上标 μ 代表新物理来自与 μ 子的耦合。情形一对应夸克左手流与轻子矢量流耦合, 情形二对应夸克左手流与轻子左手流耦合, 情形三对应夸克轴矢流与轻子矢量流耦合。考虑到标准模型是一个手性规范场论, 情形二尤其值得关注。本论文所分析的模型正是基于此拟合结果。

目前处理 B 反常的模型主要有 Z' 模型^[60-65], 标量 leptoquark 模型^[66], 矢量 leptoquark 模型^[67-70], 复合 Higgs 模型^[71-73] 等。本论文将讨论的模型属于第一类。相对其它模型而言, Z' 模型结构简单, 参数空间较小, 易于在大型对撞机上进行直接验证。最简单的 Z' 模型直接假设 Z' 玻色子与费米子的耦合可改变夸克味道且违背轻子普适性。有别于此, 我们将考虑一个紫外完备的模型, 在 Z' 玻色子之外, 其预言了更易于在对撞机上搜寻的新类矢量粒子。

1.2 对撞机物理

对撞机实验对粒子物理研究的许多方面影响至关重要。对撞机唯象学是连接基础理论与实验的一座桥梁。粒子物理最基本的可观测量是散射截面, 其只具有概率上的意义, 因而需要大量的粒子碰撞数据。而新物理对应的截面小, 事例少, 往往淹没在本底事例中, 需要设计算法筛选事例, 压低本底, 提取信号。随着 LHC 投入使用, 大批事例重建和标记的算法在对撞机物理领域涌现并逐步应用于实验。

强子对撞机中，入射粒子是质子中的夸克或胶子，其质心在实验室参考系往往有沿 z 轴的动量，所以描述出射粒子通常使用沿 z 轴 Lorentz 提速 (boost) 不变的量，包括粒子的方位角 ϕ ，横向动量 p_T 。不同粒子快度 y 之差也是提速不变的。另外，出射粒子几乎都是极端相对论性的，质量可忽略，有时用赝快度 η 代替快度。两者定义在粒子质量为 0 时相同，而赝快度直接与极角 θ 相关。

$$y \equiv \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right), \quad \eta \equiv \frac{1}{2} \ln \left(\frac{1 + \cos \theta}{1 - \cos \theta} \right) \quad (1-22)$$

另外，定义提速不变的粒子间距如下。

$$\Delta R \equiv \sqrt{(\Delta\phi)^2 + (\Delta y)^2} \quad (1-23)$$

质子-质子散射，除了产生我们感兴趣的部分子硬相互作用，比如 $gg \rightarrow H$ ，还有作为本底噪声的潜在事例 (underlying event)，即未参与硬散射的剩余部分子，以及连环相撞 (pile-up)，即同一束流中其他质子-质子对撞等软过程。计算硬过程截面，既要结合质子的部分子分布函数 (PDF)，也要考虑初态辐射，末态辐射等微扰效应和强子化等非微扰效应。尽管总 (inclusive) 截面的理论计算更容易处理，但对撞机上实际测量的通常是部分 (exclusive) 截面，即对末态施加一定限制条件后的截面。上述特性使得对撞机唯象研究高度依赖蒙特卡洛产生子进行截面计算和事例筛选。在模拟中，使用的是半经典的部分子簇射 (parton shower) 和强子化模型。

对蒙特卡洛模拟产生的数据需要设计一套处理方案进行事例重建和筛选，在尽可能保留信号事例的同时压低本底事例，最终使预期的信号显著性最大化，以此评估对撞机对新物理搜寻的灵敏度。重建对象包括夸克，胶子等部分子和 W ， Z ，Higgs 玻色子等半稳态粒子。简单的筛选判据包括末态是否含有轻子，喷注，缺失的横向能量 \cancel{E}_T ，以及这些对象的动量，质量是否超过一定阈值等。完整的筛选策略往往会综合使用各种数据分析手段，如近年来流行的机器学习技术。准确的重建算法和高效的筛选策略正是对撞机实验物理研究的重点。

1.2.1 喷注算法及子结构

在对撞机实验中，当能量很高的时候，对撞产生的一些粒子在探测器空间中会沿几乎相同的某一个方向，在一定动量范围内射出，形成喷注现象^[74-76]。它们是在 LHC 产生的最频繁，并且是在 ATLAS 和 CMS 实验测量中最复杂的一类对象。引发喷注的部分子是带有色荷的基本粒子，即夸克或胶子。一个高能的部分子会辐射出更多部分子，从而产生一束几乎准直的由夸克和胶子构成的喷射物，这个

过程就是部分子簇射。这些带有色荷的部分子经强子化转变为介子和重子，其中大部分为 π 介子， K 介子，质子和中子等，不稳定的介子进一步衰变，最终转变为探测器中观测到的光子，带电的轻子和强子，中性强子以及不可见的中微子。

顶夸克 top 的情况略有不同，电弱能标远高于 QCD 能标，因而 top 夸克衰变的时间尺度远小于部分子簇射和强子化，它会首先衰变为一个 W 玻色子和一个 bottom 夸克。如果初始的 top 夸克能量很高且 W 玻色子经历强子衰变，则最终会形成一个 top 喷注 (top-jet)。否则只会观测到一个 bottom 夸克形成的 b 喷注 ($b\text{-jet}$)。

喷注必须根据实验观测到的粒子 4-动量定义，且不唯一。喷注的定义依赖于特定的喷注算法。最早的喷注算法出现在正负电子对撞的研究中^[77]。一般要求算法满足红外和共线 ($\text{infrared and collinear, IRC}$) 可靠性^[78]。IRC 可靠是指，如果在原有喷注基础上，一个粒子辐射出一个能量很低的粒子，或劈裂成共线的两个或多个粒子，将不会改变喷注算法的结果。这源于理论和实验两反面的考虑。理论方面，在微扰 QCD 的计算中，涉及软辐射和共线劈裂的树图是发散的，需要计入相应的圈图以抵消。对于不满足 IRC 可靠性的喷注算法，两个费曼图会贡献到截然不同的喷注，导致无限大的散射截面。实验方面，探测器的能量和空间分辨率有限，并不能捕捉到极软的辐射，高度共线的粒子也表现得像单个粒子一样。

绝大多数喷注算法基于迭代的聚类过程。首先定义一个度量，据此计算所有粒子对的间距 d_{ij} ，然后将两个间距最小的粒子合并，重复这一过程直到所有粒子的间距都大于某个阈值或喷注与质子束 (beam) 的间距 d_{iB} 。最终，全部粒子会组成若干喷注。常用的 3 种喷注算法使用以下度量，其中 R 称作喷注的圆锥尺寸

表 1.3 常用喷注算法的度量

| 喷注算法 | d_{ij} | d_{iB} | α |
|--|--------------------------------------|-----------------|----------|
| Cambridge/Aachen (C/A) ^[79] | | | 0 |
| k_T ^[80] | $\min(p_{Ti}^\alpha, p_{Tj}^\alpha)$ | p_{Ti}^α | 2 |
| anti- k_T ^[81] | $\frac{\Delta R_{ij}^2}{R^2}$ | | -2 |

(cone size) 或半径。

与高能 top 夸克的情况类似，一些不稳定粒子，比如 W ， Z 玻色子， Higgs 玻色子和某些超出标准模型的假想粒子，尽管它们是色单态，但具有一定强子衰变的分支比。如果能量很高，其衰变产物也会形成一束喷射物，最终转变为一个张角较大的喷注，通常称为胖喷注。胖喷注其实是由两束或两束以上的喷注组合而成，但是由于它们方向相近，可能不易分离开来。人们可以利用喷注的内部结构将不稳定色单态粒子衰变形成的喷注与 QCD 喷注区分开来，亦可以在一定程度上区分轻夸克和胶子诱导的喷注。探索如何利用喷注内部信息对其进行分类的这一方向

被称作喷注子结构研究。

喷注是一种复合的对象，除了其本身结构的复杂性，强子对撞机上大量存在的 **pile-up** 产生的辐射与喷注中能量较低的粒子混杂在一起，也增加了喷注识别这一问题的不确定因素。尽管平均来看，**pile-up** 的能量比硬过程的对撞能量低得多，但是由于数量巨大，累积在一起，它会显著改变根据喷注重建的动力学变量，包括质量，横向动量等。**pile-up** 减轻 (**mitigation**) 于是也成为喷注研究中重要的一环。

自从 2008 年，文献 [82] 提出，可以利用子结构信息对 W , Z 玻色子关联产生的 **boosted Higgs** 玻色子进行重建，喷注子结构的研究开始引起广泛的兴趣并涌现了大批算法。子结构研究的兴起，首要原因自然是 LHC 的启用。由于对撞能量远高于此前能量最高的加速器 Tevatron，不可避免地，事例重建涉及大量 W , Z , Higgs 玻色子衰变产生的胖喷注。另一方面，ATLAS, CMS 的量能器具有更高的空间分辨率，其中电磁量能器的角分辨率最高可达 1 度，而强子量能器的角分辨率最高可达 5 度^[83]。于是我们可以得到单一喷注内部更准确的能量分布，使得子结构分析成为可能。

最早一批对子结构信息的研究，专注于标准模型过程，研究者们构造了各种 **boosted top** 夸克， W , Z , Higgs 玻色子的标记器 (**tagger**)^[84-90]。但事实上，喷注子结构分析更大的用武之地是在超出标准模型新物理的搜寻。在众多标准模型的扩展中，都预言了 TeV 量级的粒子，新粒子衰变到电弱能标的标准模型粒子，由于巨大的质量差，这些粒子具有极大的动量。尽管标准模型过程也能产生上述 **boosted** 粒子，但其产生截面很低，在 TeV 能标，大致按照 $1/s$ 的尺度下降，其中 s 是质心系能量的平方。相对而言，通过新粒子的共振衰变，**boosted** 粒子的产生截面可以有显著的增强，从而在不变质量分布的尾端提供新物理的信号。又由于电弱能标与 QCD 能标的巨大差别，它们经强子衰变形成的喷注就会具有鲜明的子结构。一般而言，具有类似下面这样的等级结构的理论都可以利用喷注子结构分析。

$$\Lambda_{\text{NP}} \gg \Lambda_{\text{EW}} \gg \Lambda_{\text{QCD}} \quad (1-24)$$

其中若没有第一个等级，则电弱能标粒子动量较小，其衰变产物将形成若干喷注，而非一个胖喷注。若没有第二个等级，则子喷注之间距离过近，相互重叠，难以分辨。式 (1-24) 中的 Λ_{EW} 也可以代表新粒子，比如 **gluino** 衰变产生的 **neutralino**，它可以通过破坏 R 宇称的算符衰变为 3 个夸克。这时子结构分析的目标就是把 **neutralino** 喷注与 QCD 喷注区分开来。

一般而言，构造的子结构变量主要有以下两种类型：

1) 喷注形状 (**jet shape**) 变量，仅考虑喷注内能量的空间分布，不考虑喷注的

重建历史，表征了一定的辐射模式。喷注质量也属于其中之一。有代表性的还包括：

a) 狭义的喷注形状^[91]，定义为：

$$\Psi(r) = \frac{\sum_{r_i < r} p_{Ti}}{\sum_{i \in J} p_{Ti}} \quad (1-25)$$

即喷注圆锥半径 r 内的横向动量标量和占喷注总横向动量标量和的比值。根据定义，喷注形状随 r 单调递增， $\Psi(0) = 0$ ， $\Psi(R_0) = 1$ 。等价地，有微分喷注形状， $\rho(r) = \frac{d\Psi(r)}{dr}$

b) 角度特征 (angularity)^[92-94]，定义为：

$$\lambda_\beta^\kappa = \sum_{i \in J} \left(\frac{p_{Ti}}{\sum_{i \in J} p_{Ti}} \right)^\kappa \left(\frac{\Delta R_i}{R_0} \right)^\beta \quad (1-26)$$

其中 $\kappa = 1$ ， $\beta = 1$ 时，退化为喷注围长 (girth)^[95]。 $\kappa = 1$ ， $\beta = 2$ 时，近似为喷注“推力” (thrust)^[96]。

2) 分叉结构 (prong) 变量，用于重建胖喷注内部较硬的子喷注，部分包含了重建的历史信息。对 boosted W ， Z ，Higgs 玻色子的识别常使用 2-分叉变量，而对 boosted top 夸克识别则使用 3-分叉变量，4-分叉及以上变量往往用于寻找新物理，如 squark 衰变， $W' \rightarrow HW$ ， $Z' \rightarrow HZ$ 等过程的强子衰变道。

a) Mass-Drop Tagger (MDT)^[82]，modified Mass-Drop Tagger (mMDT)^[97]，Soft-Drop^[98]，与 Recursive SoftDrop^[99]。这一系列算法工作的原理是，根据喷注重建的历史顺序，重新将喷注分为两部分，若子喷注的质量或横向动量满足一定的限定条件，则视原喷注为具有 2-分叉结构。Recursive SoftDrop 则是递归地使用 SoftDrop，从而实现多分叉结构的分解。以 SoftDrop 为例，算法步骤是，逆向一个 C/A 胖喷注的聚类过程，若子喷注满足，

$$\frac{\min(p_{Ti}, p_{Tj})}{p_{Ti} + p_{Tj}} > z_{\text{cut}} \left(\frac{\Delta R_{ij}}{R_0} \right)^\beta \quad (1-27)$$

则保留为合格的喷注，否则丢弃较软的子喷注继续分解过程。若分解结束仍不满足上述条件，则视为不合格喷注。 $\beta = 0$ 的特殊情况退化为 mMDT。

b) N -subjettiness^[100] 与 energy correlation function (ECF)^[101]。它们是定义在单个喷注上的一组变量，其分布对喷注的分叉数敏感。以 ECF 为例，其定义为：

$$\text{ECF}(N, \beta) = \sum_{j_1 < j_2 < \dots < j_N} \left(\prod_{a=1}^N p_{Tj_a} \right) \left(\prod_{b=1}^{N-1} \prod_{c=b+1}^N \Delta R_{j_b j_c} \right)^\beta \quad (1-28)$$

区分喷注时，通常使用相关变量的比值，包括但不限于 $ECF(N+1, \beta)/ECF(N-1, \beta)/(ECF(N, \beta))^2$ ，作为一个无量纲量，它可用于识别 N-分叉喷注。实践中，对于单分叉喷注，也可以使用前面的喷注形状变量识别，比如轻夸克与胶子喷注的区分。

衡量这些算法的表现通常借助“受试者工作特征”(receiver operating characteristic, ROC) 曲线来完成。它是用来分析二元分类器的。原始定义下，横轴表示假阳性率，纵轴表示真阳性率。ROC 曲线下面积 (area under curve, AUC) 可用于衡量分类器的性能。AUC 越大，则性能越优异。AUC=1 意味着完美的分类器，一个有效的分类器至少有 $AUC > 0.5$ 。在高能唯象领域，常见的用法是横轴表示信号的保留比例 ε_S (相当于真阳性率)，纵轴表示本底的保留比例 ε_B (相当于假阳性率) 或其等价变量，如 $1 - \varepsilon_S$, $1/\varepsilon_S$ 。使用 ROC 曲线，有助于对比不同算法的优劣。

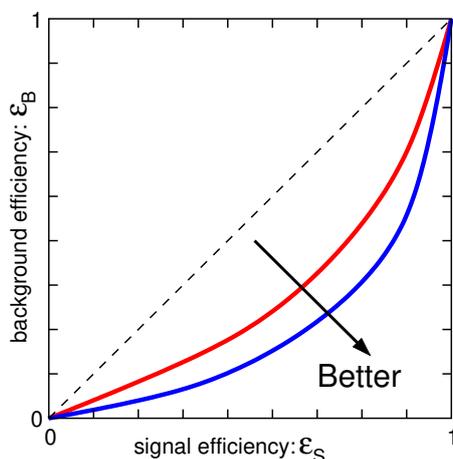


图 1.2 ROC 曲线的草图，图中蓝线的表现优于红线。图自 [102]。

人为构造的动力学变量，具有鲜明的物理意义，在对撞机实验数据的分析中效果显著。我们会展示，使用新兴的深度学习技术，采取数据驱动的方式，同样可以简单高效的完成喷注识别任务，其性能不弱于理论驱动的算法。

喷注的 4-动量是其组分的 4-动量矢量和，可以被直接拿来近似表示原始部分子的 4-动量。这种做法在 LHC 投入使用以前并没有太大问题。然而由于 LHC 的高亮度，pile-up 污染造成喷注的 4-动量远大于引发喷注的部分子 4-动量。于是，人们提出了许多专门减轻 pile-up 的算法，通称为“梳理”(groom)，比如与 Mass-Drop Tagger 同时提出的“过滤”(filter)^[82]，ATLAS 和 CMS 合作组经常使用的“修剪”(trim)^[103] 和“剪枝”(prune)^[104] 等。以 prune 为例，算法步骤是，对于一个喷注，重新使用 C/A 或 k_t 算法进行聚类，只有以下两个条件同时满足，才将新组

分合并到喷注中。

$$\frac{\min(p_{Ti}, p_{Tj})}{p_{Ti} + p_{Tj}} > z_{\text{cut}} \quad (1-29)$$

$$\Delta R_{ij} < f_{\text{prune}} \frac{2m_J}{p_{TJ}} \quad (1-30)$$

其中 m_J 和 p_{TJ} 分别为原喷注的质量和横向动量。

除了这些 groomer，还有更加细化，针对每个粒子进行分析，判断其是否属于 pile-up，且在整个事例级别而非喷注级别对 pile-up 进行减除的算法，如 Soft-Killer^[105] 和 pile-up per particle identification (PUPPI)^[106] 算法，其中后者目前已被 CMS 合作组采用。总之，经上述这些工具处理后的喷注，其 4-动量往往更接近原始部分子，从而使重建的事例具有更高的能量分辨率。

准确地重建喷注 4-动量，目的是更高效地区分信号与本底事例。首先，喷注的不变质量，作为 4-动量的导出量之一，其准确度越高，则分布的峰越窄，信号越显著。另外，若新粒子衰变到多个标准模型粒子，其中一个或多个最终形成喷注，则这些喷注的 4-动量将是重建新粒子质量不确定性的最大来源。比如，类矢量轻子可能衰变到一个 Higgs 玻色子加一个带电的轻子，再比如，双 Higgs 二重态模型中某个具有偶宇称的标量粒子可能衰变到一对 Higgs 玻色子。重建上述新粒子，其质量分辨率将取决于衰变产物中 Higgs 喷注在多大程度上还原了 Higgs 玻色子的 4-动量。

1.2.2 深度学习的运用

随着计算机硬件的算力提升，和各类高质量数据的积累，深度学习作为机器学习的一个分支获得了迅猛发展。而 LHC 的海量实验数据尤其适合借助该技术进行分析。另一方面，蒙特卡洛产生子的使用，使得唯象研究获取数据十分廉价。最近几年，深度学习领域涌现的大量模型已陆续被高能物理实验应用于事例筛选，喷注子结构分析，和提升对反常信号的灵敏度。

事例筛选属于优化问题，其目标函数就是信号显著性。很早以前，机器学习就被应用到了这一领域，比如提升决策树 (boosted decision tree, BDT)^[107]。BDT 作为一种传统的多变量分析方法曾在 Higgs 玻色子的发现中发挥重要作用^[108]。特别地，BDT 的改进版本 XGBoost^[109] 正是在一场 Kaggle 挑战赛^[110] 中开发出来优化 Higgs 玻色子搜寻的，现已广泛应用于机器学习领域。筛选事例的基础是二元分类任务，神经网络 (NN) 十分擅长处理这类问题。网络的输出是类别判定，是信号还是本底。网络的输入可以是固定数量的物理量，包括特定粒子，喷注或粒子对的数量，动量，不变质量，距离， b -tag, m_{T2} , \cancel{E}_T 等变量。

喷注的分类属于模式识别问题，正是深度神经网络 (deep neural network, DNN) 大放异彩的地方。与事例筛选不同的是，子结构分析使用哪种神经网络，取决于喷注的表示，即使用怎样的数据结构编码粒子的种类，4-动量等信息。给喷注分类，尽管较早的研究使用根据物理意义构造的变量^[111-112]，包括各粒子 4-动量的导出量，但是无法保证喷注的全部信息被有效利用。目前，喷注的表示以及相应的神经网络，有以下几种形式：1) 用图像表示喷注^[113]，以像素的亮度代表粒子能量或横向动量，可用不同图层代表粒子种类。使用卷积神经网络 (convolutional neural network, CNN) 处理喷注图像。这种表示刚好契合了量能器有限的空间分辨率。本论文重建 Higgs 玻色子的方法基于此表示。2) 用序列表示喷注^[114]，可根据横向动量排序。类似自然语言处理，使用循环神经网络 (recurrent neural network, RNN) 处理粒子序列。3) 用树状结构表示喷注^[115]，模拟喷注算法的聚类过程。使用的网络结构为递归神经网络 (recursive neural network, RecNN)。4) 用图表示喷注^[116]，图上的节点表示粒子及其属性，节点之间的边表示相应粒子的距离。使用的网络结构为图神经网络 (graph neural network, GNN)。上述方法已经被多个研究者用于 top 夸克^[117-119]， W ， Z 玻色子^[120-121]，Higgs 玻色子^[122] 的识别，轻夸克与胶子的区分^[123-124]，和 pile-up 的减轻^[125-126]。在各项任务中，神经网络的表现都达到或超过了传统喷注识别算法。

上述任务属于监督学习的范畴，即数据样本类别已知，通过训练网络使其具备分类能力。这通常需要高质量的模拟数据。而无监督学习或弱监督学习可以使我们不依赖模拟产生的样本，直接使用实验数据训练网络，使其具备发现反常或提升信号显著性的能力。一种无监督学习常用的网络结构称作自编码器 (autoencoder)。自编码器的直接作用是数据降维。它先将输入数据压缩，形成一种低维表示，再解压缩，根据低维表示尽可能忠实的重建数据。但由于压缩有损，实际输出只保留了输入的重要特征。假设一个样本大部分是标准模型的过程，只含有少量新物理或反常的过程^[127-130]。在此样本上训练以后，自编码器将可以很好的重建正常过程而对反常过程重建欠佳，从而达到发现反常的目的。

弱监督学习则能在已知存在信号的区间提升信号显著性，比如在某变量 V 的分布中寻找淹没的共振峰 P ^[131-132]。做法是将任一小区域 R_i 内的事例设为信号，附近事例设为本底，使用与 V 无关的其他变量 $\{O_j\}$ 训练一个网络 N_i ，然后将网络 N_i 判定为本底的事例剔除。假设共振峰 P 位于 R_a 以内，则经网络 $\{N_i | i \neq a\}$ 筛选后， V 的相对分布不会有明显变化，只有经网络 N_a 筛选后 P 才会突显。在文献 [131] 的例子中，此策略使信号显著性由 2σ 提高到了 7σ 。

后面我们将使用深度学习算法，完成喷注重建与识别任务。我们认为，这些算

法可以用来分析已有或即将积累的对撞机数据，帮助提高新物理的搜寻灵敏度。

神经网络普遍具有多层结构，包括输入层，隐藏层和输出层。其结构如图1.3所示。图中网络输入数据维度是3，输出数据维度是2，共两个隐藏层，网络宽度为

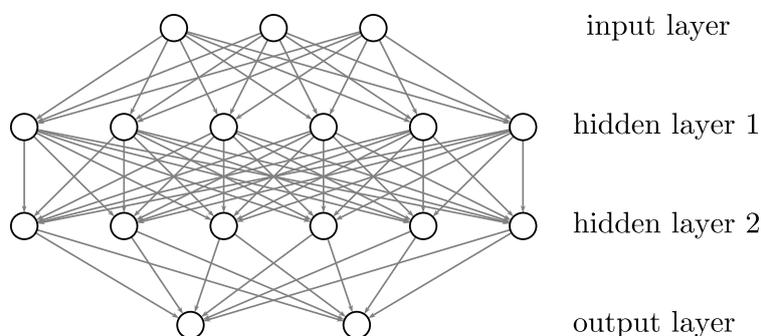


图 1.3 深度神经网络的结构示意图。

6. 实际应用中，以图像分类为例，如果一个神经网络的输入数据是图像，其分辨率为 $w \times h$ ，有 3 个颜色通道 (RGB)，则输入的维度是 $w \times h \times 3$ 。如果网络的用途是分类，则可以令输出的维度为 n ，对应 n 个类别的得分。神经网络的深度已从最早的不足十层发展到如今的可达上千层。如果有超过 1 个隐藏层，则称其为深度神经网络。每一层由一定数量代表神经元的节点组成。节点的数值 y 是从上一层节点 x 经矩阵运算后再作用一个激活函数 f 得来。矩阵运算所调用的权重 w 和偏置 b 是可训练的参数。激活函数的形式固定，用来模拟神经元对刺激的非线性相应。

$$y_j = f(w_j^i x_i + b_j) \quad (1-31)$$

常见的激活函数有单位阶跃函数 (Heaviside function)，S 型函数 (sigmoid function)，双曲正切 (tanh) 等几种。训练神经网络，特别是隐藏层数量很多的深度神经网络时，梯度传递层数过多会导致梯度消失。因而实践中，为了使网络能够正常收敛，激活函数通常使用整流线性单元 (rectified linear unit, ReLU)， $f(x) = \max(0, x)$ 。

根据万能近似定理 (universal approximation theorem)，一个前馈的神经网络只需要一个隐藏层，且包含有限个神经元，即宽度有限，并使用满足挤压性质 (将 $(-\infty, \infty)$ 映射到有限区间) 的激活函数，那么它能以任意精度近似一个欧几里得空间紧致集上的连续函数^[133-134]。然而该定理并不保证网络可以通过训练学习到合适的参数，而且对宽度的需求呈指数增长。对于宽度固定但是深度增长的网络有类似定理^[135]。实践当中，越深的网络通常对数据拟合的更好。

当用于分类任务时，常使用 softmax 函数将网络针对每个类别输出的分数 x_i

归一化为属于各类的概率。

$$q_i = \frac{e^{x_i}}{\sum_{j=1}^n e^{x_j}} \quad (1-32)$$

衡量神经网络输出的预测值 $f(x)$ 与预期值 y 的偏离，要用到损失函数 (loss function)。对于回归问题，常用 Huber 损失，定义为：

$$L_\delta(y, f(x)) = \begin{cases} \frac{1}{2}(y - f(x))^2 & |y - f(x)| \leq \delta \\ \delta|y - f(x)| - \frac{1}{2}\delta^2 & |y - f(x)| > \delta \end{cases} \quad (1-33)$$

其中 $\delta = 1$ 的特例即所谓平滑 L_1 损失，它将 L_1 (线性) 损失和 L_2 (平方) 损失结合，在保持收敛的稳定性同时避免了梯度爆炸。而对于分类问题，常用交叉熵 (cross entropy) 损失：

$$L = - \sum_i p_i \log q_i \quad (1-34)$$

其中， p_i 代表真实的概率，0 或 1，而 q_i 代表预言的概率。

给定输入数据和所需拟合结果的情况下，损失函数以各层网络参数 w 和 b 为自变量。易见，损失函数是恒大于 0 的。如果网络的输出完美地符合预期值，上述损失函数为 0。最后，通过反向传播 (本质为求导数的链式法则) 和梯度下降算法，调整参数以最小化损失函数的过程就叫神经网络的训练。使用梯度下降算法时，每次仅计算几个样本的损失函数之和，这几个样本称作一个批量 (batch)。将全部样本使用一次，称作训练一个周期 (epoch)。实际上由于网络的性能有限，且数据本身可能有瑕疵，给定的预期值可能错误等原因，损失函数几乎不会下降到 0。盲目追求损失函数的下降，还可能导致过拟合。实践中，通常将数据样本分为三部分，分别用于训练，验证和测试。在训练集上的损失函数用于训练，当验证集上的损失函数之和不再明显下降时或者反弹前停止训练，测试集则用于评估真实场景中网络的性能。

卷积神经网络是一种特化了的深度神经网络，能够处理具有网格结构的数据，而且对大规模的数据易于扩展。一种典型的二维数据就是图像。粒子在对撞机量能器上的能量沉积，也是 $\eta - \phi$ 坐标的二维函数。神经网络中用到的卷积与数学领域略有不同，可以理解为一种加权平均。对于输入的二维数据 I ，卷积运算定义为

$$S(i, j) = \sum_m \sum_n I(i + m, i + n) K(m, n) \quad (1-35)$$

其中 K 称为卷积的核，是可训练的参数， S 称为特征映射 (feature map)。卷积网

络与一般神经网络最大的不同就在于卷积核，它意味着式(1-31)中的权重 w_j^i 只对少数 i, j 非 0，对应于位置邻近的神经元，或者说 w 是一个稀疏矩阵。式(1-35)中指标 i, j 遍历同一层的神经元时， K 保持不变，意味着对于同一卷积核，大量神经元可以共用参数。这是它易于扩展的根本原因。

卷积网络中另一必要运算是池化 (pooling)，属于一种下采样方法，定义为邻近若干格点的某种统计平均。注意到，与卷积不同，池化运算的参数不需要训练。其中最常见最大池化 (max pooling)，就是取出邻域中最大值。池化有一个特性，就是对输入数据的少量平移和旋转不敏感。

一个卷积神经网络主要由卷积层，池化层，全连 (fully-connected) 层构成。大致来说，卷积层计算出不同层次的特征映射，池化层用于调整图像的分辨率 (一般是缩小)，全连层用于最后的回归或分类。三者是卷积神经网络的基本构筑模块，将它们以各种排列组合堆叠数层可以设计出不同的网络架构，从而完成各类图像识别任务，比如分类，目标检测，语义分割，和姿势估计等。有些网络还包含反卷积层用于上采样，增加图像的分辨率，这样一来，即使经过了池化，网络的最终输出仍可以保持原有图像的分辨率。

前面讲到，已有多种神经网络被用到喷注识别上，但是它们必须依赖喷注算法重建喷注。而最适合在探测器上定位喷注并将其组分粒子从本底中提取出来的神经网络无疑是卷积神经网络。原因是：

1) 已经知道，卷积神经网络易于扩展到大规模。目前只有它可以高效地处理一个事例全部粒子的 4-动量。其它网络结构的输入，只能是代表喷注组分粒子性质的数据，或者事例中的复合变量，比如喷注 4-动量，粒子间距等。

2) 循环和递归网络，需要定义数据的序列，这对单一喷注没有问题，而对一个事例中的全部粒子，无法简单自然的定义其序列。粒子在探测器上的二维分布结构却是显而易见的。

3) 卷积运算具有平移不变性，刚好契合了强子对撞机上粒子沿 z 轴的 Lorentz 提速不变性。且量能器有限的空间分辨率也与图像的像素概念相契合。需要注意的是，这要求我们使用粒子的 η ， ϕ 和 p_T 表示其属性，因为对于极端相对论性粒子，这些变量是 Lorentz 提速不变的 ($\Delta\eta$ 不变)。

4) 通过卷积神经网络实现的目标检测算法近几年有了惊人的进步，可以帮助我们找到喷注并对其分类，而语义分割算法则是聚类算法的直接替代者。问题在于将对撞机上的事例处理成神经网络可以接受的形式并优化网络的性能。

当然，卷积神经网络有其局限性，比如粒子的种类，反应顶点等信息很难整合进网络的输出，意味着它无法利用所有信息，这可能影响到它的分辨能力。不过

需要注意的是，喷注的聚类算法，同样也只利用了粒子的 4-动量信息。

1.3 本论文组织架构

本论文主要内容如下：第二章，从 Higgs 玻色子质量和电弱真空稳定性出发，我们使用 split SUSY 的两圈重整化群方程，计算超对称破缺能标的上限，并讨论在未来对撞机上搜索 gluino 的前景。第三章，受观测到的 B 物理反常启发，我们研究 $U(1)_X$ 模型的参数空间，以及在高亮度 LHC 上搜索模型中 Z' 玻色子和类矢量轻子的方案。第四章，考虑违反 R 宇称的超对称，借助喷注图像和卷积神经网络来识别强子衰变的 neutralino，重做 ATLAS 搜索 gluino 的分析，改进对 gluino, neutralino 的质量排除极限。第五章，以 Higgs 喷注为例，引入 Mask-RCNN 深度神经网络框架，以全空间的粒子横向动量为输入信息，研究其在信号与本底过程 (QCD 喷注) 区分，重建部分子 4-动量等方面对原有算法的改进。最后一章，对论文的工作进行总结。

第 2 章 Split SUSY 的超对称破缺能标上限

2.1 介绍

超对称的电弱精细调节是一个严重的问题，目前已有一些颇有前景的解释。在其中之一，称作超自然 (super-natural) SUSY^[136-137] 的情形中，如果我们采用无能标超引力 (no-scale supergravity) 边界条件^[138] 和 Giudice-Masiero 机制^[139]，MSSM 中将不会有剩余的电弱精细调节，尽管这样会得到较重的质量谱。super-natural SUSY 可以自然地解决电弱精细调节问题，于是下一个重要问题是超对称粒子质量是否有上限。特别是，如果这样的上限不超过 TeV 量级，我们可以在未来的对撞机上探测 SUSY，例如 FCC-hh (Future Circular Collider-hadron hadron) 和 SppC (Super proton-proton Collider)。这里，我们在 split SUSY 方案中考虑它。我们已经知道 SUSY 是一个破缺了的对称性，并且破缺可能发生在一个很高的能标，超过了目前 LHC 能达到的范围，未来的对撞机，如 FCC-hh 和 SppC 等也未必有探测的能力。这里，我们首先利用观测到的 Higgs 玻色子质量 m_h 作为工具来预测 sfermion 质量的上限。在数值计算中，我们以 1 个 σ 为界，从中心值改变 top 夸克的质量 m_t ，并且改变 $\tan\beta$ 的值从 1 到 60。使用最新的 top 夸克质量和强相互作用耦合常数 α_3 的数值，固定 Higgs 玻色子质量 $m_h = 122, 125, 127$ GeV，我们得到了 SUSY 破缺的能标 M_S 的范围，对于 $\tan\beta = 2$ 和 $\tan\beta = 50$ ，分别是 $[10^{5.2}, 10^{6.2}]$ GeV 和 $[10^{3.8}, 10^{4.3}]$ GeV。特别地，对于 $\tan\beta \gtrsim 4$ ，有 $M_S \lesssim 10^{1.5} \text{ TeV} \simeq 31.6 \text{ TeV}$ 。

应该再次强调的是，由于 LHC 尚未观测到任何 SUSY 信号，当前 SUSY 唯象的关键问题是超对称粒子的质量到底有多重。我们在此通过计算表明了对 $\tan\beta \gtrsim 4$ ，sfermion 质量大致低于 31.6 TeV。文献 [140] 表明在 HE-LHC 上，积累 15 ab^{-1} 的数据量可以对质量高达 3.6 TeV 的 stop squark \tilde{t}_1 以 95% 置信度进行探测。而根据文献 [141]，HE-LHC 可以对质量高达 6 TeV 的 squark 进行探测。有意思的是， $\tan\beta \gtrsim 4$ 的情形在 FCC-hh 和 SppC 上有更大的机会被发现^[142-143]。

在上述计算以外，我们还试图通过电弱真空的稳定性限制来估算 sfermion 的质量上限。即选定 Higgs 玻色子质量 m_h ，计算 Higgs 场四次耦合系数 λ 转变为负值时 SUSY 破缺的能标。在这种情况下，我们表明对于 $m_h = 123, 125, 127$ GeV，电弱真空稳定性给 M_S 的上限分别为 $10^{8.2}, 10^{8.7}$ 和 $10^{9.6}$ GeV。这些限制大致比标准模型给出的限制小一个数量级^[144-146]，其来源是额外的 Higgs-Higgsino-gaugino 耦合。由于 squark 很重，gluino 可以是长寿命的，其衰变宽度依赖于 M_S 。对于 gluino 的寿命我们将进行简短的讨论。

