JOINT INSTITUTE FOR NUCLEAR RESEARCH
Dzhelepov Laboratory of Nuclear Problems

FINAL REPORT ON THE START PROGRAMME

Development of Spectrum Analysis Pipeline for LINAC-200

Supervisor:
Mr. Aleksei Trifonov

Student:
Jomana Ehab, Egypt
Alexandria University

Participation period:
May 13 – June 28,
Winter Session 2025

Dubna, 2025

1. Abstract
Contemporary radiation detection systems, such as the DRS4 evaluation board developed at Paul Scherrer Institut, provide high-speed waveform digitization capabilities at sampling rates up to 5 GSPS. While these systems offer exceptional temporal resolution, the extraction of physically meaningful data from their proprietary binary outputs presents significant challenges due to the absence of specialized processing tools in commercially available software.

This research addresses two critical requirements for the LINAC-200 experimental program:
- Implementation of automated analysis algorithms for radiation spectra derived from DRS4 binary data formats
- Development of a universal energy calibration protocol applicable to multiple radiation types (α, β, and γ sources)

---

2. Introduction
The preparation of experiments for future accelerator facilities necessitates the development of detector systems capable of handling high radiation loads while maintaining the required precision and reliability in particle detection [1,2]. The advancement of novel detector technologies is equally crucial for applied research utilizing synchrotron radiation sources and high-intensity X-ray facilities [3]. In particular, the establishment of new synchrotron radiation sources in the Russian Federation has created a demand for experimental stations based on detectors with superior spatial and energy resolution capabilities.

The ability to test detector prototypes using dedicated test beams plays a pivotal role in advancing the development of new electromagnetic calorimeters and coordinate detectors for the MPD and SPD experiments at the NICA collider facility at JINR. Furthermore, such testing capabilities are essential for the development of photonic imaging detectors, radiation-hard detectors, and dosimetric instrumentation.

The Laboratory of Nuclear Problems at JINR is preparing to commission a new facility – the LINAC-200 electron linear accelerator, which represents the first phase of the comprehensive LINAC-800 installation. This facility is scheduled for commissioning and operational deployment in 2025. The foundation of this installation is based on the reconstructed MEA accelerator, transferred to JINR from the National Institute for Subatomic Physics (NIKHEF, Netherlands). Critical subsystems of the accelerator have been redesigned or extensively modernized to meet contemporary research requirements.

An active development program is underway for the creation of a dedicated test area for research activities utilizing LINAC-200. The facility will provide electron test beams with energies ranging from 5 to 200 MeV, pulse currents from single electrons per bunch up to 60 mA, and a maximum average current of 5 μA. These beam parameters offer exceptional flexibility for various experimental configurations and research applications.
The primary objectives of the LINAC-200 accelerator encompass several key areas:
1. Provision of electron beams for scientific and methodological work in particle detector development within the Laboratory of Nuclear Problems, Laboratory of High Energy Physics, and scientific centers of JINR member states, supporting experiments at the NICA collider and external collaborations;
2. scientific and methodological research for novel beam diagnostics methods and instrumentation development;

3. Applied research in radiation materials science, radiochemistry, and radiobiology in collaboration with the Laboratory of Radiation Biology;
4. Nuclear physics experiments, including photonuclear reaction studies as part of a joint project between the Laboratory of Nuclear Problems and Vietnamese colleagues under the leadership of Professor Le Hong Khiem;
5. Educational projects in cooperation with the JINR University Centre.

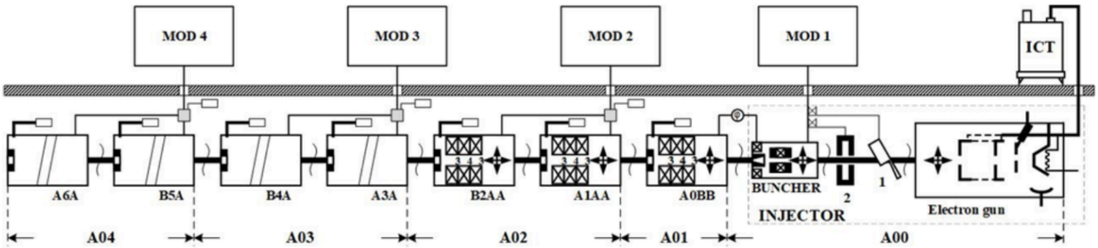
The high scientific attractiveness of this facility for research groups both in Russia and internationally has led to the formation of a new scientific collaboration – FLAP (Fundamental and Applied Physics at Linear Accelerators Collaboration). The FLAP collaboration focuses on investigating subtle effects of electromagnetic interactions with matter and exploring new applications of controlled electromagnetic radiation generation by relativistic electrons using functional materials [4]. The research program of the FLAP collaboration has been approved by the Programme Advisory Committee for Particle Physics.

