Calorimeter simulation, channel recovery, clustering

Introductory talk

G.Mavromanolakis, University of Cambridge

Outline

- ▶ introduction
- ▶ G3-G4 calorimeter simulation
  - ▷ highlights
- ▶ dead channel recovery
  - ▷ motivation
- ▶ clustering algorithm
  - ▷ introduction
- ▶ summary – future planning

Introducing GM

▶ **previously**

: on the development and design of a tungsten/quartz fiber **c**alorimeter for the **ALICE** experiment at the CERN-LHC

▶ **since 030801**

: **CALICE** Collaboration

▶ **in this talk**

: highlight some first results

: outline fields of interest and work

: feedback is welcome

G3-G4 calorimeter simulation

▶ **.**

: work with the latest versions of BRAHMS(v3.01) and MOKKA(v2.0) porting calorimeter geometry from G4

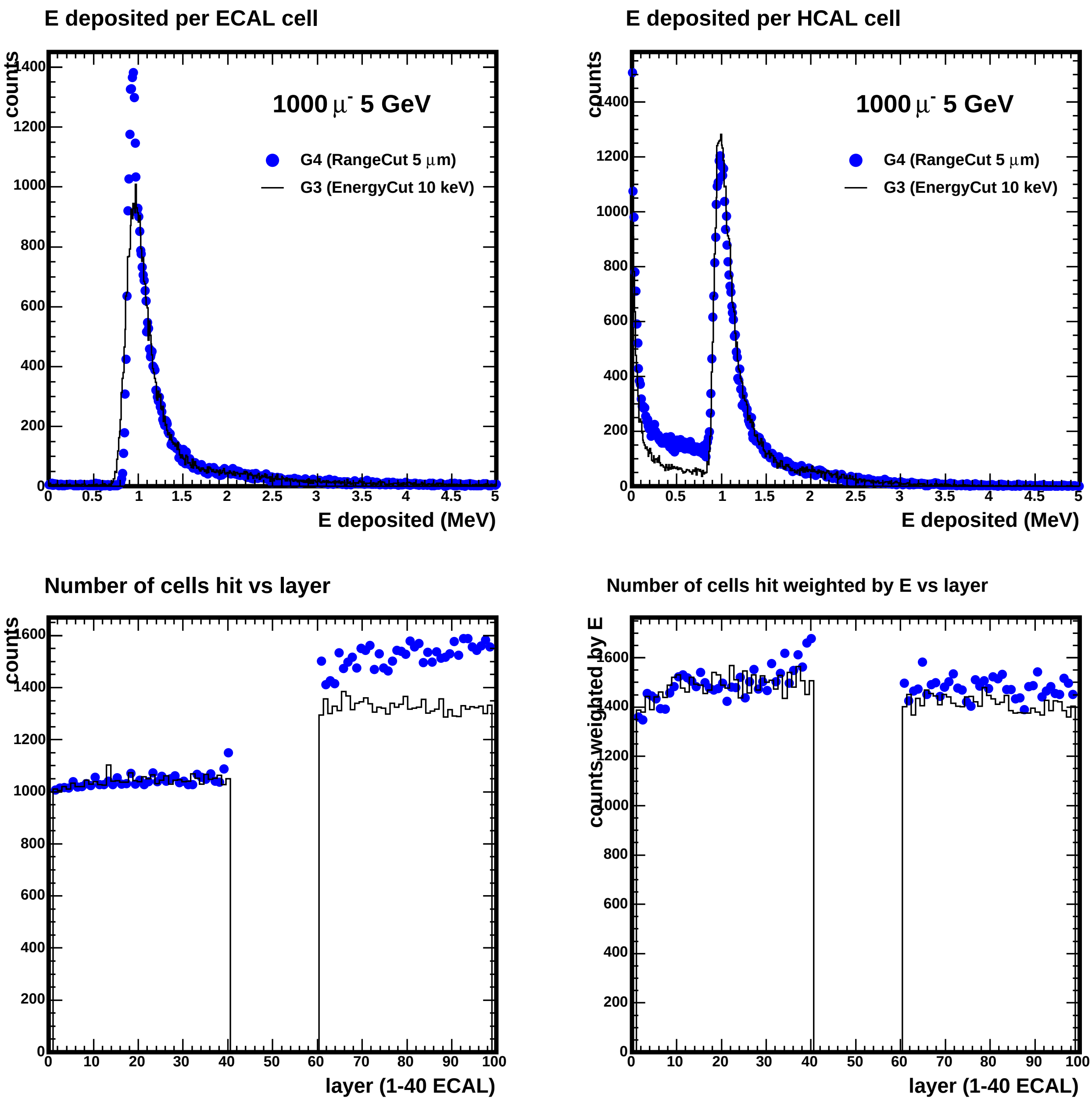
: analysis code in ROOT and C++

▶ **.**

: reproduce older results, cross-check

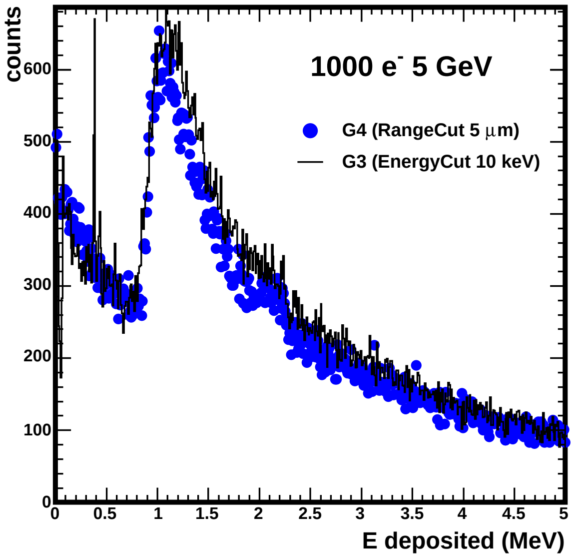
: develop Event class - tree, data in ascii grow fast, possibly develop ”root2hbook” or ”root2txt” backward utility