2.2 sfermion 质量上限

在 LHC 上发现 Higgs 玻色子基本确定了标准模型作为低能有效理论的地位，但是其它新的基本粒子的缺席给超出标准模型新物理的能标设置了越来越高的下限。在这样的情形下，我们可以设法通过已知的 Higgs 玻色子质量 m_h 和 Higgs 真空的稳定性，给出新物理能标的上限。在标准模型中，Higgs 场的自耦和系数 λ 是一个自由参数，然而超对称标准模型中， λ 并不是自由的。而是与模型中其它参数相关联。在树图阶，Higgs 玻色子的质量以 Z 玻色子的质量上限，

$$m_h < m_Z |\cos(2\beta)| \quad (2-1)$$

但是为与 LEP 给 Higgs 玻色子的质量下限^[147]兼容，必须考虑辐射修正^[148-150]。对于一个质量在 125 GeV 左右的 Higgs 玻色子，需要小混合角且质量为几个 TeV 的 stop squark 或者大混合角且质量在 TeV 量级 stop squark。在 split SUSY 中，sfermion 在 SUSY 破缺能标 M_S 处脱耦， $\lambda(M_S)$ 为定义在此能标的 Higgs 场四次耦合系数。低能标的系数可以由重整化群方程计算出来，从而得到低能有效理论。因为增加 λ 会增大 Higgs 玻色子质量， m_h 的上下限也能对 M_S 给出相应的限制。计算分为两步，在 SUSY 破缺能标 M_S 以下，用 split SUSY 的重整化群方程跑动到 1 TeV。假设 gaugino 和 Higgsino 的质量是简并的， $M_1 = M_2 = M_3 = \mu = 1$ TeV，并且超对称软破缺项中的 Higgs-stop 耦合系数 $a_t = 0$ ，因为在 split SUSY 中， $a_t/M_S \ll 1$ 。在 1 TeV 以下为标准模型。经圈图修正的 Higgs 场四次耦合系数边界条件取以下形式^[151]，

$$\lambda(M_S) = \frac{1}{4} \left[g_2^2(M_S) + \frac{3}{5} g_1^2(M_S) \right] \cos^2 2\beta \quad (2-2)$$

$$+ \Delta\lambda^{1\ell, \text{reg}} + \Delta\lambda^{1\ell, \phi} + \Delta\lambda^{1\ell, \chi^1} + \Delta\lambda^{1\ell, \chi^2} + \Delta\lambda^{2\ell} \quad (2-3)$$

其中 g_2 和 $g_1 = \sqrt{5/3}g_Y$ 分别是 $SU(2)_L$ 和 $U(1)_Y$ 的规范耦合系数。在 M_S 之上，超对称恢复，在 M_S 之下，超对称破缺。为计算 M_S 之下的 Higgs 玻色子质量，这里使用了 split SUSY 两圈的重整化群方程，加上两圈的 gaugino 和 Higgsino 贡献和一圈的标准模型 Yukawa 耦合的跑动^[41]。对于 Higgs 自耦和系数 λ ，使用两圈的重整化群方程^[152]。

在数值计算中使用的精细结构常数 α_{EM} 和 m_Z 能标的弱混合角 θ_w 如下^[153]，

$$\sin^2 \theta_w(M_Z) = 0.22332 \pm 0.00007 \quad (2-4)$$

$$\alpha_{\text{EM}}^{-1}(M_Z) = 137.03599 \quad (2-5)$$

强相互作用耦合常数和 top 夸克的极点质量是计算中极为重要的参数，它们的取值分别为^[154],

$$M_t = 173.34 \pm 0.76 \text{ GeV} \quad (2-6)$$

$$\alpha_3(M_Z) = 0.1187 \pm 0.0016 \quad (2-7)$$

我们使用了经 top 夸克辐射修正的一圈 Higgs 有效势^[155]，通过将其最小化计算 Higgs 玻色子质量，

$$V_{eff} = m_h^2 H^\dagger H + \frac{\lambda}{2} (H^\dagger H)^2 - \frac{3}{16\pi^2} h_t^4 (H^\dagger H)^2 \left[\ln \frac{h_t^2 H^\dagger H}{Q^2} - \frac{3}{2} \right] \quad (2-8)$$

这里 m_h^2 是 Higgs 玻色子质量平方， h_t 是 top 夸克的 Yukawa 耦合 $m_t = h_t v_u$ ，能标 Q 选在 Higgs 玻色子质量。

对于 $\overline{\text{MS}}$ (修正的最小减除) 方案下的 top 夸克质量 m_t ，我们使用了两圈修正的值，其与 top 夸克极点质量关系如下，

$$M_t = m_t(m_t) \left\{ 1 + \frac{4\alpha_3(m_t)}{3\pi} + \left[13.4434 - 1.0414 \sum_{k=1}^5 \left(1 - \frac{4}{3} \frac{m_k}{m_t} \right) \left[\frac{\alpha_3(m_t)}{\pi} \right]^2 \right] \right\} \quad (2-9)$$

这里 m_k 代表其它夸克的质量。另外， $\alpha_3(m_t)$ 使用了两圈的重整化群跑动。

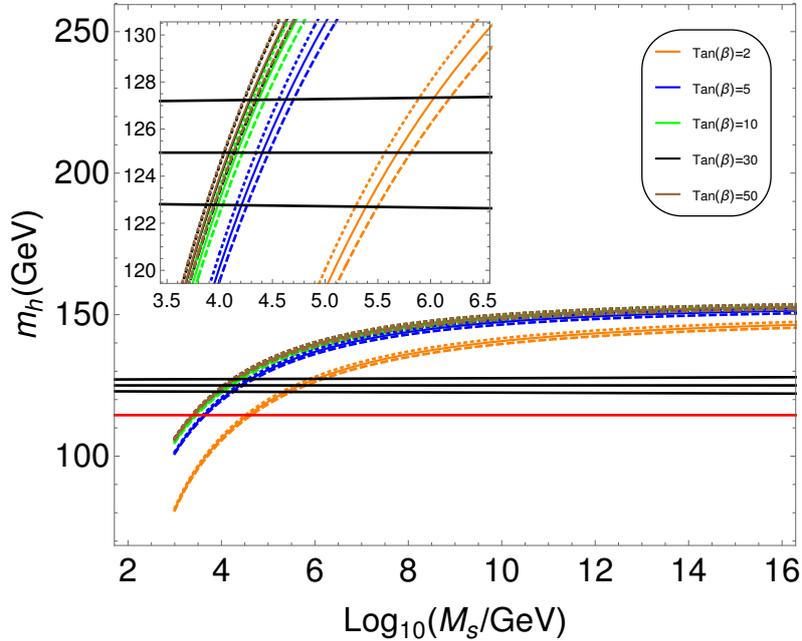


图 2.1 在 split SUSY 中，对于恒定的 $\tan \beta$ 值，语言的 Higgs 玻色子质量 m_h 随 SUSY 破缺能标 M_S 的变化。红色水平线表示 LEP 给 Higgs 玻色子的质量下限，114.5 GeV。实线，虚线和点线分别代表用 M_t ， $M_t - \delta M_t$ 和 $M_t + \delta M_t$ 计算得到的结果。

图2.1为计算结果，针对不同的 $\tan \beta$ 值，展示了 m_h 和 M_S 的关系。水平的红色

实线代表了 LEP 给 Higgs 玻色子质量的限制，黑线则代表 $m_h = 123 \text{ GeV}$, 125 GeV 和 127 GeV 。由于理论计算的不确定度^[156]，我们允许 $\pm 2 \text{ GeV}$ 的变动。实的曲线表示使用 M_t 的计算结果，而虚线表示使用 $M_t \pm \delta M_t$ 的计算结果，其中 $\delta M_t = 0.76 \text{ GeV}$ 。由于四次耦合系数依赖于 $\cos^2 2\beta$ ，对于较大的 $\tan \beta$ 值，Higgs 玻色子质量几乎保持不变。这种趋势在图 2.1 中清晰可见，当 $\tan \beta \geq 5$ 时，曲线靠的很近。对于给定的 $\tan \beta$ ，无法通过挑选 M_S 的值来产生任意的 Higgs 玻色子质量。于是，可以得到 Higgs 玻色子质量的上限，大约在 153 GeV 左右。

如前所述，split SUSY 中，所有 sfermion 具有与超对称破缺能标 M_S 在同一数量级的质量，但是 gaugino 和 Higgsino 可以很轻。从图 2.1 中我们看到，当 $\tan \beta = 2$ 时，对于 $m_h = 123 \text{ GeV}$ 到 $m_h = 127 \text{ GeV}$ ，相应 M_S 的值为 $10^{5.2} \text{ GeV}$ 到 $10^{6.2} \text{ GeV}$ 。当 $\tan \beta = 50$ 时，相应 M_S 的值为 $10^{3.8} \text{ GeV}$ 到 $10^{4.3} \text{ GeV}$ 。因此，sfermion 质量的下限和上限大约分别为 1 TeV 和 10^3 TeV 。注意到对于 $\tan \beta \gtrsim 4$ ，有 $M_S \lesssim 10^{4.5} \text{ GeV}$ ，更有可能在未来的对撞机上探测 split SUSY^[142-143]。特别地，若 gluino 的质量小于 4 TeV ，借助 3ab^{-1} 的积分亮度，对与 gluino 联合产生的第一代和第二代 squark 的搜寻能力，在压缩质量谱（gluino-LSP 质量劈裂较小）和非压缩质量谱情形下分别可达 37 和 32 TeV ^[157]。因此， $\tan \beta \gtrsim 4$ 的 split SUSY 可以在未来的 FCC-hh 和 SppC 上探测。我们得到的 M_S 上下限与文献 [40-41] 中略有不同，这是因为计算中使用了不同的 M_t 和 α_3 的数值。总之，在 LHC 搜寻 sfermion 一无所获的情形下，我们仍有机会在不远的将来找到超对称粒子。

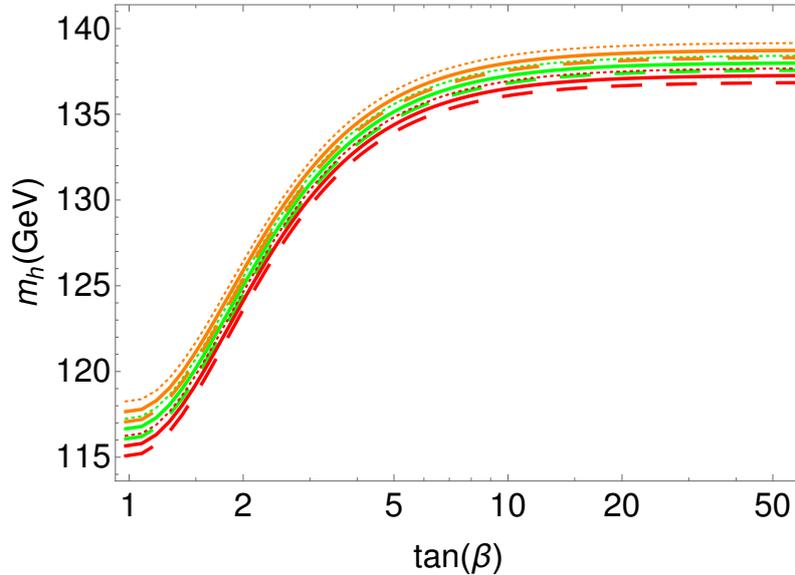


图 2.2 固定 SUSY 破缺能标 $M_S = 10^{5.2} \text{ GeV}$ 时，split SUSY 预言的 Higgs 玻色子质量 m_h 随 $\tan \beta$ 的变化。橙色，绿色，红色曲线分别对应 $M_t + \delta M_t$ ， M_t 和 $M_t - \delta M_t$ 。虚线，实线和点线分别代表 $\alpha_3 + \delta \alpha_3$ ， α_3 和 $\alpha_3 - \delta \alpha_3$ 。

图2.2展示了当 $M_S = 10^{5.2}$ GeV 时, 对不同的 α_3 数值和 top 夸克质量 M_t , Higgs 玻色子质量 m_h 与 $\tan\beta$ 的关系。橙色, 绿色和红色曲线分别代表 $M_t + \delta M_t$, M_t 和 $M_t - \delta M_t$ 。虚线, 实线和点线分别代表 $\alpha_3 + \delta\alpha_3$, α_3 和 $\alpha_3 - \delta\alpha_3$ 。实线清晰地表明, 当我们将 top 夸克质量增加一个 σ 时, Higgs 玻色子质量增加, 反之亦然。与实线相对比, 点线表明, 当我们将 α_3 的值减小一个 σ 时, Higgs 玻色子质量增加。类似地, 虚线表明, 将 α_3 的值增加一个 σ 时, Higgs 玻色子质量减小。根据我们的计算结果, 当 $\tan\beta$ 从 1 变动到 60, Higgs 玻色子质量范围是 114 到 139 GeV。

在标准模型中, 当 Higgs 玻色子质量 $m_h \lesssim 126$ GeV 时, 来自于 top 夸克质量的辐射修正会使 Higgs 场四次耦合系数 λ 在普朗克能标 M_{pl} 以下变为负值, 从而使 Higgs 势变得不稳定^[158], 这被称为稳定性问题。在 split SUSY 中, M_S 必须在 Higgs 场四次耦合系数变为负值的能标以下, 这一条件可以给出 M_S 的上限。我们在此研究中考虑了四次耦合系数精确到两圈的重整化群方程。

图2.3展示了四次耦合系数 λ 随能标 Q 的跑动。在 $m_h = 123$ GeV, 125 GeV 和 127 GeV 时, 对 λ 的计算结果在三个图分别展示。每个图中位于中间的实线对应 M_t 和 α_3 的中心值, 点线和虚线分别对应 $\alpha_3 \pm \delta\alpha_3$, 位于最上和最下的实线则对应 $M_t \pm \delta M_t$ 。从图中可知, Higgs 场四次耦合系数随 top 夸克质量增加而增加, 随 α_3 的值增加而减小, 即二者的数值变动对 λ 有截然相反的贡献。图2.3显示, 对于 $m_h = 123$ GeV, 125 GeV 和 127 GeV, Higgs 场四次耦合系数 λ 分别在能标 $Q \simeq 10^{8.2}$ GeV, $10^{8.7}$ GeV, $10^{9.6}$ GeV 时与 x 轴相交。这给出了来自于电弱真空稳定性的限制。这些限制比标准模型中的要小大概一个数量级, 因为超对称模型中有额外的 Higgs-Higgsino-gaugino 耦合。

2.3 在 LHC 上搜寻 gluino

在强子对撞机上, SUSY 的一般信号由含有缺失横向能量 E_T 的多喷注, 多轻子末态构成。潜在的典型物理过程涉及新的带色荷重粒子 (squark 和 gluino) 的成对产生, 它们再级联衰变到最轻的超对称粒子。在 split SUSY 中, 由于所有的 squark 都很重, 这些级联衰变基本不会发生。唯一可能存在的较轻末态是 bino (M_1), wino (M_2), gluino (M_3) 和 Higgsino (μ)。如果要求存在一个可行的暗物质粒子, 那么 bino, wino 和 Higgsino 的质量可能必须要在 TeV 量级^[159-161]。

对撞机上可能存在的 gluino 特征信号被认为是可以证实 SUSY 的确凿证据。split SUSY 中 TeV 量级 squark 的缺失影响了 gluino 的产生和衰变, 使其区别于 MSSM。比如, split SUSY 仍然可以通过 gg 和 $q\bar{q}$ 的湮灭成对产生 gluino, 但是与 MSSM 不同的是, 通过交换 squark 产生 gluino 的 t 道和 u 道贡献可以忽略^[162]。所

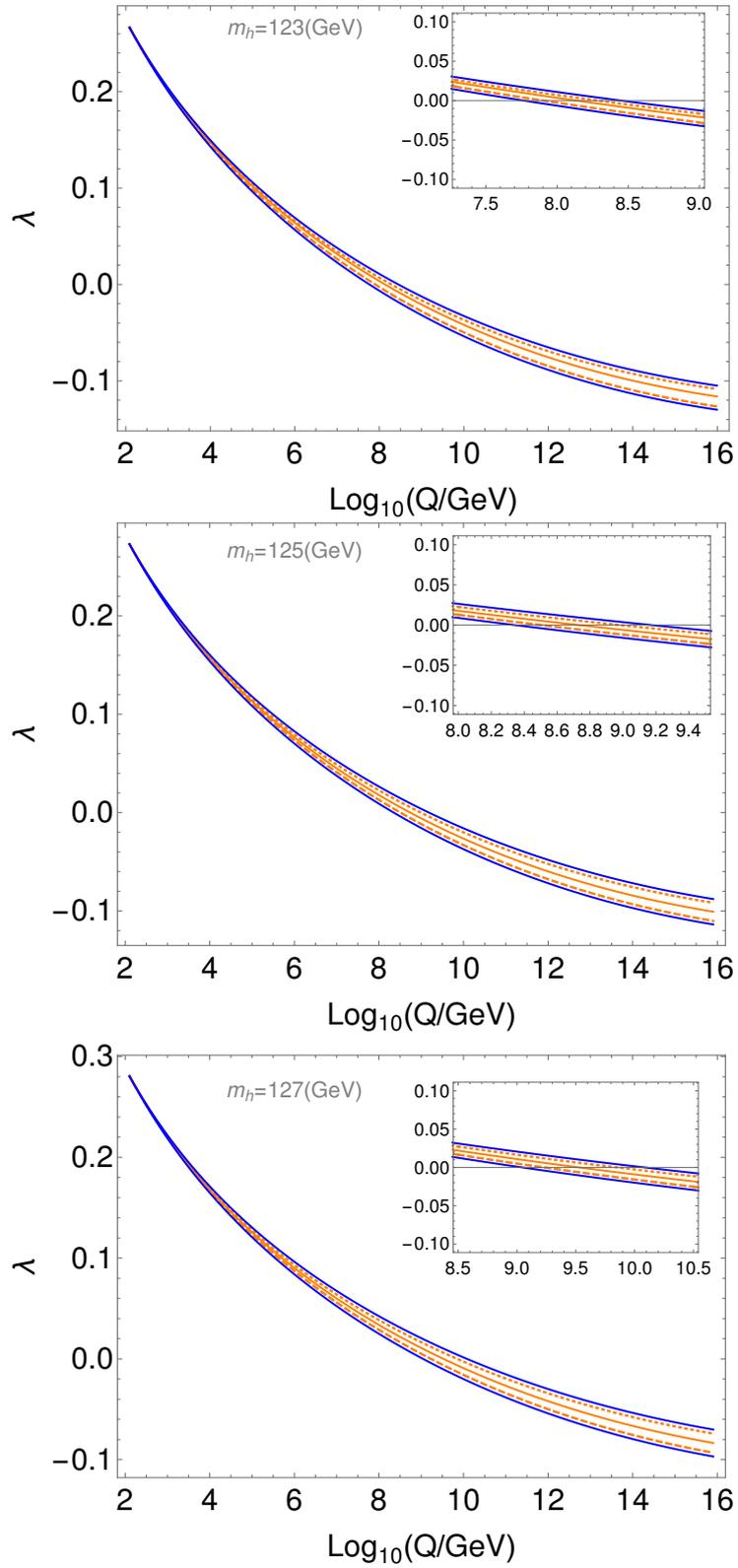


图 2.3 Split SUSY 中 Higgs 场四次耦合系数 λ 随能标 Q 的跑动。上中下图分别对应 $m_h = 123 \text{ GeV}$, 125 GeV 和 127 GeV 。橙色实线对应 $M_t = 173.34 \text{ GeV}$ 和 $\alpha_3 = 0.1187$ 。橙色点线, 虚线和蓝色实线分别对应改变 α_3 或 M_t 的数值一个 σ 时的计算结果。

以 gluino 在 split SUSY 中的总产生截面比 MSSM 中略小。

在 split SUSY 中, 可行的衰变道包括 $\tilde{g} \rightarrow q\bar{q}\tilde{\chi}^0$, $\tilde{g} \rightarrow qq'\tilde{\chi}^\pm$ 和圈图诱导的 $\tilde{g} \rightarrow \gamma\tilde{\chi}^0$ 。由于这些衰变涉及 squark, gluino 可以是长寿命的^[162-164]。最新的 LHC 直接搜寻实验已经排除了质量在 2 TeV 以下的 gluino^[165-167]。另外, LHC 对长寿命粒子的搜寻也对长寿命的 gluino 给出了限制。在文献 [168] 中, ATLAS 合作组对受 split SUSY 启发的一种简化模型中质量至高为 2.37 TeV, 平均寿命在 $O(10^{-2})$ 到 $O(10)$ 之间的长寿命 gluino 给出了排除极限。在另一项研究中^[169], CMS 合作组对同样受 split SUSY 启发, 涉及长寿命 gluino 产生与衰变的一种简化模型提供了搜寻灵敏度。他们考虑的 gluino 固有衰变长度 $c\tau_0$ 在 10^{-3} mm 到 10^5 mm 之间。对于 $c\tau_0 = 1$ mm 和亚稳态的情形, 质量分别为 1750 GeV 和 900 GeV 的 gluino 可被探测。对于 $c\tau_0 \gtrsim 1$ m, 探测灵敏度在一定程度上依赖于具体模型假设。用 M_S 和 gluino 质量 $m_{\tilde{g}}$ 表示的 gluino 寿命由下式给出^[164],

$$\tau_{\tilde{g}} \approx 4\text{s} \times \left(\frac{M_S}{10^9 \text{ GeV}} \right)^4 \times \left(\frac{1 \text{ TeV}}{m_{\tilde{g}}} \right)^5 \quad (2-10)$$

除了通常的 gluino, 即瞬发衰变的 gluino 产生的信号, 长寿命 gluino 也能在碰撞机上产生有意义的信号。平均寿命在 10^{-12} s 左右的 gluino 最有可能在抵达探测器之前衰变, 可以通过具有缺失横向能量 E_T 和偏移反应顶点的多喷注末态探测^[162]。而平均寿命在 10^{-12} s 到 10^{-7} s 的 gluino 可能会在探测器的结构体内衰变, 搜寻这样的信号会是一个棘手的问题, 因为它们很可能淹没在 QCD 本底当中。寿命 $\tau_{\tilde{g}} \gtrsim 10^{-7}$ s 的 gluino 预计会在探测器以外衰变, 这种情况下它们表现得像是稳定粒子, 需要使用稳定重粒子的搜寻方法^[170]。对我们的情况而言, 为得到 gluino 平均寿命的粗略估计, 下面假设 $m_{\tilde{g}} = 2.5$ TeV。在 $M_S \simeq 1$ TeV 时, $\tau_{\tilde{g}} \gtrsim 10^{-26}$ s, 对应于瞬发衰变的 gluino。这时 gluino 可以衰变到轻夸克和 neutralino ($\tilde{g} \rightarrow q\bar{q}\tilde{\chi}_1^0$) 或者重夸克和 neutralino ($\tilde{g} \rightarrow t\bar{t}\tilde{\chi}_1^0$)。文献 [142,162] 表明, 在 $m_{\tilde{\chi}_1^0} < 1$ TeV 的假设下, 对于这两种衰变模式, 质量分别低于 11 TeV 和 6 TeV 的 gluino 可以在 100 TeV 对撞能量的质子-质子对撞机上发现。另一方面, 对于 $M_S \simeq 10^6$ TeV 和 10^7 TeV, gluino 的平均寿命分别约为 10^{-14} s 和 10^{-10} s。寄希望于未来对于更高能对撞机数据分析效率的提升, 我们将有可能找到这种 gluino 并最终确定 SUSY 破缺能标 M_S 。

2.4 结论

利用最新的 top 夸克质量和强作用耦合系数观测值, 我们通过 Higgs 玻色子的质量研究了 split SUSY 中 sfermion 质量的上限。针对 $\tan\beta$ 的各种值计算 Higgs 玻

色子质量, 我们首先发现 $\tan\beta = 2$, $\tan\beta = 4$ 和 $\tan\beta = 50$ 的超对称破缺能标 M_G 的上限分别是 $10^{6.2}$, $10^{4.5}$ 和 $10^{4.3}$ GeV。特别是, $\tan\beta \gtrsim 4$ 的 split SUSY 可以在 FCC-hh 和 SppC 上探测。此外, 对于 $m_h = 123, 125$ 和 127 GeV, 我们表明, 从电弱真空稳定性出发, Higgs 场四次耦合系数 λ 分别在 $10^{8.2}$ 、 $10^{8.7}$ 和 $10^{9.6}$ GeV 左右变为负值。由于额外的 Higgs-Higgsion-gaugino 耦合, 这些限制比标准模型中小约一个数量级。由于 squark 很重, gluino 可以是长寿命的。我们讨论了 gluino 的寿命, 与当前的 LHC 观测结果进行了比较, 并讨论了在 FCC-hh 和 SppC 发现 gluino 和超对称破缺能标 M_G 的前景。

第3章 类矢量轻子搜寻

B 物理反常的一种解决方案是引入额外的 Z' 玻色子，它可以诱导味道改变中性流 $b \rightarrow sll$ 。它与标准模型费米子不同寻常的耦合系数可以通过类矢量粒子与标准模型费米子的混合来实现。检验此类模型的方式就是在对撞机上搜寻 Z' 玻色子或类矢量粒子。在确定搜寻策略之前，我们首先要分析模型的参数空间，考察各类实验对新粒子质量的限制，然后根据模型特有的信号过程，进行对撞机唯象研究。

3.1 $U(1)_X$ 模型

为了容纳 Z' 玻色子和类矢量粒子，可以在标准模型的基础上加入一个 $U(1)_X$ 规范对称性。 $U(1)_X$ 模型的实现方式并不唯一，比如，它可以是大统一群 E_6 的一个子群。相应的，在低能标将会引入额外的右手中微子，新的 Higgs 玻色子和其他标量粒子等。本文考虑的实现方案首先在文献 [171] 中提出，与原方案略微不同的是，由于暂时不考虑暗物质，这里简化掉一个标量场。假设所有标准模型场的 $U(1)_X$ 荷为 0。新引入一组旋量场和一个复标量场，它们在规范群 $SU(3)_C \times SU(2)_L \times U(1)_Y \times U(1)_X$ 下的表示如表 3.1 所示。

表 3.1 旋量场 Q , L 和复标量场 ϕ 的表示

| 场 | $SU(3)_C$ | $SU(2)_L$ | $U(1)_Y$ | $U(1)_X$ |
|------------|-----------|-----------|----------------|----------|
| Q_L, Q_R | 3 | 2 | $\frac{1}{6}$ | 1 |
| L_L, L_R | 1 | 2 | $-\frac{1}{2}$ | 1 |
| ϕ | 1 | 1 | 0 | 1 |

其中 $Q = (U, D)$, $L = (N, E)$ 是类矢量费米子。根据场 ϕ 的表示，可写出如下与 Higgs 场相互作用的标量势。

$$\mathcal{V}(H, \phi) = -\mu_H^2 |H|^2 - \mu_\phi^2 |\phi|^2 + \lambda_H |H|^4 + \lambda_\phi |\phi|^4 + \lambda_{H\phi} |H|^2 |\phi|^2 \quad (3-1)$$

记 v_H , v_ϕ 分别为 Higgs 场和 ϕ 的真空期望值， $\langle \phi \rangle = \frac{v_\phi}{\sqrt{2}}$, m_H , m_ϕ 分别为 Higgs 玻色子和 ϕ 玻色子的质量，则式 3-1 中参数可取为

$$\lambda_H = \frac{1}{4v_H^2} (m_H^2 + m_\phi^2 - \Delta) \quad (3-2)$$

$$\lambda_\phi = \frac{1}{4v_\phi^2} (m_H^2 + m_\phi^2 + \Delta) \quad (3-3)$$

$$\mu_H^2 = \lambda_H v_H^2 + \frac{1}{2} \lambda_{H\phi} v_\phi^2 \quad (3-4)$$

$$\mu_\phi^2 = \lambda_\phi v_\phi^2 + \frac{1}{2} \lambda_{H\phi} v_H^2 \quad (3-5)$$

其中

$$\Delta^2 = (m_\phi^2 - m_H^2)^2 - 4(\lambda_{H\phi} v_H v_\phi)^2 \quad (3-6)$$

上面的等式中， $\lambda_{H\phi}$ 为可调参数。 ϕ 的真空期望值使 $U(1)_X$ 规范对称性自发破缺。相应的 Z' 玻色子获得质量 $m_{Z'} = g_X v_\phi$ ，其中 g_X 为 $U(1)_X$ 的耦合系数。

标量场 Higgs 与 ϕ 的混合角记为 $\theta_{H-\phi}$

$$\tan \theta_{H-\phi} = \frac{1}{2\lambda_{H\phi} v_H v_\phi} (m_\phi^2 - m_H^2 - \Delta) \quad (3-7)$$

如果 $\lambda_{H\phi} \ll 1$ ，则有

$$\theta_{H-\phi} \simeq \frac{\lambda_{H\phi} v_H v_\phi}{m_\phi^2 - m_H^2} \quad (3-8)$$

根据 Q 和 L 的表示，我们可以写出如下 Dirac 质量项以及类矢量粒子与标准模型费米子的 Yukawa 耦合。

$$\mathcal{L} \supset -m_Q \bar{Q}_L Q_R - m_L \bar{L}_L L_R - \lambda_Q^i \bar{Q}_R \phi q^i - \lambda_L^i \bar{L}_R \phi \ell^i + \text{h.c.}, \quad (3-9)$$

这里 ℓ^i 和 q^i 分别表示左手轻子和夸克二重态。下文中， $\lambda_L^{1,2,3}$ 和 $\lambda_Q^{1,2,3}$ 将分别记为 $\lambda_L^{e,\mu,\tau}$ 和 $\lambda_Q^{d,s,b}$ 。有 λ_L^μ 和 $\lambda_Q^{b,s}$ 参与的项提供了 $b \rightarrow s$ 反常。 $\lambda_L^{e,\tau}$ 和 λ_Q^d 设为 0，这样可以消除任何轻子味道破坏 (LFV) 的过程并规避掉标准模型第一代费米子参与任何新物理过程的实验限制。在 ϕ 获得真空期望值以后，费米子的质量矩阵如下。

$$\mathcal{M}_U = \begin{matrix} & U_L & c_L & t_L \\ \begin{matrix} U_R \\ c_R \\ t_R \end{matrix} & \begin{pmatrix} m_Q & \frac{\lambda_Q^s v_\phi}{\sqrt{2}} & \frac{\lambda_Q^b v_\phi}{\sqrt{2}} \\ 0 & \frac{y_c v_H}{\sqrt{2}} & 0 \\ 0 & 0 & \frac{y_t v_H}{\sqrt{2}} \end{pmatrix} \end{matrix} \quad (3-10)$$

$$\mathcal{M}_D = \begin{matrix} & D_L & s_L & b_L \\ \begin{matrix} D_R \\ s_R \\ b_R \end{matrix} & \begin{pmatrix} m_Q & \frac{\lambda_Q^s v_\phi}{\sqrt{2}} & \frac{\lambda_Q^b v_\phi}{\sqrt{2}} \\ 0 & \frac{y_s v_H}{\sqrt{2}} & 0 \\ 0 & 0 & \frac{y_b v_H}{\sqrt{2}} \end{pmatrix} \end{matrix} \quad (3-11)$$

$$\mathcal{M}_N = \begin{matrix} & N_L & \nu_L^\mu \\ N_R & \begin{pmatrix} m_L & \frac{\lambda_L^\mu v_\phi}{\sqrt{2}} \end{pmatrix} \end{matrix} \quad (3-12)$$

$$\mathcal{M}_E = \begin{matrix} & E_L & \mu_L \\ E_R & \begin{pmatrix} m_L & \frac{\lambda_L^\mu v_\phi}{\sqrt{2}} \\ 0 & \frac{y_\mu v_H}{\sqrt{2}} \end{pmatrix} \\ \mu_R & \end{matrix} \quad (3-13)$$

注意到，这里的 c, s, t, b, μ 等为原标准模型下的质量本征态， $y_c, y_s, y_t, y_b, y_\mu$ 等为标准模型的 Yukawa 耦合。 Q_R 与右手夸克的混合大致正比于夸克在标准模型中的 Yukawa 耦合，因而除右手 top 夸克外的混合都被高度压低。同样地， L_R 与右手轻子的混合也被高度压低。 Q_L 与左手夸克的混合以及 L_L 与左手轻子的混合可以很大，因为它们在标准模型规范群中的表示完全相同。尽管如此，由于表示相同，标准模型费米子参与的规范相互作用在新模型中几乎不受影响，电弱精确测量对混合角的大小也几乎没有限制。在 $m_Q \gg y_{c,s,b} v_H$ 和 $m_L \gg y_\mu v_H$ 的极限下，类矢量粒子质量

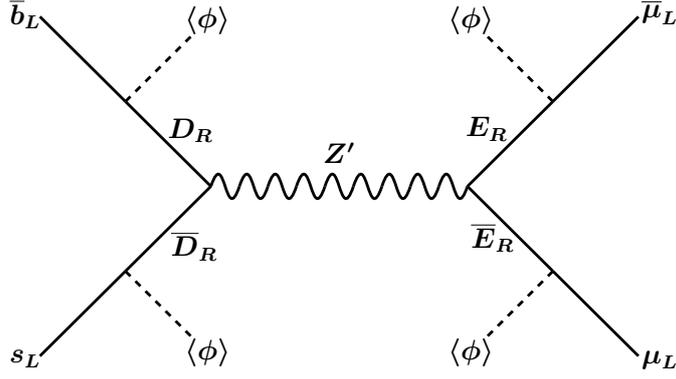
$$m_D = \sqrt{m_Q^2 + \frac{1}{2} \left(|\lambda_Q^b|^2 + |\lambda_Q^s|^2 \right) v_\phi^2} \quad (3-14)$$

$$m_U = \sqrt{m_D^2 + \frac{1}{4m_D^2} \left(|y_t \lambda_Q^b|^2 \right) v_H^2 v_\phi^2} \quad (3-15)$$

$$m_{N,E} = \sqrt{m_L^2 + \frac{1}{2} |\lambda_L^\mu|^2 v_\phi^2} \quad (3-16)$$

标准模型费米子通过与类矢量费米子混合，实现了与 Z' 玻色子的相互作用。诱导 B 反常的有效作用顶点如图3.1所示。用 $g_{f_i f_j}$ 表示 Z' 与费米子 f_i 和 f_j 的耦合系数，相关拉式量可以写作

$$\mathcal{L} \supset g_{f_i f_j} Z'_\mu \bar{f}_i \gamma^\mu P_L f_j, \quad (3-17)$$


 图 3.1 由 Z' 传递的 4 费米相互作用

其中

$$g_{bs} = \frac{\lambda_Q^b \lambda_Q^{s*} v_\phi^2}{2m_D^2} g_X, \quad g_{bb} = \frac{|\lambda_Q^b|^2 v_\phi^2}{2m_D^2} g_X, \quad g_{ss} = \frac{|\lambda_Q^s|^2 v_\phi^2}{2m_D^2} g_X, \quad g_{\mu\mu} = \frac{|\lambda_L^\mu|^2 v_\phi^2}{2m_E^2} g_X \quad (3-18)$$

