Measurement of the Cosmic-Ray Muon Lifetime

Bryce Olsen
University of California, Santa Barbara
June 2026

Abstract

This research project leverages a configuration of plastic scintillator slabs to measure the lifetime of the cosmic-ray muon. By stacking slabs in a vertical fashion, I create a setup aiming to identify and study muons that stop inside of the material with the goal of studying the amount of time it takes to detect the decay electron or positron. By reducing background-like processes and fitting an exponential to this time difference, I can get a value for the muon lifetime. A weighted average across selected runs with special configurations gave

$$\tau_\mu = 2.04 \pm 0.06_{\text{stat}} \pm 0.10_{\text{sys}}\ \mu\text{s},$$

or $\tau_\mu = 2.04 \pm 0.12\ \mu$s after combining uncertainties. The result is within about 7.1% of the Particle Data Group accepted free-muon lifetime.

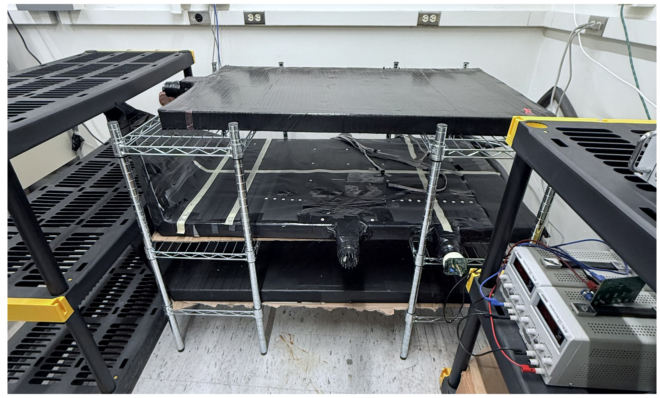
1 Introduction

Cosmic-ray muons are created when energetic particles from space interact with Earth's atmosphere. Most of them that make it to ground level pass through the apparatus, but a small fraction stop inside a scintillator slab, recording a pulse when it initially hits and a second pulse from its decay product. The goal of this project was to isolate these events and measure their characteristic decay time.

2 Setup and Identification Logic

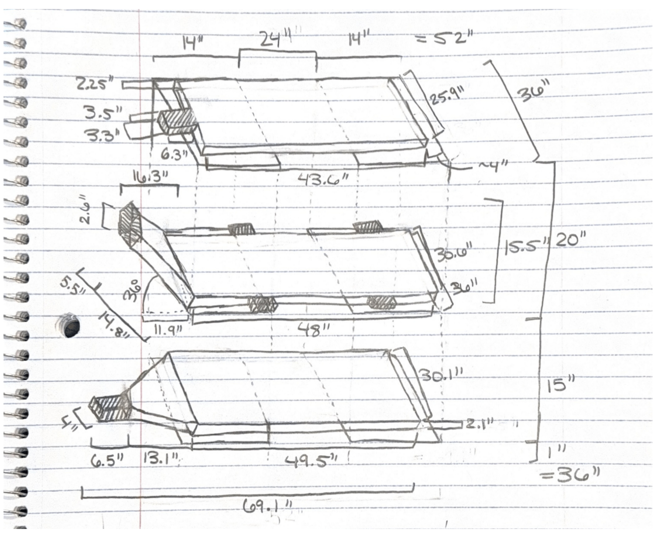
The apparatus consisted of three plastic scintillator slabs arranged vertically. Slab 1 was on top, slab 2 was in the middle, and slab 3 was on the bottom. When a charged particle passed through a slab, the scintillator emitted light, which was detected by photomultiplier tubes (PMTs) and recorded as electronic pulses. The data contained pulse times, pulse heights, pulse areas, coincidence flags, and related quantities. These can be viewed as individual event waveforms or as histograms/scatterplots of the run as a whole.

(a) Building the physical three-slab setup.

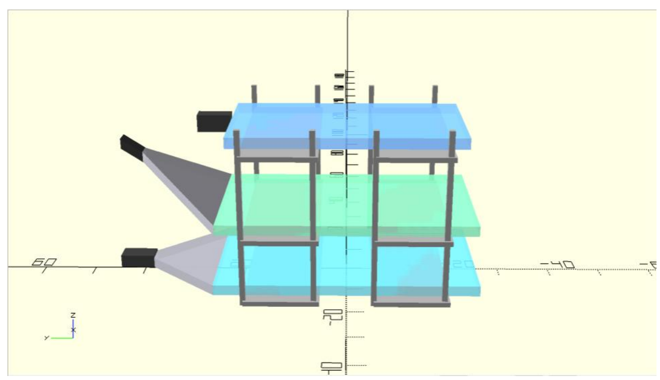


(b) Side-view model of the three-slab stack.

Figure 1: The detector geometry determines which events may represent a stopped muon.



(a) Sketch of setup with dimensions.



(b) 3D model of the slab and PMT layout.

Figure 2: Sketches and 3D models were made using physical dimensions before assembly.

In the simplified topology, the desired signal event is: Slab 1 hit + Slab 2 first pulse + Slab 2 delayed second pulse + no Slab 3 coincidence pulse. The measurement was set to trigger with any registered hit in both slabs 1 and 2, with the no-hit veto in slab 3 coming from offline filtering. The time difference between the two slab-2 pulses is the measured quantity. Naturally, the challenge is that this data set is not purely stopped muons and is dominated by background that meets these trigger conditions, which will be discussed later.

