Real-Time Event Reconstruction and Analysis in CBM and STAR Experiments

Ivan Kisel
(for CBM and STAR Collaborations)

1 Goethe-University Frankfurt, Theodor-W.-Adorno-Platz 1, 60323 Frankfurt am Main, Germany
2 FIAS Frankfurt Institute for Advanced Studies, Ruth-Moufang-Str. 1, 60438 Frankfurt am Main, Germany
3 Helmholtz Research Academy Hesse for FAIR, Max-von-Laue-Str. 12, 60438 Frankfurt am Main, Germany
4 GSI Helmholtz Centre for Heavy Ion Research, Planckstr. 1, 64291 Darmstadt, Germany
E-mail: I.Kisel@compeng.uni-frankfurt.de

Abstract. Within the FAIR Phase 0 program, the algorithms of the FLES (First-Level Event Selection) package developed for the CBM experiment (FAIR/GSI, Germany) are adapted for the STAR experiment (BNL, USA). Use of the same algorithms creates a bridge between online and offline, which makes it possible to combine online and offline resources for data processing. In this way, an express data production chain was created on the basis of the STAR HLT farm, that extends the functionality of HLT in real time up to the analysis of physics. It is important, that the express analysis chain does not interfere with the standard analysis chain. A particular advantage of express analysis is that it allows calibration, production and analysis of the data to begin immediately after they are collected. Therefore, the use of express analysis is beneficial for BES II data production and helps to speed up scientific discovery by helping to obtain results within one year after the end of data acquisition. The specific features of express data production are presented and discussed as well as the results of online production and analysis, such as real-time reconstruction of short-lived particles in the BES-II STAR environment.

1. Introduction
Within the framework of the Facility for Antiproton and Ion Research (FAIR) project, a large international centre is being constructed to study the structure and fundamental properties of matter. It will be a new generation accelerator complex that will provide unique opportunities for detailed investigations in the most interesting areas of modern science: nuclear, hadron and particle physics, atomic and anti-matter physics, high density plasma physics, and applications in condensed matter physics, biology and bio-medical sciences [1].

In the Compressed Baryonic Matter (CBM) [2] experiment with heavy ions, the highest baryon densities will be created, and the properties of super-dense nuclear matter will be investigated in various extreme states that are similar to, for example, the conditions of matter in the center of neutron stars, where matter is at the final stage of evolution before transition to the black hole. The CBM experiment will thus complement the experimental heavy-ion program...
at the LHC accelerator complex at CERN where the properties of the hot matter similar to that after the Big Bang are investigated.

The scientific program of the CBM experiment includes:

- explore properties of super-dense nuclear matter;
- search for in-medium modifications of hadrons;
- search for the transition from dense hadronic matter to quark-gluon matter;
- search for the critical endpoint in the phase diagram of strongly interacting matter;
- investigate the structure of neutron stars and the dynamics of core-collapse supernovae.

The experiment will measure rare and penetrating probes such as dilepton pairs from light vector mesons and charmonium, open charm, multistrange hyperons, together with collective hadron flow and fluctuations in heavy-ion collisions at rates up to $10^7$ collisions per second.

![Image](image_url)

**Figure 1.** A simulated central Au-Au collision at 25 AGeV energy with about 1000 charged particles in the CBM experiment.

CBM is characterized by high collision rates, large amount of produced particles, non-homogeneous magnetic fields and a very complex detector system. Event reconstruction is the most complicated and time consuming task of the data analysis in modern high-energy physics experiments. It is a key part of success in the CBM experiment with up to thousand particles per central collision (Fig. 1). An additional complication in CBM is its continuous data stream represented in form of time slices. This makes the reconstruction of such 4-dimensional data with time stamps and the search for interesting physics extremely difficult. All of the above mentioned makes necessary to develop fast and efficient algorithms for data analysis and to optimize them for running on a modern high-performance computer cluster [3].

