Contents

1 Changelog

1.1 Draft 0.5

- Added short description on the montecarlo part
- Uniform spatial distribution of avanlanche added

1.2 Draft 0.4

- Normalization factors in pulse shape corrected and description simplified

1.3 Draft 0.3

- Added recostruction section
- Add Furry distribution description
- Improved timing and pulse shape considerations

- Implemented Ole comments

1.4 Draft 0.2

1. strip/pixel readout plane description simplified
2. total collected charge implemented by histogramming bin sums instead of function integral
3. description of the simulation extended: included two candidates flow of processing (fig 1)
4. description of the data structures extended

2 Front Tracker simulation framework

The flow diagram of the Front Tracker simulation is shown in figure 1 and data structures in appendix B:

- GEANT4 is used to generate the "real hits" (both for signal and background); each real hit is identified by the parameters shown in appendix, table 1.
- Real hits are grouped in events (implemented in a `TTree` structure) which can be chained from different files.
- The Real hits can be grouped into tracks, by the "track enumerator" module.
- The digitization module (described below) scans the hits and provides "virtual strips" (one or more for each physics hit), see table 3 in appendix; virtual hits (implemented in `TTree` structure) can be chained.
- A mixer code combines (sums) signal and background virtual strips charge to generated the electronics pulses of the real strips (see table 4).
- The real strips are used to reconstruct the clusters (reconstructed hits) and then the reconstructed tracks, by the reconstruction module described below.

3 Montecarlo

The Montecarlo is based on GEANT4 framework and the code has been designed taking into account modularity and simplicity (for the end user).

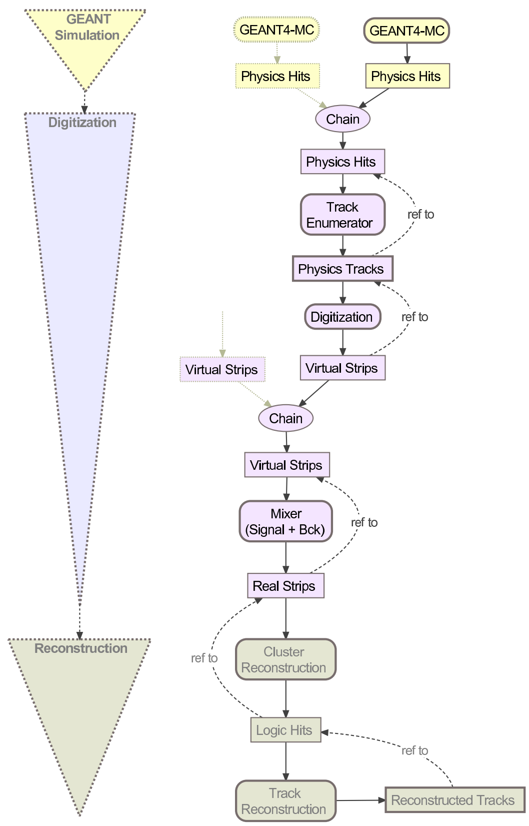

Figure 1: The flow diagram of the Front Tracker simulation framework, with the digitizatation module exploded.

It includes quite detailed models of the target, dipole magnet with field clamps, GEM trackers[1] small silicon detector and preliminary model of the HERMES RICH (plan to be used in the SIDIS experiment). Models are individual objects that inherith from a single class; they can be configured at run-time using the GEANT4 macro mechanism. Configuration parameters can be stored into a dedicated flexible database that can be stored together with the Montecarlo results into a ROOT-Tree output file.

Figure 2 shows part of a possible SBS setup for the SIDIS experiment.

The physics processes can currently be selected from the long list of GEANT4 predefined models. For background estimation the GEANT4 QGSP_BERT physics list has been chosen (it is generally suggested for most of the GEANT4 applications) with equivalent energy cut of 1 keV (configurable).

Figures 3 and 4 present some generated events for background estimation with magnetic field off and on respectively.

The Montecarlo main outputs are the hits information, that is energy loss in the sensitive materials (e.g. GEM gap) of the travelling particle, its position, momentum and time of crossing (or absorption). Each sensitive detector has a pseudo unique identifier, that tags each hit.

As mentioned, all these information are stored in ROOT-Tree files, together with the configuration parameters, which are then processed by the Digitizer discussed in details in the following section; no digitization is performed in the Montecarlo itself. In this way, the extra code is compensated by an higher flexibility; the same Montecarlo data can be used several times to test different configuration of the GEM readout as well as different background conditions.

4 Digitization

The digitization module implements the basic physics processes that generate the electronics signals in the readout strips and the corresponding pulses coming out from the readout electroncs; the starting information are the deposited energies of the primary particle in the GEM chambers of the front traker, simulated by the GEANT4 dedicated code.

4.1 Ionization

Ion pairs are generated randomly (uniform distribution) along the primary track in points (x_i, y_i, z_i). The pairs generated in the drift gap (between the drift plane

[1]The GEM chamber model includes all layers of a triple GEM as well as electronics and mechanical supports.