G3-G4 comparison

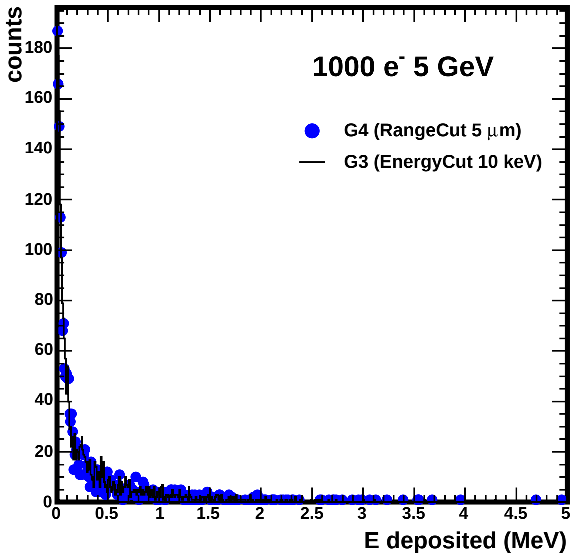


G3-G4 comparison

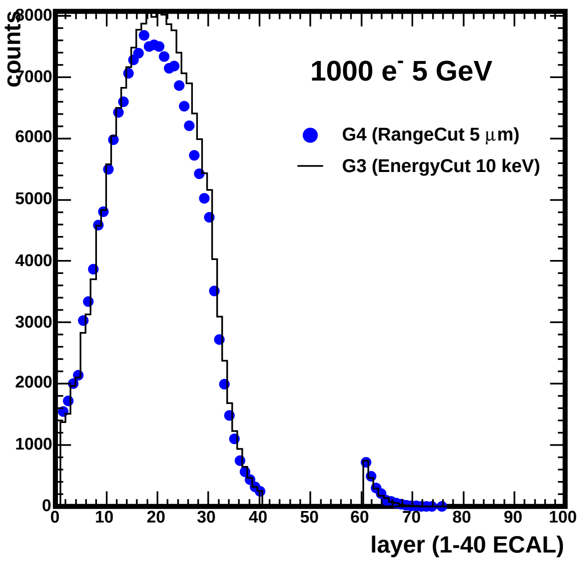
E deposited per ECAL cell



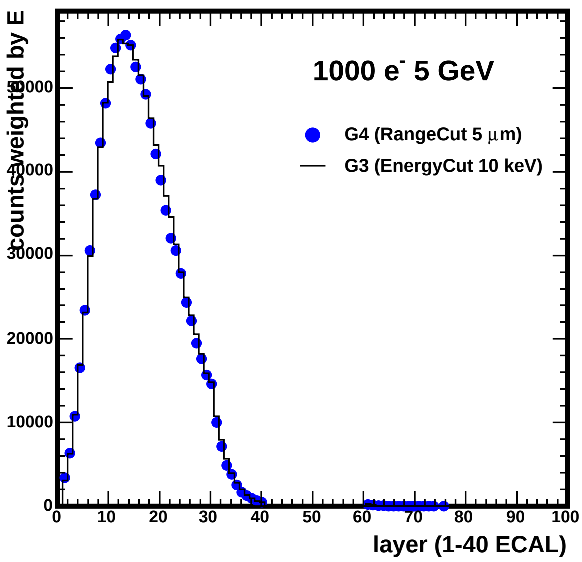
E deposited per HCAL cell



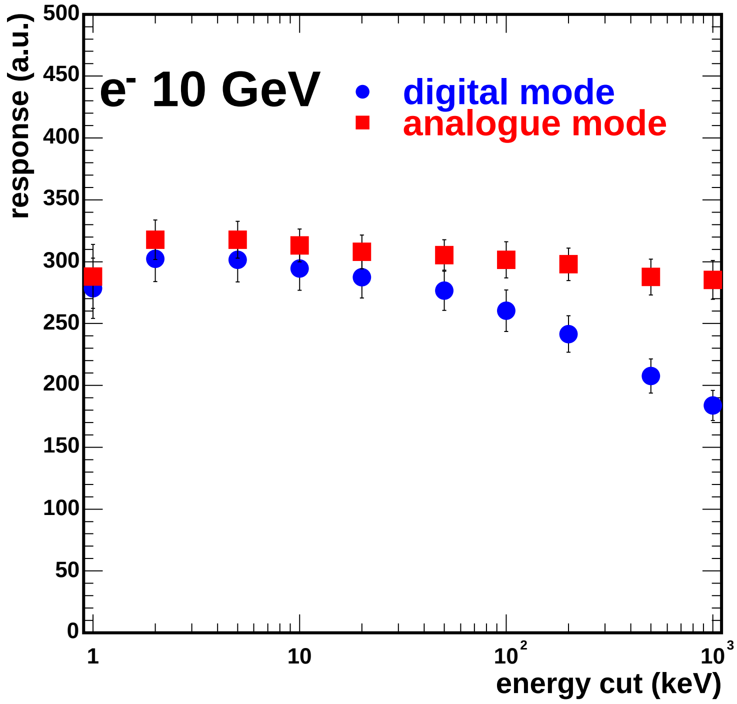
Number of cells hit vs layer



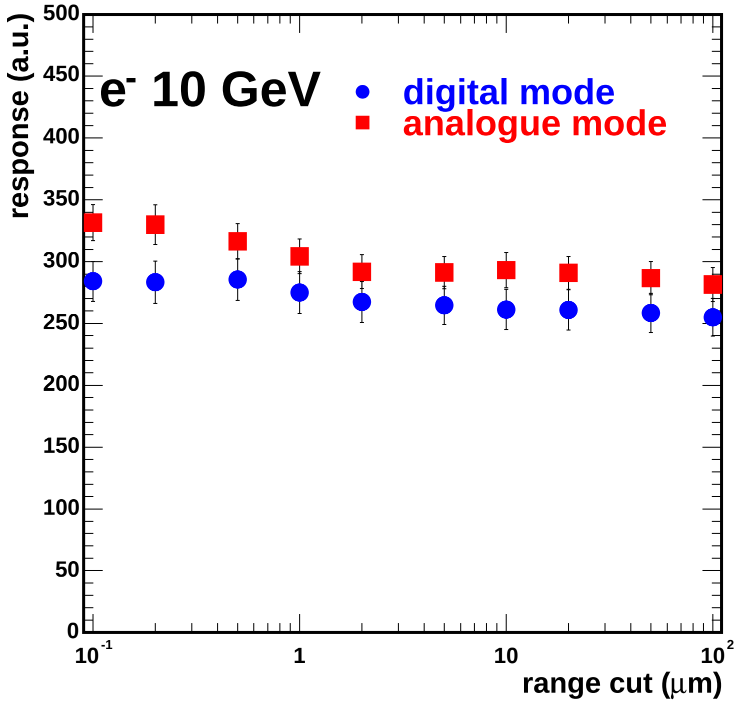
Number of cells hit weighted by E vs layer



