Nuclear Instruments and Methods in Physics Research A 834 (2016) 81–97

Contents lists available at ScienceDirect

Nuclear Instruments and Methods in Physics Research A

journal homepage: www.elsevier.com/locate/nima

A novel approach to Hough Transform for implementation in fast triggers

Nicola Pozzobon [a,b,*], Fabio Montecassiano [a], Pierluigi Zotto [a,b]

[a] Istituto Nazionale di Fisica Nucleare, Sezione di Padova, via F. Marzolo 8, 35131 Padova, Italy
[b] Dipartimento di Fisica ed Astronomia "G. Galilei", Università degli Studi di Padova, via F. Marzolo 8, 35131 Padova, Italy

ARTICLE INFO

Article history:
Received 21 December 2015
Received in revised form
9 July 2016
Accepted 10 July 2016
Available online 14 July 2016

Keywords:
L1 Trigger
Hough Transform
Telescopes of position sensitive detectors

ABSTRACT

Telescopes of position sensitive detectors are common layouts in charged particles tracking, and programmable logic devices, such as FPGAs, represent a viable choice for the real-time reconstruction of track segments in such detector arrays. A compact implementation of the Hough Transform for fast triggers in High Energy Physics, exploiting a parameter reduction method, is proposed, targeting the reduction of the needed storage or computing resources in current, or next future, state-of-the-art FPGA devices, while retaining high resolution over a wide range of track parameters. The proposed approach is compared to a Standard Hough Transform with particular emphasis on their application to muon detectors. In both cases, an original readout implementation is modeled.

1. Introduction

Particle detectors for High Energy Physics experiments, as well as for other applications, have been demanded to feature an increasingly large number of readout channels in the latest years. This requirement for higher granularity is often associated to a higher data rate to be collected and stored for analysis. Moreover, the need of high flexibility and cost containment is moving the approach to triggered readouts from custom ASICs to programmable electronic devices such as FPGAs. A new approach to real-time data analysis, either early fast triggers or high performance noise filtering systems, is then required.

The Hough Transform (HT) is a pattern recognition technique, patented in 1962 [1], originally proposed for charged track reconstruction in pictures taken at bubble chambers. Being a rather demanding algorithm, it was discarded soon, but it has been then successfully applied for decades, in the field of computer vision, for the automated recognition of shapes and features. It has been optimized in several variants for pattern recognition on digital images, after being formalized and adapted to digital implementation. For an inclusive review see, e.g., [2,3]. Recent progresses, mainly based on clever parallelization of the process, let the HT return to its original application field, elementary particle physics, in software algorithms, e.g. for track reconstruction in the time projection chamber of the Alice experiment at the LHC [4]. In this particular case, the HT is run on different CPU platforms with an execution time of the order of 1 ms per track.

FPGAs, which constantly move to higher speed and number of available resources [5–7], provide a competitive alternative as hardware accelerators for the HT, as long as it is optimized according to the specifications of the FPGA device and to the features of the curve to be detected. The necessary computing can be designed around linear operations and histograms, i.e. counters, in the parameter space, which can both be easily implemented in programmable logic. A quite recent project foresees the HT on an FPGA to improve track parameter resolution from patterns already obtained with associative memories in the new CMS tracker for the HL-LHC [8]. Another ongoing CMS upgrade project based on the HT is the Time-Multiplexed Track Trigger, embedding a HT track finder designed around a systolic array [9].

2. Study case: telescopes of position sensitive detectors

Our study case is the implementation of the HT for straight track detection in telescopes of position sensitive detectors, usually silicon devices or drift chambers, immersed in non-negligible background.

Both silicon detectors and drift chambers are affected by noise, but each drift chamber measurement carries an intrinsic left–right ambiguity: each real track is then usually accompanied by several fake companions which are rejected by offline analysis. While actual noise figures are depending on the layout and the materials of the detector, the left–right ambiguity is unavoidable in the drift chambers. Furthermore, drift chambers are used in very large area

* Corresponding author at: Dipartimento di Fisica ed Astronomia "G. Galilei", Università degli Studi di Padova, via F. Marzolo 8, Padova, 35131, Italy.
E-mail address: nicola.pozzobon@pd.infn.it (N. Pozzobon).

http://dx.doi.org/10.1016/j.nima.2016.07.020
0168-9002/

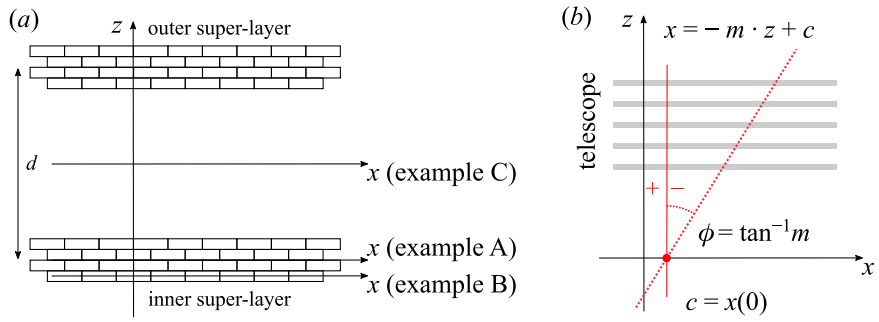


Fig. 1. (a) Layout of the CMS-like DT chamber considered as a study case. Three reference frames with different choices for $z=0$ are superimposed to the layout: such references are used in the examples of Fig. 2. (b) Representation of straight line parameters employed for the HT throughout this work in a generic layout: angles are measured with respect to the normal direction to telescope layers, and the intercept is measured on a reference plane parallel to the telescope layers.

muon detectors, and the number of measured points per track is typically lower than in silicon trackers. A small number of measured points per track is another additional challenge for good HT performance. Drift chambers are therefore an excellent study candidate for the implementation of real-time fast track reconstruction using the HT, even in an ideal situation. The results presented herein can be transposed to the analogous case of silicon detectors arrays by removing the contribution of left–right ambiguities to efficiency and resolutions, which can be expected to be better than in drift chambers. The effectiveness of HT algorithms and their robustness against spurious hits, ambiguities and noise in offline track reconstruction are known features [2]. We will instead devote our study to the development of a novel, fast and reliable approach to the HT, leading to a Level 1 Trigger algorithm.

In order to be realistic we have chosen a simplified version of the CMS muon barrel Drift Tube (DT) detector layout [10], which is a telescope of separate tracking stations, called super-layers (SL), as shown in Fig. 1(a): 8 layers of drift tubes are grouped together in two SLs, being the basic cell width $w=4.2$ cm and height $h=1.3$ cm. Given the thickness of tube walls of 1 mm, the geometrical inefficiency is 2.5% per layer at normal incidence. The distance between the inner and outer SL is $d=23.5$ cm.

Any muon crossing the detector will then provide only $O(10)$ correct input points, associated to the same amount of left–right ambiguities. An additional difficulty of the chosen layout is posed by the use of two rather distant tracking stations, causing measured points to be concentrated in space, rather than being uniformly distributed as, intuitively, would be the best situation. Anticipating the specific solution to this problem, which is addressed in Section 3.1, the designed algorithm exploits the peculiar grouping of sensitive layers with a dedicated partitioning of the input dataset. Despite being developed for a particular layout and a particular detector, the proposed algorithm can be applied to other telescope configurations, adapting it to different geometries. For instance, in an evenly distributed group of detector layers, different sets of layers can be easily generated artificially by an adequate, non necessarily consecutive, channel grouping.

An advantage of choosing the CMS DT muon detector is given by the existing L1 DT trigger [11], which is currently operating and can provide a comparison test-bench for any new development. The current muon trigger in the barrel of the CMS detector is based on a synchronous local trigger specifically designed around the described layout. Local straight track segments, being the magnetic field fully contained in the iron yoke, are built by two ASICs: the Bunch and Track Identifier (BTI) and the Track Correlator (TRACO). The BTI processor exploits the staggered-cell layout of each super-layer to reconstruct hits from the drift-times with a synchronous track fitter. The TRACO processor associates triggers from the BTI outputting track segments identified by a position in the chamber, i.e. the local x coordinate of the intersection of the track segment with the reference plane in the middle of the chamber, and a local direction, i.e. the angle between the track segment and the normal to the reference plane. Furthermore the current trigger and readout systems will be outdated at the HL-LHC and will need replacement, hence the Hough Transform could provide an alternative to this approach. The test bench is then given by the comparison between the performance of the current BTI-TRACO trigger chain and the performance of a newly proposed algorithm implementing the HT, aiming at least at the preservation of the current trigger resolution and efficiency.

We shall consider the true drift time as measured from the time of passage of the muon through the telescope, supposing that this instant can be precisely identified by a pre-trigger. If the drift time is the actually measured quantity, conversion to position is straightforward and the drift chamber can be then considered like a pure position measurement device. Unfortunately this is not the case for LHC muon triggers, where the time frame is unknown and needs to be identified by the trigger algorithm. Since our goal is the development of a fast HT algorithm applicable to purely position sensitive devices, this detector dependent complication, will not be taken into account.

A Standard Hough Transform (SHT) for finding track segments in a CMS DT chamber is presented in Section 3, addressing the peculiarities of a telescopic detector layout, while an original and more compact implementation of the HT is then proposed in Section 4 to address the need of downsizing the SHT histogram.[1] All the geometrical features of the algorithm that can affect efficiency and resolution, including the choice of coordinate system, the quality of input data, and the extraction of track parameters, will be addressed. We have not considered hits generated by electromagnetic background because they are specific of the DT chamber material. Anyway, the measured average fraction of spurious hits associated to a muon track is $O(5\%)$ per cell [12], representing a higher order correction with respect to the left–right ambiguities which are strongly correlated to every measured hit. Final efficiencies and resolutions will then be somewhat slightly biased, but general features could safely be retained. This study, even if fixing the parameters of the algorithm when needed, coherently with the aim of reproducing the performance of current CMS DT trigger, will be of more general interest, showing pathways to the tuning of the algorithm for different layouts and target performance, and general results, not depending on the specific study case, will be emphasized.

[1] Patent pending on key features of the algorithm.

3. Testing the standard hough transform

3.1. Choice of coordinate system and definition of track parameter space

The HT is based on the search for accumulations in the space of the parameters $\mathbf{p}$ used to describe a shape C. Measured points $\{\mathbf{x}_i\}$ are transformed into a bundle of curves in the parameter space, according to the shape equation $C(\mathbf{x}, \mathbf{p}) = 0$. Their intersection in the parameter space identifies the parameter set $\mathbf{p}^*$, which describes the only curve through the measured points in the coordinate space. The core of digital HT is an array of counters associated to a grid in the parameter space, whose granularity is tuned according to the desired efficiency and resolution. The cell in the grid with the highest number of counts (called "votes" in computer vision literature) is chosen as the one defining the best parameter set.

The chosen layout of the telescope layers naturally points to the choice of the straight line description in a cartesian reference frame instead than normal parameters defined by Duda and Hart [13], helping to reduce the complexity of the SHT, since the recurrent computation of inverse trigonometric functions, which are otherwise mandatory, is avoided. Fig. 1(b) shows the convention adopted in this work. Let z be the reference coordinate measured along the direction normal to the telescope, and x and y two orthogonal coordinates in the plane of a layer, that can be measured by the telescope. A slope-intercept line description fits the purpose of defining a viable implementation of a real-time track finder in the xz projection. As for each measured point (x_i, z_i) the coordinate z_i is fixed by the detector layer, such description cannot be degenerate for tracks crossing the telescope, i.e. when the angle ϕ measured with respect to $\hat{z}$ is smaller than critical values depending on the particular layout. The track vector is then defined by the slope $m = \tan\varphi$, i.e. the deviation from the telescope axis, and the intercept $c = x(z = 0)$ on a reference plane:

$$x = -m{\cdot}z + c. \tag{1}$$

Fig. 2 shows an example of the SHT of a straight muon track segment simulated with GEANT4 [14,15] in one SL of a CMS DT chamber, using the previously described convention. The straight line through the i-th measured point meets the condition

$$x_i = -m{\cdot}z_i + c, \tag{2}$$

which can be inverted, so that, for each measured point, we can write the equation of a straight line in the parameter space

$$c = x_i + m{\cdot}z_i. \tag{3}$$

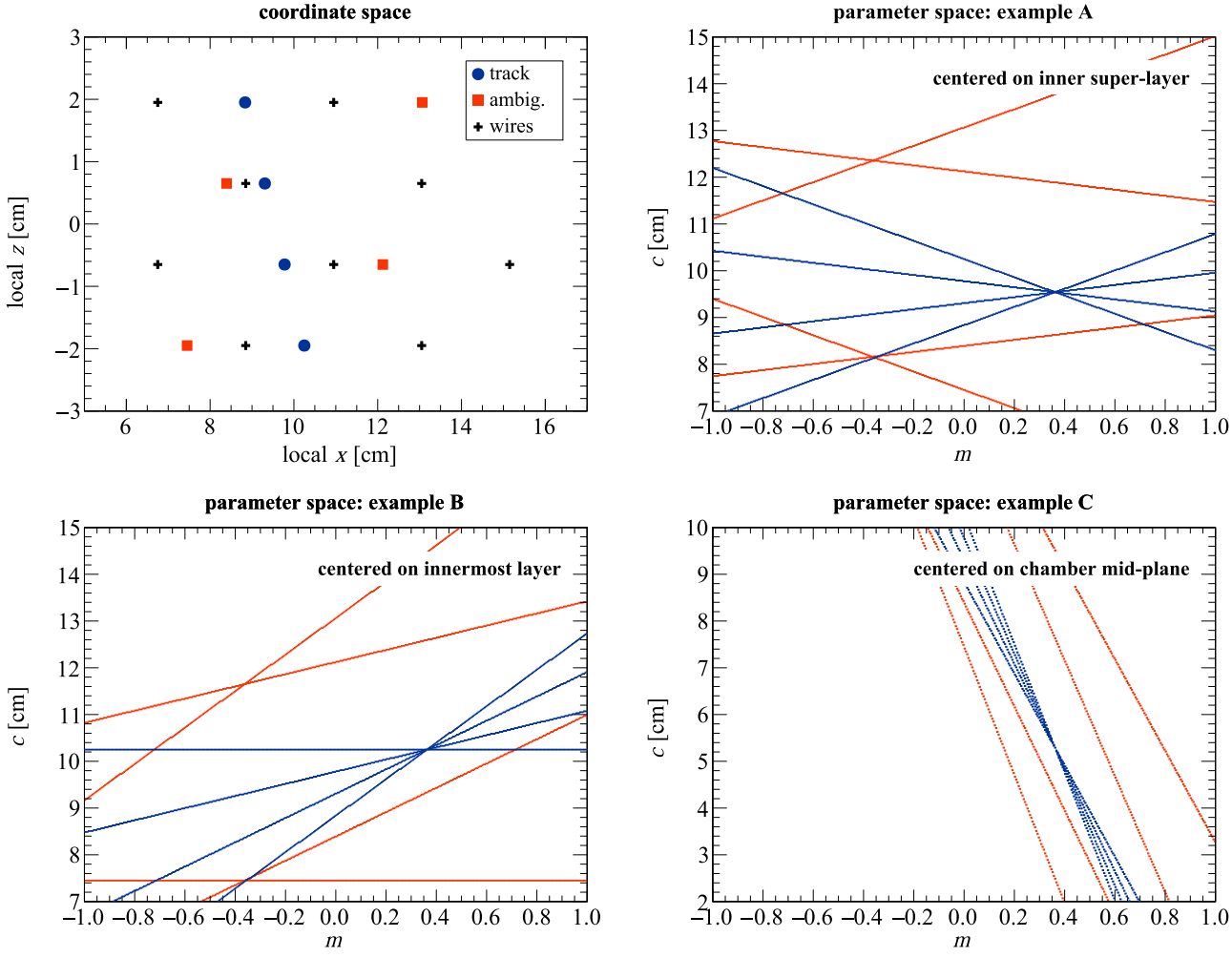


Fig. 2. Example of the HT of a straight muon track segment simulated with GEANT4 in the inner super-layer of a CMS DT chamber. Actual hit coordinates (x_i, z_i) are marked by blue circles; left–right ambiguities are shown as orange squares, symmetric with respect to the actual hits across the wires, drawn as small crosses (top, left). The HT parameter space of this example is shown as a high-granularity scatter plot for different choices of the z coordinate origin according to Fig. 1, with blue (darker) lines representing true hits and orange (fainter) lines representing the left–right ambiguities (example A) z_i is measured with respect to the middle of the super-layer; (example B) z_i is measured from the lowest layer of the telescope; (example C) z_i is measured from the middle of the DT chamber. Accumulation points are where lines converge.