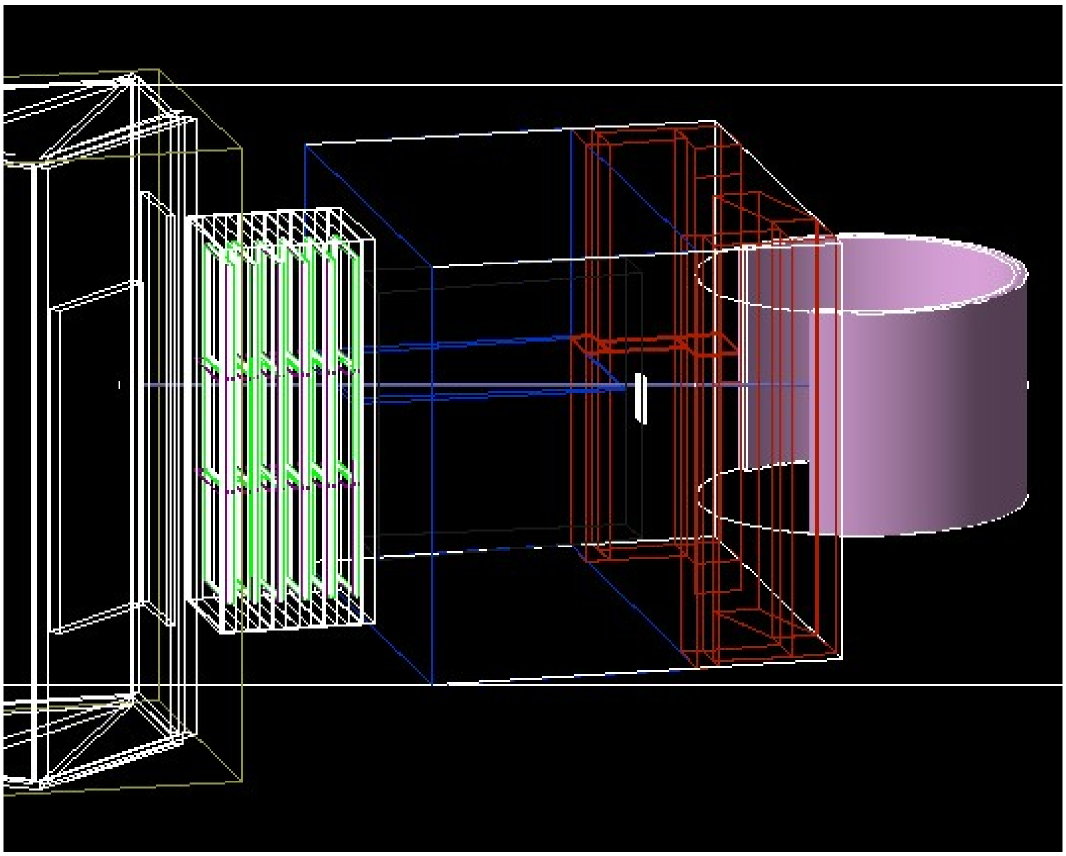

Figure 2: Part of the SBS configuration in the SIDIS experiment: from right to left: scattering chamber (with target inside), small Silicon Strip Detector, Dipole magnet with field clamps, GEM tracker with 6 Chambers, and part of the HERMES RICH detector.

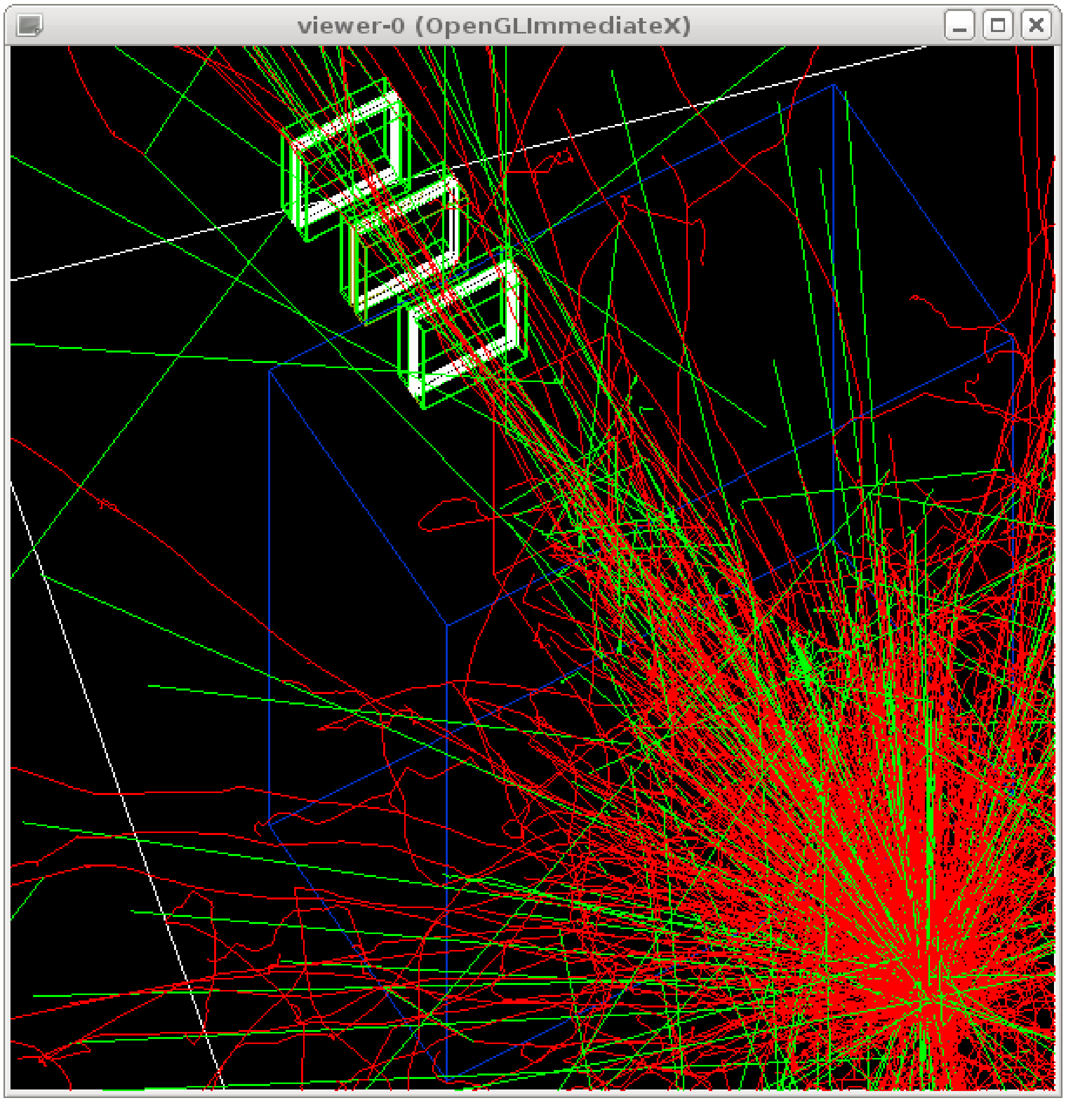

Figure 3: Events of the GEANT4 background simulation, no magnetic field.

