

BESIII电磁量能器的模拟与刻度

第五小组 张松洋 王富阳 张中华
汇报时间：2026.6.18

目录

- 1.模拟与刻度概述
- 2.EMC模拟
- 3.EMC刻度
- 4.总结与展望

1.模拟与刻度概述

模拟是正问题：从物理信息出发，推断探测器响应

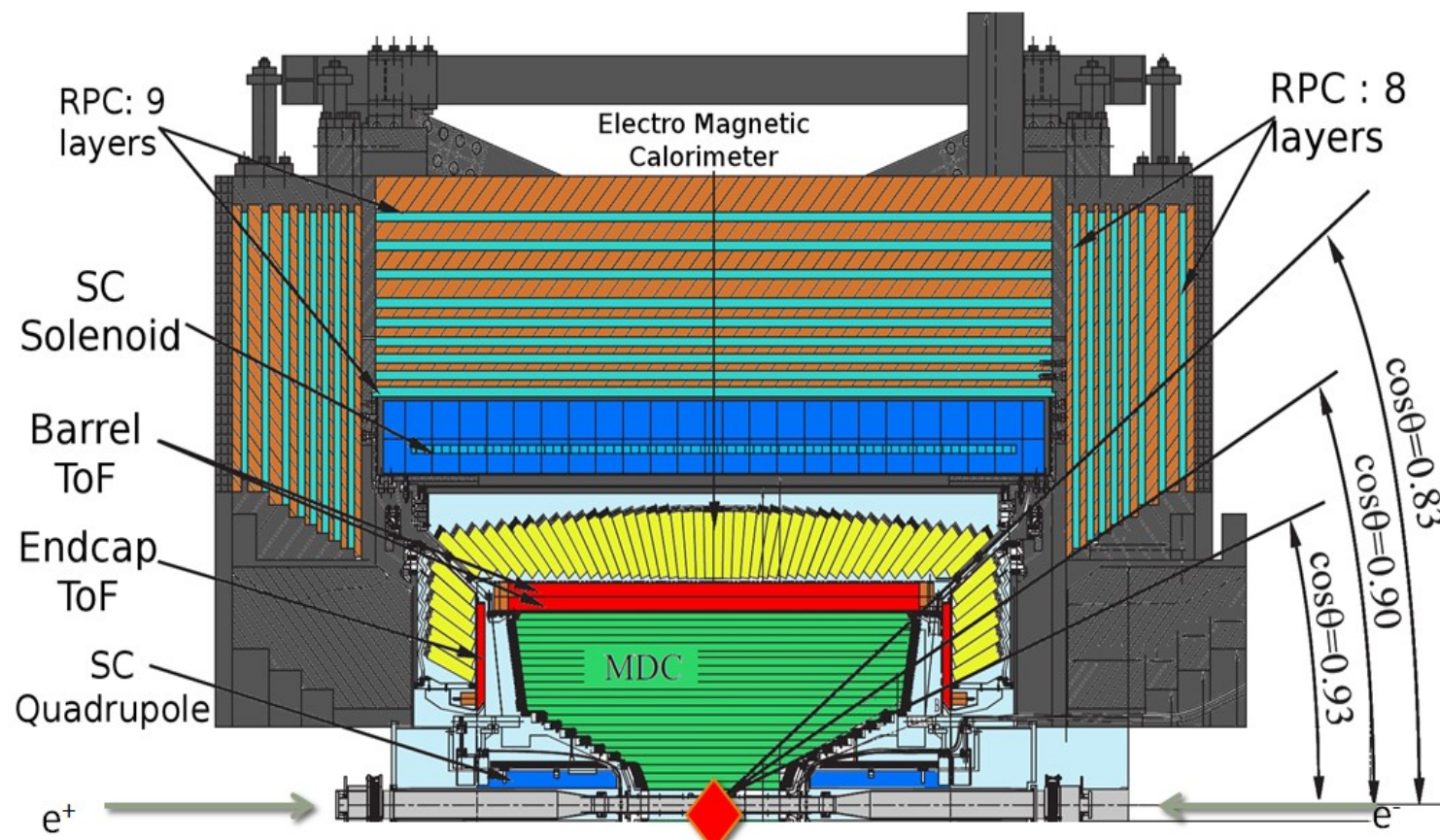
已知物理过程，通过蒙特卡洛模拟推演探测器输出。若模拟正确，输出应与真实实验符合良好，即“正中靶心”。模拟能够帮助理解探测器行为、验证分析算法、分析本底与信号、优化实验设计。

刻度是反问题：从探测器信号出发，反推物理信息

当模拟预期与真实数据存在系统性偏差（预期十环，实测八环），就需要刻度（校准）。刻度利用已知的“标准靶心”（如 π^0 质量）测量偏差，修正测量系统，使结果真实可靠地反映物理信息。