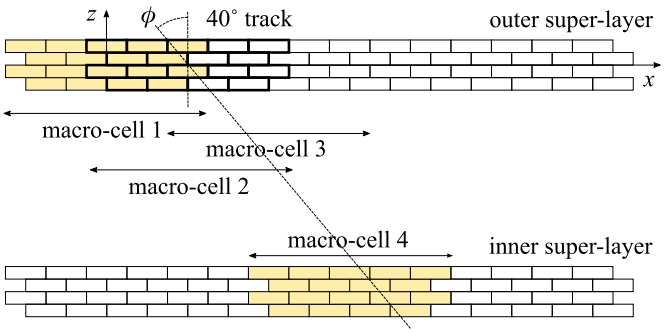


Fig. 3. Definition of the macro-cells used to group hits together for the HT. The picture shows partial superimposition between consecutive macro-cells 1 (colored) and 2 (thick contour lines). A macro-cell in the inner SL is compared to a set of 5 macro-cells (± 2 macro-cells) in the outer one by a chamber scanner associating macro-cells. In the sketched example macro-cell 4 in the inner SL is compared to macro-cells 2 to 6 in the outer SL. Only matching macro-cells with a minimum number of hits are processed. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The aligned points, belonging to the real track, are mapped into lines converging on the parameters fitting the alignment, while non aligned points are mapped into non-converging lines: left–right ambiguities in drift tubes are naturally resolved by the method. Ghost track candidates arising from ambiguities can be filtered by means of quality requirements, i.e. a minimum number of votes in the parameter space, as they generally feature fewer entries in the corresponding (m,c) HT histogram bin.

The SHT histogram is therefore a $n_m \times n_c$ matrix of counters mapping all the lines corresponding to transformed measured points. We shall use square brackets $[p]$ to indicate the histogram bin size for the track parameter p. The required histogram size is driven by the range of the parameters to be detected and by the required resolution. An excessively small bin size $[m, c]$ does not allow to find accumulations since the device resolution makes data too scattered, while a too large bin size results in a more difficult and less precise evaluation of the track parameters. In our case, the poor amount of input data can even turn into an accumulation which goes undetected if the votes corresponding to the parameters $\mathbf{p}^*$ are shared by neighboring bins. These problems require an optimization and are specifically addressed in Section 3.5. However, the finite precision of the detector and the unavoidable contribution of multiple scattering and energy loss effects imply that the size of the accumulation region is larger than a single bin, even after optimization.

Moreover, the accumulation region is better defined if the slopes of the lines in the parameter space are evenly distributed. Being such slope fixed by the positions z_i of the telescope layers, hence the requirement of evenly distributed slopes turns in having equally separated track samplings. This is not the case of the full CMS DT chamber, being this condition valid only within each super-layer. The good detection of the accumulation is therefore affected by the definition of the reference plane for the z measurement, i.e. where $z=0$ is chosen. Fig. 2 shows also the effect of different reference planes choices for $z=0$ in case only a single super-layer is used to build the HT histogram. If the layers are symmetrically distributed about the reference plane, the slopes of the lines in the parameter space span a wider range, and the best parameter set can be identified more easily. The choice made in this study is to maximize the angle between the lines in the parameter space, while retaining the measurement of each z_i starting from the middle of the telescope: the implemented workaround to this problem is described in Section 3.3.

3.2. Readout modeling

Track parameters, defined as in Section 3.1, span a wide range of values in the specific case of the CMS DT chambers. Muons can cross the chamber at angles larger than 45°, while the intercept c spans the whole chamber length, which can be larger than 2 m. The resolution of the current trigger is determined by the TRACO algorithm to be ≈ 3 mrad for the local angle ϕ and to be ≈ 1.4 mm for the intercept c [11], therefore similar figures are assumed as a goal for this study.

An original readout is modeled to ensure that only a limited range of parameters is being evaluated for each candidate track, in particular the intercept c, also featuring enough redundancy to limit the geometrical inefficiencies. Only the general notation "HT" is used in the present Section, being the proposed readout scheme independent of the actual HT implementation, either the SHT or the compact one introduced and described in Section 4.

Hits are assigned to macro-cells, according to a channel identification number. Each macro-cell groups 18 tubes together, and is left–right symmetric, as sketched in Fig. 3. Macro-cells are partly superimposed to each other, so that each tube is then assigned to either 2 or 3 macro-cells in odd or even layers, respectively. The macro-cell size and aspect ratio allow to limit the number of input data to a manageable number for each HT, being the range of values that can be assumed by the parameter c reduced by the macro-cell width to about 17 cm. The macro-cell shape is chosen to accept track segments with angles up to $\phi \simeq \pm 50°$ with respect to the $\hat{z}$ direction. The partial superimposition of macro-cells ensures that any track segment with $|\phi| < 50°$ is fully contained within at least one of them. For the purpose of evaluating the size of the HT histogram in this study, we limit the slope m to the range between -1.1 and $+1.1$.

A track candidate crossing both super-layers can be evaluated with a limited set of channels by using only two macro-cells, one per super-layer. Two kinds of HTs can be produced, the first taking inputs from one macro-cell in one super-layer (1-SL HT), the second taking inputs from two macro-cells from different super-layers (2-SL HT). In the latter case, one must ensure optimal coverage of track parameters, and the following scheme is chosen. Macro-cells with hits collected in at least two wires are selected by scanning one super-layer, let it be the inner SL. Each selected i macro-cell in the inner SL is associated to the 5 macro-cells in the outer one at positions between $i-2$ and $i+2$, as shown in Fig. 3. A macro-cell pair is then used as input of the 2-SL HT if signals are collected in at least 6 wires combining both macro-cells.

The c axis of the 1-SL HT histogram is centered on the macro-cell, while it can be displaced by up to 2 macro-cell width in the 2-SL HT. The c axis spans ± 2 and ± 4 tube widths in the 1-SL HT and the 2-SL HT, respectively.

The redundancy obtained with the described macro-cell superimposition and chamber scanning allows to reduce the effect of scattered accumulations close to boundaries between c bins, which result in a wide but low peak.

The actual c axis center is shifted by half bin in case of odd i macro-cell index, to enhance the chances of centering the accumulation off the boundaries of a (m,c) bin: if a track segment is contained in both the superimposed macro-cells, at least one of the two 1-SL HT histograms would show a clear accumulation well centered on a (m,c) bin.

3.3. Combined use of hough transforms

Section 3.1 already anticipated that the accumulation is better defined with evenly spaced z_i measurements, hence the telescope of the 1-SL HT is made of 4 layers and z origin is set at the middle

of the super-layers, while it is at the middle of the chamber when a telescope of 8 layers is used for the 2-SL HT.

Being the track slope m independent of the choice for $z=0$, it is possible to implement a combined use of the 1-SL HT and the 2-SL HT as follows. A track segment is searched for in each pair of macro-cells by means of 3 HTs: one 1-SL HT in each macro-cell and one 2-SL HT in the union of both. Bins counting above a pre-defined threshold are projected onto the m axis, defining a boolean array or m-bitset: the i-th bit is set high if a (m_i, c) bin over threshold is found. Thresholds are defined from the minimum number of aligned hits required, which is 3 or 4 in the 1-SL HT (corresponding to 4 m-bitsets, one per threshold per SL), and 6 in the 2-SL HT (corresponding to a 5th m-bitset). These threshold values let the SHT histogram to be implemented as a matrix of 3-bits counters, with values larger than 6 set as overflows.

These 5 m-bitsets can be combined together with boolean logic, defining different qualities, as in Table 1: a quality flag is assigned if there is a common subset of m-bits which are set high in all the three m-bitsets used to define the particular quality. Qualities are exclusive and only the highest one is retained when multiple assignments are found.

Fig. 4 shows an example of an 8/8 quality straight muon track segment simulated with GEANT4 in a CMS DT chamber, including all the SHT histograms and the combination of two 4/4 and one 6/8 m-bitsets. Details of the parameter evaluation are discussed in Section 3.4.

3.4. Extraction of the track parameters

Given the small number of input hits and the target resolution of $O(10^{-3})$, fine binning of the m axis is required: the procedure described in Section 3.3 to obtain m-bitsets implies, then, that several bits set high are expected.

Some clustering is performed in order to clearly separate accidental accumulations, due to ambiguities or to the m bin size, that can then affect also the identification of c. Optimization studies show that clusters in a m-bitset must have a minimum width of 0.01 in m and a minimum separation of 0.05 not to be merged together. Smaller size clusters are considered as accidental accumulations in the HT and are therefore rejected.

The best slope m^* is then extracted by averaging on the index of the m-bits set high in each cluster. For each such obtained m^*, the c hypothesis is calculated according to (3) and a 1D histogram, i.e. an array of counters, is filled. This is equivalent to retrieving the corresponding (m^*, c) slice of the SHT histogram. The resulting c-histogram must actually show accumulations, i.e. at least one bin must count at least twice: c^* is extracted by averaging over the maxima of this histogram, without any clustering.

Table 1
Quality flags for the combined use of the 1-SL HT and the 2-SL HT in a pair of macro-cells, in decreasing order. Qualities are named after the number of aligned hits and the number of possible alignments. Thresholds 1 and 2 are used in the two 1-SL HT, while threshold C is used in the 2-SL HT. Quality 6/8L, introduced in order to maximize efficiency, is a loose alignment of 6 hits, 4 of which in the same super-layer. In this case, threshold 1 is implemented as an "anti-threshold": this means that in the 2-SL HT m-projection there are bins counting 6 or more, in one of the two 1-SL HT m-projection there are bins counting 4 or more, while the remaining m-bitset is empty. Quality 4/4 is assigned to the 1-SL HT results with 4 aligned hits. Threshold 1 and threshold 2 are interchangeable between super-layers.

Quality flag	Threshold 1	Threshold 2	Threshold C
8/8	4	4	6
7/8	3	4	6
6/8	3	3	6
6/8L	$\bar{3}$	4	6
4/4	–	4	–

Section 3.5 addresses the problem of optimization of the $[c]$ and $[m]$ bin size. Anticipating some of the results described therein, the required $[c]$ bin size in the 2-SL HT is too large for a direct estimate of the track intercept. For this reason, exploiting the large lever arm d between the two super-layers, the half-sum of the values c_{in} and c_{out}, determined by the corresponding inner and outer 1-SL HT, is chosen for qualities higher than 6/8, while c is assigned directly from the only valid 1-SL HT for qualities 4/4 and 6/8L. To avoid any confusion in the following, the x_0 parameter is defined as the track intercept at the middle of the chamber, hence

$$x_0 = \begin{cases} c_{\text{in,out}} \mp \dfrac{d}{2}\cdot m & \text{for qualities 4/4 and 6/8L,} \\ \dfrac{1}{2}\left(c_{\text{in}} + c_{\text{out}}\right) & \text{otherwise.} \end{cases} \tag{4}$$

The difference $c_{\text{in}} - c_{\text{out}}$ is divided by the distance d between super-layers and compared to the chosen m for a consistency check:

$$m' = \frac{c_{\text{in}} - c_{\text{out}}}{d}. \tag{5}$$

Optimization studies show that if m and m' differ by more than 0.05, the candidate should be rejected. Anticipating another result of Section 3.5, one can assume a typical resolution of the track parameters extracted from the 1-SL HT histogram $\sigma_m \sim 10^{-2}$ and $\sigma_c \sim 0.6$ mm, assuming uniform distribution of c within the accumulation bin of about 2 mm width. These numbers can be used to estimate a lower bound on x_0 resolution as:

$$\sigma_{x_0}^{4/4,6/8\text{L}} \simeq \sqrt{\sigma_c^2 + \frac{d^2}{4}\sigma_m^2} \simeq 1.5 \text{ mm}, \tag{6}$$

$$\sigma_{x_0}^{6,7,8/8} \simeq \sqrt{\frac{1}{4}\sigma_c^2 + \frac{1}{4}\sigma_c^2} \simeq 0.4 \text{ mm}. \tag{7}$$

Moreover, if the track slope is calculated according to (5), for qualities higher than 6/8, its expected resolution can be estimated as well:

$$\sigma_{m'} \simeq \sqrt{\frac{1}{d^2}\sigma_c^2 + \frac{1}{d^2}\sigma_c^2} \simeq 4\cdot 10^{-3}. \tag{8}$$

Being the uncertainty on m' smaller than the one on m, given the small value of σ_c compared to the lever arm d, m' is used to measure the slope of track segments with quality higher than 6/8. Also, the uncertainty on m' is found to be approximately Gaussian and therefore easy to handle as a measure of the resolution. This is specifically addressed in Section 3.7.

A preliminary version of a duplicate cleaner is also implemented: if more than one track segment is found from the same set of input wires, only the highest quality is retained. If duplicates are still present afterwards, the first one, which is the one obtained from the largest cluster in the m-bitset, is retained.

