Fast Parameterisation of electromagnetic showers in the LAr calorimeter: Optimisation of the Frozen Shower Libraries

Matteo Agostini
DESY Summer Student Program 2007

University of Padua (IT)

September 14, 2007

Outline

Outline

Libraries Features

A Frozen Shower library store typically 1000 shower templates organised in:

- 10 energy bins (from 1 MeV to 1 GeV)
- twenty-five η bins (EMB, EMEC and FCAL η values)

Libraries Features

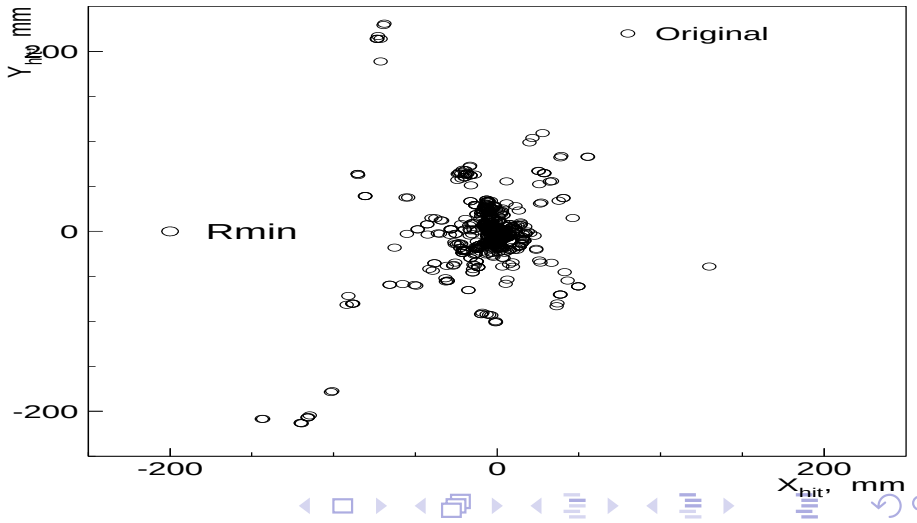
A Frozen Shower library store typically 1000 shower templates organised in:

- 10 energy bins (from 1 MeV to 1 GeV)
- twenty-five η bins (EMB, EMEC and FCAL η values)

A template is a collection of hits.
It is created by merging hits of full simulated low energy electron.

merging process:

- the two closest hits are identified
- if the distance is smaller than a defined distance R_{max} the pair is replaced by a new hit at the center of energy;



Libraries Features

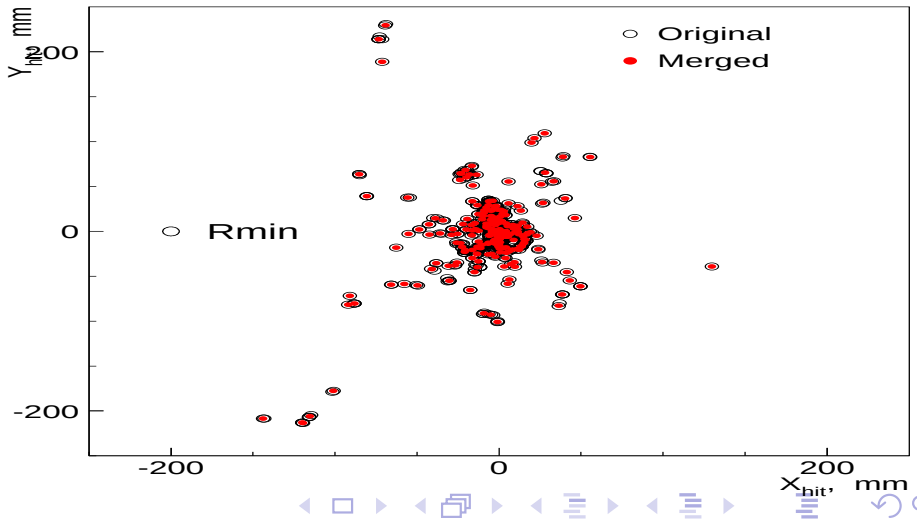
A Frozen Shower library store typically 1000 shower templates organised in:

