Report

tag: sPH-CALOCALIB-2023-001
version: 0.1
DOI: unspecified
date: August 26, 2023

Extraction of Calibrated Invariant Mass Distributions for Quark Matter 2023

Anthony Hodges with enormous help and input from the calorimeter calibrations working group

Abstract

This note describes the process and modules used to create calibrated DST's from pre-produced DST's from a set of "golden runs" chosen from the Run 23 Commissioning Data set. Documention on the derivations of these calibrations, in particular the process of calibrating the towers in the EMCal to simulation and flattening their response, as well as the MC-based calibration to the cluster position, can be found in Notes and , respectively. The creation of dead and hot maps used to mask bad towers will be described here, however. And, finally, the analysis procedure used to measure the invariant mass of di-cluster pairs will be detailed, and the resultant invariant mass distributions will be shown for different selections of p_T of the reconstructed particle.

Introduction

Measurement of neutral pions and eta mesons are a key part of assuring that the EMCal is calibrated appropriately to the electromagnetic scale. In order to reconstruct these neutral mesons, pairs of cluster (groups of EMCal towers) are selected and reconstructed into a single particle. The invariant mass of this particle can be calculated from the cluster kinematics.

Run selection

Runs are selected based on their coverage (namely based on the number of Sub-Event Buffers, or SEB's present) and the presence of a correlation between the total charge in the Minimum Bias Detector (MBD) and the total tower energy in the EMCal. The run numbers used and the location of their calibrated DST's is shown in Table 1 below. All DST's have EMCal, HCAl, ZDC, and MBD information.

Run Number	SEB's	Events	DST Location
21813	7/8	1.02M	/direct/sphenix+tg+tg01/jets/ahodges/run23_production_zvertex/21813
21796	8/8	518k	/direct/sphenix+tg+tg01/jets/ahodges/run23_production_zvertex/21796
21615	8/8	935k	/direct/sphenix+tg+tg01/jets/ahodges/run23_production_zvertex/21615
21599	8/8	916k	/direct/sphenix+tg+tg01/jets/ahodges/run23_production_zvertex/21599
21598	8/8	889k	/direct/sphenix+tg+tg01/jets/ahodges/run23_production_zvertex/21598
21891	7/8	426k	/direct/sphenix+tg+tg01/jets/ahodges/run23_production_zvertex/21891
22979	5/8	345k	/direct/sphenix+tg+tg01/jets/ahodges/run23_production_zvertex/22979
22950	5/8	343k	/direct/sphenix+tg+tg01/jets/ahodges/run23_production_zvertex/22950
22949	5/8	337k	/direct/sphenix+tg+tg01/jets/ahodges/run23_production_zvertex/22949
22951	5/8	335k	/direct/sphenix+tg+tg01/jets/ahodges/run23_production_zvertex/22951
22982	5/8	291k	/direct/sphenix+tg+tg01/jets/ahodges/run23_production_zvertex/22982
21518	8/8	120k	/direct/sphenix+tg+tg01/jets/ahodges/run23_production_zvertex/21518
21520	8/8	115k	/direct/sphenix+tg+tg01/jets/ahodges/run23_production_zvertex/21520
21889	7/8	706k	/direct/sphenix+tg+tg01/jets/ahodges/run23_production_zvertex/21899

Table 1: List of golden runs, their EMCal SEB count, approximate number of events, and DST location on SDCC

Production

This section will detail, step-by-step, the production process in terms of the subsystem reconstruction modules used. This production starts with pass 1 DSTs, right above the raw-data format of the prdf (PHENIX Raw Data File). The base DST's have towers (all subsystems in these DSTs are treated with the TowerInfoContainer object designed by Tim Rinn and Antonio da Silva and thus will be referred to as towers) with their energy or charged (in the case of the MBD) in uncalibrated ADC's.

0.1 Tower-by-Tower Calibrations

The first phase of the production process is carried out by the CaloTowerCalib module, which takes in each tower in the EMCal and HCal (MBD and ZDC object *could* feasibly have calibrations applied by this module, but do not) and applied a calibration factor. The EMCal's calibrations are derived in detail in Note, and essentially flatten the towers' response and calibrate them to the simulation's energy scale. The HCal's energy scale is set via cosmics taken in 1008 as in [1].

