APPROVED

JINR DIRECTOR

_______________/_________________

"_____" _________________ 2023 г.

PROJECT PROPOSAL FORM

Opening of a research project within the Topical plan of JINR

1. General information on the research project of the theme

1.1 Theme code

05-6-1119-2014/...

1.2 Project

1.3 Laboratory

Meshcheryakov Laboratory of Information Technologies

1.4 Scientific field

Networking, computing, computational physics

1.5 Title of the project/LRIP subproject

Mathematical methods, algorithms and software for modeling physical processes and experimental facilities, processing and analyzing experimental data

1.6 Project leader

S.V. Shmatov

1.7 Project deputy leaders

A.S. Ayriyan, N.N. Voytishin

2 Scientific case and project organization

2.1 Annotation

The project is aimed at organizing and providing computational support for the physics research programs implemented with the participation of JINR, the development of mathematical methods and software for modeling physical processes and experimental facilities, processing and analyzing experimental data in the field of elementary particle physics, nuclear physics, neutrino physics, condensed matter, radiobiology, etc. The particular attention will be paid to the creation of systems for the distributed processing and analysis of experimental data as well as information and computing platforms to support research at JINR and other world centers.

The main areas of work are mathematical and computational physics to support the JINR large research infrastructure projects, and first of all the experiments at the NICA accelerator complex and the Baikal-GVD neutrino telescope. Further cooperation will also be continued with the experiments at the largest world accelerator centers (CERN, BNL, etc.), experiments in the field of neutrino physics and astrophysics, radiobiological research programs. The possibility of using the developed methods and algorithms within other fundamental science and applied projects is being considered.

2.2 Scientific case (aim, relevance and scientific novelty, methods and approaches, techniques, expected results, risks)

In accordance with the Seven-Year Plan for the Development of JINR for 2024−2030 and the JINR Long-Term Development Strategy up to 2030 and beyond, the main priorities of JINR research in the "Networking, Computing, Computational Physics" direction include the provision of mathematical, algorithmic and software support for experimental and theoretical studies conducted at JINR, and, in particular:

- development of algorithms based on recurrent and convolutional neural networks for the tasks of machine and deep learning and Big Data analytics, designed primarily for solving various tasks in particle physics experiments, including the NICA megaproject and neutrino experiments;
- creation of modern research tools for international collaborations (NICA, the JINR neutrino program, LHC experiments);
- elaboration of scalable algorithms and software for processing multi-parameter, multidimensional, hierarchical data sets of exabyte magnitude;
- development of information and computing systems for experimental data analysis and processing in the field of radiobiology.

These priorities determine the main goals and objectives of the presented project.

Goals and objectives

The main goal of the project is to do research and development works in the field of mathematical and computational physics for experiments in elementary particle physics, physics of the atomic nucleus, condensed matter physics, radiobiology and applied research within the JINR Topical Plan. The project provides for the creation and development of mathematical methods and proper software, research tools for international collaborations (NICA, the JINR neutrino program, LHC experiments, etc) and the implementation of these methods and algorithms for applied research. The main areas of work are:

- development and application of modern methods and algorithms for modeling physical processes and experimental facilities;
- development and application of modern methods and algorithms for physical object reconstruction, event visualization and data analysis;
- development of a program, together with the staff of JINR and other Russian and world centers, for experiments aimed to study the fundamental properties of matter, their preparation and implementation;
- creation and development of information and computing systems for experimental data processing and analysis, participation in the organization and processing of experimental data and analysis of physics information, including Open Data of world experiments, using GRID technologies and computer centers of distributed computing, in particular, the JINR MICC, including the “Govorun” supercomputer;
- creation of modern research tools for international collaborations (NICA, the JINR neutrino program, LHC experiments), including software for the control of data quality, characteristics and calibration of experimental facilities within the JINR Topical Plan;
- creation and development of information and computing systems for radiation research in life sciences;
- application of the developed methods and algorithms for applied purposes.

The main strategy is to use common solutions and methods for different experiments to develop software for physics simulation, data processing and analysis, as well as computing systems. All developments are aiming at realizing the actual research tasks of the JINR experimental programs, analytical or numerical calculations will be carried out within the framework of physical scenarios that can be proved experimentally.

Project Structure

The project includes the following directions of the development and implementation of methods for simulation, experimental data processing and analysis:

- Simulation of physical processes and experimental facilities:
  - analytical and numerical calculations of physical processes, software optimization, including tuning and adaptation of physics event generators;
  - mass modeling and creation of event databases;
  - participation in the creation of computer models of experimental facilities and simulation of elementary particles passing through them based on GEANT4 (and others) and fast simulation of the response of the detectors.
- Reconstruction of physical objects and analysis of experimental data:
  - development of algorithms, including those based on recurrent and convolutional neural networks for machine and deep learning tasks, and creation of corresponding software for the reconstruction of physical objects (tracks, particles, clusters, etc.) and physical processes;
  - development of methods and algorithms for data analysis, including statistical analysis;
  - adaptation of existing software for specific experiments, reconstruction and analysis of experimental data;
  - analysis of Open Data of experiments, in particular, experiments at the LHC;
  - conducting a global analysis of data from various experiments (in particular, a combined analysis of data from accelerator and astrophysical experiments in search for candidates for the role of dark matter).
- Support and development of the software environment for experiments
  - optimization of the data structure, development of methods for collecting, processing and storing data for the experiments at NICA;
  - support and development of databases;
  - creation of software for event visualization (simulation and experimental data).

I. Simulation of physical processes and experimental facilities

Analytical computing, the modeling of physical processes using Monte Carlo methods and numerical estimates of cross sections, of the number of expected events, the kinematic distributions of observables made on their basis are an integral part of the development and implementation of any research program. At present, for research in the field of physics of elementary particles and the atomic nucleus, a large number of software packages that enable predictive calculations based on modern theoretical concepts (models) have been created. Such software systems, which are often called event generators, are sufficiently versatile and can be used in various experimental programs; therefore, the tasks of creating and developing event generators are common to many experiments. Moreover, for experiments performed under the same kinematic conditions, it is possible to use a unified database of the Standard Model (SM) and within various scenarios beyond it, for example, for the kinematic conditions of the NICA and CERN accelerator complexes.

Another important direction of work of the project is the modeling of physics facilities: the creation of their geo-model, the simulation of the passage of elementary particles through them, i.e., obtaining the response of detector systems and its digitization. As a rule, such modeling is carried out using the Geant4 package (although other approaches can be implemented for specific experimental conditions[1]). In this direction, the development of the physics models of the package, on the basis of which the interaction of elementary particles with the detector substance is simulated, as well as the creation of spatial models of specific detectors and experimental facilities (for example, SPD at NICA and CMS at the LHC) and the implementation of the entire chain of complete event modeling from the primary act of interaction of primary particles to the digital signal of the data acquisition system, is of particular value.

[1] For example, Baikal-GVD

Development of the FTF (Fritiof) and QGSM (Quark-Gluon-String-Model) hadron models of the Geant 4 package. Nucleus-to-nucleus collisions at high energies produce a large number of stable and unstable particles, as well as light nuclei and anti-nuclei. Their registration and measurement of their kinematic characteristics are expected at the existing RHIC (USA) and LHC (Geneva, Switzerland) experimental facilities, and are planned at the NICA accelerator complex (JINR, Russia). The registration and measurement are and will be performed using various electronic devices containing various materials. The accounting for the interactions of produced particles with materials is implemented using various computer software systems, namely, packages. One of these packages is the Geant4 package, which was created at CERN (Geneva, Switzerland) within the relevant international collaboration. It models quite well the passage of stable particles through various media. However, experimental studies require taking into account the interactions of unstable, long-lived particles such as charmed hadrons, hyperons, anti-nuclei and anti-hyper-nuclei. To solve this problem, it is necessary to synthesize various theoretical approaches and computational methods.

The goal to implement the formation of the above unstable particles and the simulation of their strong interactions with nuclei in the FTF (Fritiof) and QGSM (Quark-Gluon-String-Model) hadronic models of the Geant4 package is set within the project.

The formation of light nuclei and light anti-hypernuclei has been observed in the experiments at the RHIC and the LHC, in particular, by the ALICE collaboration (CERN). Experimental collaborations intend to continue these studies. However, it is necessary to solve a methodological issue, i.e., to take into account the strong interactions and rescattering of these particles in the materials of the detectors.

It is expected that the multiple production of charmed particles will take place at the accelerator complexes that are being designed, namely, the FCC (Future Circular Collider) and the ILC (International Linear Collider). In this case, it will be needed to take into account the absorption of charmed particles by the detector materials.

The novelty of the research is the transfer and adaptation of methods and approaches used for the simulation of interactions of stable particles and nuclei with nuclei, the simulation of collisions of charmed particles with nuclei, hyper-nuclei, anti-nuclei and anti-hyper-nuclei.

The main method for calculating the cross sections of the interactions of particles and nuclei with nuclei is the Glauber approximation. It involves setting the cross sections of elementary collisions and the structure of the target nuclei. This approach was tested in application to antiproton-nucleus cross sections in the work[2]. As applied to the problem under consideration, it is necessary to specify the cross sections of the interaction of charmed particles, hyperons, and anti-hyperons with nucleons. As a first approximation, the approximation of the additive quark model (AQM) will be considered.

To specify the structure of light hypernuclei, it is proposed to use the results obtained within the chiral effective field theory (χEFT). The Geant4 Barashenkov-Glauber-Gribov parameterization was chosen as the basis for calculating nuclear cross sections. The modeling of inelastic intranuclear interactions will be performed in the same way as it is done in the FTF and QGSM models of the Geant4 package with the necessary changes (see work[3]). The birth of charmed particles in elementary interactions was carried out in Geant4 in 2019 with our participation (Geant4 release 10.7).

