XIV International Conference on Calorimetry in High Energy Physics (CALOR 2010) IOP Publishing
Journal of Physics: Conference Series **293** (2011) 012024 doi:10.1088/1742-6596/293/1/012024

Fast Simulation of Showers in the H1 SpaCal Calorimeter

Sasha Glazov
DESY, Notkestrasse 85, Hamburg, 22607 Germany
E-mail: `glazov@mail.desy.de`

Abstract. A method for the fast simulation of particle showers in the H1 lead/scintillating-fiber calorimeter is presented. The method uses a shower library technique in which the detector response is simulated based on a collection of stored showers for different particle types. The library is created using the GEANT programme. The fast simulation is compared to the data collected by the H1 experiment.

Simulation of high energy physics experiments requires large computational resources. A significant fraction of time is typically spent for the simulation of showers in calorimeters. Different methods have been developed to accelerate the simulation of showers. These include parametrisation methods [1], shower libraries [2] and modified shower libraries, so called frozen showers [3]. This note describes a “shower library” based simulation of the H1 calorimeter, more details of this method can be found in [4].

A ”shower library” is a collection of GEANT [5] energies in the calorimeter for a given particle type, energy and location at the calorimeter surface. The structure of the shower library depends on the construction of the calorimeter and particle kinematics. The shower library provides description of the detector response approaching the quality of the GEANT simulation while the computing time is much reduced.

In the H1 detector [6], which is dedicated to the study of the deep inelastic scattering process using colliding electron and proton beams of 25.5 GeV and 920 GeV, respectively, from the HERA accelerator in Hamburg, the shower library is implemented for the lead/scintillating-fibre calorimeter (so called ”Spaghetti calorimeter” - SpaCal) [7]. The SpaCal calorimeter contains electromagnetic and hadronic sections. The electromagnetic section of the SpaCal consists of 1192 cells with an active volume of $4.05 \times 4.05 \times 25$ cm^3 each. A transverse view of the calorimeter is given in figure 1. The cells are made of grooved lead plates and scintillating fibers with a diameter of 0.5 mm. The scintillation light of each cell is converted into an electric pulse using photomultiplier tubes. The active length of the electromagnetic SpaCal corresponds to 27.5 radiation lengths and about one hadronic interaction length. The Moliere radius of the SpaCal is 2.5 cm. The angular coverage of the calorimeter, which is measured with respect to the proton beam direction[1], is $153^\circ < \theta < 177.8^\circ$. The electromagnetic energy resolution can be parametrised as $7.5\%/\sqrt{E} \oplus 2\%$. The spatial resolution of the calorimeter in the transverse plane is about 3.5 mm. The amount of material from the interaction point to the calorimeter

---

[1] The z axis of the right-handed coordinate system used by H1 is defined to lie along the direction of the incident proton beam and the origin to be at the nominal ep interaction vertex.

Published under licence by IOP Publishing Ltd

surface varies between 0.6 and 1.0 radiation length.

The hadronic part of the SpaCal is made of 136 cells of $12 \times 12 \times 25$ cm^3 providing one nuclear interaction length. The fibers are of the same type as in the electromagnetic section but have a larger diameter of 1 mm.

