Elements: Improving tools based on data-description standards for gigabyte-scale data sets

March 31, 2019

Amy Roberts

University of Colorado Denver

Project Summary

Elements: Improving tools based on data-description standards for gigabyte-scale data sets

Overview: The PI proposes to improve a set of data-analysis tools that require a description of the data format - rather than requiring a particular data format - to handle the gigabyte-scale datasets common in nuclear physics. The long-term goal of this work is to increase the accessibility of science analysis for both active researchers and science students.

In pursuit of this goal, this proposal has three objectives: (1) Improve existing software that provides access to data in any user-described format so that it works well with gigabyte-scale datasets. (2) Build community adoption and support of these tools. (3) Increase access to analysis of scientific data sets for undergraduate students, early-career researchers, and scientists who do not have access to dedicated software support. These objectives will be met as follows:

(1) Improving standards-based analysis software: Scientists who wish to analyze custom-format data can already do so with existing open-source software, e.g. kaitai-struct. However, this software has poor performance for files larger than a gigabyte, limiting its usefulness to the nuclear physics community. The PI proposes to improve this software to allow responsive analysis of gigabyte-scale datasets in python. This work is feasible thanks to libraries built and maintained by the IRIS-HEP collaboration that provide an intuitive interface to data structures optimized for rapid access to large, event-based data sets.

(2) Building community: The PI proposes yearly workshops to bring together developers and scientists for training, user feedback, and development of this software.

(3) Increasing Access: To fully participate in nuclear physics research, students must have extensive scientific computing skills. Mentor networks help some students through this maze but are not equally available. The PI requests funding for undergraduate students to develop and test documentation and training materials for the standards-based analysis tools and for fundamental scientific computing concepts.

Intellectual merit: This work supports the science effort of the Super Cryogenic Dark Matter Search, an experiment that seeks to better understand the nature of dark matter. This work supports additional high-priority science in the nuclear physics community such as the origin of the elements in the universe, fundamental symmetry testing, and radiation monitoring for national security.

Broader impacts: The PI proposes this work because it would be immediately useful to her own work on dark matter and because software that can provide easy access to any user-described data format would meet an urgent need of communities dealing with gigabyte-scale data. The PI requests funds to develop and field-test documentation for this software to maximize its usefulness to the community.

Right now, the necessity for local development of analysis code restricts participation in science analysis. With a small set of well-documented community tools based on data-description standards, along with open-access material introducing the computing concepts required to use those tools, the PI hopes to increase research access to a much wider group of individuals and increase the time expert scientists can spend doing science.

Contents

1 Introduction

1.1 The Problem

The PI proposes improvement and community building for a common set of tools to make it easier for scientists to access their data.

In nuclear physics, in neuroscience, in remote sensing applications, custom data formats are commonplace both for existing and newly-recording datasets. This fractures the software ecosystem. Even though the nuclear physics community performs largely similar analyses on largely similar datasets, groups typically develop their software locally. Poor documentation of these isolated code bases can have real, negative impact on science progress.

S has a data set that would be valuable to include in an analysis that combines different types of carbon scattering data. Because the data was taken with a polarized beam, it would provide a valuable constraint to her global fit. But she can't include it because the only code that can read the data won't compile on her machine. She's not familiar enough with Fortran to fix it and the original author died a few years ago. She can write a paper without the data, but the science is much less interesting.

P would like to add an additional detector to his setup – this would allow him to set a stronger limit on a reaction rate that effects how stars create heavy elements. A colleague offered to lend him a digitizer to instrument his extra detector, but the data acquisition code he has is for CAEN instruments and the loaner digitizer is from XIA. After burning a weekend trying and failing to adapt his code to handle the XIA data, he decides to try the experiment without the extra detector. It won't be as useful without angular information but it might give a reasonable constraint on the integrated cross section.

All Q wants to do is look at the data from a detector used to search for dark matter. He knows where the files are, but when he tried to open it in Word it looked like a bunch of garbage symbols. Three frustrating days later, he now knows the data is "binary" and he's downloaded "source code" that should allow him to look at the data. When none of the commands in the readme file seem to work, he asks his professor for help, who tells him to talk to the postdoc, who apparently doesn't respond to emails. It takes almost an entire summer, but Q does eventually figure out how to compile the code, run the code, and look at a single event trace.

1.2 Building a Solution

One way to solve this problem is to define a standard way to describe a data format. This allows development of tools that provide access to that data based on its description. Scientists can then keep their custom-format data and by providing a description of that data gain access to analysis tools that are supported by a broader community. The work I propose here works towards this solution in two ways:

The long-term goal of this work is to build and create community buy-in for a set of tools for the nuclear physics community to access binary data. The intent is to address one piece of a common software need in the community that is repeatedly solved in isolation and that costs the entire community unnecessary time and limits access to data and science. In pursuit of this goal, this proposal has four objectives: (1) Define a data-description language that allows scientists to describe any binary data format. (2) Improve existing software that automatically builds a convenient data-access library for any user-described data so that it works well with gigabyte-scale datasets. (3) Build community and support adoption of these community tools. (4) Increase access to analysis of original scientific data sets for undergraduate students, graduate students and postdocs, early career researchers, and scientists whose labs do not have access to dedicated software support. These objectives will be met as follows:

(1) Defining a data-description standard: The PI will use two already-existing standards that both have broad user support, DFDL and Kaitai-Struct. Kaitai-Struct will be used and where necessary further defined as the software created around this standard is better-suited to the needs of data analysis.