3.5. Evaluation of the benchmark histogram granularity

The main drawback of a SHT implementation for a real-time track finder is the large histogram size needed to obtain good resolution over a wide range of track parameters. The choice of the histogram granularity, i.e. the histogram bin size, addressed in this Section, is then a relevant issue in order to understand the effects on efficiency and resolution. About 40000 straight muon track segments were simulated with GEANT4 in a CMS DT chamber, with local ϕ uniformly distributed between -45° and $+45^\circ$: this sample is referred to as "test sample" in the text. The inclusive efficiency of finding at least one valid track segment per simulated muon over the full ϕ range is the main figure of merit used to guide the

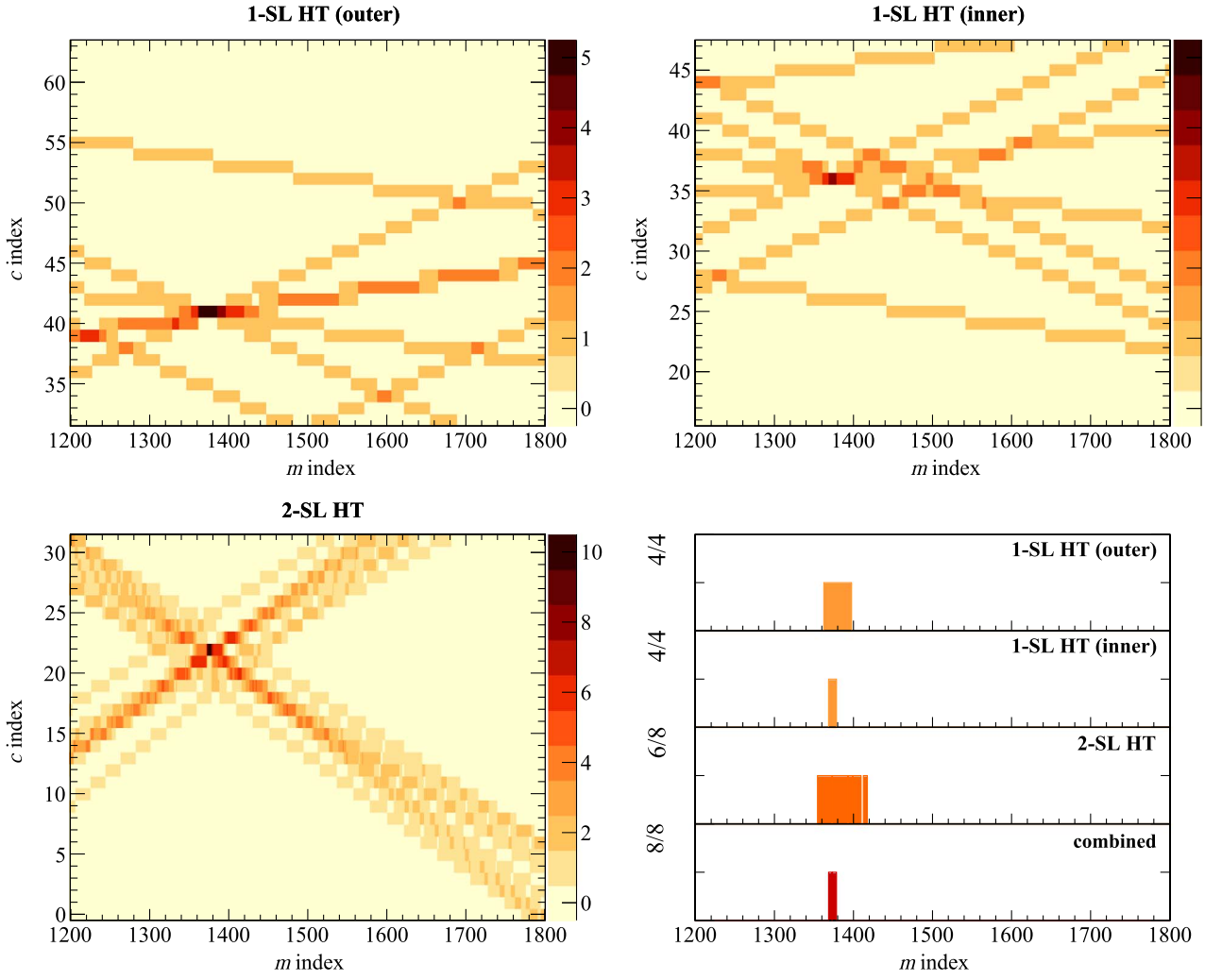


Fig. 4. Example of the identification of an 8/8 quality straight muon track segment simulated with GEANT4 in a CMS DT chamber. SHT histograms in the two 1-SL HT are shown (top) together with the 2-SL HT (bottom, left). The measurement of m is improved by combining together the 4/4 m-bitsets from the two 1-SL HT with the 6/8 m-bitset from the 2-SL HT (bottom, right).

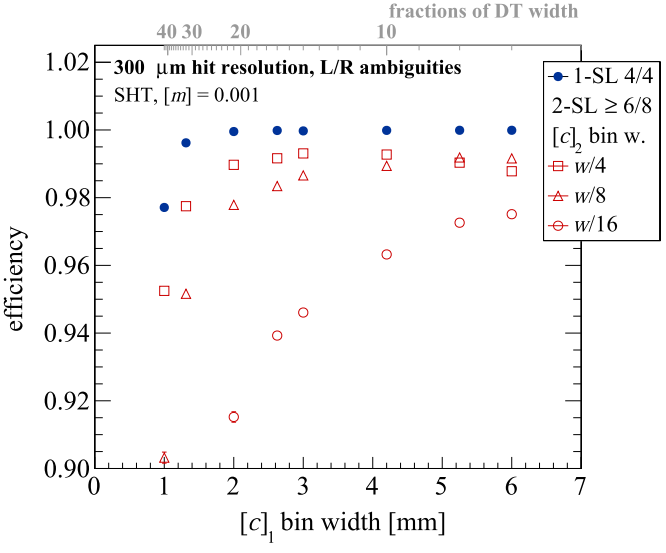


Fig. 5. Efficiency of track segment finding with the SHT in the test sample, shown as a function of $[c]_1$, for $[m] = 10^{-3}$. The hit resolution is simulated by a Gaussian smearing with $\sigma = 0.3$ mm and left–right ambiguities are included for each smeared hit. Full dots show the efficiency to find 4/4 quality track segments, hollow markers show the efficiency of finding track segments with qualities higher than 6/8, for different values of $[c]_2$.

granularity choice, accompanied by the resolution of track parameters. The size of the test sample implies that uncertainties on efficiency are of the order of 10^{-4} applying binomial statistics. The granularity $[m]$ of the slope is chosen to be 10^{-3} for the first evaluation, and the effect of its degradation is presented later.

While, in order to allow consistent comparison among projected histograms, the $[m]$ granularity must be the same for all histograms, the intercept granularity $[c]$ can be different for the 1-SL HT and the 2-SL HT, hence the notations $[c]_1$ and $[c]_2$ are used for the two cases, respectively. Chosen values for testing $[c]_1$ and $[c]_2$ are integer fractions of the drift cell tube w, ranging from 1 mm to 10.5 mm ($w/42$ to $w/4$). Only few relevant $[c]_2$ series will be shown throughout this document for improved clarity of pictures.

The effect of the resolution of the hit position measurement is simulated by a Gaussian smearing with $\sigma = 0.3$ mm, which is a typical uncertainty for drift chamber measurements, even if slightly worse than the resolution of the reconstructed hit position in the CMS DT chambers of $\approx$0.25 mm [16]. Then each hit, after Gaussian smearing, is mirrored with respect to its wire and both the true and the ambiguous positions are used as inputs.

The efficiency figures are reported in Fig. 5. As expected the overall trend shows that efficiencies are higher for larger bin choices. High efficiency comes with the drawback of worse resolution in measuring x_0, as shown in Fig. 6 for the extreme cases of 4/4 and 8/8 qualities. In this Figure only the sub-sample of

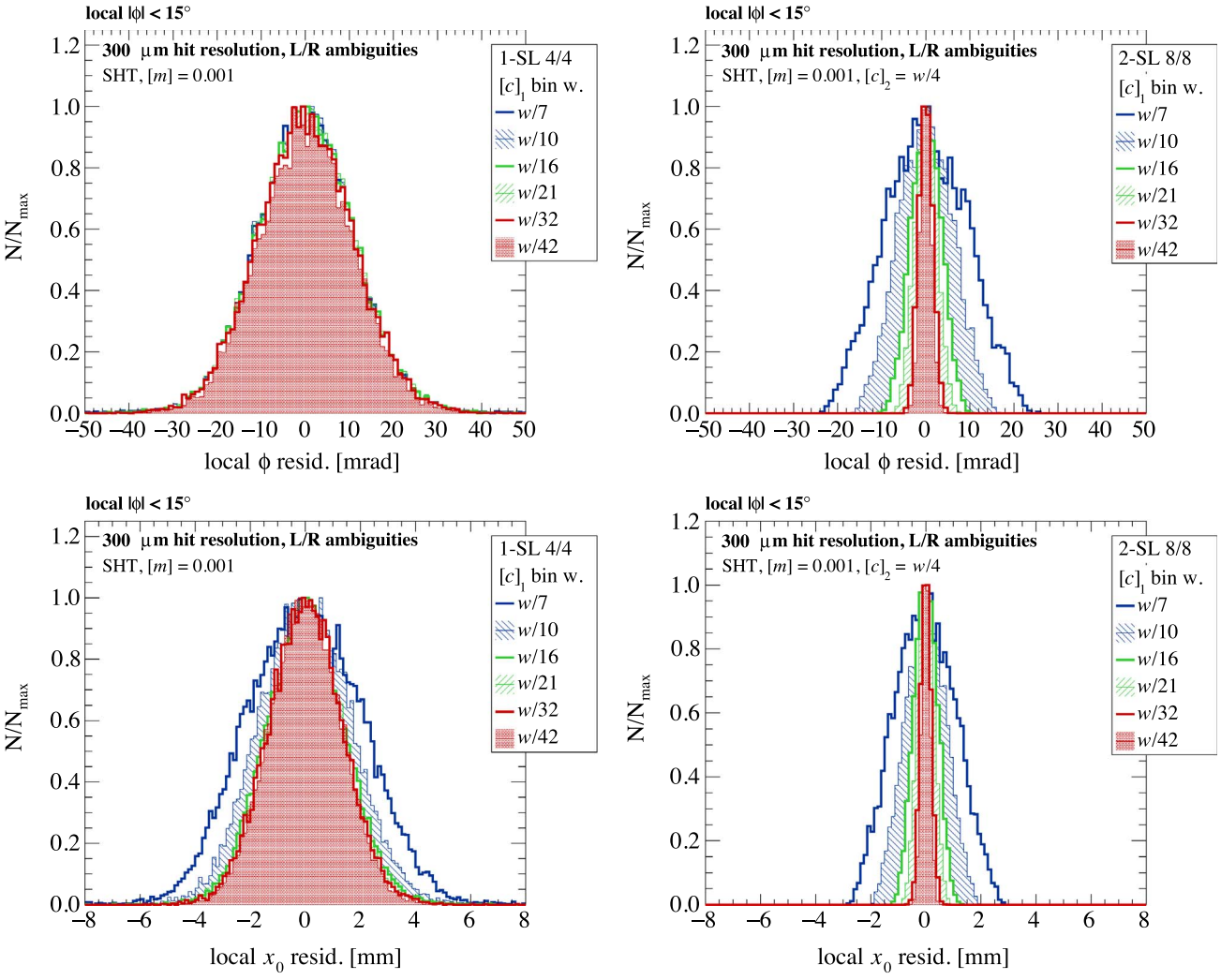


Fig. 6. Residuals of the angle ϕ (top row) and intercept x_0 (bottom row) estimated using the SHT in the test sample subset of muon tracks with $-15° < \phi < +15°$, for qualities 4/4 (left) and 8/8 (right). The hit resolution is simulated by a Gaussian smearing with $\sigma = 0.3$ mm and left–right ambiguities are included for each smeared hit. Being $[m] = 10^{-3}$ and $[c]_2 = w/4 = 10.5$ mm, the distributions are shown for different values of $[c]_1$. Distributions are scaled to the same maximum, and do not account for the actual population of the sample.

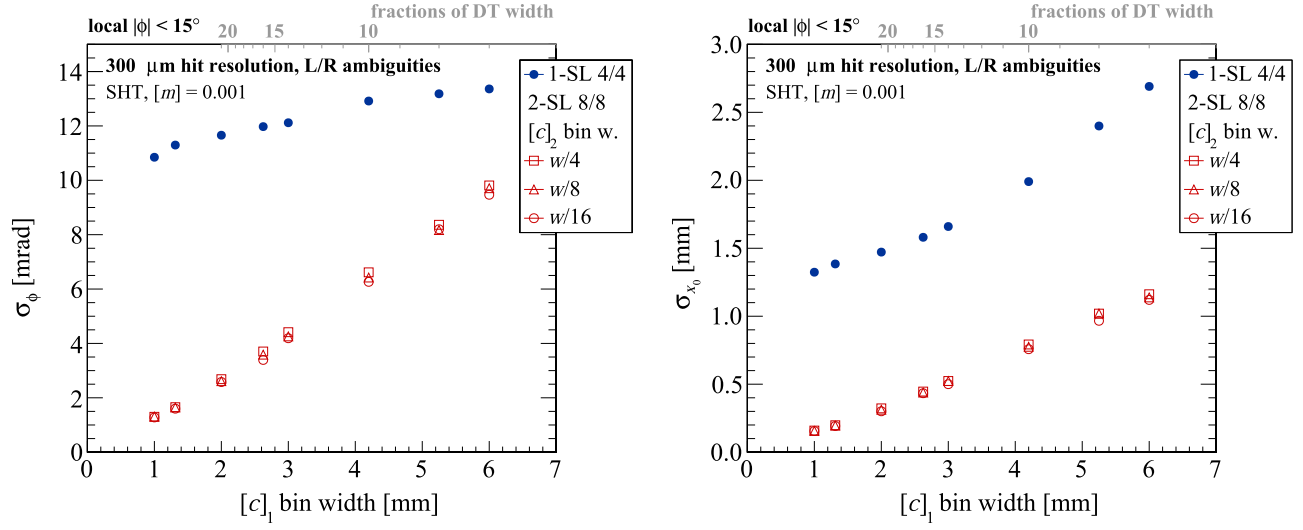


Fig. 7. Width parameter σ of the Gaussian-modeled core of the distributions of residuals of the estimated ϕ (left) and x_0 (right) using the SHT in the test sample subset of muon tracks with $-15° < \phi < +15°$, shown as a function of $[c]_1$, for $[m] = 10^{-3}$. The hit resolution is simulated by a Gaussian smearing with $\sigma = 0.3$ mm and left–right ambiguities are included for each smeared hit. Full dots show the resolution of 4/4 quality track segments, hollow markers show the resolution of 8/8 quality track segments, for different values of $[c]_2$.