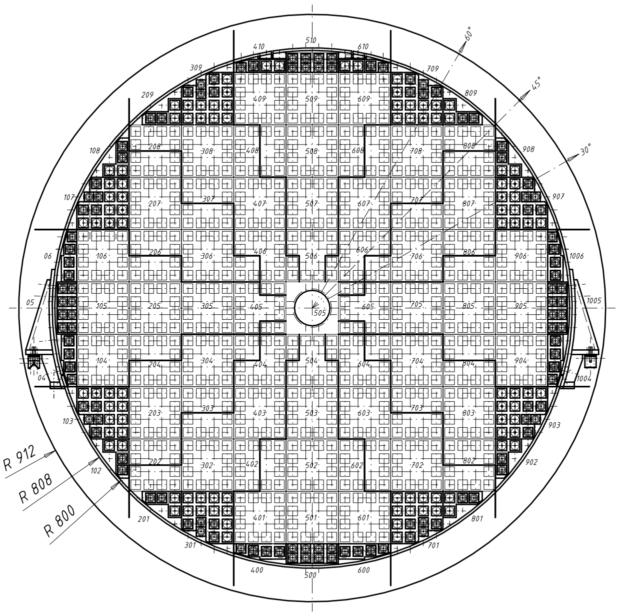


Figure 1. Transverse view of the SpaCal. Small boxes indicate individual cells. They are combined in 4 x 4 groups into super-modules.

In the simulation using the shower library a generated particle is traced through the detector components up to the calorimeter surface. At this point, if the impact position of the particle is far enough from the calorimeter boundaries, such that the shower is expected to be fully contained in the calorimeter, the shower library is applied: instead of using the GEANT package to simulate the calorimeter response, a suitable pre-simulated shower is selected from the shower library. The selected shower is then corrected for a difference between the actual particle kinematics and the kinematics of the particle used for the shower creation. Corrected energy deposits are added to the calorimeter response. The quality of the simulation based on the shower library therefore depends on how well it represents kinematics of simulated particles, how many showers are stored for each grid point and how large residual corrections are. Since the computer memory is limited, an additional important feature of the shower library is the packing scheme.

The shower library grid is binned in the particle energy, E, in the particle position with respect to the calorimeter surface, x and y, and in the particle direction angles, θ_x and θ_y, with respect to the lepton beam direction which is perpendicular to the calorimeter surface. The binning in energy is spaced uniformly in $\log E$, it should be fine enough compared to the varioation of the energy resolution as a function of energy. The grid should start from low energy, $\simeq$ 100 MeV, to describe showers starting before the calorimeter surface. The granularity of the position binning parallel to the calorimeter surface should be fine enough compared to the calorimeter spacial resolution. Symmetries of the calorimeter can be used to reduce the number of position bins. A reduction is possible for a large (vs shower size) calorimeter with translation symmetry. In this case it is sufficient to use binning inside the cell with the largest energy deposit ("hottest cell"). The granularity of the angular binning usually does not need to be too fine, if the angles are sufficiently small. A small angle $\theta_{x,y}$ leads to only a moderate (proportional to $\cos\theta_{x,y}$) increase of the shower size which is usually not very important. A larger effect is introduced by an effective shift of the shower x, y-position. This effect can be compensated for using an effective shower depth as described in [4].

The shower library is created using the GEANT simulation for particles selected with energies, positions and angles as defined by the shower library grid. For a given grid point, typically several showers are stored.

For simulation using the shower library, the bin which best fits the particle kinematics is selected. The energy bin is selected using the following algorithm:

(i) For a given particle energy, E_{gen}, find the nearest higher and lower energy grid points, E_h and E_l ($E_l < E_{gen} \leq E_h$).

(ii) With a probability proportional to the "distance" from a grid point to E_{gen}, randomly select an E_h or E_l bin. The "distance" can be measured linearly or logarithmically. For the logarithmic measure, which is used for the SpaCal shower library, the probability to select the lower grid point is:

$$P_l = \frac{\log \frac{E_h}{E_{gen}}}{\log \frac{E_h}{E_l}} \tag{1}$$

(iii) For the selected bin (h or l), take the shower from the shower library and, before adding to the calorimeter response, re-scale the energies of deposits as $E_{gen}/E_{l,h}$ ($E_{l,h}$ stands for E_l or E_h).