(2) Improving standards-based analysis software: Scientists who wish to generate a convenient library that allows them to read custom-format data can already do so using the existing Kaitai software. However, this software has poor performance in python for files larger than a gigabyte and this significantly limits the usefulness of this software to the nuclear physics community, particularly as data acquisition systems move to storing digitized pulses rather than pre-processed data. The PI proposes to create another Kaitai target that allows scientists to auto-generate a library that targets the awkward-array interface built and maintained by the IRIS-HEP collaboration. This data structure is designed to give rapid data-analysis access to the high energy physics community and is routinely tested on data sets that exceed several hundred gigabytes.

(3) Building community: The PI proposes to hold yearly workshops to bring together developers and scientists for training, user feedback, and software development. In addition, the PI has requested funds for graduate and undergraduate stipends and travel awards to foster community participation.

(4) Increasing Access: Many students who have the interest and ability to participate in science analysis are blocked by a lack of training in scientific computing. The often-complex scientific computing ecosystem is poorly represented in most physics curriculum and even computer science courses do not usually prepare students for the variety of systems they need to navigate to do typical data analysis. The PI requests funding for undergraduate students to develop and test documentation and training materials for fundamental scientific computing concepts and skills in addition to testing the documentation for the standards-based analysis tools.

1.3 Long-term research goals

1.4 The Team

This paragraph needs to be added after we do decide on the sections This project description is organized as follows. Sections **??**–**??** summarize the science and hardware of the SuperCDMS SNOLAB experiment. Section **??** describes the underground facilities to be used in the pre-operations portion of this proposal. Section **??** describes the plan of work of this effort, and Section 3 describes the schedule. Section 4 gives an overview of the management for this proposed grant, and Section 5 describes the broader impacts of the work. Finally, Section 6 summarizes the results from previous support.

2 Plan of Work

2.1 Introduction

2.2 Science-driven

This project serves the immediate needs of researchers in the dark matter community and the experimental nuclear physics community by providing a common toolset for analyzing data in any format.

The PI expects that

- Multiple research projects across the NSF directorate will be more productive because they can use existing, documented tools rather than building their own. The PI intends to estimate this impact with citations from scientific papers.
- Increased involvement of undergraduate researchers in science analysis due to improved documentation and an extended support network. The PI intends to measure this through undergraduate involvement in her own lab, community surveys, and tracking community forum data.
- Several example analyses will be publicly released, with accompanying documentation and support information for pre-requisite computing skills. The intent is for these educational materials to be accessible to someone with no domain knowledge. The PI believes these training materials will be an equitable training resource.

2.3 Innovation

A common limitation of data-analysis software that is entirely home-grown is that it does not scale as data grows and changes - a human has to update or rewrite the code if the requirements change substantially.

This library makes heavy use of existing libraries. The benefit is that as those libraries improve and scale to larger data sets, this software inherits that improvement.

In both cases, significant human effort is needed to adapt the software to changing data and needs. But by leveraging well-supported, open-source libraries, the burden is shifted away from an individual scientist and towards an active community that is highly motivated to solve similar problems. This library serves as sugar to allow scientists with many different data formats to take advantage of these popular libraries.

The danger to this approach is the same - if these libraries lose community support or focus on very different problems then this library will lose relevance over time. To mitigate this risk, the PI is focusing on integrating with a library supported by the IRIS-HEP collaboration and with pandas, which enjoys extreme popularity in the data science community.

2.4 Close collaboration among stakeholders

The PI proposes to engage both cyberinfrastructure experts and the experimental nuclear physics community by (1) working closely with pilot experiments to build software that works effectively for scientists analyzing event-based data, (2) holding yearly workshops intended to foster interaction between the scientists using the software and cyberinfrastructure developers, and (3) attending conferences that will allow outreach to the scientific community (for example, the Low Energy Community Meeting) and the cyberinfrastructure community (for example, CHEP).

The PI has working relationships with scientists in the SuperCDMS collaboration and the XIA corporation, both of which are interested in exploring the proposed software as solutions to analysis needs.

In addition, the IRIS-HEP collaboration is interested in this work as it would extend their awkward-array library to a broader audience. Awkward array was developed by Jim as part of DIANA/HEP (http://diana-hep.org, OAC-1450377) and that effort will continue with IRIS-HEP (http://iris-hep.org, OAC-1836650). Collaborating with IRIS-HEP gives us access to experienced cyberinfrastructure developers who have focused on developing software suitable for terabyte-scale data.