- 10 energy bins (from 1 MeV to 1 GeV)
- twenty-five η bins (EMB, EMEC and FCAL η values)

A template is a collection of hits.
It is created by merging hits of full simulated low energy electron.

merging process:

- the two closest hits are identified
- if the distance is smaller than a defined distance R_{max} the pair is replaced by a new hit at the center of energy;



Outline

Improvement of the clustering algorithm

Old algorithm:

- the distance between hits is calculated in each step
  $\Rightarrow$ the computing time grows as N_{hits}^3

New algorithm:

- sorted lists are used to identify the closest pair of hits
- when a new hit is created the next closest pair of hits is calculated approximately in a linear computing time
  $\Rightarrow$ the computing time grows as N_{hits}^2 in the worst case!

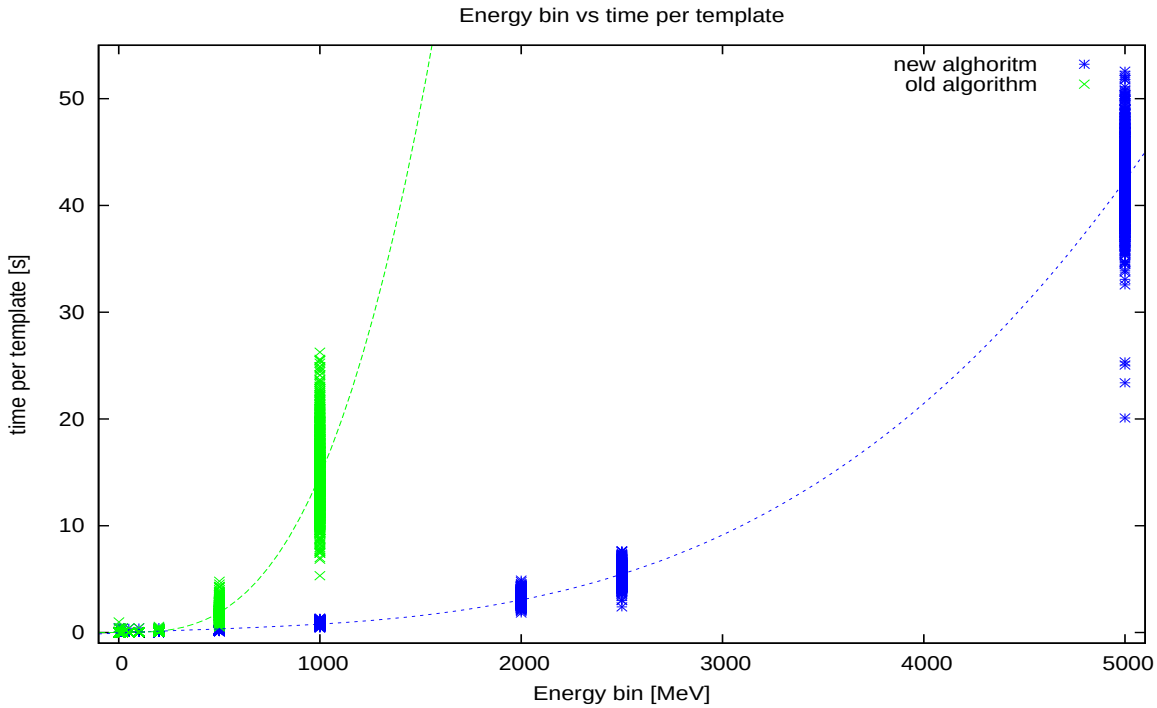
To create a default library:

- the old alg takes $\backsim$ 20 h: 8 to simulate the events + **12** to create the lib
- the new alg takes $\backsim$ 10 h: 8 to simulate the events + **2** to create the lib

$\Rightarrow$ the new algorithm is **6** times faster!

