

2024年度考核报告

刘春秀

实验物理中心 软件组

2024年11月

岗位职责和主要负责工作

岗位职责:

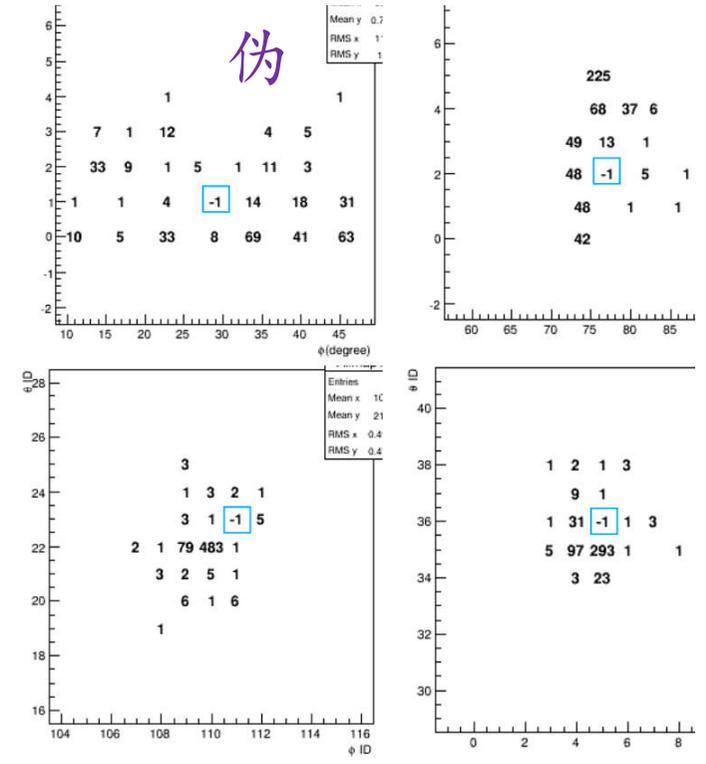
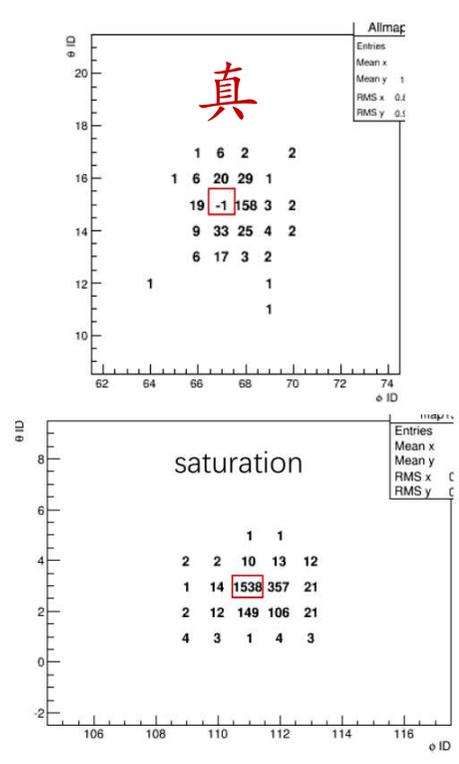
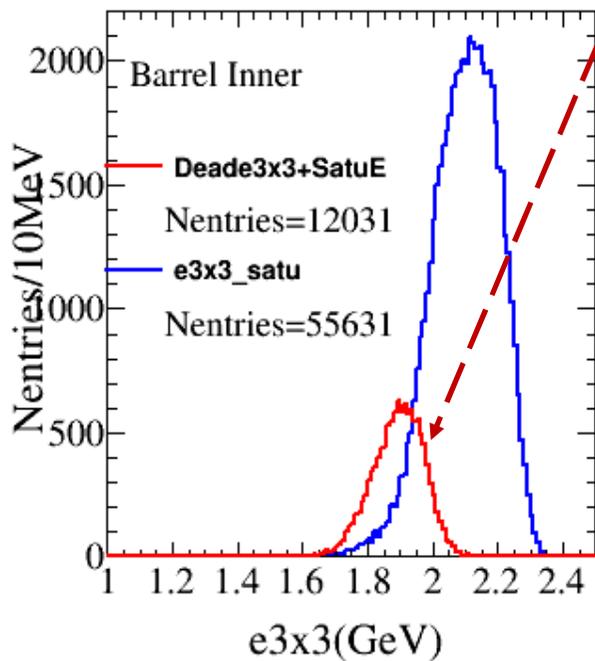
- EMC软件和数据处理

主要工作:

- EMC性能研究
 - 基于机器学习方法研究EMC特殊死道真伪鉴别
- 软件和数据处理
 - EMC特殊死道模拟的升级
 - 基于机器学习方法的EMC电子学反常探测
 - EMC time cut
 - EMC簇射能量修正: round16 ψ'' @boss711, round17 ψ'' @boss712
 - EMC数据质量检查
 - EMC离线刻度
 - TOF能量重建的能量刻度
 - CGEM对EMC的影响

基于机器学习方法研究EMC特殊死道真伪鉴别(I)

- 特殊死道现象：EMC电子学过饱和 → 很大信号 → 芯片工作反常 → 信号丢失 → 特殊死道
- 问题：EMC特殊死道的存在，影响某些物理分析，例如Dimuon样本挑选（不用muon counter）等
- 关键：准确判定特殊死道，可用饱和能量修正
- 意义：不仅能消除特殊死道对物理分析的影响，还能进一步完善特殊死道的模拟参数
- 方法：基于机器学习的分类方法研究特殊死道真伪鉴别



基于机器学习方法研究EMC特殊死道真伪鉴别(II)

- 创建训练、测试样本集
 - 用EmcFindSpecialDead算法找到‘类’特殊死道的showers
 - 用cut方法，选出‘真’特殊死道样本标记为1，其它标记为0
 - 数据集特征量：10个
 - ntheta, nphi, Esatu, DeadE3x3, DeadE5x5, No1E, No2E, No3E, No4E, No5E
 - Barrel 和Endcap 分别进行模型训练
 - Barrel: ~69% real samples --- '1'
 - Endcap: ~10% real samples --- '1'
- Scikit-Learn二分类器
 - 随机梯度下降（SGD）分类器
 - 随机森林（Random Forest）分类器
 - 支持向量机（SVM）分类器
- 混淆矩阵（Confusion Matrix, CM）评估训练模型
 - 对于不平衡数据集，评估分类器更好的办法是查看混淆矩阵

误差分析:

基于机器学习方法研究EMC特殊死道真伪鉴别(III)

Barrel

- 假阳性率、假阴性率都很低
- Barrel samples are very pure.

