Simulations with the GEANT4 software: optimization of the “cut in range” threshold parameter in different regions of the detector

CALET TIM 2020

Sandro Gonzi

Università degli Studi di Firenze and INFN

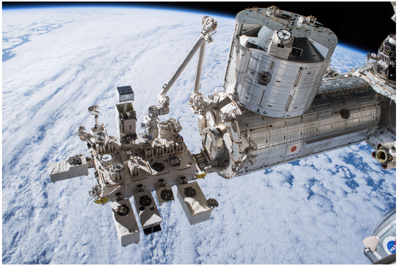


Table of contents

Table of contents

GEANT4 *production cuts*

Premise

The GEANT4 simulations use a **parameter** that sets the *production threshold of secondary particles*. This parameter is usually set to a ***default*** value.

The traditional approach in Monte Carlo softwares is to set a ***tracking cut-off*** in *energy*:

- particles are **stopped** when this energy is reached and the **residual energy is deposited** at this point;
- it may lead to **inaccurate results** in the **stop position** and in the **energy deposit** of the particle (dependence on the *kind of particle* and on the *material crossed*).

GEANT4 doesn't use a ***tracking cut-off*** in *energy*:

- all particles are mapped to **zero energy** (or until they exit the *world* volume or they are destroyed in interactions);

GEANT4 uses a ***production cut***, i.e. it makes a cut by deciding whether a **secondary particle** should be **produced** or not.

Correction of the infrared divergences

It is applied only to: *bremsstrahlung* **photons**, ionization **electrons** and **positrons** and, in addition, *to low energy* **protons** from elastic scattering of hadrons.

GEANT4 *cut in range*

GEANT4 sets, through a parameter, a **production threshold** in *range*:

- the threshold is a **distance** and not an energy;
- threshold value: ***default* = 1 mm**
- particles that fail to travel within the set *range* **are not produced**.

Furthermore, the production threshold in *range*:

- is by *default* **unique** and it sets the **spatial accuracy** of the simulation;
- is **internally converted** into an **energy threshold**, depending on the *kind of particle* and on the *material crossed*.

cut in range

A balance is sought between:

- **accuracy**: need to have a reliable simulation;
- **performance**: by lowering the threshold, the infrared divergences increase slowing down the simulation.

GEANT4 *cuts per region*

Complex detectors may however contain several **subdetectors**: a single *default* production threshold in *range* may **not be appropriate** for all the components of the detector.

The **GEANT4** user can define specific **regions** and assign a **different production threshold to each region**.

The proposed study was born from the need to verify that the threshold in *range* set by *default* (1 mm) is not **too high** for some CALET regions and therefore:

- fewer secondary particles would be produced;
- we would have a systematic error not contemplated;
- we should understand if there is an effect on the *backscattering*.

Goal

To make some **test** by varying the production threshold in *range* in the subdetectors and understand if it makes sense to process data by **optimizing the threshold** or to evaluate a **systematic error** *a posteriori*.

Table of contents

Procedure

1. choose a CALET geometry for simulations:
   - **CALET geometry from the CAD project**
2. choose the subdetectors on which to perform the test:
   - **IMC** and **TASC**
3. identify the materials for which we want to change the production threshold in *range* in the chosen subdetectors and create some specific test regions:
   - *layers* of **tungsten**, *layers* of **scintillating fibers** and **PWO** crystals
4. assign to these regions some different production threshold values in *range*:
   - $\frac{1}{2}$, $\frac{1}{10}$ and $\frac{1}{100}$ + $\frac{1}{1000}$ e $\frac{1}{10000}$ of the vertical dimension of the test region
5. choose a test configuration:
   - **1,000,000** of **electrons vertically** simulated (by respect to the higher surface of CALET) and random position on a **square** of size **2 m** × **2 m** centered at a distance of **2 m**, with **energy** of **100 GeV** and **500 GeV**
6. run the simulations by combininig the threshold requirements in **12 configurations**:
   - **1 original** and **11 modified**, for each energy
7. compare the results for a **variable** sensitive to the variation of the selections for **two different cuts** typical of the electrons analysis:
   - ***IMC-SC***, $\boldsymbol{R_E}$ and $\boldsymbol{F_E}$ variables

