arXiv:2611.30394v1 [physics.ins-det] 23 Nov 2026

拟提交至 JINST

用于高亮度大型强子对撞机长寿命粒子探测的基于 FPGA 的先进 μ 子触发算法

D. Estrada，[1] J. Prado、P. Leguina、A. Cardini、E. Aller 和 S. Folgueras 代表 CMS 合作组[2]。

奥维耶多大学，*C. San Francisco 3*，西班牙奥维耶多；
阿斯图里亚斯空间科学与技术研究所（*ICTEA*），*C. de la Independencia 13*，西班牙奥维耶多

电子邮箱：destrada@cern.ch

摘要：CMS 一级触发器必须在严格的延迟和硬件资源限制下实时重建 μ 子。对于发生辐射能量损失的高能 μ 子，传统算法的效率会降低；对于位移轨迹，其表现也可能不理想，从而限制了对长寿命粒子（LLP）等奇异信号的灵敏度。第二阶段升级引入了先进的现场可编程门阵列，使得能够采用更复杂的重建算法。本文提出两项相辅相成的进展，以应对这些挑战。在桶部区域，一种专用算法在第一个处理阶段通过漂移管击中模式探测簇射。在随后的滤波阶段，将簇射原语与传统触发原语相关联；目前的目标是辐射簇射高动量 μ 子，同时也着眼于更广泛的 LLP 信号。模拟研究表明，簇射分类准确率约为 80%，辐射 μ 子的估计匹配效率超过 80%。固件原型表明，簇射探测延迟约为 20 纳秒，初步匹配逻辑的延迟低于一个束团交叉周期。与此同时，本文还研究了一种面向硬件的优化方法，以支持未来在桶部—端盖重叠区开展基于图神经网络的位移 μ 子重建。使用基准数据集上的 GraphSAGE 代理模型，通过结合整数定量化、二的幂次缩放和数据驱动的位宽优化，工作流实现了 (75.0 ± 1.1)% 的准确率，以及 19 个时钟周期（在 360 兆赫下为 52.8 纳秒）的确定性延迟，数字信号处理器利用率仅为 20%。这些结果为利用 CMS μ 子数据开展实现研究奠定了基础。

关键词：触发算法、模式识别、粒子鉴别方法。

[1]通讯作者。[2]版权所有 © 2026 欧洲核子研究中心，代表 CMS 合作组。允许复制本文或其部分内容但须遵循知识共享署名 4.0 国际许可协议的规定。

目录

1 引言 1 2 CMS 桶部的缪子簇射标记 2 3 面向硬件的位移 μ 子重建图神经网络优化 3 4 结论与展望 5

1 引言

在大型强子对撞机（LHC）上识别长寿命粒子（LLP），可拓展紧凑型缪子螺线管（CMS）实验 [1, 2] 对标准模型（BSM）之外新物理的探索范围。根据衰变位置和末态的不同，LLP可能产生位移缪子、位移强子喷注，或在缪子系统中形成类似簇射的活动。这些特征不同于 prompt 粒子的特征，可能无法被针对标准轨迹优化的传统触发算法 [3–5]捕获。

在高亮度大型强子对撞机（HL-LHC）上，CMS 一级触发器（L1T）必须以 40 兆赫的束团交叉（BX）频率处理探测器数据，并在 12。5 μ秒的延迟预算内，以最高 750 千赫的速率筛选事例。第二阶段升级引入了处理能力更强的现场可编程门阵列（FPGA），使得在这些限制下能够运行更复杂的重建算法。对于本文研究的算法，全局缪子触发器（GMT）上游的处理窗口约为 4 μ秒。因此，低延迟和可预测的资源使用量是关键设计要求 [6, 7]。

本研究针对桶部μ子系统及桶部—端盖重叠区域中的两项互补挑战。在桶部，簇射可能源于长寿命粒子衰变等非常规特征，也可能由高动量 μ 子的辐射能量损失引起。两者都会使缪子重建更加复杂，进而影响对超出标准模型物理的灵敏度以及对 TeV 级 μ 子的探测效率 [8, 9]。在重叠区域，缪子径迹寻找采用重叠区μ子径迹寻找器（OMTF）。该算法将探测器信息与预定义的参考模式进行匹配；然而，这种方法限制了其对具有非提示径迹的位移缪子的探测效率。 [10–13]。

