# Tunka-Rex Virtual Observatory

**V. Lenok,**[a,*] **D. Kostunin,**[b] **O. Kopylova, P. Bezyazeekov,**[c] **D. Wochele,**[a] **F. Polgart,**[a] **S. Golovachev,**[d] **V. Sotnikov**[d] **and E. Sotnikova**[e] **for the Tunka-Rex Collaboration**
(a complete list of Tunka-Rex authors can be found at the end of the proceedings)

[a] *Karlsruhe Institute of Technology, Institute for Astroparticle Physics, D-76021 Karlsruhe, Germany*

[b] *DESY, 15738 Zeuthen, Germany*

[c] *Applied Physics Institute, Irkutsk State University, 664020 Irkutsk, Russia*

[d] *JetBrains Research, 194100 St. Petersburg, Russia*

[e] *Sobolev Institute of Mathematics, 630090 Novosibirsk, Russia*

*E-mail:* contact@tunkarex.info

Tunka-Rex (Tunka Radio Extension) was a detector for ultra-high energy cosmic rays measuring radio emission for air showers in the frequency band of 30-80 MHz, operating in 2010s. It provided an experimental proof that sparse radio arrays can be a cost-effective technique to measure the depth of shower maximum with resolutions competitive to optical detectors. After the decommissioning of Tunka-Rex, as last phase of its lifecycle and following the FAIR (Findability — Accessibility — Interoperability — Reuse) principles, we publish the data and software under free licenses in the frame of the TRVO (Tunka-Rex Virtual Observatory), which is hosted at KIT under the partnership with the KCDC and GRADLCI projects. We present the main features of TRVO, its interface and give an overview of projects, which benefit from its open software and data.

*Presenter

PoS(ICRC2021)421

## 1. Introduction

In the era of large astrophysical facilities, which have a lifecycle exceeding decades and produce a huge amount of heterogeneous data the one of the most important question is related to the conservation of scientific knowledge acquired by them. The attempts of solution of these issues, which actually relate to the different fields of human knowledge, have resulted into formulation of FAIR (Findability — Accessibility — Interoperability — Reuse) data principles[1].

One of the most successful example of publication of cosmic-ray data in open access is the KCDC (KASCADE Cosmic Ray Data Centre)[2] project [1], which contains data from the particle detectors, namely energy deposits and arrival times at each KASCADE-Grande station. Contrary to particle detectors, digital radio technique is relatively young and has its own specifics [2] and thus, publication of data from cosmic-ray radio detectors aims at slightly different features. Since the methods for the radio detectors is actively developed over last years, it is important to provide access to the lower levels of data, i.e. amplitudes of electromagnetic fields on the antennas (one can find the same approach in data releases from gravitational wave[3] or fast radio burst detectors[4].

We develop the first framework for the quick and easy access to the data of radio detector designed for the cosmic-ray measurements taking into account features of this particular technique and focusing on the most important requirements from the field. Tunka-Rex Virtual Observatory (`TRVO`) is based on this framework and includes radio data from the Tunka-Rex detector[5] [3] operating in 2012-2019 in the frame of TAIGA facility [4]. In this work we describe the actual status of `TRVO`, its development since first prototype [5], the structure of the data and architecture of the software.

## 2. Tunka-Rex detector and data

Tunka Radio Extension (Tunka-Rex) was a digital antenna array operating at the Tunka Advanced Instrument for cosmic rays and Gamma Astronomy (TAIGA) observatory in 2012-2019. In Fig. 1 one can see the layout of the facility before Tunka-Rex decommission and note that the cosmic-ray setups are grouped in clusters: 19 clusters in a dense core and 6 satellite clusters. Each core cluster is equipped with 3 Tunka-Rex antenna stations, while satellite clusters contain one antenna station, each, and no Tunka-Grande scintillators.

At its last stage, Tunka-Rex consisted of 57 antenna stations located in the dense core of TAIGA (1 km$^2$) and 6 satellite antenna stations expanding the sensitive area of the array to 3 km$^2$. Tunka-Rex has been commissioned in 2012 with 18 antenna stations triggered by the air-Cherenkov array Tunka-133. In the following years, Tunka-Rex was upgraded several times. The TAIGA facility was enhanced by the Tunka-Grande scintillator array providing a trigger for Tunka-Rex since 2015. One can see the timeline of the Tunka-Rex development in Fig. 2.