Test regions

Given the vertical dimensions of the test regions:

- *layer* of **tungsten**: **TUNG = 0.7 mm**
- *layer* of **scintillanting fibers**: **FIB = 0.96 mm**
- **PWO** crystal: **PWO = 20 mm**

we have run the simulations with the following **thresholds**:

Configuration	Name	Cut in TUNG [mm]	Cut in FIB [mm]	Cut in PWO [mm]	Cut everywhere [mm]
0	ori	1.0	1.0	1.0	1.0
1	mod_TUNG_1	0.35	1.0	1.0	1.0
2	mod_TUNG_2	0.07	1.0	1.0	1.0
3	mod_TUNG_3	0.007	1.0	1.0	1.0
4	mod_TUNG_3_FIB_1	0.007	0.48	1.0	1.0
5	mod_TUNG_3_FIB_2	0.007	0.096	1.0	1.0
6	mod_TUNG_3_FIB_3	0.007	0.0096	1.0	1.0
7	mod_TUNG_3_FIB_3_PWO_1	0.007	0.0096	0.5	1.0
8	mod_TUNG_3_FIB_3_PWO_2	0.007	0.0096	0.1	1.0
9	mod_TUNG_3_FIB_3_PWO_3	0.007	0.0096	0.01	1.0
10	mod_TUNG_3_FIB_3_PWO_4	0.007	0.0096	0.001	1.0
11	mod_TUNG_3_FIB_3_PWO_5	0.007	0.0096	0.0001	1.0

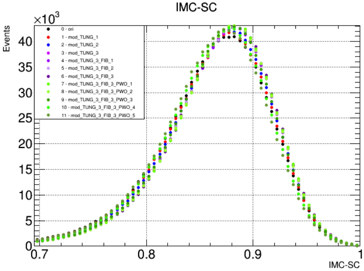
Table of contents

Table of contents

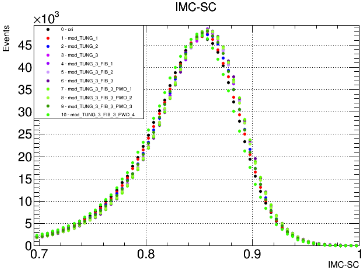
IMC-SC variable

The ***IMC-SC*** (***IMC Shower Concentration***) variable is connected to the shower profile in the IMC and is directly involved in the selection process on *acceptance* for electrons.

IMC-SC is defined as the energy deposit in the last IMC layer within 1 Molière radius of the tungsten (equal to 9 scintilling fibers) from the reconstructed particle direction, divided by the total energy deposit in the last IMC layer.



vertical electrons - 100 GeV



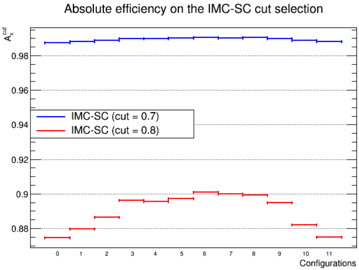
vertical electrons - 500 GeV

The plot profile shifts by lowering the production threshold in *range* in the IMC and TASC.

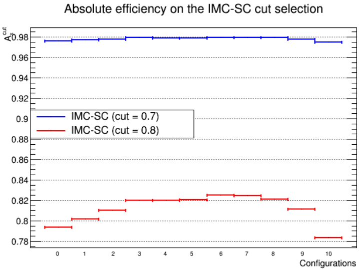
Absolute efficiency on the selection

It is the ratio between the recorded number of events that pass a cut N_x^{cut} and the total number of events N_x in each configuration x, for the following cuts: ***IMC-SC*** $\geq$ **0.7** and ***IMC-SC*** $\geq$ **0.8** (realistic selections for the BDT algorithm):

$$A_x^{cut} = \frac{N_x^{cut}}{N_x}, \qquad x = 1, \dots 11(10)$$



vertical electrons - 100 GeV



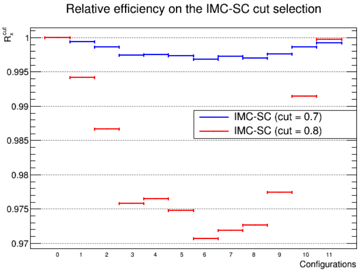
vertical electrons - 500 GeV

The difference with the original configuration increases by raising the cut on the ***IMC-SC*** variable.