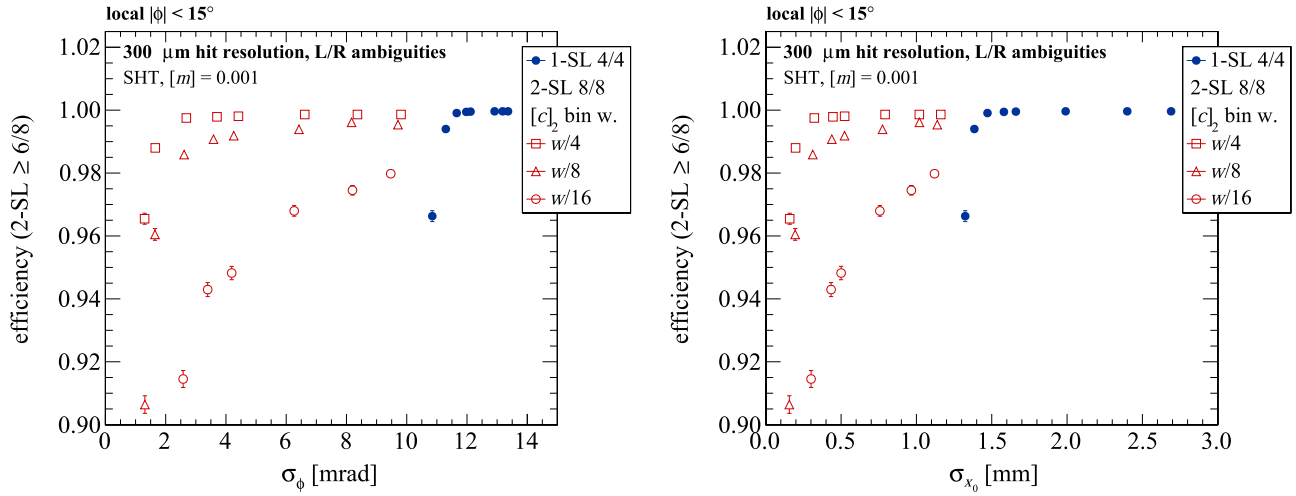


Fig. 8. Efficiency of the track segment finding with the SHT drawn against the resolution of ϕ (left) and x_0 (right) in the test sample subset of muon tracks with $-15° < \phi < +15°$, for $[m] = 10^{-3}$. The hit resolution is simulated by a Gaussian smearing with $\sigma = 0.3$ mm, and left–right ambiguities are included for each smeared hit. Full dots show the efficiency and resolution in case of 4/4 quality track segments, hollow markers show the efficiency of qualities larger than 6/8 and the resolution of their dominant component which is 8/8, for different values of $[c]_2$.

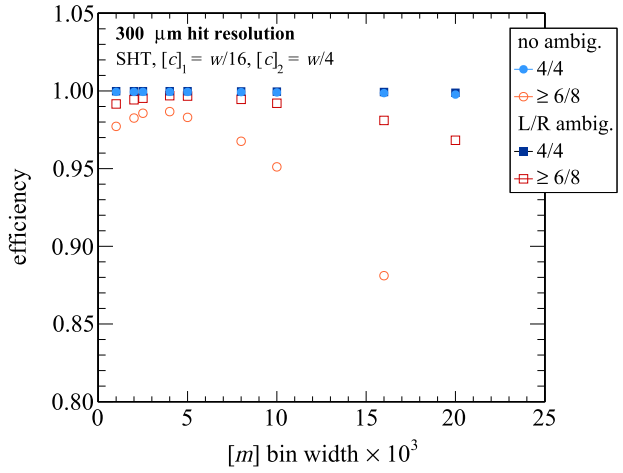


Fig. 9. Efficiency of track segment finding with the SHT in the test sample, shown as a function of $[m]$, for $[c]_1 = w/16 = 2.625$ mm and $[c]_2 = w/4 = 10.5$ mm. The hit resolution is simulated by a Gaussian smearing with $\sigma = 0.3$ mm without considering left–right ambiguities (round markers), while left–right ambiguities are included for each smeared hit (square markers). Full dots show the efficiency to find 4/4 quality track segments, hollow markers show the efficiency of finding track segments with qualities higher than 6/8.

tracks with $|\phi| < 15°$ is shown, as the resolution is ϕ dependent and small angles are the worst case. The limit of 15° is chosen to ensure a reasonable statistical significance. The resolution on the parameters is measured by the variance of a Gaussian function modeling the core of the distribution of residuals. Fig. 7 shows a summary of the resolution of ϕ and x_0 in the subset of tracks with generated ϕ between $-15°$ and $+15°$. The resolution in ϕ is of $O(10)$mrad, slowly increasing with $[c]_1$ for the 4/4 quality, while, in the case of the 8/8 quality, it is approximately linear with $[c]_1$ and independent of $[c]_2$. The resolution in x_0 is approximately linear with $[c]_1$ for both qualities and independent of $[c]_2$ for 8/8 quality. This behavior is expected, given that both $m = \tan\varphi$ and x_0 are obtained via linear combinations of c_{in} and c_{out}, which are derived in the 1-SL HT and do not depend on $[c]_2$ in first approximation, as also implied in (7) and (8). The bin size $[c]_2$ is only used to restrict the range of possible values of m and the width of the hot region in the SHT histograms ensures its impact on the measurement of c_{in} and c_{out} is negligible.

The choice of the benchmark bin size pair $([c]_1, [c]_2)$ for the SHT histograms is then driven by the compromise between efficiency and resolution, which is required to be comparable to the one featured by the current L1 DT local trigger. Hence, quality 4/4 track segments must be compared to the BTI resolution, which is approximately 30 mrad in ϕ and 1.4 mm in x_0, while quality 8/8 track segments must be compared to the TRACO resolution, which is approximately 3 mrad in ϕ and 1.4 mm in x_0 [11]. Fig. 8 shows how the efficiency of track segment finding is related to the resolution of track parameters σ_ϕ and σ_{x_0}.

Combining Figs. 8 with 7, one finds that an angular resolution better than 4 mrad for the 8/8 quality can be obtained with almost full efficiency in the ϕ range between $-15°$ and $+15°$ if $[c]_1 = w/16 = 2.625$ mm and $[c]_2 = w/4 = 10.5$ mm. This is our benchmark choice which leads to a resolution in x_0 better than 0.5 mm. The efficiency in the full range is higher than 0.999. The chosen value of $[c]_1$ also determines the resolution of the 4/4 quality, which is $\approx$ 12 mrad in ϕ and $\approx$1.6 mm in x_0. All these figures are rather close to the target values or even better. The angular resolution improves at angles larger than $|\phi| = 15°$ while the resolution in x_0 slightly worsens, but still within 0.6 mm and 2.0 mm for the 8/8 and 4/4 qualities, respectively. Another viable option would have been to choose $[c]_1 = w/14 = 3$ mm, but a segmentation of the parameter space in powers of 1/2 of the typical length is preferable when integer calculation is to be implemented in a programmable logic device.

In the proposed approach, a track segment in the whole DT chamber telescope is obtained combining three SHT histograms. Given the benchmark choice of $[c]_1$ and $[c]_2$, the 1-SL and 2-SL HT histograms are made of 64 and 32 rows, respectively. The small value of $[m]$ used so far brings up to 2200 the number of columns needed to cover the full range of $|m| < 1.1$. The three SHT histograms are then composed by a total amount of 352000 3-bit counters[2] to be implemented, which cannot fit in any current FPGA.[3] This size can be reduced with larger bin width $[m]$, at a

[2] This total is obtained with a span of 64 bins in c_1 and 32 bins in c_2. This number will then be corrected by the size of the 5 m-bitsets and by the 1D histograms used to calculate c_1 and c_2.

[3] In fact, looking at the data sheets of the state-of-the-art products of major FPGA manufacturers, including the ones foreseen in the forthcoming year, the typical number of CLB LUTs is roughly 3–4 times the total size of the three histograms. The typical fraction of actually used LUTs in research designs, however, is usually smaller by one order of magnitude than the total amount of available ones [6]. Therefore the available resources cannot include the necessary routing even in

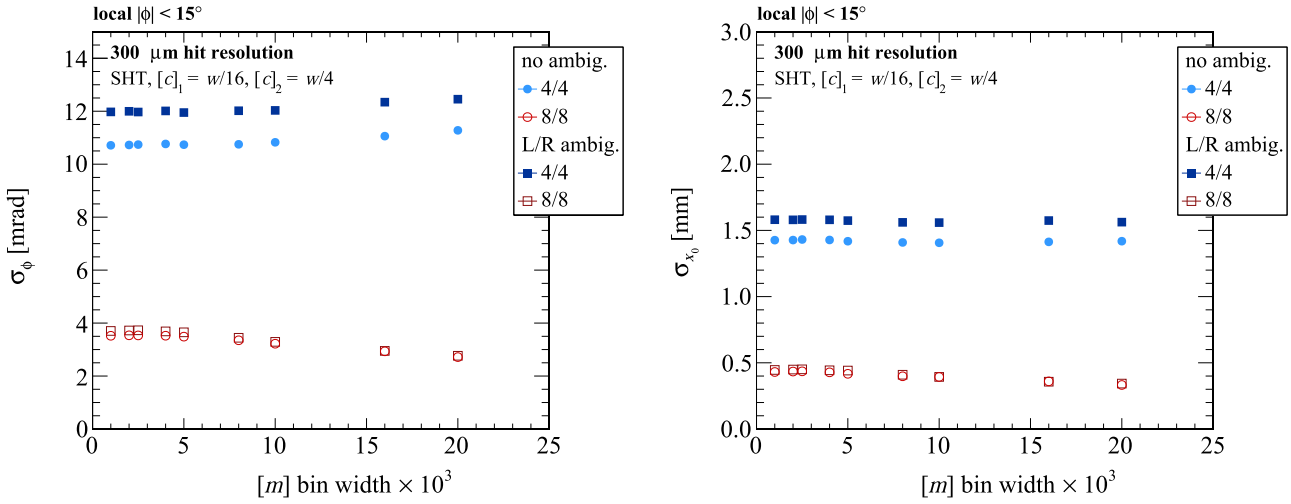


Fig. 10. Width parameter σ of the Gaussian-modeled core of the distributions of residuals of the estimated ϕ (left) and x_0 (right) using the SHT in the test sample subset of muon tracks with $-15° < \phi < +15°$, shown as a function of $[m]$, for $[c]_1 = w/16 = 2.625$ mm and $[c]_2 = w/4 = 10.5$ mm. The hit resolution is simulated by a Gaussian smearing with $\sigma = 0.3$ mm (round markers), and left–right ambiguities are included for each smeared hit (square markers). Full dots show the resolution of 4/4 quality track segments, hollow markers show the resolution of 8/8 quality track segments.

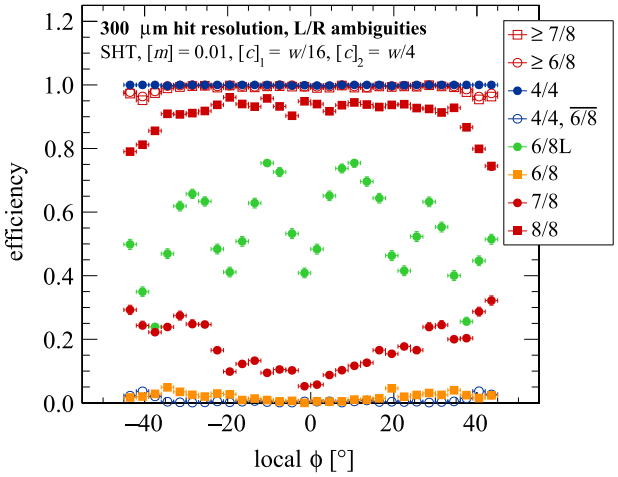


Fig. 11. Efficiency of finding track segments with different qualities in the test sample, using the proposed approach to a combined use of three SHTs, and with the benchmark histogram granularity defined by $[m] = 0.01$, $[c]_1 = w/16 = 2.625$ mm, and $[c]_2 = w/4 = 10.5$ mm, as a function of the generated track angle in the local reference frame.

negligible cost in terms of efficiency and resolution, due to the large width of the accumulation region above threshold in the parameter space. The efficiency of the 1-SL HT and the 2-SL HT with the chosen values of $[c]_1$ mm and $[c]_2$ mm is shown for different choices of the $[m]$ bin size in Fig. 9. The efficiency of finding track segments with quality 4/4 is preserved up to $[m] = 0.01$, being even slightly higher in the range $[m] = 0.004$ to $[m] = 0.005$. The higher efficiency in this range is due to two effects: small values of $[m]$ do not allow for enough accumulation in the HT parameter space, being entries too scattered around; on the other hand, when $[m]$ is too large, the accumulation is shared between neighboring bins, as already anticipated. A comparison with the case without ambiguities shows that 4–5% of high quality results are actually due to accidental accumulations of ambiguities if $[m] = 0.01$ is used. Despite being non-negligible, such a fraction can be accepted for a trigger.

(*footnote continued*)
an extremely optimized FPGA implementation of the algorithm.

The corresponding resolution of ϕ and x_0 is shown in Fig. 10. Two different behaviors are observed for qualities 4/4 and 8/8: both the ϕ and x_0 resolutions worsen with increasing $[m]$ in the 1-SL HT while they become better in the case of the 2-SL HT. The variation in case of quality 4/4 track segments is however contained to $O(1)$mrad in ϕ and $O(0.1)$ mm in x_0. The choice of the benchmark value for $[m]$ is then driven by the decrease in efficiency: $[m] = 0.01$ is chosen as it is the largest value with an efficiency higher than 0.99 for finding tracks with quality higher than 6/8. This choice allows to reduce the number of 3-bit counters needed to implement the SHT histograms by a factor 10, down to 35200 per macro-cell pair, which, although still large, is an achievable number in state-of-the-art FPGAs, yet cost unfriendly.

The final efficiency of finding track segments with different qualities, in the proposed approach to a combined use of three SHTs, and with the benchmark histogram granularity is shown in Fig. 11. Quality 4/4 track segments from both super-layers are shown inclusively. Almost all track segments can be identified with at least 7 counts, which are partly due to accidental accumulation of ambiguities close to the anode wire. The combined effect of the disposition of wires in space and chamber scanning is visible in the efficiency of 6/8L quality track segments, as peaks and dips appear periodically because of hit distribution between neighboring macro-cells. The redundant scheme allows to recover these hits in the neighboring macro-cell with a 7/8 or higher quality track segment. The chance of finding a 4/4 quality track segment that is not combined to an alignment of at least 3 hits in the complementary SL is negligible (4/4, $\overline{6/8}$ in the picture). It is worth recalling that the present study does not take into account the behavior of tracks at the edge of the chamber yet.