2.5 Building on existing, recognized capabilities

The proposed work builds on existing capabilities and communities in several ways:

- **The PI proposes to use already-existing data description languages.** The languages Katai Struct and the Data Format Description Langauage both have active communities and tools that work with data when provided a description. Kaitai Struct is the target for the proposed work because (1) it is more human-readable than than the XML-based DFDL, and (2) Katai Struct generates code libraries that allow users to load their data into the programming environment of their choice; DFDL currently works by providing an XML or JSON equivalent of the binary data. While this is a powerful approach because any language with an XML or JSON parser can now read the data, it also produces a secondary data file that is an order of magnitude larger than most binary files. This makes DFDL, in its current state, unusable for scientists with gigabyte-scale data sets as it would make the required storage space for analysis prohibitively expensive.

- **The PI proposes to use already-existing infrastructure for the data-analysis library.** Scientists who would like to avoid writing custom software to read their binary data can already use the Kaitai Struct compiler to generate libraries to read their data in python, C++, and a multitude of other languages. The advantage is that there is substantial support documentation and an active community available for troubleshooting. The disadvantage is that the current Kaitai Struct python compiler stores the data in a structure that does not provide adequate speed performance for gigabyte-scale data sets. By improving the existing Katai Struct compiler software, we can build a science-ready analysis library and scientists can benefit from the existing community support and documentation.

- **Use a supported and optimized data structure** for the improvements to the Kaitai Struct compiler. The "awkward-array" library was developed by DIANA/HEP and is now supported by IRIS-HEP and is part of a set of libraries designed to provide flexible data-analysis tools for the high-energy physics community. The awkward-array data structure is optimized for fast queries on an event-based data set and as such is ideal for the majority of nuclear physics data. By choosing this data structure as the target, we bring the optimized and convenient analysis environment of awkward-array to any scientist who describes their data with the Katai Struct language.

- **Provide analysis tools for the python environment and training materials that take advantage of the python ecosystem.** Python is a popular analysis environment in the field of big-data and has enjoyed significant adoption in the scientific community; enough so that python support is compiled in the dominant high-energy physics software, ROOT, by default. By providing a python library for data analysis, scientists can make use of a full ecosystem that supports data analysis: numpy for convenient array manipulation; scipy for fitting; matplotlib for producing publication-quality figures; and even numba for easy compilation of code that needs to run fast. This entire environment is easily installed - even for users without administration privileges - through the Anaconda Python distribution. There are many free and paid programming environments that are availble, notably the Jupyter environment. Code written in this environment is particularly nice as a tutorial because it is rendered nicely on github, gitlab, and interactive notebooks can be opened in one click through binder. By providing a small set of introductory documentation, scientists can benefit from the effort the python community has put in to lower the barrier for use.

2.6 Project plans, and system and process architecture

The timing of the proposed work is driven by the proposed, yearly workshops that focus on (1) teaching scientists how to use the tools to access their data, (2) working with scientists to perform their analyses in the python environment, (3) identifying improvements needed for the software to be easy to learn and useful in analysis, and (4) bringing developers into close contact with the science community using their tools. Each workshop will result in an updated roadmap for the software.

Thus, the workshops - and software releases that include testing, documentation, and example analyses - are the primary milestones of the proposed work.

The work for each yearly cycle can be broken down into the following categories: development of basic computing skills learning material; development of the data-access library; planning and execution of the workshop; and a community-driven update of the roadmap. See Table 2.6 for details on who will perform this work.

The minimum requirements of the work determine the work plan and are the following:

1. If students or staff move on to other positions, their replacements should be able to get up to speed in a month or less.
2. Someone with no domain knowledge but reasonable persistence should be able to run the example analysis within a week.
3. Someone with no domain knowledge but reasonable persistence should be able to analyze their own data within a month.
4. A scientist who uses the access-data library to obtain a science result should know how to cite the software.
5. A scientist experiencing trouble using the software should be able to determine how to get help quickly (within five minutes of searching).
6. A scientist who wishes to improve the code should be able to quickly determine how to contact the developers and how to change, test, and push the code.

The scope of the proposed software is relatively modest: copy an existing framework and adapt it so that it stores data in awkward-array structures rather than slower, dictionary structures. The development and testing of this code will take time - but reference code for similar work exists, there is robust community support, and there is a developer guide that gives specific instructions for developers who wish to extend the existing Kaitai Struct code in this way. The proposed work is feasible because it connects two libraries that are both designed for this purpose.

The majority of the proposed work is in making this software easy to use for scientists, and making it easy for the community to participate in the direction and development of the software. This requires robust documentation for both users and developers. Users will require installation instructions, instructions for using the library, and guidance on how to adapt the examples for their own analysis needs. Developers will need additional documentation: instructions for changing the code and testing the code, and instructions and guidance on contributing their changes to the project.

To meet the minimum requirements, the workplan involves creating of initial documentation and an automated testing suite by the Professional Research Associate and example analysis created by the Master's student.

Undergraduates will begin either by working on new scientific computation skills or improving or adding to material of already-developed scientific computation skills. Students who join the lab currently have two first projects to chose from: working through an introductory lab on water simulation, or working through an example analysis of gamma-spectroscopy data. The students try to perform their work using existing documentation; the PI provides guidance when this is inadequate. This provides an opportunity for students to design improvements and learn the basics of contributing to a code repository. In addition to providing valuable training for the students, this process identifies gaps in the documentation that are often invisible to experts.