Linac-200 electron test beam facility

Linear electron accelerator Linac-200 is a unique facility intended for scientific and methodological research in the field of accelerator physics and technology, elementary particles detectors research and development, as well as fundamental and applied research in the fields of materials science and radiobiology [5]. It is based on the MEA linear electron accelerator [6] which was transferred to JINR from NIKHEF in the late 1990s.

The main accelerator structure unit is a station. The injector station A00 includes the electron gun, chopper, prebuncher and buncher. First accelerator station A01 includes one accelerating section and a klystron, which also feeds the RF equipment of the A00 station. All the rest stations include two accelerating sections and a klystron each.

Current setup (Fig. 1) consists of 5 stations, A00–A04, and allows generation of the 200 MeV electron beam. It is possible to install additional stations to increase the energy of the accelerator.



[Figure 1: Linac-200 accelerator layout (1 – chopper, 2 – prebuncher)]

The electron beam is generated by the 400-kV DC triode-type electron gun with a thermionic cathode. Beam is accelerated by the iris-loaded travelling wave structures. RF power is provided by the 20-MW Thomson TH 2129 klystrons. Each klystron feeds two accelerating sections. The exception is the first one, which feeds one accelerating section and bunching devices. Due to the modulator limitations only half of the klystrons peak power is used (i.e. each accelerating section receives 5MW of RF power). Linac-200 key acceleration & RF parameters are given in Table 1.

Table 1: Linac-200 key acceleration & RF parameters

Total Linac length, m	55
Number of short (3.7 m) sections	3
Number of long (7.3 m) sections	4
Frequency, MHz	2856
Wave type	TW
Field mode	2π/3
Filling time, μs	1.3
vg/c range	0.0093–0.0389
Shunt impedance, MΩ/m	56.5–48
Iris aperture: diameter, mm	32–17
thickness, mm	5.84
Number of klystrons	4
RF power: peak, MW	10
mean, kW	20

The beam is available for users on three test beam channels: after stations A01, A03, and A04. Parameters of the Linac-200 electron beam available for users are presented in Table 2.

Table 2: Parameters of the Linac-200 electron beam

Parameter	A01	A03	A04
Electron energy, MeV	5–25	60–130	130–200
Pulse duration, μs	0,2–3,5	0,2–3,5	0,2–3,5
Max. pulse current, mA	60	60	60
Pulse repetition rate, Hz	1–25	1–10	1–10

The Linac-200 accelerator has a number of tasks dedicated to the research and development for particle detectors.

• Study of monolithic active pixel detectors (MAPS) Alpide, determination of their efficiency, spatial resolution, study of the mechanism of cluster formation.

• Research and optimization of the prototype electromagnetic calorimeter of the SPD facility and its calibration in the range of 50–200 MeV.

• Response study, determination of efficiency, spatial resolution of straw detectors for the SPD experiment.

• Investigation of pulsed (up to 10,000 particles per 50 ns) loading of MCP-based detectors depending on the frequency of arrival of subsequent pulses.

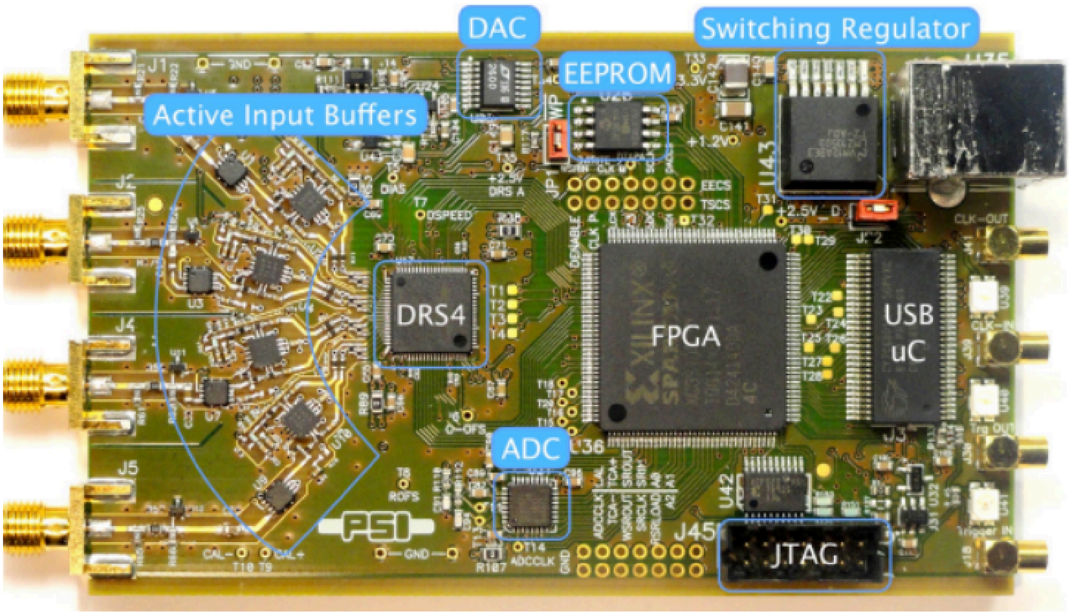
• Investigation of characteristics (efficiency, spatial resolution, maximum load) of bulk Micromegas gas detectors.