Residuals of estimated track parameters with respect to the generated values are shown in Fig. 12 for the benchmark case. Distributions are fit with a Gaussian function for qualities 4/4, 7/8 and 8/8, and the obtained σ is indicated in the picture.

3.6. *Effect of the single hit resolution*

The effect of different single hit resolutions is then investigated for the benchmark choice of the SHT histogram granularity. The evaluation procedure is the same described so far, changing the width of the Gaussian smearing and repeating the study for both the cases with and without left–right ambiguities. The resulting

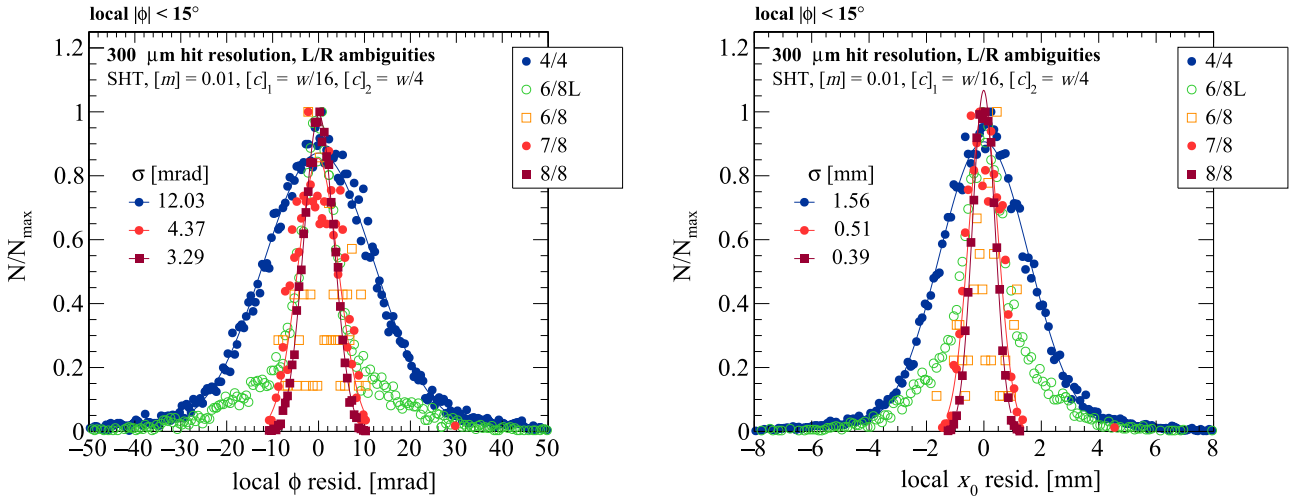


Fig. 12. Resolution of track parameters ϕ (left) and x_0 (right) for different qualities in the test sample subset of muon tracks with $-15° < \phi < +15°$, using the proposed approach to a combined use of three SHTs, and with the benchmark histogram granularity. Distributions are scaled to the same maximum, and do not account for the actual population of the sample. Error bars were removed for improved clarity of the picture.

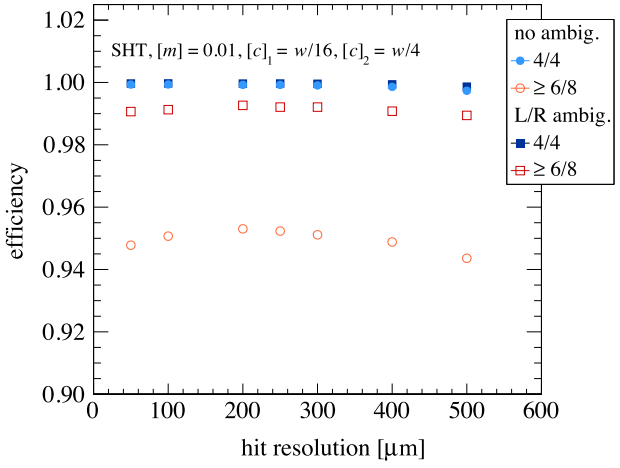


Fig. 13. Efficiency of track segment finding with the SHT in the test sample, shown as a function of the simulated hit Gaussian resolution σ, for the benchmark histogram granularity. The convention used for markers is the same as in Fig. 9.

efficiency is shown in Fig. 13. The effect of ambiguities in recovering an inefficiency of about 5% is clearly visible. The efficiency slightly depends on the single hit resolution only if left–right ambiguities are not included. Obviously, a different optimization of $[m]$ and $[c]_{1,2}$ may then be needed if a different detector, e.g. a silicon strip telescope, providing non-ambiguous measurements, is chosen.

The resolution of track parameters is shown in Fig. 14 as a function of the single hit resolution. The resolution of quality 4/4 track segments worsens with increasing uncertainty on single hit measurement, while it is slightly improving with quality 8/8 track segments. The difference between the inclusion and exclusion of left–right ambiguities can be appreciated mainly for 4/4 quality track segments: the resolution is worse in presence of ambiguities. The ϕ resolution trend shows that the direct evaluation of m from the m-bitset is better than the recalculation according to (5) only for very precise hit position measurements for the chosen histogram granularity. Indeed the quality 8/8 resolution should be similar to the one of the quality 4/4 track segments, because the $[m]$ bin size is the same, and it is better only well below $\sigma = 100\,\mu$m.

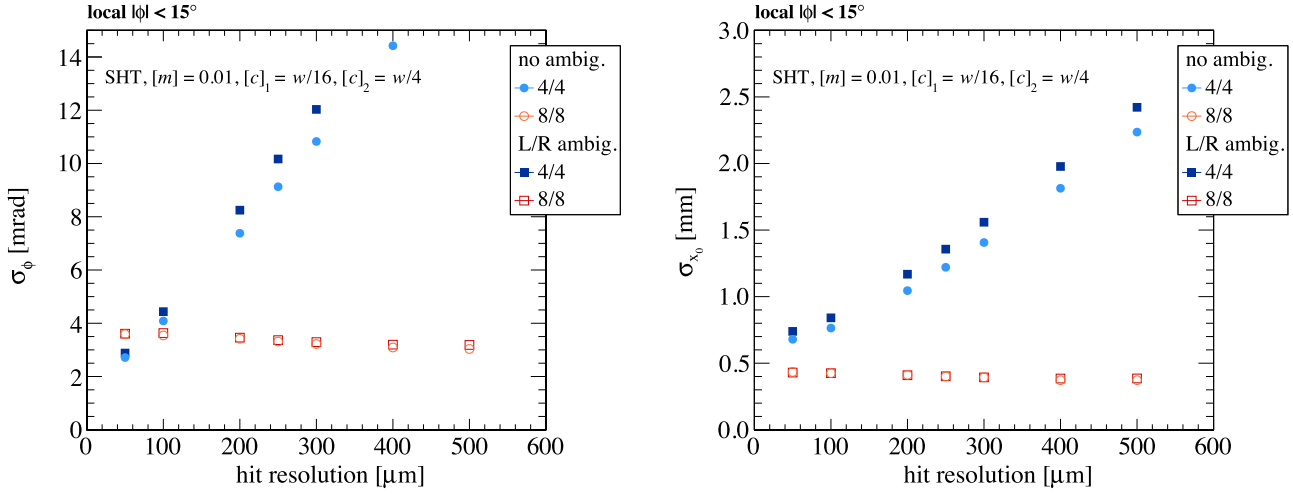


Fig. 14. Width parameter σ of the Gaussian-modeled core of the distributions of residuals of the estimated ϕ (left) and x_0 (right) using the SHT in the test sample subset of muon tracks with $-15° < \phi < +15°$, shown as a function of the simulated finite hit Gaussian resolution σ, for the benchmark histogram granularity. The convention used for markers is the same as in Fig. 10.

3.7. Effect of choices, adopted in the proposed readout, in the combination of the 1-SL HT and the 2-SL HT and in the extraction of track parameters

In this Section, the effects of the main non-trivial choices adopted throughout the algorithm in order to find the best efficiency and parameter estimate are briefly reviewed, comparing efficiencies and resolutions for alternative possibilities in the benchmark case. In particular, the investigation is focussing on:

a) the effect on the efficiency of shifting the histogram center by half c bin with respect to the neighboring and partially superimposed macro-cell,
b) the effect on resolution of finding clusters in the m-bitset rather than averaging on the whole m-bitset,
c) the effect on efficiency and resolution of the consistency check between the m resulting from the m-bitset and the recalculated value $m' = (c_{in} - c_{out})/d$ according to (5),
d) the effect on resolution of choosing m' instead of the value of m resulting from the m-bitset.

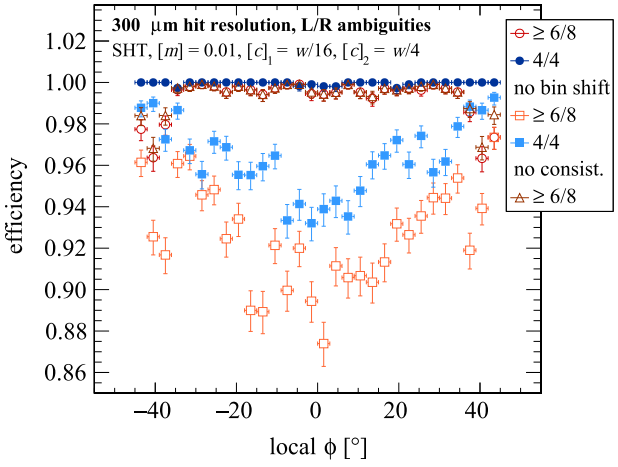


Fig. 15. Efficiency of finding track segments with different qualities, using the proposed approach to a combined use of three SHTs in the test sample, and with the benchmark histogram granularity defined by $[m] = 0.01$, $[c]_1 = w/16 = 2.625$ mm, and $[c]_2 = w/4 = 10.5$ mm, as a function of the generated track angle in the local reference frame. The complete algorithm (round markers) is compared to two simplified versions, dropping the features listed as (*a*) (square markers) and *c)* (triangular markers) in the list of Section 3.7. Full markers show the efficiency to find 4/4 quality track segments, hollow markers show the efficiency of finding track segments with qualities higher than 6/8.

Fig. 15 shows the effect on the efficiency of features (*a*) and (*c*). The first choice is clearly very important to preserve the efficiency with the chosen benchmark histogram granularity. The consistency check has only a very small effect at large values of $|\phi|$.

Residuals of estimated track parameters with respect to the generated values are shown in Fig. 16. The only visible difference is the worsening of the resolution of ϕ for 8/8 quality track segments of feature (*d*), and consistent with the rough estimates of (7) and (8).

4. Towards a compact Hough Transform

4.1. Parameter reduction

The algorithm described in Sections 3.2–3.4 addresses and partially solves some of the issues presented by a fast hardware implementation of the HT for triggering purposes, but still leaves the open problem of a parallel looping on the HT histogram columns to identify the meaningful bins. Furthermore, the number of 3-bit counters needed for the SHT implementation of the proposed readout to identify a track segment in the benchmark case is still very high, even after optimization. The path to a compact implementation of the HT that allows to directly combine the two 1-SL HT with the 2-SL HT, without dumping them to m-bitsets, is the parameter reduction. The method is based on the exploitation of pairs of measurements, as suggested by Ballard [17], without really providing a possible implementation, which is instead described by Ben-Tzvi and Sandler [18]. There are relevant differences between the method proposed in this paper and the latter one, which are discussed at the end of Section 4.2.

Let us consider a pair of hits $(\mathbf{x_i}, \mathbf{x_j})$: according to (3), this pair of hits defines two lines in the parameter space

$$\begin{cases} c = x_i + m \cdot z_i \\ c = x_j + m \cdot z_j \end{cases}. \tag{9}$$

At the correct accumulation point in the parameter space, the (m,c) pair fulfilling the previous condition is unique. Hence, the difference between the intercepts of the two transformed points

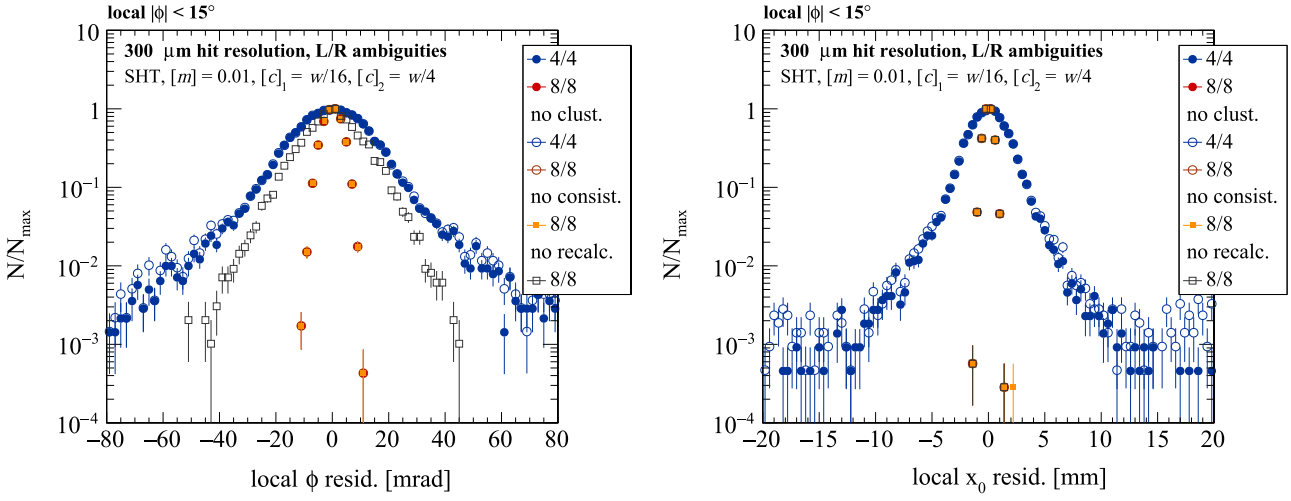


Fig. 16. Residuals of the angle ϕ (left) and intercept x_0 (right) estimated using the SHT in the test sample subset of muon tracks with $-15° < \phi < +15°$, for different qualities, and with the benchmark histogram granularity defined by $[m] = 0.01$, $[c]_1 = w/16 = 2.625$ mm, and $[c]_2 = w/4 = 10.5$ mm. The complete algorithm (full round markers) is compared to three simplified versions, dropping the features listed as b) (hollow round markers), c) (full square markers) and d) (hollow square markers) in the list of Section 3.7. Distributions are scaled to the same maximum, and do not account for the actual population of the sample.

should be null at the correct m:

$$\Delta c = c_i - c_j = (x_i - x_j) + m\cdot(z_i - z_j) \approx 0 \quad (10)$$

where c_i and c_j are the intercepts evaluated using $\mathbf{x}_i$ and $\mathbf{x}_j$ and the candidate m. The actual difference, despite being expected to be null, might not be exactly zero, because of the discretization of parameter space and the physical processes scattering the hit.[4]

4.2. Compact HT histogram and extraction of track parameters

A compact implementation of the HT (CHT) can be designed around $(m, \Delta c)$ histograms, centered on $\Delta c = 0$. Furthermore, if the value of $[\Delta c]$ is large enough to accommodate also the almost null differences, the CHT histogram can be one-dimensional, thus reducing by a large amount the size of the histogram, which is already close to be an m-bitset such as described in Section 3.3. No column dumping is required in this case, as a mere threshold is enough to transform the CHT histogram into the m-bitset.