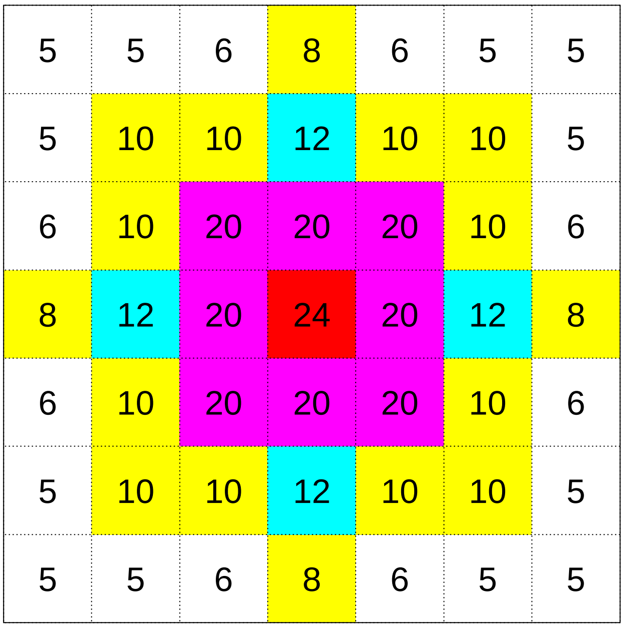
This procedure adequately reproduces the variation of the energy resolution as a function of energy correctly up to first order in $\log E_h/E_l$.

To obtain a good description of the calorimeter response and avoid statistical biases, the shower library should contain sufficient amount of showers. In order to avoid relatively slow disk access it is preferential to have the showers needed for all grid points buffered in the computer physical memory. In the case of multidimensional binning (in x, y, θ_x, θ_y, E and particle type) and fine calorimeter granularity (hundreds of energy deposits per shower), the total memory requirement can be rather high.

A usual strategy is to use buffers of packed showers. One buffer covers one complete shower library which covers all grid points and stores at least one shower per grid point. Several buffers are stored in a disk file. The showers from one buffer reside in memory, and are reused several times; a new buffer is read from the disk file after the repetition count exceeds some fixed value.

Electromagnetic showers are typically small in size and they can be stored using energy deposits in the cells in a fixed box around the hottest cell. To pack shower information effectively, it is usual to store the total energy of the shower as an uncompressed real number; the energy deposits are stored in terms of a bit packed fraction of the total energy. The fraction of energy of the hottest cell should be stored with the highest precision (e.g. with 24 bits) while for peripheral cells, just few bits can be sufficient. This packing scheme is illustrated in figure 2. For broad, asymmetric hadronic showers, many cells may contain no energy deposit. In this case, it is more efficient to store information only from the cells which contain energy.

The shower library for electromagnetic particles contains 12 bins in energy ranging logarithmically from 0.1 GeV up to 32 GeV, 10 bins in x and 10 bins in y which cover the front face of the hottest cell, 8 bins in θ_x and 8 bins in θ_y ranging from -24.8° to 24.8° and two



5	5	6	8	6	5	5
5	10	10	12	10	10	5
6	10	20	20	20	10	6
8	12	20	24	20	12	8
6	10	20	20	20	10	6
5	10	10	12	10	10	5
5	5	6	8	6	5	5

Figure 2. Bit packing scheme for showers from the electromagnetic SpaCal shower libraries. The showers are stored using a fixed 7 × 7 cells template around the shower centre. Fraction of the shower energy contained in the cell at the shower centre is stored using 24 bits, while for outer cells 5 bits are used.

bins for particle type, for electron/positron[2] and photon. The total number of grid bins is

$$12 \cdot 10 \cdot 10 \cdot 8 \cdot 8 \cdot 2 = 153600,$$

which corresponds to one buffer of the shower library.

The shower library for hadrons contains 10 bins in energy from 0.1 GeV up to 20 GeV, 5 bins in x and 5 bins in y which cover the front face of the hottest cell, 4 bins in θ_x and 4 bins in θ_y from $-12.1°$ to $12.1°$ and 9 bins in particle type. The total number of grid bins is

$$10 \cdot 5 \cdot 5 \cdot 4 \cdot 4 \cdot 9 = 36000,$$

which corresponds to one buffer of the shower library.