Endcap

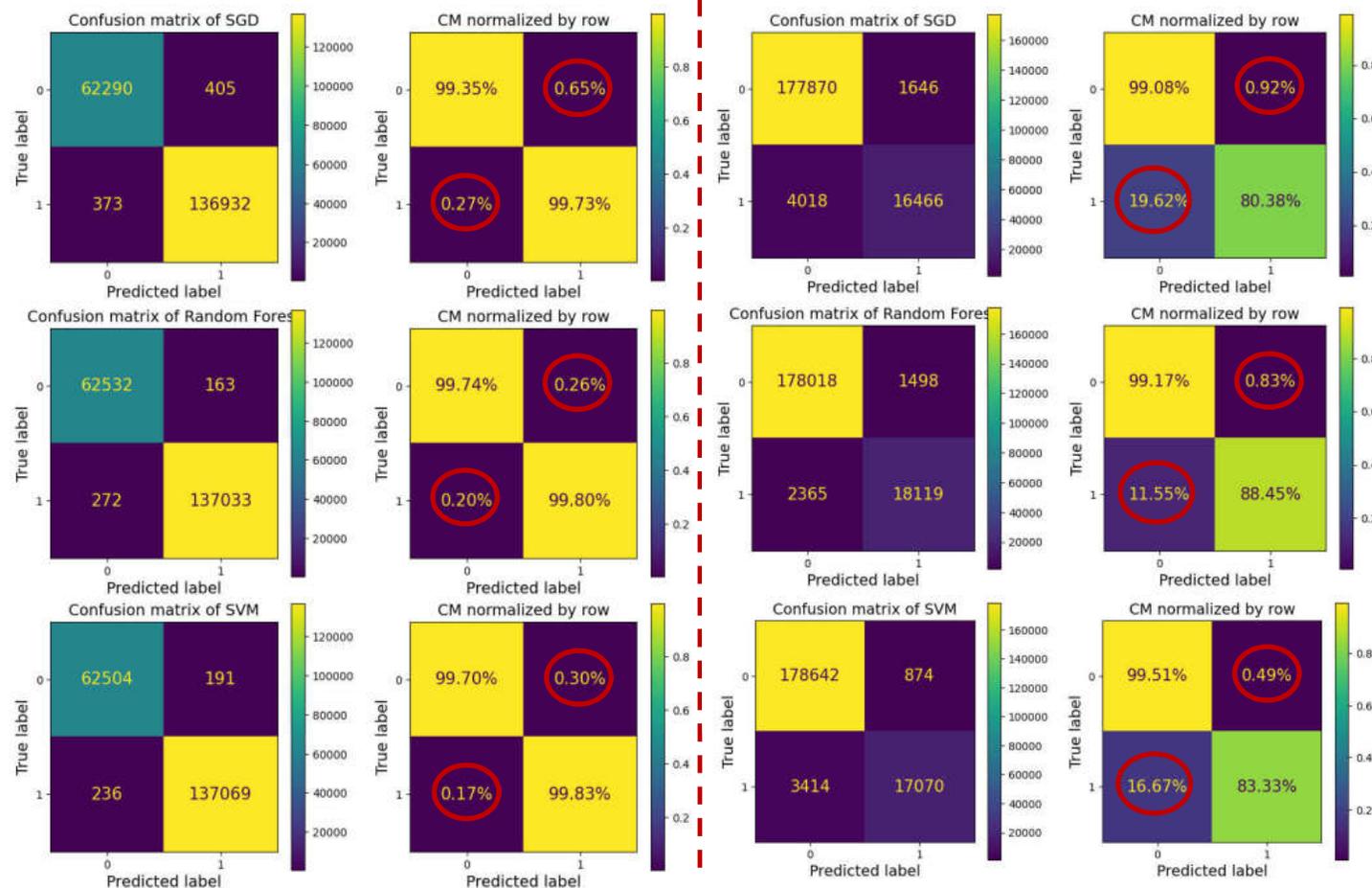
- 假阳性率<1%，假阴性率~10%-20%
- 说明Endcap样本中，真特殊死道样本的纯度不够好

5x5digimap

| | | | | |
|----|----|-----|----|---|
| 1 | 6 | 2 | | 2 |
| 1 | 6 | 20 | 29 | 1 |
| 19 | -1 | 158 | 3 | 2 |
| 9 | 33 | 25 | 4 | 2 |
| 6 | 17 | 3 | 2 | |

下一步计划:

- 优化特征量 → 更好的训练模型。
- 创建新的数据集，采用matrix形式的特征量，每个元素包含位置和能量信息。
- 研究假阳性和假阴性事例的特征 → 创建纯度更高的数据集，完善训练模型。
- Save训练好的模型 → 操作系统Alma Linux 9的集群上加载模型 → 特殊死道预测 → 特殊死道能量修正。