Each Tunka-Rex antenna station consists of two perpendicular active Short Aperiodic Loaded Loop Antennas (SALLA) [6] pre-amplified with a Low Noise Amplifier (LNA). Signals from the
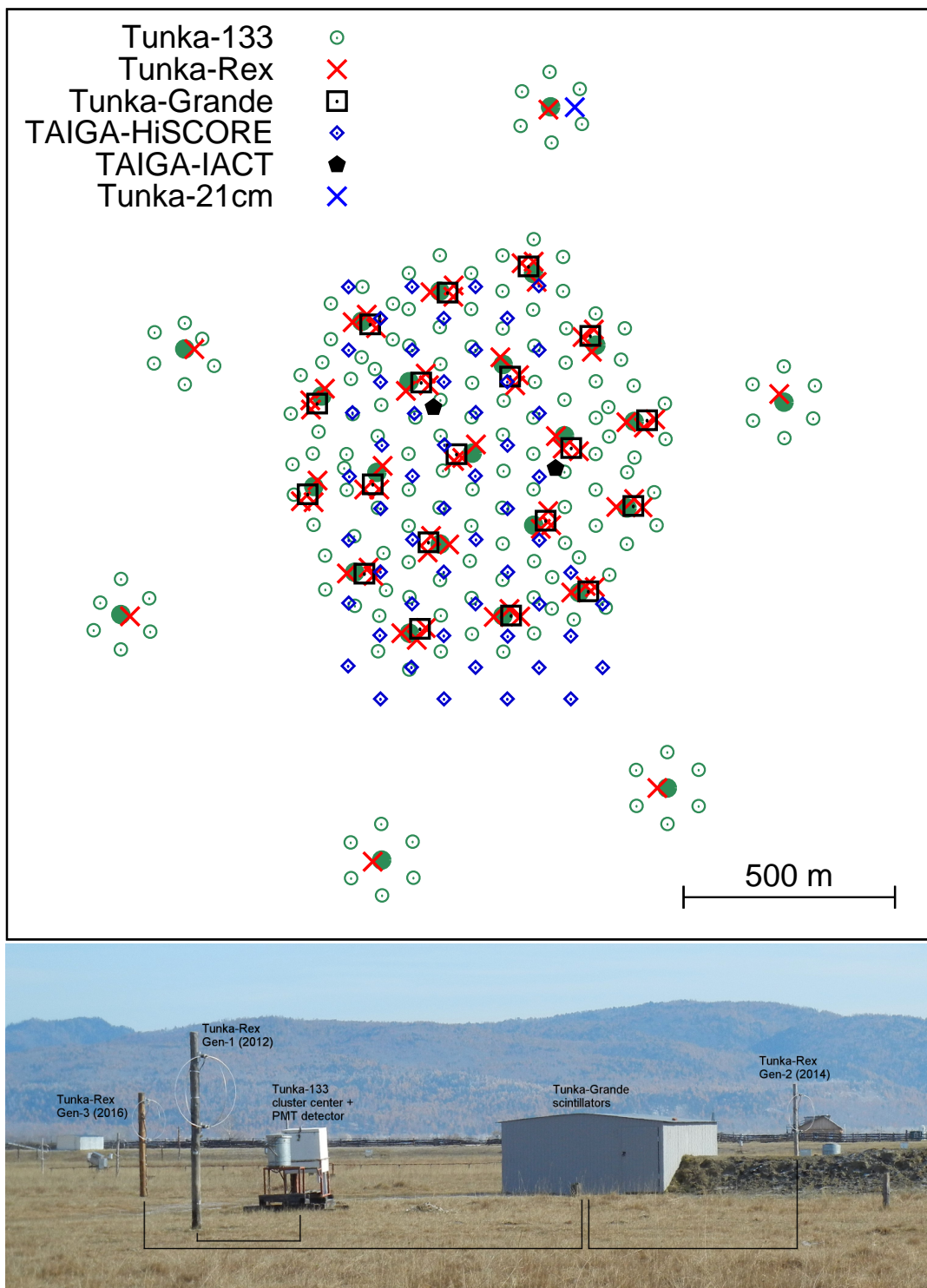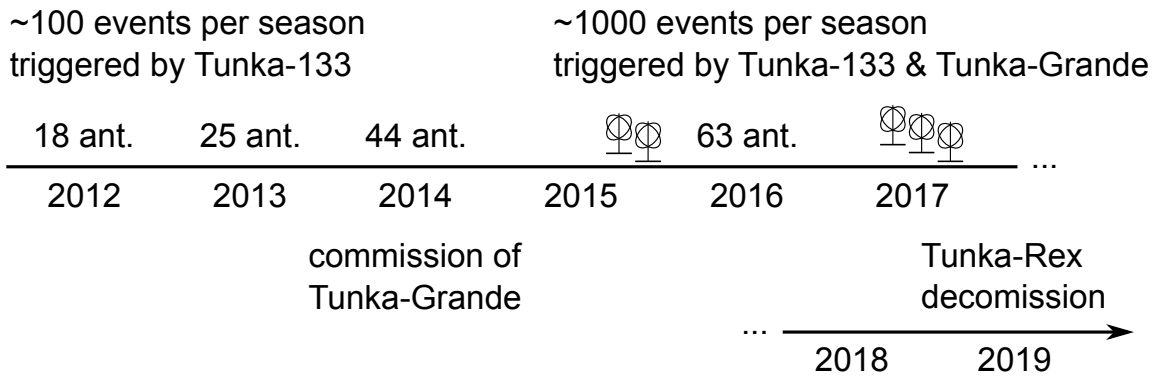
---