The shower library for electromagnetic particles is built of showers containing the energy response of 49 cells from the electromagnetic section and 25 cells from the hadronic section. The shower library for hadrons is built of showers containing the energy responses from all cells in the electromagnetic and hadronic sections (1328) with non-zero energy. The packing of the energy information reduces the size for one buffer to about 8 MB.

The shower library for electromagnetic particles is used on average 30 times for a single electron from the interaction point. This is caused by the shower development in the material in front of the calorimeter. The repetition count for reading a new buffer is set to the total number of shower in the buffer, 153600, thus one buffer is used to simulate about 5000 electrons. There

---

[2] A detailed study of the showers produced by positrons and electrons shows that they are virtually identical; therefore no additional binning distinguishing electrons and positrons is introduced.

are in total 7 buffers used for the SpaCal simulation. Since the shower development up to the SpaCal is random and the resulting SpaCal shower is build from many showers from the shower library, the number of showers in the library is sufficient for the H1 analyses which study up to $\sim 10^7$ scattered electrons.

The main quantity used for particle identification in the SpaCal is the transverse size of the shower parametrised by a cluster radius. The determination of the cluster radius starts with the calculation of the cluster centre-of-gravity which is defined as a weighted average over the positions of the cluster cell centres:

$$X_{cluster} = \sum_i x_i w_i \qquad Y_{cluster} = \sum_i y_i w_i. \tag{2}$$

Here x_i, y_i are the x and y coordinates of the i-th cell centre and w_i defines the weight of cell i. Two definitions are used in H1 to calculate w_i, based on square root and logarithmic energy weighting:

$$w_{i,\mathrm{sqrt}} = \frac{\sqrt{E_i}}{\sum_j \sqrt{E_j}}, \tag{3}$$

$$w_{i,\log} = \frac{\max(0, w_{\mathrm{cut}} + \log(E_i/E_{\mathrm{cluster}}))}{\sum_j \max(0, w_{\mathrm{cut}} + \log(E_j/E_{\mathrm{cluster}}))}, \tag{4}$$

where E_i corresponds to the energy reconstructed in the cell i, and w_{cut} defines the logarithmic cut-off parameter for the SpaCal taken to be 4.85.

The cluster radius is then calculated as a weighted sum over the distances, R_i, between the centre of each cell, i, and the cluster centre-of-gravity. Corresponding to the two definitions of weighting, two cluster radius calculations are employed at H1. These are, so called,

$$ECRA = \sum_i R_i w_{i,\mathrm{sqrt}}, \tag{5}$$

and the logarithmic cluster radius,

$$R_{log} = \sqrt{\sum_i (R_i w_{i,\log})^2}. \tag{6}$$

SpaCal clusters corresponding to the scattered positrons, from e^+p interactions, are compared between the H1 data and simulation. For this comparison data collected by the H1 collaboration in the year 2000 are used. For these data, positrons with energy of 27.5 GeV collided with protons with energy of 920 GeV. The simulation of e^+p collisions uses the DJANGO event generator [8].

A pure sample of scattered positron clusters is obtained by requiring the energy of the cluster to exceed 15 GeV; this avoids hadronic background contamination. The selected cluster is required to be in the region of the SpaCal far from the inner and outer acceptance edges.

Figures 3 and 4 show comparisons between the H1 data and Monte Carlo simulation using the shower parametrisation [1] and between data and Monte Carlo simulation using the shower library. The comparisons are made for the R_{log} and ECRA estimators. Using the shower library a good description of the cluster radius is achieved.

Both simulation methods, using the shower library and the shower parametrisation, are about ten times faster than the full GEANT simulation for the scattered positron energies in the tested range.

An implementation of the shower library for the simulation of the H1 detector is presented. The software to generate, store and read back the shower library information is described. The shower library is created for the lead/scintillating-fibre calorimeter SpaCal.