Perfomance of the new algorithm

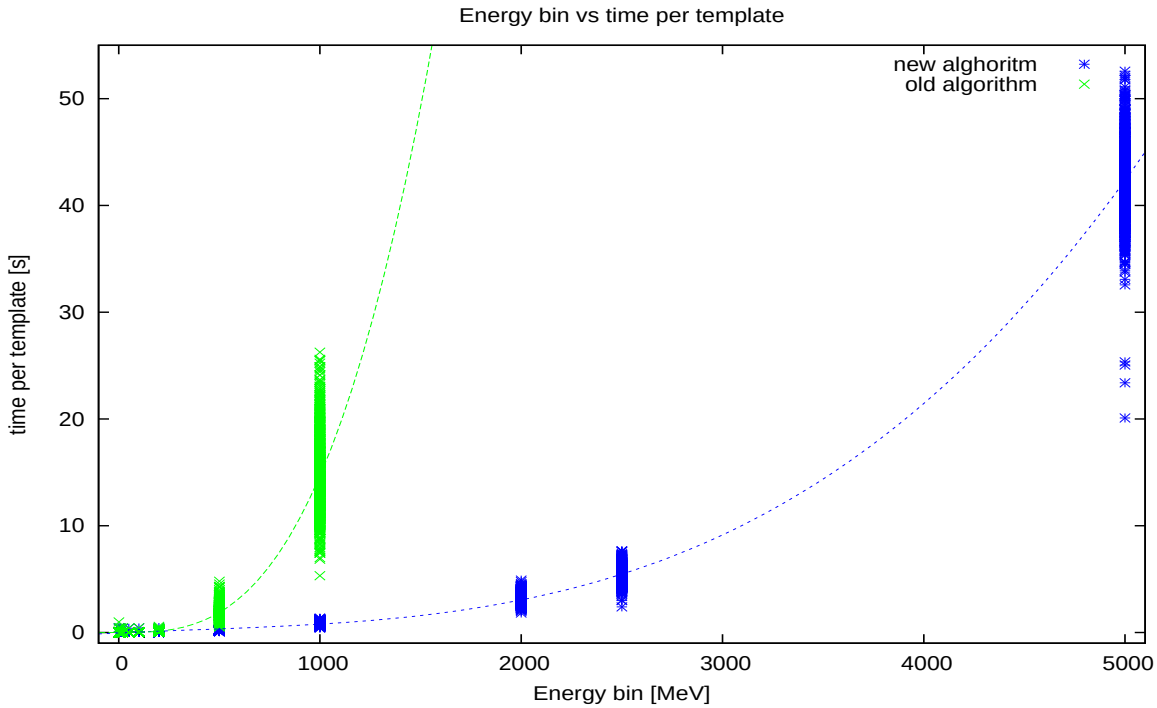
Time needed to generate and record a template vs energy bin



- to generate 1 GeV energy bin with new algorithm is factor $\backsim$ 20 faster
- FS libraries created with two different algorithms are the same

Perfomance of the new algorithm

Time needed to generate and record a template vs energy bin



- to generate 1 GeV energy bin with new algorithm is factor $\backsim$ 20 faster
- FS libraries created with two different algorithms are the same

New viable studies:

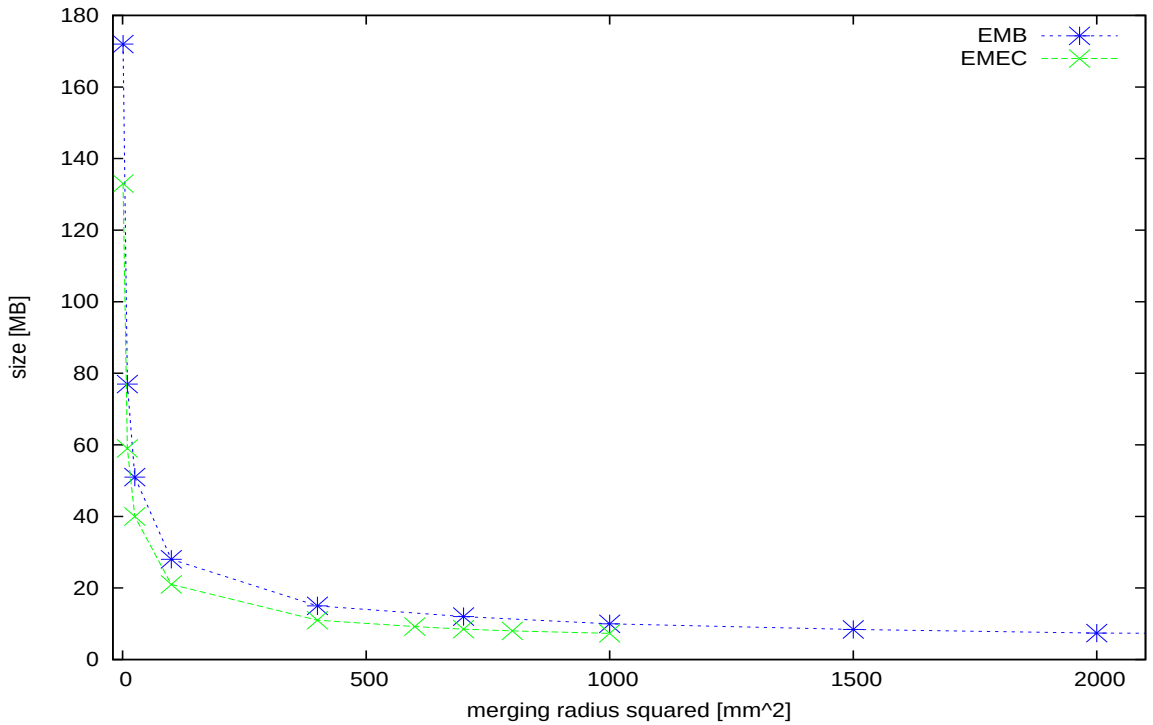
- R_{max} optimisation;
- libraries extended to higher energies

Outline

Optimisation of the clustering radius

Purpose:

- to reduce the libraries size
  $\Rightarrow$ to speed up computing time by decreasing the total number of the hits;
- to not sacrify the accuracy of the shower description



- the defualt $R_{max} = 5$ mm is equals to the size of the strips in the S1 (in η units)
- $R_{max} = 5$ mm $\Rightarrow$ default library size $\backsim 50MB$
- as far as the library size is concerned increasing R_{max} above 20 mm it doesn't bring much

The important quantities

In order to check the shower profile and the energy distribution three quantities are used:

- deposited energy [MeV]
- $\Delta\phi = R_{xy} \frac{\sum_i E_i(\phi_i - \phi_0)}{E_{tot}}$ [mm]
- $\Delta\eta = R_{xy} \frac{\sum_i E_i(\eta_i - \eta_0)}{E_{tot}}$ [mm]