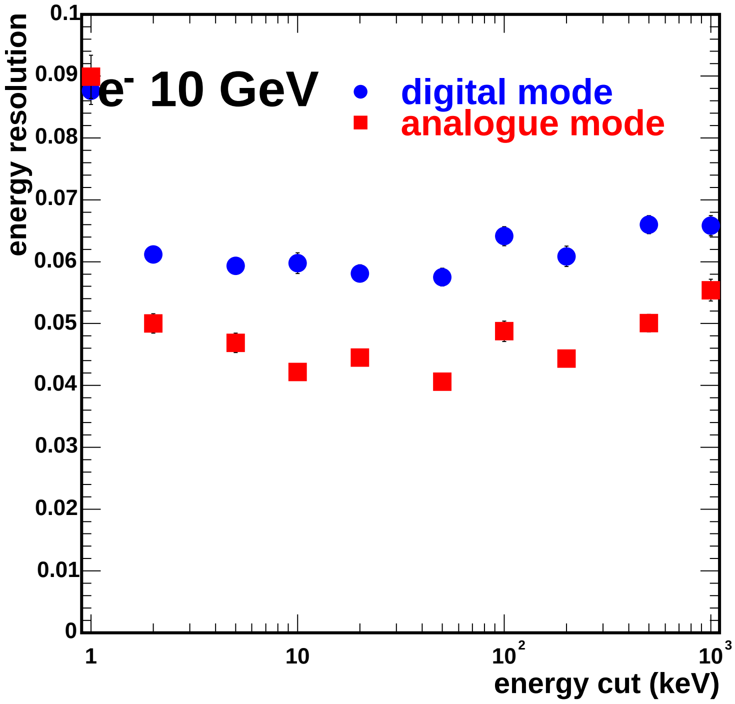
response vs G3 control parameter (energy cut)



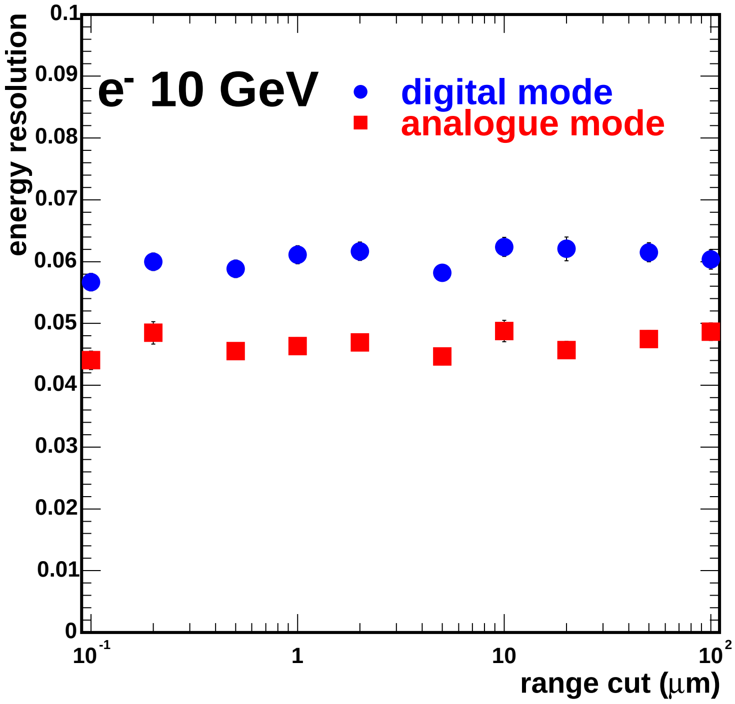
response vs G4 control parameter (range cut)



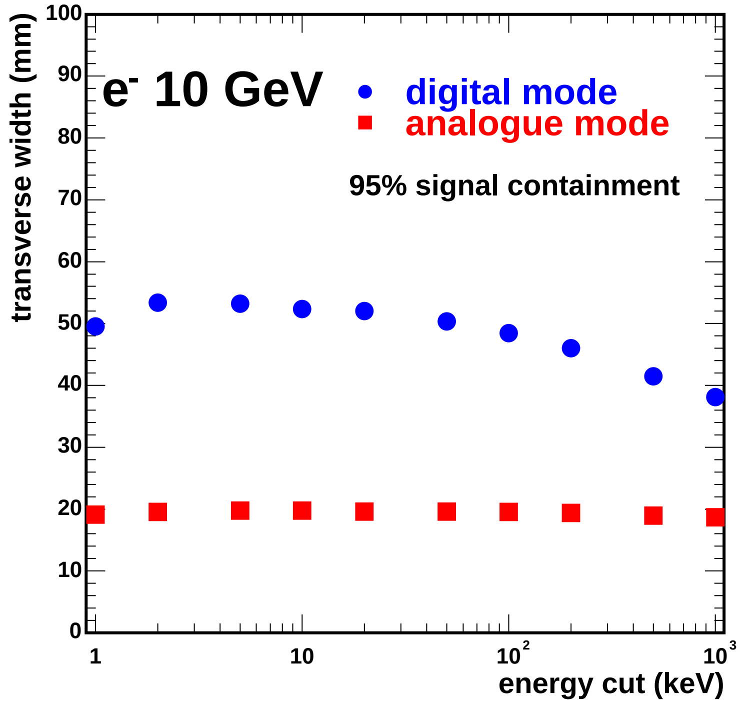
resolution vs G3 control parameter (energy cut)



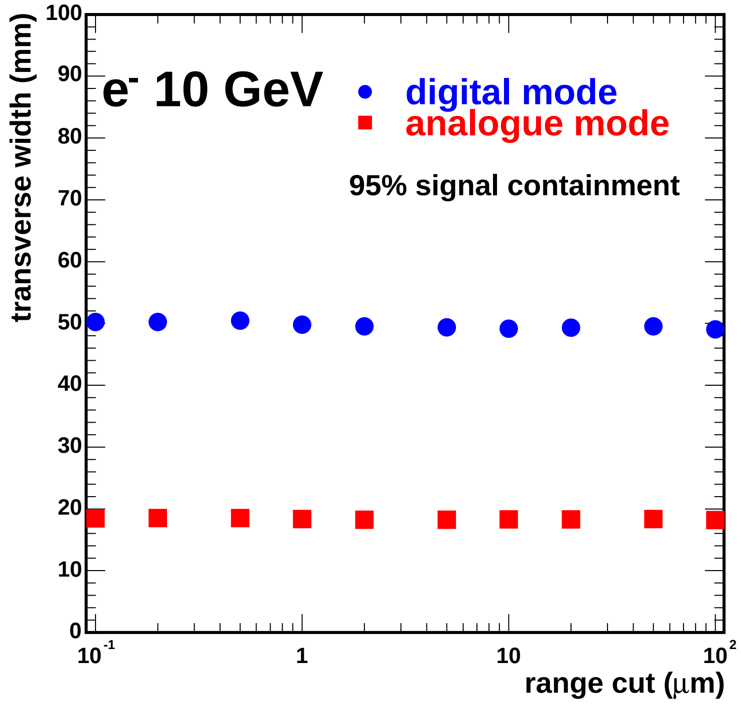
resolution vs G4 control parameter (range cut)