This initial work is expected to result in improvements and additions to web-accessible tutorial materials. In addition, this training will provide a foundation that will allow the students to attempt the following actions:

- Successfully follow a simple example analysis using the data-access library.
- Successfully follow a tutorial to make and share a change to the library documentation.
- Successfully follow a tutorial to make, test, and share a change to the library source code.

Work	Who	Notes
Roadmap & workplan development	PI, PRA, Master's student	
Basic skills materials	Undergrads	
Data-access library development		
code dev	Professional Research Assistant	this includes testing, user documentation, and dev documentation
contributing	Professional Research Assistant	this includes automated testing and contribution guidelines and instructions
example analyses	Master's student	
documentation testing	Undergraduates, Master's student	documentation testing may feed back into additional skill documetation
Workshop		
Recruiting	PI	
Organization	Professional Research Assistant, Undergraduates	
Pre-workshop analysis coordination	PI, Master's student, PRA	
Community-driven update of roadmap	PI, PRA, Community	

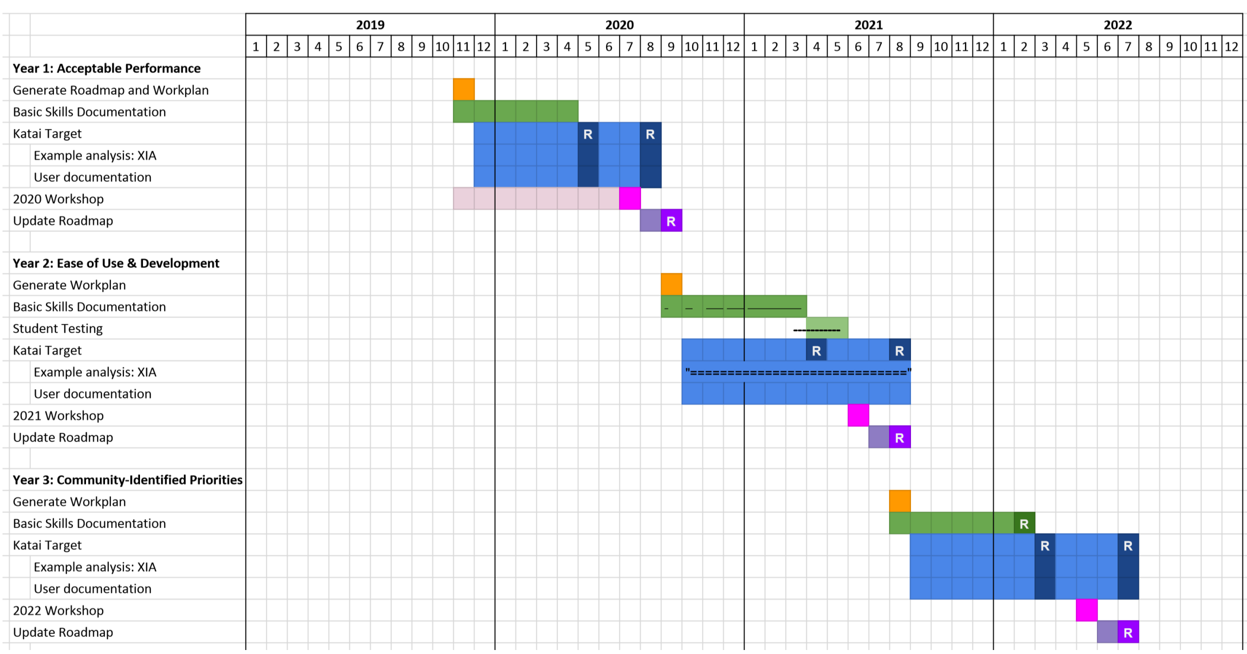

Figure 1: Schedule for the proposed work.

Architecture of the software: The architecture of the Kaitai Struct compiler that targets an awkward-array data structure will follow that of the existing Katai Struct software. Implementing a Kaitai Struct compiler for a new language requires

1. Writing a "runtime library" that provides a standard stream interface in the target language. For example, one of the functions every language needs to have defined is a method that returns the size of the file or string stream. By writing a "size" function for the language of interest that follows the Katai Struct API, the code generation becomes simpler.
2. Writing a "compiler" that translates Kaitai Struct concepts, implemented in Scala, into the target language.
3. Writing a test runner for the new language.

The proposed work targets the python environment, and there is already a Kaitai Struct python compiler. The "compiler" for the python implementation, however, stores data in native-python data structures that provide inconveniently slow access to standard queries on large, gigabyte-scale data sets.

However, the changes that need to be implemented to instead store the data in the faster awkward-array data structure are restricted to the compiler code. The runtime library provides a convenient interface for reading data from a file or stream - this code only cares about the file system interface and does not need to change. The python test interface will need to be updated as the access syntax for the data will change slightly.

Although there is opportunity for improving the speed of the data load, this development will instead focus on adhering to the existing format and style of Kaitai Struct. The goal is to make the existing Kaitai Struct community useful to scientists who work with gigabyte-scale data sets; waiting for a few minutes for the data to load is not ideal but is typical of many locally-built solutions. We can address the more-critical issue of rapid data queries while staying well within the existing framework of Kaitai Struct and intend to do so for the initial implementation of the software.