其余 $g_{f_i f_j}$ 或者为 0, 或者与上述耦合系数近似相等。只要类矢量粒子质量远大于对应的标准模型费米子, 则 $g_{\nu_\mu \nu_\mu} = g_{\mu\mu}$, $g_{cc} = g_{ss}$, $g_{tt} = g_{bb}$, $g_{tc} = g_{bs}$ 。注意到, g_{bb} , g_{ss} 与 g_{bs} 是相关的, 下面的等式永远成立。

$$g_{bb} g_{ss} = |g_{bs}|^2 \quad (3-19)$$

3.2 现有实验对参数空间的限制

有效拉式量

$$\mathcal{H}_{\text{有效}} = -\frac{g_{bs} g_{\mu\mu}}{m_{Z'}^2} (\bar{s} \gamma_\mu P_L b) (\bar{\ell} \gamma^\mu P_L \ell) + \text{h.c.} \quad (3-20)$$

将全局拟合的结果 $C_9^\mu = -C_{10}^\mu = -0.63$ 带入式 (1-16), 并与式 (3-20) 对比可得

$$\frac{m_{Z'}^2}{g_{bs} g_{\mu\mu}} \simeq 947 \text{ TeV}^2 \quad (3-21)$$

除此最佳拟合点之外, 低能标实验对 $U(1)_X$ 模型的参数空间还有两个限制。一个来自 $B_s - \bar{B}_s$ 混合 (或振荡)^[172]。图3.2展示了标准模型中占主导的 top 夸克贡献部分。另一个来自中微子三叉戟产生 (neutrino trident production)^[173]。这种形象的叫法来源于图3.3中的费曼图。

$B_s - \bar{B}_s$ 混合和中微子三叉戟产生的测量结果分别对 Z' 玻色子与 bottom,

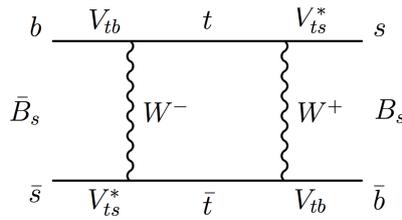


图 3.2 标准模型中 $B_s - \bar{B}_s$ 混合的领头阶贡献，图自文献 [174]。交换 top 夸克与 W 玻色子可得另一费曼图。

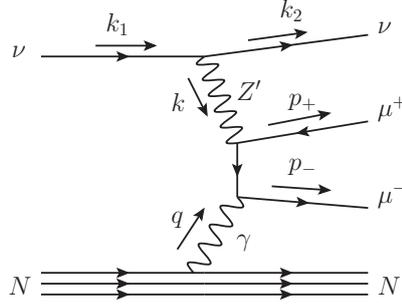


图 3.3 Z' 玻色子对中微子三叉戟产生的领头阶贡献，图自文献 [175]。将 $\mu^+\mu^-$ 调换可得另一费曼图。标准模型中 Z, W 玻色子对此皆有贡献。

strange 夸克和 μ 子的耦合施加了一定限制，

$$\frac{m_{Z'}}{g_{bs}} \gtrsim 244 \text{ TeV} \quad (3-22)$$

$$\frac{m_{Z'}}{g_{\mu\mu}} \gtrsim 0.47 \text{ TeV} \quad (3-23)$$

结合式 (3-22)，式 (3-23) 与式 (3-21)，可得另外两个不等式。

$$\frac{m_{Z'}}{g_{bs}} \lesssim 2015 \text{ TeV} \quad (3-24)$$

$$\frac{m_{Z'}}{g_{\mu\mu}} \lesssim 3.88 \text{ TeV} \quad (3-25)$$

由于 $m_{Z'} = g_X v_\phi$ ， $g_{\mu\mu} < g_X$ ，故

$$v_\phi \lesssim 3.88 \text{ TeV} \quad (3-26)$$

$$m_{Z'} \lesssim 3.88 g_X \text{ TeV} \quad (3-27)$$

根据式 (3-18)，式 (3-25) 和式 (3-26)，得到

$$m_E \lesssim 2.74 |\lambda_L^\mu| \text{ TeV} \quad (3-28)$$

如果加上满足微扰性的要求，则 $\lambda_L^\mu < \sqrt{4\pi}$ 。进一步地，如果要求 g_X 在跑动

到普朗克能标以前不遭遇朗道极点 (Landau pole), 利用 g_X 的一圈 β 函数,

$$\begin{aligned}\beta(g_X) &= \frac{g_X^3}{48\pi^2} Q_X(\phi)^2 + \sum_{f_i} \frac{g_X^3}{12\pi^2} Q_X(f_i)^2 \\ &= \frac{11}{16\pi^2} g_X^3\end{aligned}\quad (3-29)$$

可以得到 g_X 的上限,

$$g_X(m_{Z'}) < \left(\frac{11}{8\pi^2} \log \frac{M_{\text{Pl}}}{m_{Z'}} \right)^{-\frac{1}{2}} \simeq 0.45 \quad (3-30)$$

其中, $M_{\text{Pl}} = 1.22 \times 10^{19} \text{ GeV}$ 为普朗克能标。

综上, 新引入的 Z' 玻色子和类矢量轻子 E 的质量都存在上限。

$$m_{Z'} \lesssim 1.8 \text{ TeV} \quad (3-31)$$

$$m_E \lesssim 9.6 \text{ TeV} \quad (3-32)$$

事实上, 类矢量夸克的质量 m_D 同样存在上限, 不过限制很弱, 大约是

$$m_D \lesssim 77 \text{ TeV} \quad (3-33)$$

注意到, 与上型 (up-type) 类矢量夸克 U 混合后, top 夸克质量会变得稍小于 $\frac{1}{\sqrt{2}} y_t v_H$

$$m_t = \frac{y_t v_H}{\sqrt{2}} \sqrt{1 - \frac{1}{2m_D^2} |\lambda_Q^b|^2 v_\phi^2} \quad (3-34)$$

但是这仍然在目前的实验不确定度之内。

3.3 在 LHC 上寻找类矢量轻子

新粒子存在质量上限, 这为在对撞机上直接搜寻它们奠定了基础。对于 Z' 玻色子, LHC 上最直接的可观测量就是 $\sigma(pp \rightarrow Z') \times \text{BR}(Z' \rightarrow \mu^+ \mu^-)$ 。上一章得到的各种参数限制可以用来约束这一截面的大小。定性来说, 更大的 $g_{f_i f_j}$ ($f_i = c, s, t, b$) 意味着更大的 $\sigma(pp \rightarrow Z')$, 但是由于 Z' 衰变到夸克的截面变大, $\text{BR}(Z' \rightarrow \mu^+ \mu^-)$ 会相应变小, 反之亦然。因而对撞机上通过 Z' 玻色子共振产生的双 μ 事例数, 对 $g_{f_i f_j}$ 并不是简单的依赖关系。一方面, 通过在参数空间随机撒点的方式可以确定出这一截面的大小范围。另一反面, 也可以近似计算出其上下限。下文中以 q_i 指代夸克。根据 QCD 的因子化定理, 质子-质子对撞中, 总产生截面 $\sigma(pp \rightarrow Z')$

可以形式地写成 $q_i q_j \rightarrow Z'$ 的贡献之和,

$$\sum g_{q_i q_j}^2 C_{q_i q_j} \quad (3-35)$$

其中 $C_{q_i q_j}$ 依赖于 $m_{Z'}$ 和质子对撞能量。而分支比 $\text{BR}(Z' \rightarrow \mu^+ \mu^-)$ 由 Z' 玻色子衰变到 q_i 夸克和 μ 子的宽度确定

$$\text{BR}(Z' \rightarrow \mu^+ \mu^-) = \frac{\Gamma(Z' \rightarrow \mu^+ \mu^-)}{\Gamma(Z' \rightarrow \mu^+ \mu^-) + \sum \Gamma(Z' \rightarrow q_i q_j)} \quad (3-36)$$

其中, 除衰变到 top 夸克外, 忽略费米子质量,

$$\Gamma(Z' \rightarrow q_i q_j) = \frac{g_{q_i q_j}^2}{8\pi} m_{Z'} \quad (3-37)$$

$$\Gamma(Z' \rightarrow \mu^+ \mu^-) = \Gamma(Z' \rightarrow \nu_\mu \bar{\nu}_\mu) = \frac{g_{\mu\mu}^2}{24\pi} m_{Z'} \quad (3-38)$$

对于包含 top 夸克的末态,

$$\Gamma(Z' \rightarrow t\bar{t}) = \frac{g_{\mu\mu}^2}{8\pi} m_{Z'} \left(1 - \frac{m_t^2}{m_{Z'}^2}\right) \sqrt{1 - \frac{4m_t^2}{m_{Z'}^2}} \quad (3-39)$$

$$\Gamma(Z' \rightarrow t\bar{c}) = \frac{g_{tc}^2}{8\pi} m_{Z'} \left(1 - \frac{m_t^2}{m_{Z'}^2}\right)^2 \left(1 + \frac{m_t^2}{2m_{Z'}^2}\right) \quad (3-40)$$

对于质量远大于 top 夸克的 Z' 玻色子,

$$\sigma(pp \rightarrow Z' \rightarrow \mu^+ \mu^-) = \sum g_{q_i q_j}^2 C_{q_i q_j} \frac{g_{\mu\mu}^2}{3 \sum g_{q_i q_j}^2 + 2g_{\mu\mu}^2} \quad (3-41)$$

根据上式可以确定 σ 取上下限的近似条件。

上限: $g_{\mu\mu}$, g_{ss} ($= g_{cc}$) 分别取各自的上限。此时, 其余 $g_{q_i q_j}$ 可以忽略, 由于 $g_{\mu\mu} \gg g_{q_i q_j}$, 分支比 $\text{BR}(Z' \rightarrow \mu^+ \mu^-) \simeq 50\%$, 另外 50% 主要衰变到中微子。

下限: 由 $g_{q_i q_j}$ 之间的相互关系和均值不等式以及式 (3-24) 可知

$$\sigma(pp \rightarrow Z') = g_{bs}^2 \left(C_1 + C_2 t^2 + \frac{C_3}{t^2} \right) \gtrsim (C_1 + 2\sqrt{C_2 C_3}) \left(\frac{m_{Z'}}{2015 \text{ TeV}} \right)^2$$

其中 $C_1 = C_{bs} + C_{tc}$, $C_2 = C_{tt} + C_{bb}$, $C_3 = C_{cc} + C_{ss}$, $t = g_{bs}/g_{ss} = g_{bb}/g_{bs}$ 。此时, 仍然有 $g_{\mu\mu} \gg g_{q_i q_j}$, $\text{BR}(Z' \rightarrow \mu^+ \mu^-) \simeq 50\%$ 。

上述估计也得到了撒点法的印证, 最终结果见图3.4。

ATLAS 合作组和 CMS 合作组都做过利用双轻子道在 LHC 上搜寻 Z' 玻色子的分析^[176-177]。其中 ATLAS 公开给出的是综合双轻子道的排除线, 而 CMS 单独给出了双 μ 子道的结果, 因而这里考虑 CMS 的结果。利用 LHC 运行在 13 TeV 积

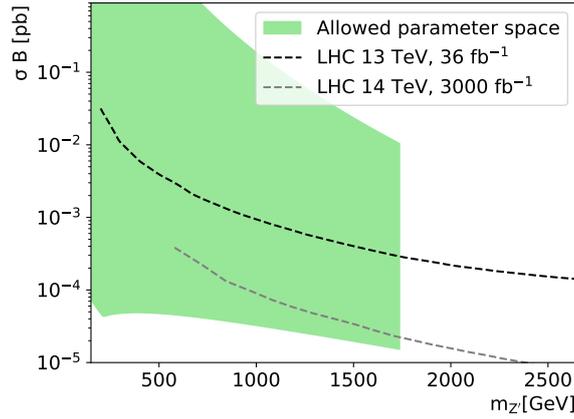


图 3.4 $\sigma(pp \rightarrow Z' \rightarrow \mu^+ \mu^-)$ 的范围，用绿色阴影表示。黑色虚线表示 CMS 合作组给出的 95% 置信度排除线，根据 LHC 运行在 13 TeV 的 36 fb^{-1} 数据得出。灰色虚线为外推至 14 TeV 和 3000 fb^{-1} 的结果。

累的 36 fb^{-1} 数据得到的排除线见图 3.4 中黑色虚线。刚才估计得到的截面上限已被排除，而下限还远未达到。利用这一条排除线，我们可以进一步地约束 $g_{q_i q_j}$ 。当 $g_{\mu\mu}$ 达到式 (3-25) 允许的下界时，较大的 g_{bb} 或者 g_{ss} 将使双 μ 子的产生截面超出排除线，据此可得 g_{bb} 和 g_{ss} 的上界。考虑到它们与 g_{bs} 的关联，即 $g_{bb}g_{ss} = g_{bs}^2$ ，直接可得其下界。完整结果在图 3.5 中展示，同时展示的还有 Z' 玻色子衰变分支比的上下界。当 $g_{\mu\mu}^2 \gg 3 \max(g_{bb}^2, g_{ss}^2)$ 时， Z' 玻色子主要衰变至轻子。由图可知，在 $m_{Z'} < 1.5 \text{ TeV}$ 时，轻子衰变道占比至少为 90%。

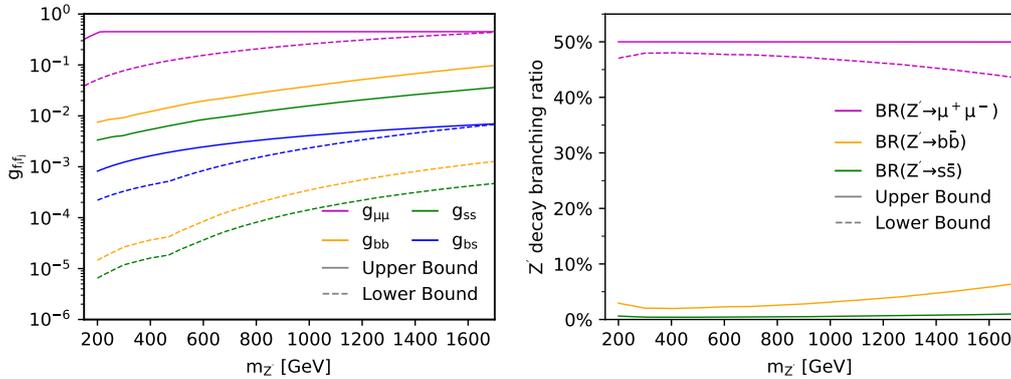
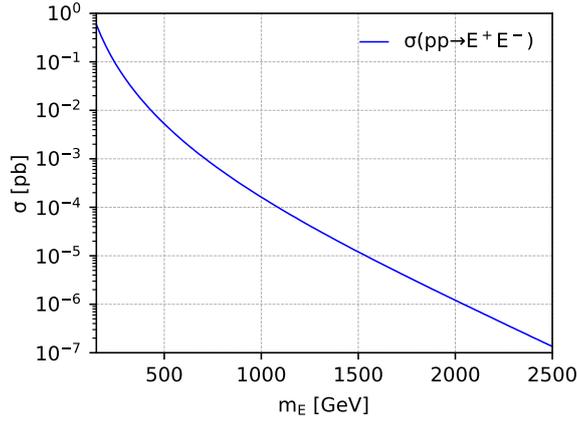


图 3.5 左： Z' 玻色子与标准模型费米子的耦合所满足的限制。右： Z' 玻色子衰变分支比所满足的限制。实线和虚线分别表示上下界。注意到 g_{bb} ， g_{ss} ， g_{bs} 不是相互独立的，不能同时达到上界或下界。右图中，由于 $\text{BR}(Z' \rightarrow b\bar{s})$ 的上下界与 $\text{BR}(Z' \rightarrow b\bar{b})$ ， $\text{BR}(Z' \rightarrow s\bar{s})$ 的下界远远小于 1%，故未示出。

为了探究升级后的 LHC，即高亮度 LHC (HL-LHC) 在双 μ 子道搜寻 Z' 玻色子的前景如何，我们使用一种专门为共振搜寻设计的外推方法^[178-179] 将排除线外推至对撞能量 14 TeV，积分亮度 3000 fb^{-1} 。结果见图 3.4 中的灰色虚线，HL-LHC 对双 μ 子道搜寻 Z' 玻色子的灵敏度可提升 1 个数量级。要注意的是，外推结果只


 图 3.6 HL-LHC 上带电的类矢量轻子 E^\pm 成对产生截面。

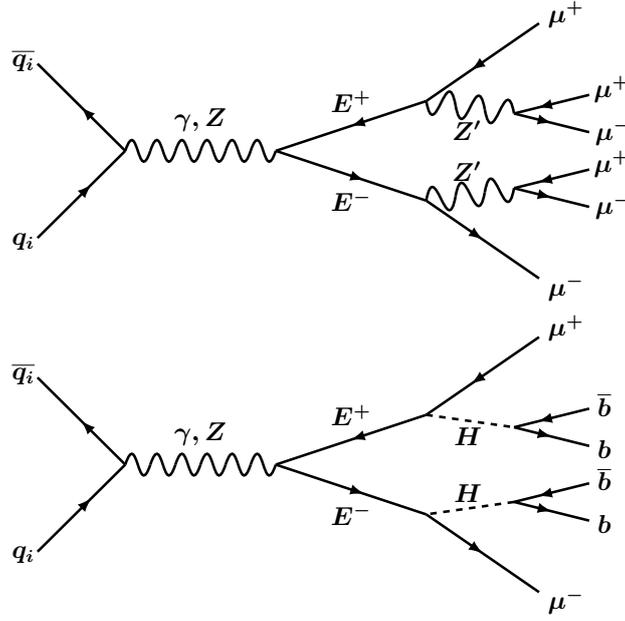
能作为未来对撞机搜寻灵敏度的粗率估计。根据文献 [178]，外推所得数据与完整分析的差距在两倍以内。尽管如此，必然有大片参数空间无法覆盖。特别是，即使很轻的 Z' 玻色子，比如几百 GeV，也有可能避开共振峰的搜寻。所以需要其他策略来验证模型，我们指出，对带电的类矢量轻子 E^\pm 的寻找可以作为一种补充的手段。

首先是 E^\pm 的产生截面较大。如前面提到的，类矢量轻子与左手轻子处于相同的标准模型表示，因而具有同样的产生道，在 LHC 上可以成对产生，其截面只被质量压低。在 LHC 对撞能量 14 TeV 时， E^\pm 通过 Drell-Yan 过程成对产生的截面大小如图 3.6 所示。目前，LHC 在对撞能量 13 TeV 积累了 160 fb^{-1} 的总数据量。考虑到对撞能量由 13 TeV 提升到 14 TeV 后， E^\pm 产生截面只有小幅增长，由图 3.6 可以大致推算，在积分亮度提高约 20 倍的情况下，相同事例数对应的 m_E 提高了约 500 GeV。与图 3.4 对比可知，在 $m_{Z'}$ ， m_E 同属亚 TeV 量级时，搜寻 E^\pm 会更有优势。当然，双类矢量轻子的衰变末态信号会更加复杂，需要进一步分析以证明其优势。

3.3.1 衰变道

类矢量轻子的寻找曾在多种理论框架下被研究过。本文的特别之处在于，利用 B 物理实验等给出的限制和 $U(1)_X$ 模型本身的约束，搜寻的衰变道不同以往，且具有很强的针对性。一般假设类矢量轻子衰变有 Zl ， $W\nu_l$ 和 Hl 等几个道，其中 $l = \mu, \tau$ 。它们大部分在我们的模型中不存在或分支比极低。 $U(1)_X$ 模型给出描述 E^\pm 与 W ， Z 玻色子相互作用的以下拉式量。

$$\mathcal{L} \supset g \frac{\lambda_L^\mu v_\phi}{4m_E} \left(\frac{y_\mu v}{\cos\theta_W m_E} Z_\mu \bar{\mu}_R \gamma^\mu E_R + \frac{(y_\mu v)^2 m_L}{m_E^3} W_\mu^+ \bar{\nu}_\mu \gamma^\mu E_L \right) + O(y_\mu^3) + \text{h.c.} \quad (3-42)$$


 图 3.7 LHC 上搜寻类矢量轻子 E^\pm 两个主要的道。

$E^\pm \rightarrow \mu^\pm Z$ 衰变宽度正比于 y_μ^2 , $E^\pm \rightarrow \nu W^\pm$ 衰变宽度正比于 y_μ^4 , 因而这两个衰变道都被显著压低。另外, $E^\pm \rightarrow NW^\pm$ 的两体衰变相空间不存在, 因为类矢量轻子二重态的质量差与 m_μ 同一量级。

$$\Delta m_L = \frac{\lambda_L^\mu v_\phi y_\mu v}{2m_E} = \frac{\lambda_L^\mu v_\phi}{\sqrt{2}m_L} m_\mu \quad (3-43)$$

真正值得注意的衰变道只剩 $E^\pm \rightarrow \mu^\pm H$ 和 $E^\pm \rightarrow \mu^\pm Z'^{(*)}$ 。其中 Higgs 玻色子随后主要衰变至 bottom 夸克对, 分支比约为 58%。而 Z' 玻色子, 在壳或者不在壳, 随后都主要衰变至双 μ 子 (双中微子具有同样分支比, 但是不利于对撞机搜寻), $m_{Z'} < 1.5 \text{ TeV}$ 时, 分支比至少为 45%。通过 $E^\pm \rightarrow \mu^\pm Z'^{(*)} (\rightarrow \mu^+ \mu^-)$ 这一过程, 对撞机上可产生多达 6 μ 子的末态, 即成对产生的类矢量轻子 $E^- E^+$ 都通过 Z' 玻色子衰变到 μ 子。这一信号基本没有来自标准模型的本底^[180], 即使 $\mathcal{O}(10)$ 的事例数也具有非常高的信号显著性。另一种对称衰变的模式是 $E^- E^+$ 都通过 Higgs 玻色子衰变到 bottom 夸克和 μ 子。两个衰变道对应的费曼图如图 3.7 所示。来自 s 道 Z' 玻色子的贡献被忽略了, 理由是, Z' 玻色子与夸克的耦合强度远远小于光子 γ 和 Z 玻色子, 除非 $m_{Z'}$ 非常大, 而非常大的 $m_{Z'}$ 会显著压低 Drell-Yan 过程的振幅。

我们的 $U(1)_X$ 模型, 引入了额外的标量粒子 ϕ 。理论上可能存在 $E^\pm \rightarrow \mu^\pm \phi$ 这一衰变道。 m_ϕ 在模型中为自由参数, 通常假设 $m_\phi > m_E$, 这个道不打开, 而 $E^\pm \rightarrow \mu^\pm \phi^* (\rightarrow f_i f_j)$ 三体衰变宽度正比于 $\sin^2 \theta_{H-\phi}^4$ 。即使 $m_\phi < m_E$, 这个道只会

给出与 $E^\pm \rightarrow \mu^\pm H$ 类似的信号。因而不考虑 $E^\pm \rightarrow \mu^\pm \phi$ 的贡献。 $E^\pm \rightarrow W^{\pm(*)} N$ 或 $E^\pm \rightarrow W^\pm N^{(*)}$ 也可以通过三体衰变参与进来，但是它们的末态 $f_i f_j N$ 或 $W^\pm H \nu_\mu$ 相对 3μ 子末态，相空间被高度压低，因而也不在计算范围内。

下面计算两个衰变道的分支比。首先是两体衰变 $E^\pm \rightarrow \mu^\pm Z'$ 的宽度

$$\Gamma(E^\pm \rightarrow \mu^\pm Z') = \frac{g_{Z'}^2}{32\pi m_E^3 m_{Z'}^2} (m_E^2 + 2m_{Z'}^2)(m_E^2 - m_{Z'}^2)^2 \quad (3-44)$$

其中

$$g_{Z'} = \frac{\lambda_L^\mu m_L m_{Z'}}{\sqrt{2} m_E^2} \quad (3-45)$$

然后是两体衰变 $E^\pm \rightarrow \mu^\pm H$ 的宽度，

$$\Gamma(E^\pm \rightarrow \mu^\pm H) = \frac{g_H^2}{32\pi m_E^3} (m_E^2 - m_H^2)^2 \quad (3-46)$$

其中

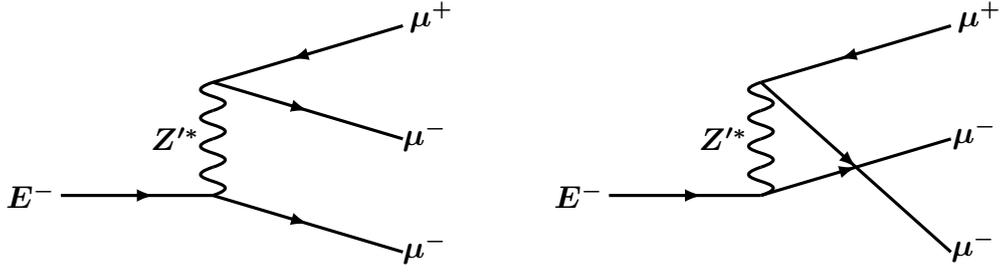
$$g_H = \frac{\lambda_L^\mu}{\sqrt{2}} \sin \theta_{H-\phi} \frac{m_L}{m_E} \quad (3-47)$$

注意到，尽管系数中都有 m_L ，但是它只影响到 E^\pm 衰变的总宽度，对分支比没有影响。

接着是三体衰变 $E^\pm \rightarrow \mu^\pm Z'^* \rightarrow \mu^\pm f_i f_j$ 的宽度

$$\begin{aligned} \Gamma(E^\pm \rightarrow \mu^\pm Z'^* \rightarrow \mu^\pm f_i f_j) &= \frac{g_{Z'}^2 (2g_{\mu\mu}^2 + 3 \sum g_{q_i q_j}^2)}{768\pi^3 m_E^3 m_{Z'} \Gamma_{Z'}} \\ &\left\{ [(m_E^2 + 2m_{Z'}^2)(m_E^2 - m_{Z'}^2)^2 + 3(m_E^2 - 2m_{Z'}^2)m_{Z'}^2 \Gamma_{Z'}^2] \left(\arctan \frac{m_E^2 - m_{Z'}^2}{m_{Z'} \Gamma_{Z'}} + \arctan \frac{m_{Z'}}{\Gamma_{Z'}} \right) \right. \\ &\left. + m_{Z'} \Gamma_{Z'} \left[-2m_E^2(m_E^2 - 2m_{Z'}^2) + m_{Z'}^2 (3m_E^2 - 3m_{Z'}^2 + \Gamma_{Z'}^2) \ln \frac{m_{Z'}^2 (m_{Z'}^2 + \Gamma_{Z'}^2)}{(m_E^2 - m_{Z'}^2)^2 + m_{Z'}^2 \Gamma_{Z'}^2} \right] \right\} \\ &+ \Delta\Gamma(E^\pm \rightarrow 3\mu) \end{aligned} \quad (3-48)$$

其中 $\Delta\Gamma(E^\pm \rightarrow 3\mu)$ 是来自于 $E^\pm \rightarrow 3\mu$ 的额外贡献。 $E^\pm \rightarrow \mu^\pm f_i f_j$ 等过程，除了 f_i 代表 μ 子的情况下，只存在 1 个树图。而 $E^\pm \rightarrow 3\mu$ 存在 2 个树图，如图 3.8 所示。形式上， $\Delta\Gamma(E^\pm \rightarrow 3\mu)$ 来自图 3.8 的右图振幅模平方与左右两图干涉项之和。


 图 3.8 $E^\pm \rightarrow 3\mu$ 的费曼图, $E^\pm \rightarrow \mu^\pm f_i f_j$ 只有类似左边的 1 个图。

$$\begin{aligned} \Delta\Gamma(E^\pm \rightarrow 3\mu) &= \frac{g_{Z'}^2 g_{\mu\mu}^2}{256\pi^3 m_E^3} \\ &\int_{m_{Z'}^2 - m_E^2}^{m_{Z'}^2} \frac{1}{x^2 + m_{Z'}^2 \Gamma_{Z'}^2} \left\{ (x + m_E^2 - m_{Z'}^2) [x(3x + m_E^2 - 5m_{Z'}^2) + 2m_{Z'}^2 \Gamma_{Z'}^2] \right. \\ &- 2m_{Z'} \Gamma_{Z'} [x^2 + m_{Z'}^2 (2m_E^2 - 4m_{Z'}^2 + \Gamma_{Z'}^2)] \left(\arctan \frac{x + m_E^2 - 2m_{Z'}^2}{m_{Z'} \Gamma_{Z'}} + \arctan \frac{m_{Z'}}{\Gamma_{Z'}} \right) \\ &+ [x(x - 2m_{Z'}^2)(x + m_E^2 - 2m_{Z'}^2) + m_{Z'}^2 \Gamma_{Z'}^2 (x + m_E^2 - 4m_{Z'}^2)] \\ &\left. \ln \frac{m_{Z'}^2 (m_{Z'}^2 + \Gamma_{Z'}^2)}{(x + m_E^2 - 2m_{Z'}^2)^2 + m_{Z'}^2 \Gamma_{Z'}^2} \right\} dx \end{aligned} \quad (3-49)$$

上面式子中 $\Gamma_{Z'}$ 的计算只考虑了直接衰变道标准模型粒子且忽略了包括 top 夸克在内的费米子质量。

$$\Gamma_{Z'} = \frac{m_{Z'}}{24\pi} (2g_{\mu\mu}^2 + 3 \sum g_{q_i q_j}^2) \quad (3-50)$$

在 $m_{Z'} < m_E$ 的情况下, 式 (3-48) 通过窄宽度近似 (narrow-width approximation) $\Gamma_{Z'} \ll m_{Z'}$, 退化到式 (3-44)。

6μ 子的总截面由 m_E , $m_{Z'}$, $g_{f_i f_j}$ 和混合角 $\theta_{H-\phi}$ 控制。由于在我们感兴趣的参数空间范围内, $g_{\mu\mu} \gg g_{q_i q_j}$, 事实上截面几乎不受 $g_{q_i q_j}$ 影响。 $\theta_{H-\phi}$ 越大, 则 $E^\pm \rightarrow \mu^\pm H$ 截面越大, $E^\pm \rightarrow \mu^\pm Z'$ 分支比越低, 因而 6μ 子的总截面与 $\theta_{H-\phi}$ 负相关。注意到图3.5是以 $m_{Z'}$ 为单一变量的。当考虑类矢量轻子质量变化时, 根据式 (3-18), $g_{\mu\mu}$ 还受以下不等式控制。

$$g_{\mu\mu} = \frac{|\lambda_L^\mu| m_{Z'} \sqrt{m_E^2 - m_L^2}}{\sqrt{2} m_E^2} < \frac{\sqrt{2\pi} m_{Z'}}{m_E} \quad (3-51)$$

图3.9左半部分展示了在对撞能量 14 TeV, 积分亮度 3000 fb^{-1} 的 LHC 上, 对于不同的混合角 $\theta_{H-\phi}$, 预期的 6μ 子事例数。实线和虚线表示事例数上界和下界的等值线。事例数的上下界显然都高度依赖于 m_E 。一个引人注意的特点是, Z' 玻

色子在壳的情况下，事例数几乎不随 $m_{Z'}$ 变化，因为此时带电的类矢量轻子对 E^\pm 几乎全部衰变到轻子末态，其中约 25% 为 6μ 子，约 75% 为 4μ 子加 2μ 子中微子 ν_μ ，约 25% 为 2μ 子加 4μ 子中微子 ν_μ ，其余为强子衰变。一旦 Z' 玻色子不在壳，事例数会突然下降，这显然是由于三体衰变的相空间压低。事例数的下界在 $m_E < m_{Z'}$ 时下降的尤其明显，表明在此区域 $E^\pm \rightarrow \mu^\pm Z'^*$ 道的搜寻灵敏度不足，需要其它方式来补充。

图3.9展示的结果并没有考虑部分子簇射和探测器响应等效应。对于轻子末态，簇射效应十分微弱。我们这样估算 CMS 探测器识别 6μ 子末态的效率：

1) 对每一个满足 $p_T > 10 \text{ GeV}$ ， $|\eta| < 2.4$ 的 μ 子，识别率为 95%，其余情形为 0。

2) 要求任意两个 μ 子的间距 $\Delta R > 0.4$ ，否则认定为不识别。

图3.10展示了由这两条筛选条件确定的最终识别效率。这一筛选过程利用 MADGRAPH5_AMC@NLO 2.6.0^[181] 完成，简化了部分参数，使得 Z' 的衰变宽度固定为 $\Gamma_{Z'}/m_{Z'} = 1\%$ 。大体上 40% - 50% 的总事例数可以通过筛选。考虑到图3.9的上下界，以及极低的本底污染。预计在满足 $m_{Z'} < m_E$ 的参数空间范围内，质量高达 1400 GeV 的类矢量轻子可被探测到或者排除掉。

图3.9的右半部分展示了分支比 $\text{BR}(E^\pm \rightarrow \mu^\pm H)$ 的上下界。在 $m_{Z'} < m_E$ 的区域，衰变由 $E^\pm \rightarrow \mu^\pm Z'$ 主导，因而 $\text{BR}(E^\pm \rightarrow \mu^\pm H)$ 的上下界几乎重合。整体而言， $\text{BR}(E^\pm \rightarrow \mu^\pm H)$ 与比值 $m_{Z'}/m_E$ 和混合角 $\theta_{H-\phi}$ 正相关。目前与 Higgs 玻色子相关的精确测量任然允许有一定大小的混合角 $|\sin \theta_{H-\phi}| \lesssim 0.3$ ^[182]。假如标量场部分混合较大，比如 $\sin \theta_{H-\phi} > 0.1$ ，而且 Z' 不在壳，即 $m_{Z'}/m_E > 1$ ，那么 $E^\pm \rightarrow \mu^\pm H$ 的衰变模式将压倒 $E^\pm \rightarrow \mu^\pm Z'^*$ 成为主导。数值上，即使 $\sin \theta_{H-\phi} = 0.05$ ， $m_{Z'}/m_E = 1$ ， $E^\pm \rightarrow \mu^\pm H$ 仍有不少于 50% 的分支比。在下一节，我们将仔细分析这一衰变模式的对撞机信号。

在上文的分析中，我们从理论的微扰性出发，要求 $g_X < 0.45$ ， $|\lambda_Q^{s,b}|, |\lambda_L^u| < \sqrt{4\pi}$ 。原则上，这些条件可以放松或进一步加强，比如有些文章中假设 Yukawa 耦合普遍小于 1。这些参数上界的变化会改变前述的一部分结果。首先，新引入粒子质量的上界对耦合系数的上界是线性依赖的， $m_{Z'} \propto g_X$ ， $m_E \propto |\lambda_L^u|$ ， $m_D \propto \sqrt{|\lambda_Q^s \lambda_Q^b|}$ 。除此之外，改变 $\lambda_Q^{s,b}$ 的上界没有影响。而改变 λ_L^u 的上界会影响不等式 (3-51)，从而改变图3.9。变化发生在 $m_{Z'} < \frac{\sqrt{2}g_X}{\lambda_L^u} m_E$ 的有限区域内，而且无足轻重，不改变主要结论。提高 g_X 的上界对图3.9的影响会更加显著，改变主要发生在 $m_{Z'} > m_E$ 的区域，此处 $E^\pm \rightarrow \mu^\pm Z'^*$ 的振幅得到增强，可能会削弱 $E^\pm \rightarrow \mu^\pm H$ 这一搜寻道的灵敏度。

