Use of cluster shape information in silicon tracker

Ferenc Siklér

February 16, 2006

Chapter 1

Introduction

Finding charged tracks with silicon tracker in a high density environment is rather difficult in terms of efficiency and fake rate. Track fitting uses hits associated by the previous track finding phase, thus the proper functioning of this latter is essential.

In the actual track finding method seeding starts from pixels. Seeds are produced by combining pixel hits from the two innermost layers. If they are compatible, a further check is made on the existence of another compatible hit in the third pixel layer. Due to the combinatorial nature of the process, the number of possible combinations grows with track multiplicity N as N^2.

The incoming charged particle leaves energy, charge in the channels (pixels or strips) of the detector. Neighboring channels are composed to form a cluster. The reconstructed cluster is regarded as a single hit. Up to know only the position of a hit, a kind of center of gravity for the cluster, has been used for track finding and fitting. It turns out that the length, direction and the average deposited energy of the cluster contain valuable information, as well.

This report consist of two parts. The first part explains how to extract the parameters of a cluster. They can be easily transformed to track parameters. The second part will discuss the possible applications of the obtained track parameters, such as:

- fast vertex finding with clusters only, before track finding
- reduction of compatible cluster seeds, resulting in faster tracking, higher efficiency and lower fake rate; in fact, track finding in track parameter space becomes possible
- V0 finding
- determination of the average energy loss of a particle (dE/dx)

Of course these advantages come at a price: a proper modelling of the detector is necessary with the use of some powerful tools from numerical analysis.

Although most of the material shown here deals with barrel pixels, the results are easily expanded to forward pixels, with some additional work even to strips.

The plots accompanying this study have been made for 1000 minimum bias p+p events and a single central Pb+Pb event.

1.1 Coordinate systems

In the global coordinate system the z-axis is along the beam direction. Both x and y axes are in the bending plane, they can also be decomposed to radial (e.g. p_T) and azimuthal components. The electric field $\mathbf{E}$ is radial, the magnetic field $\mathbf{B}$ is in beam direction, the Lorentz-shift due to $\mathbf{E} \times \mathbf{B}$ is azimuthal.

The local coordinate system is attached to detector units. The x-axis is along azimuthal direction, y-axis is parallel with the beam direction, z-axis is radial. The relations of local axes to global directions and fields are summarized in Table 1.1.

Local axis	Local direction	Global direction	Field direction
x	width	azimuthal	Lorentz-shift $\mathbf{E} \times \mathbf{B}$
y	height	beam	Magnetic $\mathbf{B}$
z	thickness	radial	Electric $\mathbf{E}$

Table 1.1: Local axes and directions and the corresponding global directions with fields.

For the sake of full use of space some of the detector units are "flipped". In case of un-flipped units the local z-axis points outwards in radial direction, while for flipped ones it points inwards. The direction of the electric field $\mathbf{E}$ is always in local z direction. As a result, the measured Lorentz-shift is always in positive azimuthal direction.

1.2 Capabilities

The parameters of a cluster can be transformed to track level. A charged particle at creation point can be described by four geometrical parameters:

- polar angle θ, which is connected to pseudo-rapidity by $\cot\theta \equiv \sinh\eta$
- creation point along the beam line z_0
- initial azimuthal angle ϕ
- signed curvature $q\kappa$, related to $p_{_T} = 0.3B/q\kappa$

If the detector unit was segmented along the beam direction the parameters $\cot\theta$ and z_0 can be extracted. If the detector unit was segmented in the azimuthal angle the parameters ϕ and $q\kappa$ can be obtained.

Barrel strips mostly[1] lie in beam direction (segmented in azimuthal angle), thus they allow to measure ϕ and $q\kappa$. Barrel pixels are segmented in both directions, they allow to measure $\cot\theta$, z_0, ϕ and $q\kappa$.

Relation between the average energy loss and particle momentum exists, based on an energy loss model ("Bethe-Bloch" curve), but only if the particle type, its mass, is assumed. Since it would introduce bias in track finding, it was kept to be a free parameter. As a result of this choice, the projection of the path of particle onto the plane of the detector unit is *well-defined* only if the size of the cluster is greater than or equal to three channels. In other words, the box containing the cluster has to be at least three channels wide in the desired direction. Still, even with box size of one or two channels the estimation of cluster parameter *intervals* is possible.