Dead Map Application

There are three types of towers that are masked during the production: dead, cold, and hot towers. The definition of each type of tower and what qualifies them to their designation is in the list below. Note that this is specifically for the EMCal.

- Hot Towers: these towers record a number of hits greater than 4σ above the mean number of hits reported across the calorimeter. During the QA process, distributions of hits in towers at a give energy are made, and towers that exceed this 4σ threshold are excluded. 4σ was chosen as the cutoff which is strict enough to exclude notable outliers, but relaxed enough as to not carve huge chunks out of the EMCal's acceptance. The how tower selection process also considers statistics at the IB (interface board) and sector level. IB's and Sectors that have over $> 70\%$ of their towers labelled as hot are masked in their entirety. This cut is made to exclude IB's or sectors which might accidentally be in high gain mode.
- Cold Towers: Essentially the opposite of hot towers, these towers under report hits in a given energy range when compared to the average across all other towers. A 3.5σ cut is used here to account for the asymmetric response distribution.
- Dead Towers: These are typically towers in SEB's which are not in a given run.

Unfortunately, the production-level code meant to produce hot and dead maps on the fly is still under development, and has a tendancy to miss some hot and dead regions. This missing happens specifically because of the codes reliance on fiducial cuts to remove extreme outliers and prevent them from contaminating calculation of the standard deviations. For this production, the fiducial cuts were not harsh enough, leaving it notable outliars on the low and high end. These outliars are taken out by a secondary piece of code that removes the hottest ten towers as well as any towers beneath 40% the mean number of hits in the calorimeter for a given run.

One of the key aspects of this tower masking is information stored in the CDBTTree. Towers are marked as 0 for being good, 1 for being dead or cold, and 2 for being hot. Deadmask application is handled by two macros:

- DeadHotMapLoader first loads the map while

- TowerInfoDeadHotMask is responsible for iterating over each tower and declaring it good or bad. Good towers are simply untouched by the masker, but bad towers have their energy set to zero as well as their "time" asset (the sample number corresponding to the peak position of the waveform in the PRDF set to -11 or -10 for dead and hot towers, respectively. This distinction was requested by the Jet Structure Topical Group.

The CDBTTrees used to mask the towers are located here temporarily before upload to the CDB by Chris Pinkenburg:

```
EMCal: /gpfs/mnt/gpfs02/sphenix/user/ahodges/macros_git/analysis/checkMaps/macro/
    finalMaps/{RUNNUMBER}CEMCbadTowerMapCDBTree_v1.root
OHCal: /gpfs/mnt/gpfs02/sphenix/user/ahodges/macros_git/analysis/checkMaps/macro/
    finalMaps/{RUNNUMBER}HCALOUTbadTowerMapCDBTree_v1.root
IHCal: /gpfs/mnt/gpfs02/sphenix/user/ahodges/macros_git/analysis/checkMaps/macro/
    finalMaps/{RUNNUMBER}HCALINbadTowerMapCDBTree_v1.root
```

Watch out for the line numbers when copying and pasting, and insert the corresponding run number where {RUNNUMBER} appears.

Impact of EMCal Tower Maps

Fig. 1 shows the impact of the EMCal maps on uncalibrated data (that is, the initial DST's produced with no calibrations applied) from Run 21813. The top left is the initial masking rendered by the masking software. One can see its scale is still being set by a few residual hot towers that biased the threshold because the initial fiducial cut was not harsh enough. The top right figure shows the impact of remove the ten hottest towers, which one can see clears up the distribution considerably. Finally, the bottom panel shows the effect of removing the cold towers left over.