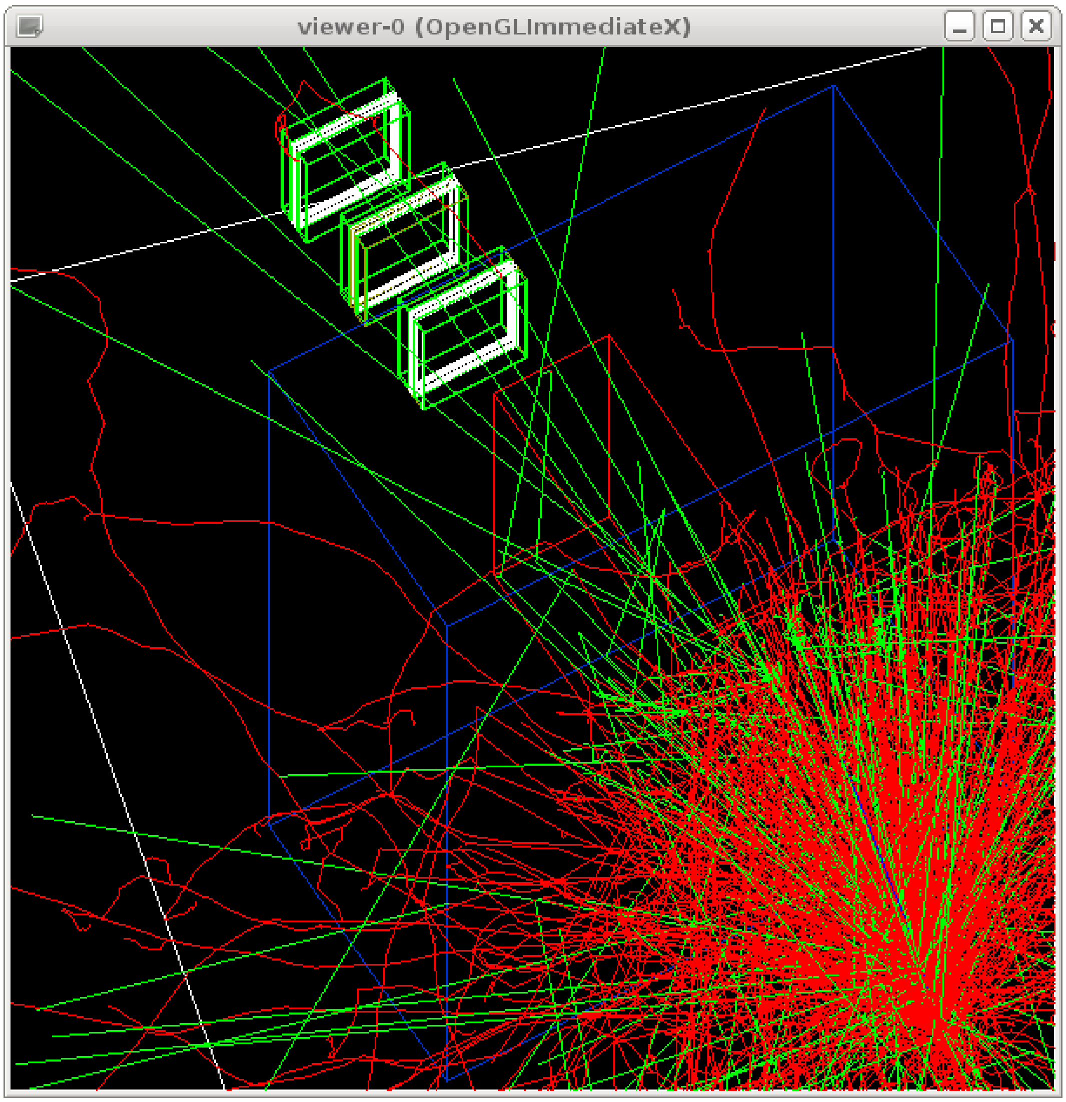

Figure 4: Events of the GEANT4 background simulation, 1.7 T magnetic field, which clean up a lot of low energy charge tracks.

and the first GEM) are multipled enough to provide an observable signal in the readout plane[2].

The average number of ion pairs $\overline{n}_{ion}$ generated by the primary particle is given by:

$$\overline{n}_{ion} = \Delta E / W_i$$

where ΔE is the energy loss by the primary particle, W_i is the effective average energy needed to produce one ion pair in the gas (for Ar $W_i^{Ar} = 26$ eV)[3]

The number of ion pairs originating the hit n_{ion} follows the Poisson distribution with mean parameter $\overline{n}_{ion}$

4.2 Diffusion and Drift

Drift and diffusion times (speeds) are different for positive ions and electrons.

Drifting toward the GEM holes and to the readout plane depends on the electrostatic field. Typical value of the average drift velocity is $v_d = 5 \div 6$ cm/μs, and therefore the time to reach the readout plane is: $t_{ro} \sim L/v_d$ where L is the distance from the center of the drift region to the readout plane (that is the average distance traveled by the secondary electrons generated in the drift region)[4].

The secondaries will diffuse perpendicular to the electrostatic field and will drift along the electrostatic field direction. The diffusion is basically described by the diffusion coefficient D (depending on the gas) which relates the standard deviation σ_s of the spatial charge distribution at time t originated from a point at time 0:

$$\sigma_s(t) = \sqrt{2Dt}$$

Diffusion coefficient in argon for electrons is $200 \div 300$ cm^2/s

We can likely assume that the spatial distribution of the charge in the readout plane is a sort of gaussian with a standard deviation of:

$$\sigma_s = \sigma_s(t_{ro} = L/v_d) = \sqrt{2DL/v_d} \tag{1}$$

Therefore the original ionization points (x_i, y_i, z_i) due to the drifting of the electrons end up in the projection (x_i^h, y_i^h) on the readout plane.

[2]Pairs generated in the other gaps have less chance to be detected due the smaller multiplication.

[3]Replacing ΔE with dE/dx, the specific energy loss, one get the number of ion pairs per unit path.

[4]More realistically, perhaps, L is the distance between the ionization point along the track in the drift region and the corresponding projection in the readout plane and therefore depend on the generated initial pair

4.3 GEM multiplication

Each GEM multiply the secondary electrons (and ions) by an average factor $\overline{g}$ (related to the first Townsend coefficient α by $\overline{g} = n/n_0 e^{\alpha x}$, being x the path where the inelastic processes responsible of the multiplication occur).