The capabilities, the efficiency of the method, is shown in Fig. 1.1. The asymmetry in the $p_{_T}$ dependence present between positive and negative particles is due to the effect of Lorentz-shift. It widens the cluster envelope in azimuthal direction for negatives, while narrows for positives. As a result of this interesting effect, clusters belonging to negative particles are much more usable than those belonging to positives. It is especially true for low $p_{_T}$, where the efficiency for negatives varies from 90% to 30% in the range of $p_{_T} = 0$ - 1 GeV/c. The difference vanished for higher $p_{_T}$, giving a steady 20%.

1.2.1 $p_{_T}$ dependence

The length of the cluster in azimuthal direction as function of $p_{_T}$

$$l_x = d\frac{r}{\sqrt{4\rho^2 - r^2}} + \delta_L \tag{1.1}$$

[1] In case of double strip layers the stereo part is rotated by 5 degrees, so those strips will not be perfectly aligned with the beam direction

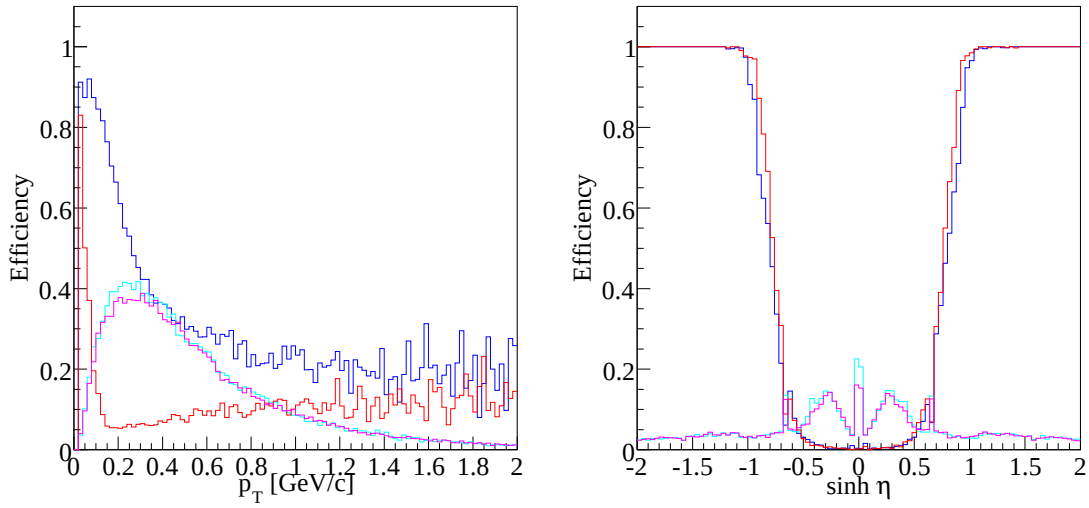

Figure 1.1: The fraction of pixel barrel clusters having at least three channels wide envelope in the corresponding direction. In other words the plots show the efficiency of extracting ϕ, $q\kappa$ and $\cot\theta$, z_0, respectively. Left: The efficiency for extracting parameters in azimuthal direction as function of p_T of the particle. The red histograms show the positive, the blue histogram shows the negative particles. Right: The efficiency for extracting parameters in the beam direction as function of $\sinh\eta$. The magenta and light blue histograms show the number distribution of positive and negative particles as function of p_T and $\sinh\eta$.

where $d = 300$ μm is the thickness of the silicon, $\delta_L = 127$ μm is the Lorentz-shift, $|\rho| = p_T/(0.3B)$. The pitch in x direction is $h_x = 100$ μm. The length n_x in pitch units, as function of p_T will be

$$n_x = \frac{l_x}{h_x} = 3.0\frac{r}{\sqrt{4\left(\frac{p_T}{0.3B}\right)^2 - r^2}} + 1.27 \tag{1.2}$$

2.4 Deposited energy

The energy loss model is discussed in detail in Chapter A. The typical energy loss ξ[keV] is

$$\xi = \epsilon x, \qquad \epsilon = \frac{K}{2} z^2 \frac{Z}{A} \rho \frac{1}{\beta^2} = \frac{178 \text{ keV/cm}}{\beta^2} \tag{2.9}$$

where $K = 307.075$ keV g^{-1} cm^2, $Z = 14$, $A = 28.0855$, $\rho = 2.33$ g/cm^3 for silicon. Futhermore the ionization energy $I = 169$ eV, the plasma energy $\hbar\omega_p = 17$ eV. Note that 1 ADC ≈ 0.5 keV.

