Finding the needle in the haystack: a charmonium trigger for the CBM experiment

T Ablyazimov, V Friese and V Ivanov

1GSI Helmholtzzentrum für Schwerionenforschung GmbH, Darmstadt, Germany
2LIT Joint Institute for Nuclear Researches, Dubna, Russia
3National Research Nuclear University “MEPhI”, Moscow, Russia

E-mail: t.ablyazimov@gsi.de

Abstract. Charmonium is one of the most interesting, yet most challenging observables for the CBM experiment. CBM will try to measure charmonium in the di-muon decay channel in heavy-ion collisions close to or even below the kinematic threshold for elementary interactions. The expected signal yield is consequently extremely low - less than one in a million collisions. CBM as a high-rate experiment shall be able to cope with this, provided a suitable software trigger can be implemented for online data selection. Since the latter will be performed exclusively on CPU, the performance of the algorithm is crucial for the maximal allowed interaction rate – and thus the sensitivity – and/or for the size of the CBM online cluster FLES (First-Level Event Selector). In this report we discuss the CBM charmonium trigger, its implementation on the FLES, and its performance.

1. Introduction

CBM is a heavy-ion experiment currently under construction at the FAIR centre in Darmstadt, Germany. It will investigate strongly interacting matter at high net-baryon densities by nuclear collisions in the energy range 2 – 11 GeV/nucleon. One of the most challenging observables for CBM is the measurement of charmonium close to its production threshold, where the cross section is extremely small - less than one J/ψ is expected in a million collisions. Consequently, CBM will be able to cope with very high interaction rates of up to 10^7 collisions per second. Because of the high rates and complex trigger signatures, which are difficult to implement in hardware, CBM chose a readout concept based on free-running, self-triggered front-end electronics. As a consequence, the necessary online data selection must be performed in software. A data rate suppression by a factor of at least 300 is required.

Charmonium will be detected through its decay into muon pairs. The identification of muons in CBM will be performed by the muon system MUCH, an active absorber system consisting of alternating absorber and detection layers, placed downstream of the dipole magnet hosting the main tracking system STS (Fig. 1, left). It contains five absorbers of varying thickness and four detection stations between the absorbers, each made of three layers of GEM detectors. Behind the last absorber, the TRD detector with four MWPC layers provides additional coordinate measurements (Fig. 1, right). The detection principle is that muons pass through the system while hadrons and electrons are absorbed. The active absorber setup allows to follow the muon trajectories through the absorber system and to connect them to the tracks found in the main tracking system STS inside the magnetic dipole field, which determines their momenta.
The software task is to filter the raw data stream with respect to candidates for $J/\psi$ events, characterised by two muons of opposite charges passing through the absorber system. The algorithm must deliver the required suppression in data rate and be fast enough to be applied in real-time. The approach described here uses data from the MUCH and TRD systems only, i.e., it does not require prior reconstruction of tracks in the STS system. The selection criteria for $J/\psi$ candidates are based on geometrical properties of the muon trajectories in the MUCH system; thus, the first step is the reconstruction of these trajectories from the measured hit data.

2. Track reconstruction in MUCH

2.1. Strategy

Muons from $J/\psi$ decays practically originate from the primary vertex, which we can approximate by the target centre. They have preferentially large momenta because of the large Q value of the decay. So, they pass the entire muon system and are registered in the TRD. Finally, the muon system is placed outside of the dipole magnet; only the first stations experience a small stray magnetic field. Thus, the trajectories inside the system can be approximated by straight lines. These properties determine the strategy for track reconstruction detailed in the following: we look for tracks registered in all stations of the system pointing back to the target centre.

A natural linear discrete model for the trajectory deducible from a set of points (hits) is a broken line. Each segment of such a line connects two hits belonging to adjacent stations. So, we need one space-point in each station as input for the track search. This means that in each station, at least one hit out of three detection layers is required. In the current implementation, we use the hits from the middle layer of each station, neglecting the measurements in the other two layers. Only for the TRD, hits from all four layers are considered.

The track reconstruction approach is inspired by the model often referred as Cellular Automaton (CA). The basic idea is to reduce combinatorics by rejecting most combinations of input data based on local criteria only, without involving extensive input data patterns. In our case track segments, connecting hits on adjacent stations, are considered cells of the Cellular Automaton.

In the ideal CA all segment computations would be done independently. It would be
impractical from the performance point of view, however, not to take into account the fact that the amount of hits registered by a detecting station rapidly decreases with distance from the target. So we employed an approach which could be called “asymmetric cellular automaton”, in which the information from the less loaded stations is used to reduce combinatorics on the others. It appears natural to start reconstruction of the $J/\psi$ decay muons from the last station (TRD). We choose instead to start from the 4th MUCH station, because we would like the reconstruction routine to be applicable not only for charmonium, but also for low-mass muon pairs, the partners of which are not likely to pass through the last absorber. So, our track reconstruction is done in two steps: 1. reconstruct a track starting from station 4 using the CA approach; 2. check whether the track can be extended to the TRD station. Only the latter tracks are used for triggering on $J/\psi$.