2. First Level Event Selection

The First Level Event Selection (FLES) package [4, 5] of the CBM experiment is intended to reconstruct online the full event topology including tracks of charged particles and short-lived particles. The FLES package consists of several modules: Cellular Automaton (CA) track finder, kalman Filter (KF) track fitter, KF Particle Finder and physics selection. In addition, a quality check module is implemented, that allows to monitor and control the reconstruction process at all stages. The FLES package is platform and operating system independent. The package is portable to different many-core CPU architectures, vectorized using SIMD (Single Instruction, Multiple Data) instructions and parallelized between CPU cores. All algorithms are optimized with respect to the memory usage and the speed.
2.1. Cellular Automaton (CA) track finder

The 4-dimensional (4D, space and time) Cellular Automaton (CA) track finder [5, 6] takes as input hit measurements from the tracking detector in the form of a time-slice, which includes time and spacial measurements. The track finding procedure starts with combining the hits into triplets, combination of three hits on adjacent stations. The triplet structure was chosen, since it allows to estimate the momentum of a particle, which could produce it. The triplets with two common hits are combined into track candidates. The track candidates should survive a dedicated selection based on the track length and calculated $\chi^2$-value to be accepted to the reconstructed tracks.

Input time information is used in the algorithm to the same extent and in similar manner as it is done with the spacial coordinates. The same logic is used while constructing triplets: the hits in the triplet should belong to the same particle, therefore they should correlate not only in space, but also in time. The resulting track reconstruction efficiencies for the cases of event-by-event analysis (so-called 3D analysis) as well as for the 4D case (with included time measurement, as well as 3-dimensional spacial information) while reconstructing time-slices are similar thus there is no efficiency degradation in the much more complicated case of time slices. The same is valid for the speed of the 4D CA tracks finder with respect to the 3D case.

2.2. Kalman Filter (KF) track fit

High precision of the parameters of particle trajectories (tracks) and their covariance matrices is a prerequisite for finding rare signal events among hundreds of thousands of background events. Such high precision is usually obtained by using the estimation algorithms based on the Kalman filter (KF) method. High speed of the reconstruction algorithms on modern many-core computer architectures can be accomplished by: optimizing with respect to the computer memory, in particular declaring all variables in single precision, vectorizing in order to use the SIMD instruction set and parallelizing between cores within a compute node.

2.3. 4D event builder

![Diagram](image)

**Figure 2.** Reconstructed tracks in time slices clearly represent groups, which correspond to the original events with 85% of single events, no splitted events, and final event building is done at the vertexing stage using TOF information.

After all tracks are found and their parameters are reconstructed, the tracks are grouped into events. This is done by clustering tracks based on their time parameters in the area of the target. The left distribution of Fig. 2 shows hits within a time slice for $10^7$ interaction rate. One can see that the traditional grouping of hits into events at this stage is impossible. The track distribution at the middle shown against the same hits displays grouping of tracks belonging to the same event. The right distribution with different hits shows different clusters of tracks, close in time in the target area. One can see that already at this stage it is possible to build events with efficiency more than 85%. The task of event building is finalized at the stage of
searching the primary vertex, where it is possible to additionally use the proximity of tracks in space, as well as more accurate time measurements of the TOF detector.

2.4. KF Particle Finder — a package for reconstruction of short-lived particles

Today the most interesting physics is hidden in the properties of short-lived particles, which are not registered, but can be reconstructed only from their decay products. A fast and efficient KF Particle Finder package [4, 7], based on the Kalman filter (hence KF) method, for reconstruction and selection of short-lived particles is developed to solve this task. A search of more than 100 decay channels has been currently implemented (Fig. 3).

In the package all registered particle trajectories are divided into groups of secondary and primary tracks for further processing. Primary tracks are those, which are produced directly in the collision point. Tracks from decays of resonances (strange, multi-strange and charmed resonances, light vector mesons, charmonium) are also considered as primaries, since they are produced directly at the point of the primary collision. Secondary tracks are produced by the short-lived particles, which decay not in the point of the primary collision and can be clearly separated. These particles include strange particles ($K^0_S$, $K^0_L$) and multi-strange hyperons ($\Xi$ and $\Omega$) and charmed particles ($D_0$, $D_s^\pm$, $D_s^0$ and $\Lambda_c$). After that tracks are combined according to the block diagram in Fig. 3. The package estimates the particle parameters, such as decay point, momentum, energy, mass, decay length and lifetime, together with their errors. The package has a rich functionality, including particle transport, calculation of a distance to a point or another particle, calculation of a deviation from a point or another particle, constraints on mass, decay length and production point. All particles produced in the collision are reconstructed at once, that makes the algorithm local with respect to the data and therefore extremely fast.