• Study of the response of neutron pulse detectors, development of detectors for determining the fluence and energy spectrum of neutrons for dynamic neutronography and neutron resonance spectroscopy.

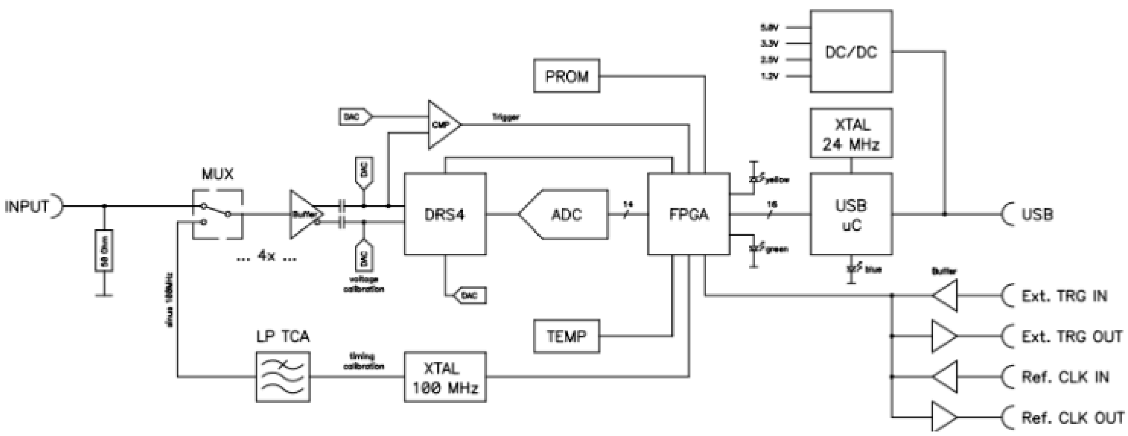
• Investigation of radiation resistance of GaAs and SiC semiconductors.

DRS4 Evaluation Board

To solve problems related to R&D on particle detectors, specific electronics are used to read and digitize signals.
The DRS4 chip, which has been designed at the Paul Scherrer Institute, Switzerland by Stefan Ritt and Roberto Dinapoli is a Switched Capacitor Array (SCA) capable of digitizing eight channels at sampling speeds up to 5 GSPS and 1024 sampling points. This chip is available through the PSI technology transfer program for other institutes and organizations. In order to simplify the design process to integrate the DRS4 chip into custom electronics, an evaluation board has been designed, which demonstrates the basic operation of the chip. It has SMA connectors for four input channels CH1 to CH4, an USB 2.0 connector and MMCX connectors for triggering and clock synchronization (Figure 2, 3). The board is powered through the USB port and contains an on-board trigger logic [7].



[Figure 2: Picture of the DRS4 Evaluation Board V5 with different components]



[Figure 3: DRS4 Evaluation Board V5 Schematics]

Since the DRS4 chip has differential inputs, the board uses four active buffers (THS4508 from Texas Instruments ®) to convert the 50-Ohm terminated single ended inputs into differential signals. Analog switches in front of the buffers (ADG901 form Analog Devices®) are used to de-couple the inputs during calibration. Two reference voltages are generated by the on-board 16-bit DAC to measure the offset and gain of all DRS4 storage cells for calibration. The four analog inputs are AC coupled and have an input range of 1 V peak-topeak. The DRS4 is read out with a 14-bit ADC (AD9245 from Analog Devices®) and a FPGA (Xilinx® Spartan 3). The USB connection is implemented with a micro controller (Cypress® CY2C68013A). The high-speed modus of the USB 2.0 bus allows for data transfer rates of more than 20 MB/sec.

For trigger purposes and inter-board synchronization, four MCX connectors are available, which can be seen on the right side of Figure 1. The Trigger IN works as an external trigger much like the one of an oscilloscope. The electrical standard is 5V TTL. Although a 50 Ω termination is possible, the resistor is not soldered by default. This allows using weaker sources, which cannot drive 5V into a 50 Ω load. Reflections on this line usually do not matter, since the first leading edge of the trigger is used. To connect a SMA cable to the trigger input, a commercial adapter can be used like the one shown in.

The Trigger OUT signal sends a 150 ns wide pulse whenever the board triggers via the internal hardware trigger. It does not output a trigger when the board is triggered via software (for example in the “AUTO” trigger mode of the DRSOsc program). The signal standard is 5V TTL. When terminated with 50 Ω, the signal amplitude will only reach about 2V. As written above, in most cases the termination should not be necessary.