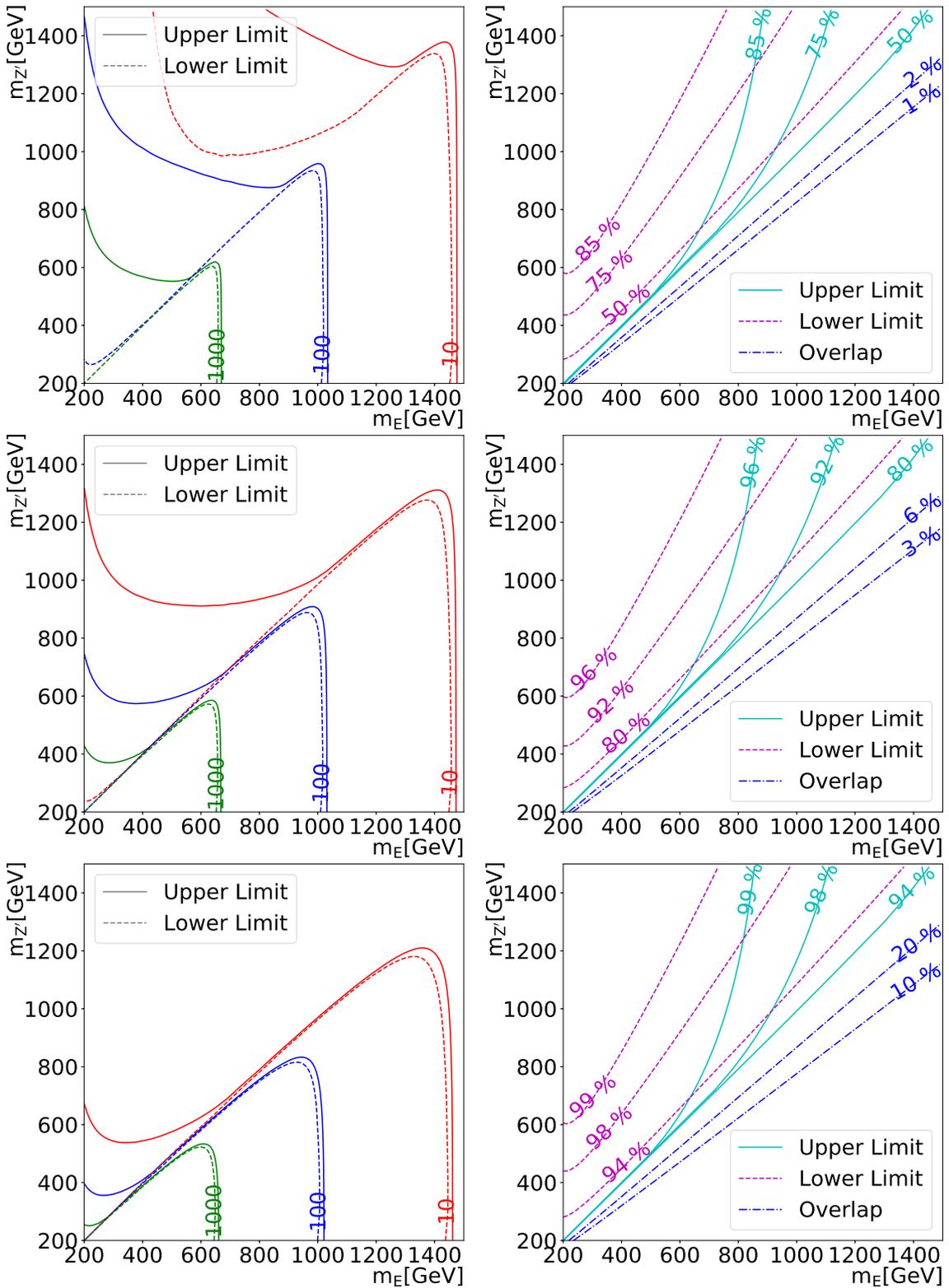
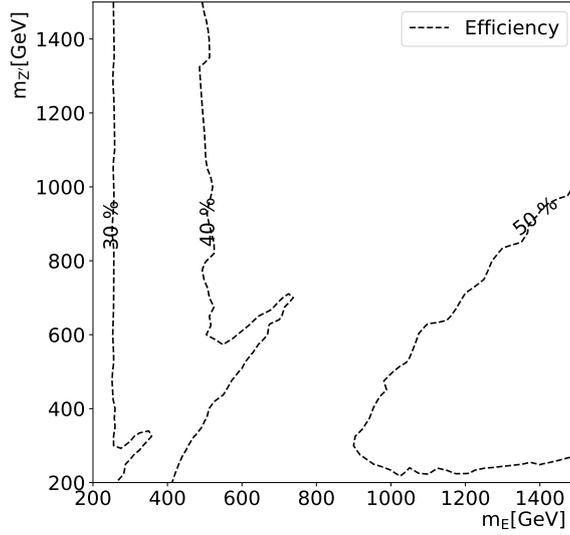


图 3.9 左：预期的 6μ 子事例数。右：预期的分支比 $BR(E^\pm \rightarrow \mu^\pm H)$ 。实线和虚线分别表示上界和下界的等值线，点划线表示两者几乎重合。从上至下，标量场的混合角分别设定为 $\sin\theta_{H-\phi} = 0.05, 0.1, 0.2$ 。由于图中显示区域 $g_{\mu\mu} \gg g_{q_i q_j}$ ，基本上，左边的上界对应右边的下界，反之亦然。注意到，这一说法并不严格成立，尤其在 $m_{Z'}$ 较大区域。


 图 3.10 6μ 事例筛选效率。

3.3.2 信号与本底模拟

信号和本底事例的生成使用 MADGRAPH5_AMC@NLO 2.6.0，其中类矢量轻子的和 Higgs 玻色子的衰变由 MADSPIN^[183] 处理，然后使用 PYTHIA8^[184] 完成部分子簇射，强子化和强子的衰变，最后使用 DELPHES 3.4.0^[185] 对探测器的响应进行快速模拟。探测器模拟部分选用了 ATLAS 配置， b -tag 的效率设定为 70%，对 charm 夸克和轻夸克喷注的误识别率分别设定为 15% 和 0.8%^[186]。喷注重建由 FASTJET 3.2.1^[187] 负责。信号选为 $m_E \in [150, 1500]$ ，间隔 25 GeV，共计 55 个基准点。这一信号的主要本底过程包括 $t\bar{t}$ ， $t\bar{t}b\bar{b}$ ， $t\bar{t}H$ 和 $t\bar{t}Z$ ，它们在 14 TeV 的 LHC 上精确到次领头阶的产生截面^[188-191] 如表 3.2 所示。

表 3.2 14TeV LHC 上本底过程截面

| 本底过程 | $t\bar{t}$ | $t\bar{t}b\bar{b}$ | $t\bar{t}H$ | $t\bar{t}Z$ |
|---------|------------|--------------------|-------------|-------------|
| 截面 [pb] | 933 | 2.636 | 0.611 | 1.121 |

3.3.3 事例重建与选取

我们使用了两种方式重建 Higgs 玻色子，分别适合不同的相空间。

1) 使用 anti- k_t 算法重建喷注，半径 $R = 0.4$ 。要求横向动量 $p_T > 20$ GeV 的喷注至少有 4 个，其中至少 3 个被 b -tag。将这三个 b 喷注加上其余喷注中 p_T 最大的一个进行两两组合，构成两个备选 Higgs 喷注。设两个喷注的质量分别为 m_{H_1} ， m_{H_2} 。则质量不对称性最小的组合定义为最终的 Higgs 喷注。不对称性 A 由下式

衡量。

$$A = \frac{m_{H_1} - m_{H_2}}{m_{H_1} + m_{H_2}} \quad (3-52)$$

这种方式重建的 Higgs 玻色子记为“标准 Higgs”。

2) 使用 C/A 算法重建胖喷注，半径 $R = 1.4$ 。要求胖喷注满足 Mass-Drop Tagger^[82]，同时包含至少 1 个被 b -tag 的子喷注。这种方式重建的 Higgs 玻色子记为“子结构 Higgs”。

前者适用于重建能量较低的 Higgs 玻色子，如 $p_T < 300 \text{ GeV}$ ，因为此时衰变成 bottom 夸克对分开的间距较大。后者适用于重建能量较高的 Higgs 玻色子，或称 boosted Higgs 玻色子。此时由于 bottom 夸克对和辐射出来的胶子都分布在一个夹角较小的锥体内，难以将子喷注很好的剥离。重建好的 Higgs 玻色子随即与末态中的两个 μ 子进行两两组合。与上述重建标准 Higgs 的方式类似，质量不对称性最小的组合定义为最终的类矢量轻子。

除了重建的类矢量轻子对的质量，双 μ 子系统的超横向质量 (stransverse mass) 也是区分信号与本底的有力判据^[192]，其定义如下：

$$m_{T2}(\mu_1, \mu_2) \equiv \min_{\mathbf{p}_{T1} + \mathbf{p}_{T2} = \sum \mathbf{p}_T^j} [\max(m_T(\mathbf{p}(\mu_1), \mathbf{p}_{T1}), m_T(\mathbf{p}(\mu_2), \mathbf{p}_{T2})))] \quad (3-53)$$

其中，横向质量 m_T 的定义为

$$m_T^2(\mathbf{p}(\mu_i), \mathbf{p}_{Ti}) = (E(\mu_i) + \sqrt{\mathbf{p}_{Ti}^2 + m_H^2})^2 - (\mathbf{p}(\mu_i) + \mathbf{p}_{Ti})^2 \quad (3-54)$$

式 (3-54) 中 $m_H = 125 \text{ GeV}$ ，式 (3-53) 中的指标 j 跑遍所有 Higgs 喷注的组分粒子。

假设一种新粒子 Λ 成对产生，它的衰变产物是一个可见粒子和一个不可见粒子 χ ， m_{T2} 这一变量原本就是为测量 Λ 这类新粒子的质量而设计的。其分布特点是

$$m_\chi \leq m_{T2} \leq m_\Lambda \quad (3-55)$$

对于 $E^- E^+ \rightarrow \mu^- \mu^+ H H$ 这一过程，由于 Higgs 玻色子重建的效率和精度不佳，我们可以把 Higgs 玻色子当作不可见粒子，从而利用 m_{T2} 的分布特性区分信号与本底。

图3.11展示了几种有代表性的变量分布。我们发现：

1) 对于质量比较重的类矢量轻子，使用子结构方法可以更有效的重建 Higgs 玻色子。大部分本底过程中的假 Higgs 喷注质量小于 125 GeV 。

2) m_{T2} 变量分布特点显著, m_H 为分布的下界。对于本底, 由于构成变量的粒子是由 **top** 夸克衰变而成, m_{T2} 小于 $m_H + m_t$ 。对于信号, m_{T2} 的上界为类矢量轻子质量, 可以远大于 $m_H + m_t$ 。这一特性使 m_{T2} 可以高效地区分信号与本底。

3) 对于较重的类矢量轻子, 信号过程能标远大于本底过程, 导致 Higgs 玻色子和类矢量轻子的横向动量分布趋向高能。

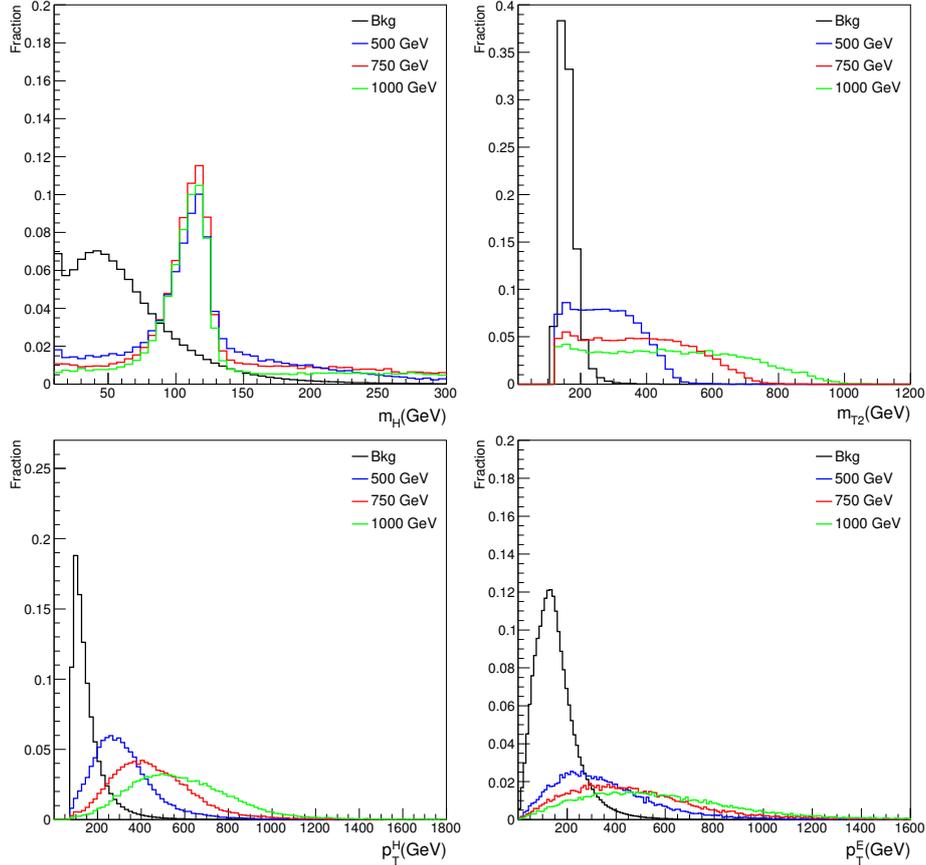


图 3.11 本底与信号事例的各变量分布, 选取的信号中 m_E 分别为 500, 750, 1000 GeV。从左到右, 从上到下分别为 Higgs 玻色子质量, 双 μ 子的 m_{T2} , Higgs 玻色子横向动量和类矢量轻子横向动量。其中的 Higgs 玻色子均为“子结构 Higgs”。展示的都来源于粒子对中较硬的一个。

事例筛选过程如下。

1) 要求末态至少有 4 个喷注和 2 个符号相异且赝快度 $|\eta| < 2.5$ 的 μ 子。这里的喷注为半径 $R = 0.4$, 横向动量 $p_T > 20$ GeV, 赝快度 $|\eta| < 2.5$ 的 **anti- k_t** 喷注。定义 3 个筛选 μ 子的信号区间 $SR_i, i=1, 2, 3$, 如表 3.3 所示, 分别适合一定的类矢量轻子质量范围。

2) 要求双 μ 子的 $m_{T2} > 300$ GeV。要求存在 2 个重建的“标准 Higgs”或者“子结构 Higgs”, 满足 $90 \text{ GeV} \leq m_H \leq 130 \text{ GeV}$ 。2 个 Higgs 玻色子应包含至少 3 个被 **b-tag** 的子喷注。在上述每个信号区间基础上, 再定义 2 个筛选 Higgs 玻色子

表 3.3 筛选 μ 子的 3 个信号区间

| 信号区间 | SR1 | SR2 | SR3 |
|------------|------------------------|-------------------------|-------------------------|
| 主要 μ 子 | $p_T > 80 \text{ GeV}$ | $p_T > 150 \text{ GeV}$ | $p_T > 250 \text{ GeV}$ |
| 次要 μ 子 | $p_T > 30 \text{ GeV}$ | $p_T > 80 \text{ GeV}$ | $p_T > 150 \text{ GeV}$ |

横向动量的信号区间 SR_{ij} , $j=1, 2$, 如表3.4所示。

表 3.4 筛选 Higgs 玻色子横向动量的 2 个信号区间

| 信号区间 | SR _{i1} | SR _{i2} |
|--------------|-------------------------|-------------------------|
| 主要 Higgs 玻色子 | $p_T > 200 \text{ GeV}$ | $p_T > 350 \text{ GeV}$ |
| 次要 Higgs 玻色子 | $p_T > 150 \text{ GeV}$ | $p_T > 250 \text{ GeV}$ |

3) 对上述每个信号区间, 根据重建 Higgs 玻色子的方法和类矢量轻子的质量定义 4 个信号区间 SR_{ijk} , $k=1, 2, 3, 4$, 如表3.5所示。

表 3.5 筛选类矢量轻子质量的 4 个信号区间

| 信号区间 | SR _{ij1} | SR _{ij2} | SR _{ij3} | SR _{ij4} |
|-----------|-------------------------|-------------------------|-------------------------|-------------------------|
| Higgs 类型 | 子结构 Higgs | 子结构 Higgs | 子结构 Higgs | 标准 Higgs |
| 主要类矢量轻子质量 | $m_E > 350 \text{ GeV}$ | $m_E > 450 \text{ GeV}$ | $m_E > 650 \text{ GeV}$ | $m_E < 300 \text{ GeV}$ |

55 个信号过程和 4 种事例组成的本底过程分别经过 24 个信号区间的筛选得到信号事例数 s 和本底事例数 b , 信号显著性由下式计算^[193]。

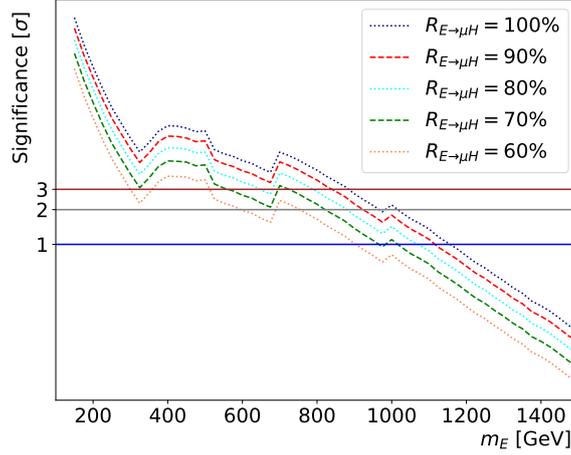
$$\mathcal{S} = \sqrt{2((s+b) \ln(1 + \frac{s}{b}) - s)} \quad (3-56)$$

对每一个给定 m_E , 选取信号显著性最大值对应的信号区间作为搜寻方案。表3.6列出了 3 个基准点的筛选流程和最终选用的信号区间。搜寻 500, 750, 1000 GeV 的类矢量轻子使用的筛选方案分别属于信号区间 SR212, SR322 和 SR323。表中所列数字为 14TeV LHC 上的截面, 单位 fb。表中结果的前提是分支比 $BR(E^\pm \rightarrow \mu^\pm H)=100\%$ 。

图3.12展示了信号显著性随类矢量轻子的质量和衰变分支比 $BR(E^\pm \rightarrow \mu^\pm H)$ 的变化。注意到, 若分支比由 100% 减少至 r , 则信号事例数减少至原本的 r^2 倍。由于质量相邻的信号筛选条件可能分属不同信号区间, 曲线呈现了一定间断。可以看到, 对撞能量 14 TeV, 积分亮度 3000 fb^{-1} 的 LHC 对 1 TeV 的类矢量轻子至多有 2σ 的搜寻灵敏度。

表 3.6 3 个基准点上信号与本底的筛选流程

| m_E [GeV] | 500 | 本底 (SR212) | 750 | 本底 (SR322) | 1000 | 本底 (SR323) |
|--|-------|------------|-------|------------|--------|------------|
| $N_{\text{喷注}} \geq 4 \ \& \ N_{\mu} \geq 2$ | 0.98 | 419 | 0.15 | 61 | 0.32 | 61 |
| $m_{T2} \geq 300 \text{ GeV}$ | 0.44 | 15 | 0.10 | 5.3 | 0.024 | 5.3 |
| $N_H = 2, \ b\text{-tag} \geq 3$ | 0.073 | 0.16 | 0.020 | 0.040 | 0.0060 | 0.040 |
| m_E | 0.073 | 0.13 | 0.020 | 0.033 | 0.0059 | 0.019 |


 图 3.12 使用 $E^- E^+ \rightarrow \mu^- \mu^+ HH$ 这一过程在 HL-LHC 上搜寻类矢量轻子的信号显著性随质量和分支比的变化。

3.4 结论

B 介子半轻子衰变分支比与标准模型预言的不符很有可能暗示了新物理的存在。假如这一异常由一个 Z' 玻色子导致，那么它与标准模型费米子的耦合既要改变夸克味道，同时违反轻子普适性。引入与标准模型费米子混合的类矢量粒子，这种形式的耦合可以由 Yukawa 相互作用自然产生。 B 反常的全局拟合提示我们，混合发生在费米子的左手部分。以此为依据建立的一种 $U(1)_X$ 模型可以在高亮度 LHC 上进行检验。首先，根据 B 反常和 $B_s - \bar{B}_s$ 振荡的观测，以及理论的微扰性，我们推测出了新粒子的质量上限，其中 $m_{Z'} \lesssim 1.8 \text{ TeV}$ ，类矢量轻子 $m_{N,E} \lesssim 9.6 \text{ TeV}$ ，类矢量夸克 $m_{U,D} \lesssim 77 \text{ TeV}$ 。尽管 Z' 玻色子的质量上限更低，但计算表明，在 LHC 上通过双轻子道搜寻 Z' 玻色子无法排除这一模型，原因是 Z' 与夸克的耦合强度较弱， Z' 在强子对撞机上的产生截面可以很小。CMS 合作组用双 μ 子末态寻找重粒子的分析，进一步限制了 $U(1)_X$ 模型中 Z' 玻色子与 μ 子， μ 子中微子，第二代，第三代夸克的耦合。 $m_{Z'} < 1.5 \text{ TeV}$ 的 Z' 玻色子，轻子衰变分支比大于 90%。将 CMS 的当前结果外推至高亮度 LHC 情形，可使搜寻灵敏度提升一个数量级，但是在 Z' 的任何质量区域都不能覆盖全参数空间。相反，类矢量轻子可以通过 Drell-Yan 过程在 LHC 上大量产生。在带电的类矢量轻子质量 $m_E > m_{Z'}$ 时，通过 $E^\pm \rightarrow \mu^\pm Z' (\rightarrow \mu^\pm \mu^\mp)$ 衰变道，可产生 6 μ 子末态。这一信号几乎没有标准模

型本底。未来 LHC 对 m_E 的排除极限可达 1400 GeV。如果 $m_E < m_{Z'}$, $E \rightarrow \mu H$ 的分支比可能会占主导, 尤其是 Higgs 场与新的标量场 ϕ 的混合角比较大的情况。对这一衰变模式, 在分支比 $\text{BR}(E^\pm \rightarrow \mu^\pm H) \in [60, 100]\%$ 时, 高亮度 LHC 对质量范围 [800, 1000] GeV 的带电类矢量轻子 E 有两个 σ 的搜寻灵敏度。

第4章 在 LHC 上搜寻违反 R 宇称的超对称

4.1 介绍

R 宇称守恒的 SUSY 可以在对撞机上通过寻找具有多喷注和缺失横向动量的末态来探测。在 LHC 的当前阶段，gluino/squark 的质量已经被限制到几个 TeV 之上^[165,169]。

然而， R 宇称在 SUSY 模型中并非强制的。与 R -守恒的情形相对比，一些 R 宇称破坏 (RPV) 的情形在对撞机实验上受到的约束仍然较弱^[194-198]。特别地，对于 R 宇称破坏的算符 $U^c D^c D^c$ ，由于 LHC 上大量的 QCD 过程，对相关参数的限制很弱。文献 [199] 研究了算符 $U^c D^c D^c$ 对 LHC 上搜寻 stop 和 sbottom 的影响，发现质量约在 500 GeV 的 stop 和 sbottom 仍然无法完全排除。原因之一是 R 宇称破坏的情形是在简化模型下研究的，一些特定信号的信息没有被完全探索。

在强子 R 宇称破坏过程中，boosted 超对称粒子的衰变产物会在探测器上形成一束胖喷注。来自胖喷注子结构的信息被发现可以提升搜寻灵敏度，比如 neutralino 喷注子结构^[200] 或者 stop 喷注子结构^[201-202]。为表征喷注子结构，传统做法是在喷注上定义一些高层级的动力学变量，如 Mass-Drop 和 N -subjettiness。但是某些低层级的信息可能被丢掉了。另一方面，可以通过将喷注在探测器量能器上的能量沉积根据 $\eta-\phi$ 坐标的分布表示成喷注的数码图像，并利用最近的计算机视觉技术对其进行标记。

这里我们使用一个卷积神经网络来提取 neutralino 喷注的子结构信息，并使用 BDT^[107] 来提升 LHC 上对 R 宇称破坏的 SUSY 搜寻的灵敏度。考虑的具体信号过程是 gluino 成对产生，gluino 衰变成两个夸克和一个 neutralino，neutralino 随后通过强子的 R 宇称破坏算符 $U^c D^c D^c$ 衰变为三个夸克，如图4.1所示。卷积神经

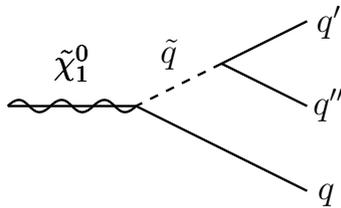


图 4.1 neutralino 通过 RPV 算符 $U^c D^c D^c$ 衰变的费曼图。

网络的主要任务是将信号过程中 boosted neutralino 喷注与 QCD 喷注中区分开来。注意到，neutralino 的质量是一个未知参数。我们将会展示，当训练神经网络与测试神经网络使用的 neutralino 质量不同时，标记 neutralino 喷注效率的变化。为了

给神经网络之前，喷注图像还会经过一定的预处理。然后，这些数据通过两个卷积层和一个最大池化层的两次迭代传递，其中每个节点采用 ReLU 激活函数加入非线性因素。每个卷积层中卷积核（或称为过滤器）的大小和总数都是自由参数。实际上，我们需要反复试验才能找到最佳选择。在图中，在迭代的第一步，输入数据经由 64 个大小为 6×6 的过滤器进行两次卷积运算，注意到两次卷积的过滤器并不相同，接着进行最大池化运算，过滤器大小为 2×2 ，步长为一个单位。在迭代的第二步中，过滤器的大小减小到 3×3 。每个卷积层和池化层中的过滤器总数与第一次迭代相同。然后将卷积与池化后得到的特征图展平成为一个向量并由全连接网络读取。在全连接网络的隐藏层中定义了 512 个神经节点，其中同样采用 ReLU 激活函数。最终输出层包含两个具有 sigmoid 激活函数的节点。每个节点的输出值在 $[0, 1]$ 之间，可用于表征信号或本底的概率。

4.3 CNN 的训练与测试

我们的目标是使用 CNN 从夸克和胶子的喷注图像中识别出 neutralino 的喷注图像，以便末态中包含 neutralino 的信号过程可以与本底进一步分离。为了使我们的 CNN 成为不依赖于任何特定产生过程的通用 neutralino 喷注识别器，用于训练的信号事例样本，其末态只包含一个可见的 neutralino，这一 neutralino 随后衰变为三个夸克。在整个工作中，MADGRAPH5 模拟了信号和本底事例的硬散射过程以及 neutralino 的衰变。PYTHIA8 用于模拟部分子簇射和强子化。探测器效应由 DELPHES3 进行模拟，并选用了 ATLAS 的参数配置。喷注的聚类算法则通过 FASTJET 实现。我们的 CNN 使用深度学习库 Keras 在 Python 中实现。

训练和测试样本的生成和处理如下。首先，在 SUSY 模型中，由 $pp \rightarrow \tilde{\chi}_1^0 \tilde{\chi}_2^0$ 过程生成带有单个可见 neutralino 喷注的信号事例，其中一个 neutralino 通过 $U_2^c D_2^c D_3^c$ 算符衰变为三个夸克 $\tilde{\chi}_2^0 \rightarrow bcs$ 。这里假定 $\tilde{\chi}_1^0$ 是稳定的，因此在探测器内不留任何痕迹。这是生成独立于过程的 neutralino 喷注进行训练和测试的一个技巧。在下一部分中将考虑完整模型， $\tilde{\chi}_1^0$ 是衰变为三个夸克，即 bcs 的 LSP。作为基准，我们选定 $\tilde{\chi}_2^0$ 的质量为 100 GeV。它的横向动量必须满足 $p_T(\tilde{\chi}_2^0) > 200$ GeV，以保证其衰变产物在探测器上表现为喷注。此外很显然的是，如果改变 neutralino 的极角（或赝快度），neutralino 喷注图像将发生变化。为了考虑这种影响，我们生成了两类信号事例样本：一类要求 $|\eta(\tilde{\chi}_2^0)| < 0.1$ ，记作中部样本，另一类允许更大的赝快度， $|\eta(\tilde{\chi}_2^0)| < 2.5$ ，记作广角样本。其次，在 SUSY 模型中，训练和测试中的本底事例由 $pp \rightarrow j\tilde{\chi}_1^0\tilde{\chi}_1^0$ 生成，其中 j 可以是夸克或胶子，而 $\tilde{\chi}_1^0$ 在探测器中是稳定的。与信号事例的生成一样，要求 j 的横向动量满足 $p_T(j) > 200$ GeV，并且定义两类本

底样本，分别要求 $|\eta(\tilde{\chi}_2^0)| < 0.1$ 和 $|\eta(\tilde{\chi}_2^0)| < 2.5$ 。应该注意的是，在训练和测试阶段，产生信号和本底事例时，Pythia8 中已关闭了初态辐射和多粒子相互作用，目的是抑制它们对信号喷注图像的污染，以使得 CNN 可以更有效地抓住喷注的重要特征。在后面考虑实际的 gluino 搜寻时将包括这些影响。第三，在信号和本底事例中，都通过 anti- k_t 算法重建喷注，其圆锥尺寸 $R = 1.0$ 。目标喷注的最小横向动量应为 100 GeV，此要求相比部分子的要求更宽松，因为我们发现有时由于大角度劈裂，重建的喷注可能比引发喷注的部分子更软。如果不存在横向动量 $p_T(j) > 100$ GeV 的喷注，则事例将被丢弃。如果一个事例中有多个喷注满足 $p_T(j) > 100$ GeV，则选取横向动量最大的喷注。对于信号事例，我们还要求选定的喷注与部分子 $\tilde{\chi}_2^0$ 的距离 $R < 1.0$ 。

在此阶段，每个事例都被指定了至多一个喷注。对于信号事例，该喷注预计将是 neutralino 喷注，对于本底事例，该喷注预计将是 QCD 喷注。尽管对喷注的指定不保证是百分百正确的，但这对于神经网络的训练并没有多少干扰，只要标注错误的样本数量足够小。接下来，我们需要将喷注信息转换为网格图像。给定一个喷注，根据 $\eta - \phi$ 的相对坐标建立一个网格，步长为 0.1×0.1 ，大小为 30×30 ，网格的中心落在喷注最硬的组分粒子 $\eta - \phi$ 坐标上。根据网格和喷注组分粒子的信息，我们可以为喷注图像定义三个不同的层（或称通道）：

- 1) 显示所有组分粒子的能量分布的层，其中属于同一网格单元的粒子，将能量累加起来；
- 2) 与第一层类似，但是仅考虑带电粒子的能量；
- 3) 计算每个单元中带电粒子数量的层。

注意到，严格来说，第三层的定义按照喷注算法的标准属于 IRC 不可靠。不过这里是对已定义好的喷注进行处理，不存在矛盾，而且实践中，这种定义方式对喷注识别有效^[204]。由于发现 CNN 在处理 $[0, 1]$ 之间的数字时最有效，因此每层中的所有数字都除以该层中的最大值，例如第一层中单元格的最大能量。与某些研究中的做法不同，我们将不再应用任何其它图像预处理步骤，例如旋转和翻转，因为发现它们会降低 CNN 的性能，这与文献 [118] 中的发现相同。

为了避免过度训练问题，我们生成了 100 万个信号和本底事例用于训练。为了有效地利用数据集，训练过程需要 30 个周期 (epoch)，即全部样本使用 30 次。另外生成了 100 万个信号和本底事例用于测试。

CNN 中有许多自由参数只能通过反复试验进行优化，包括卷积层中卷积核的大小和数量，随机丢掉节点 (dropout) 的比率，全连接网络隐藏层中节点的数量和 Adam 算法中的学习速率等。我们发现 CNN 的性能仅轻度依赖于这些参数。在

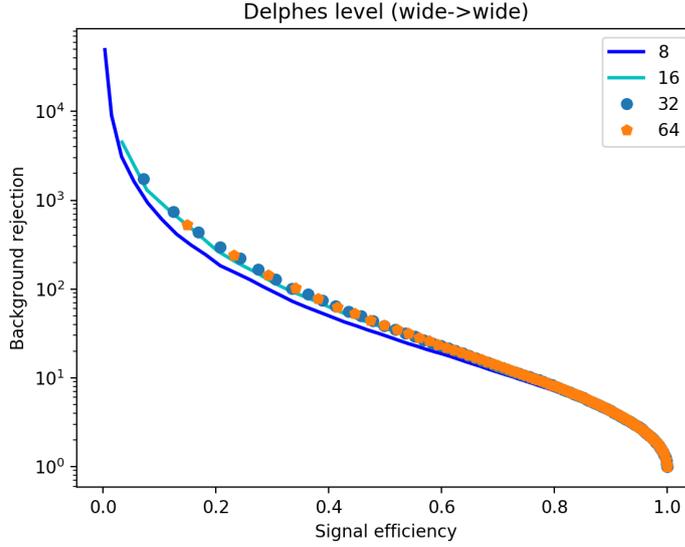


图 4.3 改变卷积核数量对 CNN 性能的影响。

图4.3中，显示了卷积层中卷积核数量分别为 8，16，32，和 64 时 CNN 的性能。所有卷积层均采用相同数量的卷积核。卷积核数量大于 16 的 CNN 性能表现几乎同样优秀，而且比卷积核数量为 8 的 CNN 稍好。卷积核的大小设为 6×6 ，因为我们发现卷积核大小为 2×2 和 4×4 的 CNN 表现更差。卷积层和池化层中 **dropout** 的比率设为 0.25，而全连接网络的 **dropout** 比率设为 0.5。隐藏层中的节点数为 512。学习速率设置为 0.001。这些参数的选取将在整个工作中保持不变。此处可训练参数的数量 ($\sim 6.5 \times 10^6$) 大于训练样本的大小（其实这在深度神经网络的训练中十分常见，我们已经使用了 **dropout** 手段来防止过拟合）。即使如此，但由于以下两个原因，我们不进一步减少 CNN 的参数规模。首先，我们尝试使用了参数数量要小得多的 CNN，在卷积层中有 8 个卷积核，全连接网络的隐藏层中有 64 个节点，参数数量约为 $\sim 1 \times 10^5$ ，其性能略逊于图4.2中所示的 CNN。其次，经过训练的 CNN 已在独立事例样本上进行了测试，其准确性与训练样本上相近。因此，CNN 并没有过度训练。

注意到我们事实上定义了两个 CNN，分别针对中心样本和广角样本进行了训练和测试。图4.3中显示的是广角样本上训练的 CNN 应用于广角样本测试的结果。在图4.4中，为了表征喷注图像特征对喷注速度依赖性的依赖性，我们展示了这两个 CNN（在所有卷积层均具有 64 个卷积核）在不同样本上的性能。毫无疑问，中部喷注 ($|\eta| < 0.1$) 比广角喷注 ($|\eta| < 2.5$) 更容易标记。在中部样本上经过训练的 CNN 无法在广角样本中标记 **neutralino** 喷注，主要是因为 CNN 在中部样本中捕获的特征相对于广角样本比较匮乏。另一方面，在广角样本上训练的 CNN 在标记中部样本的 **neutralino** 喷注方面表现良好，尽管它比直接在中部样本上训练的 CNN