2. EMC模拟：在计算机中模拟真实探测器



CsI (TI) 晶体, 6240块, 28cm长 (15.1 X_0)
覆盖93%立体角, 总重~25吨

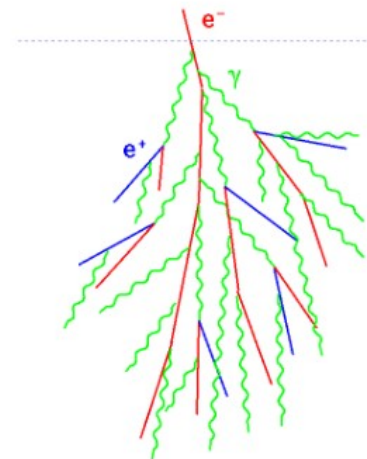
2. EMC模拟：在计算机中模拟真实探测器

基于 BOOST + Geant4 框架

- 关键流程：1. 粒子输运 (Geant4) 负责从物理上产生“击中 (Hit)”
2. 数字化 (BOOST) 负责将“击中”转化为我们最终读出的数据

利用蒙特卡罗方法，在计算机里完整重现一个“真实的物理事件”一步步变成“电子学信号”的过程。

高能电子或 γ 光子在介质中会产生电磁簇射（韧致辐射，电子对效应），收集在EMC中的沉积能量即可确定粒子的总能量。



粒子输运过程：

核心——步长竞争机制：
$$r = F(l) = 1 - e^{-\Sigma * l} \rightarrow l = -\frac{\ln(r)}{\Sigma}$$

r 为均匀分布的随机数，反解出步长 l ，在EMC中，不同效应截面的步长不同，谁的步长最短，谁就发生。这种机制确保了每个过程的发生概率严格正比于它的截面。

朴素抽样逻辑： $f(l)$ 表示飞行 l 距离的PDF

$$f(l) = \sigma e^{-\sigma l} \quad P(\text{过程 } i) = \frac{\sigma_i}{\sigma} \quad P(l, i) = f(l) \times \frac{\sigma_i}{\sigma} = (\sigma e^{-\sigma l}) \times \frac{\sigma_i}{\sigma} = \sigma_i e^{-\sigma l} \quad \int_0^{\infty} P_{\text{sample}}(l, i) dl = \sigma_i \int_0^{\infty} e^{-\sigma l} dl = \frac{\sigma_i}{\sigma} \quad 5$$

2.EMC模拟：在计算机中模拟真实探测器

步长竞争机制的可行性^[1]：与直接采样在物理上等价

- **Distribution of interactions:**

- The probability that the particle makes the $0 \rightarrow \ell$ travel without interaction from process i is:

$$S(\ell; \sigma_i) = 1 - C(\ell; \sigma_i) = \exp(-\sigma_i \cdot \ell)$$

- The probability that the particle survives the $0 \rightarrow \ell$ travel for **ALL** processes is then

$$\begin{aligned} S(\ell; \sigma_1) \times S(\ell; \sigma_2) \times \dots \times S(\ell; \sigma_n) &= \exp(-\sigma_1 \cdot \ell) \times \exp(-\sigma_2 \cdot \ell) \times \dots \times \exp(-\sigma_n \cdot \ell) \\ &= \exp(-\sigma \cdot \ell) \\ &= S(\ell; \sigma) \end{aligned}$$

- **So the distribution of ℓ is the same !**

- **Relative frequencies of processes:**

- The probability for process i to happen is:

$$\begin{aligned} \mathbb{P}(\ell_i < \ell_{1\dots j \neq i \dots n}) &= \int_0^\infty \mathbb{P}(\ell_1 > x) \dots \mathbb{P}(\ell_{j \neq i} > x) \dots \mathbb{P}(\ell_n > x) \times f(x; \sigma_i) dx \\ &= \int_0^\infty S(x; \sigma_1) \dots S(x; \sigma_{j \neq i}) \dots S(x; \sigma_n) \times \sigma_i \cdot \exp(-\sigma_i \cdot x) dx \\ &= \int_0^\infty \exp(-\sigma_1 \cdot x) \dots \exp(-\sigma_{j \neq i} \cdot x) \dots \exp(-\sigma_n \cdot x) \times \sigma_i \cdot \exp(-\sigma_i \cdot x) dx \\ &= \sigma_i / \sigma \end{aligned}$$

- **So the same probability for process i to be applied !**

[1]Agostinelli S, et al. Geant4 simulation toolkit[J]. NIM A, 2003, 506: 250-303.

