Test beam data analysis

Felix Sefkow, Katja Krüger
DESY

E-JADE Analysis meeting
Uni Tokyo, 23.2.2018

Plan

- Analysis strategy
- Basic hardware checks
- Electromagnetic performance
- Hadronic performance
- Algorithms and shower physics

General strategy

- For high level analysis - physics and algorithms - we must ensure that the detector is understood
- “Understood" means that we are able to model it and reproduce its characteristics by means of simulations

- Electromagnetic processes can be modelled with less uncertainty than hadronic showers

- 1st step: tests with noise, LEDs and muons
- 2nd step: tests with electrons
- 3rd step: hadrons
- Only then: higher level studies

Noise

- So-called pedestal data
  - from baseline of the pulse on an oscilloscope
- Method: random triggers
  - in new prototype: triggers by neighbouring cells in same ASIC
- Mean value of pulse height: zero point of energy scale
  - **pedestal** to be subtracted from all amplitudes from now on
  - otherwise ratios between amplitudes, e.g. signal / noise, cannot be formed and calibrations not applied
- Width of pulse height distribution:
  - very small: **dead** channel
  - very large: **noisy** channel
  - look at distribution of widths of distributions to find out what is small and what is large
- Exclude dead and noisy channels from all subsequent steps

LED data

- Low intensity LED light to observe single photo-electron spectra
  - distance between peaks is proportional to **gain** of SiPM
- Due to spread in LED light intensity, need to scan amplitudes of LED calibration pulse voltage (VCALIB) in oder to have useful amplitude in each channel
- Investigate distribution of results and check for outliers
  - bad fits
  - noisy channels
  - SiPM problems
- Optimise procedures and define treatment of **outliers**
  - default values or exclusion - but do not ignore them!

- LED data with larger amplitudes provide **inter-calibration** between low gain and high gain of amplifier
  - switches automatically for each hit, according to signal amplitude
  - in special runs can read both low gain and high gain simultaneously
  - not the same as inter-calibration of physics prototype modes, which had different pulse lengths
  - procedures still under development: look out for surprises

Muon data

- Muons (minimum ionising particles) define the energy scale of each individual read-out cell
- After calibration the most probable value of the MIP pulse height distribution should be 1 by construction
- Easier said than done:
  - for cosmics the pulse height depends on the track length in the cell, i.e. on the incident angle
  - for radio-active sources, on the energy spectrum of the the beta decay and on the trigger condition
  - for beams, there are
    - contaminations by hadrons which induce showers
    - delta rays and secondary particles from the absorber
- For beam data the event selection has to be optimised for statistics versus purity
- Guidance from simulation can help to judge how close the situation is to idealised conditions

Simulation

- Ingredients (detector):
  - geometry
  - material description
  - modelling of electronics effects
- Cross-checks
  - the Monte Carlo should be calibrated: MPV(MIP) = 1
  - the geometry should be checked using event displays
- Ingredients (beam):
  - particle type
  - energy (momentum)
  - material upstream
  - transverse beam profile
    - for muons only affects distribution of hits, for electrons and hadrons also amplitudes

MIP calibration and light yield

- Obtain MIP calibration values (**ADC counts per MIP_{MPV}**)
- Optimise procedure and define treatment of outliers
  - exclude or use default
- Extract light yield = MIP / gain = **no. of pixels per MIP**
  - check for **outliers** again - there should not be any
- If MIP fits do not work, the average LY provides a better guess than the average
  - i.e. $MIP_{default}$ = <LY> * gain better than $MIP_{default}$ = <MIP>
  - why?
- If calibration done, re-run to check **convergence**
  - result may change due to re-calibration of thresholds and impact on track selection
  - do not forget to apply proper calibration and thresholds in event display

- Now we have a tracker. Let us make a calorimeter.