Development of the DCM-QGSM-SMM heavy ion collision generator. Nuclear physics modeling is one of the main tools at the stages of planning, creation and analysis of the obtained data in modern experiments on the study of physical processes occurring in heavy ion collisions. It is planned to continue the development of the DCM-QGSM-SMM heavy ion collision generator to include new processes, such as environmental effects, for a more consistent description of experimental data. The scientific novelty

[2] Antinucleus-nucleus cross sections implemented in Geant4. V. Uzhinsky, J. Apostolakis, A. Galoyan et al., Phys. Lett. B 705 (2011) 235

[3] Recent Developments in Geant4, J. Allison, … A. Galoyan, … V. Uzhinsky, ... et al., Nucl. Instrum. Meth. A 835 (2016) 186)

lies in the main goal of the NICA project, for the implementation of which modeling is used: the search for a new state of matter in heavy ion collisions in the energy range achievable at the collider.

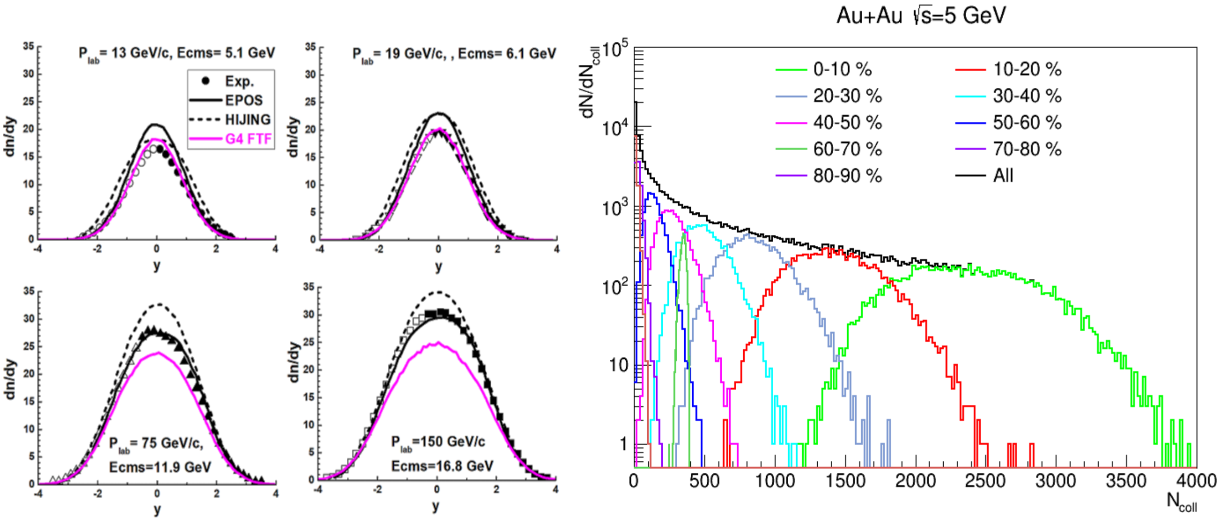

Fig. 1. Comparison of the NA61/SHINE π−meson rapidity distributions for experimental and simulated data (Geant4) (left). Development of the DCM-QGSM-SMM Monte Carlo generator for the NICA experiments (right).

Analytical and numerical methods for calculating neutron-proton systems under strong compression at the NICA SPD facility. There are theoretical indications that various phase transitions can occur in highly compressed nuclear matter at moderate temperatures. Traditional nuclear physics, as well as astrophysical limitations, has not yet provided direct evidence for the existence of such transitions. It is expected that future experiments at the NICA facility will at least partially solve this problem. However, due to the large theoretical uncertainty in this area of research, the interpretation of the experimental data that will be obtained at the NICA MPD facility may be highly ambiguous. Therefore, the experimental study of various transformations in low-nucleon systems is of particular interest. The purpose of this work is to obtain the above data at the NICA SPD facility in experiments on recording the well-defined characteristics of inelastic d+d interactions. The theoretical prediction of Matveev and Sorba about the presence of an impurity of 6-q systems in the deuteron by measuring the momenta of secondary protons in the reaction d + d → d + p + n is also supposed to be experimentally verified. The tasks set, if they can be solved, will make it possible to obtain fundamentally new information about the properties of nuclear matter. In particular, dibaryons that have not been reliably detected so far can be registered. Within the project, analytical and numerical methods for calculating the kinematics of excitation reactions and the subsequent decay of light dibaryons in the reactions d+d→d+d*→d+n+p, as well as in the reactions d+d→d+n+p of the direct knockout of protons, taking into account intranuclear motion and binding energy, are developed. In addition, it is planned to simulate these processes using the GEANT SPD program, which enables to assess the achievability of the accuracy required to solve the tasks.

Simulation of the processes of production of particles-candidates for the role of dark matter and processes with the violation of the lepton number. The Large Hadron Collider (LHC), designed as a machine of discovery, provides a unique opportunity to search for signals of new physics. The priority tasks of the JINR groups in experiments at the LHC include the search for candidates for the dark matter (DM) particles, verification of predictions of low-energy gravity scenarios on a scale of a few TeV, extended Higgs and gauge models, etc.

The DM candidates are searched for in the processes of inclusive production of pairs of opposite-signed leptons and in the channels with pairs of fermions (leptons or heavy quarks), neutral gauge bosons, Higgs bosons, etc. associated with a large amount of the missing transverse energy (MET). The results of such searches can be interpreted in the context of various DM scenarios – simplified DM models (with one Dirac DM particle and one mediator), extended Higgs models of the 2HDM+a/s type, etc. In general, events associated with MET allow testing a wide range of DM models.

One of the brightest signals of physics beyond the Standard Model (SM) are the processes that violate the lepton number (Lepton Flavor Violation, LFV). For example, studies of pairs of leptons with different flavor states (eμ, eτ, μτ) make it possible to test various physics scenarios beyond the SM, i.e., models with an extended Higgs and gauge sector allowing LFV, supersymmetric theories with R-parity violation, scenarios of multidimensional low-energy gravity with quantum black holes, etc[4]. A signature with a pair of leptons of the same flavor and the same charge allows one to test scenarios of the extended Higgs sector, which gives non-doublet Higgs representations (triplet Higgs representations and doubly charged Higgs states in composite Higgs boson models, “Little Higgs”, etc.), and some options for simplified descriptions of the interactions of dark matter (DM) with ordinary SM matter[5].

As part of the physics research program for the CMS experiment at the LHC and in cooperation with the VBLHEP and BLTP team, it is planned to perform a number of experimental analyses to search for such states. To do this, it is necessary to calculate the cross sections of their production, the corresponding simulation and estimate the expected number of events in a wide range of model parameters that were not covered by previous searches (see, for example, Fig. 2). It will also enable, in case no significant deviation is observed with respect to the SM background expectations, to obtain new unique limitations on the parameters of physics scenarios beyond the SM[5]. For modeling, a whole set of event generators will be used, namely, Pythia8, QBH, MadGraph5_aMC@NLO with the integrated FeynRules library of matrix elements, etc., The tests of the theoretical scenarios are also possible with open data from collaborations at the LHC[6].

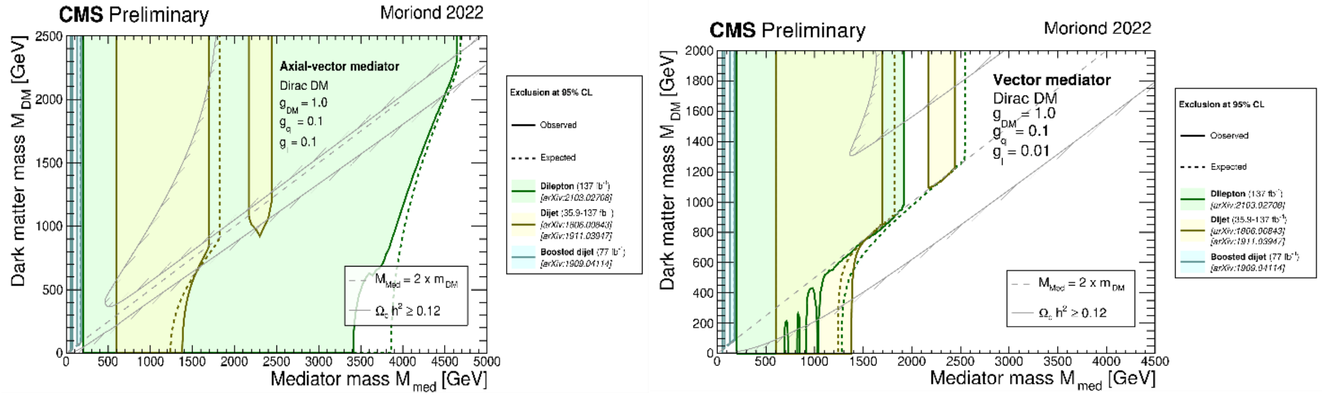

Fig. 2. Limits (95% C.L.) on the masses of particles-candidates for the role of dark matter particles m_{DM} and the particle-carrier of interaction with the dark sector M_{med}. Data from the combined channel for the production of a pair of jets and a pair of leptons in events recorded in the CMS experiment at the LHC. The shaded area corresponds to the closed mass values for the vector (on the right) and pseudovector (on the left) transporters.

II. Reconstruction of physical objects and analysis of experimental data

This section provides for the development of algorithms, including those based on recurrent and convolutional neural networks for machine and deep learning tasks, designed primarily for processing and analyzing data from particle physics experiments for the NICA megaproject, neutrino experiments, and CERN experimental programs.