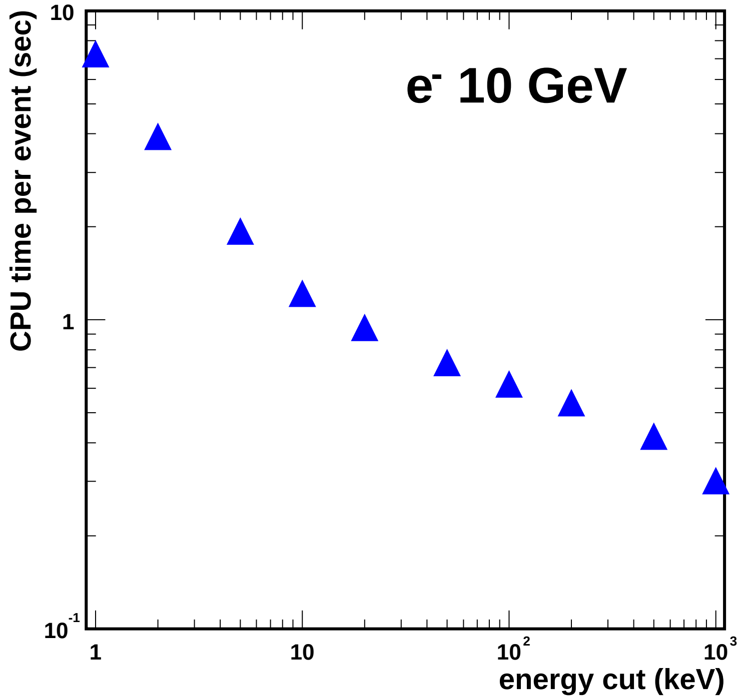
width vs G3 control parameter (energy cut)



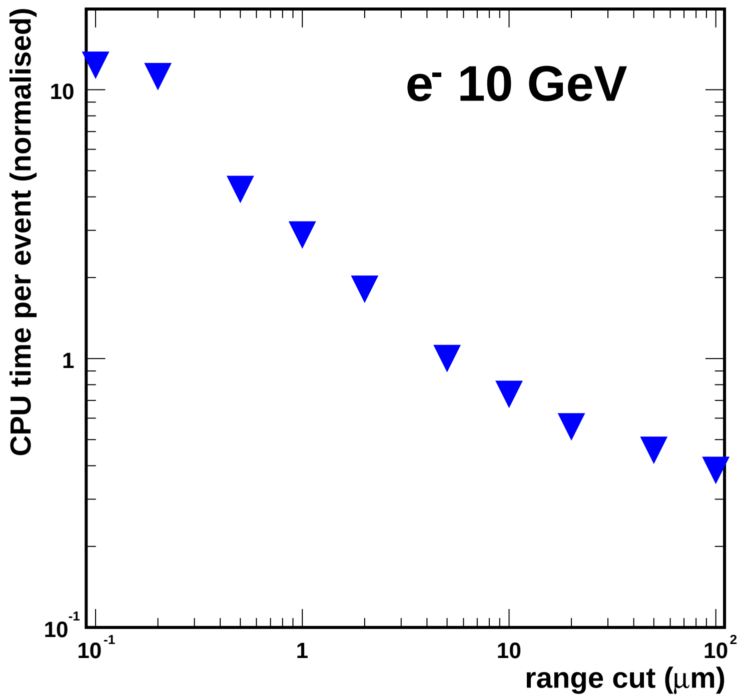
width vs G4 control parameter (range cut)

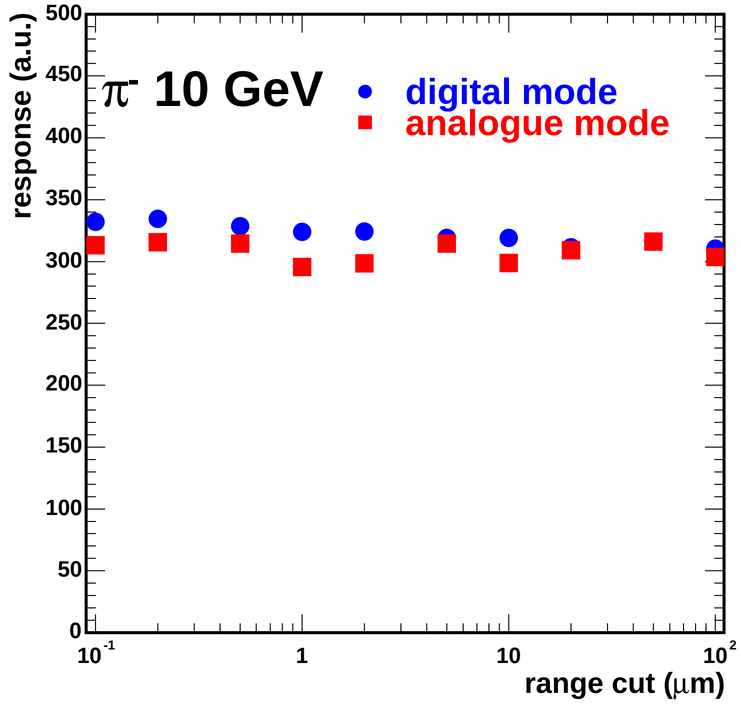


CPU time vs G3 control parameter (energy cut)