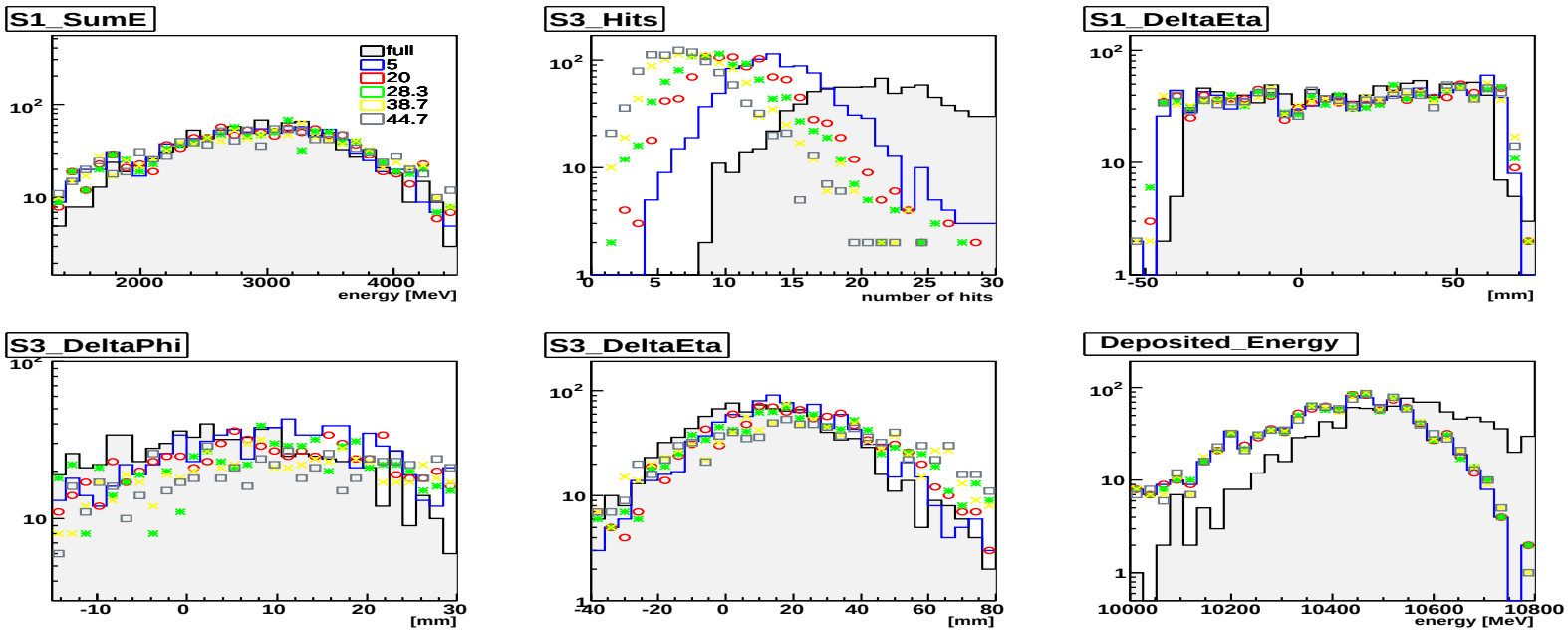
where ϕ_0 and η_0 are ϕ and η of the initial particle and the sum is above all the hits produced in the calorimeter

$R_{xy} \backsim 1.5$ in the LAr Calorimeter

Optimisation of the clustering radius

Purpose:

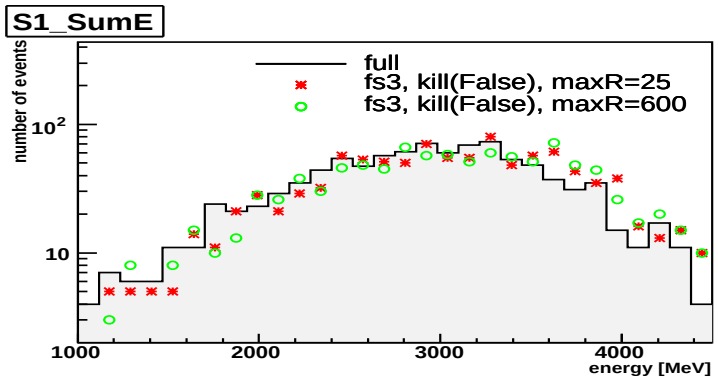
- to reduce the libraries size
  $\Rightarrow$ to speed up computing time by decreasing the total number of the hits;
- to not sacrify the accuracy of the shower description