本文介绍了两项针对上述问题的进展：一种桶部簇射标记算法，将击中多重性检测与传统 μ 子触发原语（TP）进行几何匹配；以及一种面向硬件的优化方法，用于 OMTF 中位移 μ 子重建的图神经网络（GNN）推理。簇射标记研究包含模拟结果和 FPGA 实现；GNN 研究则使用基准图数据集和 GraphSAGE 代理模型评估优化工作流。

2 CMS桶部系统中的μ子簇射标记

基于簇射的触发已用于在缪子系统中筛选非常规特征。第三轮运行期间，CMS高多重性触发器（HMT）利用端盖阴极条室（CSC）中的击中多重性，筛选与长寿命粒子衰变相符的事件 [14]。

同样，在桶部，簇射可能源自强子量能器之外的长寿命粒子衰变，包括由中性长寿命标量粒子、胶微子或标量夸克产生的位移强子喷注，以及重中性轻子（HNL）衰变 [4, 5]。高能μ子也可能产生簇射，包括宇宙线μ子以及来自超出标准模型矢量玻色子衰变的μ子，例如 $Z' \to \mu\mu$和 $W' \to \mu\nu$；这些μ子在钢制磁通回磁轭或其他探测器材料中发生韧致辐射等辐射相互作用时会产生簇射（见图1 左）。这些簇射会增加命中多重性，并可能导致触发原语生成出现歧义，从而降低常规缪子重建的性能；已有研究表明，这会降低一级触发效率。该效应随μ子动量增加，在中央桶部（赝快度为（η）且位于$[-9, 9]$ [8]的CMS区域）中，动量达到10% ，在1TeV 时约为。因此，专用簇射标记有助于识别发生簇射的μ子，并提高对非常规特征和高能μ子的探测灵敏度。

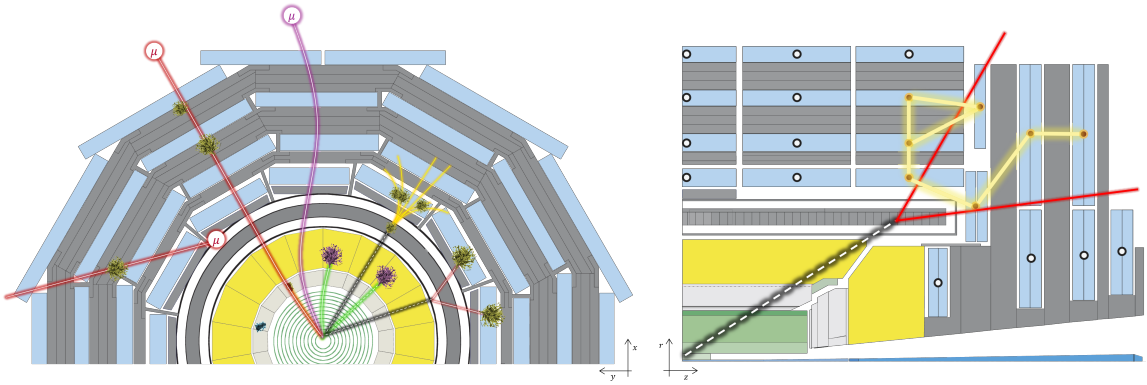


图1。左：CMS缪子系统中类似簇射的特征示例，包括高能μ子的辐射能量损失（红色）以及可能的长寿命粒子衰变所伴随的簇射（黄色）。右：用于重叠区域基于图神经网络的缪子重建的探测器测量概念图表示；节点表示探测器短段，边则编码兼容短段之间的几何关系。

第二阶段桶部缪子触发器将局部命中重建与多站关联分开处理。其架构由30块定制的ATCA BMTL1板卡组成，用于处理漂移管（DT）和阻性板室（RPC）数据，并通过运行于AMD/Xilinx Virtex UltraScale+ VU13P现场可编程门阵列上的解析方法（AM）生成局部TP [15]。输出发送至12块桶部滤波（BF）板卡，每块接收来自10块BMTL1板卡的数据，以整合不同扇区和轮的信息 [6, 7]。因此，BF阶段可以在将候选项转发至GMT之前，关联常规TP与簇射活动。

所提出的簇射标记方法由两个部分组成。首先，在 BMTL1 阶段与 AM 并行，直接从原始漂移管命中中检测簇射活动。其次，在 BF 阶段，将生成的簇射原语（SP）与常规 TP 相关联，以识别

感兴趣的特征。目前的实现侧重于辐射高动量缪子，而更广泛的目标包括 LLP 衰变。

鉴别算法会监测400 纳秒窗口（16BX，即 DT 电子漂移时间的最大值）。当累计击中数超过可配置阈值时，算法会生成一个 SP，其中包含首次击中 BX、最小和最大导线标识符，以及导线击中计数剖面向量。这种紧凑的表示方式保留了簇射的关键信息，包括时间和空间信息，而无需将完整的击中列表传送至下游 [9, 16]。