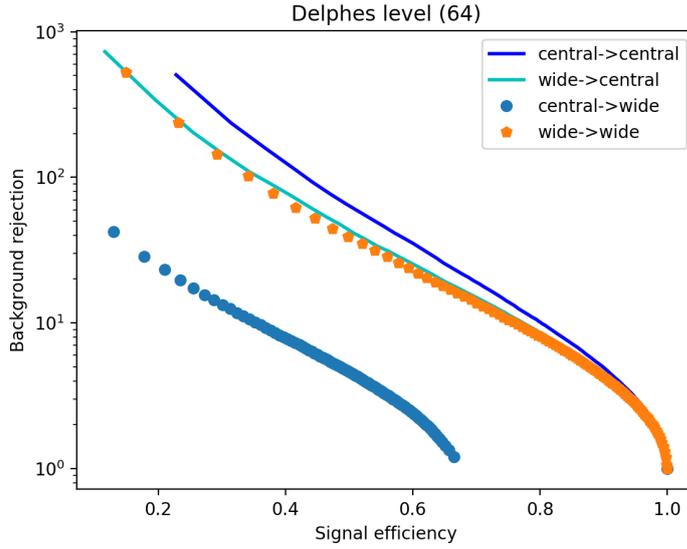


图 4.4 在两组样本上训练的 CNN 分别应用到另外两组测试样本上的性能表现，共有 4 种组合方式，即中部 \rightarrow 中部，广角 \rightarrow 中部，中部 \rightarrow 广角，广角 \rightarrow 广角。

稍差。这意味着我们不必将分析局限于中心区域的目标喷注。这在大型强子对撞机进行实际信号搜索时更有用处，因为可以保留更多信号事例。在下面我们将继续使用经过广角样本训练和测试的 CNN，每个卷积层中的卷积核数量固定为 64。

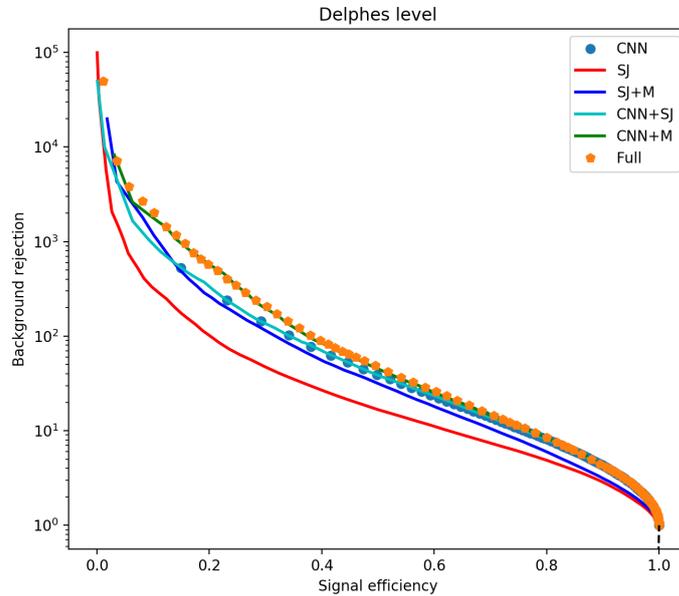


图 4.5 不同方法区分 $neutralino$ 喷注和 QCD 喷注的性能对比。SJ 表示 N -subjettiness，M 表示不变质量。不同变量的结合分析用 BDT 处理。

N -subjettiness 是一种可以表征喷注多分支结构的动力学变量，被证明是一种

有效的区分喷注的判据，已被广泛用于一般的喷注标记。给定整数 N ，将喷注重新聚类成 N 个子喷注，定义 N -subjettiness 如下，

$$\tau_N = \frac{\sum_i p_{Ti} \min_{k \in N} (\Delta R_{k,i})}{\sum_i p_{Ti} R_0} \quad (4-1)$$

其中 R_0 为喷注直径， p_{Ti} 为第 i 个组分粒子的横向动量， $\Delta R_{k,i}$ 为第 k 个子喷注与第 i 个组分粒子的间距。具有 N 分叉结构的喷注 $\tau_N \simeq 0$ ，而对所有 $I < N$ ， $\tau_I \gg 0$ 。通常有 $\tau_{N-1} > \tau_N$ 。事实上，实际使用的是更具分辨能力的 N -subjettiness 比值， τ_N/τ_{N-1} ，比如 boosted W 喷注使用 τ_2/τ_1 来标记，boosted top 喷注使用 τ_3/τ_2 来标记。我们将 CNN 的性能与 N -subjettiness 进行比较，结果示于图 4.5。一般使用变量 τ_N/τ_{N-1} 标记具有 N -分支结构的喷注。在我们的例子中，neutralino 喷注的子结构可以用 τ_3/τ_2 来标记。 N -subjettiness 的性能由图中的红色实线表示。我们发现，平均来说，CNN 的性能（以蓝色圆点表示）比 N -subjettiness 的性能好若干倍。此外，喷注的不变质量也是一个有力的判别依据，且与 N -subjettiness 不相关。为了结合两个变量的判别能力，我们采用了 BDT 方法。它在重建的一半 neutralino 喷注和 QCD 喷注上训练，并在其余的样本上进行了测试。 N -subjettiness 和喷注不变质量二者结合的性能由图中蓝色实线给出，它显示出了与单独使用 CNN 相似的标记效率。

同时有必要弄清的是我们的 CNN 是否足够智能，以至于学习到了 N -分支结构和喷注不变质量。这可以通过观察它们组合以后的标记效率得出。在图 4.5 中，CNN + N -subjettiness (SJ) 和 CNN + 喷注不变质量 (M) 的性能分别用青色和绿色实线表示。三者结合的性能用橙色圆点表示。此处同样使用了 BDT 方法。CNN + SJ 没有比单独的 CNN 表现出更多改进。而在将喷注不变质量纳入考虑之后，标记效率提高了几倍。因此，我们可以得出结论，CNN 能获取喷注中叉状结构的完整信息，但是目前的方法无法直接从喷注图像中提取喷注不变质量。原因之一是图像预处理步骤不保持 Lorentz 对称性，因此喷注不变质量的信息被破坏掉了。

在上述研究中，所有事例样本中的 neutralino 质量均设定为 100 GeV。在实践中，使用 CNN 方法的目的是提高信号发现的灵敏度，neutralino 质量是未知的参数。使 CNN 训练样本具有与我们想要探测的信号相同的 neutralino 质量是不现实的。一种方法是在训练样本中加入具有所有质量 neutralino 的事例^[205]，适用性更广泛，但需要耗费的计算资源更多。另一种方法是训练多个 CNN，每个 CNN 针对一个选定的 neutralino 质量，并将这些 CNN 应用于质量范围不确定的 neutralino 的标记。则使用最接近真实 neutralino 质量的样本训练的 CNN 应该能够最有效地

标记信号。在固定的 neutralino 质量上训练的 CNN 的通用性可以在图4.6中看到。左图显示了 CNN 在 neutralino 质量为 $[70, 150]$ GeV 的测试样本中的性能，其中 CNN 仅通过 $m_{\tilde{\chi}_2^0} = 100$ GeV 样本进行训练。我们发现，在 $[90, 125]$ GeV 范围内，变化的 neutralino 质量不会对灵敏度有太大损害，但是 CNN 更容易受到较低 neutralino 质量的影响。另一方面，如果将 CNN 与喷注不变质量 (CNN + M) 结合使用， neutralino 质量变化对灵敏度的负面影响会更小。右图显示了 CNN 与不变质量的组合在 neutralino 质量不同的样本上的性能。来自喷注不变质量的信息有助于提高标记效率，可以弥补 CNN 的不足，尤其是在轻 neutralino 的质量区域。总而言之，加入喷注不变质量不仅可以提高标记效率，而且可以扩展 CNN 的应用范围。在实际信号搜索中应将它们一起使用。

4.4 应用 CNN 在 LHC 上搜寻 gluino

显示了 CNN 方法的强大功能和一定的普适性后，我们准备将其应用在 RPV gluino 的搜索。文献 [206] 同样尝试使用 CNN 改进这一过程的搜索灵敏度，不过使用了整个事例图像作为 CNN 的输入。信号过程是 gluino 的成对产生，其中每个 gluino 均衰变为两个夸克和一个 neutralino 。 neutralino 通过强子 RPV 算符衰变为三个夸克。ATLAS 合作组曾搜索过此信号^[207]，对于质量 $m_{\tilde{\chi}_1^0} \sim 100$ GeV 的 neutralino ，质量 $m_{\tilde{g}} \lesssim 1.1$ TeV 的 gluino 已被排除。搜索中主要的本底过程是 QCD 多喷注事例。我们将展示 CNN 如何帮助改善 ATLAS 的 gluino 搜索结果。在此之前，我们首先需要对信号和本底进行重新分析。

QCD 多喷注过程使用 MADGRAPH5 模拟，模拟只计入了领头阶贡献。高阶 QCD 修正会改变喷注关于数量，赝快度和横向动量的分布，但是只能间接影响喷注子结构。比如，由于较高的喷注数量而造成目标喷注之间存在更多的污染；喷注分布更加分散，从而平均而言具有更大的赝快度和更小的横向动量。我们的结果对这些影响并不敏感，因为 PYTHIA8 对部分子簇射过程的模拟已经包含了所有领头阶对数 (leading logarithmic) 贡献，而我们的 CNN 能够在很宽的赝快度和横向动量范围内标记喷注。根据 ATLAS 分析中所采用的筛选条件，我们仅考虑在部分子级别具有 4 个或 5 个喷注的过程，每个喷注的横向动量应满足 $p_T > 200$ GeV 和 $|\eta| < 2.0$ 。在 Pythia8 中开启初态辐射和末态辐射后，可以获得具有更高喷注数量的事例。

信号事例的生成也是在领头阶进行的。信号过程的基准点是： neutralino 质量 $m_{\tilde{\chi}_1^0} \in [50, 200]$ GeV，gluino 质量 $m_{\tilde{g}} \in [1, 2]$ TeV，质量间隔 50 GeV。

我们将 ATLAS 的分析重做如下：

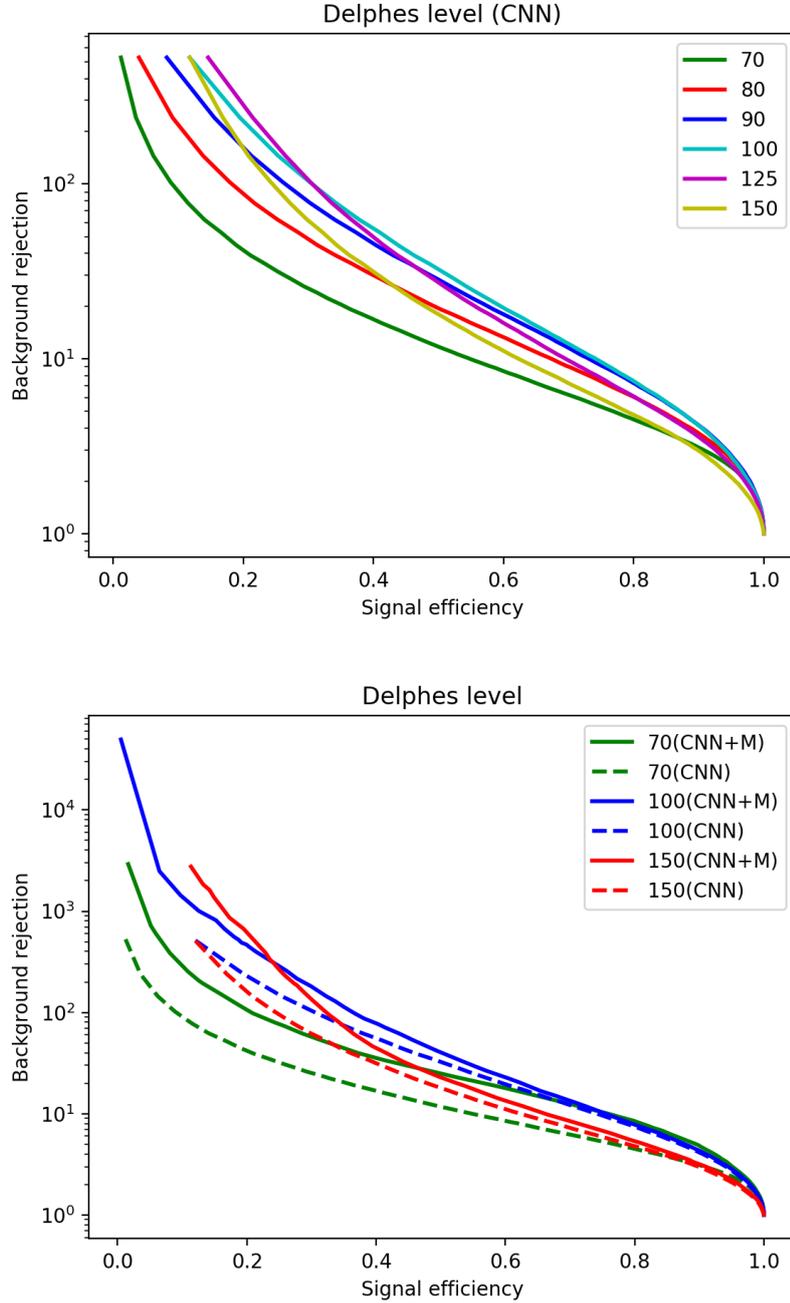


图 4.6 上图显示了 CNN 应用到 neutralino 质量不同的样本上时对信号与本底的区分能力。下图显示了当结合了 CNN 输出的信息与喷注不变质量之后性能的提升。两张图使用的 CNN 都是在 $m_{\tilde{\chi}_2^0} = 100 \text{ GeV}$ 的事例样本上训练的。图例中的数字表示测试样本中 neutralino 的质量。

1) 对于每个事例, 通过 $\text{anti-}k_T$ 算法以半径参数 $R = 1.0$ 重建胖喷注, 即大半径喷注。每个大半径喷注都应用 trim 算法处理, 要求 trim 后的子喷注半径参数为 $R_{\text{subject}} = 0.2$, 最小横向动量为原喷注的 5%。经过 trim 的大半径喷注需要满足 $p_T > 200 \text{ GeV}$ 和 $|\eta| < 2.0$ 。该分析仅选择包含至少四个 trim 过的大半径喷注 ($N_{\text{jet}} \geq 4$) 的事例, 其中主要的喷注, 即横向动量最大的喷注应满足 $p_T > 440 \text{ GeV}$ 。

2) 同时, 利用 $\text{anti-}k_T$ 算法以半径参数 $R = 0.4$ 重建每个事例的小半径喷注, 它们必须满足 $p_T > 50 \text{ GeV}$ 和 $|\eta| < 2.5$ 。这些喷注用于计算末态中被 $b\text{-tag}$, 即标记为 bottom 夸克的喷注数量 (N_b)。 $b\text{-tag}$ 的效率设定为 70%, 对 charm 夸克和轻夸克喷注的误标记率分别设定为 15% 和 0.8%。

3) 为每个事例定义两个判别变量: 总喷注质量 M_J^Σ , 定义为四个主要的大半径喷注 trim 后不变质量的标量和,

$$M_J^\Sigma = \sum_{i=1}^4 m_{j_i}^{\text{trimmed}} \quad (4-2)$$

以及两个主要的大半径喷注 trim 后的赝快度之差 $|\Delta\eta_{12}|$ 。

4) 定义四个信号区间, 示于表 4.1。表中, 倒数第二列为实验观测到的符合筛选条件的事例个数, 最后一列为预计的本底事例数, 由标准模型计算给出, 其不确定度来源见于文献 [207]。

表 4.1 ATLAS 分析中定义的 4 个信号区间

| 信号区间 | N_{jet} | N_b | M_J^Σ | $ \Delta\eta_{12} $ | 实验观测 | 标准模型预言 |
|--------|------------------|-------|---------------------|---------------------|------|--------------------------------|
| 4jSR | ≥ 4 | - | $> 0.8 \text{ TeV}$ | < 1.4 | 122 | $151 \pm 15 \pm 17 \pm 20$ |
| 4jSRb1 | ≥ 4 | > 0 | $> 0.8 \text{ TeV}$ | < 1.4 | 46 | $61 \pm 10 \pm 6 \pm 12$ |
| 5jSR | ≥ 5 | - | $> 0.6 \text{ TeV}$ | < 1.4 | 64 | $51.4 \pm 7.7 \pm 7.2 \pm 6.5$ |
| 5jSRb1 | ≥ 5 | > 0 | $> 0.6 \text{ TeV}$ | < 1.4 | 30 | $18.2 \pm 4.2 \pm 2.5 \pm 3.0$ |

因为我们对 neutralino 质量较低的情况更感兴趣, 所以 4jSRb1 信号区间提供了灵敏度最高的探测方案。样本中只有通过了 4jSRb1 信号区间所有筛选条件的信号和本底事例被保留下来供下面分析。在模拟中, 选定的信号和本底事例数保证在 10^4 左右, 以抑制统计不确定性。在此阶段, 信号的截面可以用 $\sigma(pp \rightarrow \tilde{g}\tilde{g}) \times \epsilon^{4\text{jSRb1}}$ 来计算, 其中 $\sigma(pp \rightarrow \tilde{g}\tilde{g})$ 是在 LHC 对撞能量为 13 TeV 时, gluino 成对产生的截面, 这可以由 PROSPINO2^[208] 计算到次领头阶精确度, $\epsilon^{4\text{jSRb1}}$ 是从我们的重新分析中得到的 4jSRb1 信号区间的事例筛选效率。将表 4.1 的“标准模型预言”列中的数字除以 ATLAS 此项分析对应的积分亮度 $\mathcal{L} = 14.8 \text{ fb}^{-1}$, 即可简单地估算出来此阶段的本底截面 σ^{BG} 。

现在, 我们可以在选定的信号和本底事例中将 CNN 应用到喷注的标记上。首

先, 在每个选定的事例中, 以与训练样本相同的方式重建喷注, 即半径参数 $R = 1.0$, 横向动量 $p_T > 100$ GeV 的 anti- k_t 喷注。由于信号过程中两个 **neutralino** 喷注可能具有很高的能量, 也可能相对较软, 因此所有重建的喷注都将传递到我们的 CNN 中进行 **neutralino** 标记。每个喷注都会被分配一个得分, 表示它是 **neutralino** 的概率。然后, 根据概率大小对喷注进行排序。前三个喷注的概率分布如图 4.7 所示, 其

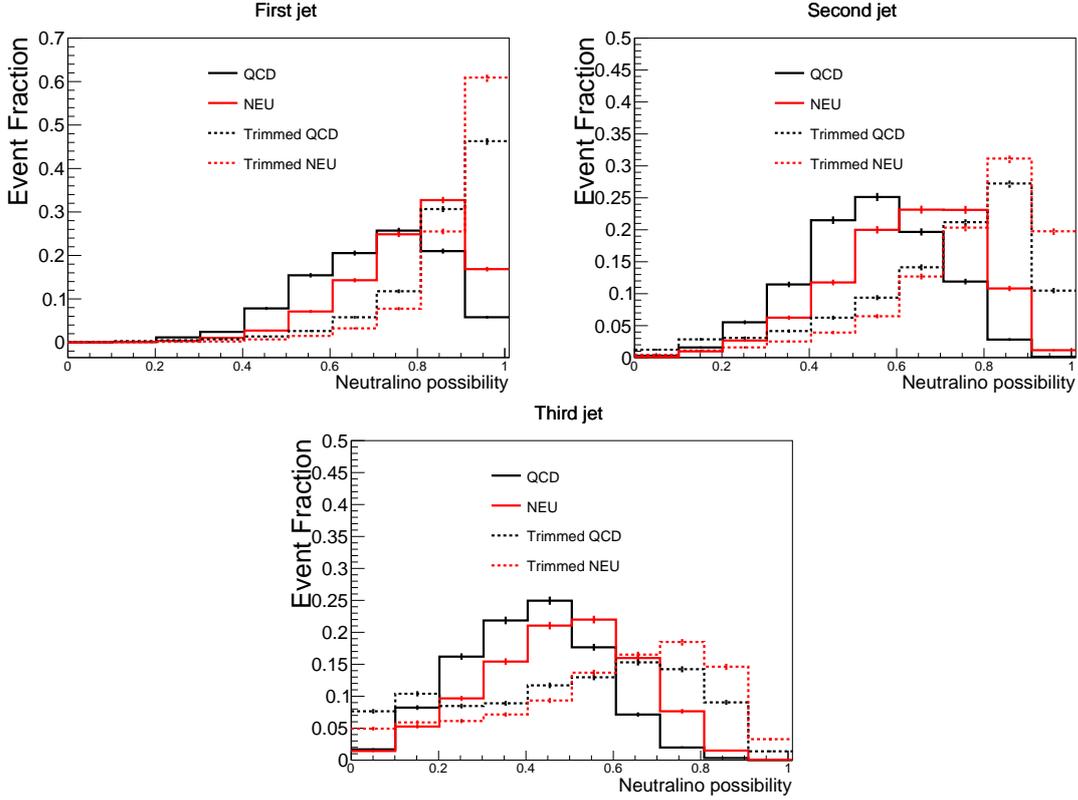


图 4.7 CNN 得分最高的前三个喷注的分数分布。红线代表信号过程, 黑线代表本底过程。虚线和实线分别表示喷注是否经过 trim。示例中的 **gluino** 和 **neutralino** 质量分别为 $m_{\tilde{g}} = 1.5$ TeV 和 $m_{\tilde{\chi}_1^0} = 100$ GeV。

中 **gluino** 质量和 **neutralino** 质量分别设置为 1.5 TeV 和 100 GeV。我们可以看到, 平均来说, 信号事例中的喷注比本底事例中的喷注从 CNN 得到的分数更高。这可以帮助我们进一步区分信号和本底。另一方面, 也容易看出, 即使本底事例中的 QCD 喷注也可能获得相对较高的 CNN 分数。这表明, 筛选后的信号和本底事例中的 **neutralino** 喷注和 QCD 喷注比训练样本中的更加难以区分。困难主要源于选定事例中喷注之间的严重污染。这些污染倾向于形成具有多分叉结构的 QCD 喷注, 这也同时降低了 N -subjettiness 的辨别能力。图 4.7 中的虚线表示 CNN 对 trim 后的喷注标记的效率, trim 参数与 ATLAS 分析中的相同。可以看到, 平均而言, trim 这一步骤同时提高了信号与本底中喷注的得分, 但是并没有提高两者的相对差异, 在此阶段对信号本底的区分几乎不起作用。但是要注意到, 使用 trim 后的喷注来

定义前面的筛选条件依然是十分必要的。

因为信号或本底中前三个喷注的 CNN 得分在某种程度上相关，我们再次采用 BDT 方法研究这些信息与喷注不变质量和 N -subjettiness 组合起来的辨别能力。与前面的 BDT 分析相比，现在的可用事例数少得多，大约在 10^4 量级，而输入变量更多，包括三个喷注的 CNN 分数和不变质量等。在每个给定 $gluino$ 和 $neutralino$ 质量的基准点上，都使用 5000 个信号事例和 5000 个本底事例对 BDT 进行训练，然后对其余的独立事例进行测试。对于每个 BDT，对其输出的值进行筛选将使信号和本底截面进一步降低至 $\sigma(pp \rightarrow \tilde{g}\tilde{g}) \times \varepsilon^{4jSRbl} \times \varepsilon_S^{BDT}$ 和 $\sigma^{BG} \times \varepsilon_B^{BDT}$ 。 ε_S^{BDT} 和 ε_B^{BDT} 分别表示信号和本底事例使用 BDT 筛选的效率。我们假设表 4.1 中“实验观测”列中的事例数以 ε_B^{BDT} 为因子减少，因为这些事例或者全部属于本底，或者绝大多数属于本底。至于标准模型预言的本底事例数的不确定度，来自统计误差的分量以 $\sqrt{\varepsilon_B^{BDT}}$ 为因子缩小，而其他分量以 ε_B^{BDT} 为因子缩小。我们将采用 p 值（或置信水平，confidence level）来表征信号被排除的概率，其定义为

$$p = 1 - \frac{\mathcal{P}(\mathcal{H}_{S+B})}{\mathcal{P}(\mathcal{H}_B)} \quad (4-3)$$

其中， $\mathcal{P}(\mathcal{H}_{S+B})$ 和 $\mathcal{P}(\mathcal{H}_B)$ 分别是信号加本底假设和仅有本底假设的概率。如果 p 值大于 0.95，则 \mathcal{H}_{S+B} 假设在 95% 置信水平被排除。

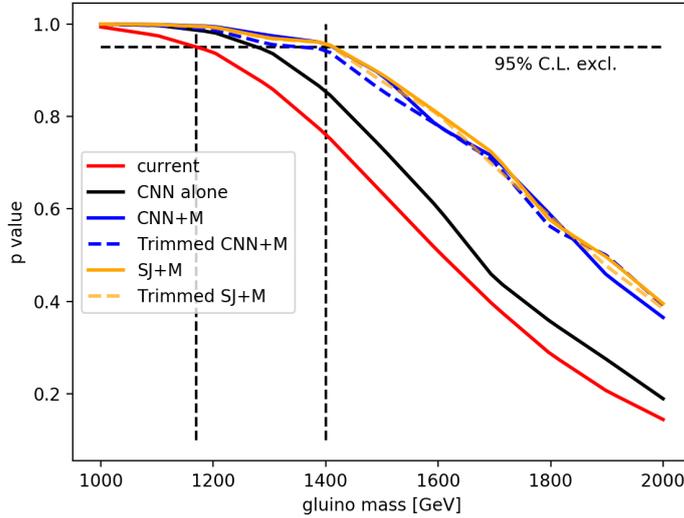


图 4.8 采用不同分析方法得到的 p -值。红色实线为原始分析，黑色实线为单独使用 CNN，蓝色和黄色分别表示 CNN 结合不变质量和 N -subjettiness 结合不变质量。实线和虚线分别表示喷注是否经过 trim。

在每种分析方法中，每个给定 $gluino$ 和 $neutralino$ 质量的基准点上，将采用最大化 p 值的 BDT 筛选条件。我们发现，在这一优化标准下，对信号的筛选效率为

$\sim 50\%$ ，对本底的筛选效率约为 10-20%。ATLAS 的原始分析，单独使用 CNN 进行标记，和结合了 CNN 输出与喷注不变质量的 BDT 分析，三种方法的 p 值显示在图 4.8 中，这里我们固定了 neutralino 质量 $m_{\tilde{\chi}_1^0} = 100$ GeV。我们对 ATLAS 分析的重做结果显示，在 95% 置信水平，gluino 质量 $m_{\tilde{g}} \lesssim 1.18$ TeV 可以排除，这与实验组的结果基本吻合。仅使用 CNN 标记喷注的信息，可以将 gluino 质量的下限推至 ~ 1.3 TeV。通过将喷注不变质量添加到 BDT 中，将使质量 $m_{\tilde{g}} \lesssim 1.4$ TeV 的 gluino 被排除。为了进行比较，我们还显示了使用 N -subjettiness 和喷注不变质量进行分析的 p 值，其表现并没有优于考虑 CNN 与不变质量的方法。此外，虚线对应于分析前首先对喷注进行 trim 的 p 值。事实证明，trim 喷注在此时无助于改善对信号和本底的辨别能力，这印证了上面从图 4.7 推出的观点。

我们已经演示过，使用 CNN 来标记未 trim 的喷注，并结合喷注不变质量的信息，可以给 RPV gluino 的搜索提供灵敏度更高的方案。最后，我们展示该方法应用于 gluino 和 neutralino 质量不同的事例时对参数空间的限制。注意到，这里的 CNN 仍然是仅在 $m_{\tilde{\chi}_1^0} = 100$ GeV 的事例样本上训练过。对参数空间的排除限制示于图 4.9 的 $m_{\tilde{g}} - m_{\tilde{\chi}_1^0}$ 平面上。

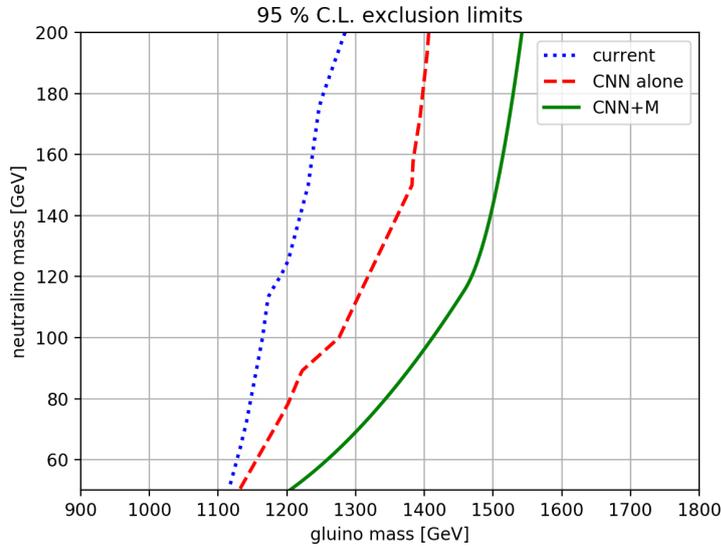


图 4.9 在 95% 置信水平对 gluino 和 neutralino 质量的排除极限。蓝色点线，红色虚线和绿色实线分别对应 ATLAS 的原始分析方法，单独使用 CNN，和同时使用 CNN 与喷注不变质量。

图中，基于 ATLAS 合作组的原始分析，仅使用 CNN 标记喷注的分析方式，和 CNN 结合喷注不变质量的分析方法，分别给出了 95% 置信水平的排除限制。我们可以观察到，较轻的 neutralino 对 CNN 的负面影响更大。对于质量小于 100 GeV 的 neutralino，其对原有分析的改善十分有限。而面对比训练样本重的 neutralino

时, CNN 对其质量改变并不敏感。这主要是因为从很重的 **gluino** 衰变而来的质量 $m_{\tilde{\chi}_1^0} \lesssim 200 \text{ GeV}$ 的 **neutralino** 具有很大的横向动量, $p_T \gtrsim 400 \text{ GeV}$ 。它的几乎所有衰变产物都在喷注重建中被保留。因此, 喷注子结构信息可以被有效利用, 除非 **neutralino** 质量过小, 以至于其子喷注开始有部分重叠。相比单独使用 CNN, 结合喷注不变质量信息可以帮助将 **gluino** 排除极限进一步提高 $\sim 100 \text{ GeV}$ 。而且, 喷注不变质量对于分辨较轻的 **neutralino** 帮助更大, 可以显著弥补 CNN 在此区域性能的不足, 这与前面从图4.6中得到的结论一致。最终的效果, 根据不同的 **neutralino** 质量, 对 **gluino** 的排除极限可以提升 $100 - 250 \text{ GeV}$ 。

4.5 结论

通常的超对称标准模型中并不考虑破坏 R 宇称的算符, 但是超对称理论本身并不强制要求 R 宇称守恒。当 R 宇称破坏, 最轻的 **neutralino** 不再稳定, 通过 $U^c D^c D^c$ 算符, 它可以衰变为三个夸克。对这一情形的实验限制仍然较弱。为了改进已有分析, 我们使用卷积神经网络提取 **neutralino** 喷注子结构信息, 使用提升决策树进行多变量分析。

将量能器单元视为像素点, 喷注在探测器上的能量沉积可以表示为图像。这样就可以用卷积神经网络识别喷注。我们在末态中仅包含一个 **neutralino** 喷注的简化过程上训练网络, 以使其识别能力不依赖于 **neutralino** 的产生机制。训练好的网络可以在仅接受约 1% QCD 喷注的条件下, 以 50% 的效率识别 **neutralino** 喷注。我们发现, 识别效率在网络的参数达到一定规模后饱和。不论 **neutralino** 质量大小, 网络的表现都很好, 识别 $|\eta| \rightarrow 0$ 的喷注比 $|\eta| < 2.5$ 的喷注效率略高。在区分 **neutralino** 喷注与 QCD 喷注时, 卷积神经网络比喷注子结构变量 N -subjettiness 的表现强若干倍。但是由于喷注图像的预处理过程不保持 Lorentz 对称性, 喷注的不变质量信息没有被网络学习到。将网络输出与喷注不变质量的信息结合起来, 信号显著性可以更高, 且适用于质量变化的 **neutralino**。

为了研究神经网络的实际应用效果, 我们重做了 ATLAS 用多喷注末态搜索 **gluino** 的分析, 对应的过程为 **gluino** 成对产生, 随后 $\tilde{g} \rightarrow qq\tilde{\chi}_1^0 (\rightarrow qqg)$ 。仅保留了通过某一信号区间所有筛选条件的信号和本底事例。使用前面训练好的网络识别所有喷注, 此时由于末态中多喷注的辐射污染, 网络和 N -subjettiness 的辨别能力都有所下降。但是, 将领头三个喷注的网络输出分数与不变质量作为提升决策树的输入, **gluino** 质量的排除极限仍然被推高了。对质量 $m_{\tilde{\chi}_1^0} = 100 \text{ GeV}$ 的 **neutralino**, 质量 $m_{\tilde{g}} \lesssim 1.4 \text{ TeV}$ 的 **gluino** 被排除。对质量较小的 **neutralino**, 网络的识别能力变弱, 而对较重 **neutralino** 质量变化不敏感。根据 **neutralino** 质量变化, **gluino** 的质量

排除极限总体被推高了 $100 - 250 \text{ GeV}$ 。

第 5 章 Higgs 喷注检测与重建

喷注的重建依赖于特定的喷注聚类算法，常用的 C/A ， k_t ， $\text{anti-}k_t$ 等算法受制于固定的半径参数，有时无法很好的完成新物理模型的数据分析任务。这里我们以 Higgs 喷注的检测和重建为例，探索使用深度学习技术提升传统算法的表现。下面首先介绍所使用的神经网络架构。

5.1 Mask R-CNN

Mask R-CNN^[209] 是一种在计算机视觉领域用于目标检测和语义分割的框架。它在 2017 年被首次提出，是由基于区域的卷积神经网络（region-based CNN，R-CNN）^[210] 这一目标识别框架经历次迭代发展而来，目前已广泛应用于业界。

Mask R-CNN 由三个主要的功能模块以精巧复杂的关系整合而成，包括区域提议网络（region proposal network，RPN），基于区域的卷积神经网络，和全卷积网络（fully convolutional network，FCN）^[211]。其中 RPN 和 R-CNN 经过了改造，将单一尺度的特征映射（feature map）替换为特征金字塔网络（feature pyramid network，FPN）^[212]，用于检测多尺度目标。原始文献比较了使用不同的卷积主干架构的差异，这里我们选择 50 层的残差神经网络（ResNet-50）。因为更深的网络，ResNet-101，在应用于我们的数据时没有表现出显著的优势。

在特征金字塔网络中，每一层级的特征映射被上采样后，与主干网络中相应尺寸的层级融合以生成低一层级的特征映射。遍布特征金字塔每一层级的锚框（anchor）是一组具有固定大小和纵横比的矩形参照物，是 RPN 的前置输入之一。RPN 由两个并列的网络层组成。一个分类层判定锚框为阳性或阴性（p/n）。一个回归层调整矩形框的边界位置。两者结合，给出的矩形候选框，被称作感兴趣区域（regions of interest，RoI）。特征金字塔网络和锚框的设计使得 Mask R-CNN 可以应对多尺度的目标检测。至此，提议过程结束。接着，一个池化层从每个感兴趣区域提取特征并绘制统一大小的方形特征映射，用于后面的检测过程。两个并列的全连层对此特征映射进行分类和进一步调整边界框位置。具体到我们的应用，只有两个类别，Higgs 喷注与本底。提议阶段和检测阶段所使用的特征是共享的，这种设计大大提升了网络性能，节省了计算系统开支。最后，在预测类别与最终边界框的同时，一个小型全卷积网络针对每个感兴趣区域输出一个二值掩膜，用于判定任意像素点是否属于 Higgs 喷注。图 5.1 简要展示了 Mask R-CNN 的各主要功能模块以及数据处理的流水线。