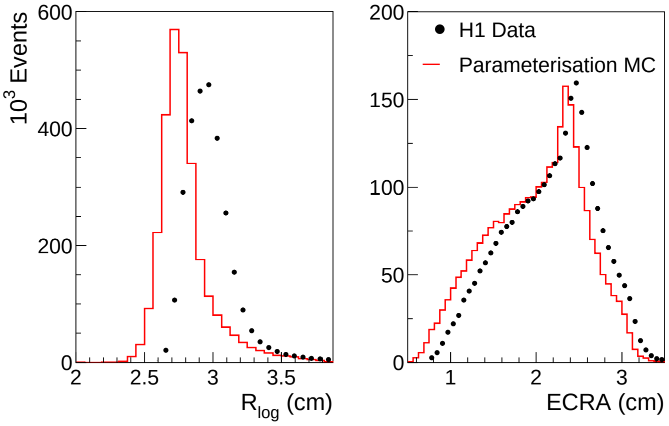


Figure 3. Comparison of H1 data (dots) with simulation using the shower parametrisation (histogram) for R_{log} (left) and ECRA (right).

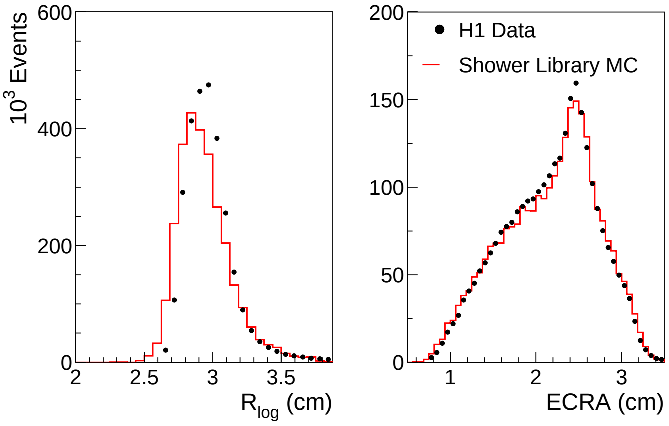


Figure 4. Comparison of H1 data (dots) with Monte Carlo simulation using the shower library (histogram) for R_{log} (left) and ECRA (right).

Monte Carlo simulation using the shower library provides good description of the cluster shapes observed in data. This reduces systematic uncertainty from electron identification efficiency which is important for the measurement of the structure function F_L [9]. The CPU time is reduced compared to the full GEANT simulation by about a factor of ten.

[1] G. Grindhammer, M. Rudowicz and S. Peters, Nucl. Inst. Methods A290 (1990) 469; G. Grindhammer and S. Peters, hep-ex/0001020.
[2] E. Turner Worcester, PhD Thesis, University of Chicago, Chicago, (2007).
[3] E. Barberio et al., J. Phys.: Conf. Ser. 160 (2009) 012082.
[4] A. Glazov, N. Raicevic, A. Zhokin, Comp.Phys.Comm, 181, 1008 (2010).
[5] R. Brun et al., GEANT3, CERN DD/EE/84-1 (1987); Application Software Group, GEANT Detector Description and Simulation Tool, CERN Program Library Long Writeup W5013, October (1994).
[6] I. Abt *et al.*, H1 Collaboration, Nucl. Instr. and Meth. A 386, 310 (1997); H1 Collaboration, Nucl. Instr. and Meth. A 386, 348 (1997).
[7] R.-D. Appuhn *et al.*, H1 SPACAL Group, Nucl. Instr. Meth. A 386 (1997) 397.
[8] G. A. Schuler and H. Spiesberger, Proc. Workshop on HERA Physics, Vol 3, eds. W. Buchmuller and G. Ingelman, Hamburg, DESY (1992), 1419; A. Kwiatkowski, H. Spiesberger and H. - J. Mohring, Comp. Phys. Comm. 69 (1992) 155.
[9] F.D. Aaron *et al.*, H1 Collaboration, Phys. Lett. B 665 (2008), 139.