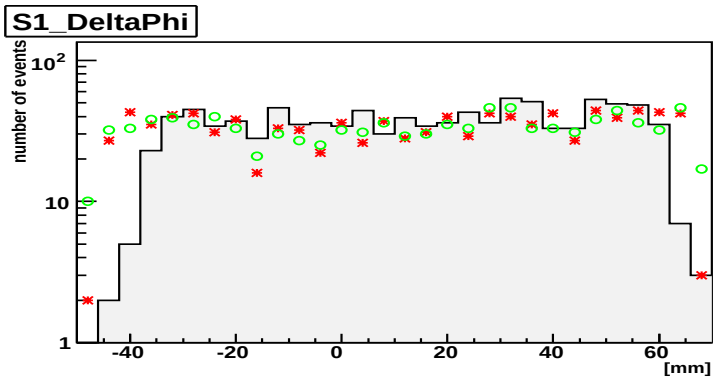


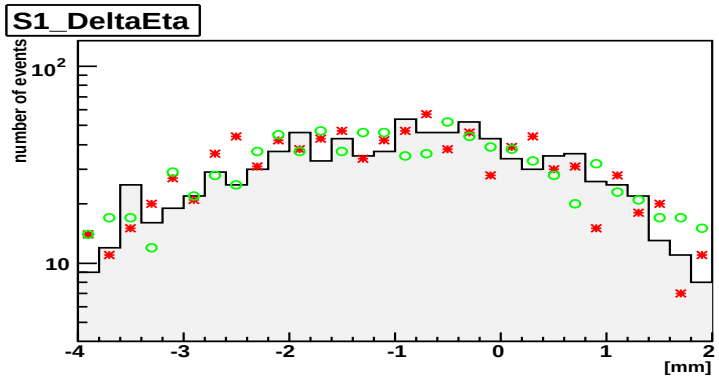
- all the series below 28.3 mm appear to be very similar
- a reasonable compromise could be $R_{max} = 24.5$ mm

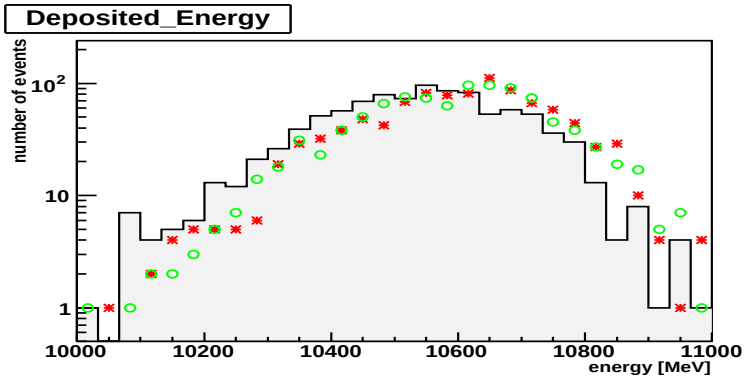
Distributions for different cluster radius R_{max}

events generated using electron and photon libraries with $R_{max} = 5$ mm and $R_{max} = 24.5$ mm