Mathematical methods and software for processing and analyzing data from the BM@N experiment. The BM@N (Baryonic Matter at Nuclotron) experiment is the first operating experiment in the NICA project. The purpose of the experiment is to study the interactions of heavy ion beams (up to gold nuclei) with a fixed target in the energy range $\sqrt{s_{NN}}$ = 2.3−3.5 GeV. To achieve the maximum

[4] M.V. Savina, S.V. Shmatov, "In Search for New Physics", in An Essay on Modern Particle Physics. Eds. V.A. Matveev and I.A. Golutvin – Dubna: JINR, 2020 - p. 133-215, ISBN 978-5-9530-0506-7.

[5] M. Savina, "DM interpretations of heavy resonances and BSM-Higgs searches in ATLAS and CMS", The Eighth Annual Conference on Large Hadron Collider Physics-LHCP2020, 25-30 May, 2020, PoS(LHCP2020) 176.

[6] https://opendata.cern.ch/

accuracy of experimental results, the efficient and accurate reconstruction of particle tracks and the additional software alignment of the spatial arrangement of the detectors that make up the BM@N setup are required. To implement the research program of the BM@N experiment, it is planned to perform the following:

- development and application of methods and algorithms with their subsequent implementation in the form of a set of problem-oriented programs for modeling and reconstructing physical events, as well as processing and analyzing experimental data for the coordinate detectors (Forward Silicon, GEM, small CSC, DCH, large CSC, Silicon Beam Tracker, Silicon Profilometer) of the BM@N experiment tracking system;
- alignment and calibration of individual detectors of the BM@N facility using the software solution being developed for the geometric alignment of STS (silicon chambers) and GEM (gas electron multipliers) track detectors; the integration of the obtained solution into the main BmnRoot software environment is expected. The alignment involves finding correction parameters for the STS and GEM track detectors of the BM@N experiment and is based on the simultaneous determination of track parameters and geometric corrections using the technique proposed by W. Blobel to reduce the dimension of the main matrix of the system of equations obtained by minimizing the functional. The scientific novelty of the research is primarily defined by the uniqueness of the detector system of the experiment itself for the correction parameters, considering the geometry of the detectors and the set of experimental data with and without considering the magnetic field.

Development of algorithms for the reconstruction and identification of particles in the MPD and SPD experiments based on machine learning methods. The main goal of this area of research is to broaden the scope of application of machine learning methods in the tasks of the experiments at the NICA accelerator complex, namely, the development of an algorithm for solving the task of particle reconstruction and identification in the MPD and SPD experiments. Today, machine learning methods are becoming increasingly popular for solving a wide range of tasks in high-energy physics, including the task of charged particle identification. This is due to the fact that machine learning methods make it possible to obtain more effective identification in those areas in which traditional methods do not provide sufficient identification accuracy. In addition, the model is trained once for each data set, depending on the geometry of the physics experiment, which can increase the rate of charged particle identification. The reconstruction of trajectories and the identification of charged particles are an important step in processing data from any high-energy physics experiment in general and in the experiments at the NICA collider. Within the project, it is planned

- to apply machine learning methods for the particle identification task in MPD based on the gradient boosting of decision trees (Fig. 3, left), the ensemble machine learning method, the Bayesian approach, in particular, the Tree of Parzen Estimators Bayesian optimization method for finding optimal hyperparameters of machine learning models, classical approaches of charged particle identification implemented in the software of the experiment, methods for assessing the quality of the machine learning model, as well as the method of comparative analysis, methods for assessing the importance of the characteristics of charged particles, the SHapley Additive explanation method for calculating the Shapley vector (Fig. 4, right) for the purpose of the further interpretation of results obtained through machine learning methods;
- within the SPD experiment, to develop algorithms for recognizing the trajectories of charged particles in the detector system of the experiment using deep neural networks. The authors of the project previously proposed two new approaches to track recognition in strip and pixel detectors. The first approach was implemented by the authors in the TrackNetv3 software package; it relies on the use of a recurrent neural network (RNN) (Fig. 4, left), which allows one to combine track extrapolation with testing the hypothesis that the set of points belongs to the true track and is compatible with a smooth curve, i.e., essentially reproduces the idea of the Kalman filter with the difference that the physical parameters describing the track are approximated by the neural network using synaptic weights determined during its training. The second approach, implemented by the authors in the RDGraphNet package, uses a graph network (Fig. 5, right) and makes it

possible to implement a global search for tracks in an event, which is especially attractive when analyzing events with a large multiplicity. These approaches have already been successfully used for track recognition in the BM@N experiment at JINR and in the BESIII experiment at IHEP CAS in China. During the project, these approaches will be adapted to search for and reconstruct elementary particle tracks in SPD data from the silicon vertex detector and the main straw tracker. The main difficulty is the adaptation of neural networks to restore tracks in drift detectors, which requires the solution of “left-right” ambiguity. The Ariadne software package will be used to prototype neural networks and study the quality of their work. Algorithms based on convolutional networks will be developed to search for clusters in the SPD electromagnetic calorimeter and quickly reconstruct π^0. To identify muons in the muon system, it is planned to either use convolutional neural networks or apply a simpler gradient boosting algorithm on trees. Creation of software for autonomous (offline) data processing with the efficient use of multi-core computing architectures.

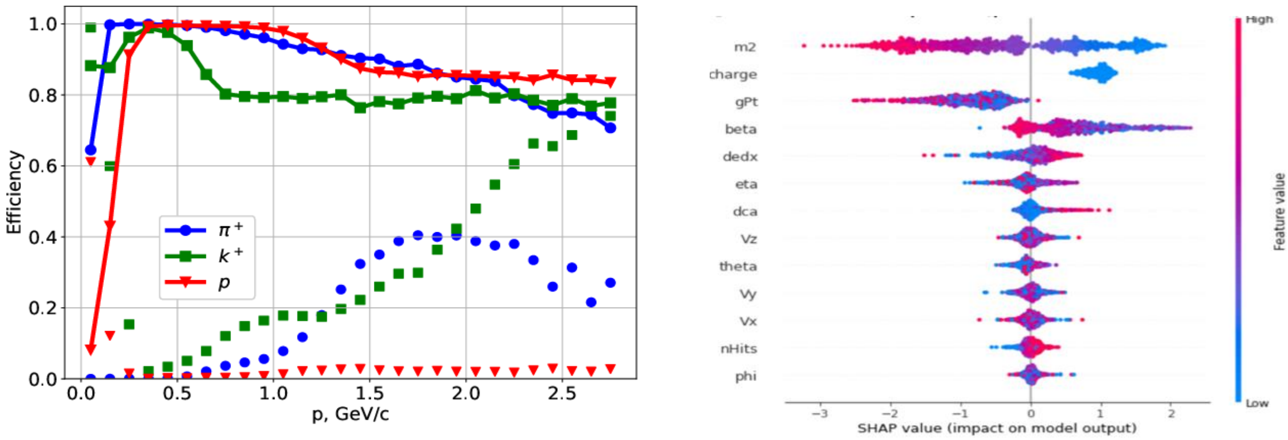

Fig. 3. Efficiency and contamination (error of the second kind) of the identification of positively charged particles in the MPD model data, obtained by the method of decision trees with gradient boosting (left), and the Shapley vector of the influence of track characteristics on the detection of pions with respect to kaons (right).

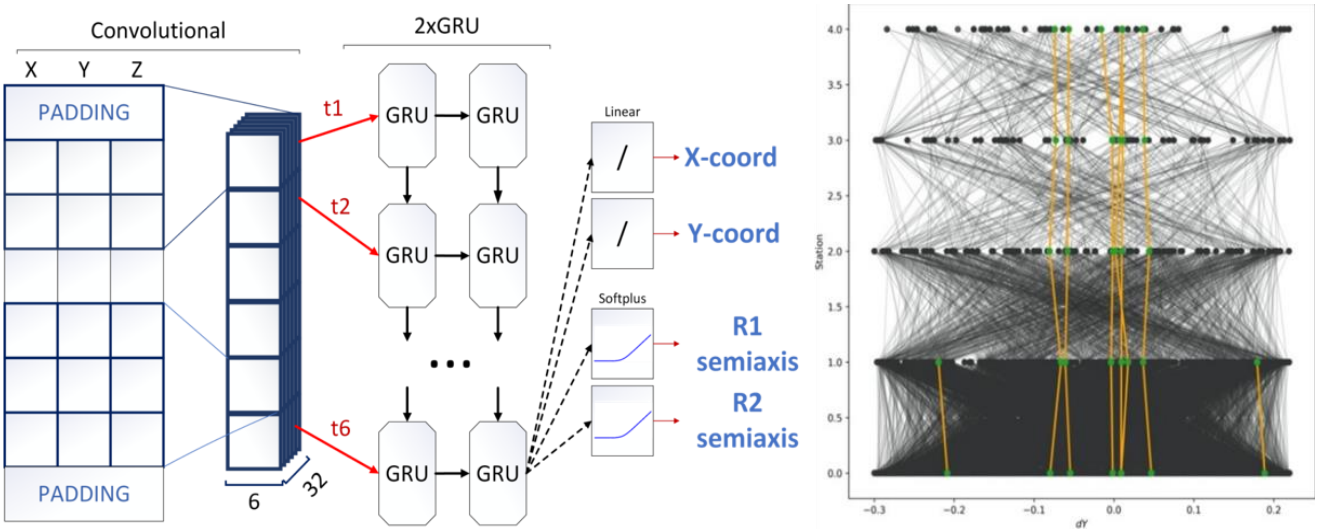

Fig. 4. Schematic representation of the model based on recurrent neural networks for the tracking task in BM@N (left) and graphical representation of the event in BM@N, where the black nodes and edges correspond to fakes, and the green nodes and yellow edges correspond to correctly found tracks (right).