A good energy deposition model should describe both the resonant and Coulomb collisions with the effects of the electronic noise. The resonant part in the fast particle approximation, as well as the noise contribution, can be described well with a Gaussian. The energy loss distribution of Coulomb collisions have a long tail towards high energies: this Landau function can also be simplified to a Gaussian, taking its most probable value as mean and converting its full width half maximum to standard deviation σ_C.

$$\overline{\Delta_r} = \xi \sum_n F_n \log \frac{2m_e c^2 E_{0n}}{\hbar^2 \omega_p^2} = 13.30\xi, \qquad \sigma_r^2 = \xi \sum_n F_n E_{0n} \log \frac{2m_e c^2 E_{0n}}{\hbar^2 \omega_p^2} = 11.43 \text{ keV} \cdot \xi \tag{2.10}$$

$$\overline{\Delta_C} = \xi \left(\log \frac{\xi}{I} + 0.20 \right), \qquad \sigma_C = \frac{\Gamma}{2\sqrt{2\log 2}} = \frac{4.02\xi}{2\sqrt{2\log 2}} = 1.71\xi, \tag{2.11}$$

$$\overline{\Delta_e} = 0 \qquad \sigma_e = 500 \cdot 3.7 \text{ eV} = 1.85 \text{ keV} \tag{2.12}$$

The energy deposition distribution is the convolution of the three processes. The result is also a Gaussian with the mean $\overline{\Delta}$ and variance σ^2 summed:

$$\overline{\Delta} = \overline{\Delta_r} + \overline{\Delta_C} + \overline{\Delta_e} = 13.50\xi + \xi \log \frac{\xi}{0.169 \text{ keV}} \tag{2.13}$$

$$\sigma^2 = \sigma_r^2 + \sigma_C^2 + \sigma_e^2 = 11.43 \text{ keV} \cdot \xi + (1.71\xi)^2 + (1.85 \text{ keV})^2 \tag{2.14}$$

As it can be seen in Fig. 2.2 the mean $\overline{\Delta}$ is proportional to the length x. In addition σ and $\overline{\Delta}$ are closely linear. It follows that an easily calculable linear model can be formulated:

$$\overline{\Delta} = \varepsilon x, \qquad \sigma = 0.11\overline{\Delta} + 4.5 \text{ ADC} \tag{2.15}$$

Here x is not necessarily the true path-length, it is enough if x is proportional to the true one. This allows the use of the drift projected planar path-length for the cluster shape optimization. Similarly ε can be different from the true average energy loss.

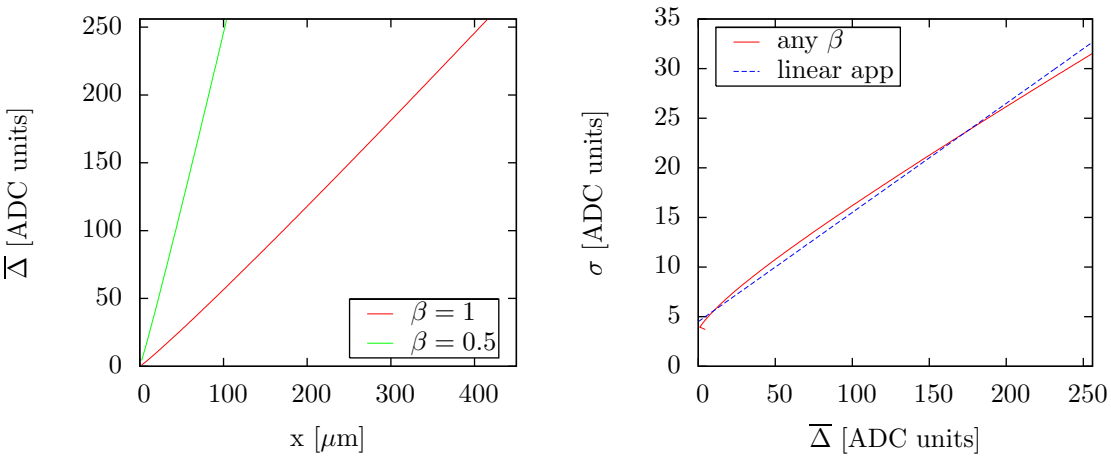

Figure 2.2: Left: The mean of deposited energy per length as function of length when $\beta = 1$. Right: The standard deviation of the Gaussian distribution as function of the mean of deposited energy when $\beta = 1$. The line corresponding to the linear relationship $\sigma = 0.11\overline{\Delta} + 4.5$ ADC is shown in dashed blue.