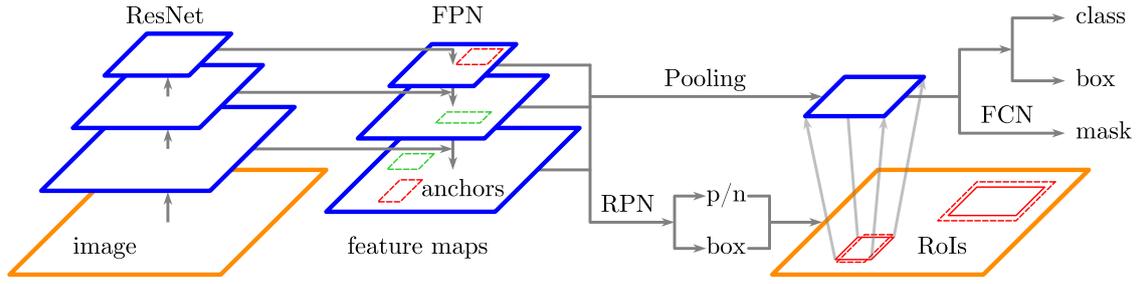


图 5.1 Mask R-CNN 的组成部分以及工作流程简图，受文献 [209,212-214] 启发而绘制。橘色方形表示输入图像，蓝色方形表示多尺度的特征映射。图中为简化只绘制了 3 个层级，实际实现中有 5 个层级，理论上可检测大小差距达 $2^5 = 32$ 倍的目标。虚线矩形表示锚框，锚框的阳性或阴性分别用红色和绿色表示。感兴趣区域用红色实线矩形表示。

5.2 数据集

我们使用 MADGRAPH5 生成了两组数据用于训练，验证和测试。每组数据包含 15 万个事例，其中 8 万个用于训练，2 万个用于验证，5 万个用于测试。每个事例中包含 1 个 Higgs 玻色子和 3 个轻夸克或胶子喷注，其中 Higgs 玻色子衰变到 2 个 bottom 夸克。然后每个事例被叠加了 50 次 pile-up。在 2 个数据集中，Higgs 玻色子分别被要求具有大于 200 GeV 和 300 GeV 的横向动量。之所以没有将横向动量限制在一个比较窄的区域内，目的就是展示这一算法的灵活性与通用性。

初态辐射会使 Higgs 玻色子的动量在衰变之前改变一点，这一部分的模拟是由 Pythia 处理的。因而一些 Higgs 玻色子的横向动量会掉到 200 GeV 或 300 GeV 以下，低于在 GadGraph 设置的阈值。我们决定保留这些事例，对于后面的测试样本也是如此处理的。两组样本中 Higgs 玻色子的 p_T 分布见图 5.2。

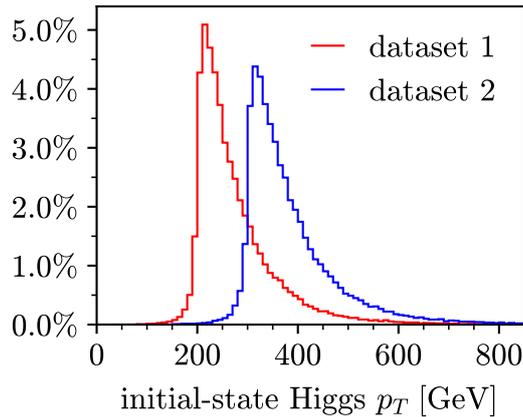


图 5.2 训练样本的 Higgs 玻色子横向动量分布。

数据是以分辨率 320×320 的灰度图形式喂给神经网络的。图片表示的是 $\eta - \phi$ 平面上 $[-\pi, \pi] \times [0, 2\pi]$ 范围内所有粒子的横向动量，包括中微子。每个像素点对应 $\Delta\eta \times \Delta\phi = 0.020 \times 0.020$ 。为了使图片尽可能利用多个尺度的横向动量，每个

像素的亮度根据整个数据集中 Higgs 衰变产物的横向动量分布做了正规化，目标是使对应于 Higgs 喷注的像素点，灰度分布在 0-255 的全区间内是平的，即 256 个色阶中的每一阶都包含了同样数量的衰变自 Higgs 的粒子。注意到对于整个事例的粒子没有类似结论。这一步骤可以看作一种图像增强手段，类似于对 HDR (high dynamic range) 图像进行全局色调映射 (tone mapping) 处理，以增强图像的局部对比度。因为我们感兴趣的粒子动量分布跨越了 3-4 个数量级，一种简单的做法是使用对数标度来绘制量能器图像。但这会导致对软辐射敏感度不足，将亮度的梯度浪费在 10 GeV 以上区域。图 5.3 左边显示了 Higgs 喷注组分粒子的 p_T 分布，右边则为据此分布设计的标度。数值上，右侧曲线等于左侧曲线的积分乘以 256，对应 8bit 色深。如此表示粒子横向动量可以最大化神经网络对 Higgs 衰变产物的能量分辨率。注意到，尽管两个样本对应的标度不同，但差别较小。原则上也可以根据两个样本的联合分布，使用统一的标度。另外，尽管该标度是针对一定能量的 Higgs 玻色子优化的，但是对于类似的物理对象，如强子衰变的矢量玻色子，top 夸克等的识别，也在一定程度上适用，这是由部分子簇射和强子化过程的能量尺度决定的。如此可以将这套方法推广到多类别喷注的识别与重建，而不仅限于 Higgs 喷注。

图 5.5 和图 5.6 展示了完整事例的探测器图像。按照前述的亮度正规化方案，大部分像素应为黑色，图中对灰度做了反转处理以提升观感。

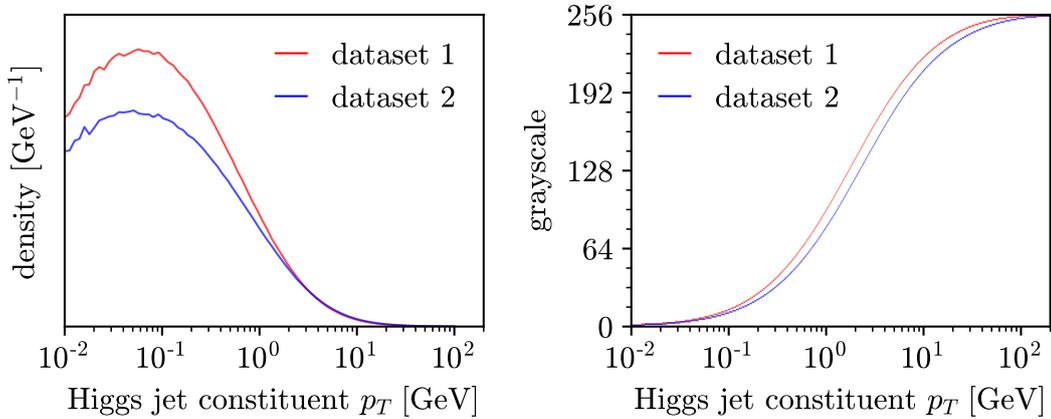


图 5.3 左：训练样本中 Higgs 喷注组分粒子的横向动量分布。右：量能器横向动量沉积对应的像素点亮度。

5.2.1 信号标定

每个对撞机图像配以一个分辨率 320×320 的二值图像作为掩膜，提供 Higgs 玻色子的位置信息（如果有多个 Higgs 玻色子，则配以多个二值图像）。信号像素点标注为 1，其余标注为 0。严格来说，只有表示 Higgs 衰变产物的单独像素点理

应被点亮，不论该粒子的横向动量是多少，这种方法最容易实现。但问题是，我们所选用的神经网络高度依赖卷积运算，不能很好的应对过于疏散的掩膜。相反地，连通的图形更适合卷积网络。尽管这类标注方式使算法更易受 **pile-up** 的影响，我们发现，由于它给出的喷注面积能比传统喷注算法小很多，**pile-up** 的影响可控。而且根据神经网络输出的掩膜，在假设 **pile-up** 分布平均的基础上对喷注进行修正，可以获得更加理想的重建效果。下面我们展示两种定义掩膜的方案。

首先对 Higgs 玻色子的衰变产物进行预筛选。将 Higgs 玻色子沿对撞机 z 轴方向做 Lorentz 提速，使其 z 方向动量为 0，仅保留能量大于 1 GeV 且位于以 Higgs 运动方向为中心的 2π 立体角内的衰变产物。如果以横向动量或粒子与 Higgs 玻色子的间距 ΔR 作为筛选依据，会使沿不同方向辐射的粒子地位不平等。这种不平等的效应在 Higgs 玻色子的快度较大或横向动量较小导致衰变产物在 $\eta - \phi$ 平面分布较广时尤为明显。

以 Higgs 玻色子的快度 y 和方位角 ϕ 作为其在 $\eta - \phi$ 平面上的坐标值，则经过预筛选的粒子散落的分布于 Higgs 玻色子周围。注意到只有快度的差 Δy 是 Lorentz 提速不变的。对衰变产物而言，由于其质量远小于能量，故快度约等于赝快度 $y \simeq \eta$ ，可相互替代。而 Higgs 玻色子质量很大，两者不可互换。在 $\eta - \phi$ 平面上，第一种掩膜定义为所有衰变粒子的凸包，即覆盖所有粒子的最小凸多边形。第二种掩膜按照以下方式构造：以 Higgs 玻色子为极点建立极坐标系，将所有衰变粒子根据极角大小依次用线段连接并首尾相连形成闭环。最终形成的辐射状不规则多边形即为掩膜。按照这种方式构造的掩膜必然小于或等于凸包。这里构造掩膜的原则是：

- 1) 掩膜的面积应该尽可能小。
- 2) Higgs 玻色子 $y - \phi$ 坐标可以被掩膜覆盖。
- 3) 保证掩膜为单连通的。

其中 1) 目的是减轻 **pile-up** 影响，2) 是用于测试时判定识别是否正确。3) 是根据所用神经网络的实测性能而附加的。全卷积网络对于分布过于稀疏分散的图形表现不佳。易见，两种掩膜满足条件 2) 3)，而对条件 1) 可能还存在优化的空间。图5.4举例展示了两种构造方式在实际数据上的结果。注意这里并没有考虑有限空间分辨率造成的像素化，实际使用的掩膜可见图5.5和图5.6中绿色轮廓线。

5.3 算法表现

使用两组数据集和两种定义方式，我们共训练了 4 个模型。训练的终止条件是网络的损失函数在验证集上不再有明显下降。最终训练的周期数为 28-30 个 epoch。

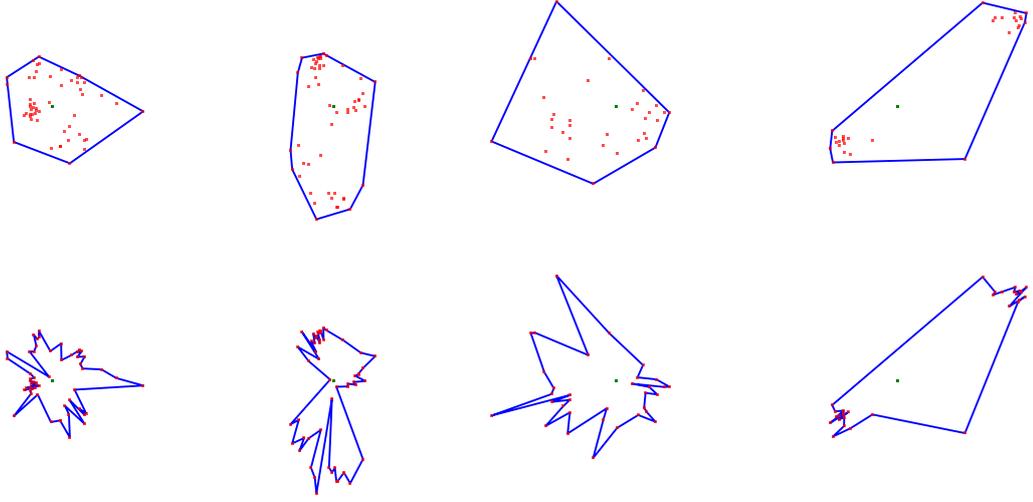


图 5.4 上：凸包形状的掩膜。下：辐射状的掩膜。绿点代表 Higgs 玻色子的 $y - \phi$ 坐标位置，红点代表筛选过的衰变产物 $\eta - \phi$ 坐标位置。

考虑到部分 Higgs 喷注的组分粒子甚至是全部粒子都落在我们构造的掩膜边界上，而卷积神经网络的输出在边界位置必然不能准确到像素级，我们将网络输出的掩膜沿边界扩大一个（凸包状）或两个像素点（辐射状）作为最终的检测结果。示例性的检测结果见图 5.5 和图 5.6。

我们从检测 Higgs 喷注的效率和 Higgs 玻色子 4-动量的重建精度等方面展示 CNN 算法在不同使用场景的表现。作为对比的基准，同样展示的还有以 Mass-Drop Tagger 和 trim 为代表的传统算法在相同数据集上的表现。Mass-Drop Tagger 的定义如下：首先用 C/A 算法以 R_0 为半径参数重建喷注 j 。然后逆向喷注的聚类过程，得到子喷注 j_1, j_2 ，其中 $m_{j_1} > m_{j_2}$ ，若它们满足以下两个条件，

$$m_{j_1} < \mu m_j \quad (5-1)$$

$$\frac{\min(p_{Tj_1}^2, p_{Tj_2}^2)}{m_j^2} \Delta R_{j_1, j_2}^2 > y_{\text{cut}} \quad (5-2)$$

则认为喷注 j 通过了筛选，否则重定义 j_1 为 j ，重复上述步骤，直至条件满足或喷注被完全分解为止。这里，我们使用了原文献中的参数 $y_{\text{cut}} = 0.09$ ， $\mu = 0.67$ 。为减轻 pile-up 影响，原文献提出了 filter 算法，即选出 3 个横向动量最大的半径 $R_{\text{sub}} = \min(0.3, R_0/2)$ 的子喷注，这样做的根据是保留两个 bottom 夸克喷注和 1 个辐射出来的胶子喷注。这里用 trim 来替代 filter，即对于一个半径 R_0 的胖喷注，将其组分重新聚类成为若干半径为 R_{sub} 的子喷注，并丢弃掉其中横向动量 $p_{T\text{sub}} < f_{\text{cut}} p_{T0}$ 的子喷注。经实验，其效果更佳。原文献中的参数选取为 $R_0 = 1.0$ ， $R_{\text{sub}} = 0.2$ ， $f_{\text{cut}} = 0.05$ 。这里，trim 的参数根据重建的 Higgs 玻色子质量分布进行

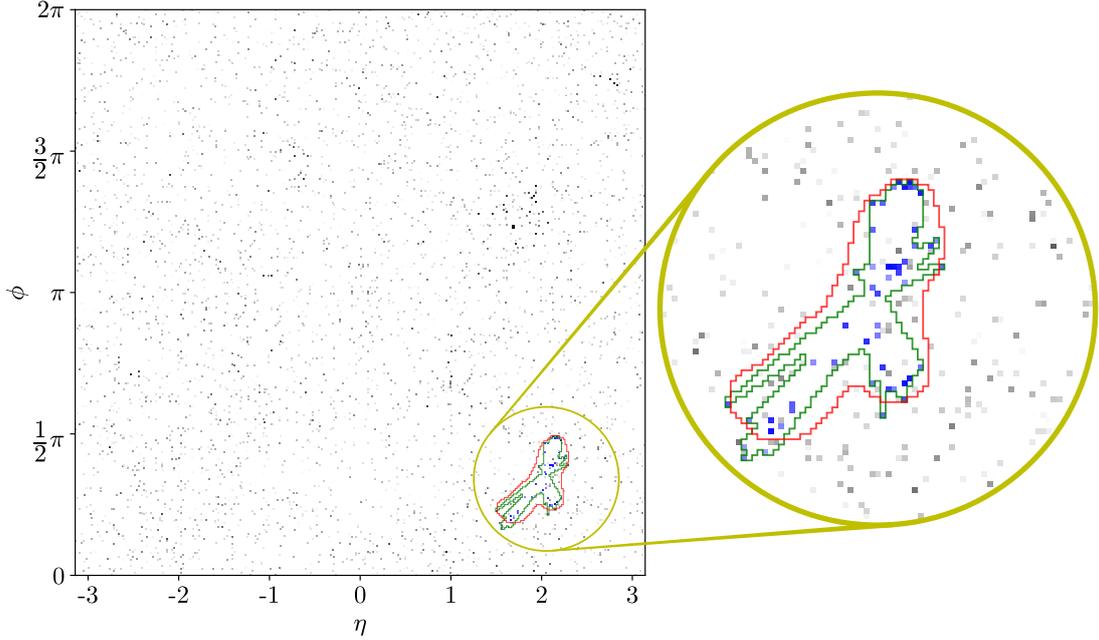


图 5.5 包含 1 个 Higgs 玻色子和 3 个 QCD 喷注的典型事例。每个像素点对应 $\Delta\eta \times \Delta\phi = 0.020 \times 0.020$ 。属于 Higgs 喷注的粒子已用蓝色加亮。黄色圆环的半径 $R = 0.8$ ，这是 C/A 喷注包围预筛选后的 Higgs 衰变产物所需最小半径。该事例中 Higgs 玻色子的横向动量 $p_T = 327$ GeV。在施行任意 groom 程序之前，这个胖喷注具有横向动量 $p_T = 417$ GeV 和质量 $m = 209$ GeV。绿色轮廓线标示了我们构造的辐射状掩膜，红色轮廓线则标示了神经网络预测的结果。为使对比更清晰，这一圆形区域被放大了 3 倍。该事例中，红色轮廓线内所有粒子重建出的喷注具有质量 $m = 118$ GeV 和横向动量 $p_T = 329$ GeV。作为对比，胖喷注经 trim 后具有质量 $m = 133$ GeV 和横向动量 $p_T = 336$ GeV。

了调整，以使 [120, 130] GeV 范围内 Higgs 玻色子数量达到最大。具体的参数选择见表 5.1。除质量分布外，重建的 Higgs 玻色子横向动量，快度和方位角的分布，共同充当评定喷注算法优劣的标准。

表 5.1 Higgs 喷注 trim 参数

| 数据集 | R_0 | R_{sub} | f_{cut} |
|-----|-------|------------------|------------------|
| 1 | 1.5 | 0.20 | 0.5 |
| 2 | 1.3 | 0.19 | 0.5 |

5.3.1 检测效率

图 5.7 展示了使用 Mask-RCNN 识别出的 Higgs 喷注质量分布与通过 MDT 和 trim 算法得到的分布对比。不论 CNN 算法还是 MDT 都可能在同一个事例中得到超过一个 Higgs 喷注。这种情况下，对于 CNN，我们选取得分最高的喷注，而对 MDT，选取横向动量最大的喷注。从图中可见：

- 1) 4 个神经网络对于 Higgs 喷注不变质量的重建均优于 MDT+trim。

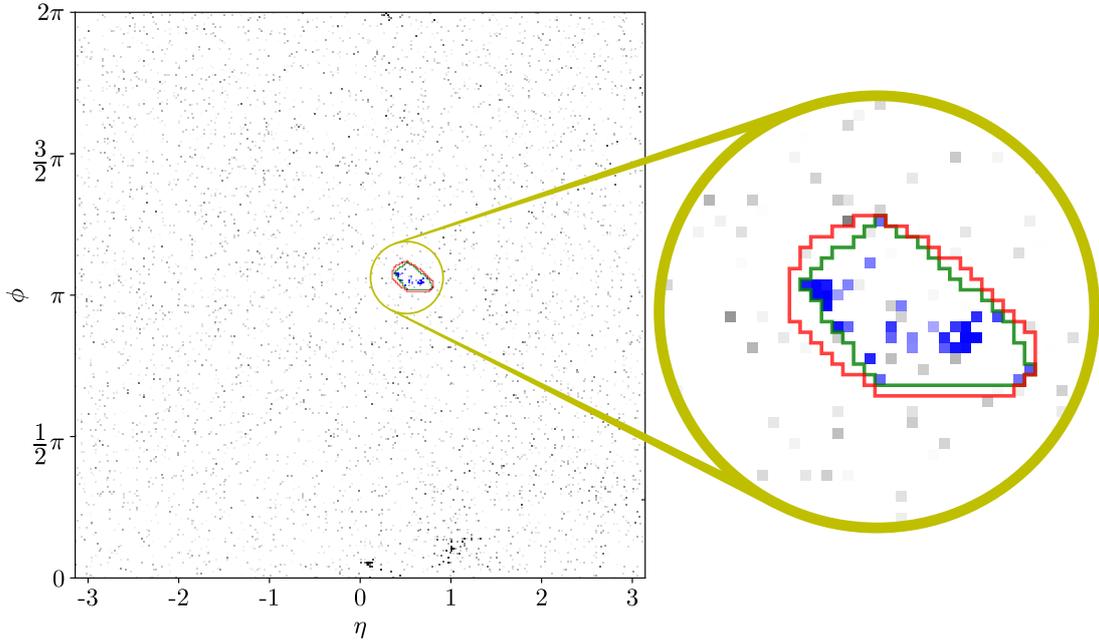


图 5.6 与图5.5类似，圆环的半径 $R = 0.4$ ，圆形区域被放大了 6 倍。信号标定和探测均使用了凸包掩膜。原始 Higgs 玻色子的横向动量 $p_T = 925$ GeV。大半径 C/A 喷注经过 trim 后的质量 $m = 127$ GeV，横向动量 $p_T = 934$ GeV。重建出的喷注具有质量 $m = 125$ GeV 和横向动量 $p_T = 927$ GeV。

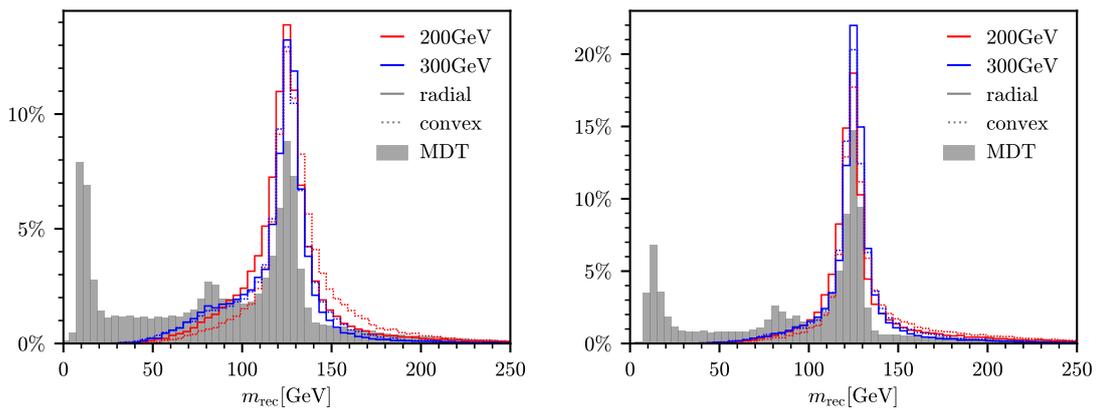


图 5.7 CNN 算法与 MDT+trim 的对比。左：测试集 1。右：测试集 2。图例中，“200 GeV”和“300 GeV”分别代表在数据集 1 和 2 上训练的网络，“radial”和“convex”分别代表使用凸包掩膜和辐射状掩膜训练的网络。灰色阴影代表经过 trim 的喷注不变质量。

2) 神经网络在与训练数据集性质相同的测试集上表现更好。在 $p_T > 200$ GeV 的数据集 1 上训练的网络，运用到 $p_T > 300$ GeV 的测试集上时，表现不如直接在 $p_T > 300$ GeV 的数据集 2 上训练的网络，反之亦然。尽管在理论上，数据集 2 应该是数据集 1 的子集。这可能与样本数不足有关。另一种可能 p_T 分布的不同造成了神经网络一定的倾向性。如果信号过程已知，可利用这种倾向性优化网络表现。但是它会损害神经网络在复合样本上的表现。

3) 使用辐射状掩膜训练的网络，表现优于凸包。

4) 神经网络对识别横向动量较小的 Higgs 喷注，提升更显著。

注意到，这里的检测效率是基于喷注的不变质量来衡量的。尽管对于 trim 算法，我们已经尽力优化了参数选取，但是更有效的 groom 算法有可能使喷注不变质量更趋于 125 GeV，从而提升传统算法的检测效率。

图5.8表示 CNN 方法相较传统算法的效率提升，定义为不变质量在 $125 \text{ GeV} \pm \Delta m$ 范围内的喷注数量比。这里只画出了训练集与测试集相匹配的情况。在数据集 1 上，辐射状掩膜的表现明显优于凸包。在数据集 2 上，这种优势不再明显。在事例筛选的过程中，往往将喷注不变质量落在一定范围内作为筛选条件。如果以 $[105, 145]$ GeV 为筛选条件，则根据图5.8，CNN 比 MDT+trim 多保留了 50% 到 80% 的信号数量。

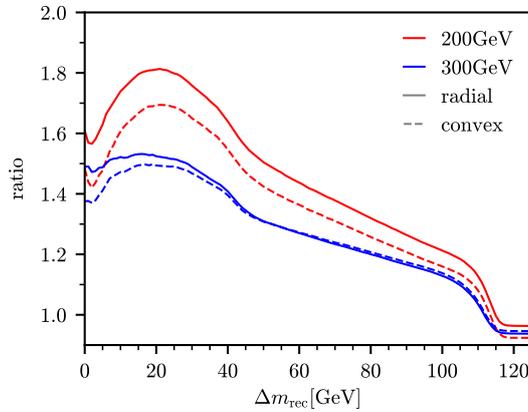


图 5.8 $m_J \in [m_H - \Delta m, m_H + \Delta m]$ 的喷注数量比值。比值越大，代表质量分布的峰越窄。

尽管该神经网络只在单一过程的数据集上训练过，即 1 个 Higgs 玻色子和 3 个 QCD 喷注，它也可以作为一个一般的 Higgs 标记器在各种事例上使用。作为演示，我们列出它在其他 3 种场景中的检测能力：

- 1) 两个 Higgs 玻色子加 3 个 QCD 喷注。
- 2) 一个 Higgs 玻色子加两个 top 夸克。

3) 一个假想的超对称过程，如下所示。

$$pp \rightarrow \tilde{t}_1^* \tilde{t}_1 \rightarrow \bar{t} \tilde{\chi}_1^0 t \tilde{\chi}_2^0 \rightarrow \bar{t} \tilde{\chi}_1^0 t H \tilde{\chi}_1^0 \quad (5-3)$$

其中 \tilde{t}_1 为轻的 stop，设其质量为 1000 GeV。 $\tilde{\chi}_1^0$ 为最轻的 neutralino，设其质量为 100 GeV， $\tilde{\chi}_2^0$ 为次轻的 neutralino，设其质量为 800 GeV。以上 3 个过程中的 Higgs 玻色子都强加了横向动量大于 200 GeV 且衰变到两个 bottom 夸克的限制。

当我们将使用辐射状掩膜训练的模型原封不动的应用到这些事例时，仍得到了可观的检测结果。作为对比，同时展示的是 MDT 方法标记正确的 Higgs 喷注。对 3 个过程使用的 trim 参数如表 5.2 所示。重建的 Higgs 玻色子质量分布如图 5.9 所示。上，中，下图依次对应三种过程。左侧对应 Higgs 玻色子横向动量 $p_T > 200$ 的测试集，使用在数据集 1 上训练的网络来检测。右侧对应 Higgs 玻色子横向动量 $p_T > 300$ 的测试集，使用在数据集 2 上训练的网络来检测。图中虚线代表 CNN 得分最高或 trim 后 p_T 最大的喷注，实线代表至少被 b -tag 两次且满足上述条件的喷注。可以看到，如果不加入 b -tag 要求，MDT 很容易受到 W 玻色子或 top 夸克的影响，使用其他子结构变量，如 N -subjettiness 可以消除此缺陷。而 CNN 将 W 玻色子或 top 夸克误识别为 Higgs 玻色子的概率明显更低。无论是否使用 b -tag，都实现了更高的识别效率。这里对 b -tag 进行了简化模拟，假设 b -tag 的正确率为 100%。

这三种代表性的过程显示，CNN 无论对于多喷注识别，还是复杂本底下的喷注识别，都可以胜任。

表 5.2 Higgs 喷注裁剪参数

| 过程 | R_0 | R_{sub} | f_{cut} |
|-----------------------------|-------|------------------|------------------|
| $HHjjj$ | 1.4 | 0.19 | 0.5 |
| Htt | 1.4 | 0.19 | 0.5 |
| $\tilde{t}_1^* \tilde{t}_1$ | 0.9 | 0.18 | 0.5 |

5.3.2 重建精度

相比聚类，标记，裁剪的算法流程，基于图像的目标检测和语义分割可以实现同等甚至更加的效率和精度，为了说明这一点，我们在图 5.10 和 5.11 中展示了重建的 Higgs 玻色子快度 y ，方位角 ϕ ，质量 m ，横向动量 p_T 等变量与原始 Higgs 玻色子的偏差。注意到这 4 个变量完全等价于粒子的 4-动量。

$$e = \sqrt{m^2 + p_T^2} \cosh y \quad (5-4)$$

$$p_x = p_T \cos \phi \quad (5-5)$$

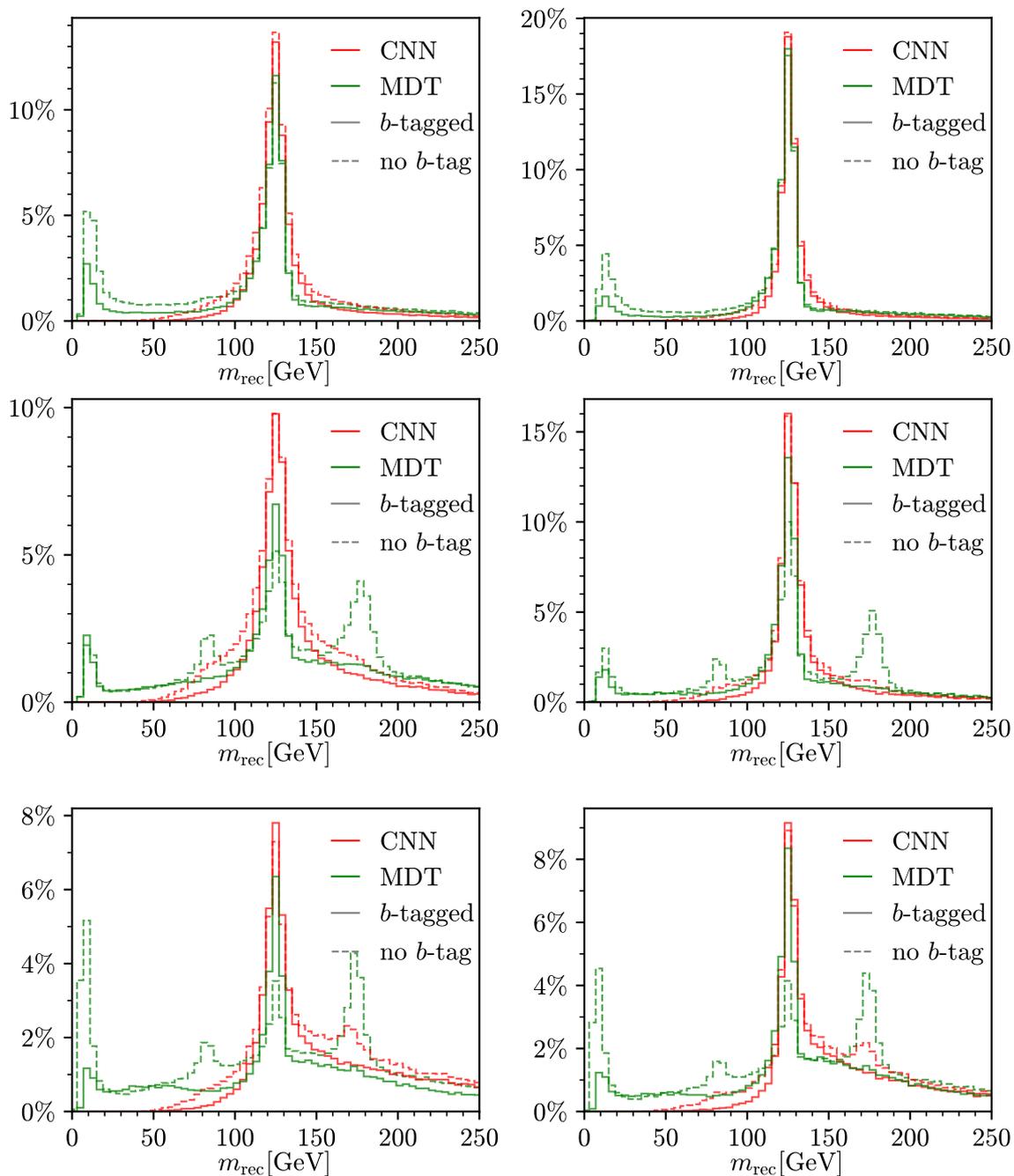


图 5.9 其它三种过程中，重建的 Higgs 喷注不变质量分布。

$$p_y = p_T \sin \phi \quad (5-6)$$

$$p_z = \sqrt{m^2 + p_T^2} \sinh y \quad (5-7)$$

其中 $\Delta m/m$ 和 $\Delta p_T/p_T$ 是相对偏差。图中只计入了标记正确的 Higgs 喷注。对于 MDT 方法，如果重建的 Higgs 玻色子与原始 Higgs 玻色子的距离小于 R_0 ，我们视其为标记正确。对于 CNN 方法，如果重建 Higgs 玻色子的 $y - \phi$ 坐标位于掩膜内，则视其为标记正确。需要注意的是，在实际的数据分析中无法使用这套判据，因为无从得知真实的原始 Higgs 玻色子 4-动量。实践当中，通常选取的是横向动量最大或神经网络得分最高的喷注。不同深度的灰色阴影和轮廓线分别代表 20%，40%，60%，和 80% 的事例数。轮廓线包围的面积越小，位置越接近中心，表示重建的精度越高。

从图中看出，平均而言，Higgs 玻色子质量的相对偏差大于其横向动量的相对偏差。对于横向动量的重建，CNN 方法表现差于 MDT 方法，主要原因是我们没有执行任何 groom 程序，尽管较小的喷注面积已经在很大程度上控制了 pile-up 的污染。横向动量的分布整体上偏大，凸包状掩膜和横向动量较低的 Higgs 更易受 pile-up 影响，因为它们的喷注面积相对更大。使用两种掩膜的 CNN 对质量的重建精度都高于 trim 算法。但是对于辐射状掩膜，对快度 y 和方位角 ϕ 的重建精度不佳。下面我们将根据 CNN 输出的掩膜对 pile-up 进行减除。注意到，上一节的检测结果都是在未减除 pile-up 的情况下得出的，原因是对质量的重建精度已经很高。

Higgs 玻色子衰变为两个夸克，在 $p_T \gg m_H$ 的条件下，

$$R_{b\bar{b}} \simeq \frac{1}{\sqrt{z(1-z)}} \frac{m_H}{p_T} \quad (5-8)$$

其中 z ， $1 - z$ 是 bottom 夸克的动量占总动量的分数。式 (5-8) 可以根据 $m_J^2 \simeq \sum_{i,j \in J} p_{Ti} p_{Tj} \Delta R_{ij}^2$ 推出。正是由于式 (5-8)，选择重建喷注的半径参数 R 时，往往以 $z = 1/2$ 时 $R_{b\bar{b}}$ 的最小值 $2m_J/p_T$ 为参考。为了尽量将喷注的组分囊括进来，以便后续分析，需要选取更大的半径参数，而后通过 groom 消除 pile-up 的影响。这增加了重建的喷注对 pile-up 污染的敏感性。而 CNN 可以对喷注图像进行语义分割，实现喷注聚类算法的功能，得到更小的喷注面积，减少对 pile-up 的敏感性。CNN 重建 Higgs 喷注的直径和面积见图 5.12。对测试集 1 和测试集 2，为了最大化识别效率，我们分别选择了半径参数 1.5 和 1.3。与之对比，CNN 给出的掩膜，其半径的平均值分别约为 0.5 和 0.7。而且，喷注聚类算法给出的喷注近似成圆形，与 CNN 输出的近似椭圆形相比，同等半径时面积更大。在图 5.12 的右侧，竖直灰色点线依次代表半径 $R=0.1, 0.2, \dots, 0.7$ 时圆形喷注的面积。使用 CNN 重建的喷注