Mathematical methods and software for muon reconstruction and the estimation of operation parameters of CMS detectors. A gradual increase in the luminosity and energy of the LHC beam directly affects the operation of the detectors and the quality of the reconstruction of the trajectories of detected particles. In this regard, there is a need for fast and accurate tracking algorithms in CMS detectors, as well as the development, testing and implementation in the experimental facility of new detectors that can work effectively in such complex data acquisition conditions. Software development for the CMS experiment will address two main directions:

- reconstruction of the cosmic muon trajectory in the setup for testing active elements of the High Granularity Calorimeter (HGCal), as well as evaluation of the efficiency of HGCal modules;
- usage of discrete wavelet analysis to recognize the coordinates of close-flying particles from overlapping signals in the Cathode Strip Chambers (CSC). Evaluation of the operation parameters of CSC detectors and of the rate of background particles for different types of experimental data.

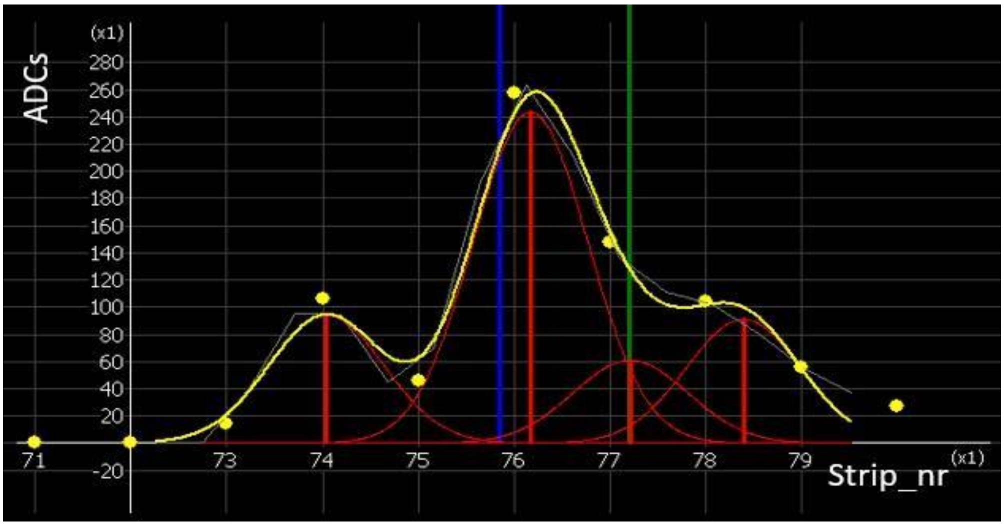

Fig. 5. Recognition of overlapping signals on the CSC layer. Blue is the coordinate reconstructed by the standard approach, red is the coordinates detected by continuous wavelet analysis, green is the simulated muon trajectory.

Reconstruction of cosmic rays at the TAIGA gamma-ray observatory. The TAIGA observatory is designed to study cosmic gamma radiation and charged cosmic rays in the energy range 10^{13}–10^{18} eV. Galactic cosmic rays are accelerated to energies $E \sim Z\times10^{15}$ eV in the processes of supernova explosions. The fundamental questions for this energy range are still unanswered. First of all, this is the question of sources of galactic cosmic rays with energies of about 1 PeV (the region of the classical "knee" in the energy range 10^{15}–10^{17}eV), which is the most probable limit in galactic accelerators.

The energy of gamma rays from such accelerators should be ~300 TeV, and several such photons have recently been measured. Thus, the knee energy range is an important area for understanding the origin, acceleration, and propagation of cosmic rays in our galaxy, but it is also a transition region from galactic to extragalactic cosmic rays. The cosmic ray flux beyond 10 TeV decreases rapidly, requiring a large effective detector area. The TAIGA gamma-ray observatory combines several atmospheric Cherenkov telescopes (IACT) with a network of TAIGA-HiSCORE wide-aperture Cherenkov optical detectors. This makes it possible to expand the detector area to several square kilometers and significantly suppress the background from charged cosmic rays. The development and software implementation of algorithms for the reconstruction of cosmic rays and high-energy gamma rays on data from the TAIGA-HiSCORE and TAIGA-IACT detectors are an urgent task. This, in turn, will require the use of the Monte Carlo simulation of the background signal for its effective suppression, as well as methods of artificial neural networks, machine learning and cellular automata for the reconstruction of cosmic rays.

Track reconstruction in the proton digital calorimeter for proton therapy. At the moment, the use of proton therapy in medicine in the treatment of tumor patients has great development and prospects. There is a need to develop more accessible equipment for the affordable implementation of this method in the work. At JINR, work to develop a high-speed digital proton tomograph on a medical accelerator is underway. The technical characteristics of the equipment being developed should allow obtaining high-resolution images based on the reconstruction of proton tracks. It is proposed to reconstruct tracks in the proton digital calorimeter using a mathematical apparatus based on a cellular automaton.

Processing and analysis of neutron noise of the IBR-2M reactor. The goal is to improve methods for processing reactor data in the task of increasing the level of nuclear safety of the IBR-2M reactor, namely, the development of methods for approximating and smoothing functions from data based on the basis element method (BEM) for processing and analyzing reactor neutron noise. In recent decades, MLIT has developed a new 4-point approach to solving these tasks, i.e., BEM. It is based on the idea of P.L. Chebyshev on the approximation of a function on a segment of limited length (rather than at a single point). It allows suppressing errors and creating stable algorithms for processing experimental data. In the methods of BEM-piecewise-polynomial approximation (PPA) and root-mean-square PPA (RMSPPA), the regularization of the problem occurs due to the choice of the structure of the internal connection of the variable with the control parameters. On this basis, methods and algorithms that effectively suppress errors and provide a high rate of convergence of data processing have been developed. On the basis of BEM polynomials, there have been developed a number of effective methods for processing experimental data, such as the high-order piecewise polynomial approximation method (BEM-PPA); the method of root-mean-square PPA (MBE-RMSPPA) of high orders; the method of 2D-approximation of curves of complex topology (BEM-2DSA); a method for numerically solving the Cauchy problem, suitable for solving stiff problems using an explicit “predictor-corrector” scheme (BEM-PC), etc. High-order RMSPPA provides a more accurate result than a cubic spline, with a much smaller number of nodes.

Software complex for data processing of the YuMO small-angle neutron scattering spectrometer. The method of small-angle neutron scattering is an effective method for studying fundamental problems in various fields of science, the most important feature of which is the possibility of analyzing the structure of disordered systems. The project of a new unique large-area position-sensitive detector, being implemented at FLNP, not only expands the possibilities for increasing data collection (almost a thousand times), but also opens up new qualitative opportunities for studying anisotropy on the nanoscale, expanding the sample environment, etc. A key element in application of the method is experimental data processing. To solve a whole range of tasks with the YuMO spectrometer, the software package for the fast and efficient processing of experimental data is fundamental and decisive. The purpose of the proposed study is the development of algorithmic support and a software package for data processing of the YuMO small-angle neutron scattering spectrometer at the IBR-2 reactor, which is extremely popular and relevant. The scientific novelty of the research, at the same time, is determined by both the uniqueness of the YuMO small-angle spectrometer and specific ring detectors for detecting thermal neutrons, as well as by a position-sensitive detector of unique geometry and a large area for collecting experimental data.

The basis for the development of the software package is a combination of specialists in small-angle neutron scattering, creators of a multi-detector system, including the position-sensitive detector, and programmers working in this direction and, in fact, having created a software package for the multi-detector system without the position-sensitive detector and supporting this package for more than 20 years. It is supposed to convert the methodological developments into a unified software package, considering the increase in the volume of files by a thousand times with the processing of the results with a small delay (up to several minutes) after the completion of measurements of the next sample.

III. Support and development of the software environment for experiments

Modern physics facilities produce a huge amount of experimental data, the processing of which requires significant computing resources. For example, experiments at the Large Hadron Collider have so far collected about 1 EB of data, which are processed using 900,000 CPU cores, combined in the Worldwide LHC Computing Grid. Along with experiments at the LHC, other facilities with comparable scales of expected data flows and high requirements for data processing systems, both in high-energy physics and in other fields of science (astronomy, thermonuclear fusion, etc.), are currently being built or designed. The MPD and SPD facilities being created at the NICA collider are among them.

The purpose of this section of the project is the elaboration, commissioning, maintenance and development, during the life cycle of the experiment, of a software environment of information and computing systems and services to ensure the high-throughput processing of the collected and simulated data of the experiments of the NICA project based on the JINR Multifunctional Information and Computing Complex (MICC). The specified complex of systems and services of the intermediate level should provide the possibility of implementing various models (scenarios) of data processing, depending on the needs of the experiment; ensure that available computing and storage resources are used efficiently enough in accordance with the agreed policy of the experiment. In addition, the development of information and computing systems for the analysis and processing of experimental data in the field of radiobiology, life sciences, ecology and the creation of a number of information services for experiments at the LHC are expected.

The scientific novelty in the field of computing and application software includes:

- implementation of multi-stage data processing within online processing using artificial intelligence technologies;
- implementation of fundamentally new methods, models and processing processes related to the use of artificial intelligence technologies in a distributed heterogeneous computing environment, for example, the organization of learning processes and retraining of neural networks; distribution, cataloging and storage of neural networks used in high-throughput data processing;
- adaptation of "traditional" processing methods to the architectures of computing infrastructures that change over time.

At the same time, a number of information systems and services created during the project may be required by other experiments that face the need to process data in a distributed environment. First of all, these include authentication and authorization systems, access interfaces to computing resources and data storage resources, a service for mass data transfer between processing centers.