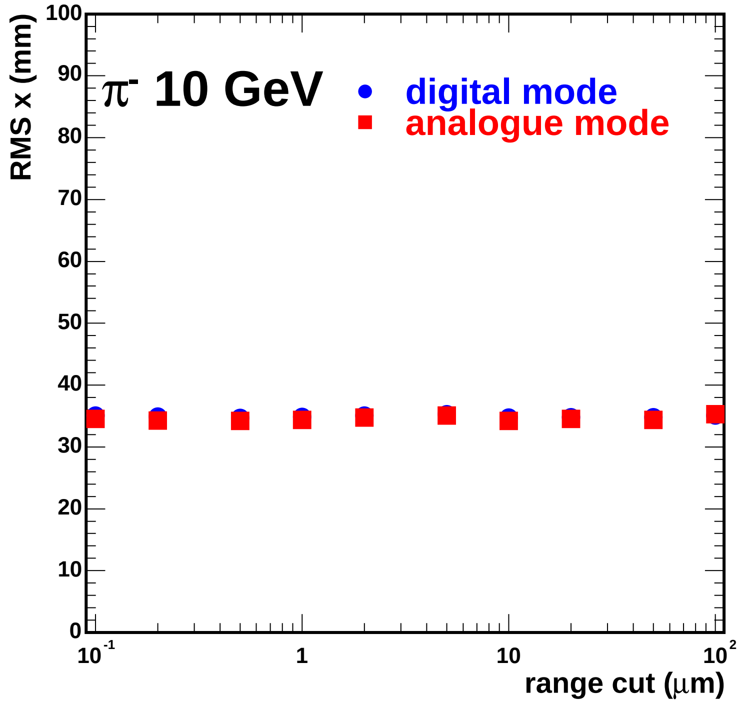


CPU time vs G4 control parameter (range cut)





response (a.u.)
π⁻ 10 GeV
digital mode
analogue mode
range cut (μm)



RMS x (mm)
π⁻ 10 GeV
digital mode
analogue mode
range cut (μm)

Dead channel recovery

▶ **fact**

: production cost considerations may favour dead channel fraction around 5%

▶ **questions**

: how this affects performance

: can we compensate - recover dead channels

▶ **study - solution**

: study detector performance vs dead channel fraction

: develop dead channel recovery scheme

besides conventional methods, potential application of neural network based techniques