在 BF 中，对于辐射缪子标记，算法会针对每个 SP 评估来自相邻扇区的 TP，并应用时间切选以排除不兼容的候选对象。在几何上，簇射被建模为连接其最远导线位置的线段，而每条 TP 轨迹则被视为一条直线。如果 TP 轨迹在时间窗口内与簇射线段相交，则会将其标记为簇射相关，以供下游重建使用 [9]。这种方法将几何测试简化为适用于 FPGA 实现的运算。

模拟研究显示，簇射识别的总体准确率约为 80%，并优先确保尽可能减少漏检的簇射事件；估算的局部 TP 率约为 AM 算法的十分之一。几何匹配算法在标记簇射高动量缪子方面的估算效率超过 80%。

初始的簇射鉴别 FPGA实现约使用 2.5 k 个查找表（LUT）和 2.5 k 个触发器（FF），延迟为 20 纳秒。对于 BF 匹配逻辑，固件原型每个时钟周期以 480 兆赫的频率处理最多 48 对 TP–SP，实现亚BX延迟，约使用 4.3 k 个 LUT 和 5.6 k 个 FF。目前版本的该设计覆盖单个轮中的 3 个扇区；仍需将其扩展至完整的 BF覆盖范围。

一项初步的并行研发工作将 SP密度剖面作为深度神经网络（DNN）的输入，以区分类强子簇射与背景活动。该方法有望支持桶部位移强子型长寿命粒子衰变的筛选。基线簇射生成设计已在停机前切片测试期间集成至 CMS，目前正在利用宇宙线数据评估其性能。

3 面向硬件的优化：用于位移 μ 子重建的图神经网络

在桶部–端盖过渡区（$0.83 < |\eta| < 1.24$），OMTF 将 DT、CSC 和 RPC 短段结合起来，并使用朴素贝叶斯分类器将其与预定义的参考模式进行匹配 [10,11]。虽然这种方法对即时缪子有效，但不太适用于长寿命粒子衰变产生的位移轨迹。图神经网络可以利用探测器测量之间的空间关系，提供另一种方案 [12, 13]。

在图表示中（见图 1 右），节点表示探测器短段，包括层信息及其 η 和 ϕ 坐标。边连接几何上兼容的短段，并编码相对角度差（$\Delta\eta$ 和 $\Delta\phi$） [12]。重建任务是识别兼容的短段，排除由堆积产生的组合，并推断缪子轨迹参数。

GraphSAGE 是一种归纳式架构，通过聚合来自邻近节点的信息来更新节点表示 [17]。这种消息传递结构为整合各站点的局部测量值提供了基础。然而，其不规则数据依赖以及

算术强度给 FPGA 实现带来了挑战，尤其是在一级触发器严格的延迟和资源限制下 [13, 18]。

本研究将硬件可行性评估与物理学专用模型的开发分开进行。我们在 Cora 引文数据集上评估了一个两层 GraphSAGE 固件代理模型 [19], , 该数据集中的节点代表出版物，边代表引文。为匹配目标应用的规模，提取包含八个节点、扩展至两跳的固定大小子图，并将原始的 1433 维节点特征投影到片外的 16 个特征（这部分不计入推理延迟测量）。该代理模型为评估消息传递的时序和资源使用情况提供了受控环境，但不评估 CMS 的物理学性能 [13]。

传统的 GraphSAGE 会将聚合后的邻居信息与根节点的自身项结合起来 [17], , 而固件代理模型省略了这一部分。这种简化移除一个权重矩阵，减少乘加运算的数量，消除中间拼接缓冲区，并简化固定大小循环，以便在高级综合（HLS）中对完整数组进行展开 [13]。

从 PyTorch 浮点模型出发，所采用的优化流程包括固定尺寸图提取、纯整数训练后量化（PTQ）、二次幂尺度近似、数据驱动的位宽优化，以及与 Python 整数参考模型 [13] 进行逐位精确验证。以下讨论聚焦于二次幂缩放和位宽优化概念，然后再介绍结果。