[1] https://www.go-fair.org/
[2] https://kcdc.iap.kit.edu/
[3] https://www.gw-openscience.org/
[4] https://chime-frb-open-data.github.io/
[5] https://www.iap.kit.edu/tunka-rex/

**Figure 1:** The layout of the TAIGA setups as of 2019. The antenna stations are depicted as crosses (the Tunka-21cm array is depicted by a single marker due to scale of the map). The layout and hardware configuration was changes several times during the past years. *Bottom:* Photo of a single cosmic-ray cluster of the TAIGA facility. Lines mark the cable connections.

~100 events per season
triggered by Tunka-133

~1000 events per season
triggered by Tunka-133 & Tunka-Grande

| 18 ant. | 25 ant. | 44 ant. | | 63 ant. | | ... |
|---------|---------|---------|------|---------|------|-----|
| 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | |

commission of
Tunka-Grande

Tunka-Rex
decomission

... 2018    2019

**Figure 2:** The antenna array has been commissioned in 2012 with 18 antenna stations triggered by the Tunka-133 air-Cherenkov detectors. Since the commissioning of Tunka-Grande in 2014-2015, Tunka-Rex additionally receives a trigger from Tunka-Grande (during daytime measurements). Starting from 2018, we are working on the public access of Tunka-Rex software and data.
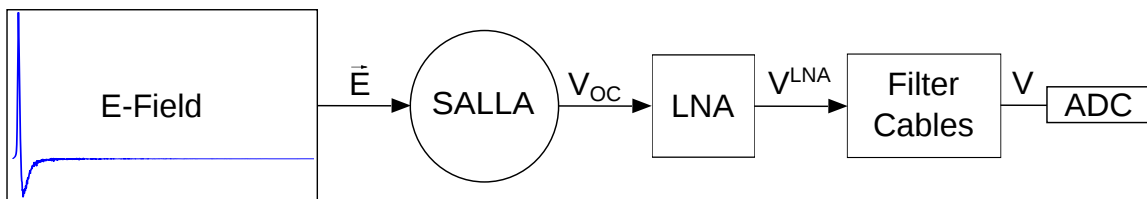


**Figure 3:** Scheme depicting the instrument response function (IRF). The incoming radio signal is received by the antenna, passes through electronics and cables, and is recorded by the ADC. For details see Ref. [3].

antenna arcs are transmitted via 30 m coaxial cables to an analog filter-amplifier, which cuts the frequency band to 30-80 MHz. The filtered signal is then digitized by the local data acquisition system (DAQ) with a 12 bit-sampling at a rate of 200 MHz; the data are collected in traces of 1024 samples each. Each element of this signal chain has been calibrated under laboratory conditions, which resulted in the instrument response function (IRF) defining the resulting digital traces recorded by the DAQ (see Fig. 3) For the reconstruction of the original signal, the inverse IRF is convoluted with the raw data. This convolution defines the *data layers (DL)* explained below.