: evaluate recovery efficiency and performance improvement

Illustration of the problem

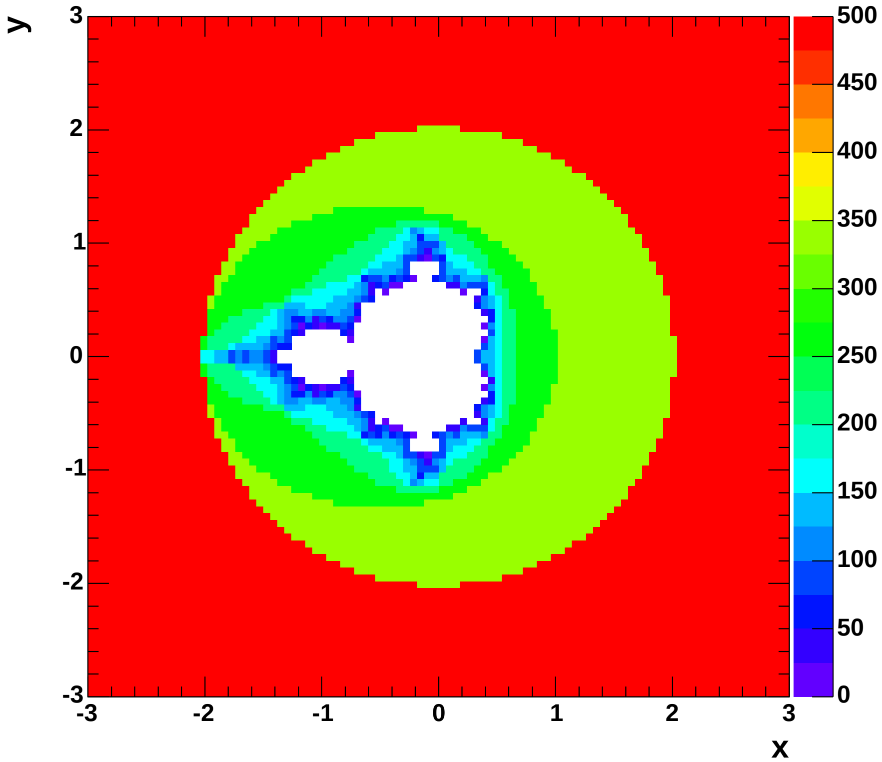


Figure 1: original picture, 10000 channels

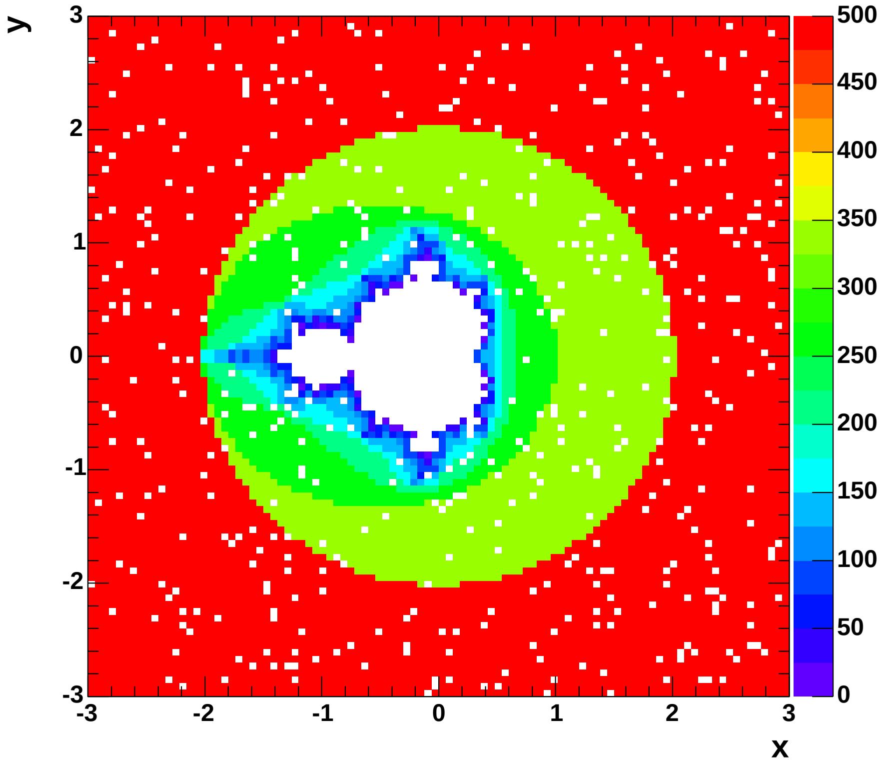


Figure 2: picture with 5% dead channels

Illustration of the solution

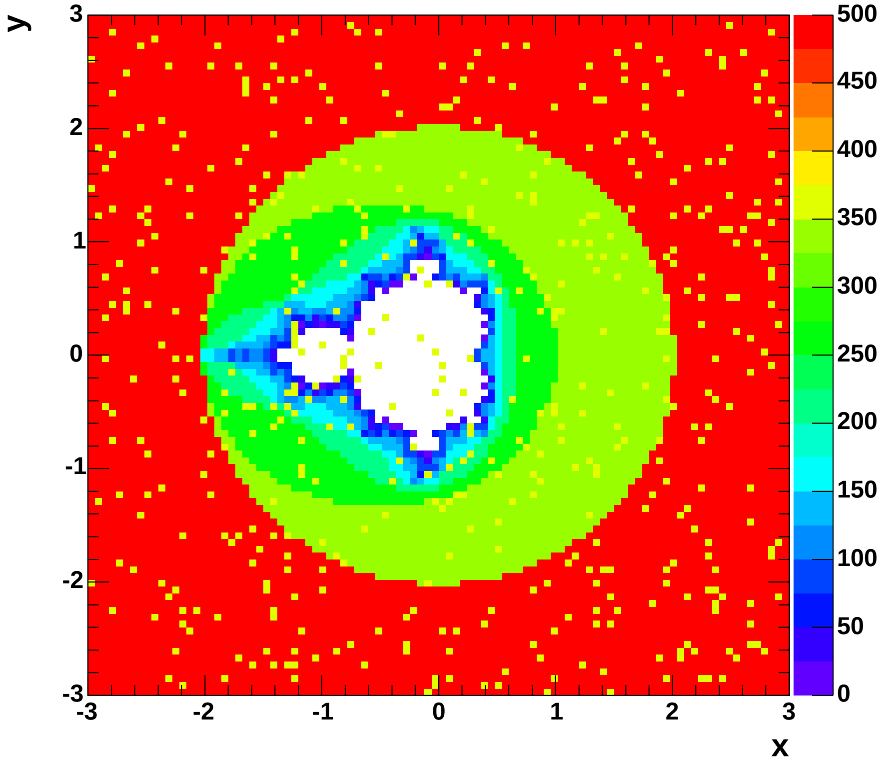


Figure 3: "recovery scheme" step 1

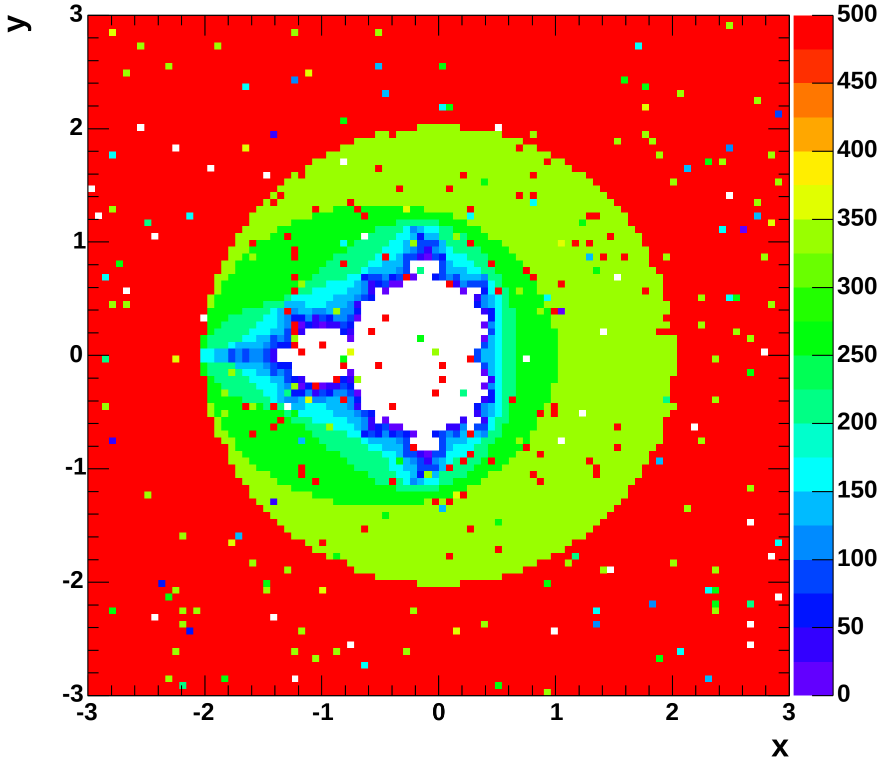


Figure 4: "recovery scheme" step 2

Illustration of the solution

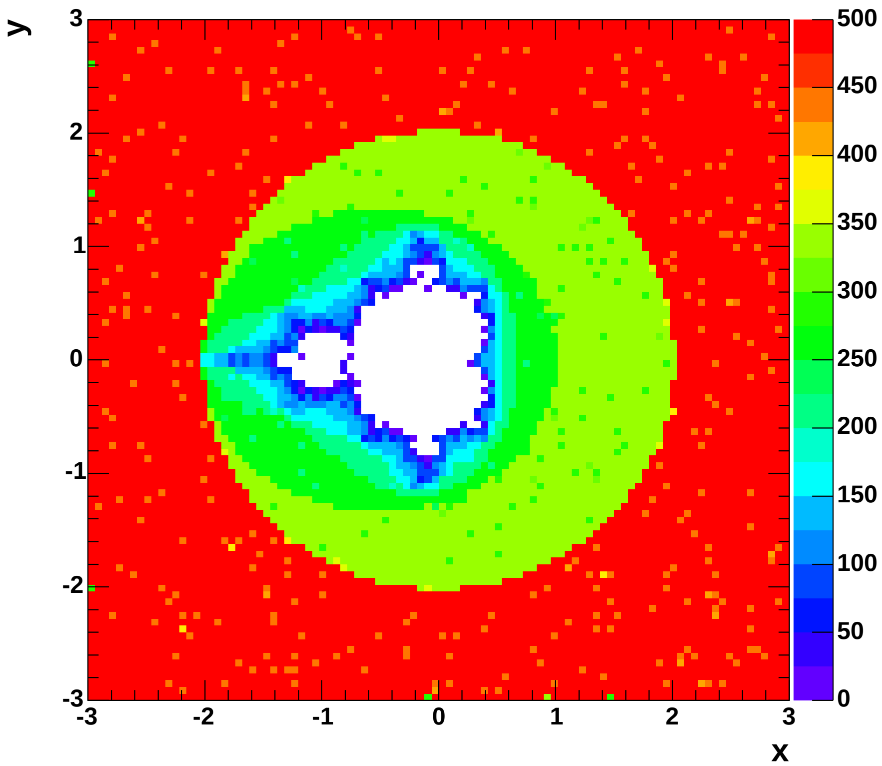


Figure 5: “recovery scheme” step 3

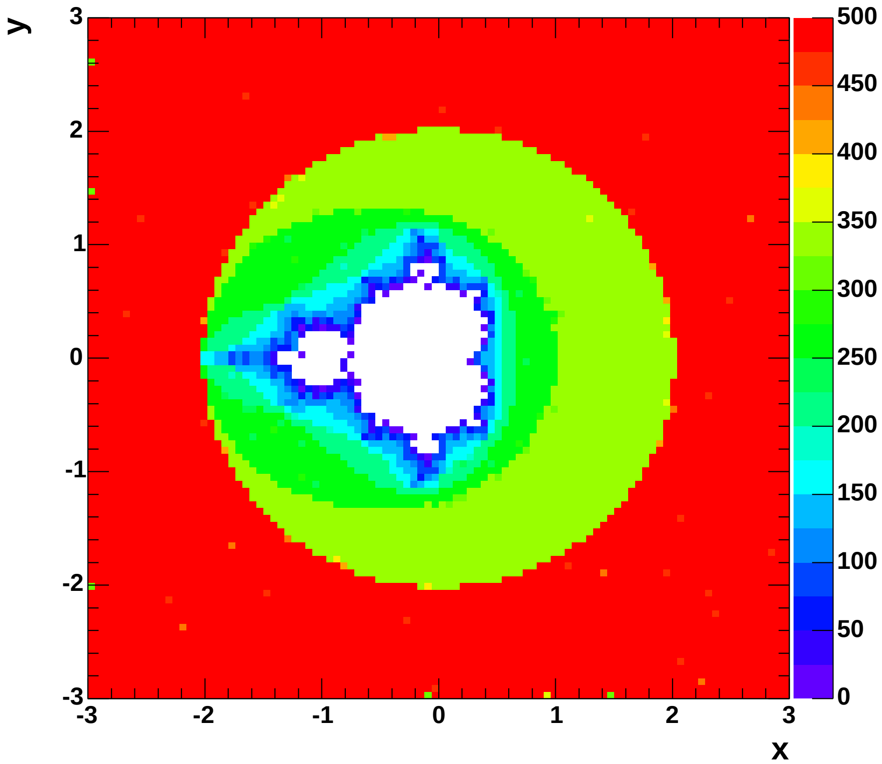
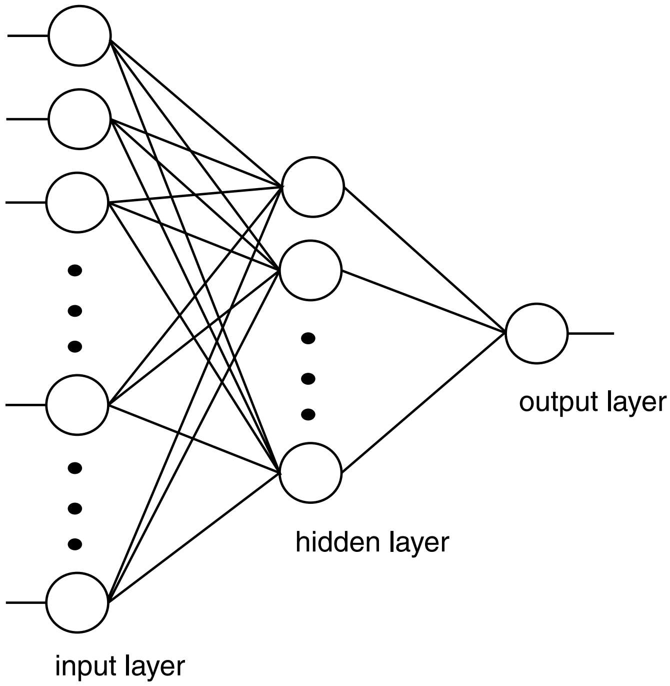


Figure 6: recovered picture



- I_i^k : total input to neuron i of layer k