If we find that data-load times are a significant issue for the nuclear physics community then we will consider more substantial changes to the Kaitai Struct compiler and runtime libary.

```
def size(self):
    # Python has no internal File object API function to get
    # current file / StringIO size, thus we use the following
    # trick.
    io = self._io
    # Remember our current position
    cur_pos = io.tell()
    # Seek to the end of the File object
    io.seek(0, SEEK_END)
    # Remember position, which is equal to the full length
    full_size = io.tell()
    # Seek back to the current position
    io.seek(cur_pos)
return full_size
```

```
uint64_t kaitai::kstream::size() {
    std::iostream::pos_type cur_pos = m_io->tellg();
    m_io->seekg(0, std::ios::end);
    std::iostream::pos_type len = m_io->tellg();
    m_io->seekg(cur_pos);
    return len;
}
```

Architecture of the user documentation: User documentation should make it possible for users with little to no domain knowledge to use the data-access library for science. Documentation for the use of the library will be stored as text files in the repository with the code. The files will be written in markdown syntax to improve their readability; this will also render them nicely on cloud-based repository hosts such as github and gitlab. The following documentation will be provided:

1. How to get help with questions or issues about the library.
2. How to install the library and its dependencies.
3. An overview explaining what the user will need to provide (data and a description of the data) and what the library will provide (software to read that data).
4. A tutorial walking through the use-case of a scientist looking at simple data with a custom format.
5. Links to additional resources detailing more complex data formats and more complex analyses.
6. Citation guidelines.

Architecture of the developer documentation. Documentation intended to facilitate development of the code will be stored in the repository alongside the code. Text files referenced in the top-level README file will detail, for every repository,

1. How to install, develop, and test the code for individuals who wish to make changes.
2. How to contribute changes back to the project. This will provide instructions on the version control practices used by the repository maintainers and instructions for implementing the tests required for changes to be considered for merging with the main code base.

Architecture of the basic scientific computing skills documentation: Documentation of basic computational skills and concepts will have several possible forms: (1) Text and images, (2) tutorial videos, (3) jupyter notebooks, (4) printable images that illustrate a focused concept, and (5) links to recommended resources such as Software Carpentry tutorials.

All materials will be licensed with a permissive, open-source license such as CC-BY or MIT. The source for all the materials will be publicly available through a public host such as github or gitlab and will be archived on a content-tracker such as the Open Science Framework or Figshare. Videos will be released on YouTube and licensed CC-BY.

All materials will be disseminated using a static site generated by Antora. Antora is specifically designed for documentation and allows a user to specify a set of repositories containing text files formatted in the Asciidoc markdown language to build a single, searchable documentation site. Because Antora generates a static site, free hosting services are readily available. This solution allows my students to focus on creating material to explain core concepts and practice interacting with version control rather than spending time wrestling with web development.

The topics students choose to document are largely student-led, with some guidance from the PI. Spring 2019 marks the inception of this project, and the concepts chosen by students for illustration have focused on (1) tutorial-format guide for installing python and running a basic python-based analysis of gamma spectroscopy data, (2) instructions for using a docker container to simplify installation of a complex software environment, and (3) a poster explaining what an executable file is.

Documentation that will be provided in this format alongside the scientific computing resources will include

1. instructions on where to get help with the material and how to provide feedback and and file bug reports
2. instructions for those who wish to contribute to the documentation
3. instructions for the deployment of the documentation

2.6.1 Security

2.6.2 Trustworthiness

2.6.3 Provenance

2.6.4 Reproducibility

2.6.5 Usability

2.6.6 Adaptability

2.7 Deliverables

2.8 Metrics

2.9 Sustained and sustainable impacts

3 Schedule

Figure 2 gives a broad view of the activities of the SuperCDMS collaboration centered on the SuperCDMS SNOLAB experiment, including the DOE/NSF G2 Project, the pre-operations and commissioning period covered by this proposal, the actual operation of SuperCDMS SNOLAB (also detailed in the Experimental Operations Plan), and the R&D program that will naturally lead to upgrades for SuperCDMS SNOLAB that allow it to probe the low-mass dark matter region down to the neutrino floor.

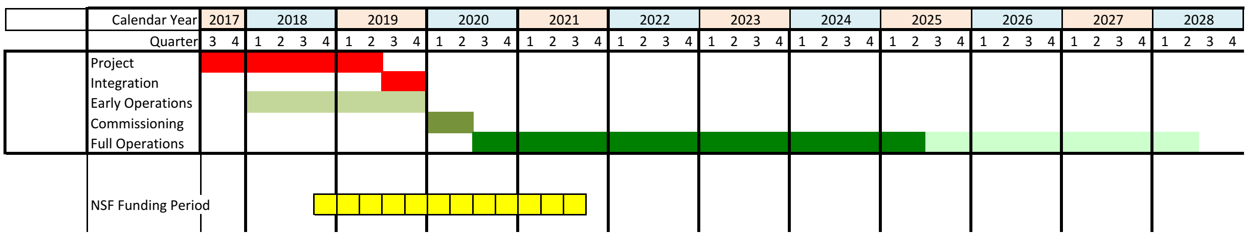

Figure 2: Schedule for the SuperCDMS SNOLAB Experiment.