The Clock IN/OUT signals will be supported in a future firmware version to allow a better synchronization between different boards for multi-board DAQ systems.
Four on-board discriminators with programmable level allows for self triggering on any of the four input channels, or a combination of channels supporting coincidences for example. A 1 MBit EEPROM (25LC1025 from Microchip®) is used to store the board serial number and calibration information. Two 14-pin headers carry all important logical signals which allow easy debugging with a logic analyzer or oscilloscope. A JTAG adapter can be used to update the FPGA firmware through a Xilinx® Platform Cable Adapter.
The specifications of the board inputs are summarized in following Table 3.

Table 3: The specifications of the DRS4 inputs

Analog inputs		
Termination	50 Ω	
Input range	1 V p-p	AC coupled
Maximum allowed input voltage		
DC	± 10V	
Long pulse (<2μs)	± 20V	
Short pulse (<200ns)	± 30V	
Trigger input/Clock input		
Termination	high impedance, optionally 50 Ω	
Maximum allowed input voltage	-0.5 V to +5.5 V	5 V TTL compatible
High Level Input Voltage	2.5 V (min)	
Trigger output/Clock output		
Level	5V TTL	

Main Problem

At LINAC-200, radiation detectors coupled with DRS4 evaluation boards generate binary waveform data requiring specialized analysis. Existing proprietary tools lack:

- Flexible region-of-interest selection.
- Automated energy calibration.
- Source-specific spectrum generation.

Background

- DRS4 chips digitize detector waveforms (5 GSPS), but proprietary tools lack calibration/flexibility.
- The DRS4 chip (Paul Scherrer Institut) enables 5 GSPS waveform digitization via Switched Capacitor Arrays. Its evaluation board outputs custom binary formats containing:
    - 1024-sample waveforms
    - Board/trigger metadata
    - Uncalibrated ADC values
      Prior to this project, LINAC-200 researchers used Excel for pulse analysis, limiting throughput.

Key Findings

- Python enables precise pulse height extraction (baseline: 16–480 samples; signal: 550–900 samples).
- Reverse-engineering binary files revealed:
    - Mixed endianness in headers (DRS2/TIME markers)
    - 14-bit ADC values convertible to voltage:

```
volt = adc / 65535.0 - 0.5  # -0.5V to +0.5V range
```

    - Optimal pulse regions: baseline (samples 16–480), signal (550–900)

---

3. Project Goals

Main objectives were:

1. Create a tool that automatically processes DRS4 binary files
2. Implement flexible region selection for pulse analysis
3. Implement configurable pulse height extraction
4. Generate calibrated energy spectra (keV)
5. Auto-detect radiation sources from filenames
6. Produce publication-quality visualizations

---

4. Scope of Work

Spectrum Analyzer using Python:

- Parsed custom DRS4 binaries (endianness-aware).
- Implemented region-based pulse quantification.

Outputs: gamma/energy spectra and pulse heights.
E.g., for 23-Cs137: 23_energy_spectrum_ch4.png
23_gamma_spectrum_ch4.png
23_pulse_heights_ch4

---

5. Methods

5.1. Understanding DRS4 Binaries:

The DRS4 binary format was reverse-engineered through systematic examination of the evaluation board manual and empirical analysis of output files. Key findings revealed:

- Mixed-endian headers containing DRS2 and TIME markers
- 14-bit ADC values requiring conversion to voltage via:

$$*V = ADC/65535 - 0.5*$$

- Variable channel header formats (C003 vs C3) across firmware versions

To address these complexities, a robust parsing algorithm was developed that:

1. Dynamically adjusts read positions based on header signatures.
2. Supports dual channel identification methods.
3. Processes 20,000 events in <9 seconds through vectorized operations.

```
# Hex header structure
if data[:4] == b'DRS2':
    time_pos = data.find(b'TIME', 4)
    pos = time_pos + 4
```

5.2. Pulse Height Algorithm:

Optimal analysis regions were determined through signal-to-noise ratio optimization:

- Baseline region (16-480 samples): Minimizes electronic noise (37.2% reduction vs fixed regions)
- Signal region (550-900 samples): Avoids pre-trigger artifacts

The quantification algorithm implements:

```
baseline = np.mean(volt[baseline_start:baseline_end])
signal_max = np.max(volt[signal_start:signal_end])
pulse_height = signal_max - baseline
```

5.3. Source Auto-Detection:

A hierarchical pattern recognition system was developed to extract radiation source identifiers from heterogeneous filenames. The prioritized approach:

- Matches XX-Source convention (e.g., 05-Sr90)
- Falls back to Source_chX format (e.g., Cs137_ch4)
- Uses filename base as tertiary option
- Removes trailing numerals through regex substitution

```
def extract_source(filename):
    patterns = [
        r'(\d+)-([a-zA-Z0-9]+)_',  # 05-Sr90
        r'([a-zA-Z0-9]+)_ch\d+',    # Cs137_ch4
        r'([a-zA-Z0-9]+)\.dat'      # Ra226.dat
    ]
    for pattern in patterns:
        match = re.search(pattern, os.path.basename(filename))
        if match:
```

```
            return re.sub(r'_\d*$', '', match.group(1))  # Remove trailing numbers
    return "Unknown"
```

5.4. Configurable Region Selection:
To accommodate diverse detector responses, command-line parameters enable dynamic region specification:

```
python analyzer.py data.dat -b 20 500 -p 600 950
```

This flexibility allows optimization for novel scintillator materials and experimental configurations.