In addition, simultaneous reconstruction in the KF Particle Finder of different decay channels of the same particle, including also decays with a neutral particle in the final state, makes it

![Figure 3. Block diagram of the KF Particle Finder package. The particle parameters, such as decay point, momentum, energy, mass, decay length and lifetime, together with their errors are estimated using the Kalman filter method.](image-url)
possible to calculate the efficiency of the reconstruction of rare particles and reliably estimate their systematic errors.

Figure 4. The FLES package provides clean probes of various stages of the collision: results of the search for short-lived particles are shown for 5M central AuAu UrQMD events at 10 AGeV with realistic PID.

The use of the Kalman filter at all stages of particle reconstruction allows in many cases to get rid almost completely of the combinatorial background and to obtain clean sets of particles, which can serve as probes of various stages of the collision (Fig. 4).

2.5. Deep learning for quark-gluon plasma detection
In addition to the macroscopic inverse approach [8] we investigate the microscopic inverse approach by using artificial neural networks to classify processes in heavy ion collisions. We have created two types of neural networks: fully connected (FC) and deep convolutional (CNN) neural networks. These networks were then used to identify quark-gluon plasma simulated within the Parton-Hadron-String Dynamics (PHSD) microscopic off-shell transport approach for central Au+Au collision at a fixed energy.

For FC networks we use a 64-neuron fully-connected hidden layer with batch normalization, Leaky Rectified Linear Unit (LReLU) activation and dropout. The number of neurons is chosen empirically and is fixed to allow comparison of FC neural networks with one, two and three layers. Batch normalisation and dropout are used to reduce overfitting and therefore improve overall performance. LReLU is used as it performs similarly to the most commonly used Rectified Linear Unit (ReLU) activation function but avoids dead neuron issues.

The CNN consists of two three-dimensional convolutional layers, each followed by a max pooling layer, and two sequential fully-connected layers.
The Fig. 5 shows that the accuracy on the validation set rapidly increases for all four network architectures up to the fifth epoch, when the rise slows down and the curves level off. At the same time, the precision on the training set continues to go up until it reaches 100%, which suggests that overfitting occurs after the fifth epoch. Nevertheless, the fully-connected networks reach 80% precision while the convolutional neural network attains the best performance of more than 90% accuracy.

3. Express reconstruction and analysis in STAR

The STAR (Solenoidal Tracker At RHIC) experiment [9] at the RHIC (Relativistic Heavy Ion Collider) facility of the Brookhaven National Laboratory (BNL, USA) is designed to study nuclear matter under extreme conditions of relativistic heavy ion collisions, including hadron production and search for signs of quark-gluon plasma formation and its properties.

Very important for RHIC is the possibility to collide ions, covering the range of baryon chemical potential ($\mu_B$) from 20 to 420 MeV, which corresponds to a wide range of energies (so called Beam Energy Scan, BES). The BES results once again confirmed evidence of QGP discovery in the upper RHIC energy $\sqrt{s_{NN}} = 200$ GeV. The results of the search for the critical point and the first-order phase boundary have narrowed the region of interest to collision energies below $\sqrt{s_{NN}} = 20$ GeV. Thus, the phase II of the Beam Energy Scan (BES-II), scheduled for 2019-2020 [10], covers the energy range $\sqrt{s_{NN}} < 20$ GeV. Current lattice QCD calculations suggest that key features of the phase diagram like the critical point and the first-order phase transition lie within $\mu_B$ of the BES Phase-II program.

There are three main improvements to the STAR detector for the BES-II phase: inner TPC (iTPC), end-cap Time-Of-Flight (eTOF), and event plane detector (EPD). These new detector systems cover pseudo-rapidity in the range of $1 < |\eta| < 5.4$. While iTPC and EPD cover both sides of STAR, eTOF is only installed on one side. The upgrade of iTPC and eTOF allows better identification and acceptance of particles. The EPD provides event plane reconstruction and centrality detection. All three detector systems improve significantly the statistical accuracy of STAR measurements in the low energy range close to the CBM energy range.

The FLES package of the CBM experiment is ported to the STAR experiment. Since 2013 (online) and 2016 (offline) the CA track finder is the standard STAR track finder for data production. Use of CA provides 25% more $D_0$ and 20% more $W$. The KF particle finder provides a factor 2 more signal particles than the standard approach in STAR. The package is
used in HLT and its integration into the official STAR repository for use in the standard physics analysis is currently in progress.