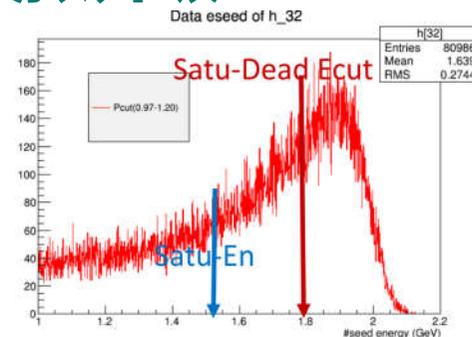


EMC软件和数据处理：EMC特殊死道模拟升级

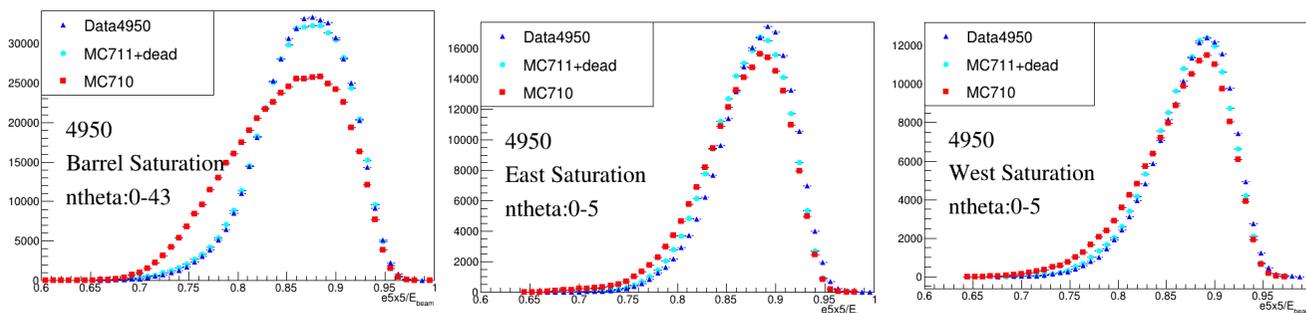
特殊死道模拟的假设：



- Energy measurement of i^{th} crystal: E_i is the crystal energy deposited
 1. The energy deposited: $E_i \leq \text{Satu_En}$
 2. Saturation energy : $\text{Satu_En} < E_i \leq \text{Satu_DeadEcut}$
 3. Set into Dead channel: $E_i > \text{Satu_DeadEcut}$
- 特殊死道能量阈值 Satu_DeadEcut : 第1套来自4950data@2021
- 不同批次数据饱和能量有差异，那么特殊死道能量阈值也可能不同。
- 再假设：对于同一电子学通道，特殊死道能量阈值与饱和能量有特定比值 $C_i = \frac{\text{Satu_DeadEcut}}{\text{Satu_En}}$ ，
 对不同数据特殊死道阈值可计算为 $C_i \times \text{Satu_En}$ （当前数据饱和能量）
- 软件升级 `EmcCalibConstSvc-00-01-19`，boss7.1.3发布
- 对于不同批次数据，这个假设还需进一步验证



Energy distribution of seed crystal without electr-saturation



To evaluate the consistency of the two histograms (h_1, h_2): correlation, the criterion is 1.

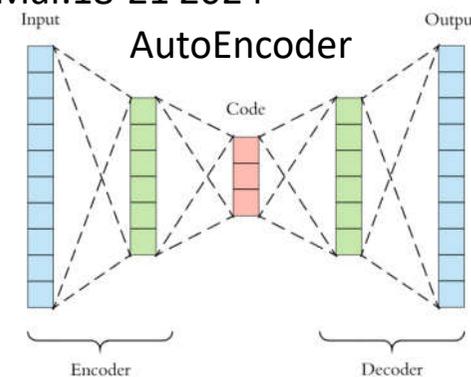
$$d(h_1, h_2) = \frac{\sum_i (h_1(i) - \bar{h}_1)(h_2(i) - \bar{h}_2)}{\sqrt{\sum_i (h_1(i) - \bar{h}_1)^2 \sum_i (h_2(i) - \bar{h}_2)^2}}$$

| | Barrel | East Endcap | West Endcap |
|----------------|----------|-------------|-------------|
| w/o Dead Sim. | 0.946958 | 0.981917 | 0.9888 |
| with Dead Sim. | 0.999568 | 0.990281 | 0.996387 |

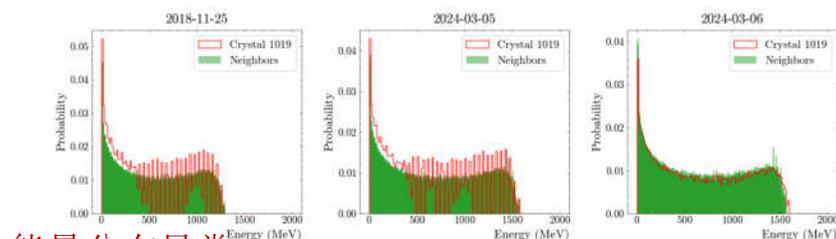
基于机器学习方法的EMC电子学反常探测

Anomaly Detection on EMC using Machine Learning @Phys.&Soft. Workshop Mar.18-21 2024

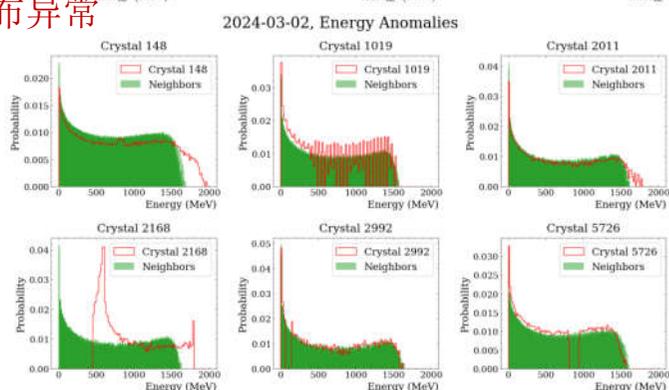
- EMC有6240晶体，传统方法很难实现对每一路电子学状态进行实时监测
- 自编码器能够自动学习数据特征，适合这种异常检测
- 利用EMC digi 能量、时间直方图信息，开发了一个有效的基于机器学习的EMC电子学monitor方法
- 每天例行检查，成功地完成了本轮的取数任务
- 发现了1019号晶体隐藏多年的EMC电子学信息反常，更换插件后恢复正常



Model Structure



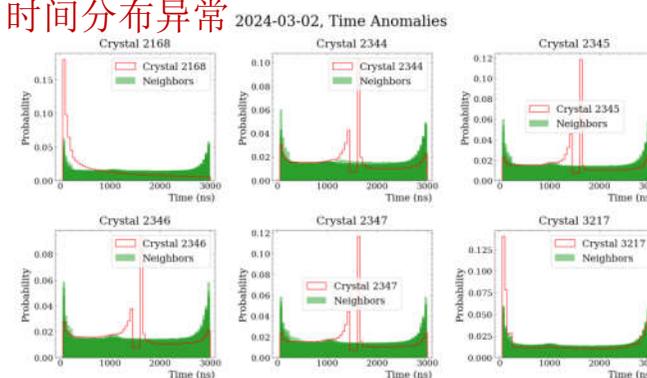
能量分布异常



"Neighbors": Normal crystals around the anomaly.