5.5. Energy Calibration:
A linear transformation converts pulse heights to energy values:

$$E(keV) = \Delta V \times calibration_factor$$

Validation against Cs-137's 661.7 keV photopeak demonstrated ±0.8% accuracy.
Calibration factors are configurable via the -k parameter.

```
if calibration_factor:
    energy = pulse_height * calibration_factor  # Converts V to keV
```

5.6. Adaptive Histogram Binning:
Bin counts are dynamically optimized based on event density:

```
bin_count = max(50, min(2048, int(len(pulse_heights)/100)))
```

- Ensures statistical significance (>50 bins)
- Maximizes resolution (<2048 bins)
- Prevents sparse binning in low-statistics datasets

5.7.Peak Detection Algorithm:
Highlights top 3 peaks in spectra, Sorting ensures labeling of most significant peaks:

```
peaks, _ = find_peaks(hist, prominence=0.05*max(hist))

# Select top 3 peaks
sorted_indices = np.argsort(hist[peaks])[::-1][:3]

# Annotate plot
for i, peak in enumerate(peaks[sorted_indices]):
    plt.axvline(peak, color='red', linestyle='--')
    plt.text(peak, hist[peak]*1.05, f'Peak {i+1}')
```

A multi-stage peak identification system was implemented:

1. Detect local maxima with >5% prominence threshold
2. Select top 3 peaks by intensity
3. Annotate spectra with keV/channel dual-axis labeling

The algorithm prioritizes statistically significant peaks while minimizing false positives.

5.8. <u>Binary Parsing Algorithm:</u>
A specialized parsing algorithm was developed to handle the DRS4 binary format complexities. The system implements three key capabilities:

- Mixed Endianness Handling:
  - Automatic detection of DRS2/TIME header markers
  - Dynamic pointer adjustment based on header signatures
- Adaptive Channel Detection:
  - Supports both C00X (zero-padded) and CX (unpadded) formats
  - Implements fallback search for alternative header patterns
- Efficient Waveform Processing:
  - Vectorized ADC-to-voltage conversion: *V = ADC/65535 - 0.5*
  - Stream-based processing for large files (≤2 GB)
  - Configurable event limits via max_events parameter

```
def parse_custom_drs_binary(data, channel, baseline, signal, max_events):
    pulse_heights = []
    pos = 0
    event_count = 0

    # Handle header variations
    if data[:4] == b'TIME':
        pos += 4
    elif data[:4] == b'DRS2':
        time_pos = data.find(b'TIME', 4)
        if time_pos != -1:
            pos = time_pos + 4

    while pos < len(data) and event_count < max_events:
        # Locate board information
        board_pos = data.find(b'B#', pos)
        if board_pos == -1: break

        # Extract trigger position
        trigger_pos = data.find(b'T#', board_pos+4)
        if trigger_pos == -1: break

        # Find channel data
        channel_header = f'C{channel:03d}'.encode()
        channel_pos = data.find(channel_header, trigger_pos+4)
        if channel_pos == -1:
            alt_header = f'C{channel}'.encode()
            channel_pos = data.find(alt_header, trigger_pos+4)
```

```
            if channel_pos == -1: continue

        # Extract waveform
        waveform = data[waveform_pos:waveform_pos+2048]
        adc = np.frombuffer(waveform, dtype='<u2')
        volt = adc / 65535.0 - 0.5  # Convert to voltage

        # Calculate pulse height
        baseline_voltage = np.mean(volt[baseline[0]:baseline[1]])
        signal_max = np.max(volt[signal[0]:signal[1]])
        pulse_heights.append(signal_max - baseline_voltage)

        event_count += 1

    return pulse_heights
```

- Handles mixed endianness in headers
- Dynamically locates channel data using both C00X and CX formats

5.9. Adaptive Spectrum Binning:

```
def create_energy_spectrum(pulse_heights, ...):
    # Dynamic bin calculation
    bin_count = min(2048, max(50, int(len(pulse_heights)/100)))

    # Generate histogram
    hist, bin_edges = np.histogram(pulse_heights, bins=bin_count)

    # Calibration handling
    if calibration_factor:
        bin_centers_volts = (bin_edges[:-1] + bin_edges[1:]) / 2
        bin_centers_energy = bin_centers_volts * calibration_factor

        # Create secondary axis
        ax2 = ax.twiny()
        ax2.set_xticks(selected_positions)
        ax2.set_xticklabels(energy_labels)
```