Year 1: 7/2018–6/2019
During the first year the focus will be on calibration of the small HV test detectors from Stanford University with the Northwestern ADR at TUNL. These detectors will then be taken to NEXUS, to perform cross-calibration measurements in preparation for the large detector calibrations. The first pre-production detectors arrive at CUTE in January 2018, and the performance and background measurement campaign begins soon after.

Year 2: 7/2019–6/2020
The second year is dominated by runs of pre-production detectors at NEXUS for calibration and at CUTE for performance testing and background measurements. The first production SuperCDMS SNOLAB HV tower arrives at CUTE in 12/2018, with testing commencing in 1/2019. Performance and run optimization of all six detectors of the HV tower will continue until 6/2020.

Year 3: 7/2020–6/2021
There is no activity planned for NEXUS on this grant in the third year. At CUTE, the SuperCDMS SNOLAB HV tower science run will be taking data until 11/2019, when the tower will be taken out of CUTE to prepare it for installation in the SuperCDMS SNOLAB experiment. Commissioning activities for SuperCDMS SNOLAB will be the main focus of this grant in the last six months of the award, along with the collaboration-wide effort in analysis and publication of all the data products from this work.

4 Management

The management team of this NSF collaborative proposal will be chaired by the Northwestern PI and consist of all the PI's and Co-PI's of the proposal. They will interact with the NEXUS, CUTE, and SNOLAB facility managers to coordinate all the activities delineated in this proposal, as well as with the SuperCDMS Collaboration for science analyses and training of students and postdocs.

The work funded by this proposal is part of the full set of SuperCDMS SNOLAB operations and commissioning activities, some of which are outside the scope of this proposal and are funded by DOE or Canadian funds. The management team for this proposal will be in close coordination with all parties involved in the SuperCDMS SNOLAB Project and Operations, as well as R&D, and other SuperCDMS Collaboration activities. These multi-agency collaboration-wide activities will be coordinated by a team consisting of the SuperCDMS SNOLAB Project and Operations Managers, the SuperCDMS Spokesperson, and the NSF Operations PI. This team will meet regularly to coordinate activities.

The early operations phase of the SuperCDMS SNOLAB experiment covered by this proposal is when it is both necessary and appropriate to ramp up the operations management and support functions for SuperCDMS SNOLAB, which should all be in place for the commissioning phase of the experiment in order to ensure a smooth transition to the operations phase as quickly as possible.

The detailed planning for commissioning of SuperCDMS SNOLAB will be led by a scientist identified to be the SuperCDMS SNOLAB Commissioner, with this role naturally evolving into that of a Run Coordinator during the operations phase. Details of the operations phase functions and management (which are outside the scope and period of performance of this project) are given in the SuperCDMS SNOLAB Experimental Operations Plan.

4.1 Facilities Management

Since much of the work in this proposal occurs in dedicated underground facilities, we outline here the management of these facilities and their relationship to this grant's management.

4.1.1 NEXUS

The NEXUS facility will receive primary management oversight from the Northwestern PI. In addition, SuperCDMS collaborators at Fermilab will provide support to interface with Fermilab as needed. A fraction of an FTE will be provided by this grant for an operations manager, to oversee the activities delineated in this proposal.

4.1.2 CUTE

The CUTE facility will receive primary management oversight from our collaborators at Queen's University. This grant's management team will coordinate with Queen's to oversee the CUTE activities delineated in this proposal.

4.1.3 SNOLAB

As mentioned above, the commissioning of SuperCDMS SNOLAB is an activity coordinated at the collaboration level. The PI of this grant will be part of the Operations Management team and will work closely with the SuperCDMS SNOLAB Operations Manager to coordinate and integrate the NSF-sponsored commissioning work with the overall multi-agency commissioning effort. SNOLAB conducts a Gateway review process for projects which includes pre-operations activities, operation of the experiment, and decommissioning. For Gateway 1 and 2, the Project Director is the primary contact for review of the design and fabrication of the SuperCDMS SNOLAB experiment. For Gateway 3 and 4, the Operations Manager is the primary contact for review of the experiment pre-operations and operations. The Collaboration Canadian PI is involved in all interfaces between SuperCDMS SNOLAB project and operations and SNOLAB management. This grant's PI will be in close coordination with the Project Director, the Operations Manager and the Canadian PI to oversee the commissioning activities supported by this proposal.

5 Broader Impacts

The fundamental goal of the proposed work is to make data easily accessible to individuals who want to answer science questions.

This work is immediately useful to the PI's own collaboration and - if the work is successful - useful to practicing scientists who can more-easily extract information from their datasets.

But this work extends beyond the community of active scientists to scientific learners.

Technical Development CDMS and SuperCDMS have developed, and continue to advance detector technologies that have significant impact both inside and outside of the dark matter community. These technologies include electrothermal feedback, transition-edge sensors (TESs); detection and utilization of athermal phonons; large, kilogram-scale phonon-mediated germanium and silicon-based detectors; arrays of large-scale, low-noise SQUIDs; and usage of LPN HEMPTs for ionization readout.