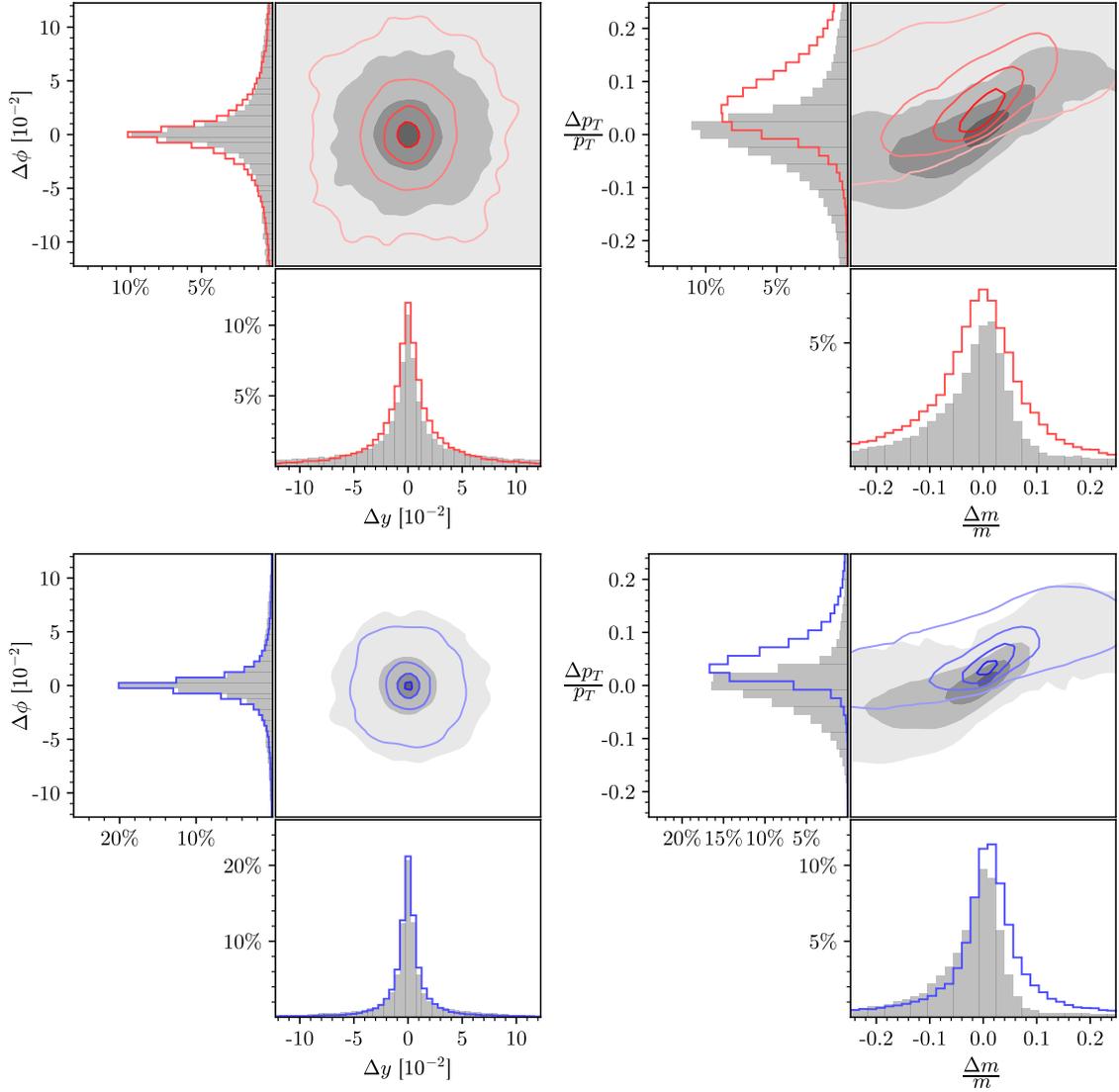


图 5.10 使用辐射状掩膜训练的网络识别 Higgs 喷注，各变量的偏差分布。上：测试集 1，下：测试集 2，左： $y - \phi$ 平面，右： $m - p_T$ 平面。

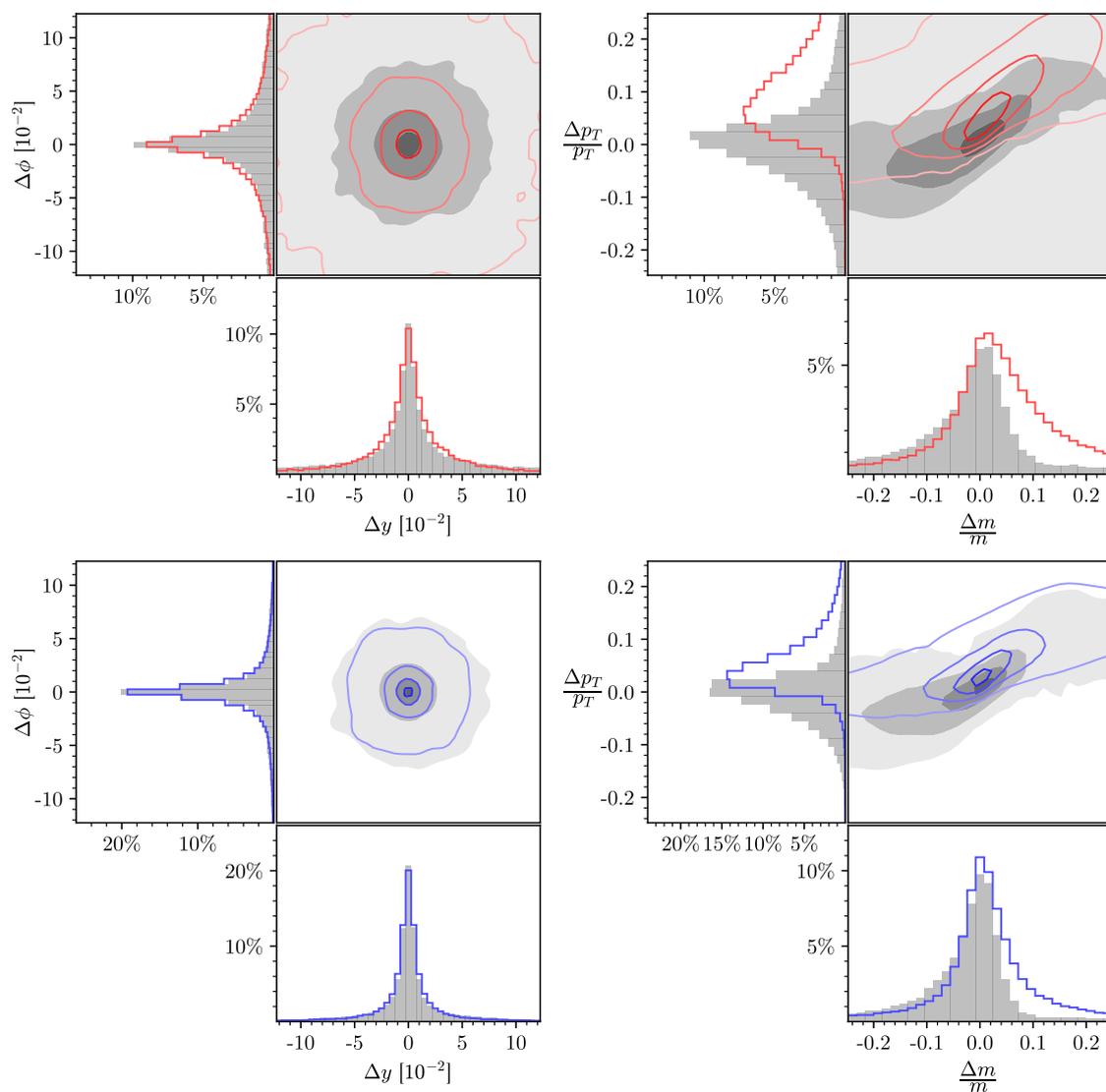


图 5.11 使用凸包掩膜训练的网络识别 Higgs 喷注，各变量的偏差分布上：测试集 1，下：测试集 2，左： $y - \phi$ 平面，右： $m - p_T$ 平面。

面积远小于 C/A ，*anti-kt* 等算法的优势显而易见。

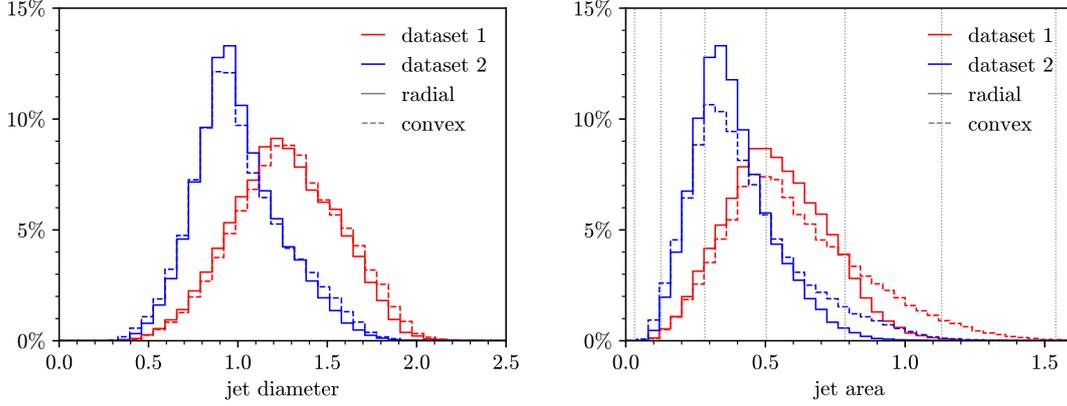


图 5.12 CNN 算法得到的喷注直径和面积分布。

对于 *pile-up* 污染，我们采取如下方式进行减除。假设 *pile-up* 的 p_T 在 $\eta-\phi$ 平面平均分布（根据模拟的数据，在每次事例伴随 50 次 *pile-up* 的情况下，该假设近似成立）。分布密度 ρ 暂时设为自由参数。根据 CNN 输出的掩膜，从喷注 4-动量中减去 *pile-up* 贡献的 4-动量，将剩余 4-动量作为修正的 Higgs 喷注。对于修正的质量和横向动量，有下列一阶近似：

$$\delta p_T \simeq -\rho A \quad (5-9)$$

$$\delta m^2 \simeq -\rho p_T A \langle \Delta R^2 \rangle \lesssim \rho p_T A \frac{\Delta R_{\max}^2}{2} \quad (5-10)$$

式中 A 为掩膜面积， ΔR 代表掩膜中一点与 Higgs 玻色子的间距， ΔR_{\max} 为间距的最大值，注意到，它并不等于掩膜直径的一半，因为 Higgs 并不一定位于喷注的中心位置。 $\langle \Delta R^2 \rangle$ 为平均到整个掩膜的 ΔR^2 值。

在事例模拟中，我们假设每个事例叠加 50 次 *pile-up*，对应的 ρ 值约等于 55–60 GeV。但在实际操作中，我们发现选取参数 $\rho = 35$ GeV，且假设减掉的 *pile-up* 不变质量为 0，能获得更好的结果。原因之一是，我们使用 CNN 得到的喷注并没有完整的包含 Higgs 玻色子的全部衰变产物。另一个原因是，用平均分布的 *pile-up* 代替离散分布的 *pile-up* 时，实际上高估了需要减除的横向动量和不变质量。上述的 *pile-up* 减除方案并非没有瑕疵，尽管如此，这套算法仍然取得了很好的实际效果。注意到，减除使用的是 CNN 输出掩膜的全部信息，而非上面的近似结果。修正后 Higgs 喷注各个变量的偏差示于图 5.13 和图 5.14。易见，所有变量的精准度都有所提高。

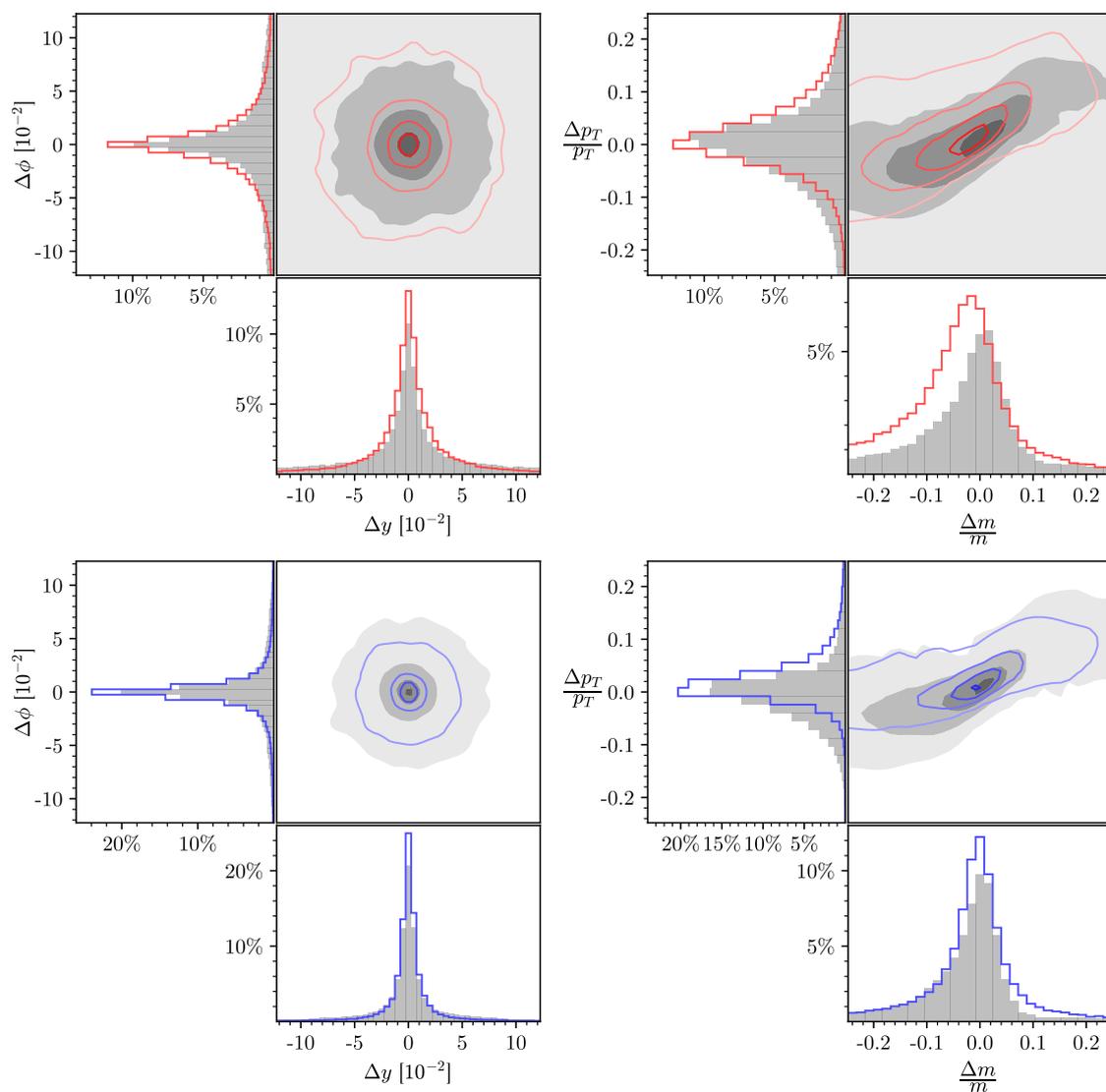


图 5.13 减除 pile-up, 修正后各变量的偏差分布。对应辐射状掩膜。上: 测试集 1, 下: 测试集 2, 左: $y - \phi$ 平面, 右: $m - p_T$ 平面。

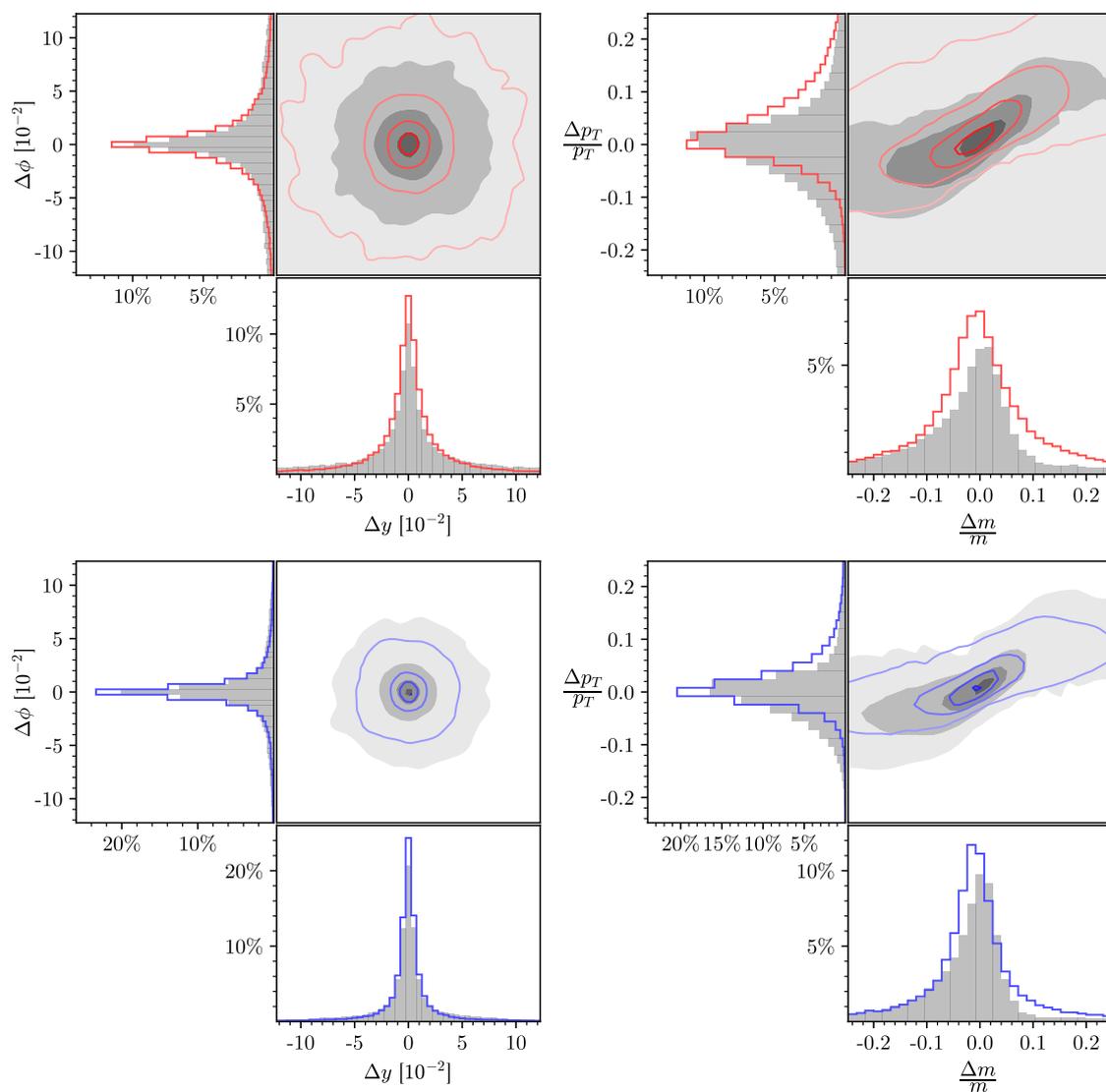


图 5.14 减除 pile-up, 修正后各变量的偏差分布。对应凸包状掩膜。上: 测试集 1, 下: 测试集 2, 左: $y - \phi$ 平面, 右: $m - p_T$ 平面。

5.3.3 区分能力

考虑到 CNN 方法的高检测率，难免担忧它会误识别，即有很高的假阳性率。这会导致它无法排除本底事例，使算法实用性大打折扣。因而下面将检查它对信号和本底的区分能力。信号事例取自数据集 1，本底事例选择 $ttjj$ ，即两个 top 夸克加两个 QCD 喷注作为代表。理想条件下，网络赋予每个喷注的分数即足以作为判别真伪的依据。从 0 到 1 连续改变阈值，即可得到一条 ROC 曲线，显示出信号与本底的区分程度。然而仅仅使用分数为判据的效果不佳，表明这一神经网络还不具备足够的能力一步到位完成此任务。使用更复杂的喷注表示和更多网络输入分支，也许可以做到这一点。这里，我们在网络输出的基础上添加一些子结构变量，训练一个 BDT 来实现信号与本底的区分。

纳入 BDT 的变量包括：喷注的横向动量 p_T ，质量 m ，横向动量与质量的比值 p_T/m ，喷注组分粒子中最大的横向动量值， N -subjettiness（这里同时用到了 τ_2/τ_1 和 τ_3/τ_2 ），喷注 p_T 加权半径 R_{weight} （与喷注 girth 基本等价）：

$$R_{\text{weight}} = \frac{\sum_i p_{Ti} \Delta R_i}{\sum_i p_{Ti}} \quad (5-11)$$

以及一个喷注形状变量，记作 f_{center} ，定义为喷注中心 $\Delta R < 0.1$ 区域内横向能量沉积相对于 $\Delta R < 0.2$ 区域内总能量的比例。

CNN 方法重建的喷注额外使用了神经网络输出的分数作为变量。CNN* 在此基础上加入了掩膜的直径和面积两个变量。完整的变量使用情况见表 5.3。

表 5.3 BDT 变量

| 变量 | MDT | CNN | CNN* |
|---------------------|-----|-----|------|
| p_T | ✓ | ✓ | ✓ |
| m | ✓ | ✓ | ✓ |
| p_T/m | ✓ | ✓ | ✓ |
| $p_{T,\text{max}}$ | ✓ | ✓ | ✓ |
| f_{center} | ✓ | ✓ | ✓ |
| R_{weight} | ✓ | ✓ | ✓ |
| τ_2/τ_1 | ✓ | ✓ | ✓ |
| τ_3/τ_2 | ✓ | ✓ | ✓ |
| 得分 | | ✓ | ✓ |
| 直径 | | | ✓ |
| 面积 | | | ✓ |

不同输入的 BDT 对包含 Higgs 玻色子的信号事例与 $ttjj$ 本底事例区分的程度

见图5.15。图中虚线代表未经 b -tag，实线代表至少有两个 b -tag。这里对 b -tag 的统计，按照喷注或子喷注是否覆盖了 bottom 夸克部分子的 $\eta - \phi$ 坐标来判定，且要求 bottom 夸克的横向动量 $p_T > 20 \text{ GeV}$ 。在此基础上，又设定 b -tag 的效率为 70%。除了 CNN 相比 MDT 有明显提升以外， b -tag 对信号与本底的区分至关重要。综合的双 b -tag 的效率约为 35%，所以 ROC 曲线无法覆盖 $\epsilon_S > 0.35$ 的区域。

与 ROC 曲线类似，显著性改进特征（significance improvement characteristic, SIC）曲线也常用来衡量算法筛选信号事例的性能。由于它与信号显著性的提升直接相关，可以更直观地反映算法性能并用来优化筛选条件。从图5.15可知，在保留大约 10% 信号事例的情况下，本底压低至不足 10^{-4} 。相比使用 MDT 喷注训练的 BDT，使用 CNN 方法可以提升信号显著性 1 倍以上。

5.4 结论

对撞机上新物理的寻找，离不开喷注的准确识别和重建。如果新粒子共振衰变的产物中包含喷注，那么显然的，重建的喷注 4-动量越接近原始部分子 4-动量，新粒子的质量测量越准确。我们以 Higgs 喷注为例，探索了新的喷注算法。在前面的章节中可以看到，类矢量粒子的重建就主要受制于 Higgs 喷注。事实上，包括双 Higgs 二重态模型在内的数个模型^[215]都预言了会衰变成两个 Higgs 玻色子的标量粒子。Randall-Sundrum 模型中自旋等于 2 的 Kaluza-Klein 引力子有相同衰变模式。这些新粒子的质量分辨率完全取决于 Higgs 喷注的重建精度。

通常，分析喷注时，首先要用 C/A 或 anti- k_t 等喷注算法对粒子聚类，然后用 trim, prune 等 groom 算法去除 pile-up 影响。喷注半径参数固定，对于重建横向动量分布较广的喷注缺乏灵活性，且喷注面积较大，易受 pile-up 污染。即使是使用各类深度神经网络识别喷注的研究，依旧离不来喷注算法。我们将整个探测器上粒子的横向动量分布编码为图像，尝试使用深度神经网络框架 Mask-RCNN 进行 Higgs 喷注的识别与重建。作为对比，使用 Mass-Drop Tagger 和 trim 完成相同任务。在 $H + 3j$ 事例中，重建质量 $|m - m_H| < 30 \text{ GeV}$ 的范围内，该算法多识别了 50-80% 的 Higgs 玻色子。在单一样本上训练的 Higgs 识别器具有很好的通用性，对于包含两个 Higgs 玻色子或 top 夸克的末态，同样表现优异。利用网络输出的掩膜进行 pile-up 减除以后，重建的 4-动量准确度明显提升。区分 $H + 3j$ 信号过程与 $ttjj$ 本底过程，信号显著性提升超 1 倍。神经网络的优势在于对横向动量分布较广的喷注能灵活聚类。

尽管该研究目前只具有概念验证的作用，距离实际运用到对撞机实验有相当的距离，但是深度学习展示出的强大功能让我们有理由相信，这项技术对于在对

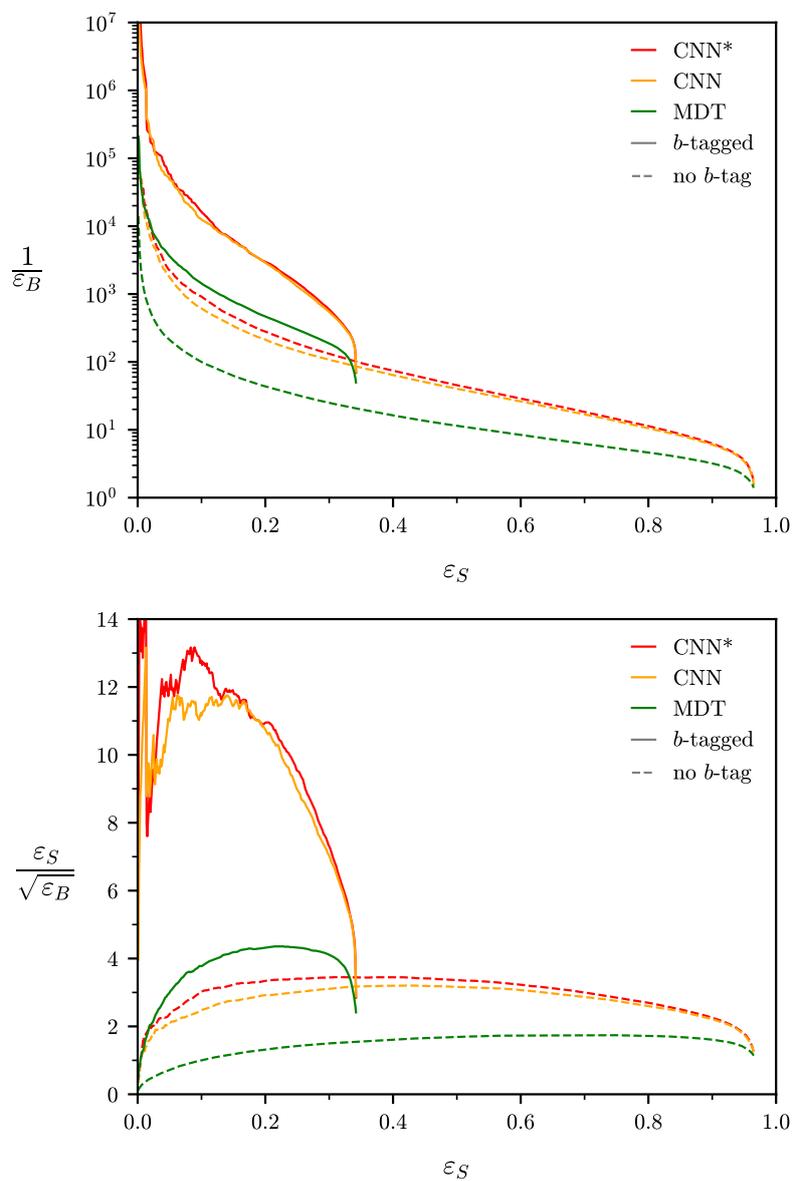


图 5.15 上: ROC 曲线。下: SIC 曲线, 对信号显著性提升最大的地方出现在 $\epsilon_S \simeq 0.09$, $\epsilon_B \simeq 5 \times 10^{-5}$ 。

撞机上探索新物理，特别是通过观测共振衰变等高度依赖喷注 4-动量重建精度的手段寻找新粒子，将大有应用前景。

第 6 章 总结

众所周知，尽管标准模型的预言与几乎所有粒子物理实验的观测相符，它不可能是描述物质世界的终极理论。探索超出标准模型的新物理是高能物理学界的共同追求。对粒子物理学的研究而言，不论是出于完善理论的动机扩展标准模型，还是为解释实验观测到的疑似反常提出唯象模型，只要理论预言在更高能标存在新粒子，最终都要接受对撞机实验的直接检验。本论文主要研究对撞机寻找不同新物理模型的理论限制和技术方法，考虑了 3 种扩展标准模型的理论在对撞机上的检验可能性，包括 split-SUSY，违反 R 宇称守恒的 SUSY， $U(1)_X$ 模型，以及一种数据分析手段，即借助卷积神经网络处理包含喷注的事例。

粒子物理学家们原本对在 LHC 上发现超对称寄予厚望，然而目标信号的缺席将超对称粒子的质量下限一再推高，也迫使人们转而考虑理论更复杂的版本。本文从 Higgs 玻色子质量和电弱真空稳定性出发，使用最新的 top 夸克质量和强相互作用耦合常数的观测值，计算了 split-SUSY 情形下，超对称破缺能标和 squark 质量的上限。对 $\tan\beta = 2$ ， $\tan\beta = 4$ ， $\tan\beta = 50$ 的情形分别得出 $M_S = 10^{6.2}$ ， $10^{4.5}$ ， $10^{4.3}$ GeV。特别是， $\tan\beta \gtrsim 4$ 的 split-SUSY 有可能在未来的 100 TeV 质子-质子对撞机，如 FCC-hh 和 SppC 上检验。对 $m_h = 123$ ，125，和 127 GeV，Higgs 场四次耦合系数 λ 分别在 $10^{8.2}$ ， $10^{8.7}$ ，和 $10^{9.6}$ GeV 变负，由于额外的 Higgs-Higgsino-gaugino 耦合，这些限制比标准模型小一个数量级左右。另外，我们讨论了在未来对撞机上搜索不同寿命的 gluino 以探索超对称破缺能标的前景。

近些年来，对撞机实验上持续观测到 B 介子半轻子衰变与标准模型预言的不符，这些过程与味道改变中性流 $b \rightarrow sll$ 密切相关。本文研究了可解释 B 物理反常的 $U(1)_X$ 模型。根据 B 反常的全局拟合结果和 B_s 振荡的观测数据等，给出新粒子的质量上限， $m_{Z'} \lesssim 1.8$ TeV， $m_E \lesssim 9.6$ TeV。指出直接搜索 Z' 玻色子的局限性：由于 Z' 与夸克耦合较弱，不论其质量大小，在高亮度 LHC 上通过双轻子道搜寻其共振衰变都有可能一无所获。而带电的左手类矢量轻子 E ，与 γ ， Z 的耦合强度同标准模型轻子一致，产生截面只有质量压低。计算了 $m_{Z'}$ 和标量场混合角 $\theta_{H-\phi}$ 对 E 衰变分支比的影响。据此，对 $m_E > m_{Z'}$ ， $m_E < m_{Z'}$ 两种情形，分别利用 6μ 子末态和双 μ 子加 $4b$ 夸克末态，分析了未来高亮度 LHC 对 m_E 的搜寻灵敏度，得到了 1.4 TeV 和 1 TeV 的探测极限。

目前对于超对称中的强子 R 宇称破坏项，限制依然比较弱。在此情形下，neutralino 衰变为 3 个夸克，可形成具有子结构的喷注。提取事例中的喷注，用图像表

示能量分布，借助卷积神经网络，可以将 **neutralino** 喷注与 QCD 喷注进行一定程度地区分，其效果不弱于传统的子结构变量， N -subjettiness。结合喷注不变质量的信息，该方法即可用于识别质量在较广范围内变动的 **neutralino**。将此方法运用到 **gluino** 的搜寻当中，并利用提升决策树分析事例中的多维信息，可以改进 ATLAS 合作组使用多喷注末态搜索 **gluino** 的分析。对质量 $m_{\tilde{\chi}_1^0} = 100$ GeV 的 **neutralino**，质量 $m_{\tilde{g}} \lesssim 1.4$ TeV 的 **gluino** 可以被排除。根据 **neutralino** 的质量不同，这一方法可以将 ATLAS 合作组对 **gluino** 的质量排除极限提高 100 到 250 GeV。

对撞机上寻找重的新粒子涉及衰变产物的重建。这其中，喷注的重建精度往往限制了重建的新粒子质量分辨率，而喷注重建受喷注聚类算法和 **pile-up** 污染的制约。本文以 **Higgs** 喷注的识别和重建为例，探索使用深度神经网络对整个事例的数据进行分析。用图像来编码末态粒子在整个探测器的横向动量分布，利用 **Mask-RCNN** 架构的语义分割和目标检测能力，代替喷注聚类算法和喷注子结构变量的功能。该方法有很好的泛化能力，对不同事例中不同能量的 **Higgs** 喷注都取得了更好的识别和重建效果。采用特有的 **pile-up** 减除方案后，重建的喷注 4-动量更接近初态 **Higgs** 玻色子的 4-动量。对于 **Higgs** 玻色子加多个 QCD 喷注的事例，以喷注不变质量为筛选判据，可以增加多达 50% 的探测效率。使用提升决策树，还可以更好的区分信号与本底，最多实现一倍以上的信号显著性提高。该研究目前具有概念验证的作用，距离实验采纳有相当距离。但是通过研究可以发现，深度学习方法在对撞机唯象研究领域具有巨大潜力。日后在 LHC 上和未来的对撞机上寻找新物理，这些方法应用前景广阔。

参考文献

- [1] Schwartz M D. Quantum field theory and the standard model. Cambridge University Press, 2013.
- [2] Aad G, et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 2012, 716:1-29.
- [3] Chatrchyan S, et al. Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC. *Phys. Lett. B*, 2012, 716:30-61.
- [4] Ade P, et al. Planck 2015 results. XIII. Cosmological parameters. *Astron. Astrophys.*, 2016, 594:A13.
- [5] Canetti L, Drewes M, Shaposhnikov M. Matter and Antimatter in the Universe. *New J. Phys.*, 2012, 14:095012.
- [6] Peccei R. The Strong CP problem and axions. *Lect. Notes Phys.*, 2008, 741:3-17.
- [7] Fukuda Y, et al. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 1998, 81: 1562-1567.
- [8] An F, et al. Observation of electron-antineutrino disappearance at Daya Bay. *Phys. Rev. Lett.*, 2012, 108:171803.
- [9] Langacker P. Grand Unified Theories and Proton Decay. *Phys. Rept.*, 1981, 72:185.
- [10] Martin S P. A Supersymmetry primer. *Perspectives on supersymmetry. Vol.2*, 2010:1-153.
- [11] Ellwanger U, Hugonie C, Teixeira A M. The Next-to-Minimal Supersymmetric Standard Model. *Phys. Rept.*, 2010, 496:1-77.
- [12] Witten E. Search for a Realistic Kaluza-Klein Theory. *Nucl. Phys. B*, 1981, 186:412.
- [13] Randall L, Sundrum R. An Alternative to compactification. *Phys. Rev. Lett.*, 1999, 83:4690-4693.
- [14] Randall L, Sundrum R. A Large mass hierarchy from a small extra dimension. *Phys. Rev. Lett.*, 1999, 83:3370-3373.
- [15] Polchinski J. String theory: volume 1. Cambridge University Press, 1998.
- [16] Polchinski J. String theory: volume 2. Cambridge University Press, 1998.
- [17] Branco G, Ferreira P, Lavoura L, et al. Theory and phenomenology of two-Higgs-doublet models. *Phys. Rept.*, 2012, 516:1-102.
- [18] Pati J C, Salam A. Lepton Number as the Fourth Color. *Phys. Rev. D*, 1974, 10:275-289.
- [19] Peccei R, Quinn H R. CP Conservation in the Presence of Instantons. *Phys. Rev. Lett.*, 1977, 38:1440-1443.
- [20] Bertone G, Hooper D, Silk J. Particle dark matter: Evidence, candidates and constraints. *Phys. Rept.*, 2005, 405:279-390.
- [21] King S F. Models of Neutrino Mass, Mixing and CP Violation. *J. Phys. G*, 2015, 42:123001.
- [22] Abe K, et al. Search for proton decay via $p \rightarrow e^+\pi^0$ and $p \rightarrow \mu^+\pi^0$ in 0.31 megaton \cdot years exposure of the Super-Kamiokande water Cherenkov detector. *Phys. Rev. D*, 2017, 95(1):012004.