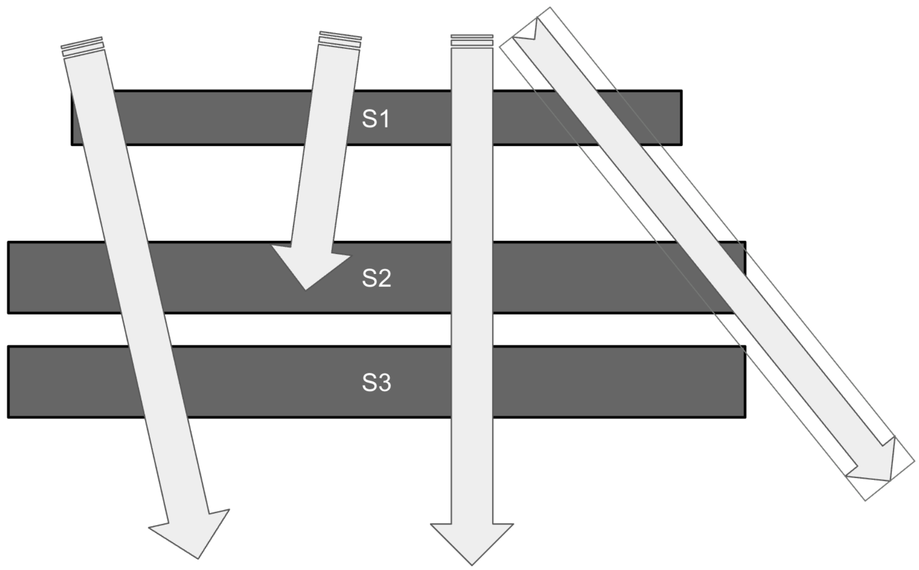


Figure 3: Examples of through-going, stopped, and geometric-miss topologies. A pulse in slabs 1 and 2 without a pulse in slab 3 is consistent with a stopped-muon candidate, but it can also be produced by a geometric miss. This ambiguity motivates the background studies and offline cuts.

3 Measurement Principle

The event class I want to study is a cosmic-ray muon that passes through slab 1, enters slab 2, stops in slab 2, decays after some time Δt, and produces a second pulse in slab 2 from the decay product. For many events, the distribution of delay times should follow an exponential decay shape,

$$N(t) = Ae^{-(t-t_{\min})/\tau} + B.$$

where τ is the lifetime. $N(t)$ describes the predicted number of counts in a histogram bin at delay time t. The lower fit edge, $t_{\min}$, cuts out early-time prompt contamination. The parameter A sets the size of the decay-like signal at the start of the fit range. The constant term B accounts for random delayed pulses that remain after cuts.

4 Timing reliability

A timing measurement is only meaningful if the recorded pulse times can be interpreted consistently. To find the signal propagation speed in the cables, I performed a cable delay experiment using a pulse generator and oscilloscope. The test compared known cable-length differences with measured delays. The analysis included Gaussian fits to timing peaks and line fits to check consistency of timing behavior. The fitted slope yielded a cable signal speed of $v = (0.640 \pm 0.006)c$. This value is not the muon speed; it is the propagation speed of electrical signals through the cable, which is expected to be less than c because the signals travel through material rather than vacuum.

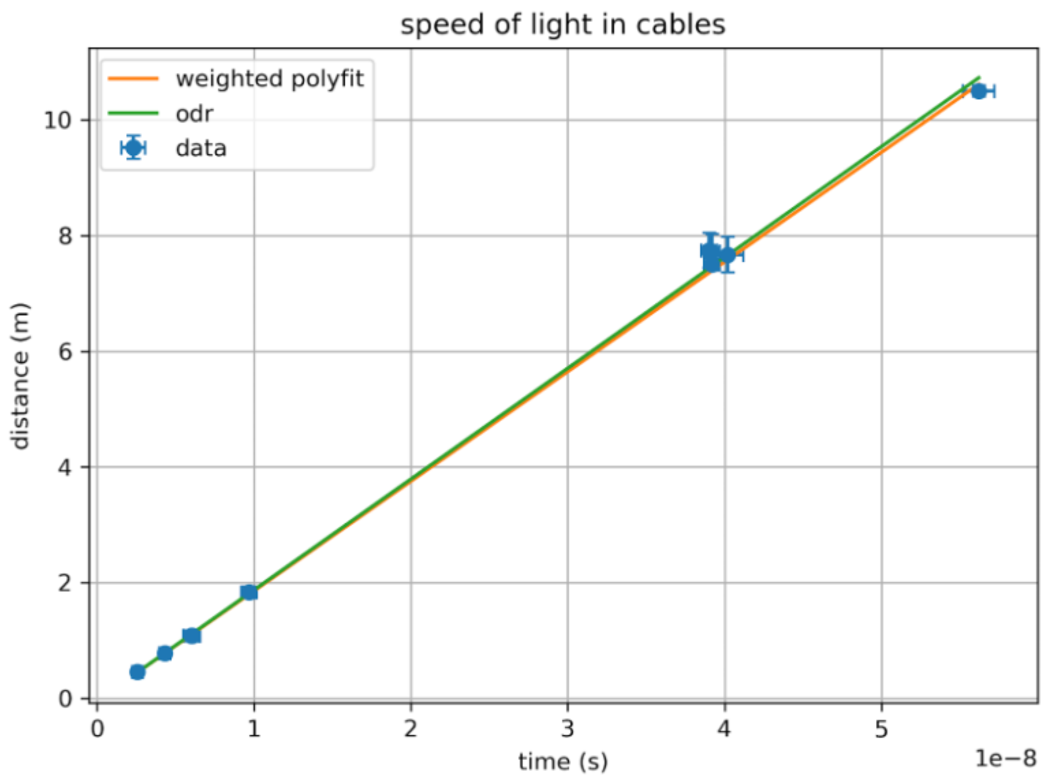


Figure 4: Cable delay study used to test the reliability of timing measurements. It also gave practice with the similar core tasks used later.