The field of direct detection of dark matter has benefited from these advances. The EDELWEISS and CRESST collaborations have both implemented some of these concepts. In addition, SuperCDMS pioneering of very-low-mass WIMP detection has had a strong impact on the field. There are now many R&D efforts focused on very-low-mass WIMP detection. Multiple of these programs are run by PIs who were primarily engaged in liquid noble direct detection. The concepts developed by SuperCDMS have had far reaching impacts benefiting even direct detection experiments using the liquid nobles.

The broader, basic-science community has benefited from SuperCDMS technology advances. The TES technology is now a core technology for multiple cosmic microwave background experiments including APEX, South Pole Telescope, and PolarBear. It is being rapidly adapted for far-infrared instruments including Super-SCUBA, single-photon spectroscopy, and large X-ray calorimetry experiments at NIST and NASA GSFC including the forthcoming Micro-X, which will launch the first TES into space. In addition, the TES technology is beginning to impact industry as TESs are being adapted for use in electron microscope sensitive surface element analysis and are being investigated for use in quantum coherence and computing.

The direct impact of SuperCDMS research is underscored by the number of graduate students impacted and the number and quality of the SuperCDMS publications. In the past three years, SuperCDMS has produced 12 Ph.D. theses and 11 science publications not including conference proceedings. The impact of these publications is underscored by their many citations (over 3,000 since 2011).

Undergraduate, Graduate, and Postgraduate Education Direct detection of dark matter is multidisciplinary and therefore creates a broad spectrum of opportunities for students to explore and advance in many fields. From year to year, numerous undergraduates, graduate students and postdoctoral researchers from the SuperCDMS NSF-funded institutions have benefited from these collaborative opportunities and have been guided to develop precise simulations of semiconductor-based detectors, sensitive analysis techniques, statistical applications, and application of advanced software techniques. These opportunities combined with regular mentoring have afforded many high-visibility talks for students and postdocs.

Sadoulet is the director of Berkeley Connect in Physics, a mentoring program within the physics department that accepts undergraduate students at all levels. The course is a small seminar class led by a physics graduate student, under the guidance of the director. The goals of the program are to help students develop understanding, community, and career preparedness that go beyond what traditional courses provide. Interactions with graduate students and faculty play a large role throughout the semester. Under Sadoulets leadership, this course has significantly gained in enrollment with every succeeding semester since its launch in Spring 2014.

The Physics Department at CU Denver maintains an undergraduate-only program committed to providing students with hands-on research experience at this urban campus. The Department operates in close collaboration with the Physics Department at Metropolitan State University of Denver, and students from both institutions participate in the CU Denver groups activities, broadening the reach of the research program beyond a single institution. Underrepresented groups and first-generation college students are well represented in the PIs group, and many of his mentees have proceeded to graduate school in physics and other disciplines, or embarked on careers in industry at companies such as Northrop Grumman. Currently, there are eight undergraduates in the CU Denver group majoring in Physics, Electrical Engineering, and Mechanical Engineering.

Southern Methodist U. typically funds one to three undergraduate students to work with the Cooley group each year. In the past five years, the eight undergraduates who have participated in the groups research have either entered graduate programs in physics or engineering or embarked on careers in industry at companies such as Texas Instruments and AT&T.

K-12 Education SuperCDMS institutions provide strong support for K-12 education through programs for both teachers and students. Too numerous to describe in detail in this proposal, these programs take place in local schools, on campus and directly in the lab. At the core of many of the programs is engagement of students from disadvantaged backgrounds and underrepresented groups, including women.

CU Denver, U. Florida, SMU, and USD engage middle school students in physics learning through lab tours, classroom visits and science camp. Northwestern U., UC Berkeley, U. Florida, TAMU, SMU, and SD School of Mines and Technology support a range of high school programs, including summer research internships for students and teachers, hands-on research projects for underserved high school students of color, mini-courses to prepare Native American students for postsecondary education in science and engineering, interviews with physicists conducted by high school students, and structured physics courses for high school teachers.

6 Results from Prior NSF Support

In this section we describe the results from prior NSF support.

PHY-1809769
Collaborative Research: The SuperCDMS SNOLAB Experiment
Period of Support: 8/2018–7/2021
Amount of Support: $340,000
Publications: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11]
Data products from this work are available at the SuperCDMS collaboration's publications website [12].

6.0.1 Intellectual merit

This grant supports students and scientists working on an experiment that addresses one of the most fundamental problems of modern science, the nature of dark matter. The SuperCDMS SNOLAB experiment will achieve world-leading sensitivity for dark matter searches in the 1–10 GeV/c^2 mass range.

Analyses of CDMS II data demonstrated the power of improved analysis techniques [7, 4], and provided limits for alternate dark matter models [8]. It also revealed a possible signal on Si [13], the analysis and publication of which was lead by the Figueroa Group, then at MIT. Operation of 15 SuperCDMS "iZIP" detectors at Soudan since 2012 demonstrated the detectors' rejection capabilities [14] and yielded world-leading sensitivity to low-mass DM [15, 11, 2].