- [23] Braine T, et al. Extended Search for the Invisible Axion with the Axion Dark Matter Experiment. *Phys. Rev. Lett.*, 2020, 124(10):101303.
- [24] Jiang H, et al. Limits on Light Weakly Interacting Massive Particles from the First 102.8 kg \times day Data of the CDEX-10 Experiment. *Phys. Rev. Lett.*, 2018, 120(24):241301.
- [25] Cui X, et al. Dark Matter Results From 54-Ton-Day Exposure of PandaX-II Experiment. *Phys. Rev. Lett.*, 2017, 119(18):181302.
- [26] Mitsou V A. SUSY searches in ATLAS and CMS. *PoS*, 2019, CORFU2019.
- [27] ATLAS. Search for resonances decaying to photon pairs in 3.2 fb^{-1} of pp collisions at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector. ATLAS-CONF-2015-081, 2015.
- [28] CMS. Search for new physics in high mass diphoton events in proton-proton collisions at 13TeV. CMS-PAS-EXO-15-004, 2015.
- [29] ATLAS. Search for scalar diphoton resonances with $15.4 \sim \text{fb}^{-1}$ of data collected at $\sqrt{s}=13 \text{ TeV}$ in 2015 and 2016 with the ATLAS detector. ATLAS-CONF-2016-059, 2016.
- [30] CMS. Search for resonant production of high mass photon pairs using 12.9 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$ and combined interpretation of searches at 8 and 13 TeV. CMS-PAS-EXO-16-027, 2016.
- [31] Aaij R, et al. Measurement of Form-Factor-Independent Observables in the Decay $B^0 \rightarrow K^{*0} \mu^+ \mu^-$. *Phys. Rev. Lett.*, 2013, 111:191801.
- [32] The rise of the standard model: A history of particle physics from 1964 to 1979. Cambridge University Press, 1997.
- [33] Cabibbo N. Unitary Symmetry and Leptonic Decays. *Phys. Rev. Lett.*, 1963, 10:531-533.
- [34] Kobayashi M, Maskawa T. CP Violation in the Renormalizable Theory of Weak Interaction. *Prog. Theor. Phys.*, 1973, 49:652-657.
- [35] Pontecorvo B. Inverse beta processes and nonconservation of lepton charge. *Sov. Phys. JETP*, 1958, 7:172-173.
- [36] Maki Z, Nakagawa M, Sakata S. Remarks on the unified model of elementary particles. *Prog. Theor. Phys.*, 1962, 28:870-880.
- [37] Georgi H, Glashow S. Unity of All Elementary Particle Forces. *Phys. Rev. Lett.*, 1974, 32: 438-441.
- [38] Barbier R, et al. R-parity violating supersymmetry. *Phys. Rept.*, 2005, 420:1-202.
- [39] Wells J D. PeV-scale supersymmetry. *Phys. Rev. D*, 2005, 71:015013.
- [40] Arkani-Hamed N, Dimopoulos S. Supersymmetric unification without low energy supersymmetry and signatures for fine-tuning at the LHC. *JHEP*, 2005, 06:073.
- [41] Giudice G, Romanino A. Split supersymmetry. *Nucl. Phys. B*, 2004, 699:65-89.
- [42] Aaij R, et al. Differential branching fractions and isospin asymmetries of $B \rightarrow K^{(*)} \mu^+ \mu^-$ decays. *JHEP*, 2014, 06:133.
- [43] Aaij R, et al. Measurements of the S-wave fraction in $B^0 \rightarrow K^+ \pi^- \mu^+ \mu^-$ decays and the $B^0 \rightarrow K^{*}(892)^0 \mu^+ \mu^-$ differential branching fraction. *JHEP*, 2016, 11:047.
- [44] Aaij R, et al. Angular analysis and differential branching fraction of the decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$. *JHEP*, 2015, 09:179.

-
- [45] Aaij R, et al. Angular analysis of the $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decay using 3 fb^{-1} of integrated luminosity. JHEP, 2016, 02:104.
- [46] Wehle S, et al. Lepton-Flavor-Dependent Angular Analysis of $B \rightarrow K^* \ell^+ \ell^-$. Phys. Rev. Lett., 2017, 118(11):111801.
- [47] Bordone M, Isidori G, Pattori A. On the Standard Model predictions for R_K and R_{K^*} . Eur. Phys. J., 2016, C76(8):440.
- [48] Aaij R, et al. Test of lepton universality using $B^+ \rightarrow K^+ \ell^+ \ell^-$ decays. Phys. Rev. Lett., 2014, 113:151601.
- [49] Aaij R, et al. Test of lepton universality with $B^0 \rightarrow K^{*0} \ell^+ \ell^-$ decays. JHEP, 2017, 08:055.
- [50] Graverini E. Flavour anomalies: a review. J. Phys. Conf. Ser., 2019, 1137(1):012025.
- [51] Glashow S, Iliopoulos J, Maiani L. Weak Interactions with Lepton-Hadron Symmetry. Phys. Rev. D, 1970, 2:1285-1292.
- [52] Altmannshofer W, Paradisi P, Straub D M. Model-Independent Constraints on New Physics in $b \rightarrow s$ Transitions. JHEP, 2012, 04:008.
- [53] Alok A K, Bhattacharya B, Kumar D, et al. New physics in $b \rightarrow s \mu^+ \mu^-$: Distinguishing models through CP-violating effects. Phys. Rev., 2017, D96(1):015034.
- [54] Capdevila B, Crivellin A, Descotes-Genon S, et al. Patterns of New Physics in $b \rightarrow s \ell^+ \ell^-$ transitions in the light of recent data. JHEP, 2018, 01:093.
- [55] Altmannshofer W, Stangl P, Straub D M. Interpreting Hints for Lepton Flavor Universality Violation. Phys. Rev., 2017, D96(5):055008.
- [56] D'Amico G, Nardecchia M, Panci P, et al. Flavour anomalies after the R_{K^*} measurement. JHEP, 2017, 09:010.
- [57] Geng L S, Grinstein B, Jäger S, et al. Towards the discovery of new physics with lepton-universality ratios of $b \rightarrow s \ell \ell$ decays. Phys. Rev., 2017, D96(9):093006.
- [58] Ciuchini M, Coutinho A M, Fedele M, et al. On Flavourful Easter eggs for New Physics hunger and Lepton Flavour Universality violation. Eur. Phys. J., 2017, C77(10):688.
- [59] Alok A K, Bhattacharya B, Datta A, et al. New Physics in $b \rightarrow s \mu^+ \mu^-$ after the Measurement of R_{K^*} . Phys. Rev., 2017, D96(9):095009.
- [60] Altmannshofer W, Straub D M. New Physics in $B \rightarrow K^* \mu \mu$? Eur. Phys. J., 2013, C73:2646.
- [61] Gauld R, Goertz F, Haisch U. On minimal Z' explanations of the $B \rightarrow K^* \mu^+ \mu^-$ anomaly. Phys. Rev., 2014, D89:015005.
- [62] Buras A J, Girrbach J. Left-handed Z' and Z FCNC quark couplings facing new $b \rightarrow s \mu^+ \mu^-$ data. JHEP, 2013, 12:009.
- [63] Crivellin A, Hofer L, Matias J, et al. Lepton-flavour violating B decays in generic Z' models. Phys. Rev., 2015, D92(5):054013.
- [64] Bečirević D, Sumensari O, Zukanovich Funchal R. Lepton flavor violation in exclusive $b \rightarrow s$ decays. Eur. Phys. J., 2016, C76(3):134.
- [65] Altmannshofer W, Gori S, Pospelov M, et al. Quark flavor transitions in $L_\mu - L_\tau$ models. Phys. Rev., 2014, D89:095033.
- [66] Bečirević D, Fajfer S, Košnik N, et al. Leptoquark model to explain the B -physics anomalies, R_K and R_D . Phys. Rev. D, 2016, 94(11):115021.

-
- [67] Fajfer S, Košnik N. Vector leptoquark resolution of R_K and $R_{D^{(*)}}$ puzzles. *Phys. Lett. B*, 2016, 755:270-274.
- [68] Calibbi L, Crivellin A, Li T. Model of vector leptoquarks in view of the B -physics anomalies. *Phys. Rev. D*, 2018, 98(11):115002.
- [69] Fornal B, Gadam S A, Grinstein B. Left-Right SU(4) Vector Leptoquark Model for Flavor Anomalies. *Phys. Rev. D*, 2019, 99(5):055025.
- [70] Di Luzio L, Greljo A, Nardecchia M. Gauge leptoquark as the origin of B-physics anomalies. *Phys. Rev. D*, 2017, 96(11):115011.
- [71] Niehoff C, Stangl P, Straub D M. Violation of lepton flavour universality in composite Higgs models. *Phys. Lett.*, 2015, B747:182-186.
- [72] Carmona A, Goertz F. Lepton Flavor and Nonuniversality from Minimal Composite Higgs Setups. *Phys. Rev. Lett.*, 2016, 116(25):251801.
- [73] Carmona A, Goertz F. Recent B physics anomalies: a first hint for compositeness? *Eur. Phys. J. C*, 2018, 78(11):979.
- [74] Hanson G, et al. Evidence for Jet Structure in Hadron Production by $e^+ e^-$ Annihilation. *Phys. Rev. Lett.*, 1975, 35:1609-1612.
- [75] Blazey G C, et al. Run II jet physics. *Physics at Run II: QCD and Weak Boson Physics Workshop: Final General Meeting*, 2000:47-77.
- [76] Ellis S, Huston J, Hatakeyama K, et al. Jets in hadron-hadron collisions. *Prog. Part. Nucl. Phys.*, 2008, 60:484-551.
- [77] Serman G F, Weinberg S. Jets from Quantum Chromodynamics. *Phys. Rev. Lett.*, 1977, 39:1436.
- [78] Salam G P. Towards Jetography. *Eur. Phys. J. C*, 2010, 67:637-686.
- [79] Dokshitzer Y L, Leder G D, Moretti S, et al. Better jet clustering algorithms. *JHEP*, 1997, 08:001.
- [80] Catani S, Dokshitzer Y L, Seymour M, et al. Longitudinally invariant K_t clustering algorithms for hadron hadron collisions. *Nucl. Phys. B*, 1993, 406:187-224.
- [81] Cacciari M, Salam G P, Soyez G. The Anti- $k(t)$ jet clustering algorithm. *JHEP*, 2008, 04:063.
- [82] Butterworth J M, Davison A R, Rubin M, et al. Jet substructure as a new Higgs search channel at the LHC. *Phys. Rev. Lett.*, 2008, 100:242001.
- [83] Chatrchyan S, et al. The CMS Experiment at the CERN LHC. *JINST*, 2008, 3:S08004.
- [84] Kaplan D E, Rehermann K, Schwartz M D, et al. Top Tagging: A Method for Identifying Boosted Hadronically Decaying Top Quarks. *Phys. Rev. Lett.*, 2008, 101:142001.
- [85] Thaler J, Wang L T. Strategies to Identify Boosted Tops. *JHEP*, 2008, 07:092.
- [86] Plehn T, Salam G P, Spannowsky M. Fat Jets for a Light Higgs. *Phys. Rev. Lett.*, 2010, 104:111801.
- [87] Plehn T, Spannowsky M, Takeuchi M, et al. Stop Reconstruction with Tagged Tops. *JHEP*, 2010, 10:078.
- [88] Cui Y, Han Z, Schwartz M D. W-jet Tagging: Optimizing the Identification of Boosted Hadronically-Decaying W Bosons. *Phys. Rev. D*, 2011, 83:074023.

-
- [89] Khachatryan V, et al. Identification techniques for highly boosted W bosons that decay into hadrons. *JHEP*, 2014, 12:017.
- [90] CMS. Boosted Top Jet Tagging at CMS. CMS-PAS-JME-13-007, 2014.
- [91] Seymour M. Jet shapes in hadron collisions: Higher orders, resummation and hadronization. *Nucl. Phys. B*, 1998, 513:269-300.
- [92] Berger C F, Kucs T, Sterman G F. Event shape / energy flow correlations. *Phys. Rev. D*, 2003, 68:014012.
- [93] Almeida L G, Lee S J, Perez G, et al. Substructure of high- p_T Jets at the LHC. *Phys. Rev. D*, 2009, 79:074017.
- [94] Larkoski A J, Thaler J, Waalewijn W J. Gaining (Mutual) Information about Quark/Gluon Discrimination. *JHEP*, 2014, 11:129.
- [95] Rakow P E, Webber B. Transverse Momentum Moments of Hadron Distributions in QCD Jets. *Nucl. Phys. B*, 1981, 191:63-74.
- [96] Farhi E. Quantum chromodynamics test for jets. *Phys. Rev. Lett.*, 1977, 39:1587-1588.
- [97] Dasgupta M, Fregoso A, Marzani S, et al. Towards an understanding of jet substructure. *JHEP*, 2013, 09:029.
- [98] Larkoski A J, Marzani S, Soyez G, et al. Soft Drop. *JHEP*, 2014, 05:146.
- [99] Dreyer F A, Necib L, Soyez G, et al. Recursive Soft Drop. *JHEP*, 2018, 06:093.
- [100] Thaler J, Van Tilburg K. Identifying Boosted Objects with N-subjettiness. *JHEP*, 2011, 03:015.
- [101] Larkoski A J, Salam G P, Thaler J. Energy Correlation Functions for Jet Substructure. *JHEP*, 2013, 06:108.
- [102] Marzani S, Soyez G, Spannowsky M. Looking inside jets: an introduction to jet substructure and boosted-object phenomenology. Springer, 2019.
- [103] Krohn D, Thaler J, Wang L T. Jet Trimming. *JHEP*, 2010, 02:084.
- [104] Ellis S D, Vermilion C K, Walsh J R. Techniques for improved heavy particle searches with jet substructure. *Phys. Rev. D*, 2009, 80:051501.
- [105] Cacciari M, Salam G P, Soyez G. SoftKiller, a particle-level pileup removal method. *Eur. Phys. J. C*, 2015, 75(2):59.
- [106] Bertolini D, Harris P, Low M, et al. Pileup Per Particle Identification. *JHEP*, 2014, 10:059.
- [107] Roe B P, Yang H J, Zhu J, et al. Boosted decision trees, an alternative to artificial neural networks. *Nucl. Instrum. Meth.*, 2005, A543(2-3):577-584.
- [108] Chatrchyan S, et al. Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC. *Phys. Lett.*, 2012, B716:30-61.
- [109] Chen T, Guestrin C. Xgboost: A scalable tree boosting system. Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, 2016:785-794.
- [110] Adam-Bourdarios C, Cowan G, Germain-Renaud C, et al. The higgs machine learning challenge. *Journal of Physics: Conference Series*, 2015, 664(7):072015.
- [111] Baldi P, Sadowski P, Whiteson D. Searching for Exotic Particles in High-Energy Physics with Deep Learning. *Nature Commun.*, 2014, 5:4308.
- [112] Baldi P, Sadowski P, Whiteson D. Enhanced Higgs Boson to $\tau^+\tau^-$ Search with Deep Learning. *Phys. Rev. Lett.*, 2015, 114(11):111801.

-
- [113] Cogan J, Kagan M, Strauss E, et al. Jet-Images: Computer Vision Inspired Techniques for Jet Tagging. *JHEP*, 2015, 02:118.
- [114] Guest D, Collado J, Baldi P, et al. Jet Flavor Classification in High-Energy Physics with Deep Neural Networks. *Phys. Rev. D*, 2016, 94(11):112002.
- [115] Louppe G, Cho K, Becot C, et al. QCD-Aware Recursive Neural Networks for Jet Physics. *JHEP*, 2019, 01:057.
- [116] Henrion I, Brehmer J, Bruna J, et al. Neural message passing for jet physics. *Deep Learning for Physical Sciences Workshop at the 31st Conference on Neural Information Processing Systems (NeurIPS)*, 2017.
- [117] Almeida L G, Backović M, Cliche M, et al. Playing Tag with ANN: Boosted Top Identification with Pattern Recognition. *JHEP*, 2015, 07:086.
- [118] Kasieczka G, Plehn T, Russell M, et al. Deep-learning Top Taggers or The End of QCD? *JHEP*, 2017, 05:006.
- [119] Butter A, et al. The Machine Learning Landscape of Top Taggers. *SciPost Phys.*, 2019, 7:014.
- [120] de Oliveira L, Kagan M, Mackey L, et al. Jet-images — deep learning edition. *JHEP*, 2016, 07:069.
- [121] Baldi P, Bauer K, Eng C, et al. Jet Substructure Classification in High-Energy Physics with Deep Neural Networks. *Phys. Rev. D*, 2016, 93(9):094034.
- [122] Lin J, Freytsis M, Moutl I, et al. Boosting $H \rightarrow b\bar{b}$ with Machine Learning. *JHEP*, 2018, 10:101.
- [123] Cheng T. Recursive Neural Networks in Quark/Gluon Tagging. *Comput. Softw. Big Sci.*, 2018, 2(1):3.
- [124] ATLAS. Quark versus Gluon Jet Tagging Using Jet Images with the ATLAS Detector. *ATL-PHYS-PUB-2017-017*, 2017.
- [125] Komiske P T, Metodiev E M, Nachman B, et al. Pileup Mitigation with Machine Learning (PUMML). *JHEP*, 2017, 12:051.
- [126] Arjona Martínez J, Cerri O, Pierini M, et al. Pileup mitigation at the Large Hadron Collider with graph neural networks. *Eur. Phys. J. Plus*, 2019, 134(7):333.
- [127] Hajer J, Li Y Y, Liu T, et al. Novelty Detection Meets Collider Physics. *Phys. Rev. D*, 2020, 101(7):076015.
- [128] Farina M, Nakai Y, Shih D. Searching for New Physics with Deep Autoencoders. *Phys. Rev. D*, 2020, 101(7):075021.
- [129] Heindel T, Kasieczka G, Plehn T, et al. QCD or What? *SciPost Phys.*, 2019, 6(3):030.
- [130] Cerri O, Nguyen T Q, Pierini M, et al. Variational Autoencoders for New Physics Mining at the Large Hadron Collider. *JHEP*, 2019, 05:036.
- [131] Collins J H, Howe K, Nachman B. Anomaly Detection for Resonant New Physics with Machine Learning. *Phys. Rev. Lett.*, 2018, 121(24):241803.
- [132] Collins J H, Howe K, Nachman B. Extending the search for new resonances with machine learning. *Phys. Rev. D*, 2019, 99(1):014038.
- [133] Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators. *Neural Networks*, 1989, 2(5):359 - 366.

- [134] Cybenko G. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 1989, 2(4):303 - 314.
- [135] Lu Z, Pu H, Wang F, et al. The expressive power of neural networks: A view from the width. *Advances in Neural Information Processing Systems* 30, 2017:6231-6239.
- [136] Leggett T, Li T, Maxin J A, et al. Confronting Electroweak Fine-tuning with No-Scale Supergravity. *Phys. Lett. B*, 2015, 740:66-72.
- [137] Du G, Li T, Nanopoulos D, et al. Super-Natural MSSM. *Phys. Rev. D*, 2015, 92(2):025038.
- [138] Lahanas A, Nanopoulos D V. The Road to No Scale Supergravity. *Phys. Rept.*, 1987, 145:1.
- [139] Giudice G, Masiero A. A Natural Solution to the mu Problem in Supergravity Theories. *Phys. Lett. B*, 1988, 206:480-484.
- [140] Cid Vidal X, et al. Report from Working Group 3: Beyond the Standard Model physics at the HL-LHC and HE-LHC. Report on the Physics at the HL-LHC, and Perspectives for the HE-LHC, 2019:585-865.
- [141] Collaboration C. CMS at the High-Energy Frontier. Contribution to the Update of the European Strategy for Particle Physics. CMS-NOTE-2012-006, 2012.
- [142] Cohen T, Golling T, Hance M, et al. SUSY Simplified Models at 14, 33, and 100 TeV Proton Colliders. *JHEP*, 2014, 04:117.
- [143] Golling T, et al. Physics at a 100 TeV pp collider: beyond the Standard Model phenomena. *CERN Yellow Rep.*, 2017(3):441-634.
- [144] Degrassi G, Di Vita S, Elias-Miro J, et al. Higgs mass and vacuum stability in the Standard Model at NNLO. *JHEP*, 2012, 08:098.
- [145] Tang Y. Vacuum Stability in the Standard Model. *Mod. Phys. Lett. A*, 2013, 28:1330002.
- [146] Buttazzo D, Degrassi G, Giardino P P, et al. Investigating the near-criticality of the Higgs boson. *JHEP*, 2013, 12:089.
- [147] Barate R, et al. Search for the standard model Higgs boson at LEP. *Phys. Lett. B*, 2003, 565: 61-75.
- [148] Casas J, Espinosa J, Quiros M, et al. The Lightest Higgs boson mass in the minimal supersymmetric standard model. *Nucl. Phys. B*, 1995, 436:3-29.
- [149] Carena M, Espinosa J, Quiros M, et al. Analytical expressions for radiatively corrected Higgs masses and couplings in the MSSM. *Phys. Lett. B*, 1995, 355:209-221.
- [150] Haber H E, Hempfling R, Hoang A H. Approximating the radiatively corrected Higgs mass in the minimal supersymmetric model. *Z. Phys. C*, 1997, 75:539-554.
- [151] Bagnaschi E, Giudice G F, Slavich P, et al. Higgs Mass and Unnatural Supersymmetry. *JHEP*, 2014, 09:092.
- [152] Barger V D, Berger M, Ohmann P. Supersymmetric grand unified theories: Two loop evolution of gauge and Yukawa couplings. *Phys. Rev. D*, 1993, 47:1093-1113.
- [153] Tanabashi M, Hagiwara K, Hikasa K, et al. Review of particle physics. *Phys. Rev. D*, 2018, 98: 030001.
- [154] First combination of Tevatron and LHC measurements of the top-quark mass. ATLAS-CONF-2014-008, CDF-NOTE-11071, CMS-PAS-TOP-13-014, D0-NOTE-6416, FERMILAB-TM-2582-E, 2014.

-
- [155] Ellis J R, Ridolfi G, Zwirner F. On radiative corrections to supersymmetric Higgs boson masses and their implications for LEP searches. *Phys. Lett. B*, 1991, 262:477-484.
- [156] Allanach B, Djouadi A, Kneur J, et al. Precise determination of the neutral Higgs boson masses in the MSSM. *JHEP*, 2004, 09:044.
- [157] Ellis S A R, Zheng B. Reaching for squarks and gauginos at a 100 TeV p-p collider. *Phys. Rev. D*, 2015, 92(7):075034.
- [158] Degrandi G, Di Vita S, Elias-Miro J, et al. Higgs mass and vacuum stability in the Standard Model at NNLO. *JHEP*, 2012, 08:098.
- [159] Wang F, Wang W, Yang J M. Dark matter constraints on gaugino/Higgsino masses in split supersymmetry and their implications at colliders. *Eur. Phys. J. C*, 2006, 46:521-526.
- [160] Masiero A, Profumo S, Ullio P. Neutralino dark matter detection in split supersymmetry scenarios. *Nucl. Phys. B*, 2005, 712:86-114.
- [161] Pierce A. Dark matter in the finely tuned minimal supersymmetric standard model. *Phys. Rev. D*, 2004, 70:075006.
- [162] Hewett J L, Lillie B, Masip M, et al. Signatures of long-lived gluinos in split supersymmetry. *JHEP*, 2004, 09:070.
- [163] Kilian W, Plehn T, Richardson P, et al. Split supersymmetry at colliders. *Eur. Phys. J. C*, 2005, 39:229-243.
- [164] Gambino P, Giudice G, Slavich P. Gluino decays in split supersymmetry. *Nucl. Phys. B*, 2005, 726:35-52.
- [165] Aaboud M, et al. Search for squarks and gluinos in final states with jets and missing transverse momentum using 36 fb^{-1} of $\sqrt{s} = 13 \text{ TeV}$ pp collision data with the ATLAS detector. *Phys. Rev. D*, 2018, 97(11):112001.
- [166] Sirunyan A M, et al. Search for supersymmetry in multijet events with missing transverse momentum in proton-proton collisions at 13 TeV. *Phys. Rev. D*, 2017, 96(3):032003.
- [167] Sirunyan A M, et al. Search for new phenomena with the M_{T2} variable in the all-hadronic final state produced in proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$. *Eur. Phys. J. C*, 2017, 77(10):710.
- [168] Aaboud M, et al. Search for long-lived, massive particles in events with displaced vertices and missing transverse momentum in $\sqrt{s} = 13 \text{ TeV}$ pp collisions with the ATLAS detector. *Phys. Rev. D*, 2018, 97(5):052012.
- [169] Sirunyan A M, et al. Search for natural and split supersymmetry in proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$ in final states with jets and missing transverse momentum. *JHEP*, 2018, 05:025.
- [170] Perl M L, Kim P C, Halyo V, et al. The Search for stable, massive, elementary particles. *Int. J. Mod. Phys. A*, 2001, 16:2137-2164.
- [171] Aristizabal Sierra D, Staub F, Vicente A. Shedding light on the $b \rightarrow s$ anomalies with a dark sector. *Phys. Rev.*, 2015, D92(1):015001.
- [172] Altmannshofer W, Straub D M. New physics in $b \rightarrow s$ transitions after LHC run 1. *Eur. Phys. J.*, 2015, C75(8):382.
- [173] Altmannshofer W, Gori S, Pospelov M, et al. Neutrino Trident Production: A Powerful Probe of New Physics with Neutrino Beams. *Phys. Rev. Lett.*, 2014, 113:091801.

- [174] Cowan G A. Measurements of CP violation in B mixing through $B \rightarrow J/\psi X$ decays at LHCb. PoS, 2015, EPS-HEP2015:543.
- [175] Altmannshofer W, Gori S, Pospelov M, et al. Neutrino Trident Production: A Powerful Probe of New Physics with Neutrino Beams. Phys. Rev. Lett., 2014, 113:091801.
- [176] Aaboud M, et al. Search for new high-mass phenomena in the dilepton final state using 36 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector. JHEP, 2017, 10:182.
- [177] Sirunyan A M, et al. Search for high-mass resonances in dilepton final states in proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$. CMS-EXO-16-047, 2018.
- [178] Thamm A, Torre R, Wulzer A. Future tests of Higgs compositeness: direct vs indirect. JHEP, 2015, 07:100.
- [179] Allanach B C, Gripaios B, You T. The case for future hadron colliders from $B \rightarrow K^{(*)} \mu^+ \mu^-$ decays. JHEP, 2018, 03:021.
- [180] Kumar N, Martin S P. Vectorlike Leptons at the Large Hadron Collider. Phys. Rev., 2015, D92(11):115018.
- [181] Alwall J, Frederix R, Frixione S, et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. JHEP, 2014, 07:079.
- [182] Dupuis G. Collider Constraints and Prospects of a Scalar Singlet Extension to Higgs Portal Dark Matter. JHEP, 2016, 07:008.
- [183] Artoisenet P, Frederix R, Mattelaer O, et al. Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations. JHEP, 2013, 03:015.
- [184] Sjöstrand T, Ask S, Christiansen J R, et al. An Introduction to PYTHIA 8.2. Comput. Phys. Commun., 2015, 191:159-177.
- [185] de Favereau J, Delaere C, Demin P, et al. DELPHES 3, A modular framework for fast simulation of a generic collider experiment. JHEP, 2014, 02:057.
- [186] ATLAS. Optimisation of the ATLAS b -tagging performance for the 2016 LHC Run. ATL-PHYS-PUB-2016-012, 2016.
- [187] Cacciari M, Salam G P, Soyez G. FastJet User Manual. Eur. Phys. J., 2012, C72:1896.
- [188] Czakon M, Fiedler P, Mitov A. Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$. Phys. Rev. Lett., 2013, 110:252004.
- [189] Kardos A, Trócsányi Z. Hadroproduction of t anti-t pair with a b anti-b pair using PowHel. J. Phys., 2014, G41:075005.
- [190] Kardos A, Trocsanyi Z, Papadopoulos C. Top quark pair production in association with a Z-boson at NLO accuracy. Phys. Rev., 2012, D85:054015.
- [191] Kulesza A, Motyka L, Stebel T, et al. Associated $t\bar{t}H$ production at the LHC: Theoretical predictions at NLO+NNLL accuracy. Phys. Rev. D, 2018, 97(11):114007.
- [192] Lester C G, Summers D J. Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders. Phys. Lett., 1999, B463:99-103.
- [193] Cowan G, Cranmer K, Gross E, et al. Asymptotic formulae for likelihood-based tests of new physics. Eur. Phys. J., 2011, C71:1554.

-
- [194] Allanach B, Gripaos B. Hide and Seek With Natural Supersymmetry at the LHC. JHEP, 2012, 05:062.
- [195] Durieux G, Smith C. The same-sign top signature of R-parity violation. JHEP, 2013, 10:068.
- [196] Bhattacharjee B, Evans J L, Ibe M, et al. Natural supersymmetry's last hope: R-parity violation via UDD operators. Phys. Rev. D, 2013, 87(11):115002.
- [197] Diglio S, Feligioni L, Moutaka G. Stashing the stops in multijet events at the LHC. Phys. Rev. D, 2017, 96(5):055032.
- [198] Buckley M R, Feld D, Macaluso S, et al. Cornering Natural SUSY at LHC Run II and Beyond. JHEP, 2017, 08:115.
- [199] Li J, Li T, Zhang W. Least constrained supersymmetry with R -parity violation. Phys. Rev. D, 2019, 99(3):036011.
- [200] Butterworth J M, Ellis J R, Raklev A R, et al. Discovering baryon-number violating neutralino decays at the LHC. Phys. Rev. Lett., 2009, 103:241803.
- [201] Bai Y, Katz A, Tweedie B. Pulling Out All the Stops: Searching for RPV SUSY with Stop-Jets. JHEP, 2014, 01:040.
- [202] Bhattacharjee B, Chakraborty A. Study of the baryonic R-parity violating MSSM using the jet substructure technique at the 14 TeV LHC. Phys. Rev. D, 2014, 89(11):115016.
- [203] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016:770-778.
- [204] Komiske P T, Metodiev E M, Schwartz M D. Deep learning in color: towards automated quark/gluon jet discrimination. JHEP, 2017, 01:110.
- [205] Baldi P, Cranmer K, Faucett T, et al. Parameterized neural networks for high-energy physics. Eur. Phys. J. C, 2016, 76(5):235.
- [206] Bhimji W, Farrell S A, Kurth T, et al. Deep Neural Networks for Physics Analysis on low-level whole-detector data at the LHC. J. Phys. Conf. Ser., 2018, 1085(4):042034.
- [207] ATLAS. Search for massive supersymmetric particles in multi-jet final states produced in pp collisions at $\sqrt{s} \approx 13$ TeV using the ATLAS detector at the LHC. ATLAS-CONF-2016-057, 2016.
- [208] Beenakker W, Hopker R, Spira M. PROSPINO: A Program for the production of supersymmetric particles in next-to-leading order QCD. 1996.
- [209] He K, Gkioxari G, Dollár P, et al. Mask r-cnn. Proceedings of the IEEE international conference on computer vision, 2017:2961-2969.
- [210] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation. 2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014:580-587.
- [211] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015:3431-3440.
- [212] Lin T, Dollár P, Girshick R, et al. Feature pyramid networks for object detection. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017:936-944.
- [213] Girshick R. Fast r-cnn. 2015 IEEE International Conference on Computer Vision (ICCV), 2015: 1440-1448.

- [214] Ren S, He K, Girshick R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2017, 39(6):1137-1149.
- [215] Aad G, et al. Combination of searches for Higgs boson pairs in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Lett. B*, 2020, 800:135103.

致 谢

感谢导师陈少敏教授对我博士期间研究工作和论文写作的指导。感谢李田军教授在我学习和研究过程中答疑解惑和指引研究方向。感谢何红建教授将我引入粒子物理研究的大门。他们的言传身教将使我终生受益。感谢师兄李金勉副教授带领我进入对撞机物理的领域并提供计算资源，感谢合作者郭俊博士，牛家树博士，Shabbar 博士，Waqas 博士和张文星师妹在研究中对我的启发。感谢肖瑞卿师兄，王雨晨师弟，郭磊同学，郭子溢同学对我学习和生活上的帮助。

声 明

本人郑重声明：所呈交的学位论文，是本人在导师指导下，独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本学位论文的研究成果不包含任何他人享有著作权的内容。对本论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明。

签 名： _____ 日 期： _____

个人简历、在学期间发表的学术论文与研究成果

1991年3月26日出生于吉林省吉林市。

2010年9月考入清华大学物理系物理学专业，2014年7月本科毕业并获得理学学士学位。

2014年9月免试进入清华大学工程物理系攻读物理学博士学位至今。

发表的学术论文

- [1] Jun Guo, Jinmian Li, Tianjun Li, Fangzhou Xu, Wenxing Zhang. Deep learning for R-parity violating supersymmetry searches at the LHC. *Phys. Rev. D* 98, 076017 (2018).
- [2] Fang-Zhou Xu, Wenxing Zhang, Jinmian Li, Tianjun Li. Search for the vectorlike leptons in the $U(1)_X$ model inspired by the B -meson decay anomalies. *Phys. Rev. D* 98, 115033 (2018).
- [3] Jia-Shu Niu, Tianjun Li, Fang-Zhou Xu. A simple and natural interpretations of the DAMPE cosmic-ray electron/positron spectrum within two sigma deviations. *Eur. Phys. J. C* 79, 125 (2019).
- [4] Waqas Ahmed, Adeel Mansha, Tianjun, Shabbar Raza, Joydeep Roy, Fangzhou Xu. The Upper Bounds on Supersymmetry Breaking Scales in the Split Supersymmetry from the Higgs Mass and Electroweak Vacuum Stability. *Prog. Theor. Exp. Phys.* 2020, 043B08 (2020)