The distinguishing feature of the broadband radio detectors is that they can be used both for radio astronomy and astroparticle purposes (e.g. ultra-high energy neutrino and cosmic-ray detection) depending on the configuration and operation mode. For example, the core of the LOFAR antenna array has been successfully applied for cosmic-ray detection [7]; meanwhile the proposed air-shower array GRAND aims also at astronomy goals [8].

## 3. Implementation

As for the prototype, the production version of TRVO[6] is deployed at IAP KIT. We have selected PostgreSQL[7] for backend storage and implemented communication with database via

---

[6]https://gitlab.iap.kit.edu/tunkarex/trvo
[7]https://www.postgresql.org

SQLAlchemy[8]. All events are stored in the single table, with one row per radio trace (i.e. both channels from single antenna are stored in two rows). We have developed interface and basic functions for the data selection and analysis, e.g. in the tutorial notebook at the JupyterLab of IAP KIT[9] on can find the application of autoencoder for the radio traces denoising [9].

### 3.1 Structure of the data

In this section we provide a general description of the Tunka-Rex data types, their structure, and their connection with the hardware of the experiment and observed phenomena. As described above, raw Tunka-Rex data consist of traces recorded for each antenna from the DAQ buffer after receiving an external trigger. Below the description of the main table with radio traces:

- `uuid`: unique identifier of the entry. We decided to use UUID [10] to have an opportunity for combination of data from different detectors without ID overlapping.

- `trigger`: trigger configuration of the array, in case of Tunka-Rex it is either Tunka-133 or Tunka-Grande.

- `run_id`: identifier of run, in case of Tunka-Rex we use date in format YYYYMMDD, i.e. unique run is identified by combination of `trigger` and run_id.

- `event_id`: identifier of the event within single run starting from zero (non-unique identifier)

- `station`: identifier of the antenna station, enumerated with the following convention: 1-25 (1st generation), 31-49 (2nd generation), 61-79 (3rd generation)

- `timestamp`: float number of the GPS time of the event with nanosecond precision

- `traces`: serialized arrays (two channels or three electric-field components) each with at least 1024 elements (non-upsampled Tunka-Rex traces)

As will be described below, `DL0-2` differs only in the way of the representation of the `traces` field.

### 3.2 Calibration and supplementary data

Antenna[10] and electronics[11] calibration data and software are published in Tunka-Rex software repository. We plan integrate this software with `TRVO` in the nearest future. The supplementary data (e.g. observation conditions provided by TAIGA meteo) can be extracted via GRADLCI services [11].

### 3.3 Data layers

In this section we describe the naming conventions for the data layers in the `TRVO` `DL0-2` are organized in the standard structure described above: Station + Calibration + Supplementary data, while the `DL3+` can have additional entries, e.g. cosmic-ray events, radio bursts, etc.

---

[8]https://www.sqlalchemy.org/

[9]tutorials/jbr/trvo/trvo_icrc.ipynb at https://jupyter.iap.kit.edu

[10]https://gitlab.iap.kit.edu/tunkarex/antenna

[11]https://gitlab.iap.kit.edu/tunkarex/calibration

**Data Layer 0** consists of raw traces recorded by the ADCs, i.e. arrays containing values in the range [0;4095]. These data are intended to be used in case of recalibration/debugging of the instrument and are not recommended for the external application.

**Data Layer 1** consists of the traces containing voltages at the antenna stations (i.e. antenna-induced voltages) obtained after unfolding the raw traces from the hardware response of Tunka-Rex amplifiers, filters, and cables. From these values the electrical field at the antenna station can be reconstructed using the specific antenna pattern and direction of incoming radio wave.

**Data Layer 2** consists of the traces containing voltages converted to the values of electrical field at the antenna stations. Depending on the data release, the electrical fields will be calculated for air-shower events (`DL2-AIRSHOWER`), for astronomical objects (`DL2-ASTRONOMY`), or for any other kind of measurements, e.g. background, RFI, etc (`DL2-OTHER`).

**Data Layer 3+** contains high-level reconstruction of radio data, i.e. quantities obtained after sophisticated processing and analyzing of radio traces. These data can be represented in tables, histograms, FITS files, etc.