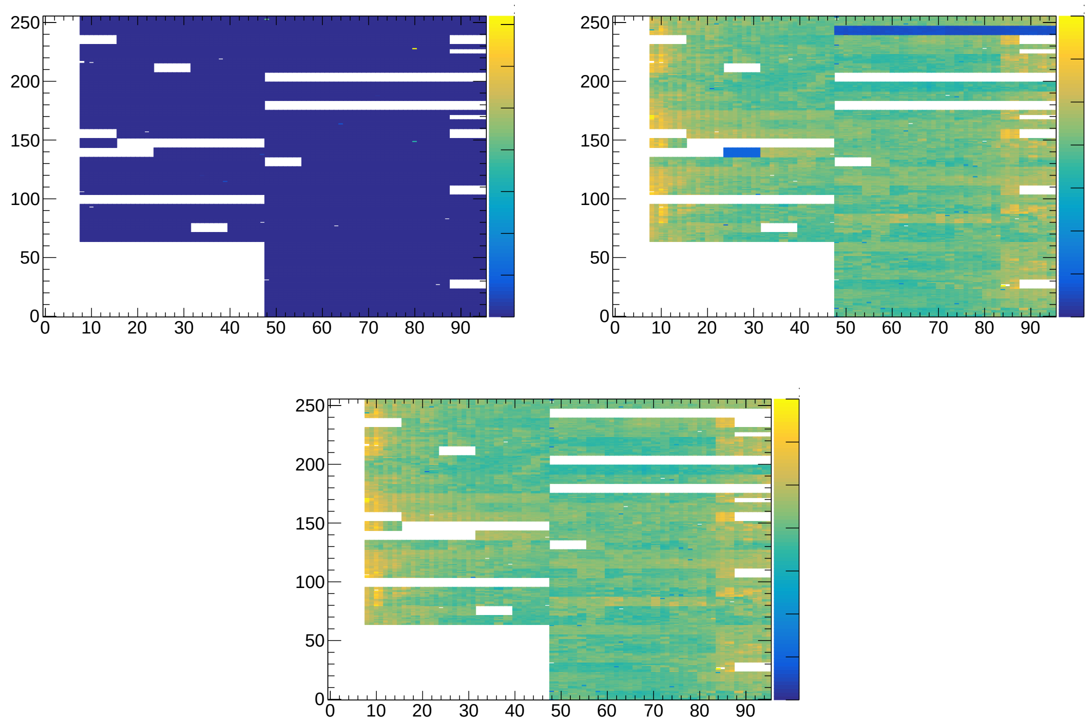

Figure 1: Impact of various stages of the tower masking. Top left is the initial tower masking, top right is the removal of the ten hottest towers, and the bottom is the removal of the cold towers.

Clusterization

Clusterization, carried out by the RawClusterBuilderTemplate module, is the process whereby groups of towers with energy above a set threshold are grouped into extended objects known as clusters. In a similar way that a jet reconstruction algorithm is meant to return a region of contiguous energy from dozens of particles, clusterization is meant to do the same, but for single particles, namely photons and electrons. The sPHENIX EMCal clusterizer is a port of the PHENIX EMCal clusterizer [2]. The minimum tower energy necessary for a tower to be used in clusterization is currently 30MeV, but further studies will be carried out to confirm this is the optimal value.

Cluster Position Dependent Energy Correction

After clusterization, the clusters' energies are corrected to account for two things. The first are the nonuniformities within the towers themselves that can effect their response to the energy embedded by photons and electrons. This is meant to flatten the overall cluster energy response as a function of ϕ and η and is discussed in detail in , and the second is the overall energy scale. Both are handled by firing single photons into the calorimeter in simulation and taking the ratio between the reconstructed cluster and the truth photon from which it originated. In the original study, truth photons are generated with p_T between $20-21$GeV, and the calibration constants are derived from clusters originating from these photons. However, in a study over a test sample with a broader truth photon momentum input, a p_T-dependent non-closure can be seen, such as that seen in Fig. 2. To correct for this, the cluster energy is again divided by the residual, which is flattened in η to account for the fact that the tower-by-tower corrections already account for η-dependent effects.