5 Background Classification

A major part of the project was identifying backgrounds that matched the trigger conditions before fitting the lifetime spectrum. The trigger and recorded pulses do not automatically separate true stopped muons from other event classes. It is important to define event classes to account for later.

The important point is that not every background was completely removed. Instead, the analysis used cuts to reduce the most obvious contamination and included a constant background term in the fit to account for residual random delayed pulses. The spread of reasonable cut choices was then used as the dominant estimate of systematic uncertainty.

Event class	Measurement / mitigation
Stopped-muon candidate Cosmic-ray muon passes through slab 1, enters slab 2, stops in slab 2, then decays; no pulse in slab 3.	Signal class. The two slab-2 pulses define $$\Delta t = t_{11} - t_{10}.$$ This class should produce the exponential distribution used to extract the muon lifetime.
Geometric miss Real cosmic-ray particle passes slabs 1 and 2 but misses slab 3 because of detector geometry; may imitate a stopped muon. Some events also include an unrelated second slab-2 pulse.	Reduced by limiting the possible solid angle for misses using smaller scintillator panels. Measured with panel configurations that isolate misses, allowing pulse-height, timing, and rate distributions to be compared with the main sample. The rate of unrelated second pulses in geometric-miss events estimates how often this background enters the final sample. Residual contribution is included in the constant background term.
Accidental coincidence Unrelated pulses overlap in time or produce a random second slab-2 pulse inside the lifetime window. Can occur from two unrelated pulses, a real/background first pulse followed by an unrelated second pulse, or an unrelated slab-2 trigger during a real muon event.	Measured by studying time separations outside the expected muon-decay region and looking for characteristic timing or waveform behavior. Treated mainly through the constant background term B in the final exponential fit.
Noise, radiation, and shower particles Electronics noise, ambient radiation, or shower particles create non-signal pulses. Some are small, while others may be large enough to pass basic cuts.	Reduced with V_{10} and V_{11} thresholds to remove pulses noticeably smaller than muon-like pulses. Radioactive-source tests were used to amplify and study this background directly, helping identify behavior that should be removed from the final sample.
Equipment or reconstruction failure Slab 3 fails to record a pulse, trigger times are shifted or unexpected, PMTs produce extra pulses, or false pulses mimic the muon or decay electron.	Mitigated by rejecting unexpected raw-data behavior. False timings are removed with expected raw-time windows, and false pulses are reduced using pulse-size and pulse-shape behavior. Since the final dataset is large, strict cuts can remove likely false positives without strongly affecting the result.

Table 1: Summary of the desired stopped-muon event class and the main background classes considered in the muon lifetime analysis.

6 Data-taking modes

Before the final lifetime fit, I studied how the detector and electronics responded under different configurations. A variety of data-taking slab configurations were used to target different classes of events. In doing so, small panels made out of the same scintillator material, labeled Jackson panels (or "Jacksons") by my lab group, were used to restrict or study the possible particle paths through the detector. Because they are smaller than the slabs, using them as event triggers restricts the accepted angular range and helps study geometric backgrounds.

During all of these modes, data are collected from any pulses in any of the 3 slabs or either of the 2 Jackson panels, even when they are not used in the trigger logic. This was useful because muons triggered by one part of the setup could still leave information in other detector elements, allowing me to compare event behavior across modes after the data were saved. The main analysis still comes from the data in slabs 2 and 3 as described by the measurement principle. I eventually calculated the lifetime from several of these modes for comparison, and used the spread of fitted lifetime values to estimate my systematic uncertainty.

Most of the final statistical sample came from the high-rate slab mode without Jackson panels. This mode produced the most data because of the larger slab area. The Jackson-panel modes were used mainly as checks: they helped test how angular acceptance, geometric misses, and smaller-area triggers affected the candidate sample, but they did not dominate the final weighted mean.