Assuming that the electron attachment and molecular dissociation are negligible, in a uniform electric field, the distribution of the number of secondaries follows quite well the Polya distribution[3]:

$$f_{Polya}(n) = \frac{b}{\overline{n}} \frac{1}{(b-1)!} \left(\frac{b(n-1)}{\overline{n}} \right)^{b-1} \exp\left(-\frac{b(n-1)}{\overline{n}} \right)$$

where $\overline{n}$ is the mean avalanche size (that is $\overline{g}$), $b = (1+\theta)$. θ depends on the electrostatic field and the gas properties by $\theta = kW_i\alpha/E$, being k a constant and E the electrostatic (constant) field. Basically θ accounts for the variation of α with n: $\alpha(n) = \alpha(1+\theta/n)$. The Polya distribution becomes a Poisson distribution for $b \to \infty$, while is a Furry distribution for $\theta \to 0$ (that is $b \to 1$):

$$f_{Furry}(n) = \frac{1}{\overline{n}} \exp\left(-\frac{n}{\overline{n}} \right).$$

The latter limit is probably more suitable for a first approximation in GEM (due to the large electrostatic field E in the GEM holes).

Let us consider both cases:

4.3.1 Poisson amplification

Assuming that the multiplication is a Poisson process on each GEM foil, in case of n_{GEM} GEM foils, the total average gain $\overline{G}$ is the product of the gain of the single foil $\overline{g}_i$, that is:

$$\overline{G} = \prod_{1}^{n_{GEM}} \overline{g}_i$$

The distribution of the gain G is not a Poisson distribution (the fluctuation on the multiplication σ_G is much larger than $\sqrt{G}$ and is of the order of the statistical propagation of the fluctuation); however it is similar to a Gaussian distribution (central limit theorem).

In fact

$$\sigma_G \sim \overline{G}/\sqrt{\overline{g}_0} \tag{2}$$

where $\overline{g}_0$ is the mean gain of the first GEM foil (see figure 6).

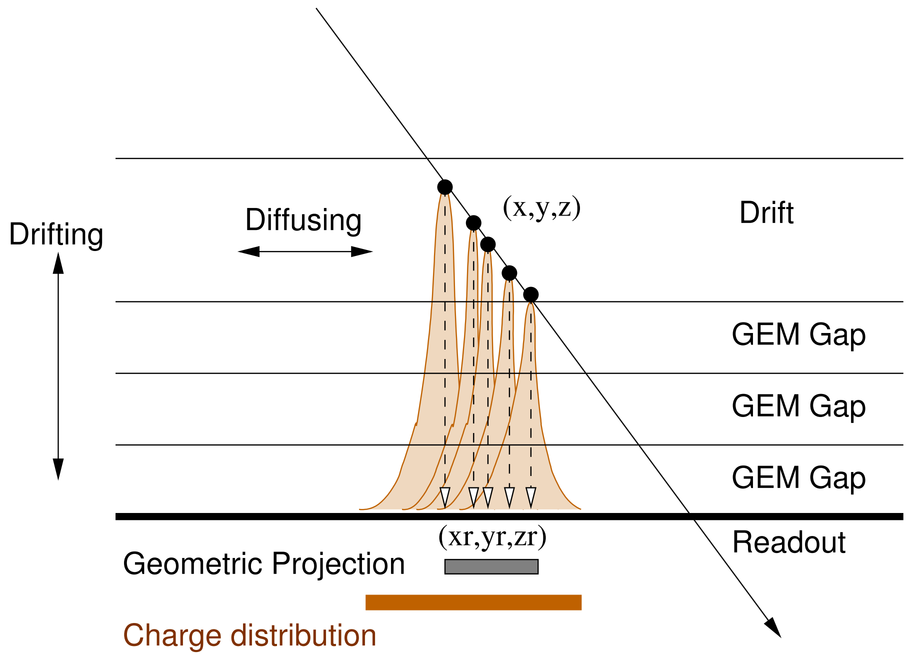

Figure 5: Very schematic view of the ionization, diffusion, drift and multiplication processes in the GEM chamber.

In this case, we can approximate the gain G_i (number of collected electrons from the single ion pair i) to a gaussian distribution with mean $\overline{G}$ and sigma σ_G, that is:

$$G_i(x_G) = \exp\left\{-\left(x_G - \overline{G}\right)^2 / (2\sigma_G^2)\right\} \quad (3)$$

4.3.2 Furry Amplification

In this case we assume that the distribution of the electron avalanche after a single GEM foil is the Furry distribution (see above).

The RMS, after several GEM foils, is similar to the mean value, with the gain parameter replaced by the actual gain (see fig. 7), that is $\sigma_G \sim \overline{G}$.

In this case the gain G_i is distributed almost like a Furry:

$$G_i(x_G) \sim \exp\left(-\frac{x_G}{\overline{G}}\right) \quad (4)$$

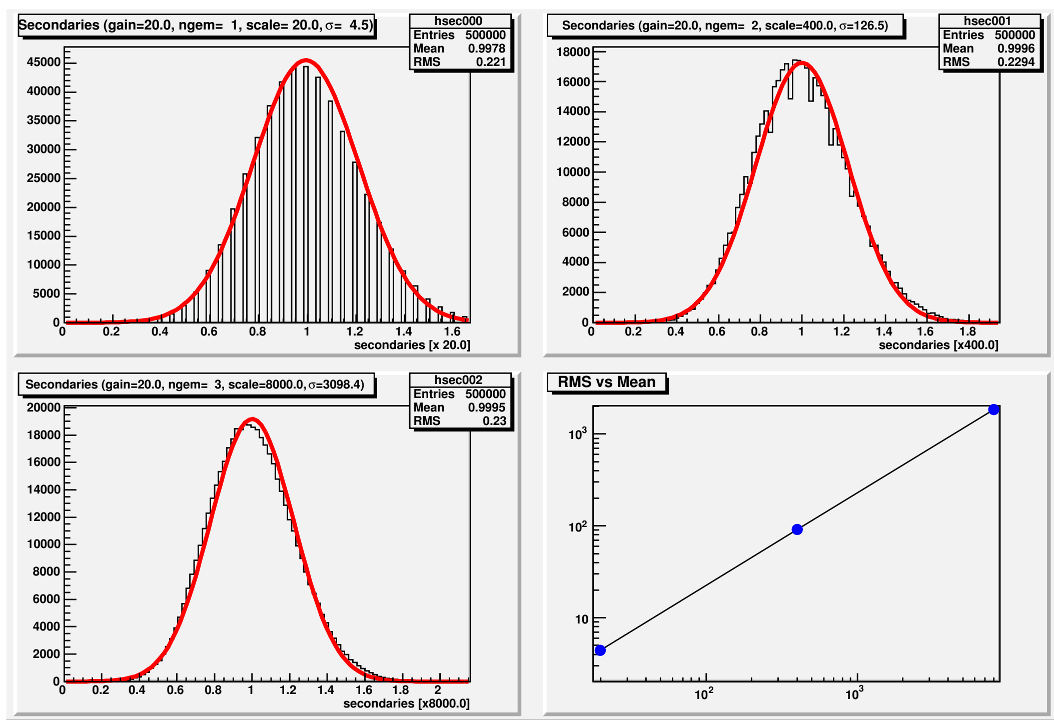

Figure 6: Distributions of the number of electrons after 1 (top-left), 2 (top-right) and 3 (bottom-left) GEM foils, coming from a single secondary pair generated in the drift; equal gain of 20 per GEM foil and Poisson distribution is assumed; in red the gaussian fit. In the bottom-right the RMS of the distribution versus the mean value. Note that the predicted RMS of a pure Poisson would be $\text{RMS}_{Poisson} = \sqrt{8000} = 89.4$ while the RMS predicted by the simulation is $\text{RMS}_{stat} \sim 8000/\sqrt{20} = 1788$. This seems to depend only on the first gain.

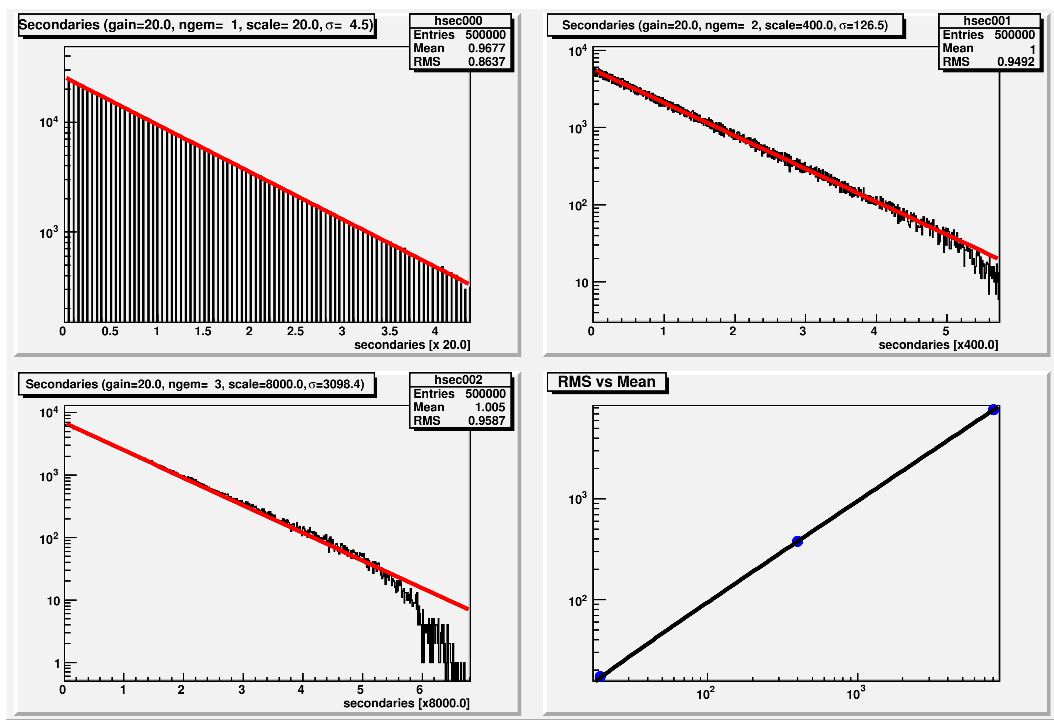

Figure 7: Distributions of the number of electrons after 1 (top-left), 2 (top-right) and 3 (bottom-left) GEM foils, coming from a single secondary pair generated in the drift; equal gain of 20 per GEM foil and Furry distribution is assumed; in red the exponential fit ($y_{3\times GEM} = (8.8361 \pm 0.0019) \cdot \exp(-(1.0163 \pm 0.0013)x/g_3)$). In the bottom-right the RMS of the distribution versus the mean value. Note that the predicted RMS of a pure Poisson would be $\mathrm{RMS}_{Poisson} = \sqrt{8000} = 89.4$ while the RMS predicted by the simulation is $\mathrm{RMS}_{stat} \sim 7700$ very similar to the mean gain (8000).

4.4 Charge Collection

In the first approach, the spatial distribution of the hit charge collected was assumed to be the sum of the gaussian distributions centered at each projection (x_i^r, y_i^r, z_i^r)[5] of the original pair production points (x_i, y_i, z_i) in the drift gap:

$$G_{hit}(x, y) = \sum_{i=1}^{n_{ion}} G_i \exp\left\{-\left((x - x_i^r)^2 + (y - y_i^r)^2\right)\right\} / (2\sigma_s^2(i)) \qquad (5)$$

where G_i is given by eq. 3 or 4 and σ_s by eq. 1

This approach did not reproduce the proper charge sharing between the two layer of the COMPASS like 2dimensional readout plane.

We therefore replaced the gaussian distribution by a constant uniform distribution which demonstrated to reproduce quite well the measured charge distribution from literature [4].

Therefore equation 5 has been replaced by[6]:

$$G_{hit}(x, y) = \sum_{i=1}^{n_{ion}} G_i \cdot H\left((f \cdot \sigma_s^2(i)) - ((x - x_i^r)^2 + (y - y_i^r)^2)\right) \qquad (6)$$

where $H()$ is the properly normalize Heaviside step function (zero for negative values, 1 for positive) and $f = 3$.

Implementation details of the response of the strips of the COMPASS like 2D readout plane are presented in A.