$$I_i^k = \sum_{j=1}^{N^{k-1}} w_{ji}^{k-1} O_j^{k-1} \quad , \quad k > 1 \tag{1}$$

- O_i^k : neuron output, w_{ji}^{k-1} : connection weight with neuron j of layer $(k-1)$

$$O_i^k = f(I_i^k + w_{i0}^k) \quad , \quad k > 1 \tag{2}$$

- $f(x) \equiv$ *sigmoid activation function*, e.g. $f(x) = \frac{1}{1+e^{-x/T}}$

- number of free parameters to be optimized

$$N_{weights} = \sum_{k=2}^{H+2} (N^k + N^k N^{k-1}) \tag{3}$$

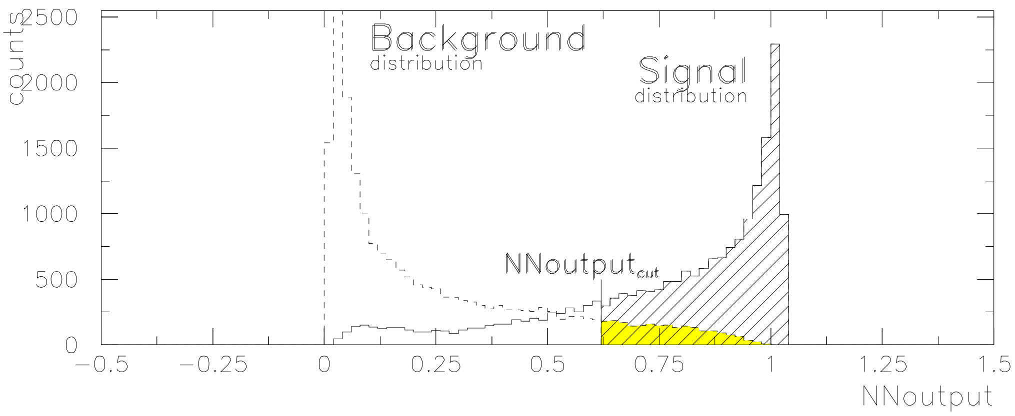


Figure 7: typical signal and background distribution as a function of NNoutput. T he hatched area contains the $N_{signalNN}$ signal events that are above the selection cut NNoutput_{cut}. The colored area contains the contaminating $N_{signallikeNN}$ background event s.