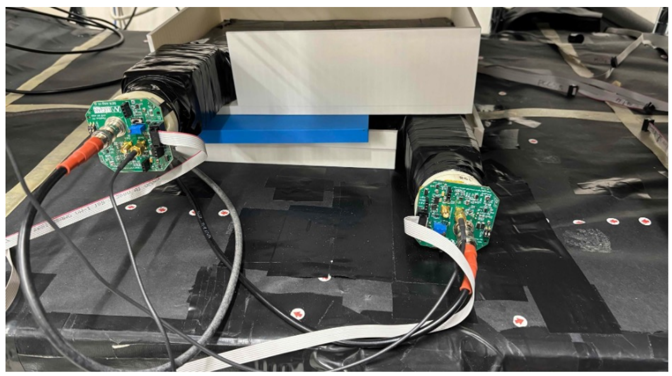
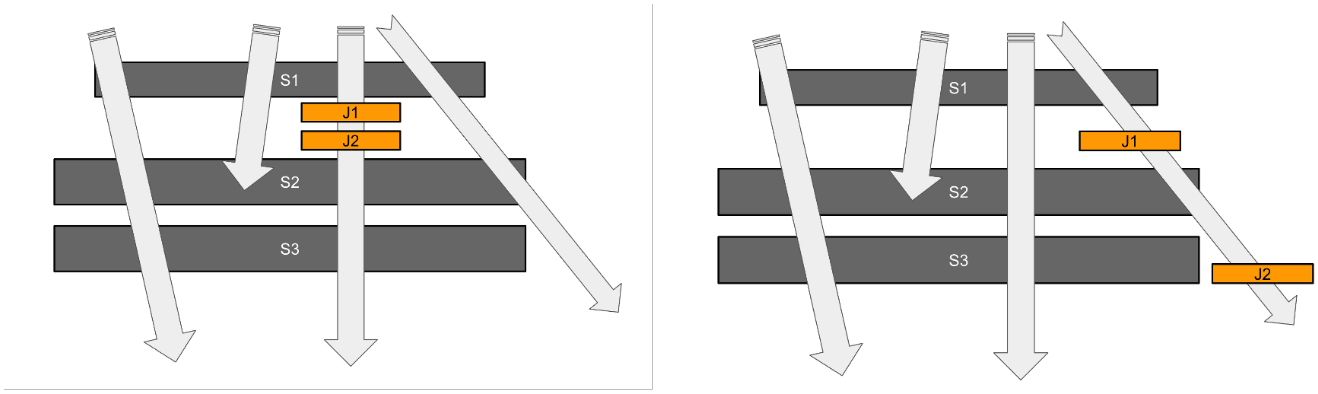


Figure 5: Jackson panels used in various positions for different modes. This image shows them stacked.

Configuration mode	**Purpose**
Trigger on slab 1 and slab 2. No Jackson used.	Main high-statistics mode. Yields a larger dataset due to the larger surface area of the slabs. Slightly less pure because there is a larger solid angle in which the muon may miss slab 3 geometrically, satisfying the candidate condition without stopping in reality.
Trigger on Jackson 1 above slab 1 and Jackson 2 between slab 1 and slab 2. Jacksons are aligned vertically.	Same basic idea as the slab mode, but with a smaller accepted solid angle. Smaller dataset due to smaller surface area; more pure due to configuration.
Trigger on Jackson 1 and Jackson 2 aligned vertically on top of each other between slab 1 and slab 2.	Compromise mode. Allows more data than the stricter Jackson geometry while only slightly increasing possible miss angles. Useful for comparing trigger rate and purity.
Trigger on Jackson 1 aligned vertically with the three slabs and Jackson 2 placed adjacent to slab 3.	Geometry-miss check. Designed to make geometric misses easier to identify and study, so this background could be compared with the main dataset.

Table 2: Summary of the main data-taking logic. The final measurement was dominated by the high-statistics no-Jackson slab mode.



(a) Jacksons are stacked to increase purity.

(b) Jacksons are aligned to catch geometric misses.

Figure 6: Data-taking modes using the Jackson panels.

7 Pulse Selection

Each run was saved as a ROOT file, the standard high-energy physics file format used to store detector-event data. In the Python analysis, I used `uproot` to open each ROOT file, read the relevant tree, and extract arrays for the pulse times, pulse heights, and coincidence flags. The same pipeline was then applied run-by-run: calculate the delay time, apply cuts, build a histogram, fit the lifetime spectrum, and compare the result across runs.

Pulse-height information was useful because charged particles passing through scintillator material should produce a broad Landau-like energy-loss distribution. The measured pulse-height distribution in Fig. 7 is consistent with this expected minimum-ionizing-particle shape, supporting that the selected population is dominated by charged-particle, muon-like hits, while other features in the histogram indicate pulses from additional sources. Minimum voltage cuts were then used to remove the low-amplitude region while keeping the broader muon-like population.

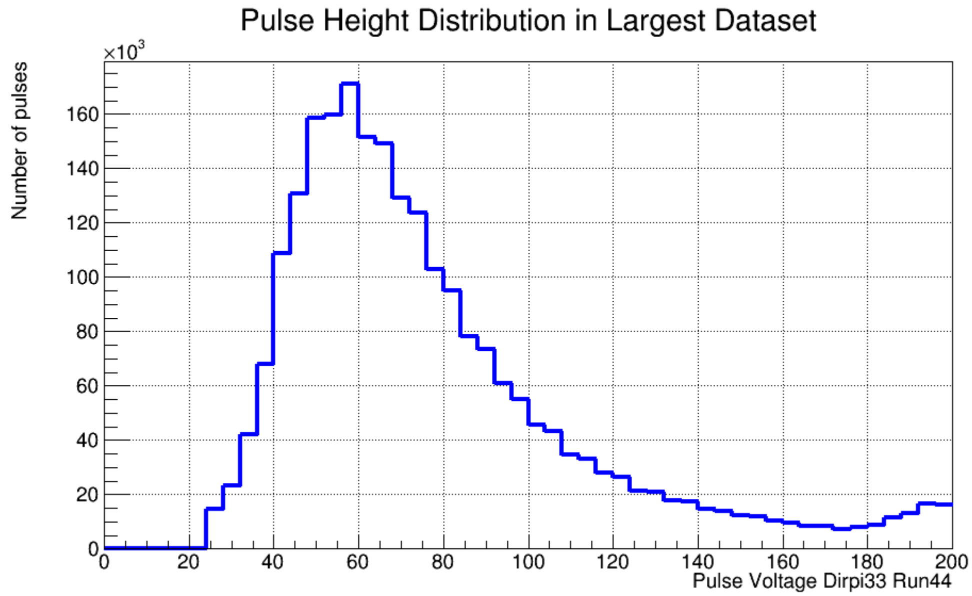


Figure 7: Measured pulse-height distribution showing the expected Landau-like shape for charged particles.