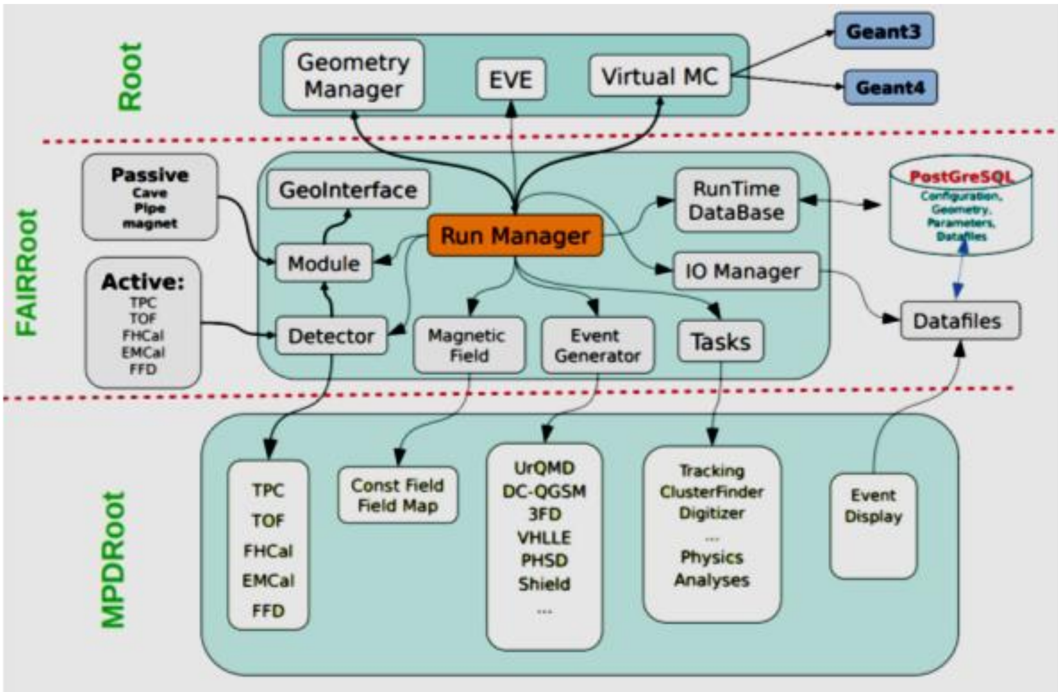

Fig. 6. Structure of the MPD software platform.

Development of the software environment for processing and analyzing data from the MPD experiment. The software environment of the experiment, i.e., mpdroot, is the main one for processing data from the MPD detector of the NICA megaproject (Fig. 6). Without its stable and correct functioning, the possibility of obtaining a reliable physics result of the experiment may be questioned. In addition, the ease, understandability and accessibility of the software environment for members of the international MPD collaboration enhances the efficiency of solving tasks of detector calibration, data acquisition, accumulation, processing and physics analysis.

The novelty of the research is determined by the uniqueness of the experimental setup, the physics problems solved in the experiment and, as a result, the uniqueness of the requirements and requests for the software environment of the experiment.

Since the environment is based on the C++ programming language, the rules of development through OOP testing will be applied during its development, with an emphasis on analyzing the requirements of users and the project. To update the system for building and distributing the mpdroot package, the system for building packages (based on ALICE aliBuild) and distributing compiled packages (CernVM-FS) will be adapted.

Creation, implementation and development of an information and computing complex for processing, analyzing and storing data for the SPD experiment (Fig. 7). Tasks in this direction include:

- development of a model for processing and organizing data in the SPD experiment;
- development of methods and approaches to filtering data in real time using artificial intelligence technologies;
- determination of technical requirements, design and creation of a software and hardware complex for organizing the high-throughput processing of data received from the data acquisition system of the SPD facility (SPD OnLine Filter);
- development of algorithms for reconstructing events in the SPD detector using machine learning methods using deep neural networks (see section II of the project), integration of the developed algorithms into the software platform of the applied software of the experiment;
- implementation of a system for distributed data processing of the SPD experiment on the resources of the collaboration participants.

The development of methods and approaches to filtering SPD data in real time will be implemented using artificial intelligence technologies, including the definition of technical requirements, design and prototyping of software and hardware solutions.

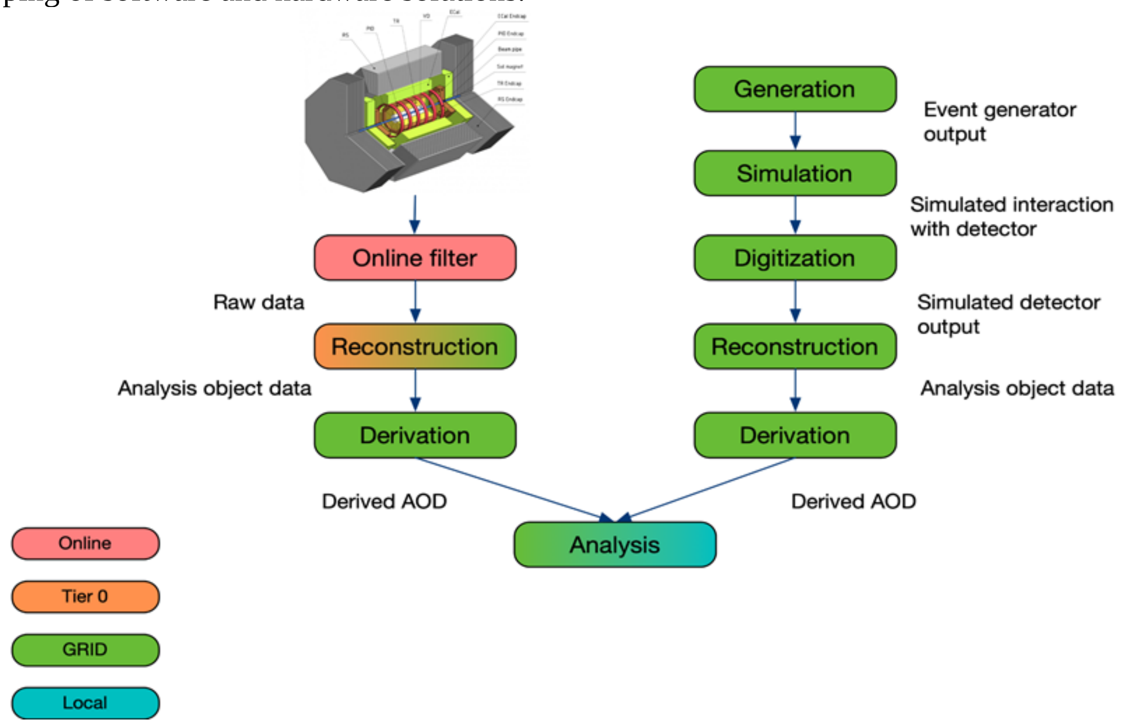

Fig. 7. General scheme for processing and analyzing data from the SPD experiment.

The fast reconstruction and filtering of data coming from the SPD facility cannot be performed on a single node. It is required to design a specialized computing complex and a control system for high-throughput processing in this specialized computing system. The data flow from the facility, which will reach values of the order of 20 GB/s, causes high demands on the performance of data storage systems and on the speed of data processing algorithms. To ensure the required data processing speed, it is proposed to use methods based on machine learning, since they have high speed and are capable of parallelization on graphics coprocessors (GPUs).

The prototype of the distributed data processing system of the SPD experiment implies the use of a hierarchical structure of distributed data processing tools with several levels (Fig. 8). The Tier1 computing center must provide long-term, high-capacity storage (including tape drives) that will have sufficient capacity to store a complete copy of the primary data and a significant amount of important derivative data. At the location of the experimental facility, the Tier1 computing center is JINR. Tier2 data centers must provide (temporary) storage with sufficient capacity to model, process or analyze a certain amount of data. With decades of data processing experience at the LHC, a set of technologies mature enough to create high-throughput distributed computing systems for high-energy physics experiments has already been developed; they can be used in the SPD experiment. Thus, for task management, it is planned to use the PANDA framework. The RUCIO package will be used for distributed data management. One can use FTS for mass data transfer. However, although the key tools are already in place, they need to be adapted to the tasks and data structure of SPD. In addition, it is required to develop a high-level orchestration system that will manage low-level services. The main task of this system will be to provide efficient, highly automated multi-stage data processing, taking into account the features of the SPD experiment.