在标准 INT8 量化中，神经元的累加结果（acc）会在进入下一层之前重新缩放，以防止数据溢出 [20]。从数学上看，这种标准的重新缩放表示为：

$$q_{\text{out}} = \text{clamp}\left((acc \cdot \alpha) \gg M\right),$$

其中，α 是定点尺度乘数，M 是二进制小数点位移。在现场可编程门阵列上，通用定点乘法需要专用数字信号处理器模块，并会增加资源占用。INT8-PO2 优化将每个尺度因子近似为最接近的二次幂（$\alpha \approx 2^p$） [13], , 从而将重新缩放简化为简单的位移：

$$q_{\text{out}} = \text{clamp}(acc \gg k), \qquad k = M - p.$$

这消除了通用重新缩放乘法，使矩阵乘法成为数字信号处理器资源的主要消耗项。

数据驱动的位宽优化可进一步降低资源占用。该方法不采用统一精度或最坏情况估算，而是分析中间信号分布，为每个信号分配能够防止溢出的最小位宽，并留出额外的安全余量。由此得到的位宽用于生成各层专用的 HLS 参数 [13]。

表 1 总结了 FPGA 综合结果。将 8位整数-二次幂量化与数据驱动的位宽优化相结合，可将推理延迟缩短至 19 个时钟周期（52.8 纳秒，频率为 360 兆赫），所用资源分别为 20% 的数字信号处理器、6% 的 FF，以及 27% 的 LUT。优化后的配置准确率达到 (75.0 ±1.1)%，而浮点参考模型的准确率为 (78.0 ± 0.8)%。Python 整数参考模型与 HLS C 仿真也显示，两者在位级上完全一致 [13]。

配置	准确率 (%)	延迟（时钟周期）	延迟（纳秒）	数字信号处理器 (%)	触发器 (%)	查找表（%）
单精度浮点数	78.0 ± 0.8	603	1675.0	283	128	123
固定点后训练量化	75.7 ± 1.0	77	213.9	56	7	6
8位整数-二次幂	75.0 ± 1.2	28	77.8	43	7	10
QAT	76.8 ± 0.9	55	152.8	55	16	16
INT8-二次幂优化	75.0 ± 1.1	19	52.8	20	6	27

表1。准确率（均值 ± 五个随机种子的标准差），延迟以时钟周期数表示，或假设时钟频率为360兆赫时以纳秒表示，以及XCVU13P FPGA资源利用率（> 100%超过单个器件的容量），频率为 360兆赫。改编自 [13]。

4 结论与展望

在 GMT 上游的延迟和 FPGA 资源限制内，研究了两种互补的方法，用于识别 CMS Phase-2 一级触发器中的非常规μ子特征。

在桶部，固件簇射识别算法能够以较低的FPGA资源使用量和亚BX延迟生成SP。模拟研究表明，该算法在簇射探测和标记辐射簇射高动量μ子方面具有良好的潜力。基准设计已在停机前切片测试期间集成到CMS中，目前正在使用宇宙线数据进行评估。

针对OMTF，已开发出一套面向FPGA的优化工作流，用于指导GNN在硬件中的实现。该工作流结合了整数定量化、二次幂缩放和数据驱动的位宽微调；在GraphSAGE代理模型上，它实现了19个时钟周期的确定性延迟和较低的DSP使用量，同时保持了接近浮点参考模型的准确率。接下来的必要步骤包括使用物理驱动的位移μ子数据集进行验证，以及在目标FPGA上实现。

桶部未来的开发工作包括将匹配逻辑扩展至完整的探测器覆盖范围，同时纳入纵向 θ 原语，并利用SP识别长寿命粒子位移顶点的特征。未来的GNN研究将把该工作流应用于真实CMS μ子数据，并探索用于径迹关联、电荷鉴别和横向动量重建的架构。

致谢

本研究得到了 ERC 项目 INTREPID（10.3030/101115353）的资助。该项目由欧盟资助。然而，文中表达的观点和意见仅代表作者本人，并不一定反映欧盟或欧洲研究委员会执行局的观点和意见。欧盟和资助机构均不对此承担责任。

参考文献

[1] CMS 合作组，欧洲核子研究组织大型强子对撞机上的 *CMS* 实验，《仪器学杂志》（*JINST*）**3**（2008）S08004。[2] CMS 合作组，面向欧洲核子研究组织大型强子对撞机第三轮运行的 *CMS* 探测器研制，《仪器学杂志》（*JINST*）**19**（2024）P05064。