This compact and looping-less implementation comes at the price of more complicated routing of signals which must be duplicated in order to account for all possible pairs of hits in different layers. Thresholds cannot be now defined by the number of aligned hits, as the CHT counts pairs of hits: 3 (4) aligned hits in a super-layer correspond to 3 (6) valid pairs, while 6 aligned hits in the 2-SL HT correspond to 15 valid pairs. Qualities are then defined according to Table 2, which is similar to Table 1, and the 2-SL HT histogram must be implemented with 5-bit counters.

Table 2
Quality flags for the combined use of the 1-SL HT and the 2-SL HT, analogous to the ones defined while evaluating the SHT, in Table 1, for the compact implementation of the proposed algorithm.

Quality flag	Threshold 1	Threshold 2	Threshold C
8/8	6	6	15
7/8	3	6	15
6/8	3	3	15
6/8L	$\overline{3}$	6	15
4/4	–	6	–

The procedure to extract the track slope after combining m-bitsets together and finding clusters is similar to the one used with the SHT. Instead in this case, the intercept c cannot be obtained from the SHT histogram, and must be recalculated from the input hits. The same procedure employed in the SHT is applied. It is worth noticing that the value of $[c]$ used for the evaluation of the intercept with a one-dimensional histogram can be different from the value of $[\Delta c]$ used to define the compatibility with zero of the difference between intercepts according to (10).

In the work of Ben-Tzvi and Sandler [18], the polar parameter of a straight line θ_0 is computed from a pair of measurements, and a two-dimensional histogram is filled by computing the radial parameter r_0 for each obtained value of θ_0, avoiding the typical sampling procedures of Hough Transform implementations. In our proposal, instead, each pair of measurements is checked for consistency with the value of m in the sampling loop: only one-dimensional histograms are filled, while the second line parameter c is extracted after the best hypothesis for m is selected.

[4] This procedure can actually be iterated in case of curves described by N parameters: N-ples of measurements can be used to reduce the number of parameters down to one when it is possible to write a function of N measurements $\{\mathbf{x}_i\}$ and one parameter p, $f(\{\mathbf{x}_i\}, p) = 0$, being (10) a special case of such functions.

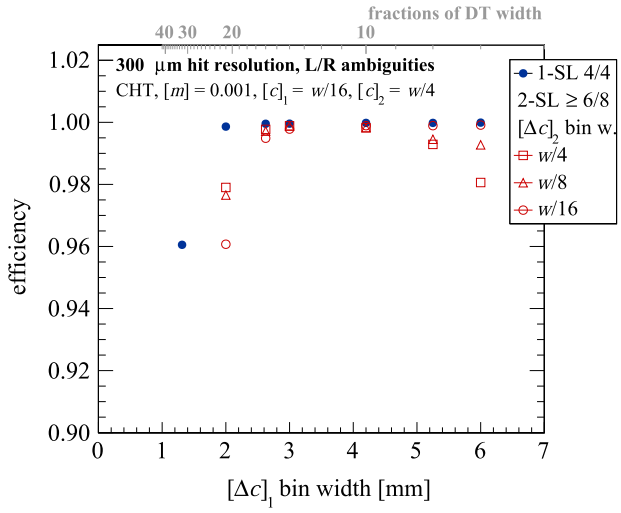


Fig. 17. Efficiency of track segment finding with the CHT in the test sample, shown as a function of $[\Delta c]_1$, for $[m] = 10^{-3}$, $[c]_1 = w/16 = 0.2625$ mm, and $[c]_2 = w/4 = 10.5$ mm. The hit resolution is simulated by a Gaussian smearing with $\sigma = 0.3$ mm. Full dots show the efficiency to find 4/4 quality track segments, hollow markers show the efficiency of finding track segments with qualities higher than 6/8, for different values of $[\Delta c]_2$.

4.3. Evaluation of the benchmark histogram granularity

The optimization of the histogram granularity is more involved for the CHT, as both $[\Delta c]$ and $[c]$ concur in determining the efficiency of track segment finding and the resolution of track parameters. Hence, the following approach is adopted:

a) given that the efficiency is dominated by $[\Delta c]_1$ and $[\Delta c]_2$, as the extracted value of c is used only in higher order corrections, $[m]$ is fixed to 0.001; $[c]_1$ and $[c]_2$ are fixed to the values coming from the optimization of the SHT; the efficiency and the resolution are evaluated as a function of $[\Delta c]_1$ and $[\Delta c]_2$[5];
b) it can be expected that the resolution depends only weakly on $[\Delta c]_1$ and $[\Delta c]_2$, then $[\Delta c]_1$ and $[\Delta c]_2$ are fixed to the values providing maximum efficiency and the resolution will be evaluated as a function of $[c]_1$ and $[c]_2$;
c) the choice of $[\Delta c]_1$ and $[\Delta c]_2$ is *a posteriori* verified for the new set of $[c]_1$ and $[c]_2$, if different from initial hypothesis;
d) $[m]$ is then increased as in the evaluation of the SHT.

It is worth noticing that, in the CHT implementation, $[c]_2$ is actually used only for the higher-order optional consistency check between c and x_0, as mentioned in Section 3.4, which is currently dropped in the running algorithm. The evaluation with respect to $[c]_2$ is included at this first stage, although it is expected that changing $[c]_2$ has no impact on the outcome.

The same test sample described in Section 3.5 was used for the evaluation of the CHT performance. Results of the evaluation are shown in Fig. 17. The efficiency is increasing with both $[\Delta c]_1$ and $[\Delta c]_2$. In the case of quality 4/4, the efficiency reaches a plateau at about 0.999 for $[\Delta c]_1 \geq w/16 = 2.625$ mm while, in the case of qualities higher than 6/8, it is higher than 0.99 for $[\Delta c]_1 \geq w/16 = 2.625$ mm and for $[\Delta c]_2 \geq w/10 = 4.2$ mm, up to a maximum of about 0.997. The efficiency of qualities higher than 6/8 starts decreasing at very large values of $[\Delta c]_2$. This is due to the

[5] $[m]$ is fixed to 0.001 in order not to be biased by SHT optimization result, being m the main parameter. Fixing $[c]_1$ and $[c]_2$ to the values coming from the optimization of the SHT is not expected to affect the efficiency, as confirmed *a posteriori* (see Fig. 19), while their impact on the resolution is being evaluated in the following steps of this procedure.

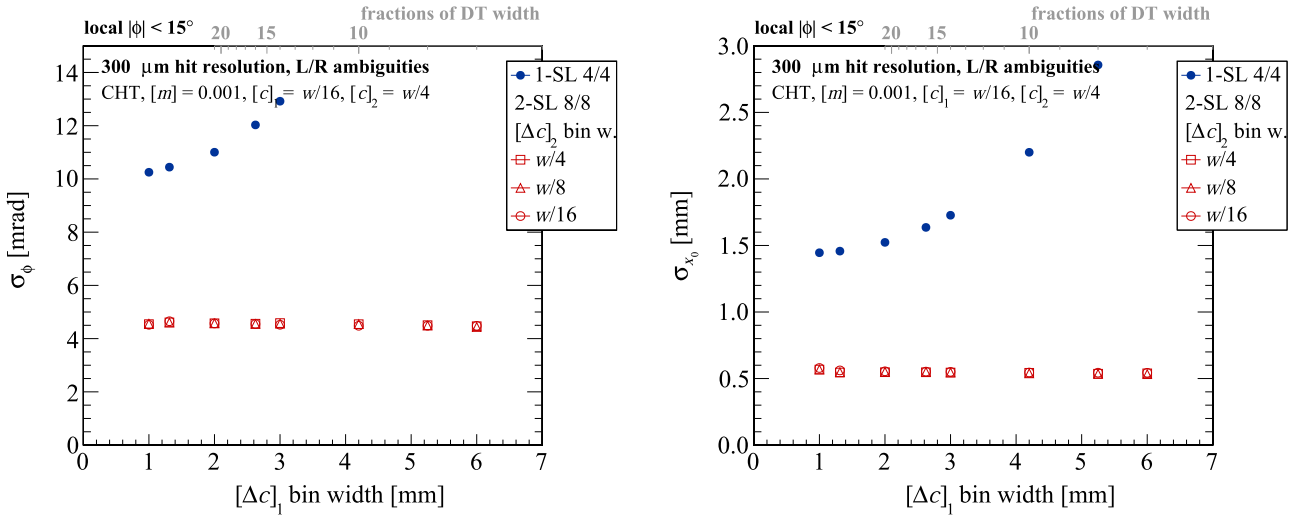


Fig. 18. Width parameter σ of the Gaussian-modeled core of the distributions of residuals of the estimated ϕ (left) and x_0 (right) using the CHT in the test sample subset of muon tracks with $-15° < \phi < +15°$, shown as a function of $[\Delta c]_1$, for $[m] = 10^{-3}$, $[c]_1 = w/16 = 0.2625$ mm, and $[c]_2 = w/4 = 10.5$ mm. The hit resolution is simulated by a Gaussian smearing with $\sigma = 0.3$ mm. Full dots show the resolution of 4/4 quality track segments, hollow markers show the resolution of 8/8 quality track segments, for different values of $[\Delta c]_2$.

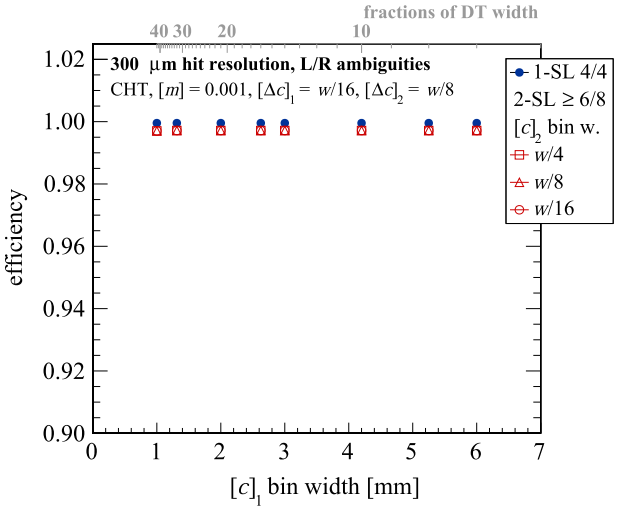


Fig. 19. Efficiency of track segment finding with the CHT in the test sample, shown as a function of $[c]_1$, for $[m] = 10^{-3}$, $[\Delta c]_1 = w/16 = 0.2625$ mm, and $[\Delta c]_2 = w/8 = 5.25$ mm. The hit resolution is simulated by a Gaussian smearing with $\sigma = 0.3$ mm, and left–right ambiguities are included for each smeared hit. The convention used for markers is the same as in Fig. 5.

presence of left–right ambiguities, meaning that an excess of accumulations due to ambiguous hits drives m, as extracted from the m-bitset, off its true value by an amount that cannot cope with additional consistency checks applied downstream of it.

The resolution of the track parameters for the subset of track segments with generated ϕ in the $[-15°, 15°]$ range is shown in Fig. 18. The resolution of quality 4/4 is strongly worsening if $[\Delta c]_1$ is large, again because m is driven off its true value by ambiguities. The resolution of quality 8/8 is independent on both $[\Delta c]_1$ and $[\Delta c]_2$. Given that both m and x_0 are calculated as linear combinations of c_{in} and c_{out} in the case of quality 8/8, it is not surprising at all that their resolution does not depend neither on $[\Delta c]_1$ nor $[\Delta c]_2$. The resolution is slightly worse than the one obtained with SHT and the same values of $[c]_1$.

Chosen benchmark values for $[\Delta c]_1$ and $[\Delta c]_2$ must be large enough to account for high efficiency and small enough to limit the effect of contamination from ambiguities. While the choice of $[\Delta c]_1 \geq w/10 = 4.2$ mm and $[\Delta c]_2 = w/16 = 2.625$ mm would be the optimal on the efficiency side, a smaller value of $[\Delta c]_1$ is required to improve the resolution of quality 4/4 track segments. Therefore, $[\Delta c]_1 = w/16 = 2.625$ mm and $[\Delta c]_2 = w/8 = 5.25$ mm are chosen, despite a slight loss in efficiency.

The chosen values of $[\Delta c]_1$ and $[\Delta c]_2$ are then fixed while different values of $[c]_1$ and $[c]_2$ are tested to improve the resolution of track parameters. The efficiency is shown as a function of $[c]_1$ in Fig. 19, which shows no dependence on neither $[c]_1$ nor $[c]_2$. The effect on resolutions of changing bin sizes $[c]_1$ and $[c]_2$ is more complicated as shown in Fig. 20. The resolution does not depend on $[c]_2$, as expected, which is therefore omitted in the rest of this evaluation of the CHT. The ϕ resolution of quality 4/4 track segments does not even slightly depend on $[c]_1$ either, on the contrary of what is shown in Fig. 7 for the SHT. This could be expected as m, in lower qualities track segments, is determined directly from the m-bitsets, which depend only on $[m]$ and $[\Delta c]_{1,2}$.

The ϕ resolution of quality 8/8 track segments is approximately linear in $[c]_1$ and slightly worse than in SHT. This means that c_{in} and c_{out} are determined with lower precision. The same reason leads to a slightly worse resolution in x_0 for quality 8/8 track segments. The relative increase in σ_ϕ and σ_{x_0} is comparable. The resolution in x_0 for quality 4/4 track segments is slightly better than in the SHT case because of the more stable resolution of the m-bitsets which partly compensates for the slightly worse resolution of c_{in} and c_{out}. Given the target resolution, it is chosen to use $[c]_1 = w/32 = 1.3125$ mm. The choice of $[\Delta c]_1$ and $[\Delta c]_2$ was further confirmed with the updated value of $[c]_1$.