4.5 Pulse formation and Timing

The shape of the analog pulse coming out from the electronics coupled to a silicon detector is approximated by ([1]):

$$v_{out} = A\frac{t}{T_P^2}e^{-t/T_p}$$

where T_p is the shaping time (~ 50 ns, which provide a total width of the signal of about 250 ns), see fig. 9. More realistically, the shape from the electronics coupled to a GEM is represented by[2]:

$$v_{out} = A\frac{\tau_0 + \tau_1}{\tau_1^2}\left(1. - e^{-(t-t_0)/\tau_0}\right)e^{-(t-t_0)/\tau_1}$$

[5] z_i^r is constant and equal to the position of the readout plane

[6] The most realistic function is something closer to the uniform distribution with smoth borders.

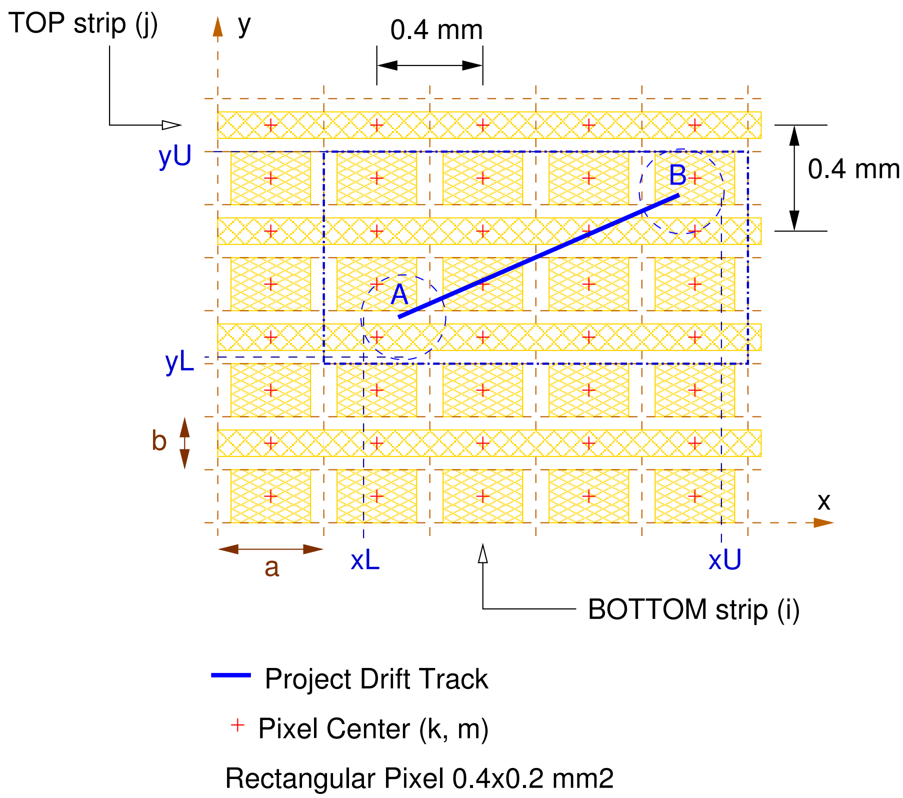

Figure 8: Readout 2D strip plane modeled by a 2D regular array of equidistant rectangular bins (assuming perfect balance of the signal between x and y); bin size is the strip pitch along one axis (a) and the semi pitch along the other (b). The segment $A - B$ represents the projection of the drift track on the readout plane. The (x_L, y_L) and (x_U, y_U) are the lower and upper limits of the area involved in the charge collection (definition is given in the text).

where t_0 is the time shift, τ_0 and τ_1 the time constant that contribute to the width of the pulse; figure 10 compares the two above expressions which appear to differ mainly in the length of the tail.

The normalization A is the total charge on the strip (the I_i or I_j of eq. 7 and 8 respectively)[7].

In our GEM implementation, the above time dependent pulses are sampled by the APV25 with a period of $t_{sample} = 25$ ns. $n_{sample} = 3$ adjacent samples are transferred from the APV25 to the DAQ system and therefore available as raw data of the single strip. In case of signal (particle in coincidence with the trigger) the 3 samples are synchronized and therefore represent always the same part of the function (with a fluctuation related to the trigger jitter).

In case of a background generated hits, the 3 adjacent samples are randomly (uniformly) distributed and therefore, the sampling is random relative to the starting time of the signal. The simulation takes into account an effective time window Δt of about $-t_{signal}$ to $n_{sample} \cdot t_{sample}$, where t_{signal} is the width of the signal[8], while t_{sample} is the sampling time.

The background sums up to the charge generated by the signal.

4.6 Digitization algorithm

The digitization algorithm implements the above modeled processes; it uses the output of the GEANT4 MonteCarlo (see table 1 in appendix) as well as the geometry parameters and physics parameters of the GEM tracker (see 2 in appendix) and produces the `TTree` output (see table 3 in appendix).

Here are summarized the main steps of the algorithm, whose flow chart like is presented in figure 11.

1. Project the track segment in the drift gap to the readout plane.

2. Assume n_{ion} (x_i^r, y_i^r) points Poisson distributed with mean $\overline{n}_{ion}$, in the projected segment.

3. Such points are assumed to be spatially uniformly distributed (along the segment); (implementation: extract n_{ion} values of length from 0 to L_s, being L_s the length of the projected segment).

[7] The normalization factor is defined up to a common factor that includes the APV25 and the ADC gains, at least.

[8] The signal width should be estimated at the over threshold time of the signal, where the threshold is the sensibility of the electronics; as first approximation one can assume the Full Width at One-Tenth Maximum.

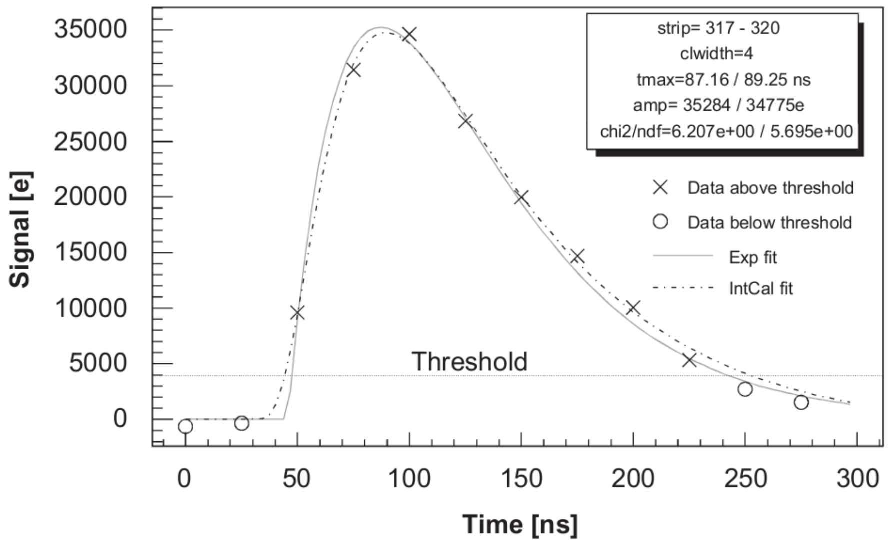

Figure 9: APV25 sampler output from [1].