2.2. Building of segments and segment chains

The process begins from the station 4 and continues until the segment ends reaches station 1. We classify a hit as “usable” if it is the end-point of a segment of a previous step, except for the first step where all hits in station 4 are considered usable. In each step, a usable hit is connected to the target centre by a straight line (Fig. 2). The intersection of this line with the adjacent upstream station is considered the extrapolation of the considered hit. Around this extrapolation, a rectangular validation area is defined by the Highland formula for multiple scattering

$$\theta_0 = 13.6 \frac{MeV}{\beta p c^2} z \sqrt{\frac{l}{X_0}} (1 + 0.038 ln(\frac{l}{X_0})),$$

(1)

where $\theta_0$ is the scattering angle, $\beta$ and $p$ are the particle velocity and momentum, respectively, and $l/X_0$ is the traversed material budget in units of radiation length. Since no information on the momentum of the tracks is available, we assume an initial energy of 3 GeV, just above the threshold energy to traverse all absorbers, and account for the subsequent energy loss in the absorbers using the Bethe-Bloch formula.

Figure 2. Building of segments. The shaded areas indicate the validation regions around the line connecting a hit with the target centre (dotted green lines). Only hits in the validation area of hits in the downstream neighbour station are considered. The red lines indicate valid segment chains.

A track candidate is defined as a chain of segments connecting station 4 and station 1. Segments not ending at station 1 are rejected.
2.3. Chain competition
Each hit on the start MUCH station can become the root of number of valid chains during reconstruction. In such cases, the chain having the least $\chi^2$ is chosen and considered as reconstructed track. The $\chi^2$ of track candidates is determined using the Kalman filtration technique [4].

A linear model is applied for the filtering. The XZ and YZ projections are handled independently. The state vectors are

$$\vec{p} = \begin{pmatrix} \xi \\ \tan \theta_\xi \end{pmatrix},$$

where $\xi$ denotes either x or y. The extrapolation matrix is

$$F = \begin{pmatrix} 1 & \Delta z \\ 0 & 1 \end{pmatrix},$$

where $\Delta z$ is the distance between two adjacent detecting stations. The weight and measurement matrices are

$$G = \begin{pmatrix} 1/\sigma_\xi^2 & 0 \\ 0 & 0 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$ 

Considering the multiple-scattering angle $\theta_0$ is not too large, we can use the approximation $\tan \theta_0 \approx \theta_0$. With this assumption, the covariance matrix becomes

$$C = \begin{pmatrix} \theta_0^2 z^2/3 & \theta_0^2 z^2/2 \\ \theta_0^2 z^2/2 & \theta_0^2 z^2 \end{pmatrix},$$

where $\theta_0$ denotes the projection of the multiple-scattering angle to the XZ and XY planes, respectively, calculated with the Highland formula, and $z$ is the thickness of the absorber between the stations.

The state vectors and covariance matrices are initialized as

$$\vec{p}_\xi = \begin{pmatrix} \xi_r \\ \xi_r/z_r \end{pmatrix}, \quad \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} = \begin{pmatrix} \sigma_\xi^2 & 0 \\ 0 & 1 \end{pmatrix},$$

where $\xi_r$ denotes either the x or y coordinate of a hit on the station from which the reconstruction starts, $\sigma_\xi$ denotes the measurement error for the corresponding hit coordinate, and $z_r$ is the z-position of the starting station.

Measurements are subsequently taken into account by updating the state vector, the covariance matrix and $\chi^2$ according to the Kalman filter formalism [4]. The final $\chi^2$ of the track is obtained when all hit coordinate measurements were included.

2.4. Track validation in the TRD
After the CA stage of the track reconstruction we have “short” tracks, which span all the MUCH stations except TRD. To find which of them have corresponding hits in the TRD, the tracks a linearly extrapolated to the first two TRD layers. A validation area in these layers is defined in the same way as used for the backward extrapolation of segments described earlier. Each combination of hits in the first and second layer within the validation area defines a straight line, which is extrapolated to the third and forth layers. If in both layers hits compatible with the extrapolation are found, the track is considered validated in the TRD and is called “long track”. The validation procedure with the TRD is illustrated in Fig 3.
3. Optimizing the track reconstruction for free-streaming input data

Online data selection implies working with data which are not subdivided into blocks corresponding to separate events, e.g., by a hardware trigger. The times of hits from different events can be interleaving. The data acquisition will deliver digital data in so-called time slices, i.e., portions of discrete measurements having their timestamps in a given time interval. A time slice contains a large number of events (O(1000) or more). Handling of such big amounts of data must be done carefully to avoid excessive runtime growth.

To account for this, our implementation tries to avoid combinations of data which certainly do not belong to the same track. We decided to use an approach inspired by the bin sort algorithm. All data are sorted into sufficiently small bins (containing only small amounts of data), which can be addressed by their coordinates in the space-time continuum.