时间分布异常



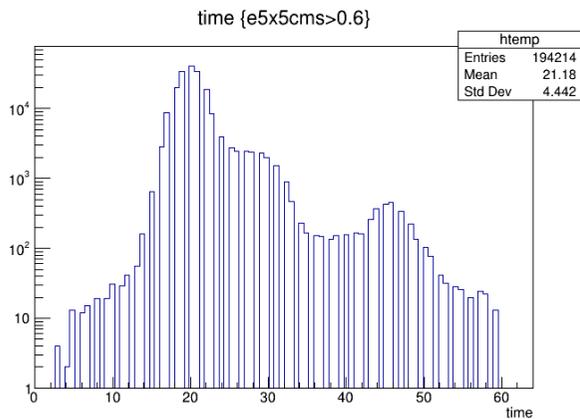
"Neighbors": Normal crystals around the anomaly.

EMC软件和数据处理: EMC time cut

- 目的: 降低beam本底的假光子对物理分析的影响
 - EMCRec中使用[0,35(1750ns)], EMC总时间窗 [0,60*50ns]
 - 物理分析使用[0,14*50ns], **DST数据中EMC time减掉t0**
- 问题: 在EMCRec时间窗外EmcTime>35存在少量物理事例, EMC没有重建。
- 影响: 例如invisible particle分析、Bhabha/dimu亮度测量 (使用MDCtrack与EMC匹配效率降低)
- 讨论: 是否放开EMCRec time cut?

run 35081 at 4.58 GeV
96%Bhabha+4%Dimu

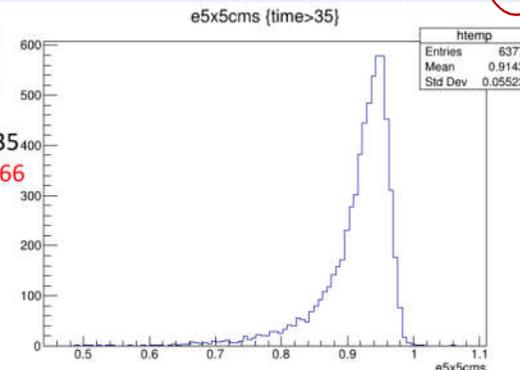
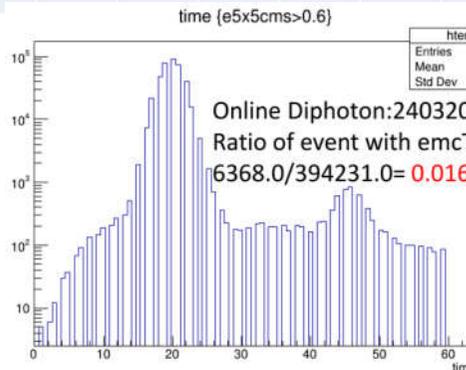
| | Without EMC | With EMC | ratio |
|------------|-------------|----------|--------|
| TimeMax=60 | 45394 | 45384 | 99.98% |
| TimeMax=35 | 45362 | 44896 | 98.97% |



240320/run81573 onlineBhabha
Ratio of event with emcTDC>35:
3666.0/194214.0= 0.018876085

从2009到2024, 每年任选一天, 统计了emcTime>35 online Bhabha 事例的占比
 $R = \text{onlineBhabha}(t>35) / \text{onlineBhabha}$ for e5x5cms>0.4

| | 2009 0320 | 2010 0320 | 2011 0320 | 2012 0320 | 2013 0320 | 2014 0320 | 2015 0320 | 2016 0320 | 2017 0320 | 2018 0320 | 2019 0320 | 2020 0320 | 2021 0310 | 2022 0310 | 2023 0310 | 2024 0320 | 2024 0326 |
|------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| R(%) | 0.1 | 1.04 | 0.57 | 0.45 | 0.62 | 0.54 | 0.64 | 0.58 | 0.65 | 1.69 | 1.79 | 0.48 | 2.0 | 1.47 | 2.5 | 1.2 | 0.86 |
| data | psip | psipp | psipp | psip | XYZ | XYZ | Rscan | 4180 | XYZ | Jpsi | XYZ | XYZ | psip | psipp | psipp | psipp | psipp |



EMC软件和数据处理：簇射能量修正 (I)

Status of Offline Software Coll. Meeting Dec.4-8, 2023

BOSS 7.1.1

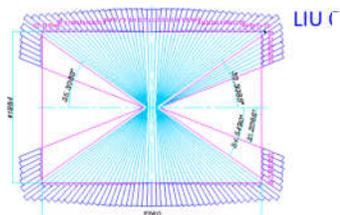
For round16 $\psi(3770)$ data

Step1: Correction using simulated single γ samples.

- Based on the symmetry geometry of $\cos\theta$
- θ id dependence 0~27 corresponding

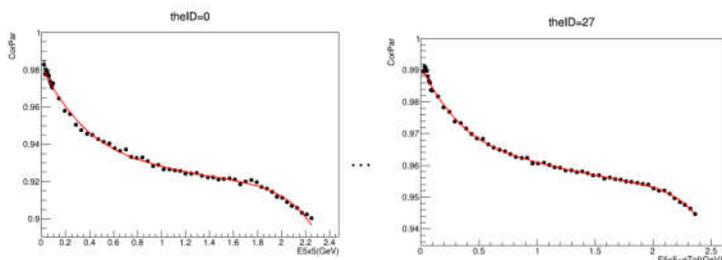
Barrel: $(E_{5\times5} + E_{TOF}) \rightarrow (E_{5\times5} + E_{TOF})/F_{corr}$