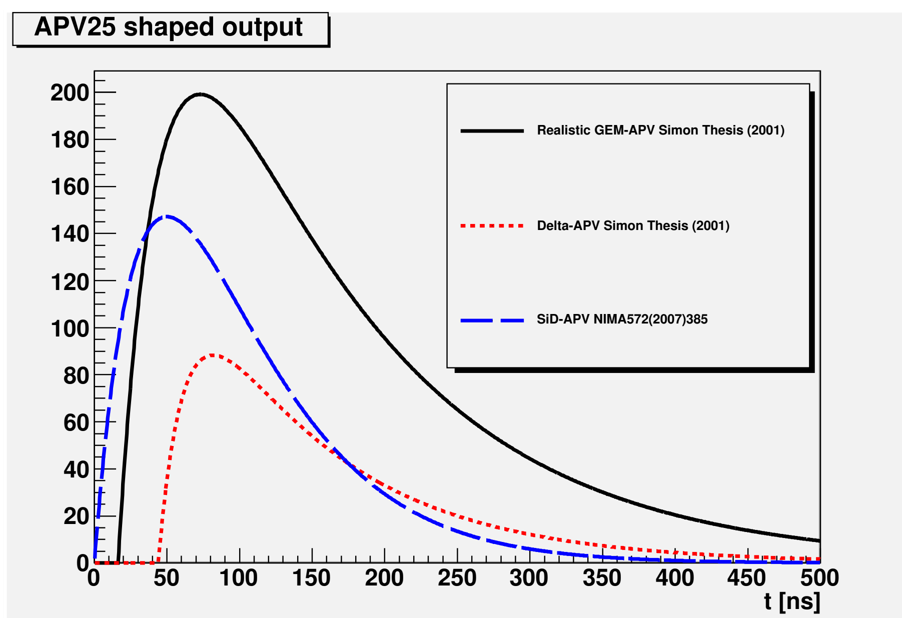

Figure 10: Different functions (not normalized) modeling the pulse shape coming out of the APV25; from [1] and [2]; see also fig. 9

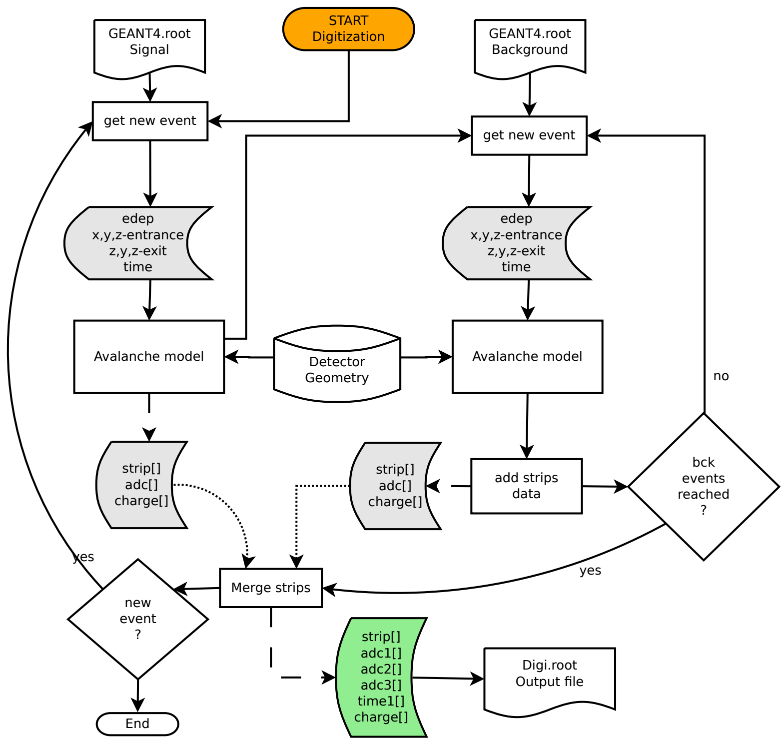

Figure 11: Digitization flow chart diagram (sort of)

4. Each point is the center of a 2D Uniform distribution extending for $3 \cdot \sigma_s$ (from eq. 1).

5. The integral (total charge for each secondary) of the gaussian is given by G distributed according to a gaussian function with mean $\overline{G}$ and σ_G given by eq. 2 or a Furry function (see eq. 4).

6. Sum up all uniform spots (see above equation 6) centered in the projection of the primary ionization pairs points.

7. Choose a rectangular window around the projected segment as suggested above in section 4.4.

8. Integrate the window pixels to get the charge of each strip in x and y (note that integration along one axis is continuous, while along the other is in steps, see fig 8) and discussion above.

9. For each strip in the window take 3 consecutive 25 ns samples of the signal according to eq. 4.5 starting at fixed time (or point in the signal, adding a random gaussian jitter).

10. Repeat all the above steps for the background hits, except step 9, which shall be replaced by:

    (a) extract the number of background hits n_{bck} according to a Poisson distribution with mean $r_{bck} \cdot \Delta t$ where r_{bck} is the background rate (in principle space dependent!) and Δt the effective time window (see 4.5).

    (b) uniformly extract n_{bck} times tr^i_{bck} between $-t_{signal}$ and $n_{sample} \cdot t_{sample}$

    (c) for each strip take 3 consecutive samples starting at the random points extracted in the previous item, relative to the “0” of the GEM signal

A Charge collection implemented

As represented in fig 8 the x-y strips of the COMPASS like 2D readout can be modeled by a 2D regular grid of size $L_x \times L_y$, each "pixel" being addressed by the indexes k, m running in the ranges $k = [0, L_x/a - 1]$ and $m = [0, L_y/b - 1]$). In this representation the horizontal strips (top-layer, the y axis) are continuous while the vertical strips (bottom-layer, the x axis) are interleaved with the horizontal strips.