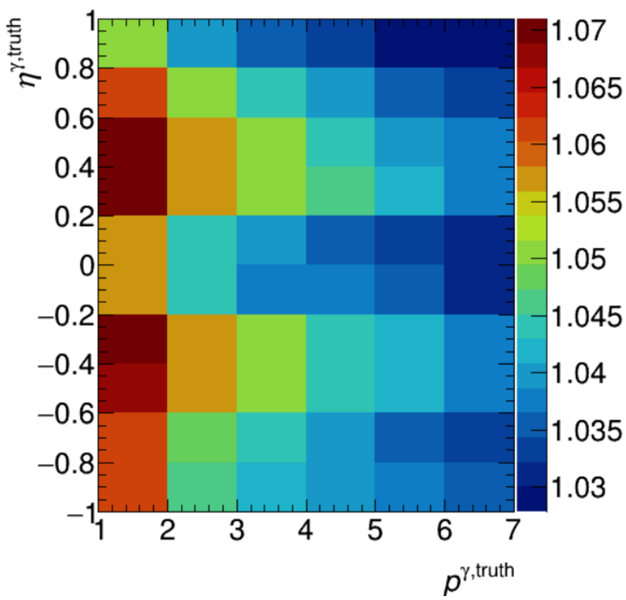

Figure 2: Residual non-uniformity from the position dependent correction

Cluster Masking

Lastly, clusters have within them masked towers are themselves excluded from the cluster pool by RawClusterDeadHotMask. This is because bad towers negatively influence the cluster energy resolution. Hot towers cause a (possibly drastic) overestimation of the cluster energy, and cold/dead towers cause an underestimation. Both also cause a worsening of the cluster spatial kinematic resolution due to a biasing in the calculation of the cluster center of mass.

Cluster Analysis

Cluster analysis and the measurement of the di-cluster invariant mass distribution is done with the Calo_Validation_Package module written by Tim Rinn. It has been modified slightly to look at two different cluster nodes. The first is the "CLUSTERINFO_CEM" node, which contains the unmasked and uncalibrated clusters, and the second is the "CLUSTERINFO_POS_COR_CEMC" which contains the corrected clusters. Invariant mass distributions from the former are shown in black, the latter in red.

The following cuts are imposed on the analysis:

Event-level Cuts

- The reconstructed event z vertex must be within ± 30cm.
- The total charge in the MBD must be less than 12,500 ADC.

The following cuts are imposed at the cluster level:

- $1\text{GeV} < E_{Cluster} < 32\text{GeV}$
- $|\eta_{Cluster}| < 0.7$
- $\chi^2_{Cluster} < 4$

The first cut is to remove clusters which might arise from background noise, and the upperbound cut on energy is a sanity check that would primarily remove clusters whose energy might be set by residual hot towers. The need for the second cut arises from the fact that towers at high rapidity in the EMCal are split between being made by UIUC and Fudan, and have different compositions, and thus responses. Even more challenging, any given high rapidity ring at a given η typically has a mixture of Fudan and Illinois blocks, making completely calibrating out these effects difficult. Thus, the cluster η limits are set to avoid these regions entirely for simplicity.

The cluster χ^2 is a value computed by the clusterizer and is meant to quantify the probability that a given cluster is from a single electromagnetic shower or something else (i.e. from overlapping showers or from hadrons). The lower the this value, the more likely a cluster came from a single electromagnetic shower.

Lastly, the following cuts are made on pairs of clusters:

- $\Delta R > 0.08$
- $\alpha = \frac{|E_1 - E_2|}{E_1 + E_2} < 0.8$
- $p_T^{Reco} > 1\text{GeV}$

Cut one, like the χ^2 is to enforce some amount of separation between the clusters to improve their energy resolution by decreasing the likelihood that they're sharing towers. Cut 2, the asymmetry cut, is meant to reject background pairs, which are more likely to be highly asymmetric, whereas the energy asymmetry between mesonic decay products is flat. And Cut three is simply a floor on the reconstructed meson momentum.