Endcap: $E_{5\times5} \rightarrow E_{5\times5}/F_{corr}$

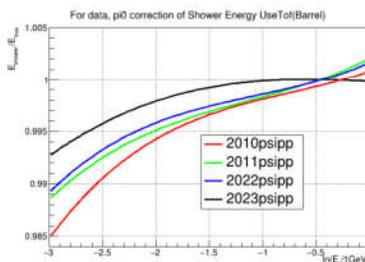
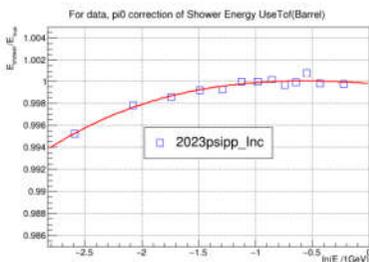
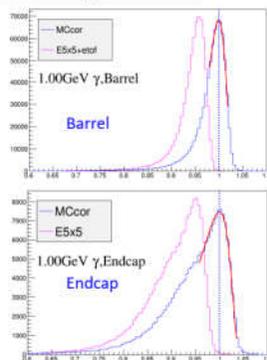


Step2: π^0 correction for data

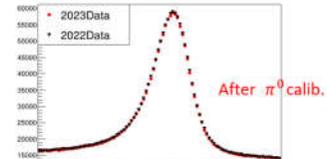
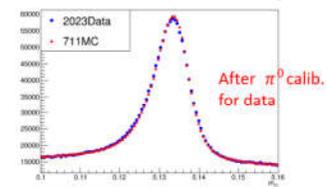
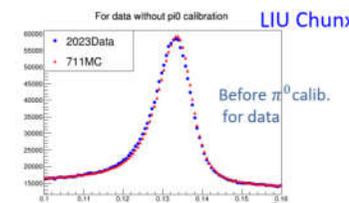
- Fine correction to improve the consistency and energy linear
- π^0 samples of real data selected with additional 2 charged tracks
- 13 energy ranges of photons, $E_{\gamma\gamma} = \sqrt{2 \cdot E_{low} \cdot E_{high} \cdot (1 - \cos\theta_{\gamma\gamma})}$



$$CorFun = \frac{p_0}{p_1 + e^{-E_{Rec}}} + p_2 + p_3 E_{Rec} + p_4 E_{Rec}^2 + p_5 E_{Rec}^3$$



$$\frac{E_{shower}(MCCor)}{E_{true}} = a_0 + a_1 \ln E + a_2 \ln^2 E + a_3 \ln^3 E$$



inclusive π^0 samples from the events with $\pi^+\pi^-K^+K^-$ + anything

| SerNo | RunFrom | RunTo | singleGammaCalib | pi0Calib | Status | SftVer |
|-------|---------|---------|--|---|--------|--------|
| 9 | 74031 | 1000000 | /cvmfs/bes3.ihep.ac.cn/CalibConst/emc/ShEnCalib/7.1.1/evsetTofCorFunctionPar2023psipp.conf | /cvmfs/bes3.ihep.ac.cn/CalibConst/emc/ShEnCalib/7.1.1/c3ptof2023psipp.txt | OK | 7.1.1 |

EMC软件和数据处理：簇射能量修正 (II)

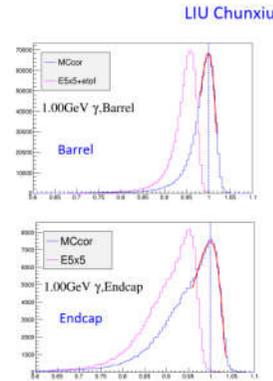
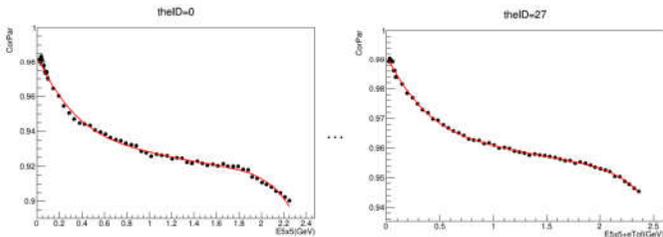
Status of Offline Software Coll. Meeting July 2-6 2024

EMC Shower Calibration of Round17 $\psi(3770)$ data

Step1: MC calibration using simulated single photon

- Based on the symmetry geometry of $\cos\theta$
- θ id dependence 0~27

Step2: π^0 calibration for real data



Barrel: $(E_{5\times 5} + E_{TOF})/F_{corr}$

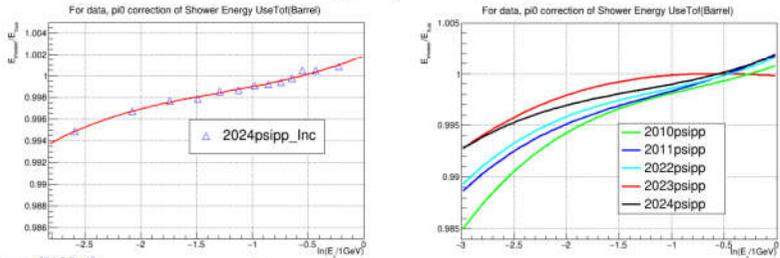
Endcap: $E_{5\times 5}/F_{corr}$

$$F_{corr} = \frac{p_0}{p_1 + e^{-E_{Rec}}} + p_2 + p_3 E_{Rec} + p_4 E_{Rec}^2 + p_5 E_{Rec}^3$$

Step1: MC calibration using simulated single photon

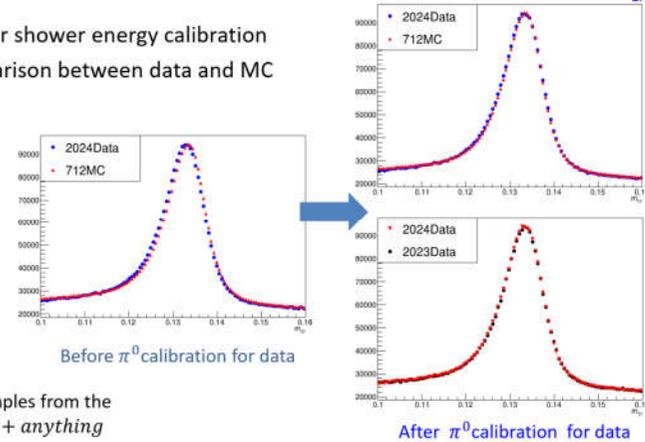
Step2: π^0 calibration for real data

- Fine correction to improve the consistency and energy linear
- π^0 samples of real data selected with additional the charged tracks=2.
- 13 energy ranges of photons, $E_{\gamma\gamma} = \sqrt{2 \cdot E_{low} \cdot E_{high} \cdot (1 - \cos\theta_{\gamma\gamma})}$



$$\frac{E_{shower}(MCCor)}{E_{true}} = a_0 + a_1 \ln E + a_2 \ln^2 E + a_3 \ln^3 E$$