Electrons

- Unfortunately also electron beams are not 100% pure
  - and Cerenkov based particle ID is not 100% efficient
- Scan the events, look for
  - **contaminations** by hadrons and muons
  - additional particles in the beam line
  - soft garbage
- Conceive cuts to suppress unwanted contributions
  - define **fiducial volume**: do not include more cells than necessary for measuring electrons
  - reject events using all available information,
    - topology, outer and rear part
- Verify with simulations that the cuts do not bias response and resolution
  - i.e. there is no effect on pure electrons
  - indirectly select what you want

Electron observables

- All to be compared with **simulations**
- **Response** = energy in units of MIPs = sum over cells in fiducial volume
  - mean as expected? stable in time? independent of impact point?
    - use centre of gravity
  - distribution has no unexpected tails or shoulders?
- More detailed look: **longitudinal profile**
  - sensitive to dead channels and mis-calibrations
  - contaminations
- Even more details: **cell energies**
  - careful! this depends rather strongly on impact point
  - use either tight cuts on c.o.g. or MC with accurately tuned beam profile
- Reproducing cell energy spectra is hard, but on the other hand, if they match, everything else does, too
  - **radial profiles** last
- For a first pass, concentrate on shower centre
  - deviations here spoil everything else
- Apart from material, calibration, beam profile, here **saturation** corrections become relevant
  - and for the new prototype also the **inter-calibration**

Linearity and resolution

- If everything OK up to here, we can analyse performance
- **Linearity**: mean response vs beam energy
- Sensitive to
  - remaining impurities
  - imperfect saturation correction
  - noise and threshold (positive and negative offsets)
- **Resolution**: width of response distribution
- Sensitive to
  - noise (at low energies)
  - material (and electronics) description at intermediate energies
  - mis-calibrations, instabilities, and not properly modelled hardware effects (inhomogeneities) at high energies
  - parameters are inter-correlated

- Now we have a calorimeter and can do physics

Hadrons

- Hadron response depends on **particle type**
  - higher for pions than for protons
    - lower available energy for p due to baryon number conservation
    - anti-p?
- Simulated hadron response is **model-dependent**
  - 5%, locally (profiles) up to 20%
- **Leakage** introduces
  - non-linearity (negative; "saturation")
  - asymmetric response: carefully devise fit and extraction of "the" response
- **Non-compensation** (e/pi > 1) introduces (positive) non-linearity
  - effect is small for AHCAL, but not zero
- In principle: **response** is non-linear
- In principle: **resolution** does not follow $1/\sqrt{E}$ behaviour
- **Shower start** point:
  - distribution to check material and possible problems (contamination, noise)
  - profiles from start allow for more refined tests
- Be aware that cells and regions enter which were not validated with electrons
  - that is why MIPs need to be checked so carefully

Timing Analysis

- still under development, so strategy less well established than for amplitude analysis
- idea is the same, but cannot use LED information anywhere
- procedure:
  - first calibration with muons (instantaneous, small amplitudes, few hits)
  - then cross-check with electrons (instantaneous, larger amplitudes, many hits)
    - watch out for unexpected effects! (e.g. many hits on a chip lead to shift and broadening of time distribution in Spiroc2B)
  - apply to hadrons
- effects we know we need to check/correct
  - non-linearity of TDC ramp
  - time walk
  - dependence on number of hits on a chip

Traps and Pitfalls

- AT vs. ET: very few parameters can be extracted from external trigger runs (LED) and directly applied in auto-trigger data (beam)
  - pedestal is not the same
  - memory cell dependence (?)
  - inter calibration ?
- memory cell dependence
  - it’s there for many parameters (e.g. pedestal), it's a new feature for technological prototype → it’s a new feature in the software
- “features" of the readout ASIC
  - always beware of unexpected effects!
  - in order to minimise number of parameters to determine in calibration, carefully evaluate what depends on
    - chip
    - ramp number (odd or even BXid)
    - channel
    - memory cell

Higher levels

- Software compensation
  - energy
  - topology
- Two-particle separation
  - new Pandora
  - ARBOR
- Electron-pion separation

- Timing analysis
  - shower model validation
  - shower parameters vs time cuts
  - use of timing in particle flow

Technicalities

- all data will be uploaded on the grid, directly reachable from NAF and BIRD cluster at DESY
  - needs DESY computing account (not a problem)
  - alternative from Tokyo? needed?
- running the reconstruction needs the database
  - located on a server at DESY
  - is reachable from restricted number of locations outside DESY
    → check if it works from Tokyo
  - snapshot can be written to file, but then changes cannot be fast and easily made available to everyone

Summary

- A calorimeter is not black magic
- Everything can be checked and understood

- There are lots of interesting physics to come

- Muon and electron test beams form the basis of all studies

- The technological prototype has a number of new features not present in the physics prototype, so software is maybe not yet adequate everywhere
  - memory cell dependence
  - active temperature compensation
  - …