In the current version of `TRVO` we provide only DL1, since it does not require recalibration of the raw data on the one hand, and allows for flexible application of antenna pattern on the other one.

## 4. Application of the Tunka-Rex Virtual Observatory

Since the primary goal of Tunka-Rex is the detection of cosmic rays, the access to the high-level reconstruction of air showers (DL3+) has already been developed in the frame of KCDC/GRADLCI and we do not depart from this concept significantly. It is worth nothing that DL3 radio data can be used for cross-calibration of different cosmic-ray experiments, as shown in Ref. [12]. Below we discuss unique features of the Tunka-Rex archival data and their application to current and future research (one should remember, that the Tunka-Rex trigger is tuned for cosmic-ray detection and the selection from the archival data might be significantly biased and can be used only for tentative studies and prototyping).

**Studies of the radio background in the frequency band of 30-80 MHz.** Nowadays there are only few radio telescopes operating in this frequency band, moreover these telescopes operate in an interferometric mode. They correlate the radio signal using beam-forming and record the resulting correlation, while radio arrays aimed at cosmic-ray detection record full uncorrelated time series. The broadband measurement of radio background in this frequency band is of special interest to search for a possible cosmological signal from neutral hydrogen. Since this signal has a signal-to-noise ratio (SNR) of about $10^{-5}$, understanding of systematic uncertainties is crucial for this type of measurements. The Tunka-Rex child experiment, Tunka-21cm [13], was designed to test the possibility of application of cosmic-ray detectors for studies of this cosmological signal, and is a first user of `DL2-BACKGROUND` and `DL2-ASTRONOMY`. `TRVO` was also used for the development of self-triggering techniques for the future radio arrays [14].

**Searching for radio transients.** Obviously archival data can be used for searching for astronomical transients in this frequency band. The effective exposure of Tunka-Rex provides only a very small probability of detection of any kind of transients. However, the archival data can be used for the test of detection techniques for future multi-purpose detectors.

6

**Training of neural networks for RFI tagging.** It was shown, that deep learning can improve the signal reconstruction of radio detectors when using an autoencoder architecture [15–17], because neural networks are able to learn features of the background and can be used for either denoising of radio traces or tagging of traces containing special features. It is worth noting, that the present Tunka-Rex autoencoder is trained on a dataset containing less than 1% of all Tunka-Rex background traces, what promises significant improvements by using larger training samples extracted from TRVO.

**Outreach and education.** Open data implies outreach and educational activities, and we support this activities. TRVO will be used as educational platform in the outreach part of the KCDC and GRADLCI projects, as well as in others, e.g. in the first worksop of Mathematical center in Akademgorodok[12]. Last but not least, the developed framework can be applied to future arrays, e.g. Tien-Shan cosmic-ray setup [18] and GCOS [19].

## 5. Conclusion

The Tunka-Rex Virtual Observatory provides open access to the data of experiments measuring cosmic rays with radio technique. We plan to combine both astroparticle- and astronomy-related features in TRVO and provide fast and user-friendly access with the possibility of custom scripting for complex preselection and preprocessing of the data.

## Acknowledgements

## References

[1] A. Haungs et al., *The KASCADE Cosmic-ray Data Centre KCDC: Granting Open Access to Astroparticle Physics Research Data*, *Eur. Phys. J.* **C78** (2018) 741 [1806.05493].

[2] Schröder, Frank G., *Radio detection of Cosmic-Ray Air Showers and High-Energy Neutrinos*, *Prog. Part. Nucl. Phys.* **93** (2017) 1 [1607.08781].

[3] Tunka-Rex collaboration, *Measurement of cosmic-ray air showers with the Tunka Radio Extension (Tunka-Rex)*, *Nucl. Instrum. Meth.* **A802** (2015) 89 [1509.08624].

[4] D. Kostunin et al., *Tunka Advanced Instrument for cosmic rays and Gamma Astronomy*, in *18th International Baikal Summer School on Physics of Elementary Particles and Astrophysics: Exploring the Universe through multiple messengers (ISAPP-Baikal 2018) Bolshie Koty, Lake Baikal, Russia, July 12-21, 2018*, 2019 [1903.07460].