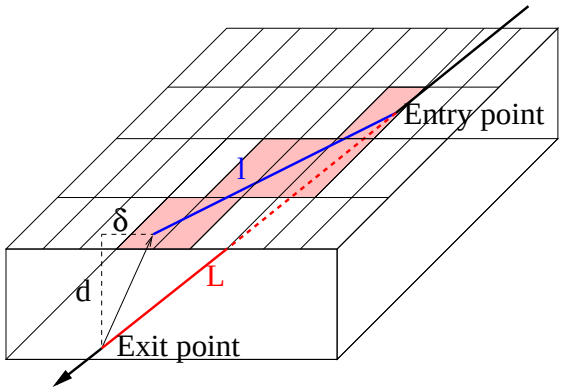

Figure 2.3:

2.5 Noise, digitization, zero suppression

Noise is assumed to have Gaussian distribution. Charge is converted to ADC values, truncated to integers, the maximal allowed value is 255. Channels are zero suppressed, only those are kept which reach the appropriate threshold.

Chapter 6

Track parameters

Depending on the orientation of the detector unit (normal of flipped) the correspondance between local planar, local spatial and global spatial systems is shown in Table 6.1. In case of "normal" situation the Lorentz-shift is subtracted from the exit point, for "flipped" situation it is subtracted from the entry point.

6.1 From cluster to track

The transformation from local to global system can be done via the rotation matrix R and the translation vector $\mathbf{t}$. It can be written for spatial vector $\mathbf{r}$ and directional vector $\mathbf{v}$ as follows.

$$\mathbf{r} \equiv (r_L, \mathbf{r_T}) = R\mathbf{r_{loc}} + \mathbf{t}, \qquad \mathbf{v} \equiv (v_L, \mathbf{v_T}) = R\mathbf{v_{loc}} \tag{6.1}$$

Geometrical relationships in the transverse plane

$$\cos(\alpha/2) = \frac{\mathbf{r_T} \cdot \mathbf{v_T}}{r_T v_T}, \qquad \sin(\alpha/2) = \frac{r_T}{2\rho}, \tag{6.2}$$

If only tracks with less than a total turn ($0 \leq \alpha < 2\pi$) are considered, obtaining α (and the curvature κ) is unambiguous. (The `acos` function returns a value defined to be between 0 and π.)

$$\alpha = 2\arccos\left(\frac{\mathbf{r_T} \cdot \mathbf{v_T}}{r_T v_T}\right), \qquad \kappa \equiv \frac{1}{\rho} = \frac{2\sin(\alpha/2)}{r_T}, \qquad q = \text{sign}\left(\mathbf{v_T} \times \mathbf{r_T}\right) \tag{6.3}$$

Relation between directions and path-lengths

$$\frac{v_L}{v_T} = \frac{\Delta z}{\alpha\rho} \tag{6.4}$$

DetUnit orientation	Lorentz-shift acts on	Entry point is at local spatial	Exit point
normal	entry point	-z/2	+z/2
flipped	exit point	+z/2	-z/2

Table 6.1: Correspondance between local planar, local spatial and global spatial systems. The particle is expected to move outwards in the global system, thus the entry point is closer to the creation point than the exit point, in the transverse plane.

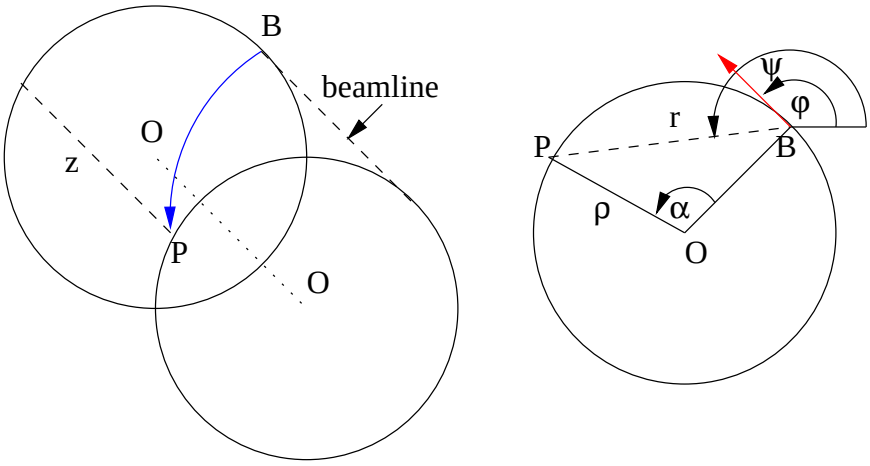

Figure 6.1:

Using the relations above the global track parameters can be calculated.

$$\cot\theta \equiv \sinh\eta = \frac{v_L}{v_T} \qquad \text{polar angle (pseudo-rapidity)} \qquad (6.5)$$

$$\phi = \arg \mathbf{v_T} + q\alpha \qquad \text{azimuthal angle} \qquad (6.6)$$

$$q\kappa = \frac{\text{sign}\,(\mathbf{v_T} \times \mathbf{r_T})_L}{\rho} \qquad \text{signed curvature} \qquad (6.7)$$

$$z - z_0 = \alpha\rho\frac{v_L}{v_T} \qquad \text{shift in beam direction} \qquad (6.8)$$

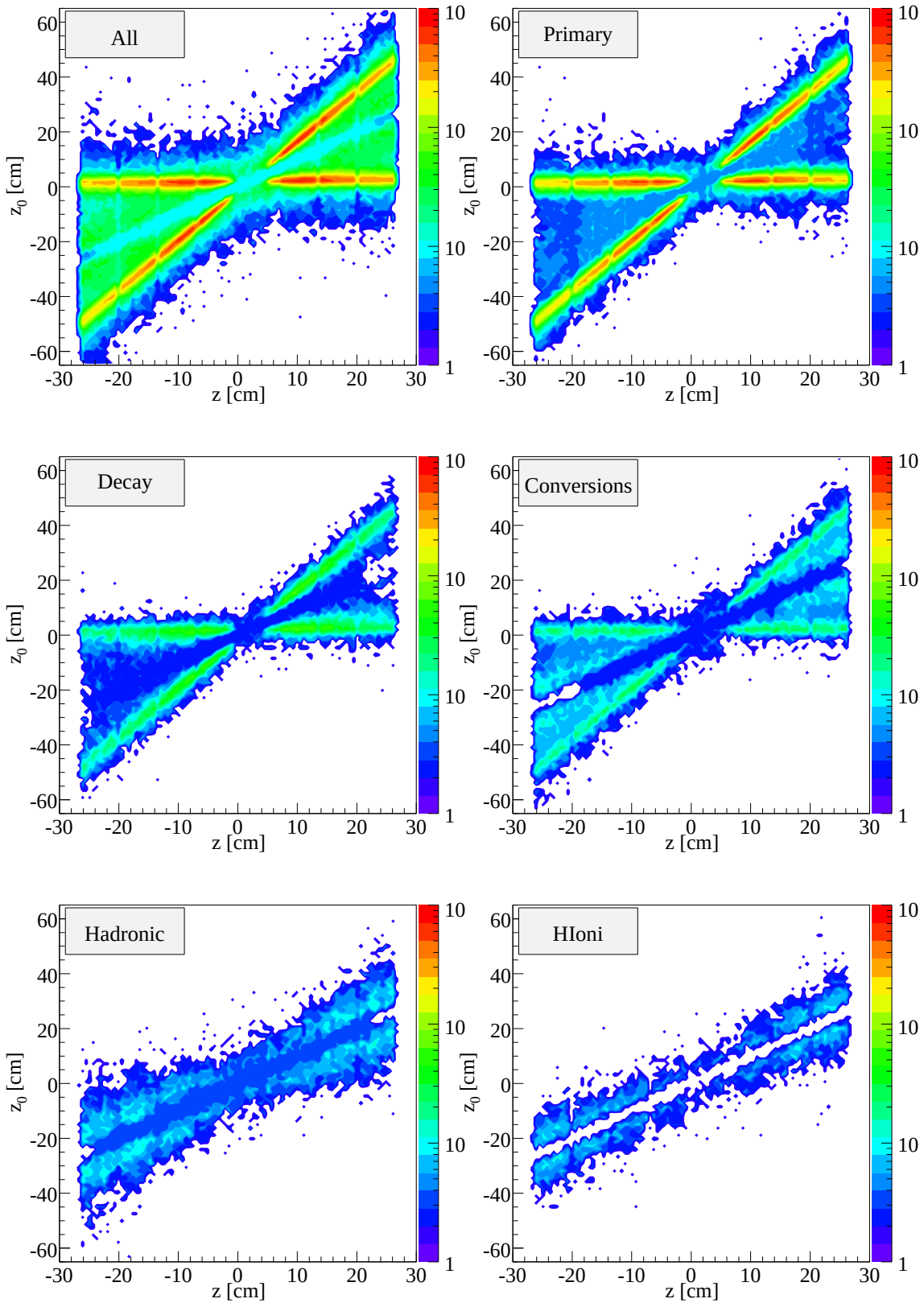

Figure 6.2: Estimated z-coordinate of the creation point for the particles as function of cluster position z, shown here for pixel barrel clusters with box size of at least 3 units in local y direction. The plots give separately the distributions for the most abundant creation processes. Note the log-scale.

Chapter 7

Vertexing

In the actual vertexing method the primary vertex is found after the tracking phase. With the vertex position at hand, tracks are refitted to get better estimation of their parameters. The proper use of cluster shape information, discussed in the previous chapters, enables the estimation of the z-position of the primary vertex $\overline{z_0}$, before track finding and fitting. This way seeding can be limited to a much smaller phase space.

$\overline{z_0}$ can be estimated reasonably well only for pixel barrel clusters, if their container box is at least 3 units wide in local x direction (global z direction). Clusters cannot be fitted unambiguously, the entry and exit points are invertible. In other words, one cannot decide if the track came from positive or negative global z. This way for a cluster there always exists two solutions, two z_0 values. One of them is the real one, the another is just a "reflection" of the first.