2.EMC模拟：在计算机中模拟真实探测器

数字化过程：

第一步，波形生成。CsI(Tl)晶体的闪烁光不是瞬间发出的，而是有一个衰减过程。衰减时间常数 $\tau=1$ 微秒，波形的数学形式为 $h(t) = 1/24 * t^4 * e^{-t/\tau}$ 。[2]这是一个先快速上升、后缓慢衰减的脉冲。

第二步，采样与提取。BESIII的电子学系统从峰值前1.5微秒开始，每50纳秒采样一次，共采60个点。采样点的最大值就是ADC值——对应能量；最大值出现的位置就是TDC值——对应时间。

第三步，噪声模拟。为了让模拟更真实，还要叠加噪声。电子学噪声是高斯分布的，直接加到每个采样点上。此外还有束流相关本底——从真实数据的随机触发中提取本底波形，与信号波形混合。

经过数字化，我们就得到了与真实数据格式完全相同的原始数据。接下来，重建算法会读取这些原始数据，通过聚类和簇分裂算法，把分散的每个晶体的Hit信息组合成有物理意义的Shower。

[2]He M., J. Phys.: Conf. Ser., 293, 012025 (2011)

3.1 单晶体刻度

由于每块晶体的光产额、光电二极管效率、电子学增益都不完全相同，即使入射光子能量相同，但晶体的读出也不一定相同，打在B晶体上读出是0.89，而打在C晶体上读出是0.9

单晶体刻度的目标，就是为每块晶体找到一个刻度常数 c_i ，将电子学读出转换为能量信息，让所有晶体的响应与预期“对齐”。

可选用Bhabha散射事件来实现。^[3]

$$E_i = \frac{ADC_i - PED_i}{e_i} * c_i$$

ADC: 电子学读出 PED: 基线 e_i : 增益系数 i:第i个晶体



Fig. 3.1.1电磁簇射示意图：粒子击中晶体，引起电磁簇射，引发横向和纵向的发展

[3] Liu C.X., J. Phys.: Conf. Ser., 293, 012048 (2011)

3.1 单晶体刻度

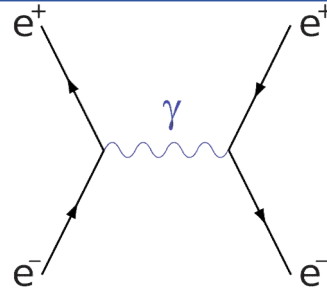


Fig. 3.1.2 Bhabha散射示意图：正负电子的两体散射

优点：

1. 两体散射，Bhabha散射运动学精确已知，可精确计算出末态能量
2. 事件的统计量大，便于实验测量

注：理论计算出的能量是电子在对撞点的能量。电子在飞行过程中，会穿过束流管、MDC外壁、TOF等物质，损失一部分能量。所以，真正通过探测器得到的“期望能量”，还需要修正：

$$E_{exp} = E_{theory} * f(E, \theta, \varphi)$$

其中， $f(E, \theta, \varphi)$ 是MC模拟给出的修正因子，体现了模拟对刻度的支撑

3.1 单晶体刻度

参数提取：最小二乘法

$$\chi^2 = \sum_{k=1}^N \frac{(E_{exp}^k - \sum_i^{5 \times 5} g_i E_i^k)^2}{\sigma_k^2}$$

其中， k 表示第 k 次事件， E_{exp}^k 表示第 k 次事件能量的期望值， E_i^k 表示第 k 次事件第 i 个晶体的能量读出值， σ_k^2 为第 k 次事件的误差， g_i 为第 i 个晶体的刻度常数

$$\frac{\partial \chi^2}{\partial g_i} = 0 \quad \rightarrow \quad \sum_j g_j Q_{ij} = R_i \quad Q_{ij} = \sum_k \frac{E_{ik} E_{jk}}{\sigma_k^2} \quad R_i = \sum_k \frac{E_{ik} E_{exp}^k}{\sigma_k^2}$$

Q_{ij} 代表晶体 i 和晶体 j 在所有事件中的关联强度， R_i 代表晶体 i 与期望能量的总关联

这是一个 6240×6240 的线性方程组。但它是稀疏的——因为每个事件只击中几十块晶体，绝大多数矩阵元为0，采用 SLAP 稀疏矩阵方法来求解这个方程组，就能得到所有晶体的刻度常数 g_i