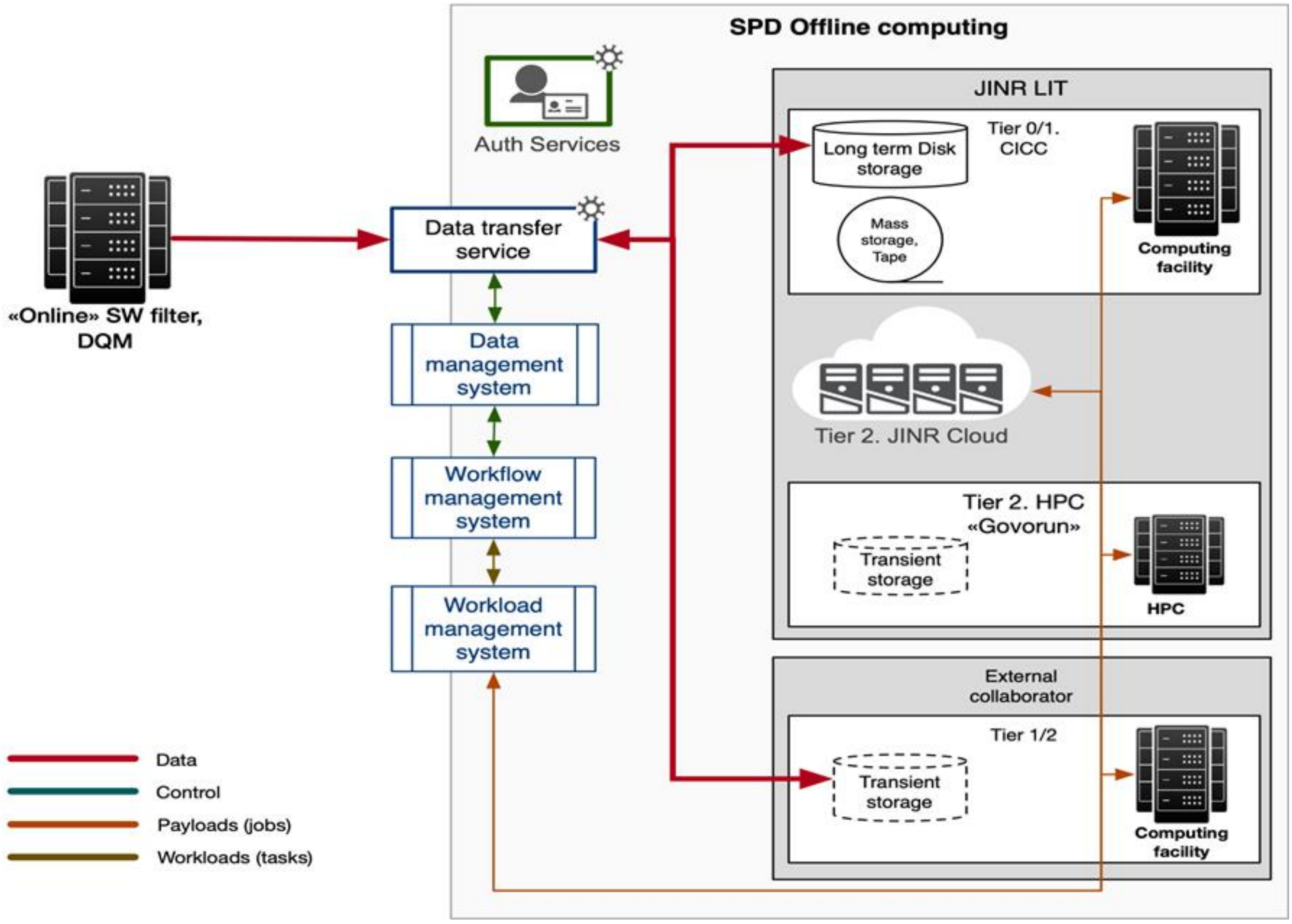

Fig. 8. Structure of the distributed data processing system of the SPD experiment.

The development of the software environment for processing and analyzing data from the SPD experiment includes the adaptation and optimization of the data structure itself, modeling and reconstruction algorithms used in the SpdRoot collaboration software, for the effective use of multi-core computers in modeling and reconstruction tasks (parallelization task). The basic functionality of this kind is already implemented in the FairRoot framework, which is the basis of SpdRoot. In addition, an alternative possibility of using the Key4HEP software toolkit created within the international consortium HEP Software

Foundation for offline processing in the SPD experiment will be explored. This will require the adaptation of the data structure, modeling and reconstruction algorithms used in the SPD experiment to the Key4HEP framework. To solve this task, it is planned to cooperate with the LPI group, which has extensive experience in creating software for reconstructing tracks, vertices and identifying particles in elementary particle detectors.

Data processing system for the BAIKAL-GVD neutrino telescope. High-energy neutrino astrophysics is a young field of research at the intersection of neutrino physics, astrophysics and high-energy physics, which received a new boost with the development of next-generation large neutrino telescopes, which is the Baikal Deep Sea Neutrino Telescope (BAIKAL-GVD). The BAIKAL-GVD project is one of the key ones for JINR, thus ensuring the importance of developing software for the data processing system for the neutrino telescope. The developed system is a unified complex modular software package that solves the tasks of data transfer, storage, processing and neutrino event reconstruction under conditions close to real-time.

Creation of specialized databases and information systems. Modern scientific research in the field of experimental high-energy physics is characterized by duration, complexity, high labor intensity, large time and financial costs, handling large amounts of data recorded during the experiment. In this regard, for any experiment, the task of automating the process of collecting, processing and analyzing experimental data is of particular relevance. The automation of a modern physics experiment is impossible without the use of information and computing software that allows one to collect, store and process a large amount of information, to manage the experiment in the process of its implementation, to simultaneously maintain a large number of the equipment of the experimental setup and perform other actions necessary for the timely receipt of high-quality physics results. Thus, work on the elaboration, implementation and development of automated information systems (AIS) that ensure the fulfillment of the above tasks, as well as work on the creation and modernization of databases (DB), which are the basis for the AIS functioning, are relevant.

The novelty of the research is ensured by the uniqueness of the addressed experiments (ATLAS, BM@N, MPD). Various databases, such as geometric, configuration, states, metadata, and time series databases, classified both by the type of information stored and by the nature of data organization, are used in almost all experiments in high-energy physics. However, despite the similarity of the tasks that arise in different experiments, considering the specifics of the experiment, different types of databases and the corresponding AIS must be developed anew for each experiment. In addition, the search and analysis of features that can be used to adapt the types of systems under consideration for different experiments is an extremely important task.

Development and creation of an information and computing system for automating the processing of data from radiobiological studies. The automation of the data processing of radiobiological studies carried out at LRB is a paramount task associated with the study of physically induced neurodegenerative changes in the central nervous system of laboratory animals. A complete understanding of the impact process and a qualitative picture of the consequences of these impacts on biosystems require the systematization and simultaneous processing of a significant amount of data related to various aspects of the manifestation of impacts. Within the project, it is proposed to create an information and computing system (ICS) that provides, based on algorithmic approaches and IT solutions, the study of morphological and functional changes in the central nervous system during physically induced neurodegenerative changes studied at the JINR Laboratory of Radiation Biology, namely, the effects of exposure to various types of ionizing radiation on the body of small laboratory animals.

The ICS being created will allow combining a set of services and tools for experimental data storage, its analysis, algorithmic blocks based on the methods and approaches of machine and deep learning, which allow analyzing video data obtained during behavioral tests and analyzing images obtained during histological studies. At the same time, the scientific novelty lies in the creation of an ICS that is unique in functionality, combining and structuring heterogeneous experimental data obtained at different experimental stages by pathomorphological methods and methods for assessing the behavior of laboratory animals into a unified information space that can provide both ease of storage and access to data, and a set of advanced (relevant) algorithmic procedures for automating data analysis (based on machine learning methods, deep learning, neural network approaches), which will detect new patterns in the development of neurodegenerative pathologies and reveal more deeply the mechanisms of the neurotoxic effect of ionizing radiation.

To solve the set task, three groups of methods based on machine learning methods and neural network approaches, statistical analysis will be developed and applied. The first group of methods includes methods for analyzing video sequences. The second group of computer vision methods is associated with the task of processing morphological data (images of histological sections of various biological tissues). The third group comprises technologies for implementing a software environment based on modern IT solutions, including web technologies, modern inference solutions and data analysis visualization components.

Development of the intelligent platform for determining the state of agricultural and decorative plants. Food security issues are of extreme importance in today's world. Despite the availability of specialized crops and modern technologies, the total crop losses from diseases and pests, according to various experts, range from 10 to 30%. The development of technologies capable of minimizing losses in agriculture is an important area of research. The purpose of this study is to develop the intelligent platform for determining the state of agricultural and decorative plants. Artificial intelligence technologies are increasingly being used in agriculture. Since 2018, MLIT has been developing a platform for identifying plant diseases based on images and text descriptions. To date, when determining plant diseases, it has been possible to achieve an accuracy of 98% on images from real life. Within the project, it is planned to search for additional options to optimize the architecture and the learning process. For this, various algorithms for finding optimal augmentation policies, loss minimization functions and the possibility of using unsupervised networks on a large number of plant images will be considered. The novelty of the study is determined by the use of methods that have not been used in the tasks of classifying images of plant diseases, namely

- application of algorithms for the automatic selection of optimal data augmentation policies (auto augmentation);
- testing of various loss minimization functions (triplet loss, arcface, cosface, sphereface);
- application of unsupervised learning algorithms for solving tasks of the replenishment of image databases and optimization of basic model architectures;
- determination of the most effective approaches for classifying images with plant diseases.

Development of the platform for monitoring and predicting the state of the environment. Issues of environmental pollution control and environmental safety are always relevant. Various state and international programs have been launched to control and monitor the environment. Today, environmental pollution control projects are beginning to use the technologies of the Internet of things, artificial intelligence and Big Data, which allows one to talk about the transition to intelligent digital environmental monitoring platforms that can generate new knowledge based on incoming and existing data, as well as independently make decisions, which previously required the involvement of experts.

Within the cooperation between MLIT and FLNP for the UN Commission program on long-range transboundary transport of air pollution ICP Vegetation, an intelligent monitoring platform to identify the most disadvantaged areas in Europe and Asia, to create regional maps and generally increase the understanding of the nature of long-term transboundary pollution by the research community, bringing together scientists from 43 countries from Europe and Asia, is being developed. The platform is based on data from the analysis of FLNP samples, which provide information on the concentration of various elements at collection sites. Their coordinates can be used to get indexes of various satellite programs. The index includes the name of the satellite program whose data is used, the size of the analyzed area, the identifier of the spectral channel in which the survey was carried out, and the mathematical function applied to the digital matrix of the resulting image. The platform implements a forecasting mechanism based on the use of machine learning in conjunction with Earth remote sensing data. The indexes are calculated using the Google Earth Engine (GEE) platform, which contains data from dozens of different satellite programs and products. In the current implementation, depending on the amount of initial data, statistical machine learning models or neural networks are used, regression and classification tasks are solved. As a result, when building global and regional maps, the accuracy of models reaches 90 - 95%. The Python programming interface is used to work with GEE, since at the first stage, the search for the most promising indexes is carried out. More than 40 different collections, several variants of data aggregation functions and sizes of the analyzed areas are used. The most promising indexes, together with data on the concentration of metals at sampling sites, are used to train the models. This data is fed into the model, and a forecast is built. Thanks to GEE and machine learning, we can track changes in the air pollution situation much faster, get detailed information in areas of interest and in areas where sampling is not possible, and in the future even partially automate the process of environmental monitoring.