The ambiguity can be resolved if we look at not only a single cluster, but more. While the reflections are scattered, the real ones will group around the true value of $\overline{z_0}$. A simple Gaussian fit of the z_0 distribution of clusters in the event could produce a reasonable estimate, its mean could be the vertex position. This is not true if the multiplicity is low and/or if the z positions of clusters are not balanced on the two side of the true $\overline{z_0}$. Outliers due to secondaries, multiply scattered primaries, etc will also distort the estimation.

Here a *robust* two-step method is described which uses the concept of statistical median. If x_j are ordered ($x_1 \leq x_2 \leq \cdots \leq x_{n-1} < x_n$) the median is

$$\tilde{x} = \begin{cases} x_{(n+1)/2} & \text{if } n \text{ is odd} \\ (x_{n/2} + x_{n/2+1})/2 & \text{if } n \text{ is even} \end{cases} \tag{7.1}$$

First, some kind of initial guess is needed. Find the median of all z_0 estimates. Here n will be even, twice the number of the available clusters. For the determination of $\tilde{x}$ a slightly different prescription will be used. $x_{n/2}$ will be chosen if the standard variation of the lower half is smaller than the upper half of the sample. Otherwise the $x_{n/2+1}$ will be taken.

$$\tilde{x}' = \begin{cases} x_{n/2} & \text{if } \sigma(x_1, \ldots, x_{n/2}) < \sigma(x_{n/2+1}, \ldots, x_n) \\ x_{n/2+1} & \text{if } \sigma(x_1, \ldots, x_{n/2}) > \sigma(x_{n/2+1}, \ldots, x_n) \end{cases} \tag{7.2}$$

With this first guess $\tilde{z_0}'$ every cluster is re-examined and one of the solutions is chosen which gives z_0 value closer to $\tilde{z_0}'$. The median of this newly produced sample will give the final estimate of the true $\overline{z_0} = \tilde{z_0}''$, using the original definition (Eq. 7.1).

Some results are shown in Fig. 7.1. The average resolution of the determination of the primary vertex is $\sigma_z \approx 0.35$ mm. This is to be compared with the size of the interaction region $\sigma_z^I = 5.3$ cm. The resolution depends on the number of usable pixel barrel clusters as $\sigma_z(n) \approx 2.9cm/\sqrt{n}$. This gives a resolution of 180 μm for a central Pb+Pb event.

In case of overlapping events the task of finding the primary vertices can also be done by the histogramming method with similar efficacy.