求解完成后，更新刻度常数： $c_i^{new} = c_i^{old} * g_i$

3.1 单晶体刻度

Bhabha calibration flow chart

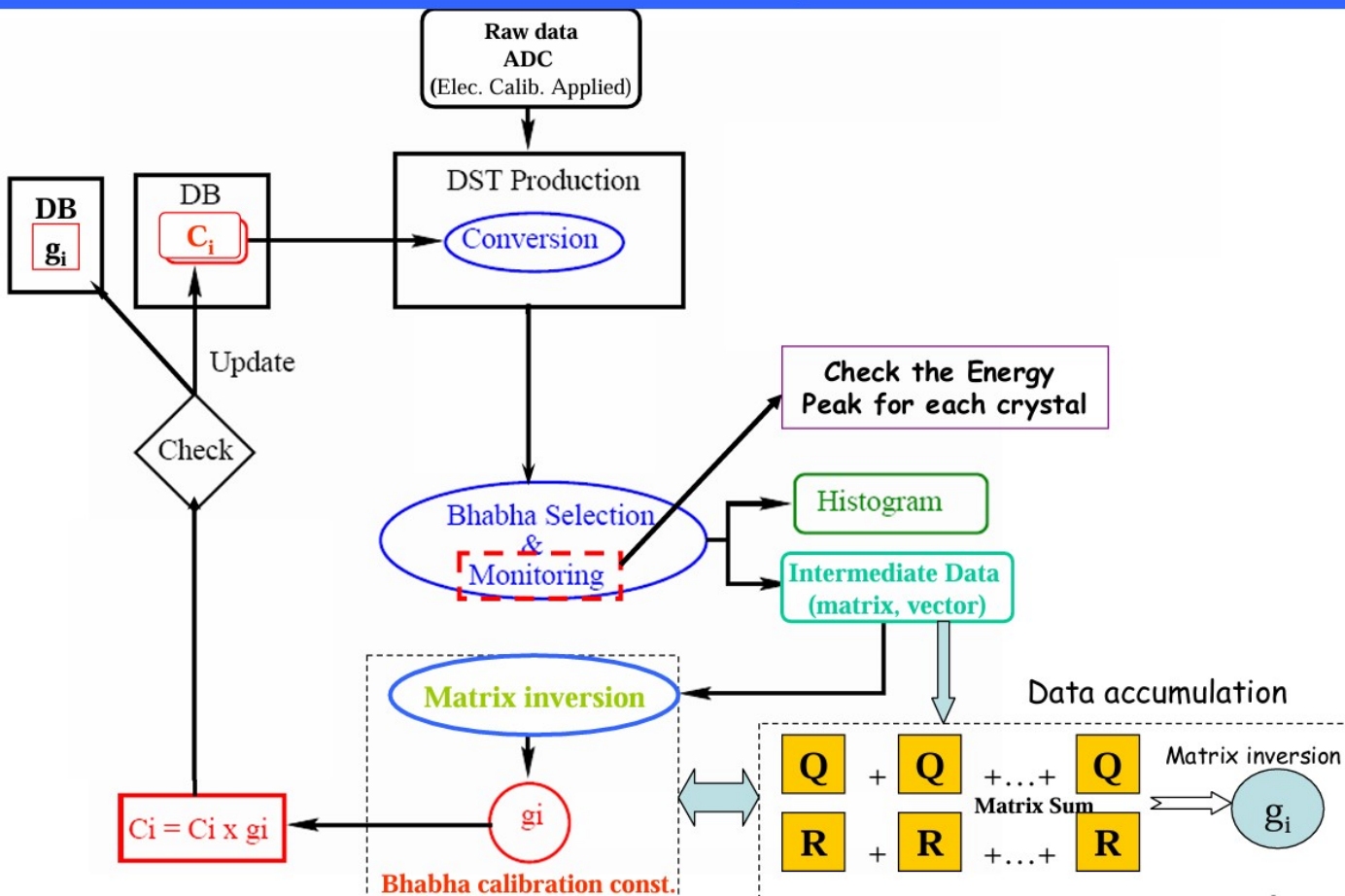
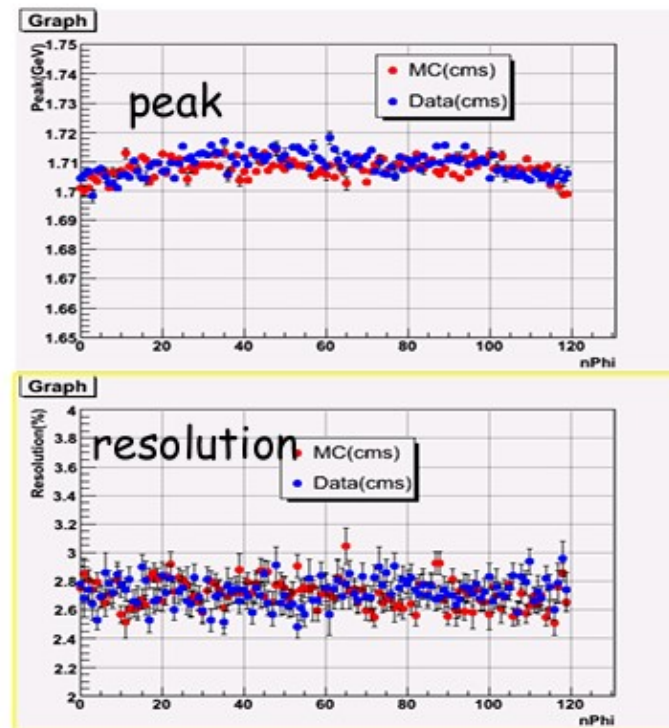