Validation checks for shower energy calibration π^0 inv. mass comparison between data and MC

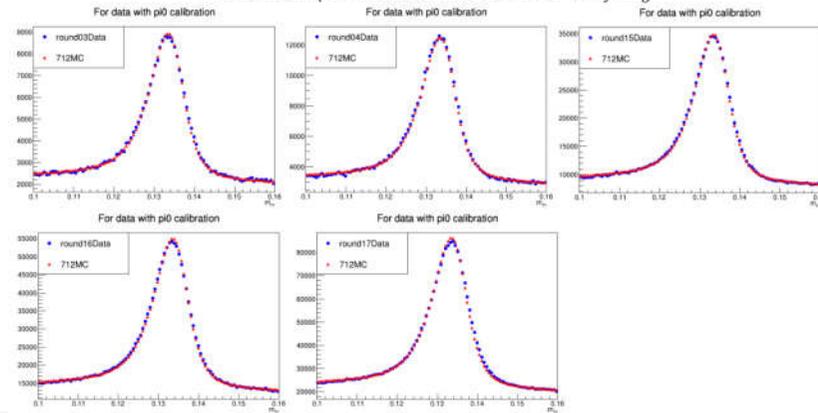


Inclusive π^0 real data samples from the events with $\pi^+\pi^-K^+K^- + anything$

@BOSS7.1.2 round03,round04,round15,round16,round17 ψ'' are OK

Check of round03,04,15,16,17 @Boss7.1.2

inclusive π^0 samples from the events with $\pi^+\pi^-K^+K^- + anything$



| SerNo | RunFrom | RunTo | singleGammaCalib | pi0Calib | Status | SftVer |
|-------|---------|---------|--|---|--------|--------|
| 10 | 78615 | 1000000 | /cvmfs/bes3.ihep.ac.cn/CalibConst/emc/ShEnCalib/7.1.2/evsetTofCorFunctionPar2024psipp.conf | /cvmfs/bes3.ihep.ac.cn/CalibConst/emc/ShEnCalib/7.1.2/c3ptof2024psipp.txt | OK | 7.1.2 |

EMC软件和数据处理: round17数据质量检查

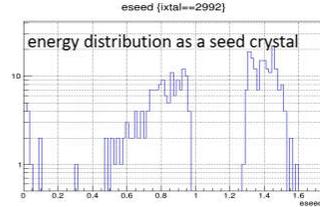
- 利用online Bhabha事例, Run by run 检查晶体状态, 保证探测器性能
- 确定死道、冒道、信号异常 → 反馈给EMC电子学 → 问题及时解决

[Status of Offline Software Coll. Meeting July 2-6 2024](#)

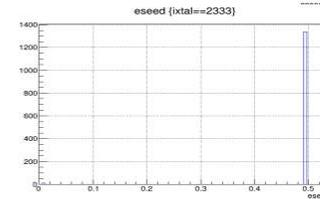
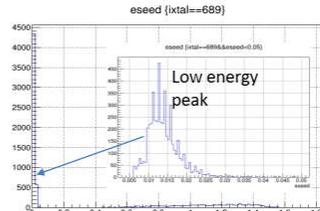
Detector Performance of Round17 data: EMC Performance

Dead and Hot Channels

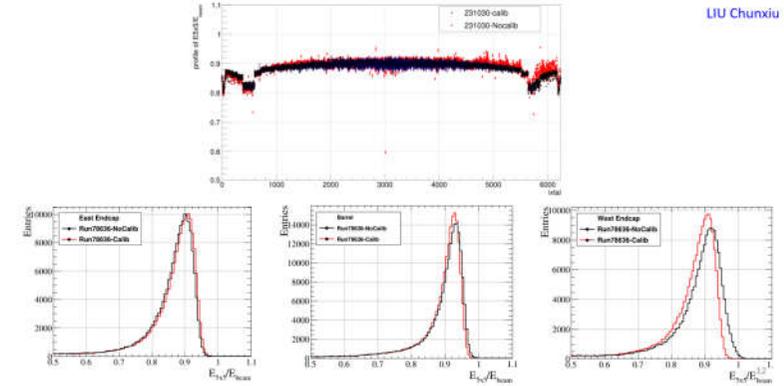
- 3 Dead channels in Barrel,
 - 3202 (Run78615-78667) : middle and high range are dead. It is OK after reinsert Q plug-in subboard.
 - 3295 (Run78684)
 - 3007 (Run78679-78868)
 - 834 (Run81015-81063)
 - 4154 (Run81020-81063)
- Hot crystals in Barrel
 - 2168 (Run78679-78684,78710,78773-78778, 81157, 81166, 81171-81172,81181-81182,81197-811228)
 - 4034 (Run80114-80116)
 - 3188 (Run81465-81470,81472-81478,81540-81563)
- Abnormal crystal 2992(Run78615-79125):
 - It is OK after replacing the main amplifier in Nov. 21 2023.
- Crystal 689 with a low energy peak(10-20MeV)
 - to be dealt with
- Crystal 2333 only a energy~500MeV (Run79172-79382)
 - is dead channel