[3] *S. Folgueras*，*INTREPID*，第*108–111*页，见 *https://www.europeandissemination.eu/article/intrepid/22858*（*2025*）。[4] *CMS* 合作组，《在$\sqrt{s}$ = 13 *TeV*质子-质子碰撞中利用位移喷注搜寻长寿命粒子》，*Phys. Rev. D 104*（*2021*）*012015*。[5] *G. Cottin*、*J.C. Helo*、*M. Hirsch* 等，《*CMS*实验中具有位移簇射特征的长寿命重中性轻子》，《高能物理学杂志》（*JHEP*）*02*（*2023*）*011*。[6] *CMS* 合作组，《*CMS*一级触发器第二阶段升级技术设计报告》，技术报告 *CERN-LHCC-2020-004*，*CMS-TDR-021*，欧洲核子研究中心，瑞士日内瓦（*2020*）。[7] *A. Zabi* 等，《高亮度大型强子对撞机上 *CMS* 一级触发器的系统设计与原型开发》，载于《粒子物理技术与仪器国际会议论文集》— *J. Phys.: Conf. Ser.*（*TIPP 2021*），第*2374*卷，第*012090*页，*2022*，*DOI*。[8] *CMS* 合作组，《$\sqrt{s}$ = 13 *TeV*质子-质子碰撞中高动量 μ 子的重建与鉴别性能》，《仪器学杂志》*15*（*2020*）*P02027*。[9] *D. Estrada* 等，《面向高亮度大型强子对撞机升级的 *CMS* 实验桶部μ子触发器中的 μ子簇射触发》，载于《欧洲物理学会高能物理会议论文集》— *PoS(EPS-HEP2025)*，第*485*卷，第*577*页，*2026*，*DOI*。[10] *K. Bunkowski* 等，《*CMS* 一级触发器重叠区μ子径迹寻找器触发算法》，载于第*14*届比萨先进探测器会议（《前沿物理的前沿探测器》）论文集—《核仪器与方法物理研究*A*辑》，第*936*卷，第*368–369*页，*2019*，*DOI*。[11] *W.M. Zabolotny* 等，《*CMS*探测器桶部-端盖重叠区μ子触发器与数据传输系统的算法及实现》，载于《粒子物理电子学专题研讨会论文集》— *JINST*（*TWEPP 2015*），第*11*卷，第*C03004*页，*2016*，*DOI*。[12] *P. Leguina* 等，《在 *FPGA* 上使用人工智能实现高亮度大型强子对撞机的 *CMS* 重叠区μ子径迹寻找器》，载于第二届欧洲基础物理人工智能会议， 9, 2025 [*arXiv:2509.*23347].[13] *A. Cardini* 等，《用于实时*FPGA*径迹寻找的图神经网络整数定量化》，载于第*25*届 *IEEE* 实时会议（*RT2026*）论文集， 2026 [*arXiv:2609.*28144].[14] *M. Abbott* 等，《*CMS*探测器中用于长寿命粒子的高多重性触发器》，《仪器学杂志》*21*（*2026*）*P07040*。[15] *G. Abbiendi* 等，《大型强子对撞机漂移管探测器中用于生成触发原语的解析方法算法》，《核仪器与方法物理研究*A*辑》（*NIMA*）*1049*（*2023*）*168103*。[16] *J. Prado* 等，《*CMS*实验桶部μ子系统中的μ子簇射标记》，载于第*26*届国际辐射成像探测器研讨会论文集— *JINST*（*IWORID2025*），第*21*卷，第*C01050*页，*2026*，*DOI*。[17] *W. Hamilton* 等，《大型图上的归纳式表示学习》，载于《神经信息处理系统进展》— *NeurIPS*，第*30*卷，第*1024–1034*页，*2017*，*https://proceedings.neurips.cc/paper_files/paper/2017/file/5dd9db5e033da9c6fb5ba83c7a7ebea9- Paper.pdf* [*arXiv:1706.*02216].[18] *Y. Iiyama*、*G. Cerminara*、*A. Gupta*、*J. Kieseler*、*V. Loncar*、*M. Pierini* 等，《用于高能物理实时粒子重建的 *FPGA* 距离加权图神经网络》，*Frontiers in Big Data* 第*3*卷—*2020*（*2021*）。

[19] P. Sen、G. Namata、M. Bilgic、L. Getoor、B. Gallagher 和 T. Eliassi-Rad，网络数据中的集体分类，《*AI Magazine*》**29**（2008）93。[20] B. Jacob、S. Kligys、B. Chen、M. Zhu、M. Tang、A. Howard 等，用于高效纯整数运算推理的神经网络量化与训练，载于*2018*年*IEEE/CVF*计算机视觉与模式识别会议（*CVPR*），（美国加利福尼亚州洛斯阿拉米托斯），第2704–2713页，IEEE计算机学会，2018年6月，DOI。