Budget justification

The financial costs of the project include the costs for international scientific and technical cooperation (participation in conferences, reception of staff members and students to perform joint work), the purchase of computer equipment (personal computers and components, GPUs for testing high-speed parallel computing) and software. The successful implementation of the project, along with free software, requires the use of paid software, both general-purpose software for supporting the workflow at workplaces (operating systems, office software, etc.), and specialized software for the organization and work with databases (such as PostgreSQL), the organization of virtual stations (e.g., VMware), the effective debugging of developed and supported within large software systems (such as TotalView).

Expected results

- Completion of the revision of the interaction generators and their development for modeling the processes of interactions of light and heavy nuclei, including at NICA energies
  - revision of the FTF and QGSM models and inclusion of the developed software modules for modeling nuclear interactions of unstable, long-lived particles, namely, charmed hadrons, bothonium, light hyper-nuclei and anti-hyper-nuclei, after comprehensive verification and testing in the Geant4 package;
  - development of the DCM-QGSM-SMM generator: taking into account the dependence of the lifetime of resonances on the density of the nuclear medium, suppression of the cross section of the production of pseudoscalar mesons in a dense nuclear medium, enhancement of the production of hyperons in a dense nuclear medium; enhancement of the dilepton yield, replacement of the Woods-Saxon model of colliding nuclei with a lattice one, which will allow including the deformation of nuclei;
  - obtaining unambiguous information about the admixture of quark states at NICA energies in the weakly excited and ground states of the deuteron;
- Completion of the revision of the interaction generators and their development for modeling the processes of production of particles-candidates for the role of dark matter, extra Higgs bosons and processes that violate the lepton number for the LHC conditions
  - revision of the space of model parameters of a number of dark matter models available for study at the LHC at a nominal energy and total integrated luminosity up to 350 fb-1, carrying out the complex modeling of the processes of formation of dark matter particles within simplified and extended Higgs models;
  - carrying out the complex modeling of processes that occur with the violation of the lepton number, using the QBH, Pythia, MadGraph generators, etc.
- Development of algorithms for the reconstruction of charged particle tracks for experimental facilities, including those at NICA and the LHC, creation of appropriate software and its application for data processing and analysis, study of the physical and technical characteristics of detector systems
  - universal algorithm based on wavelet analysis for the recognition of overlapping signals in tracking detectors of HEP experiments with the amplitude representation of the signal;
  - assessment of background particle rates and the effect of aging in the CSC detectors of the CMS experiment on various types of data and under various data collection conditions;

- optimization of the tracking algorithm in the DCH detectors of the BM@N experiment for the data set from the interaction of heavy ions with the target. Automation of obtaining the transfer function for the DCH detectors of the BM@N experiment. Improvement of charged particle trajectory reconstruction in the CSC detectors of the BM@N experiment;
- elaboration and implementation of modeling and data processing methods, as well as their development and adaptation for current configurations of a number of tracking detectors (Forward Silicon, GEM, small CSC, large CSC, Silicon Beam Tracker, Silicon Profilometer.) of the BM@N experiment tracking system;
- development of a methodology and software implementation of the system for the global alignment of detectors in the BM@N experiment;
- technique for identifying charged particles, in particular, in the MPD detector, based on machine learning methods, software implementation of the developed approaches in the mpdroot experiment processing and analysis software environment, integration of modern machine and deep learning libraries into the mpdroot software environment;

▪ Creation and development of a data processing and analysis system for the experiments at the JINR NICA accelerator complex

- formation of a version of the mpdroot package that works with clear support rules for the entire duration of the MPD experiment, with regular updates of the entire working environment (in terms of software) and software adaptation to current project requirements;
- realization, methodological confirmation and implementation of data processing algorithms using artificial intelligence elements into the application software of the experiment;
- implementation and commissioning of a set of systems and services that provide data processing for the SPD experiment in a distributed heterogeneous computing environment, with support for predictable loads for timely data processing at the first stage of the experiment;
- creation of auxiliary information services for the experiments of the NICA megaproject: logbook, e-log, a single graphical interface for managing the experiment, information service, implementation of interfaces for data conversion in the subsystems of the experiment, creation of specialized databases, monitoring systems, other services requiring general IT support;

▪ Support and development of the software environment for experiments at the LHC

- development of a prototype of the CREST project as part of the ATLAS experiment for Run 4 of the LHC, support of the operation of the Pickup Service and monitoring of the EventIndex project of the ATLAS experiment;

▪ Development of the data processing and analysis system for the reconstruction of experimental events within the JINR program in the field of neutrino physics

- debugging the data processing system on a large amount of experimental data and achieving its performance at the level of the main processing system in the BAIKAL-GVD project;
- development of software for the reconstruction of cosmic rays and high-energy gamma quanta based on data from the TAIGA-HiSCORE and TAIGA-IACT detectors;

▪ Development of algorithms and software for JINR research projects in the field of neutron physics

- development of a software package for the primary processing of small-angle experimental data from the YuMO spectrometer for a multi-detector system with a position-sensitive detector with distributed capabilities for a combination of processing types, including normalization to fluxes, adaptation to a possible change in the pulse frequency of the IBR-2 reactor, methods for accounting for background conditions and adaptation to changes in the multi-detector system of the YuMO spectrometer;

▪ Development of algorithms, software and computing platforms for radiobiological research, applied research in the field of proton therapy and ecology

- creation of software for track reconstruction in the digital calorimeter prototype for proton therapy;
- creation of an information and computing system that provides a convenient environment for storing experimental data, analyzing the results of laboratory radiobiological studies, creating data

sets, developing and applying algorithms based on methods and approaches of machine and deep learning;
- building a hierarchical data storage and processing system for the created information and computing system for laboratory radiobiological studies based on the multi-layer storage system created on the HybriLIT platform, operating on the principle from "hot" to "cold" layers; filling the algorithmic block of the information and computing system for laboratory radiobiological studies with three main modules: a module for studying the behavioral patterns of small laboratory animals exposed to physical or chemical influences, a module for histological studies and a module for the statistical complex analysis of biological data;
- implementation of the intelligent platform for determining the condition of agricultural and decorative plants, which presents a significant number of detectable diseases, provides detailed treatment plans and uses advanced artificial intelligence technologies;
- platform for monitoring and predicting the state of the environment, combining advanced data management technologies and artificial intelligence to solve environmental monitoring tasks.

Risks (SWOT Analysis)

	Positive	**Negative**
Internal	**Strengths** • wide approbation of research (at meetings of the governing bodies of JINR and a large number of international conferences); • project participants are recognized experts in various research fields (information technology, mathematical physics, theoretical and experimental physics of elementary particles and atomic nucleus, radiobiology); • experience in the field of modeling physical processes, processing and analyzing experimental data, creating information and computing systems and software libraries in the world largest research centers (JINR, CERN, Fermilab, etc.); • experience of participation in major international collaborations (WLCG, JINR, CERN, DESY, Fermilab, etc.); • balanced composition of researchers group: 8 participants under 35 years old and 9 under 39 years old, 8 doctors and 17 candidates of sciences.	**Weaknesses** • restrictions for participants from the Russian Federation to visit a number of scientific centers and international conferences; • possible partial revision of project priorities; • partial shortage of personnel in a number of areas of the project
External	**Opportunities** • development of new methods of modeling, data processing and analysis (Big Data, deep and machine learning, etc.); • emergence of new experimental data that gives a boost to the development of theoretical and computational methods for describing and calculating physical processes; • development of the JINR MICC and other world data processing centers;	**Threats** • possible changes in the timing of the implementation of research projects at JINR and other centers; • uncertain situation with the extension of the agreement on scientific cooperation between JINR and CERN, possibility of suspending scientific cooperation with other world centers

	• increase in the availability of quantum computers	

2.3 Estimated completion date

2024-2026

2.4 Participating JINR laboratories

MLIT (see table 3.2.1)

VBLHEP (V. Yu. Aleksakhin, A.A. Aparin, Yu.V. Bespalov, D.V. Budkovski, A.V. Bychkov, I.R. Gabdrakhmanov, A.S. Galoyan, K.V. Gertsenberger, V.M. Golovatyuk, D.K. Dryablov, M.N. Kapishin, V.Yu. Karzhavin, A.A. Korobitsyn, A.V. Krylov, V.V. Lenivenko, S.P. Lobastov, S.P. Merts, A.A. Moshkin, A.A. Mudrokh, D.N. Nikiforov, M. Patsyuk, O.V. Rogachevsky, V.G. Ryabov, V.V. Shalaev, S.G. Shulga, I.A. Zhizhin, V. Zhezher, A.I. Zinchenko)

BLTP (D.I. Kazakov, M.V. Savina, O.V. Teryaev, V.D. Toneev)

FLNR (M. Balasoiu, M.-O. Dima, M.-T. Dima, A.I. Ivankov, A.H. Islamov, Yu.S. Kovalev, A.I Kuklin, Yu.N. Pepelishev, Yu.L. Ryzhikov, A.V. Rogachev, V.V. Skoy, M.V. Frontasyeva)

DLNP (V.A. Bednyakov, A.P. Belova, I.A. Belolaptikov, I.V. Borina, A.N. Borodin, V. Dik, I.I. Denisenko, T.V. Elzhov, A.A. Grinyuk, A.V. Guskov, E.V. Khramov, V.A. Krylov, V.S. Kurbatov, D.V. Naumov, A.E. Pan, D. Seitova, A.E. Sirenko, M.N. Sorokovikov, L.G. Tkachev, B.A. Shaibonov, E. Sholtan, A.C. Zhemchugov, D.Yu. Zvezdov)