For each event, the key derived variable was $\Delta t = t_{11} - t_{10}$, where t_{10} and t_{11} are the raw times of the first and second slab-2 pulses. The variables V_{10} and V_{11} are the corresponding pulse heights. The variables `coinc10` and `coinc11` record whether the first or second slab-2 pulse was coincident with recorded slab-3 activity. These coincidence flags were used after data collection as offline cuts to reject events that looked through-going rather than stopped.

The final candidate-only selection required

$$\text{finite } t_{10}, t_{11}, V_{10}, V_{11}, \texttt{coinc10}, \texttt{coinc11}, \Delta t,$$

$$t_{10} > 0, \qquad t_{11} > 0, \qquad \Delta t = t_{11} - t_{10} > 0,$$

$$250 \leq t_{10} \leq 260 \text{ ticks},$$

$$75 \leq \Delta t \leq 500 \text{ ticks},$$

$$V_{10} \geq 30, \qquad V_{11} \geq 40,$$

$$\texttt{coinc10} = 0, \qquad \texttt{coinc11} = 0.$$

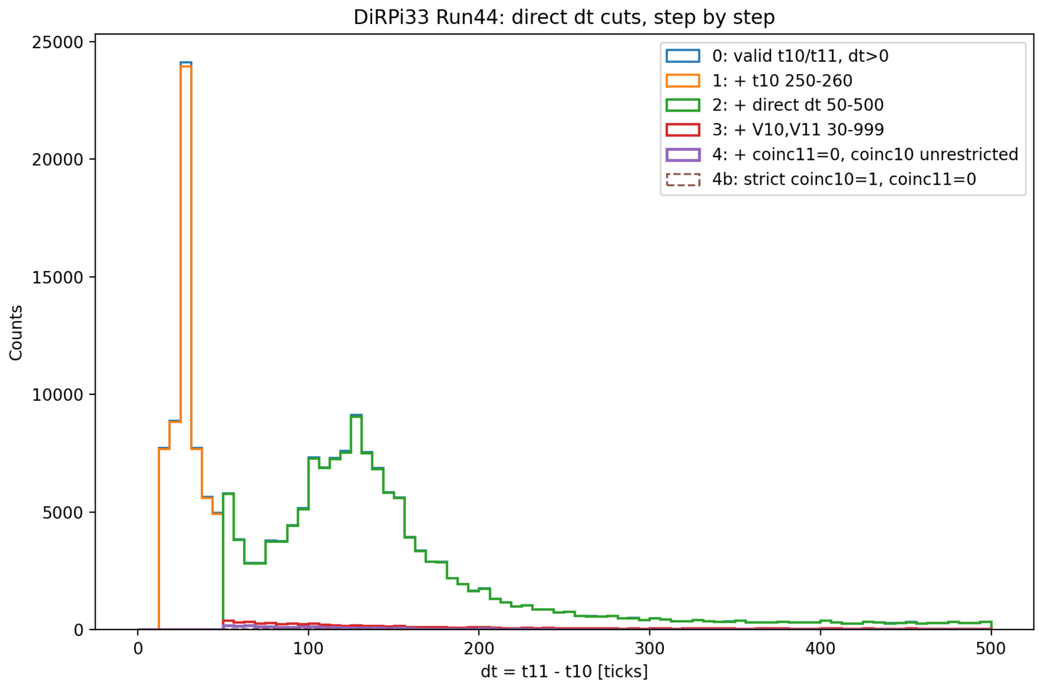


Figure 8: Example of the direct $\Delta t = t_{11} - t_{10}$ cut sequence for Run 44. The prompt and broad timing structures visible before the final cuts motivated the narrow t_{10} window, pulse-height thresholds, and offline coincidence requirements.

After the final cuts were applied to a run, the Python script histogrammed the selected Δt values and fit the distribution with

$$N(t) = Ae^{-(t-t_{\min})/\tau} + B.$$

The fit returned a value of τ and its uncertainty for that run. Since the detector timing unit was one tick $= 0.025$ μs, each fitted lifetime was converted from ticks to microseconds. Combining runs reduced the dependence on statistical fluctuations from any one dataset. The candidate-only fits from the selected runs were then combined with inverse-variance weighting,

$$w_i = \frac{1}{\sigma_i^2}, \qquad \bar{\tau} = \frac{\sum_i w_i \tau_i}{\sum_i w_i}, \qquad \sigma_{\bar{\tau}} = \sqrt{\frac{1}{\sum_i w_i}}.$$

This weighting gives more influence to runs with smaller fitted uncertainty, which generally means the higher-statistics slab-trigger runs contribute more strongly than the smaller Jackson-check runs.

8 Run Selection

For the key runs, the `dirpi33` tree corresponded to the DiRPi33 readout unit, which recorded the slab-2 and slab-3 pulse information. The final statistical measurement used selected `dirpi33` runs that passed basic checks on timing behavior, pulse selection, and fit stability. The run selection used for the final analysis consisted of larger and smaller datasets, according to the modes outlined earlier.

I performed a large scan over combinations of $\Delta t_{\min}$, $\Delta t_{\max}$, pulse-height thresholds, and binning sizes to identify settings that gave a good combination of physically reasonable selections and stable fits according to the reduced χ^2_ν value. Table 3 shows a small representative subset of that scan. The final weighted-average settings were

$$75 \leq \Delta t \leq 500 \text{ ticks}, \qquad V_{10} \geq 30, \qquad V_{11} \geq 40, \qquad 20 \text{ bins}.$$

These settings were both physically reasonable based on earlier tests identifying different background classes and gave stable fit quality across the selected runs. The lower Δt bound avoids prompt contamination and possible instrumental secondary pulses. The upper bound keeps the analysis in the region where the candidate statistics are still meaningful.