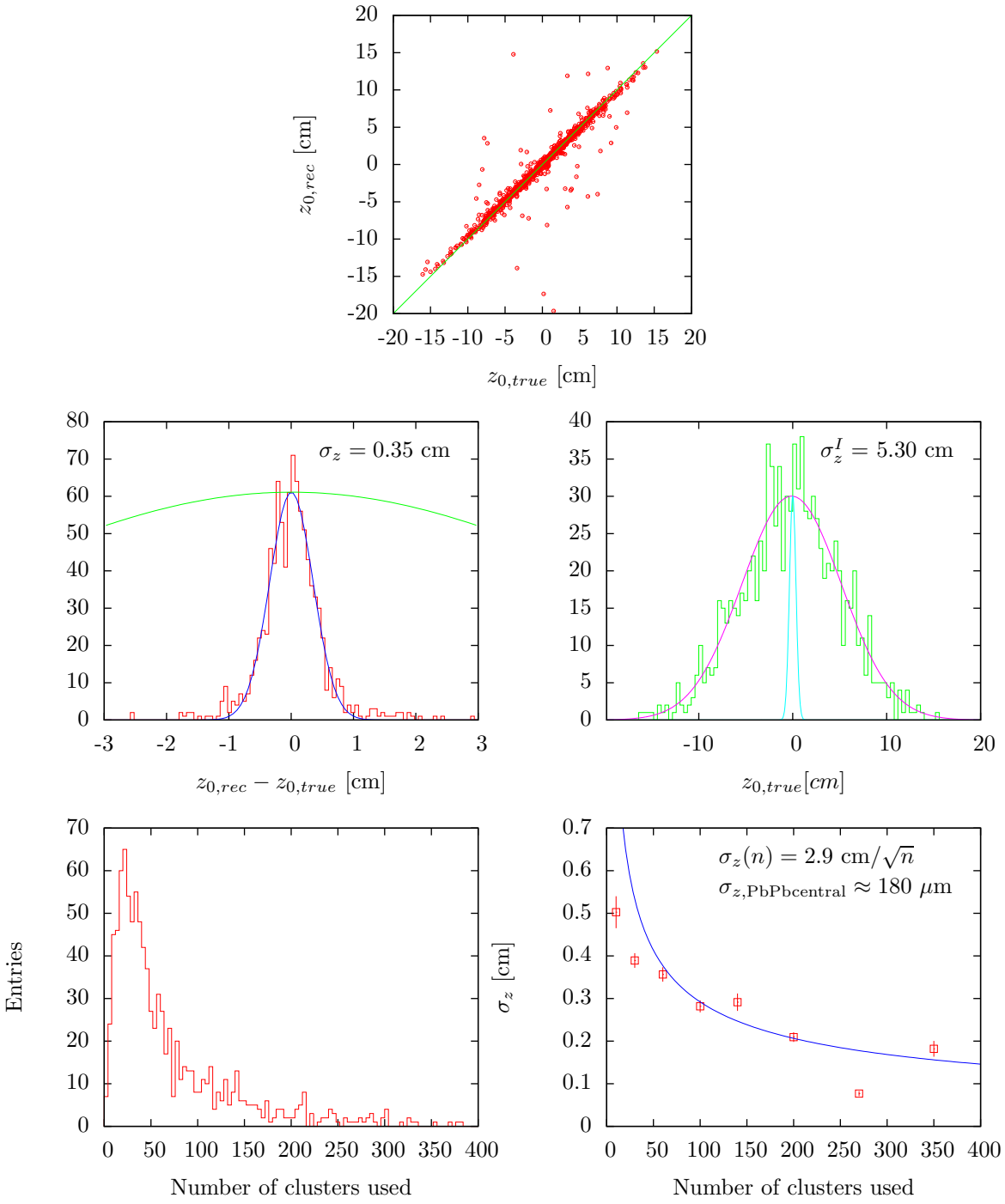

Figure 7.1: Result of vertexing using 1000 minimum bias p+p events. Upper: reconstructed z_0 vs Monte Carlo truth. Center left: distribution of z_0 residuals. Center right: distribution of the simulated z_0 positions. The light blue Gaussian curve indicates the resolution of the vertexing method. Bottom left: distribution of the number of usable clusters. Bottom right: achieved standard deviation of primary vertex determination as function of number of usable clusters. The blue curve is the result of the $1/\sqrt{n}$ fit.