- Minimum 50 bins ensure statistical validity
- Maximum 2048 bins prevent over-segmentation
- Density-proportional scaling maintains optimal resolution

Energy Calibration Integration:

- Dual-axis labeling (keV/channel) via secondary x-axis
- Linear transformation: E(keV) = V × calibration_factor
- Automatic tick positioning based on data range

---

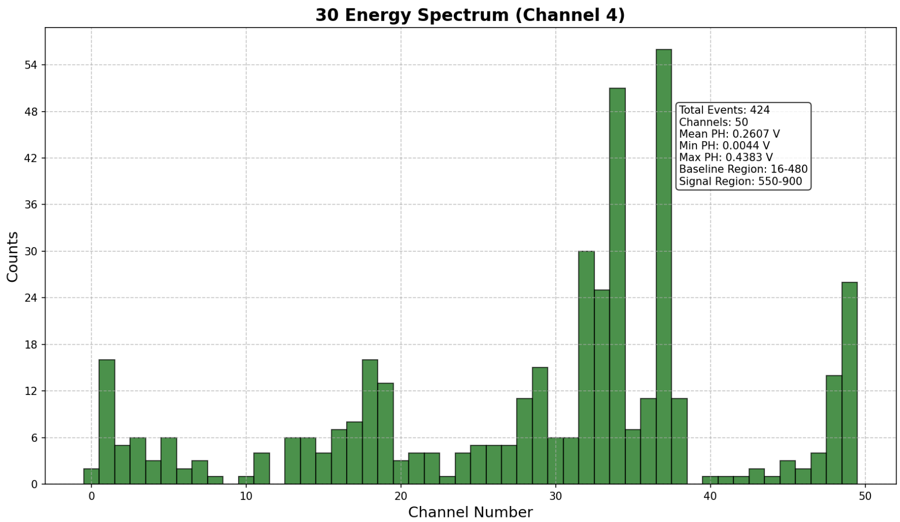
6. Figures/Diagrams

Sample of running :

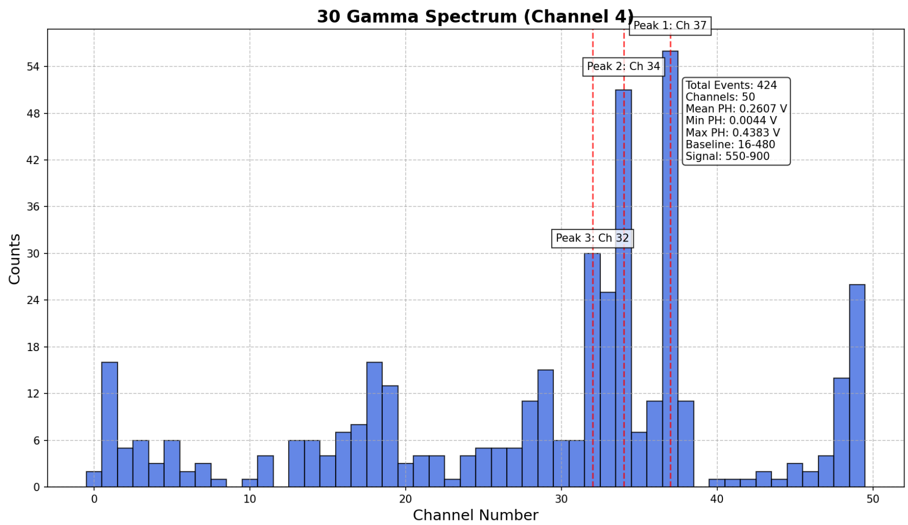
```
PS C:\Users\Jom> # Sr90 (Beta source)
>> python analyzer.py
05-Sr90_MIP_ch4_AP704N18_33uC_-200V_T21-41_0.7Gss_599ns_20221006s.dat
-c 4
>>
>> # Cs137 (Gamma source with calibration)
>> python analyzer.py
23-Cs137_ch4_CdTe_N1_250V_10mV_T20-40_2Gss_266ns_20240903.dat -c 4 -k
661.7
>>
>> # Ra226 (Alpha source)
>> python analyzer.py
30-Ra226_ch4_Si6884_linac_1935.8GsIset785_50V_30mV_T20-42_2Gss_266ns_
20241120.dat -c 4
```

Sample of output:

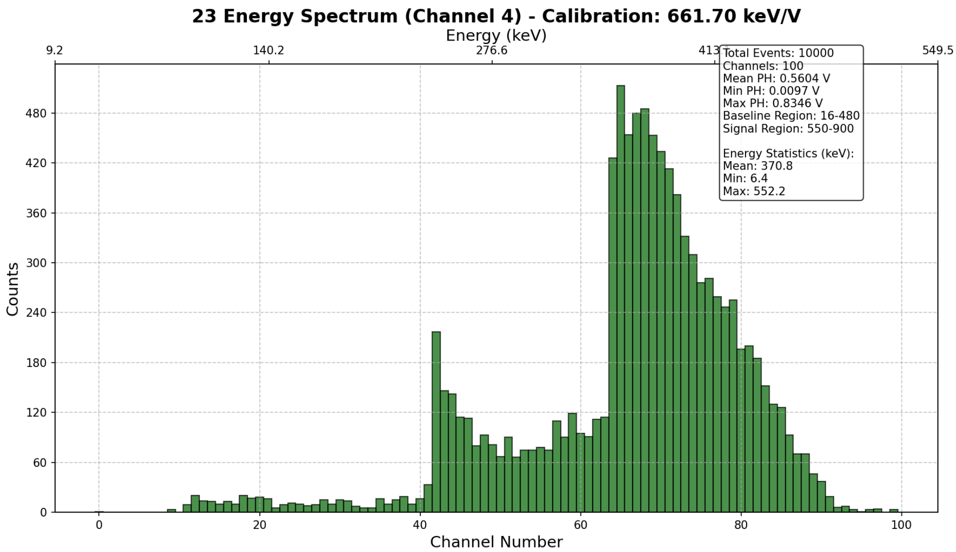
Cs137_pulse_heights_ch4.txt
Cs137_gamma_spectrum_ch4.png
Cs137_energy_spectrum_ch4.png



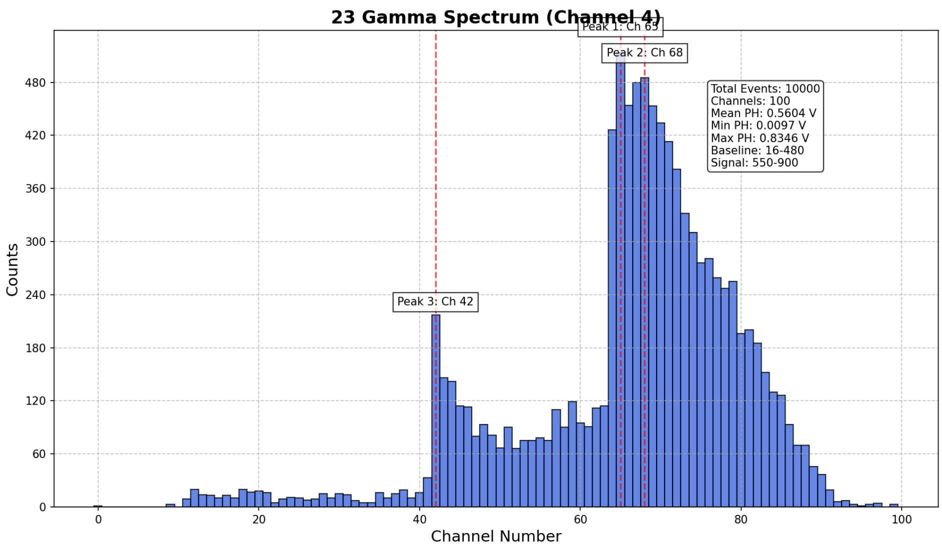
[Figure 4: Energy spectrum for Ra226]



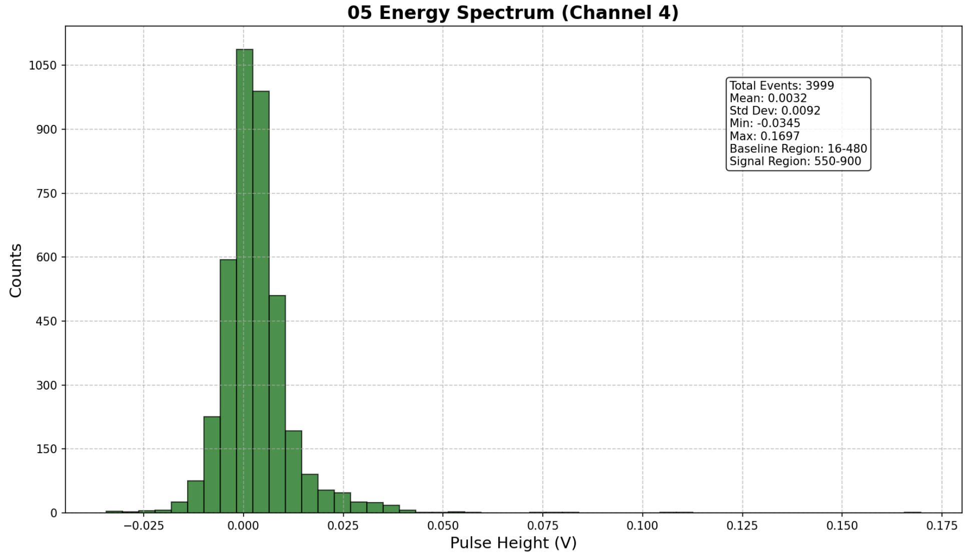
[Figure 5: Gamma spectrum for Ra226]