LIU Chunxiu

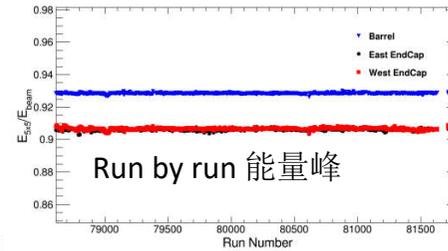


Detector Performance of Round17 data: EMC Performance

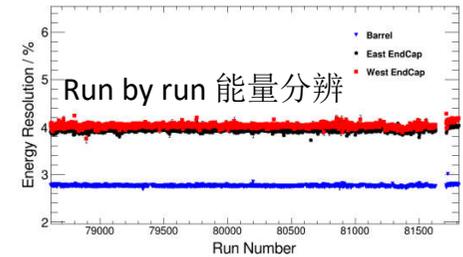


LIU Chunxiu

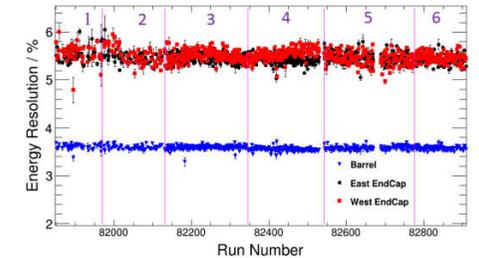
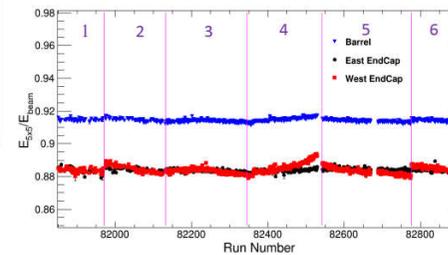
Detector Performance of Round17 data: EMC Performance



LIU Chunxiu



Psipp
3780
3800-3890Scan
3768
chic2



R Scan
6 groups of calibration constants

EMC软件和数据处理： 数据离线刻度

- 完成了round17 所有数据的EMC离线刻度（Boss7.1.1）， **124个刻度文件**
 - 选择没有电子学饱和的online Bhabha事例做刻度 / ~1-2天刻度一次
 - 对死道、冒道run做相应处理
- 根据离线能量测量，对round17 ψ'' 数据的二次刻度（Boss7.1.2）， **87个刻度文件**

round17所有数据 78615-82909 第1次刻度

共找到124条。

[首页/上一页] 1, 2, 3, 4, 5, 6, 7[下一页/尾页]

| 编号 | 起始运行号 | 结束运行号 | 文件名称 | 状态 | BOSS版本 |
|----------------------|-------|-------|---------------------------------------|----|--------|
| 1390 | 78615 | 78678 | EmcCalibConstD3202-78615-667.root | OK | 7.1.1 |
| 1634 | 78679 | 78683 | EmcCalConstH2168D3007-78679-683.root | OK | 7.1.1 |
| 1635 | 78684 | 78684 | EmcCalConstH2168D3007D3295-78684.root | OK | 7.1.1 |
| 1636 | 78685 | 78709 | EmcCalibConstD3007-78679-803.root | OK | 7.1.1 |
| 1637 | 78710 | 78710 | EmcCalConstH2168D3007-78710.root | OK | 7.1.1 |

• • •

| | | | | | |
|----------------------|-------|--------|---------------------------------------|----|-------|
| 1841 | 82132 | 82345 | EmcCalibConstIthe82132-82325-emc.root | OK | 7.1.1 |
| 1842 | 82346 | 82542 | EmcCalibConstIthe82346-82490-emc.root | OK | 7.1.1 |
| 1843 | 82543 | 82776 | EmcCalibConstIthe82543-82716-emc.root | OK | 7.1.1 |
| 1844 | 82777 | 100000 | EmcCalibConstIthe82777-82909-emc.root | OK | 7.1.1 |

round17 ψ'' 78615-81094 第2次刻度

共找到87条。

[首页/上一页] 1, 2, 3, 4, 5[下一页/尾页]

| 编号 | 起始运行号 | 结束运行号 | 文件名称 | 状态 | BOSS版本 |
|----------------------|-------|-------|--|----|--------|
| 1738 | 78615 | 78678 | EmcCalibConstD3202-78615-667-II.root | OK | 7.1.2 |
| 1739 | 78679 | 78683 | EmcCalConstH2168D3007-78679-683-II.root | OK | 7.1.2 |
| 1740 | 78684 | 78684 | EmcCalConstH2168D3007D3295-78684-II.root | OK | 7.1.2 |
| 1741 | 78685 | 78709 | EmcCalibConstD3007-78679-803-II.root | OK | 7.1.2 |
| 1742 | 78710 | 78710 | EmcCalConstH2168D3007-78710-II.root | OK | 7.1.2 |

• • •

| | | | | | |
|----------------------|-------|--------|---|----|-------|
| 1821 | 81000 | 81014 | EmcCalibConstIthe81000-81014-II.root | OK | 7.1.2 |
| 1822 | 81015 | 81019 | EmcCalibConstD834-81015-81019-II.root | OK | 7.1.2 |
| 1823 | 81020 | 81063 | EmcCalibConstD834D4154-81020-63-II.root | OK | 7.1.2 |
| 1824 | 81064 | 100000 | EmcCalibConstIthe81080-81094-II.root | OK | 7.1.2 |

EMC软件 and 数据处理： TOF能量重建的能量刻度

- 完成了round17所有数据的TOF能量刻度

- ψ'' data
Run78615-run81094 (round17)
- 3780 data
Run81095-run81272 (round17)
- 3800-3890scan data
Run81273-run81377(round17)
- 3768 data
Run81389-run81631(round17)
- Chic2 data
Run81727-run81810(round17)
- Rscan data
Run81849-run82909(round17)

| | | | | | | | | | | | |
|-------|-------|---------|----------|------------|---------------|--------------|----------------|--------------|------|---|------|
| 7.1.2 | 78615 | 79281 | 0.007348 | 0.00765373 | -0.000058913 | -0.000824422 | 0.00000242608 | -0.000413787 | NULL | 1 | Data |
| 7.1.2 | 79282 | 79897 | 0.007452 | 0.00776358 | -0.0000364943 | -0.000811473 | -0.00000407682 | -0.000465818 | NULL | 1 | Data |
| 7.1.2 | 79898 | 80484 | 0.007378 | 0.00769422 | -0.0000348561 | -0.000860385 | -0.0000222287 | -0.000420663 | NULL | 1 | Data |
| 7.1.2 | 80485 | 81094 | 0.007409 | 0.00772115 | -0.0000513117 | -0.000822463 | -0.0000100724 | -0.000452404 | NULL | 1 | Data |
| 7.1.2 | 81095 | 1000000 | 0.007401 | 0.00772049 | -0.0000345302 | -0.00088402 | -0.0000275487 | -0.000388602 | NULL | 1 | Data |
| 7.1.3 | 81811 | 1000000 | 0.007429 | 0.00773593 | -0.0000459303 | -0.000934362 | -0.00000438004 | -0.000424967 | NULL | 1 | Data |