Results

Fig. 3 shows the invariant mass distributions in 4 different meson p_T bins. One can see in each a prominent peak at approximately the π^0 mass. As mentioned previously, the black data points come from the uncalibrated cluster node, and the red come from the calibrated node. The

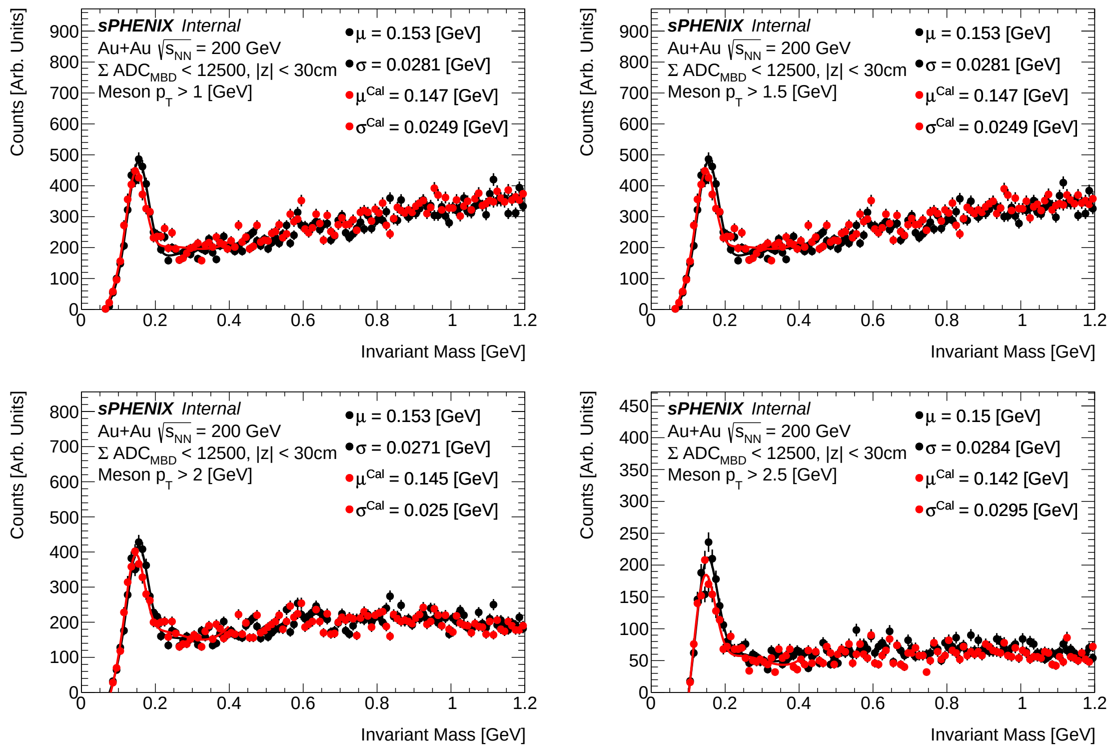

Figure 3: Invariant mass distributions in four different p_T ranges: Top left: $p_T > 1$GeV, top right: $p_T > 1.5$GeV, bottom left: $p_T > 2$GeV, bottom right: $p_T > 2.5$GeV.

difference between the two is just a few percent, but one can see the position dependent correction and residual correction do two things. They first shift the mass peak closer to the PDG value of approximately 135GeV, and on average it improved the resolution. Both the peak position and the width are extracted with a Gaussian fit with extra polynomial terms of the form:

$$c_0 + c_1 x + c_2 x^2 + c_3 x^3 + c_4 e^{(-((c_5 - x)^2)/(2c_6^2))} \tag{1}$$

The c_n terms are allowed to float, and c_5 and c_6 are the peak position and width of the π^0 peak, specifically. no attempt is made to resolve the η peak in these plots, and the fit is constrained from $50 - 400$MeV.

References

[1] Hanpu Jiang and Shuhang Li. HCal Cosmic Muon Calibration Note. 2023. 0.1

[2] G. David, E. Kistenev, S. White, C. Woody, A. Bazilevsky, S. Belikov, V. Kochetkov, V. Onuchin, and A. Usachev. Pattern recognition in the phenix pbsc electromagnetic calorimeter. In *1998 IEEE Nuclear Science Symposium Conference Record. 1998 IEEE Nuclear Science Symposium and Medical Imaging Conference (Cat. No.98CH36255)*, volume 1, pages 520–524 vol.1, 1998. doi:10.1109/NSSMIC.1998.775194. 0.1