The bin widths chosen so far imply that a high quality track segment, in order to be identified and measured, needs 4400 3-bit counters for the two 1-SL HT compact histograms, 2200 5-bit counters for the 2-SL HT compact histogram, 256 3-bit counters for the calculation of c_{in} and c_{out}, plus the m-bitsets. The size of the CHT histograms depends only on the slope bin size $[m]$, hence the effect of a bigger $[m]$ is investigated also in this case. The efficiency of the 1-SL and the 2-SL HT with the chosen values of $[\Delta c]_1$, $[\Delta c]_2$, and $[c]_1$ is shown in Fig. 21. The efficiency of finding track segments with quality 4/4 is higher than 0.999 up to $[m] = 0.016$, and up to $[m] = 0.01$ if ambiguities are included. The efficiency combining both super-layers is even slightly higher thanks to the contribution of quality 6/8 track segments.

The resolution of ϕ and x_0 with the chosen values of $[\Delta c]_1$, $[\Delta c]_2$ and $[c]_1$ is shown in Fig. 22. While the ϕ and x_0 resolution of the

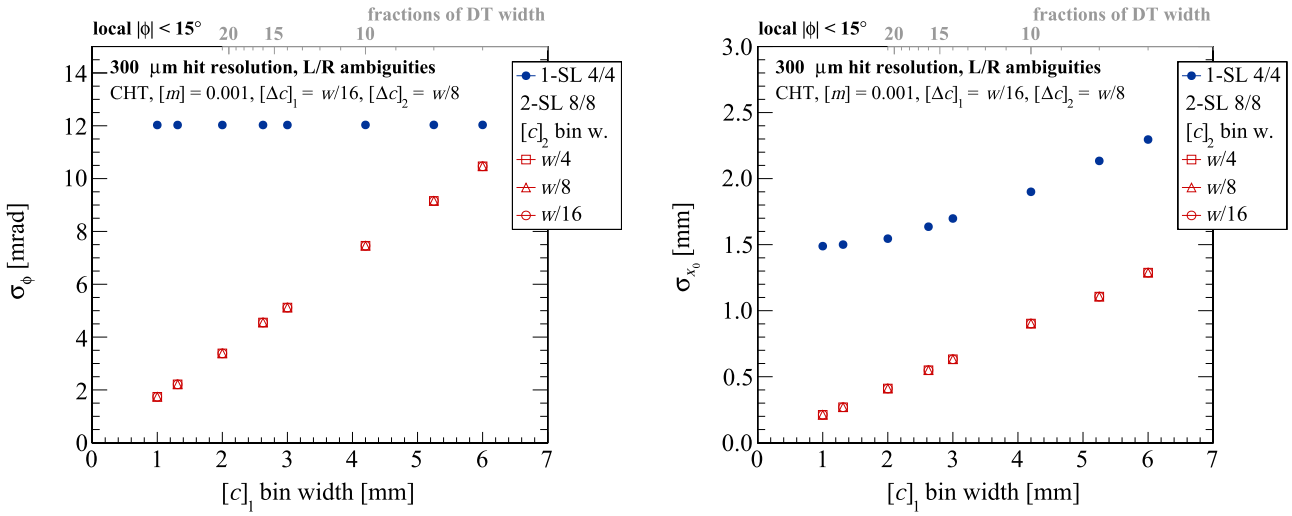


Fig. 20. Width parameter σ of the Gaussian-modeled core of the distributions of residuals of the estimated ϕ (left) and x_0 (right) using the CHT in the test sample subset of muon tracks with $-15° < \phi < +15°$, shown as a function of $[c]_1$, for $[m] = 10^{-3}$, $[\Delta c]_1 = w/16 = 0.2625$ mm, and $[\Delta c]_2 = w/8 = 5.25$ mm. The hit resolution is simulated by a Gaussian smearing with $\sigma = 0.3$ mm, and left–right ambiguities are included for each smeared hit. The convention used for markers is the same as in Fig. 7.

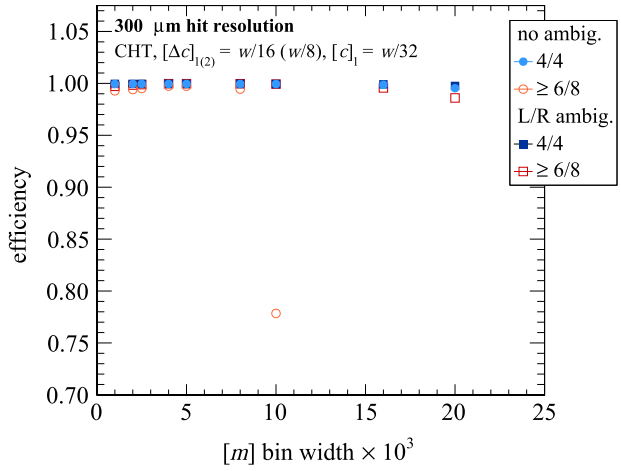


Fig. 21. Efficiency of track segment finding with the CHT in the test sample, shown as a function of $[m]$, for $[\Delta c]_1 = w/16 = 2.625$ mm, $[\Delta c]_2 = w/8 = 5.25$ mm, and $[c]_1 = w/32 = 1.3125$ mm. The hit resolution is simulated by a Gaussian smearing with $\sigma = 0.3$ mm, and left–right ambiguities are included for each smeared hit. The convention used for markers is the same as in Fig. 9.

quality 4/4 track segments is worsening with larger values of $[m]$, the resolution of the quality 8/8 track segments is constant. The variation in the case of quality 4/4 track segments is however contained to $O(1)$ mrad in ϕ and $O(0.1)$ mm in x_0, as in the SHT case. Therefore it seems to be possible to safely choose $[m] = 0.01$ also for the CHT case, reducing by a factor 10 the number of counters needed to implement the histograms. However, the high efficiency at $[m] = 0.01$ of the 2-SL HT is partly due to accidental accumulations of left–right ambiguities. In fact, if no ambiguities are used, the efficiency of segments with qualities higher than 6/8 drops from approximately 0.993 down to less than 0.8, between $[m] = 0.008$ and $[m] = 0.01$, while the efficiency of quality 4/4 track segments is higher than 0.995 in the whole $[m]$ range. The contribution of accidental accumulations of ambiguities is more important at high ϕ, as the 2-SL HT has less discrimination power between good and fake hits when the slope is measured with pairs of hits in different super-layers. A high efficiency obtained with a 20% contribution from fake hits is not acceptable, hence $[m] = 0.008$ is chosen, reducing by a factor 8 the number of counters needed to implement the histograms. In this configuration, a track segment is identified with a minimum load of the programmable devices of 275 counters per histogram, 256 3-bit counters for the calculation of c_{in} and c_{out}, plus 5 m-bitsets of the same length as the histograms. Therefore the implementation of a CHT-based Level 1 trigger can in principle be done in state-of-the-art FPGAs, at manageable cost even for large area systems such as the CMS Drift Tube detector, requiring few hundred devices.

The efficiency of finding track segments with different qualities, using the proposed implementation of a CHT combining three different histograms, with benchmark granularity defined by $[m] = 0.008$, $[\Delta c]_1 = w/15 = 2.625$ mm, $[\Delta c]_2 = w/8 = 5.25$ mm, and $[c]_1 = w/32 = 1.3125$ mm, is shown in Fig. 23. The definitions are similar to the ones used in Fig. 11 in Section 3.5. The large value of $[m]$ allows for a high efficiency, although with a non-negligible contribution of accidental accumulations of ambiguities in the 2-SL HT.

Residuals of estimated track parameters with respect to the generated values are shown in Fig. 24 for the benchmark case. The same effects already discussed for the SHT are observed.

4.4. Effect of the single hit resolution

The performance of the CHT is investigated also as a function of the single hit resolution for the benchmark granularity, as already discussed in Section 3.6 for the SHT. The effect of the left–right ambiguities mentioned in the previous paragraph is clearly visible in Fig. 25, showing the efficiency as a function of the applied Gaussian smearing width. The efficiency is stable up to a single hit resolution $\sigma = 0.4$ mm.

The resolution of the track parameters is shown in Fig. 26 as a function of the single hit resolution. The resolution of the quality 4/4 track segments becomes worse with increasing uncertainty on single hit measurement. The effect is much smaller in the case of the quality 8/8 track segments, although the resolution is worsening, in contrast with what observed in the case of the SHT. The difference between the inclusion and exclusion of left–right ambiguities can be appreciated mainly for 4/4 quality track segments, as the resolution is clearly worse in presence of ambiguities.

4.5. Effect of choices adopted in the proposed readout, in the combination of the 1-SL HT and the 2-SL HT, and in extraction of track parameters

The same investigation about the non-trivial algorithm features

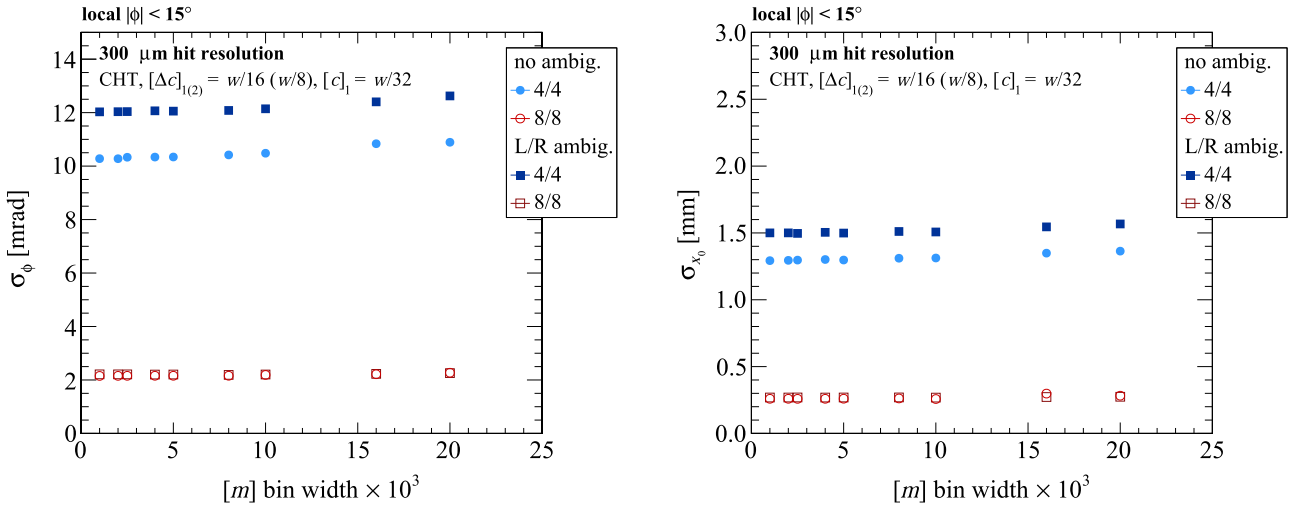


Fig. 22. Width parameter σ of the Gaussian-modeled core of the distributions of residuals of the estimated ϕ (left) and x_0 (right) using the CHT in the test sample subset of muon tracks with $-15° < \phi < +15°$, shown as a function of $[m]$, for $[\Delta c]_1 = w/16 = 2.625$ mm, $[\Delta c]_2 = w/8 = 5.25$ mm, and $[c]_1 = w/32 = 1.3125$ mm. The hit resolution is simulated by a Gaussian smearing with $\sigma = 0.3$ mm, and left–right ambiguities are included for each smeared hit. The convention used for markers is the same as in Fig. 10.

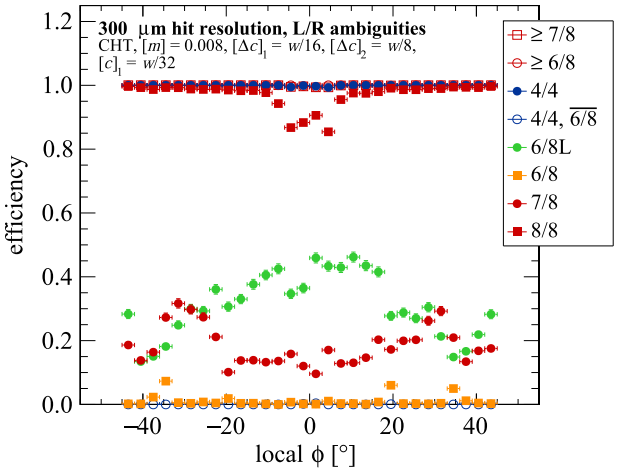


Fig. 23. Efficiency of finding track segments with different qualities in the test sample, using the proposed approach to a combined use of three CHTs, and with the benchmark histogram granularity defined by $[m] = 0.008$, $[\Delta c]_1 = w/16 = 2.625$ mm, $[\Delta c]_2 = w/8 = 5.25$ mm, and $[c]_1 = w/32 = 1.3125$ mm, as a function of the generated track angle in the local reference frame.

already presented in Section 3.7 for the SHT is described herein. Fig. 27 shows the effect of features (a) and (c) on the efficiency. Shifting the histogram center by half $[c]$ bin has effect only on the extraction of track parameters. It can be then accounted as a higher order effect on the efficiency, visible only at small angles in high quality track segments, due to a failure of the consistency check making use of already calculated values of c_{in} and c_{out}. The consistency check between on m has a minor effect.