$\Delta t_{\min}$	$\Delta t_{\max}$	$V_{10,\min}$	$V_{11,\min}$	Bins	Mean χ^2_ν
75	500	30	30	20	1.036
75	500	30	40	20	0.925
75	500	30	50	20	0.990
75	500	40	40	20	0.973
80	500	30	40	25	1.194
80	500	30	40	50	1.040
75	450	50	50	20	1.057
75	500	50	50	20	1.014

Table 3: Small representative subset of the larger settings scan.

Next, I applied my candidate-only pipeline to each selected run, and the results were combined through the inverse-variance weighted mean described above. This gave

$$\tau_\mu = 2.04 \pm 0.06_{\text{stat}}\ \mu\text{s}.$$

The systematic uncertainty was then estimated from the variation in the fitted lifetime under reasonable changes to the analysis settings. When changing each variable, I restricted the estimate to the preferred stable-analysis region with settings that preserved the same physical logic, gave stable fits, and avoided extreme binning or fit ranges.

Across this final subset, the fitted lifetime varied by approximately $\sigma_{\text{sys}} \approx 0.10\ \mu$s. This was taken as the systematic uncertainty associated with cut choice, binning, fit-range dependence, and residual background modeling.

9 Lifetime Fit

Figure 9 and Figure 10 show the candidate-only fit for Run 44, a run with trigger logic using slab 1 and slab 2, and one of the largest runs used in the analysis. This run particularly was weighted heavily in the final calculation because of its larger candidate sample and smaller fitted uncertainty.

The plotted settings used were: $80 \leq \Delta t \leq 500$ ticks, $V_{10} \geq 30, V_{11} \geq 40, 25$ bins.

This fit gives

$$\tau = 2.06 \pm 0.11\ \mu\text{s}, \qquad \chi^2_\nu = 0.957.$$

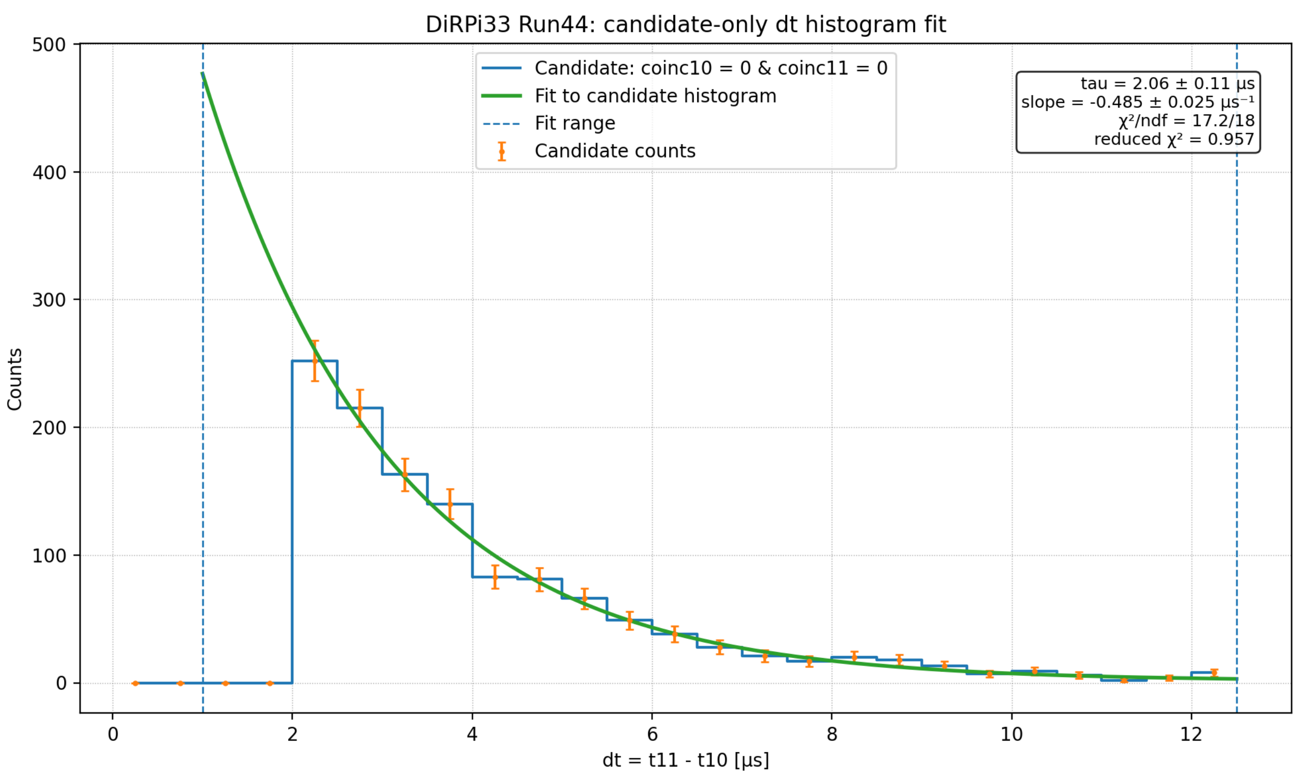


Figure 9: Candidate-only lifetime fit for relevant Run 44, largely driving the final weighted result. Linear scale.