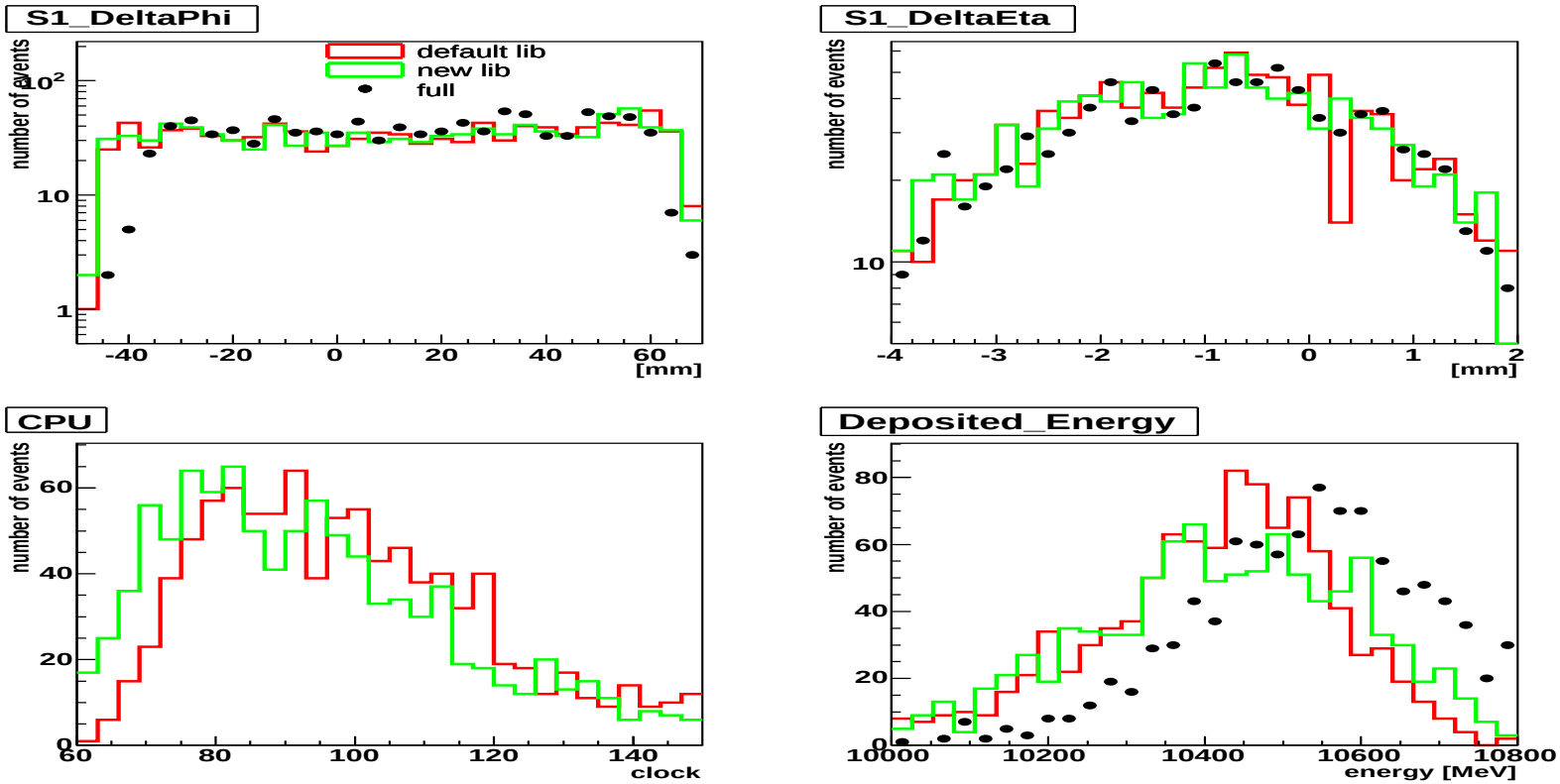


- the distributions appears to be very close
- there is also a small improvement in computing time with $R_{max} = 24.5$ mm
- The size of the new library corresponds to 20% of the old one

Outline

FS libraries extended to higher energy

events generated using the default electrons library and the new library with an additional energy bin at 2000 MeV



- the distributions appears to be very close
- The improvement in time is $\backsim$ 12% ($0.692 \pm 0.009 \rightarrow 0.614 \pm 0.009$)
- increase in library size by a factor $\backsim$ 2

Summary

- the new clustering algorithm speeds up the FS libraries generation $\Rightarrow$ new kind of studies viable
- on-going studies in order to optimize the merging radius (EMB and EMEC)
- higher energies libraries could be an important further possibility to speed up the simultion.