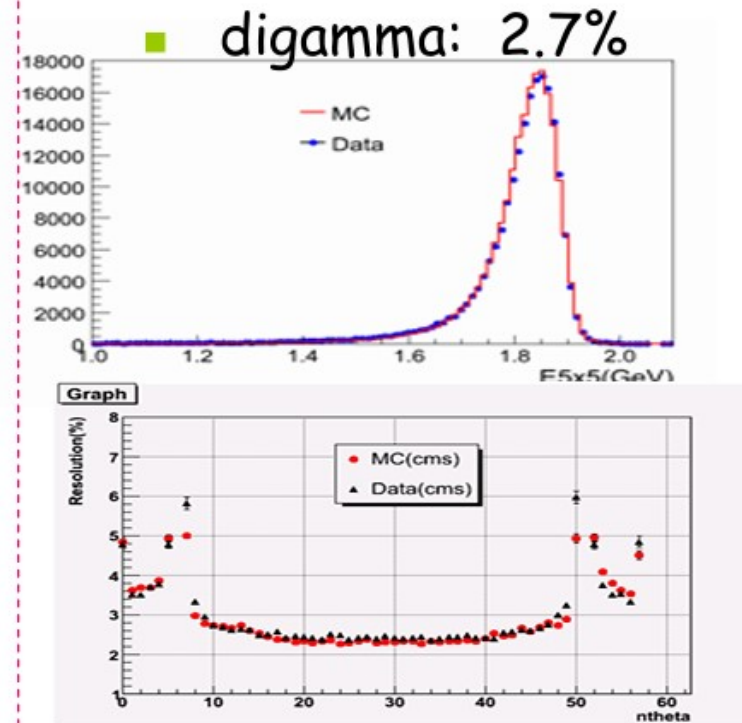
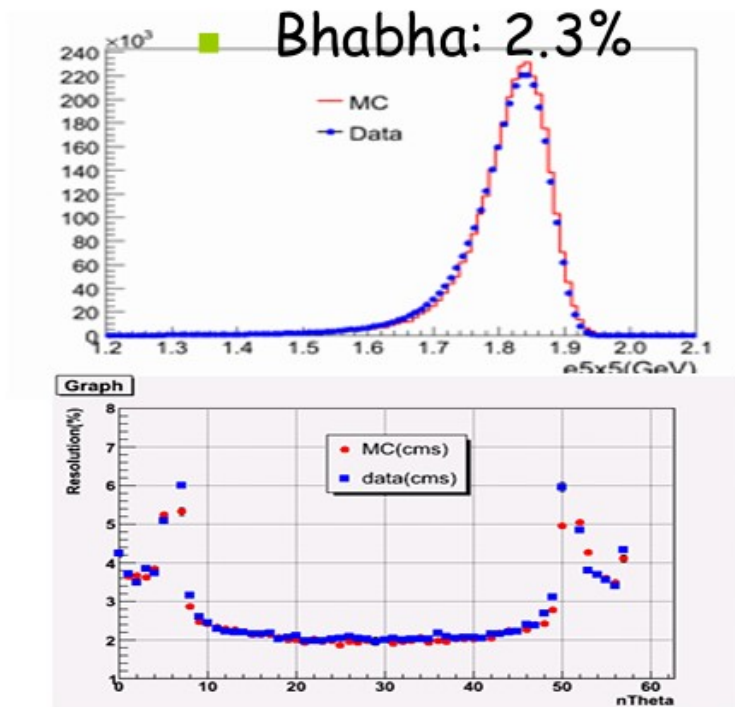
Fig. 3.1.3 刻度流程图



E5x5cms vs. phi-index
in barrel

Fig. 3.1.4 5x5晶体能量在 ϕ 方向分布

Data @3.686GeV



Energy resolution for electrons (from $e^+e^- \rightarrow e^+e^-$) 2.3% in barrel and 4.1% in endcap, for photons (from $e^+e^- \rightarrow \gamma\gamma$) reaches 2.7% in barrel and 4.2% in endcap

3.2 绝对能量刻度

经过Bhabha刻度，可以解决晶体间响应不一致的问题，但以下效应仍然存在：

1. 前端物质损失 光子/电子在到达EMC前，穿过束流管、MDC、TOF
2. 簇射横向泄漏 簇射能量扩散到 5×5 晶体范围之外
3. 簇射纵向泄漏 高能簇射从晶体后端泄漏

导致测得的物理量与模拟不符，且上述问题对于不同能量是非线性的，因此需要对整体的能标进行修正
思路类似：用一个已知的物理常数作为标尺，来校正这个系统偏差—— π^0 的不变质量