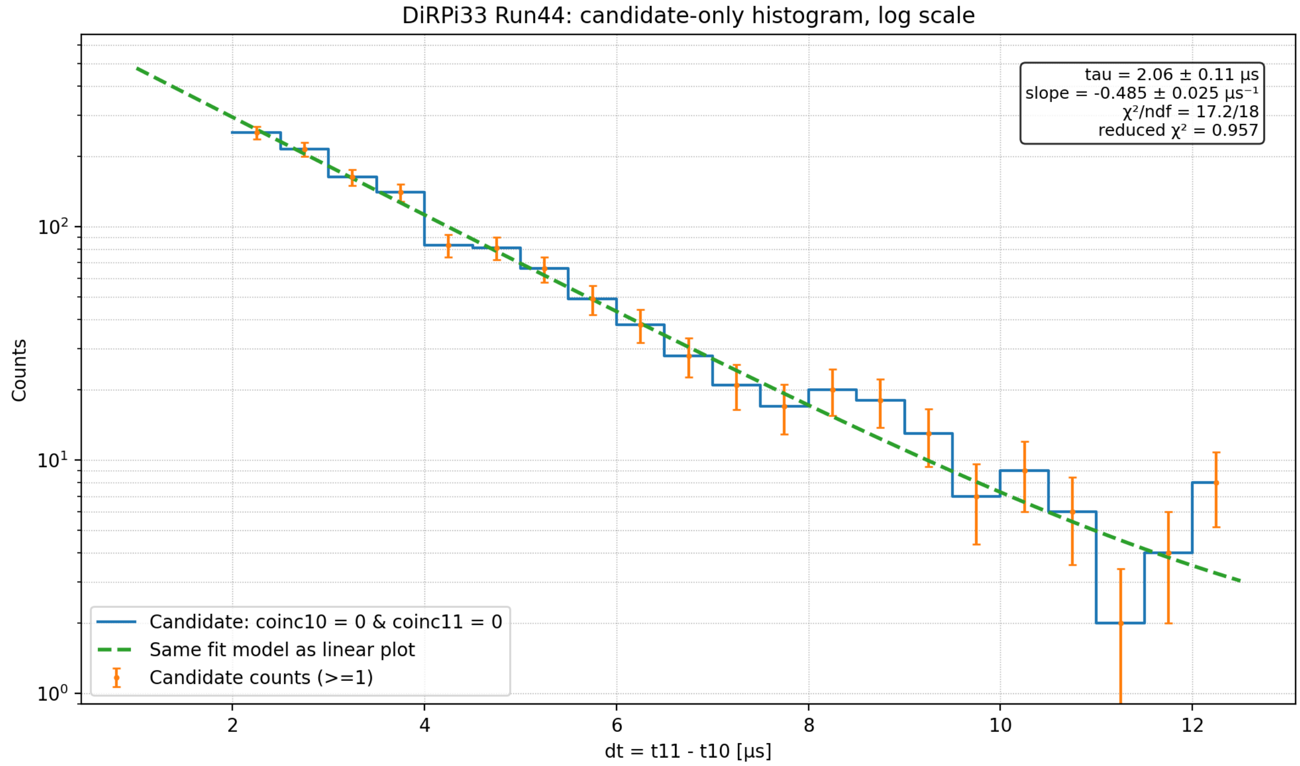


Figure 10: Candidate-only lifetime fit for relevant Run 44, largely driving the final weighted result. Log scale.

10 Background-Subtraction Cross-Check

I also programmed and tested a more explicit background-subtraction method. The idea was to form a control sample from events with recorded slab-3 coincidence activity and scale it up to match the shape of residual backgrounds in the candidate distribution. This way, subtracting the control sample histogram from the candidate sample histogram would remove the background in a subtracted histogram, and an added constant in the fit is unnecessary. In this method,

$$\text{candidate: } \texttt{coinc10} = 0 \text{ and } \texttt{coinc11} = 0,$$

$$\text{control: } \texttt{coinc10} = 1 \text{ or } \texttt{coinc11} = 1.$$

If $C(t)$ is the candidate histogram and $B(t)$ is the control-background histogram, the subtracted spectrum was

$$S(t) = C(t) - \alpha B(t), \qquad \alpha = \frac{N_{\text{candidate,tail}}}{N_{\text{control,tail}}}.$$