LRB (I.A. Kolesnikova, Yu.S. Severyukhin, D.M. Utina)

UC (A.Yu. Verkheev, B.S. Yuldashev)

2.4.1 MICC resource requirements

Computing resources	**Distribution by year**		
	1st year	2nd year	3rd year
Data storage (TB) - EOS - Tapes	 215 100	 225 100	 230 100
Tier 1 (CPU core hours)			
Tier 2 (CPU core hours)			
SC Govorun (CPU core hours) - CPU - GPU	 1000	 1000	 1000
Clouds (CPU cores)	300	300	300

2.5. Participating countries, scientific and educational organizations

Organisation	**Country**	**City**	**Participants**	**Type of agreement**
Foundation ANSL	Armenia	Yerevan	A. Tumasyan A. Ayrapetyan A. Gevorkyan	Collaboration
GSU	Belarus	Gomel	V.V. Andreev N.V. Maximenko	Collaboration

INP BSU	Belarus	Minsk	D.V. Ermak V.V. Makarenko V.A. Mosolov	Collaboration
Univ.	Great Britain	Oxford	E. Gallas	Collaboration
INFN	Italy	Genova	D. Barberis	Collaboration
NOSU	Russia	Vladikavkaz	A.K. Gutnova	Collaboration
NRC KI PNPI	Russia	Gatchina	B.T. Kim A.K. Kiryanov	Collaboration
BM@N Collaboration	Russia	Dubna	M.N. Kapishin	Collaboration
MPD Collaboration	Russia	Dubna	V.G. Ryabov	Collaboration
SPD Collaboration	Russia	Dubna	A.V. Guskov	Collaboration
ITEP	Russia	Moscow	A.N. Nikitenko	Collaboration
SINP MSU	Russia	Moscow	E.E. Boos L.V. Dudko I.P. Lohtin O.L. Kodolova S.V. Petrushanko	Collaboration
NNRU “MEPhI”	Russia	Moscow	M.V. Danilov	Collaboration
LPI RAS	Russia	Moscow	I.M. Dremin	Collaboration
IHEP	Russia	Protvino	A.V. Petrova	Collaboration
SSU	Russia	Samara	A.V. Baskakov V.A. Saleev	Collaboration
SPbSU	Russia	St. Petersburg	A.V. Bogdanov A.B. Degtyarev A.K. Zarochentsev	Collaboration
INR RAS	Russia	Moscow, Troitsk	S.N. Gninenko	Collaboration
UTA	USA	Arlington	N. Ozturk	Collaboration
INP AS RUz	Uzbekistan	Tashkent	B.S. Yuldashev	Collaboration
CEA	France	Saclay	A.Formika	Collaboration
CERN	Switzerland	Geneva	G. Avolio Sh. Roe A. Ribon	Collaboration
ALICE Collaboration	Switzerland	Geneva	M. van Leeuwen	Collaboration
ATLAS Collaboration	Switzerland	Geneva	A.Hoecker	Collaboration

CMS Collaboration	Switzerland	Geneva	P. Mc Bride	Collaboration

2.6. Key partners *(those collaborators whose financial, infrastructural participation is substantial for the implementation of the research program. An example is JINR's participation in the LHC experiments at CERN).*

3. Manpower

3.1. Manpower needs in the first year of implementation

№№ n/a	**Category of personnel**	**JINR staff, amount of FTE**	**JINR Associated Personnel, amount of FTE**
1.	research scientists	25	2
2.	engineers	8	1
3.	specialists	0	0
4.	office workers	0	0
5.	technicians	0	0
	Total:	**33**	**3**

3.2. Available manpower

3.2.1. JINR staff

No.	Category of personnel	Full name	Division	Position
1.	research scientists	A.S. Ayriyan	MLIT	head of sector
2.	research scientists	P.G. Akishin	MLIT	deputy head of division
3.	research scientists	E.I. Alexandrov	MLIT	researcher
4.	research scientists	I.N. Alexandrov	MLIT	head of sector
5.	research scientists	D.A. Baranov	MLIT	researcher
6.	research scientists	T.Zh. Bazhenyan	MLIT	researcher
7.	research scientists	Yu.A. Butenko	MLIT	junior researcher
8.	research scientists	J. Busa	MLIT	senior researcher
9.	research scientists	N.N. Voytishin	MLIT	researcher
10.	research scientists	S. Gnatich	MLIT	lead researcher

11.	research scientists	P.V. Goncharov	MLIT	intern-researcher
12.	research scientists	O.A. Grygoryan	MLIT	senior researcher
13.	research scientists	O.Yu. Derenovskaya	MLIT	scientific secretary
14.	research scientists	N.D. Dikusar	MLIT	lead researcher
15.	research scientists	V.B. Zlokazov	MLIT	lead researcher
16.	research scientists	M.I. Zuev	MLIT	researcher
17.	research scientists	V.V. Ivanov	MLIT	chief researcher
18.	research scientists	A.A. Kazakov	MLIT	head of group
19.	research scientists	Yu. V. Korsakov	MLIT	intern-researcher
20.	research scientists	B.F. Kostenko	MLIT	senior researcher
21.	research scientists	M.A. Mineev	MLIT	researcher
22.	research scientists	Zh.Zh. Musulmanbekov	MLIT	senior researcher
23.	research scientists	E.G. Nikonov	MLIT	head of sector
24.	research scientists	G.A. Ososkov	MLIT	chief researcher
25.	research scientists	V.V. Palichik	MLIT	lead researcher
26.	research scientists	V.V. Papoyan	MLIT	intern-researcher
27.	research scientists	I.S. Pelevanyuk	MLIT	researcher
28.	research scientists	D.V. Podgainy	MLIT	head of sector
29.	research scientists	D.I. Pryahina	MLIT	researcher
30.	research scientists	I. Satyshev	MLIT	junior researcher
31.	research scientists	K.V. Slizhevsky	MLIT	intern-researcher
32.	research scientists	A.G. Solovyev	MLIT	senior researcher
33.	research scientists	T.M. Solovyeva	MLIT	senior researcher
34.	research scientists	O.I. Streltsova	MLIT	senior researcher
35.	research scientists	Z.K. Tuhliev	MLIT	researcher
36.	research scientists	V.V. Uzhinsky	MLIT	lead researcher
37.	research scientists	Z.A.Sharipov	MLIT	senior researcher

38.	research scientists	S.V. Shmatov	MLIT	director of laboratory
39.	research scientists	A.V. Yakovlev	MLIT	researcher
40.	engineers	E.P. Akishina	MLIT	lead programmer
41.	engineers	A.I. Anikina	MLIT	software engineer
42.	engineers	P.I. Gluhovetz	MLIT	laboratory assistant
43.	engineers	I.S. Kadochnikov	MLIT	software engineer
44.	engineers	A.I. Kazimov	MLIT	lead programmer
45.	engineers	A.V. Nechaevsky	MLIT	lead engineer
46.	engineers	D.A. Oleynik	MLIT	lead programmer
47.	engineers	A.Sh. Petrosyan	MLIT	lead programmer
48.	engineers	S.K. Slepnev	MLIT	software engineer
49.	engineers	A.V. Uzhinsky	MLIT	lead engineer
	Итого:			

3.2.2. JINR associated personnel

No.	Category of personnel	Partner organization	Amount of FTE
1.	research scientists		
2.	engineers		
3.	specialists		
4.	technicians		
	Total:		

4. Financing

4.1 Total estimated cost of the project

450 thousands of US dollars

4.2 Extra funding sources

Project Leader ______________/_______________/

Date of submission of the project to the Chief Scientific Secretary: __________

Date of decision of the laboratory's STC: __________ document number: __________

Year of the project start: _________________

Proposed schedule and resource request for the Project

Expenditures, resources, funding sources			**Cost (thousands of US dollars)/ Resource requirements**	**Cost/Resources, distribution by years**		
				1st year	2nd year	3rd year
		International cooperation	180	60	60	60
		Materials				
		Equipment, Third-party company services	165	50	55	60
		Commissioning				
		R&D contracts with other research organizations				
		Software purchasing	105	30	35	40
		Design/construction				
		Service costs (*planned in case of direct project affiliation)*				
Resources required	**Standard hours**	Resources				
		– the amount of FTE,				
		– accelerator/installation,				
		– reactor,…				
Sources of funding	**JINR Budget**	JINR budget *(budget items)*	450	140	150	160
	Extra fudning (supplementary estimates)	Contributions by partners Funds under contracts with customers Other sources of funding				

Project Leader __________/_________________/

Laboratory Economist __________/_________________/

APPROVAL SHEET FOR PROJECT

TITLE OF THE PROJECT

Mathematical methods, algorithms and software for modeling physical processes and experimental facilities, processing and analyzing experimental data

SHORT DESIGNATION OF THE PROJECT

PROJECT CODE

THEME

05-6-1119-2014/…

NAME OF THE PROJECT LEADER

Sergei Shmatov

AGREED

JINR VICE-DIRECTOR	SIGNATURE	NAME	DATE
CHIEF SCIENTIFIC SECRETARY	SIGNATURE	NAME	DATE
CHIEF ENGINEER	SIGNATURE	NAME	DATE
LABORATORY DIRECTOR	SIGNATURE	NAME	DATE
CHIEF LABORATORY ENGINEER	SIGNATURE	NAME	DATE
LABORATORY SCIENTIFIC SECRETARY	SIGNATURE	NAME	DATE
THEME LEADER	SIGNATURE	NAME	DATE
THEME LEADER	SIGNATURE	NAME	DATE
PROJECT LEADER	SIGNATURE	NAME	DATE
APPROVED BY THE PAC	SIGNATURE	NAME	DATE