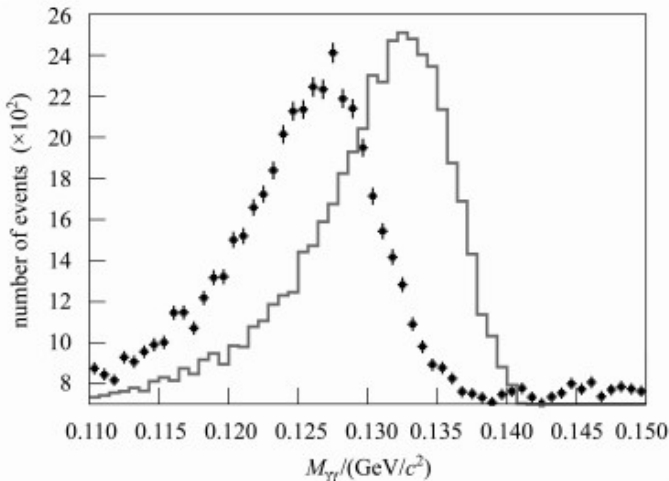


Fig. 4. Distribution of invariant mass of π^0 in inclusive $\psi(2S)$ decays in real data (points with error bars) and Monte Carlo (histogram).

Fig. 3. 2. 1 数据和MC模拟在绝对能量刻度前得到的双光子不变质量谱：数据和MC模拟的峰位有明显偏移[4]

[4] Bian et al., Chin. Phys. C 34, 72 (2010)

3.2 绝对能量刻度

考虑到光子能量沉积分布不是高斯分布，而是带低能尾巴的非对称分布，用Novosibirsk函数拟合质量峰

$$f_{\text{Nov}}(m) = A \exp \left(-\frac{\ln^2 \left(1 + t \frac{\sinh(t\sqrt{\ln 4})}{t\sqrt{\ln 4}} \cdot \frac{m-m_0}{\sigma} \right)}{2t^2} - \frac{t^2}{2} \right)$$

π^0 的不变质量计算公式：

$$m_{\gamma\gamma}^{\text{exp}} = \sqrt{2E_1E_2(1 - \cos\theta_{\gamma\gamma})}$$

设修正因子为指数形式：

$$\exp(\alpha_i) = E_{\text{shower}}/E_{\text{true}}$$

可得修正质量的关系为：

$$m_{\gamma\gamma}^{\text{corr}} = m_{\gamma\gamma}^{\text{exp}} * \exp(-\alpha_i/2 - \alpha_j/2), \text{取对数: } \ln m_{\gamma\gamma}^{\text{corr}} = \ln m_{\gamma\gamma}^{\text{exp}} - \alpha_i/2 - \alpha_j/2$$

$$C_{ij} = \ln m_{\gamma\gamma}^{\text{data}} - \ln m_{\gamma\gamma}^{\text{exp}}$$

$$\chi^2 = \sum_{i,j} \frac{(\alpha_i/2 + \alpha_j/2 - C_{ij})^2}{\sigma_{i,j}^2}$$

α_i 为待求的第*i*个能量bin的(光子1/2)刻度因子参数，利用最小二乘法则可以得到各个bin的刻度因子

3.2 绝对能量刻度

通过最小二乘后，得到了8个bin的比值，再拟合整体

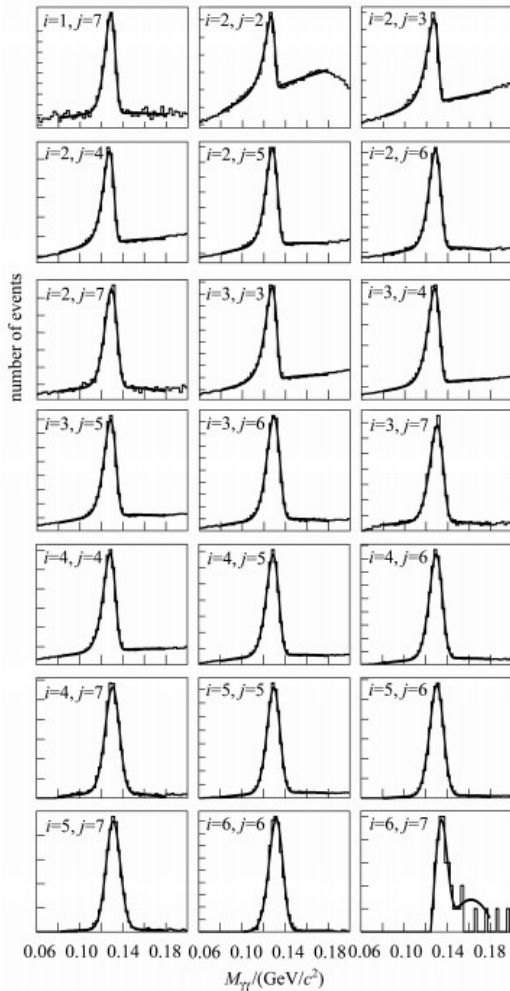


Fig. 6. The invariant mass of $\gamma\gamma$ for different energy bin combinations (ij). Both of the photons are required to be in the barrel.