For Run 44, this method gave

$$\tau = 1.85 \pm 0.13\ \mu\text{s}, \qquad \chi^2_\nu = 0.667.$$

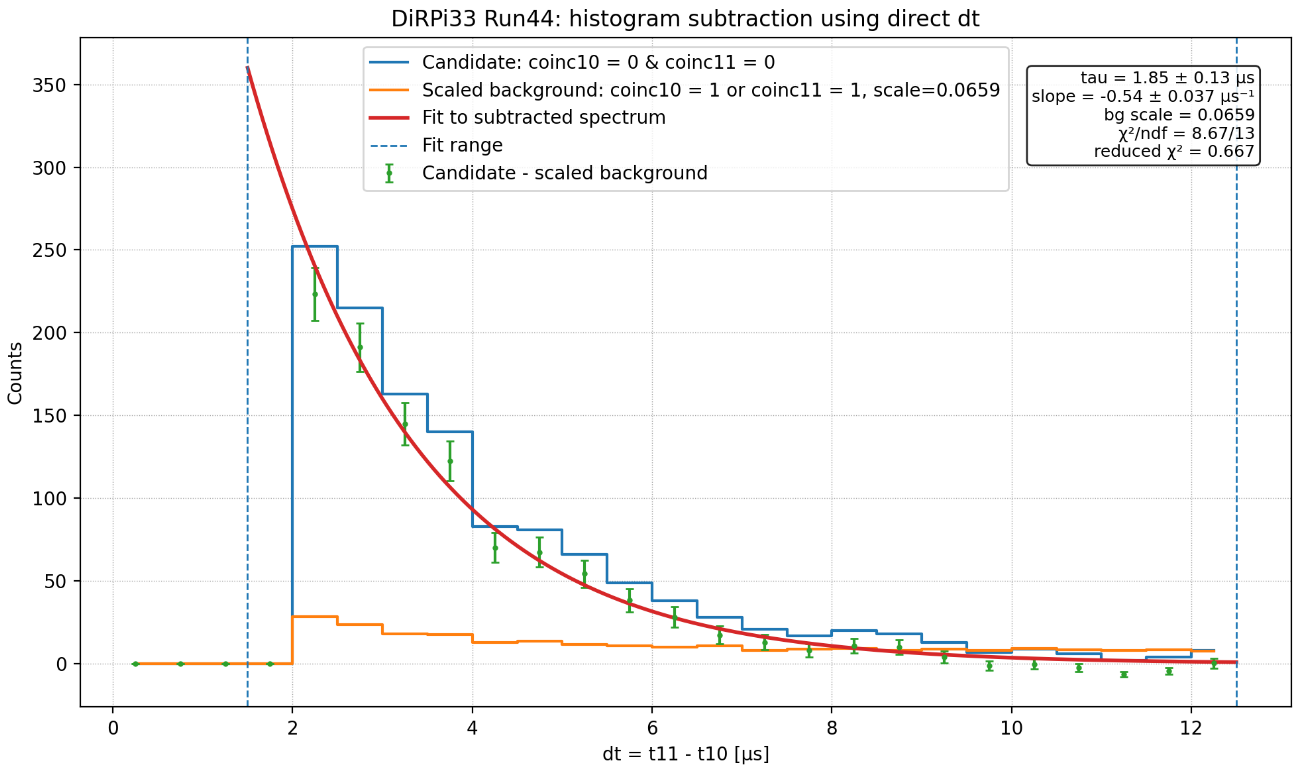


Figure 11: Background-subtraction cross-check for Run 44. The control sample was scaled to the late-time tail of the candidate sample and subtracted before fitting. The shifted lifetime relative to the candidate-only fit showed that this method was sensitive to the control-sample assumption.

The background-subtraction result was lower than the candidate-only result, showing that background modeling is a real source of systematic uncertainty. Repeating this method across selected runs gave results that were generally shifted lower than the candidate-only values, even when reasonable fit settings were used. Since the subtraction assumes that the coincidence-tagged control sample has the same shape as the residual background inside the candidate sample, I treated it as a systematic cross-check rather than using it for the final measurement.

11 Final Result

The final measured lifetime was

$$\tau_\mu = 2.04 \pm 0.06_{\text{stat}} \pm 0.10_{\text{sys}}\ \mu\text{s}.$$

Combining uncertainties in quadrature gives

$$\sigma_{\text{tot}} = \sqrt{(0.06)^2 + (0.10)^2} = 0.12\ \mu\text{s},$$

so the result can also be written as

$$\boxed{\tau_\mu = 2.04 \pm 0.12\ \mu\text{s}.}$$

The Particle Data Group accepted free-muon mean lifetime is

$$\tau_{\mu,\text{PDG}} = 2.1969811 \pm 0.0000022\ \mu\text{s}[1].$$

The percent difference between my measurement and this accepted value is

$$\frac{|2.04 - 2.1969811|}{2.1969811} \times 100\% = 7.1\%.$$

The accepted value is slightly above my combined uncertainty interval,

$$2.04 \pm 0.12\ \mu\text{s} = [1.92, 2.16]\ \mu\text{s}.$$

The PDG uncertainty is much smaller than the uncertainty of this measurement, so it does not affect the comparison. The remaining difference is plausible given the smaller selected candidate samples, the difficulty of completely isolating true stopped-muon decays from raw data, and the sensitivity to variable changes and background modeling.

12 Conclusion

This project measured the lifetime of stopped cosmic-ray muon candidates from scintillator detector data. The final value was

$$\tau_\mu = 2.04 \pm 0.06_{\text{stat}} \pm 0.10_{\text{sys}}\ \mu\text{s},$$

or 2.04 ± 0.12 μs with combined uncertainties. The result was within 7.1% of the accepted free-muon lifetime.

The most important outcome was the full experimental analysis process that unfolded in the quest for this measurement. Reasoning through each step of the measurement, testing and upgrading run configurations, hypothesizing backgrounds, choosing event cuts, and developing the Python analysis pipeline contributed to a more accurate and largely successful measurement.

The measurement may be further improved by taking more data for longer periods of time in the most consistent modes. The subtracted-background model was not very consistent with the candidate-only model, so improvements with background selection and scaling may better estimate the background in the larger dataset and give a more accurate value.

Acknowledgments

I thank Professor David Stuart and the Stuart Group for guidance, access to lab materials, and project support. I also thank the students who contributed to making the detector setup and analyzing run data.

References

[1] Particle Data Group, "Muon," *pdgLive, 2026 Review of Particle Physics.* The PDG listing reports the free-muon mean lifetime as $(2.1969811 \pm 0.0000022) \times 10^{-6}$ s. Available at `https://pdgprod.lbl.gov/pdgprod/pdgLive/Particle.action?node=S004`.