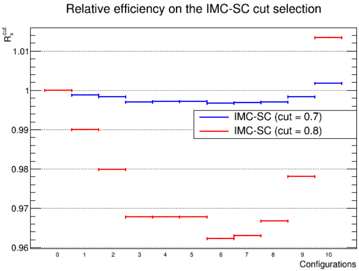
Relative efficiency on the selection

It is the ratio between the absolute efficiency of a modified configuration A_x^{cut} and the original one A_0^{cut}, for the following cuts: ***IMC-SC*** $\geq$ **0.7** and ***IMC-SC*** $\geq$ **0.8** (realistic selections for the BDT algorithm):

$$R_x^{cut} = \frac{A_x^{cut}}{A_0^{cut}}, \qquad x = 1, \dots 11(10)$$



vertical electrons - 100 GeV



vertical electrons - 500 GeV

By raising the cut on on the ***IMC-SC*** variable, we get an effect of 3-4% by lowering the production threshold on the IMC. The effect goes in the opposite direction by subsequently lowering the production threshold on the TASC $\Rightarrow$ **to be understood**.

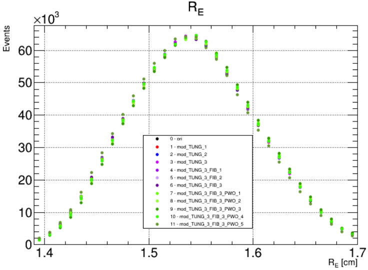
Table of contents

R_E variable

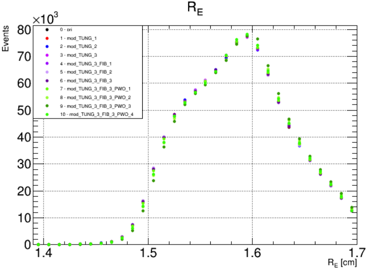
The $\boldsymbol{R_E}$ variable is connected to the *lateral* development of the shower profile in the initial part of the TASC and is directly involved in the *electron/proton discrimination* process.

$$R_E = \sqrt{\frac{\sum_j \left[\Delta E_j \cdot \left(x_j - x_c\right)^2\right]}{\sum_j \Delta E_j}}$$

where x_c is the coordinate of the reconstructed particle track in the first TASC layer, x_j and ΔE_j are the coordinate of the centre and the energy deposit in the j-th PWO log.



vertical electrons - 100 GeV



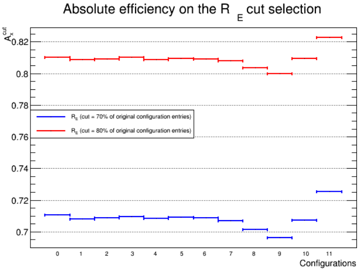
vertical electrons - 500 GeV

The plot profile shifts by lowering the production threshold in *range* in the TASC only.

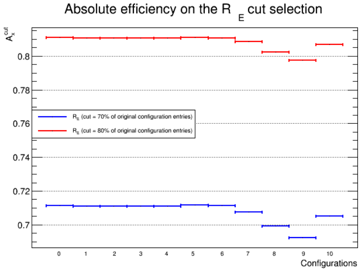
Absolute efficiency on the selection

It is the ratio between the recorded number of events that pass a cut N_x^{cut} and the total number of events N_x in each configuration x, for the cuts $\boldsymbol{R_E^{70}}$ and $\boldsymbol{R_E^{80}}$ that select the 70% and 80% of the events in the *original configuration*: $\boldsymbol{R_E} \leq \boldsymbol{R_E^{70}}$ and $\boldsymbol{R_E} \leq \boldsymbol{R_E^{80}}$ (BDT):

$$A_x^{cut} = \frac{N_x^{cut}}{N_x}, \qquad x = 1, \dots 11(10)$$



vertical electrons - 100 GeV



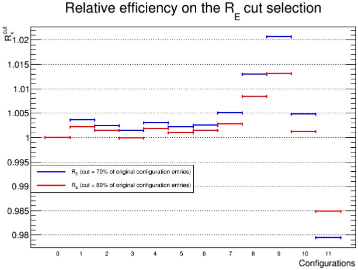
vertical electrons - 500 GeV

By raising the cut on the $\boldsymbol{R_E}$ variable, the difference with the original configuration seems to shift following the same profile,