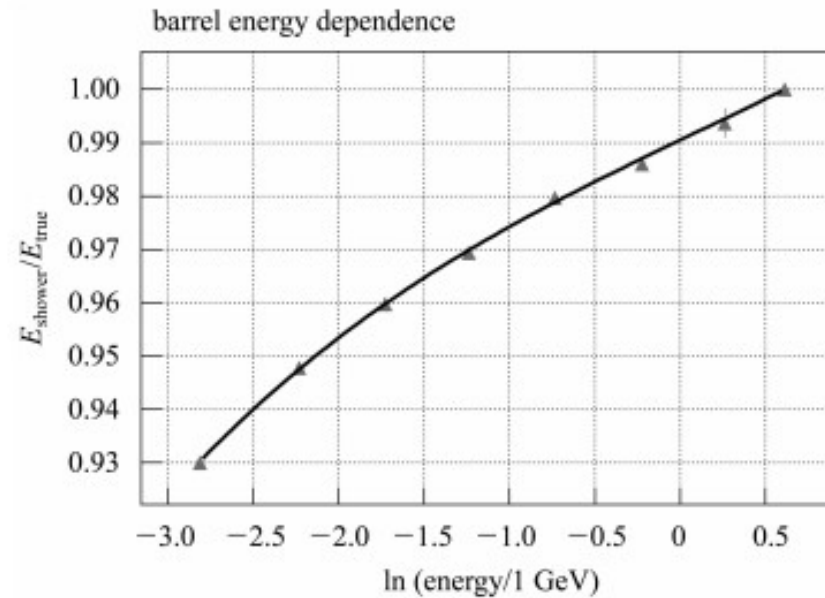


Fig. 7. Photon energy correction factors versus $\ln E$ obtained from π^0 calibration.

$$E_{shower}/E_{true} = a_0 + a_1 \ln E_{shower} + a_2 (\ln E_{shower})^2$$

3.2 绝对能量刻度

刻度效果的验证：选用其他控制样本，证明普适性

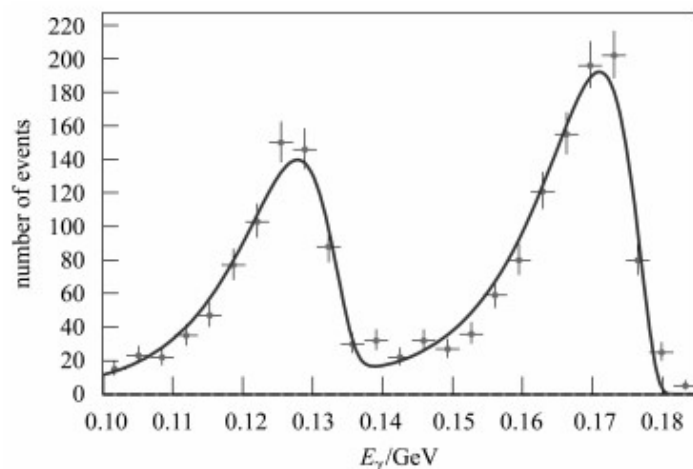


Fig. 8. Photon energy in $\psi(2S) \rightarrow \gamma\chi_{c1,2}(1P)$ with π^0 calibration.

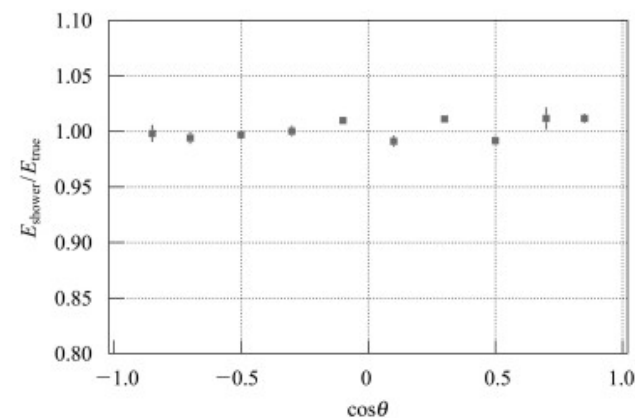


Fig. 9. Geometry dependence of the γ energy scale after π^0 calibration obtained from analyzing $\psi(2S) \rightarrow \gamma\chi_{c1}(1P)$.