Each strip can be represented by an index i (and j for the other axis); in this scheme, the j strip is formed by the pixels $\forall k, m = 2j$, while the i strip is formed by the pixels $i = k, \forall(2m+1)$.

Each pixel has center at $(x_k, y_m) = (x_0 + a \cdot k, y_0 + b \cdot m)$, being a and b the pitch along the two directions and x_0 and y_0 the offset that define the origin of the coordinates (for example $x_0 = a/2$ and $y_0 = b/2$).

The gaussian sum of eq. 5 can be defined in a rectangular window (see fig. 8) that includes the segment projection containing the (x_i^r, y_i^r) points; the coordinates of the lower-left (x_L, y_L) and upper-right (x_U, y_U) of this rectangle are:

$$x_L = x_A^r - f\sigma_s^{max},\ y_L = y_A^r - f\sigma_s^{max}$$

and

$$x_R = x_B^r + f\sigma_s^{max},\ y_U = y_B^r + f\sigma_s^{max}$$

where $\sigma_s^{max} = \sqrt{2DL_{max}/v_d}$ (L_{max} is the distance between the drift foil and the readout plane) and $x_{A,B}, y_{A,B}$ are the minimum and maximum values of x_i^r and y_i^r coordinates, while f is the coverage factor ($f = 2$ should be reasonable).

For implementation purpose, instead of (x_L, y_L) and (x_U, y_U) one considers the closest (safer) pixel centers defined by (k_L, m_L) and (k_U, m_U) as: $k_L = \mathit{floor}\,((x_L - x_0)/a)$, $m_L = \mathit{floor}\,((y_L - y_0)/b)$ and $k_U = \mathit{ceil}\,((x_U - x_0)/a)$, $m_U = \mathit{ceil}\,((y_U - y_0)/b)$ (with the constraints on k and m range of validity).

The above indexes are used to define a 2D histogram $H2(n_k, x_l, x_r; n_m, y_b, y_u)$ where $n_k = k_U - k_L + 1$ and $n_m = m_U - m_L + 1$ and borders: $x_l = x_0 + a(k_L - 1/2)$, $x_r = x_0 + a(k_U + 1/2)$ and $y_b = y_0 + b(m_L - 1/2)$, $y_u = y_0 + b(m_U + 1/2)$ and the bin content is given by the charge:

$$C(k, m) = \int_{x_k - a/2}^{x_k + a/2} dx \int_{y_m - b/2}^{y_m + b/2} dy \cdot G_{hit}(x, y)$$

and therefore the charge collected by a i-strip is given by:

$$I_i = \sum_{\forall m} C(k = i, 2m + 1) \tag{7}$$

Table 1: GEANT4 MonteCarlo physics hit parameters definition; these data are used by the digitization; when applicable, the parameters are evaluated in the local chamber. NOTE: name of the variables will be changed in agreement to the ROOT code convention.

Physics Hit	
$hcID$	chamber index
$htID$	particle ID entering the drift gap ($htID = 1$ is the primary particle)
$hPar$	particle code
(hx, hy, hz)	entrance point on the drift gap
(hxe, hye, hze)	exit point on the drift gap
$hedep$	energy deposited in the drift gap (ΔE)
(hmx, hmy, hmz)	entrance momentum on the drift gap
$(hmxe, hmye, hmze)$	exit momentum on the drift gap
$(hx_{ro}, hy_{ro}, hz_{ro})$	entrance point on the readout foil

and the j-strip is given by:

$$I_j = \sum_{\forall k} C(k, m = 2j) \tag{8}$$

B Data struture implementation

The different data structures refer to each others as shown in fig. 1 by the "ref to" dashed lines, in order to permit to compare the reconstructed tracks to the physics tracks; the implementation is done using the *Entry* of the `TTree`.

Table 2: Geometry and physics parameters of the GEM chamber used in the digitization.

L_{max}	drift to readout planes distance
W_i	effective average ionization energy of the gas mixture
D	diffusion coefficient of the gas mixture
v_{drift}	mean drift speed of the secondary electrons
$\overline{g}_i$	gain of each GEM foil (due to gas mixture and field in the GEM holes)

Table 3: Parameters definition of the digitized virtual strip

Virtual Strip	
$chamber$	chamber index
$plane$	axis index
ns	number of strips influenced by hit(s)
$strip[ns]$	strip index array
$charge[ns]$	total charge in strip
$adc[ns][3]$	ADC sampled charge
$sType$	type of strip (signal=0x1, background=0x2 ...)
$pHit$	reference to the physics hit entry (or index)
$pTrack$	reference to the physics track entry (to be implemented)

Table 4: Parameters definition of the mixed real strip (sums of virtual strips). Only 1 dimensional array used, due to some limitation/or not understood behavior in ROOT/`TTree`

Real Strip	
$digi.gem.nch$	(nch for short) number of strips with signal
$digi.gem.chamber[nch]$	chamber index
$digi.gem.plane[nch]$	plane index of the strip
$digi.gem.strip[nch]$	strip address in a given axis
$digi.gem.adc1[nch]$	ADC first sampled value
$digi.gem.adc2[nch]$	ADC first sampled value
$digi.gem.adc3[nch]$	ADC first sampled value
$digi.gem.type[nch]$	type of strip, register
$digi.gem.charge[nch]$	total charge collected in strip
$digi.gem.time1[nch]$	time of first sample

References

[1] M. Friedl et al. NIMA 572 (2007) 385-387

[2] F. Simon, "Commisioning of the GEM Detectors in the COMPASS Experiment", PhD Thesis, Physik-Department E18, Technische Universität München (2001)

[3] A.H. Cookson and T.J. Lewis, "Variations in the Townsend first ionization coefficient for gases", BRIT. J. APPL. PHYS., 1966, vol. 17, pg. 1473 (or http://iopscience.iop.org/0508-3443/17/11/312)

[4] C. Altunbas et al., "Contruction, Test and Commissioning of the Tripe-GEM Tracking Detector for COMPASS", CERN-EP/2002-008