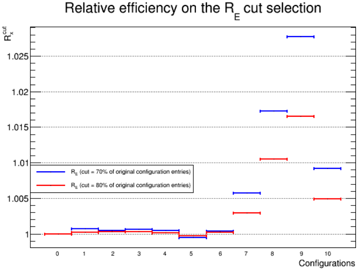
Relative efficiency on the selection

It is the ratio between the absolute efficiency of a modified configuration A_x^{cut} and the original one A_0^{cut}, for the cuts $\boldsymbol{R_E^{70}}$ and $\boldsymbol{R_E^{80}}$ that select the 70% and 80% of the events in the *original configuration*: $\boldsymbol{R_E} \leq \boldsymbol{R_E^{70}}$ and $\boldsymbol{R_E} \leq \boldsymbol{R_E^{80}}$ (BDT):

$$R_x^{cut} = \frac{A_x^{cut}}{A_0^{cut}}, \qquad x = 1, \dots 11(10)$$



vertical electrons - 100 GeV



vertical electrons - 500 GeV

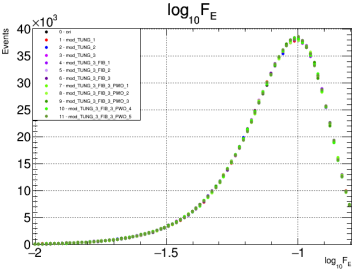
As expected, the plot profile seems to depend on the changes made on the TASC but not on the IMC: it increases and decreases with a maximum effect of 2% with the lowest cut in the R_E variable ⇒ **to be understood**.

Table of contents

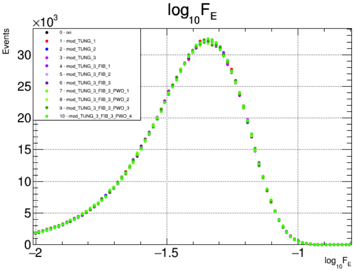
F_E variable

The F_E variable is connected to the *longitudinal* development of the shower profile in the TASC and is directly involved in the *electron/proton discrimination* process.

$\boldsymbol{F_E}$ is defined as the energy deposited in the last (bottom) TASC layer divided by the total energy deposited in the TASC. Selections are applied on the $\boldsymbol{log_{10} F_E}$ value.



vertical electrons - 100 GeV



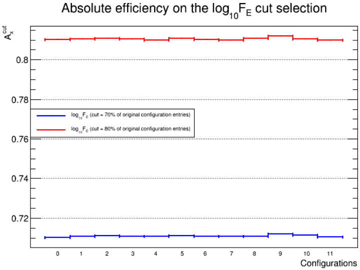
vertical electrons - 500 GeV

The plot profile doesn't shift by lowering the production threshold in *range*, both in the IMC and in the TASC.

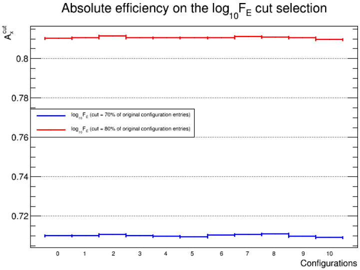
Absolute efficiency on the selection

It is the ratio between the recorded number of events that pass a cut N_x^{cut} and the total number of events N_x in each configuration x, for the cuts $\boldsymbol{log_{10}F_E^{70}}$ and $\boldsymbol{log_{10}F_E^{80}}$ that select the 70% and 80% of the events in the *original configuration*: $\boldsymbol{log_{10}F_E} \leq \boldsymbol{log_{10}F_E^{70}}$ and $\boldsymbol{log_{10}F_E} \leq \boldsymbol{log_{10}F_E^{80}}$ (BDT):