Fig. 3. 2. 2 数据和MC模拟得到的光子能量谱

Fig. 3. 2. 3 数据和MC模拟的光子能量比值角分布图，说明刻度效果良好

刻度后也并非完美

我目前的工作：测量 $D^0 \rightarrow K_s \eta, K_l \eta$ 的分支比, $\eta \rightarrow \gamma\gamma$

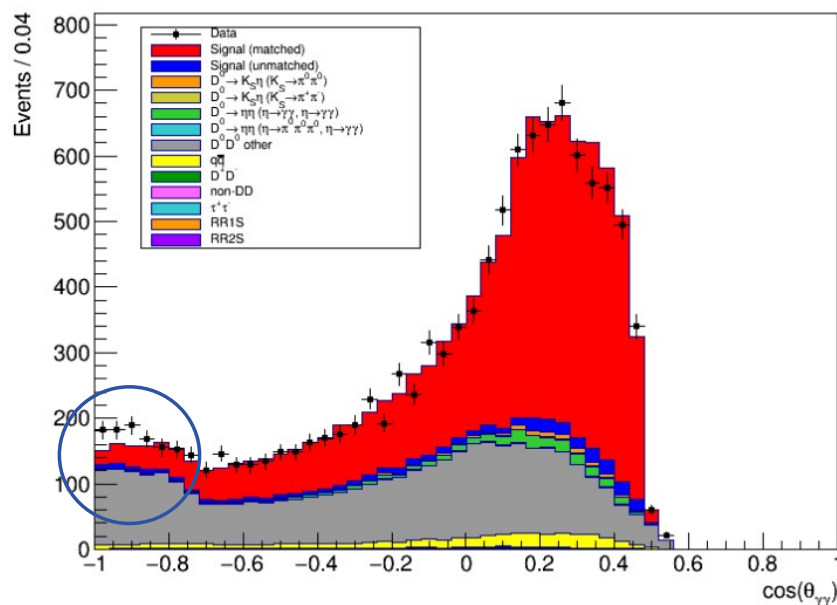
K_l 是强子，可能在EMC中发生强子簇射，该过程较为复杂，簇射产生的光子，被误组合成 η 的光子

在比较数据和MC一致性时，发现在实验室系下，两个光子方向背对背时，数据和MC符合不一致，该部分主要事例为 $D^0 \rightarrow K_l \pi^0$ 。

经过分析，发现此时两个光子一个来源于 K_l ，一个来源于 π^0 ，数据和MC符合不一致的原因很可能来源于MC模拟中对 K_l 与探测器相互作用的估计不准确

实验室系下两个光子的cos角

Photon Pair Opening Angle in Lab Frame



4.总结与展望

总结：

模拟和刻度是硬币的两面。模拟是正问题，从物理信息出发预测探测器信号，输出MC数据，为我们提供“标准答案”。刻度是反问题，从探测器信号出发反推物理信息，输出刻度常数，修正系统的偏差。二者相辅相成，最终才能保证物理结果真实可靠。

在BESIII上的模拟，基于GEANT4和BOOST模型，通过最小步长的抽样方法模拟相互作用的截面，再通过数字化过程得到HIT信息，再利用聚类算法将HIT信息转化为shower信息。

在BESIII上进行刻度，两步刻度缺一不可。第一步利用Bhabha散射刻度解决晶体间响应不一致，第二步利用 π^0 衰变修正整体能量标度，二者均使用了最小二乘的优化方法。

展望：

将无监督学习方法（Autoencoder）应用于EMC，自动检测异常晶体，相比传统方法大幅节省人力。^[4]
使用神经网络方法来替代传统的Geant4模拟。

.....

附录

宇宙线预刻度：提供初始值，依赖MC真值

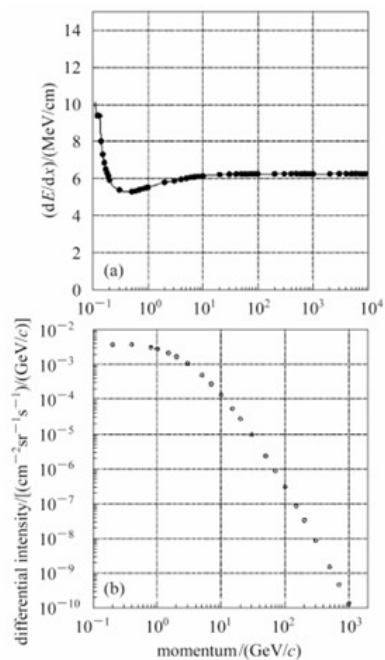


图 2 μ 在 CsI(Tl) 中的 dE/dx 随动量变化曲线(a)和宇宙线中 μ 的动量谱(b)

$$C_i^{LY} = \frac{LY_{\text{average}}}{LY_i} \quad C_i^{MC} = \frac{E_{\text{average}}(MC)}{E_{\text{average}}(data)}$$

$$C_i^{\text{cosmic}} = C_i^{LY} \times C_i^{MC}$$

LY:相对光产额

1. **边界处理精度高**: 能够准确停在几何边界上, 避免粒子跨过界线而忽略不同介质中的物理过程变化。
2. **多过程协调**: 自动协调多个物理过程的作用顺序, 保证不同过程在同一能量/距离尺度下同时被正确地模拟, 不会因为某个过程步长过长而错过其他过程的发生点。
3. **效率与准确性平衡**: 由过程自身决定步长限度, 既能捕捉小概率但重要的相互作用 (如硬碰撞), 又不至于在无相互作用的区域浪费计算资源。