Each detecting station of planar form (the MUCH stations are of this type) can be represented as a cuboid in the three-dimensional space of the x- and y- spatial coordinates and the time coordinate t. We subdivide it into smaller cuboids later referred to as 'bins'. This introduces straightforward and efficient procedures for the composition of and access to the data (hit) structures. The conceptual scheme is implemented as a hierarchy of unidimensional arrays in the program. The described small cuboids are referred to as X-bins. All X-bins with the same y and time coordinate are grouped into a Y-bin, and finally all Y bins with the same time coordinate define a T bin. This is schematically depicted in Fig. 4.

**Figure 3.** Finding TRD tracklets. The shaded area indicates the validation region defined by a hit in MUCH station 4 (not shown). Hit pairs in the first two TRD layers within the validation area are extrapolated to the third and fourth layer by a straight line. The yellow circle around the hits in layers 3 and 4 denote the hit measurement error.

**Figure 4.** Subdivision of coordinate and time space and hierarchy of binning
4. Event selection
Having reconstructed “long” tracks passing through the entire absorber system, we can define selection criteria for candidates for a $J/\psi \rightarrow \mu^+\mu^-$ decay. Obviously, a simultaneous pair of such tracks is required, where we define ”simultaneous” by satisfying

$$\Delta t \leq 4\sqrt{\sigma_{t_1}^2 + \sigma_{t_2}^2},$$

where $\Delta t = |t_1 - t_2|$. $t_i$ is the time of track $i$ and $\sigma_{t_i}$ the associated measurement error. We define the track time as the time of the hit in station 4 (the start station of the track finding procedure). The differences in time-of-flight of the two muon tracks from a $J/\psi$ decay is negligible compared to the measurement error, which is of the order of 5 ns.

This selection criterion alone is not sufficient to separate signal decays from the background, which mostly originates from muon pairs from pion and kaons decays. Thus, we apply additional criteria, which are based on specific properties of the charmonium decay. First, we can estimate the electric charge sign of a particle from its trajectory reconstructed in the muon system, making use of its bending in the magnetic field in front of MUCH. We calculate the angle between the projections of the first track segment (connecting stations 1 and 2) and the line connecting the hit in the second station with the target centre in the XZ plane (bending plane). This angle has different signs for positively and negatively charged particles as illustrated in Fig. 5. We require the candidate pairs to be of opposite charges.

Furthermore, we employ the fact that the large $Q$ value of the $J/\psi$ decay leads to large opening angles of the daughter muon pair. The criterion was formulated not in terms of opening angle but of a minimal distance between the tracks in the first MUCH station. The optimal cut value was determined by Monte-Carlo simulations to be 50 cm.

In summary, our $J/\psi$ software trigger requires a pair of oppositely charged long tracks detected simultaneously (within the time measurement error) with a minimum distance of 50 cm in the first muon station.

5. Results and summary
The performance figures for the developed algorithm are the efficiency to detect a $J/\psi$ decay, the data suppression factor, and the computation speed. These figures were evaluated by applying...
the algorithm on simulated minimum-bias Au+Au collisions at 10 A GeV. We use the UrQMD program as event generator for such collisions and embed one $J/\psi \rightarrow \mu^+\mu^-$ into each collision. The simulation includes the full detector response of the MUCH and TRD systems. We assume that a complete event (collision) is triggered if the algorithm finds a $J/\psi$ signature as described in the previous section, and all other events are rejected.

We find the trigger efficiency to be 83% (normalized to reconstructable signals, i.e. to those where both daughter muons leave hits in all TRD layers) at a background data suppression rate of 1,700. Both numbers satisfy the CBM requirements. The computation speed 4 $\mu$s per event on an Intel Xeon(R) 3.1 GHz CPU indicates that the algorithm can be applied even for extreme interaction rates on a quite limited number of computing devices.

In summary, we presented a software trigger for the selection of rare $J/\psi$ decays in heavy-ion collisions with the CBM experiment. The procedure uses data from the muon system of CBM only and thus does not require the computationally expensive full reconstruction of tracks in the main tracking system STS. The track reconstruction part of the algorithm is designed such that it can also be used for the detection of low-mass muon pairs, e.g., from the decays of $\rho$ or $\omega$, which are of physics interest as well, but do not require the extreme interaction rates that a charmonium measurement demands.

References
[1] Ablyazimov T et al., arXiv:1607.01487 [nucl-ex]
[2] Friese V (for the CBM collaboration), this volume
[3] Chattopadhyay S et al. (eds), Technical Design Report for the CBM Muon Chambers, GSI-2015-02580, http://repository.gsi.de/record/161297
[4] Hernando J A, The Kalman Filter Technique applied to Track Fitting in GLAST (1998), http://scipp.ucsc.edu/groups/fermi/software/kalman98.pdf