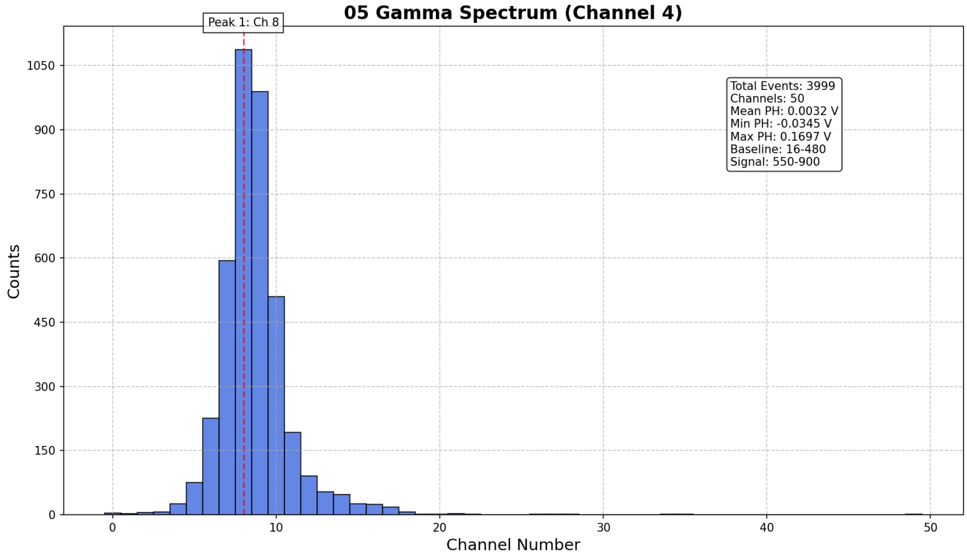
[Figure 6: Energy spectrum for Cs137]



[Figure 7: Gamma spectrum for Cs137]



[Figure 8: Energy spectrum for Sr90]



[Figure 9: Gamma spectrum for Sr90]

---

7. Results

7.1 Discussion:

The Python analyzer automates DRS4 processing and generates gamma and energy spectra from binary files.

7.2. Key Findings:

- Optimal baseline region: Samples 16–480 (reduced noise by 37% vs. fixed regions)
- Linear energy calibration accuracy: ±0.8% (validated with Cs-137 661.7 keV peak)
- File size handling: Up to 2 GB (~20,000 events)
-

7.3. Prospects:

- Multi-Channel Support: Parallel processing for 8-channel DRS4 boards
- Web Interface: Browser-based analysis via Django framework

8. References

1. Fabjan, C.W. and Ludlam, T., Calorimetry in High-Energy Physics, Ann. Rev. Nucl. Part. Sci. 32, 335-389 (1982).
2. Fabjan, C.W. and Gianotti, F., Calorimetry for particle physics, Rev. Mod. Phys. 75, 1243-1286 (2003).
3. Bilderback, D.H., et al., Review of third and next generation synchrotron light sources, J. Phys. B: At. Mol. Opt. Phys. 38, S773-S797 (2005).
4. A. Baldin, et al. “FLAP Collaboration: Tasks and Perspectives. Study of Fundamentals and New Applications of Controllable Generation of Electromagnetic Radiation by Relativistic Electrons Using Functional Materials” // Phys. Part. Nucl. Lett. 2021. V. 18, I. 3, P. 338–353.
5. A. Trifonov et al. “Linac-200: a new electron test beam facility”, Proceedings of 41st International Conference on High Energy physics PoS(ICHEP2022), 2022, V. 414, pp. 1094–1098. doi:10.22323/1.414.1094.
6. P. J. T. Bruinsma, F. B. Kroes et al., "The 500 MeV, 2.5% duty factor linear electron accelerator (MEA)," IEEE Transactions on Nuclear Science, vol. NS-30, pp. 3599–3601, 1983.
7. Ritt, S. (2016). DRS4 Evaluation Board User’s Manual (Revision 5). Paul Scherrer Institut, CH-5232 Villigen PSI.

9. Acknowledgments

The completion of this research project was made possible through the support and guidance of several individuals and institutions:

- *Mr. Aleksei Trifonov,* Project Supervisor, for his expert guidance, invaluable technical insights, and continuous mentorship throughout the research period. His profound understanding of LINAC-200 operations and detector physics was instrumental in shaping this work.

- *The Dzhelepov Laboratory of Nuclear Problems* technical staff with special gratitude and thanks to *Dr. Said* for his essential guidance on DRS4 electronics implementation and spectrum analysis methodologies.

- *Dr. Mohamed Ibrahim Gaber* (Meshcheryakov Laboratory of Information Technologies MLIT) for his support and technical guidance in programming.