$$\textit{signal efficiency: } \epsilon_s = \frac{N_{signalNN}}{N_{signal}} \tag{4}$$

$$\textit{contamination: } \epsilon_b = \frac{N_{signallikeNN}}{N_{background}} \tag{5}$$

$$\textit{signal enhancement} = \frac{\epsilon_s}{\epsilon_b} \tag{6}$$

$$(S/B)_{NN} = \frac{N_{signalNN}}{N_{signallikeNN}} = \frac{\epsilon_s}{\epsilon_b} \cdot \frac{N_{signal}}{N_{background}} = \frac{\epsilon_s}{\epsilon_b} \cdot S/B \tag{7}$$

Clustering algorithm

▶ minimal spanning tree

: a tree which contains all nodes with no circuits and of which the sum of weights of its edges is minimum

▶ properties

: unique for the given set of nodes and the chosen metric

: deterministic, no dependence on random choices of nodes

: invariant under similarity transformations that preserve the monotony of the metric

▶ MST and clustering

: *theorem 1:* any MST contains at least one edge from each link-set between P and Q partitions

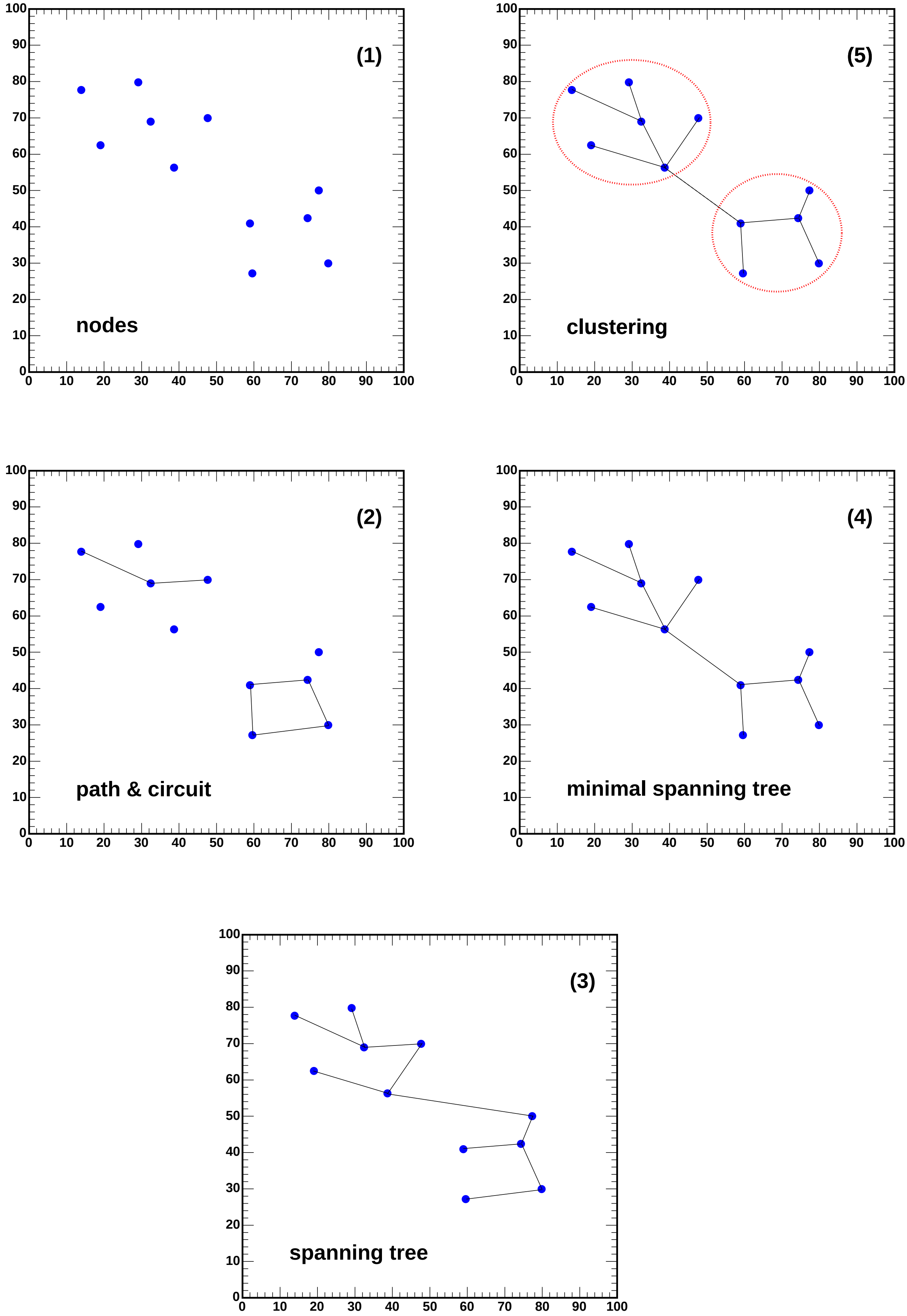
: *theorem 2:* all MST edges are links of some partition of graph

: *theorem 3:* if S denotes the nodes of graph and C is a non-empty subset of S with the property that ρ(P,Q) $<$ ρ(C,S-C) for all partitions P, Q of C, then the restriction of any MST to the nodes of C forms a connected subtree of the MST

: *theorem 4:* if T is an MST for graph G and X, Y are two nodes of G then the unique path in T from X to Y is a minimax path from X to Y

References

[1] C.T.Zahn, *IEEE Trans.Comput. C20(1971)68*
[2] J.C.Gower, G.J.S.Ross, *Appl.Statis. 18(1969)54*
[3] G.J.S.Ross, *Appl.Statis. 18(1969)103*
[4] R.C.Prim, *Bell System Techn.Jour. 36(1957)1389*
[5] K.DeWinter etal. (CHARM II Collaboration) *Nucl.Instr.Meth. A277(1989)170*
[6] N.Saoulidou *Ph.D. thesis, Univ. of Athens 2003*


(1)
nodes
(5)
clustering
(2)
path & circuit
(4)
minimal spanning tree
(3)
spanning tree

Summary – Future planning

▶ **G3-G4 calorimeter simulation**

: high priority

: make cut / particle / energy scan with G3 and G4

: studies to concentrate on prototype geometry

: identify regions where testbeam should focus to give answers

▶ **channel recovery**

: study detector performance vs dead channel fraction

: develop dead channel recovery scheme

: evaluate recovery efficiency and performance improvement

▶ **clustering algorithm**

: develop clustering algorithm based on MST approach