Figure 6. The HLT express and the standard data production and analysis workflows.

The use of the CA track finder and the KF particle finder in online extends significantly the functionality of HLT (Fig. 6). The standard calibration, production and analysis remain unchanged. HLT starts the calibration procedure as soon as data become available. The express chain makes possible physics analysis of the data as soon as the calibration is reasonable. It unifies approaches in extended (x)HLT and online (o)RCF to speed up the express workflow, and combines high competence of xHLT and oRCF experts involved in online operation. In addition, it provides physics working groups with instant and uncomplicated access to the data, like picoDST etc.

With the express calibration and alignment one can reconstruct hyperons with high significance and low level of background, as it is shown in Fig. 7. Hyperons are clearly seen at all BES-II energies: 3, 3.2, 3.9, 7.7, 9.1, 14.5, 19.6, 27 GeV. In addition, high significance allows extraction of spectra.

4. Conclusion
The CBM experiment with $10^7$ input rate will require the full event reconstruction and physics analysis of the experimental data online. As the same HPC farm will be used for offline and online processing of experimental data, the main reconstruction and analysis algorithms will work both offline and online. Errors and insufficient accuracy in online data processing, physics analysis or selection of interesting collisions by the reconstruction algorithms will lead to complete loss of all experimental data, since only the incorrectly selected data will be stored in this case. Therefore only immediate comparison of the results of online analysis with the predictions of theoretical models using ANNs can guarantee the proper operation of the whole experiment. It has been demonstrated, that the core algorithms of the FLES package, the Cellular Automaton for searching for particle trajectories (100 $\mu$s/core/track) and the Kalman Filter to estimate their parameters (0.5 $\mu$s/core/track), have a very high level of intrinsic parallelism for their fast
and efficient implementation on many-core CPU/GPU architectures. The KF particle finder package with more than 150 decay channels implemented (100 µs/core/decay) is a common platform for offline physics analysis and for real-time express analysis at $10^7$ interaction rate in CBM.

Adaptation of the FLES algorithms within the FAIR Phase-0 program to the STAR experiment with its excellent detector performance, high quality experimental data and a well established reconstruction chain is the first and successful step in preparing the FLES algorithms for reconstruction and analysis of CBM real data at Day-1. Use of the CA track finder and KF particle finder developed in the CBM experiment can be beneficial for other experiments as the experimental heavy ion physics becomes more and more challenging.

References

[1] FAIR — Facility for Antiproton and Ion Research. Green Paper. The Modularized Start Version. GSI. October 2009.
[2] CBM Collaboration, Compressed Baryonic Matter Experiment, Tech. Stat. Rep., GSI, Darmstadt, 2005; 2006 update.
[3] I. Kisel, Event reconstruction in the CBM experiment, Nucl. Instr. and Meth. A566 (2006) 85-88.
[4] I. Kisel, I. Kulakov and M. Zyzak, Standalone first level event selection package for the CBM experiment, IEEE Trans. Nucl. Sci. 60 (5) (2013) 3703-3708.
[5] V. Akishina and I. Kisel, Online 4-dimensional reconstruction of time-slices in the CBM experiment, IEEE Trans. Nucl. Sci. 62 (6) (2015) 3172-3176.
[6] V. Akishina, “4D Event Reconstruction in the CBM Experiment”, Dissertation thesis, Goethe university, Frankfurt am Main (2017).
[7] M. Zyzak, “Online Selection of Short-Lived Particles on Many-Core Computer Architectures in the CBM Experiment at FAIR”, Dissertation thesis, Goethe university, Frankfurt am Main (2016).
[8] I. Kisel, Event Topology Reconstruction in the CBM Experiment, J. Phys. Conf. Ser. 1070 (2018), 97.
[9] K.H. Ackermann et al. (STAR Collaboration), “STAR detector overview”, Nucl. Instr. Meth. A499 (2003) 624.
[10] STAR Collaboration, “Studying the Phase Diagram of QCD Matter at RHIC”, 01 June 2014. https://drupal.star.bnl.gov/STAR/files/BES_WPHI_ver6.9_Cover.pdf

Figure 7. Online search for hyperons on 200M AuAu events at 14.5 GeV (2019 BES-II express production).