$$A_x^{cut} = \frac{N_x^{cut}}{N_x}, \qquad x = 1, \dots 11(10)$$



vertical electrons - 100 GeV



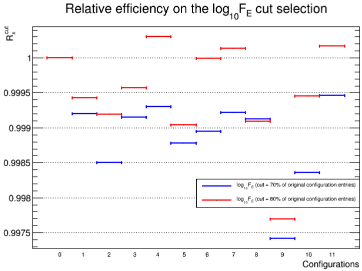
vertical electrons - 500 GeV

By raising the cut on the $\boldsymbol{log_{10}F_E}$ variable, the difference with the original configuration remains constant.

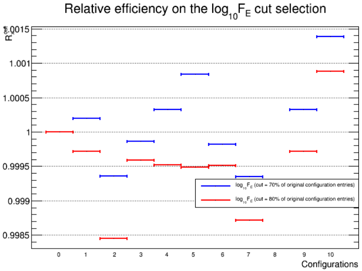
Relative efficiency on the selection

It is the ratio between the absolute efficiency of a modified configuration A_x^{cut} and the original one A_0^{cut}, for the cuts $\boldsymbol{log_{10} F_E^{70}}$ and $\boldsymbol{log_{10} F_E^{80}}$ that select the 70% and 80% of the events in the *original configuration*: $\boldsymbol{log_{10} F_E} \leq \boldsymbol{log_{10} F_E^{70}}$ and $\boldsymbol{log_{10} F_E} \leq \boldsymbol{log_{10} F_E^{80}}$ (BDT):

$$R_x^{cut} = \frac{A_x^{cut}}{A_0^{cut}}, \qquad x = 1, \ldots 11(10)$$



vertical electrons - 100 GeV



vertical electrons - 500 GeV

The effects of the cuts are negligible when the configurations vary, by suggesting that this variable does not depend on the production threshold on the TASC.

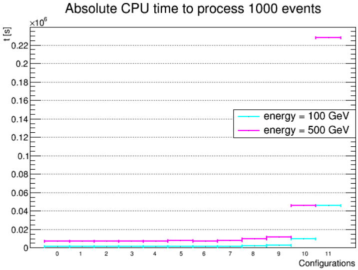
Table of contents

CPU processing time

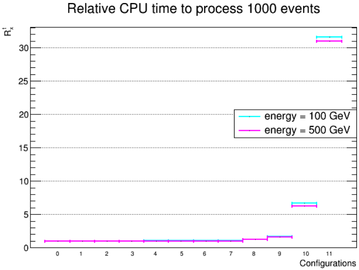
To evaluate the **CPU processing time**, a new test has been carried out on a sample of **1,000 vertical electrons** with energy of **100 GeV** and **500 GeV** for each configuration.

Relative times have been calculated as the ratio between the time recorded for each modified configurations t_x and the original one t_0:

$$R_x^t = \frac{t_x}{t_0}, \qquad x = 1, \ldots 11(10)$$



Absolute times



Relative times

The CPU processing times are sensitive to the variation of the threshold in *range* only for the unlikely values used for the last two configurations, with changes on the TASC.

Table of contents

Conclusions

Conclusions:

- the variation of the thresholds in *range* on the IMC and on the TASC comes to give effects of **a few percent** on the tests carried out;
- the reasons are not entirely clear, especially as regard the TASC: **we can't optimise the parameter** without a clear trend (**plateau** in the relative efficiency plots);
- CPU processing times remain **reasonable** if we don't overdo the changes.

Next steps:

- we could **repeat the test** by changing the kind of particle, the energy and the geometrical features of the generator but a **global approach** could not be so didactic;
- we could study the **electrons flight data** but it is not easy: electrons are higly suppressed by protons and it is difficult to select them without introducing a bias;
- we could analyse the **test beam data** for **electrons**, performed on a simplified version of the detector before the mission started.

In any case, the GEANT4 results would have to be compared with them of EPICS and FLUKA.

Goal

At the end of the work we expect to have a **tuning** of each Monte Carlo generator in order to reduce the **systematic error** for the **electron analysis**.