[5] P. Bezyazeekov et al., *Towards the Tunka-Rex Virtual Observatory*, in *3rd International Workshop on Data Life Cycle in Physics*, 6, 2019 [1906.10425].

---

[12]https://english.nsu.ru/mca/media/news/2973965/

[6]  PIERRE AUGER collaboration, *Antennas for the Detection of Radio Emission Pulses from Cosmic-Ray*, *JINST* **7** (2012) P10011 [1209.3840].

[7]  P. Schellart et al., *Detecting cosmic rays with the LOFAR radio telescope*, *Astron. Astrophys.* **560** (2013) A98 [1311.1399].

[8]  GRAND collaboration, *The Giant Radio Array for Neutrino Detection (GRAND): Science and Design*, 1810.09994.

[9]  TUNKA-REX collaboration, *Reconstruction of sub-threshold events of cosmic-ray radio detectors using an autoencoder*, *PoS* **ICRC2021** (2021) 223.

[10] P. Leach, M. Mealling and R. Salz, *A Universally Unique IDentifier (UUID) URN Namespace*, RFC 4122, RFC Editor (July, 2005).

[11] A. Haungs et al., *Status and Future Prospects of the KASCADE Cosmic-ray Data Centre KCDC*, *PoS* **ICRC2021** (2021) 422.

[12] TUNKA-REX, LOPES collaboration, *A comparison of the cosmic-ray energy scales of Tunka-133 and KASCADE-Grande via their radio extensions Tunka-Rex and LOPES*, *Phys. Lett.* **B763** (2016) 179.

[13] D. Kostunin et al., *Quest for detection of a cosmological signal from neutral hydrogen with a digital radio array developed for air-shower measurements*, *PoS* **ICRC2019** (2020) 320 [1908.06975].

[14] P. Bezyazeekov et al., *Efficiency estimation of self-triggered antenna clusters for air-shower detection*, *PoS* **ICRC2021** (2021) 273.

[15] TUNKA-REX collaboration, *Signal recognition and background suppression by matched filters and neural networks for Tunka-Rex*, ARENA 2018 proceedings (2018) [1812.03347].

[16] M. Erdmann, F. Schlüter and R. Smida, *Classification and Recovery of Radio Signals from Cosmic Ray Induced Air Showers with Deep Learning*, *JINST* **14** (2019) P04005 [1901.04079].

[17] A. Rehman et al., *Classification and Denoising of Cosmic-Ray Radio Signals using Deep Learning*, *PoS* **ICRC2021** (2021) 417.

[18] S. Shinbulatov et al., *High-mountain hybrid installation for multicomponent detection of air-showers induced by ultra-high energy cosmic rays*, *PoS* **ICRC2021** (2021) 264.

[19] J. Hörandel et al., *The Global Cosmic Ray Observatory - GCOS*, *PoS* **ICRC2021** (2021) 027.

## Full Authors List: Tunka-Rex Collaboration

P. Bezyazeekov[1], N. Budnev[1], O. Fedorov[1], O. Gress[1], O. Grishin[1], A. Haungs[2], T. Huege[2,3], Y. Kazarina[1], M. Kleifges[4], E. Korosteleva[5], D. Kostunin[6], L. Kuzmichev[5], V. Lenok[2], N. Lubsandorzhiev[5], S. Malakhov[1], T. Marshalkina[1], R. Monkhoev[1], E. Osipova[5], A. Pakhorukov[1], L. Pankov[1], V. Prosin[5], F. G. Schröder[2,7] D. Shipilov[8] and A. Zagorodnikov[1]

[1]Applied Physics Institute ISU, Irkutsk, 664020 Russia
[2]Karlsruhe Institute of Technology, Institute for Astroparticle Physics, D-76021 Karlsruhe, Germany
[3]Astrophysical Institute, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium
[4]Institut für Prozessdatenverarbeitung und Elektronik, Karlsruhe Institute of Technology (KIT), Karlsruhe, 76021 Germany
[5]Skobeltsyn Institute of Nuclear Physics MSU, Moscow, 119991 Russia
[6]DESY, Zeuthen, 15738 Germany
[7]Bartol Research Institute, Department of Physics and Astronomy, University of Delaware, Newark, DE, 19716, USA
[8]X5 Retail Group, Moscow, 119049 Russia