The algorithm for the PWO calorimeter

I.Larin
5x5 VS island

• 5x5 algorithm:
  – 1) Find maximum energy deposition cell
  – 2) declare all 5x5 area around belonging one hit

• Island algorithm:
  – 1) Find maximum energy deposition cell
  – 2) declare all simply connected area around as initial “raw” cluster
  – 3) try to split “raw” cluster into many hits
GAMS island algorithm

*(what is inside, block-scheme with original function names preserved)*

```plaintext
(data_pht)
(prepare integer arrays for involved cells:
cell address: 100*row+column,
cell energy: in [10 MeV] units)

(clus_pht)
(initial cluster search as a simply connected cell areas)

(order_pht)
(sort cell energy array in increasing order)

(loop over initial “raw” clusters:
call gams_pht for each one)

(out_pht)
(cut off low energy clusters)
```
GAMS island algorithm (continued)

Find out all local peaks in the initial “raw” cluster.

In case of many local peaks, calculate which part of energy of each cell belongs to each local peak minimizing $\chi^2$ (gradient maxima search method).

For each local peak call gams_pht (calculates local peak parameters, split it into two adjacent ones if needed).

Reconstruct up to 12 hits per initial raw cluster.

After reconstruction of all hits within initial “raw” cluster, try to process unused part of working cell array (“secondary step”).
GAMS island algorithm (continued)

Gamma_pht
(calculates local peak parameters, split it into two adjacent ones if needed)

Call mom1_pht
(calculates energy $E$; coord. $\langle X \rangle, \langle Y \rangle$ weighted with $E$; correct $E$ by artificial adding a part of shower belonging to cells which are not in cluster list)

Call chisa_pht
(calculates $\chi^2$ using:
\[
\langle e \rangle_{\text{cell}} = E(\text{hit}) \cdot \langle \text{fraction} \rangle_{\text{cell}}_{\text{expected}}
\]
\[
\sigma^2(\text{cell}) = 150 \cdot \langle e \rangle_{\text{cell}} \cdot \{1 - \langle e \rangle_{\text{cell}} / E(\text{hit})\} + 30^2 \text{[MeV}^2]\]
$\sigma$ supposed to be $\langle e \rangle_{\text{cell}} \cdot (1 + 2b / \sqrt{\langle e \rangle_{\text{cell}}})$
a, b are from $\sigma(E)/E = a + b \sqrt{E}$ formula

Moves hit center X and Y to get $\chi^2$ minimum (gradient method)

If resulting $\chi^2$ more than 50 (5 for “secondary step” hit), try to split the hit into two:
call tgamma_pht

In case of $\chi^2$ improvement by more than 6 split the hit. If energy product of two split hits times the distance between them less than cut value, merge it back
Cluster overlapping: different cases

Single cluster

Two clusters do not produce two maxima but can be distinguished

Two clusters can not be distinguished

Two clusters produce two maxima
Cluster overlapping cases

Shower density

Cluster pattern

second maximum:
Corner to corner:
dist. = √2 cells

second maximum:
Side to side:
dist. = 2 cells

No second maximum
Things to be improved:

• cluster elements (which cells belong to which hit) are not stored
• Missed channels (cells with zero signal) are not participating in $\chi^2$ calculation
• Only rough estimation of $\sigma$ for $\chi^2$ calculation
• $\chi^2$ calculation assumes Gaussian distribution of energy deposition in cells. Real distributions are not Gaussian, log-likelihood needs to be implemented

Electron scan data for total energy fraction deposited in cell as a function of $d$ - distance between hit point and cell center:

• There are a lot of hardcoded constants - the code is difficult for transferring
What is inside

• “Phot_cell” function: expected mean energy fraction of total energy deposited in the cell as a function of the distance between hit point and cell center

• Its sigma
Transverse shower profile:

• Integrated energy function:

\[ F(x, y) \equiv \frac{E_{\text{int}}(x, y)}{E_{\text{tot}}} = \int_{-\infty}^{x} dx' \int_{-\infty}^{y} dy' \rho(x', y') \]

• Parameterized shape (Lednev’s parameterization):

\[ F(x, y) = \frac{1}{2\pi} \sum_{i} a_i \cdot f_{\text{cumulative}}\left(\frac{x}{b_i}, \frac{y}{b_i}\right); \text{normalization: } \sum_{i} a_i = 1 \]

\[ f_{\text{cumulative}}(u, v) = \text{Arctg} \frac{u + v}{1 - uv} + \text{Arctg} \frac{uv}{\sqrt{1 + u^2 + v^2}} \]
Examples of set of constants obtained from the PWO electron scan data fit:

<table>
<thead>
<tr>
<th></th>
<th>a1</th>
<th>b1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.815636</td>
<td>0.91497</td>
</tr>
<tr>
<td>b1</td>
<td>0.0887125</td>
<td>0.105129</td>
</tr>
<tr>
<td>a2</td>
<td>-0.78113</td>
<td>-0.068324</td>
</tr>
<tr>
<td>b2</td>
<td>1.06357</td>
<td>894.971</td>
</tr>
<tr>
<td>a3</td>
<td>0.965494</td>
<td>0.153355</td>
</tr>
<tr>
<td>b3</td>
<td>0.729796</td>
<td>0.541627</td>
</tr>
</tbody>
</table>

Usually a2 < 0
Calculating cell energy from cumulative function difference:

\[ f_{cell} = F(x_{cell} + \frac{1}{2}l_{cell}, y_{cell} + \frac{1}{2}l_{cell}) + \]
\[ + F(x_{cell} - \frac{1}{2}l_{cell}, y_{cell} - \frac{1}{2}l_{cell}) - \]
\[ - F(x_{cell} + \frac{1}{2}l_{cell}, y_{cell} - \frac{1}{2}l_{cell}) - \]
\[ - F(x_{cell} - \frac{1}{2}l_{cell}, y_{cell} + \frac{1}{2}l_{cell}) \]
Cell energy fraction VS distance between hit point and cell center: electron scan data ($E_e \approx 4\,\text{GeV}$) and it’s fit

- Central area $\bigcirc$ ($|d| < 0.5$) (flat hat)
- Edge area (e) ($0.5 < |d| < 1$) (high gradient value)
- Tail area (t) ($|d| > 1$) (exponential decreasing)
Monte-Carlo implementation:

- Function of cell energy fraction and its sigma can be used for “fast” Monte-Carlo simulations w/o real shower propagation. This allows to save 3-4 order of magnitude of simulation time.

\[ \langle E_{cell} \rangle = E_{hit} \cdot f_{cell} \]
Other Monte-Carlo options:

• **Real shower propagation:**
  – 4-5 order of magnitude more CPU time
  – but cluster shape closer to real, in particular for irregular areas like holes, shielded cells, corners

• **Photon (and hadronic) shower libraries:**
  – Shape of cluster closer to reality
  – needs some work to create
  – operates with relatively large files and takes relatively large program memory
  – In case if hit momentum is not perpendicular to calorimeter surface needs excessive library volume and still very approximate
  – Probably good solution for hadronic part where resolution is not critical
Non γ-e particles:

- MIPs (2 parameters: mean & sigma)
- Neutral hadrons (neutrons and $K_L$)
- Charged hadrons (= MIP + Neutral hadron),
- $\pi^+$, $\pi^-$, $p$ give probably different showers at low energies
- **Important notice:** $E$-$M$ calorimeter could be used for identification of hadrons to reject background in decays reconstruction like $\eta \rightarrow \pi^+ \pi^- \gamma$ (was successfully implemented in SELEX experiment at Fermilab)
Applying GAMS island algorithm to PWO calorimeter (PrimEx2)

Mixing two events from the snake scan (one cluster is constant and another is moving with the beam)

Distance = 2 cells -> good energy reconstruction
Applying GAMS island algorithm to PWO calorimeter: probability to separate two clusters as a function of distance between hit points

First cluster: “permanent” with energy 5.1 GeV, second ("moving") with energy 1...5 GeV; Events with artificial split counted as missed

Curve a – GAMS cut settings
Curve b – changed cut settings

Decreasing due to artificial split

“transition” zone

No energy dependence for d
Energy resolution:

Original reconstruction (no mixing)

$\sigma = 2.6\%$

d = 4.3cm ... 4.7cm

$\sigma = 2.6\%$

d = 2.3cm ... 2.7cm

$\sigma = 4.1\%$

d = 1.3cm ... 1.7cm

$\sigma = 8.1\%$
• Separation threshold should be selected to not disturb clusters much
• GAMS used separation threshold to distinguish pi0s
• Algorithm tune up is in progress