EMC软件和数据处理: CGEM对EMC的影响

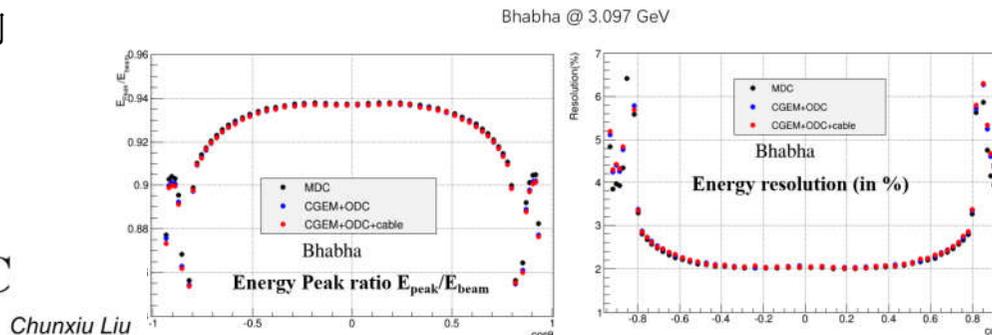
Collaboration Meeting in Summer 2024, Jul. 2 - 6, 2024

Impact to electron and positron reconstruction by EMC from CGEM long-haul cables

- CGEM 对EMC光子重建的影响
 - 光子探测效率、能量分辨
- CGEM 长电缆 对EMC的电子和光子重建的影响
 - 电子、光子能量峰位和分辨

Phys. and Soft. Workshop, Mar. 18-21, 2024

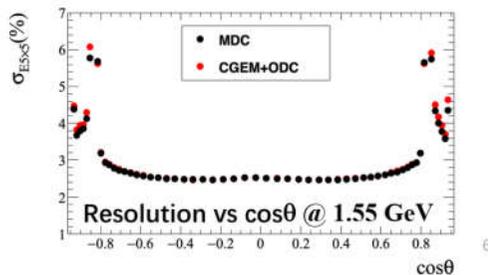
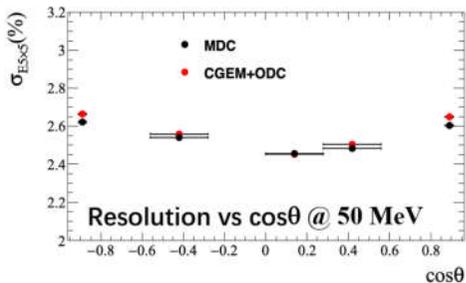
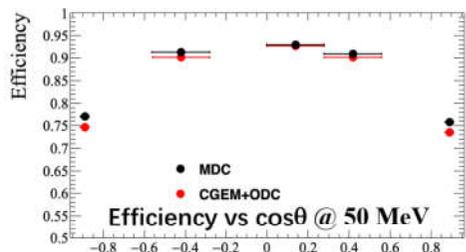
Impact to photon reconstruction by EMC



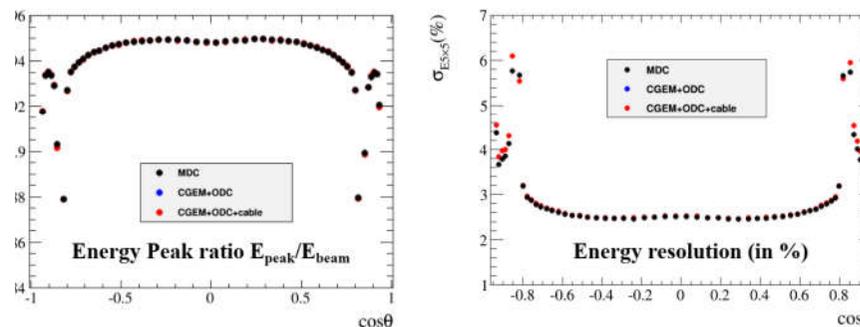
Impact to photon reconstruction by EMC from CGEM long-haul cables

- A few percent degradation
 - in both efficiency and resolution for low energy photons
 - in resolution at two endcaps for high energy photons

No big impact to photon efficiency and resolution



$e^+e^- \rightarrow \gamma\gamma$ @ 3.097 GeV



No visible effect

参加会议

1. “Integral Dose and Aging of EMC Crystals”，粒子物理实验软件与计算研讨会，2023年12月19日~21日，江门 [口头报告](#)
2. “EMC time Problem” BESIII Data Quality Meeting 2024-04-15, [口头报告](#)
3. [“2021psi \(2S\) 数据中EMC死道和特殊死道”](#)，BESIII Data Quality Meeting 2024-07-15, [口头报告](#)
4. [“特殊死道的研究”](#)，BESIII Data Quality Meeting 2024-08-12, [口头报告](#)
5. [“Check of EMC shower energy for 2024 Rscan data”](#)，BESIII Data Quality Meeting 2024-09-30, [口头报告](#)
6. BESIII Collab. Meeting and Physics & Software Workshop上 为软件总体报告 [提供EMC内容](#)
7. 粒子物理实验计算软件与技术研讨会（2024），2024年6月1日~6月4日，杭州
8. 2024 高能物理大会（青岛）

文章

- [Offline data processing system of the BESIII experiment](#) , J.h. Zou, et al, Eur. Phys. J. C (2024) 84:937 发表
- [Integral Dose and Aging of EMC Crystals in the BESIII](#), 完成Note, [文章撰写中](#)

争取项目和经费情况

- ◆参与 科技部 国家重点研发计划 重点专项 “奇特强子态及强子谱学” 项目 “软件发展与科学计算” 课题 2020YF04063004

自我评价和下年度计划

- 自我评价
 - 工作任务饱满
 - 综合业务能力强
 - 积极与组内外同事沟通、讨论，建立良好的合作关系
 - 积极参加学术活动
- 下年度工作计划
 - 累积剂量及晶体老化研究---文章撰写
 - 继续完成基于机器学习的EMC电子学特殊死道鉴别的研究、以及Note整理
 - 新数据EMC刻度、质量检查
 - 新数据EMC簇射能量的修正
 - 新数据TOF能量重建的刻度
 -等

谢谢大家！