Residuals of the estimated track parameters with respect to the generated values are shown in Fig. 28. We chose to show the plots for $|\phi|$ ranging between 15° and 30°, where the effects are more important. The resolution in ϕ of the 8/8 quality track segments is actually much worse if m is assigned directly from the m-bitset rather than being recalculated as $m' = (c_{in} - c_{out})/d$, and consistent with the rough estimates of (7) and (8). The effect of cluster finding is clearly visible in the quality 4/4 track segments, as it definitely reduces long tails in ϕ residuals, while it does not affect the measurement of c_{in} and c_{out}, resulting in no effect on the quality 8/8 track segments. There is a small effect of the m consistency check on the x_0 residuals: a badly measured slope can result in c accumulating far off the true value. This can be visualized in a SHT histogram, where a remarkably wrong value of m

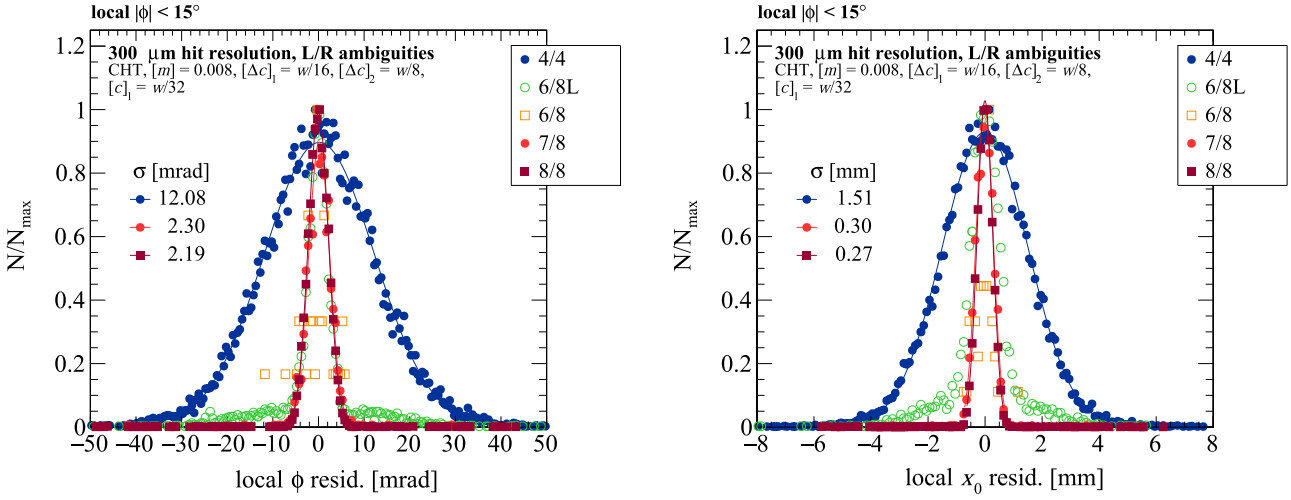


Fig. 24. Resolution of track parameters ϕ (left) and x_0 (right) for different qualities in the test sample subset of muon tracks with $-15° < \phi < +15°$, using the proposed approach to a combined use of three CHTs, and with the benchmark histogram granularity. Distributions are scaled to the same maximum, and do not account for the actual population of the sample. Error bars were removed for improved clarity of the picture.

corresponds to a vertical slice with low maxima given by the ambiguous hits. As a direct effect, the distribution of x_0 residuals has multi-peaked tails starting about 1 cm off the center. This is affecting in particular the quality 7/8 track segments.

5. Future perspective and concluding remarks

We have presented herein novel solutions for real-time track finding algorithms, based on the Hough Transform, that can be suitable for employment with telescopic arrays of position sensitive detectors. The proposed Compact Hough Transform, in particular, can reduce by approximately one order of magnitude the amount of resources needed for its implementation in the chosen example of the CMS Drift Tube muon detectors, with respect to a Standard Hough Transform.

We are currently planning the realization of a demonstrator, based on a state-of-the-art FPGA evaluation board, with particular care given to the reduction of the latency, which we could not evaluate properly yet. Moreover, we started to design a pre-trigger layer for the evaluation of the particle crossing time for the specific case of the CMS DT trigger.

Some specific aspects of the CMS DT trigger implementation were neglected in the present study, in particular the effect of correlated and random noise, and overlapping tracks at high luminosity, which are peculiar of the environment and will be subject of further specialized studies.

The key features of the proposed algorithm, i.e. the CHT and the combined use of groups of tracking layers, are of wider interest than the specific case of the CMS DT trigger presented herein. In particular, the proposed algorithm can be applied to the real-time analysis of charged particle tracks in medical imaging applications, e.g. proton Computed Tomography (see [19]). Such instrumentation features two sets of solid state tracking devices in front and behind of the patient to be examined. The detectors are in a magnetic field free zone and cannot be cooled, being very noisy. Triggering on these devices, by the online correlation of tracks reconstructed in front and behind the patient, could help to reduce the complexity of the off-line image reconstruction saving the recording of spurious hits, hopefully reducing also the required and delivered proton dose.

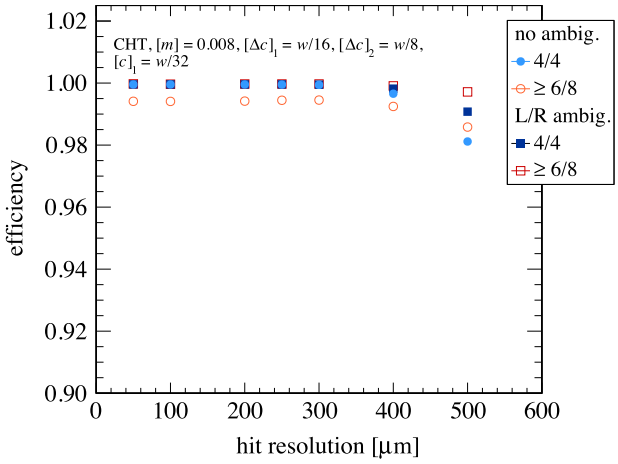


Fig. 25. Efficiency of track segment finding with the CHT in the test sample, shown as a function of the simulated finite hit Gaussian resolution, for the benchmark histogram granularity. The convention used for markers is the same as in Fig. 9.

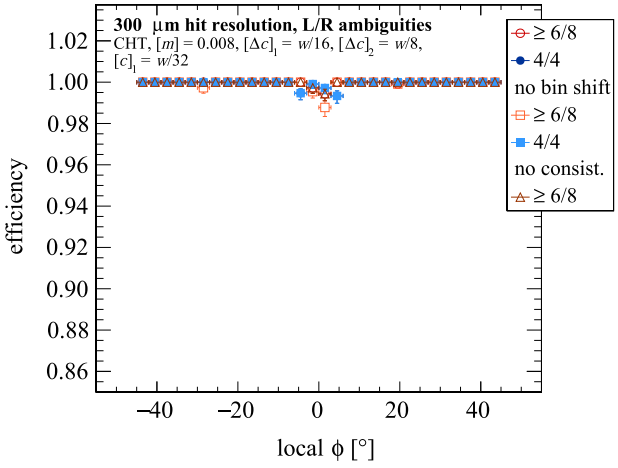
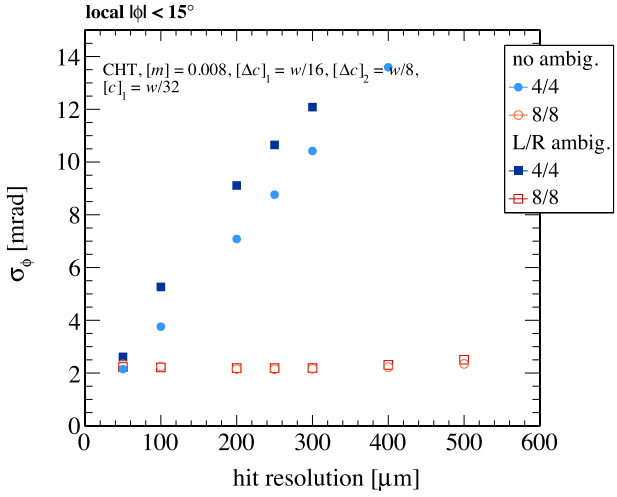


Fig. 27. Efficiency of finding track segments with different qualities, using the proposed approach to a combined use of three CHTs in the test sample, and with the benchmark histogram granularity, as a function of the generated track angle in the local reference frame. The complete algorithm is compared to two simplified versions, with the same convention as in Fig. 15.



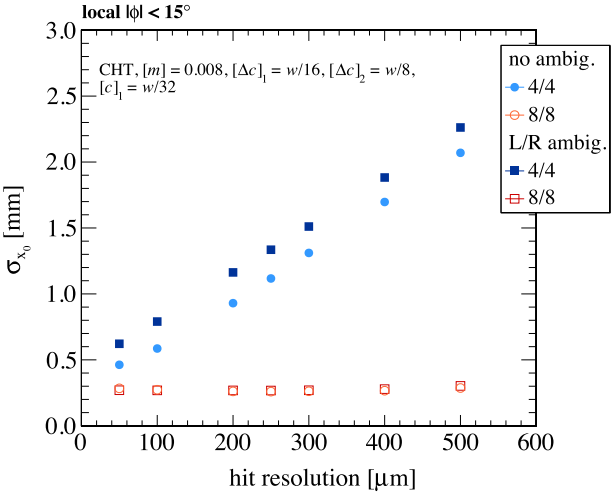


Fig. 26. Width parameter σ of the Gaussian-modeled core of the distributions of residuals of the estimated ϕ (left) and x_0 (right) using the CHT in the test sample subset of muon tracks with $-15° < \phi < +15°$, shown as a function of the simulated finite hit Gaussian resolution σ, for the benchmark histogram granularity. The convention used for markers is the same as in Fig. 10.

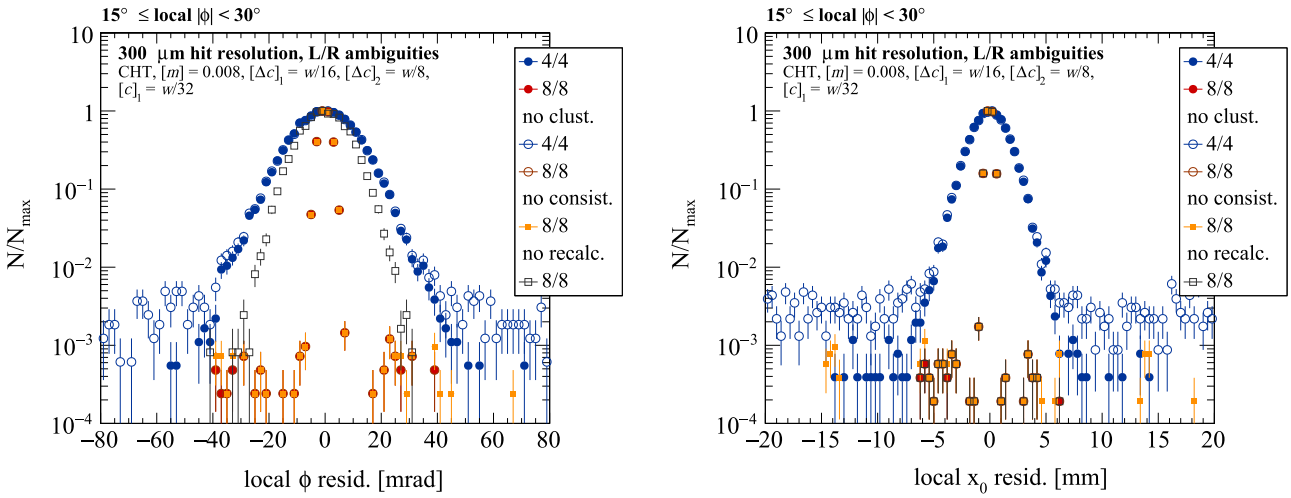


Fig. 28. Residuals of the angle ϕ (left) and intercept x_0 (right) estimated using the SHT in the test sample subset of muon tracks with $15° \leq |\phi| < 30°$, for different qualities, and with the benchmark histogram granularity. The complete algorithm is compared to three simplified versions, with the same convention as in Fig. 16.

References

[1] P.V.C. Hough, Methods and Means for Recognizing Complex Patterns, U.S. Patent 3,069,654, 1962.

[2] J. Illingworth, J. Kittler, A survey of the hough transform, Comp. Vis. Graph. Image Proc. 44 (1988) 87–116, http://dx.doi.org/10.1016/S0734-189X(88)80033-1.

[3] P. Mukhopadhyay, B.B. Chaudhuri, A survey of hough transform, Pat. Rec. 48 (2015) 993–1010, http://dx.doi.org/10.1016/j.patcog.2014.08.027.

[4] C.V. Cheshkov, Fast hough transform track reconstruction for the ALICE TPC, Nucl. Instrum. Methods A 566 (2006) 35–39, http://dx.doi.org/10.1016/j.nima.2006.05.036.

[5] M. Véstias, H. Néto, Trends of CPU, GPU and FPGA for high-performance computing, in: Proceedings of the 24th International Conference on Field Programmable Logic and Applications (FPL). ⟨http://dx.doi.org/10.1109/FPL.2014.6927483⟩, 2014.

[6] L. Shannon, et al., Technology scaling in FPGAs: trends in applications and architectures, in: Proceedings of IEEE 23rd Annual International Symposium on Field-Programmable Custom Computing Machines (FCCM). ⟨http://dx.doi.org/10.1109/FCCM.2015.11⟩, 2015.

[7] R. Boada Gardenyes, Trends and Patterns in ASIC and FPGA use in Space Missions and Impact in Technology Roadmaps of the European Space Agency (Master's thesis), TU Delft, 2012.

[8] D. Sabes, L1 track triggering with associative memory for the CMS HL-LHC tracker, JINST 9, C11014, Proceedings of the Workshop on Intelligent Trackers 2014. ⟨http://dx.doi.org/10.1088/1748-0221/9/11/C11014⟩, 2014.

[9] D. Cieri, et al., L1 track finding for a time multiplexed trigger, Proceedings of the 13th Pisa Meeting on Advanced Detectors, Elba 2015, Nucl. Instrum. Methods A 824 (2016) 268–269, http://dx.doi.org/10.1016/j.nima.2015.09.117.

[10] The CMS Collaboration, The CMS experiment at the CERN LHC, JINST 3 (2008) S08004, http://dx.doi.org/10.1088/1748-0221/3/08/S08004.

[11] P. Arce, et al., Bunched beam test of the CMS drift tubes local muon trigger, Nucl. Instrum. Methods A 534 (2004) 441–485, http://dx.doi.org/10.1016/j.nima.2004.06.169.

[12] C. Albajar, et al., Electromagnetic secondaries in the detection of high energy muons, Nucl. Instrum. Methods A 364 (1995) 473–487, http://dx.doi.org/10.1016/0168-9002(95)00474-2.

[13] R.O. Duda, P.E. Hart, Use of the hough transformation to detect lines and curves in pictures, Comm. ACM 15 (1972) 11–15, http://dx.doi.org/10.1145/361237.361242.

[14] S. Agostinelli, et al., GEANT4 – a simulation toolkit, Nucl. Instrum. Methods A 506 (2003) 250–303, http://dx.doi.org/10.1016/S0168-9002(03)01368-8.

[15] J. Allison, et al., Use of the hough transformation to detect lines and curves in pictures, IEEE Trans. Nucl. Sci. 53 (2006) 270–278, http://dx.doi.org/10.1109/TNS.2006.869826.

[16] The CMS Collaboration, The performance of the CMS muon detector in proton-proton collisions at $\sqrt{s} = 7$ TeV at the LHC, JINST 8 (2013) P11002, http://dx.doi.org/10.1088/1748-0221/8/11/P11002.

[17] D. Ballard, Generalizing the hough transform to detect arbitrary shapes, Pat. Rec. 13 (1981) 111–122, http://dx.doi.org/10.1016/0031-3203(81)90009-1.

[18] D. Ben-Tzvi, M. Sandler, A combinatorial Hough Transform, Pat. Rec. Let. 11 (1990) 167–174, http://dx.doi.org/10.1016/0167-8655(90)90002-J.

[19] S. Mattiazzo, et al., Advanced proton imaging in computed tomography, Rad. Prot. Dosim. 166 (2015) 388–392, http://dx.doi.org/10.1093/rpd/ncv197.