The first operation of a single ("CDMSlite") detector with a high voltage bias provided the world's most constraining limits for DM masses below 6 GeV/c^2 by achieving an extremely low energy threshold of 170 eV electron-recoil energy [15]. A second run of the detector reached an energy threshold for electron recoils as low as 56 eV and demonstrated the power of a fiducialization cut [2]. The Figueroa Group developed key data quality cuts for the first two CDMSlite run analyses, and is currently involved in the analysis of the third run.

6.0.2 Accomplishments

The Roberts group is funded primarily for contributions to the data acquisition and data quality systems.

A critical need facing the collaboration as we move to larger data sets is transitioning our analysis platform to computing clusters where we can submit jobs to batch queues. This requires a reworking of existing analysis software, which is primarily MatLab-based and cannot be run at SLAC.

The group began testing and documenting the installation requirements of prototype python software and has since developed the first isolated build environment, allowing reliable installation of the analysis tools across platforms.

Josh Elsarboukh has worked closely with SLAC computing division to successfully deploy this analysis environment via a web interface. This work represents an unprecedented ease of access within the collaboration and has made it possible for test facilities working on crucial R&D and calibration efforts to efficiently analyze their data.

6.0.3 Broader impacts

The SuperCDMS experimental and R&D efforts advance phonon-mediated detectors and new active veto concepts, which have already found many applications in cosmology, astronomy and industry.

This grant strongly contributes to the training of undergraduate researchers. The CU Denver Physics Department maintains an undergraduate-only program committed to providing students with hands-on research experience at this urban campus. Electrical Engineering and Mechanical Engineering students have also participated in internships in the CU Denver lab. Minorities and first-generation college students have been well represented in these internships.

References

[1] R. Agnese, A. J. Anderson, T. Aramaki, I. Arnquist, W. Baker, D. Barker, R. Basu Thakur, D. A. Bauer, A. Borgland, M. A. Bowles, P. L. Brink, R. Bunker, B. Cabrera, D. O. Caldwell, R. Calkins, C. Cartaro, D. G. Cerdeño, H. Chagani, Y. Chen, J. Cooley, B. Cornell, P. Cushman, M. Daal, P. C. F. Di Stefano, T. Doughty, L. Esteban, S. Fallows, E. Figueroa-Feliciano, M. Fritts, G. Gerbier, M. Ghaith, G. L. Godfrey, S. R. Golwala, J. Hall, H. R. Harris, T. Hofer, D. Holmgren, Z. Hong, E. Hoppe, L. Hsu, M. E. Huber, V. Iyer, D. Jardin, A. Jastram, M. H. Kelsey, A. Kennedy, A. Kubik, N. A. Kurinsky, A. Leder, B. Loer, E. Lopez Asamar, P. Lukens, R. Mahapatra, V. Mandic, N. Mast, N. Mirabolfathi, R. A. Moffatt, J. D. Morales Mendoza, J. L. Orrell, S. M. Oser, K. Page, W. A. Page, R. Partridge, M. Pepin, A. Phipps, S. Poudel, M. Pyle, H. Qiu, W. Rau, P. Redl, A. Reisetter, A. Roberts, A. E. Robinson, H. E. Rogers, T. Saab, B. Sadoulet, J. Sander, K. Schneck, R. W. Schnee, B. Serfass, D. Speller, M. Stein, J. Street, H. A. Tanaka, D. Toback, R. Underwood, A. N. Villano, B. von Krosigk, B. Welliver, J. S. Wilson, D. H. Wright, S. Yellin, J. J. Yen, B. A. Young, X. Zhang, and X. Zhao. Projected Sensitivity of the SuperCDMS SNOLAB experiment. *arXiv:1610.00006 [physics.ins-det]*, September 2016.

[2] SuperCDMS Collaboration, R. Agnese, A. J. Anderson, T. Aramaki, M. Asai, W. Baker, D. Balakishiyeva, D. Barker, R. Basu Thakur, D. A. Bauer, J. Billard, A. Borgland, M. A. Bowles, P. L. Brink, R. Bunker, B. Cabrera, D. O. Caldwell, R. Calkins, D. G. Cerdeno, H. Chagani, Y. Chen, J. Cooley, B. Cornell, P. Cushman, M. Daal, P. C. F. Di Stefano, T. Doughty, L. Esteban, S. Fallows, E. Figueroa-Feliciano, M. Ghaith, G. L. Godfrey, S. R. Golwala, J. Hall, H. R. Harris, T. Hofer, D. Holmgren, L. Hsu, M. E. Huber, D. Jardin, A. Jastram, O. Kamaev, B. Kara, M. H. Kelsey, A. Kennedy, A. Leder, B. Loer, E. Lopez Asamar, P. Lukens, R. Mahapatra, V. Mandic, N. Mast, N. Mirabolfathi, R. A. Moffatt, J. D. Morales Mendoza, S. M. Oser, K. Page, W. A. Page, R. Partridge, M. Pepin, A. Phipps, K. Prasad, M. Pyle, H. Qiu, W. Rau, P. Redl, A. Reisetter, Y. Ricci, A. Roberts, H. E. Rogers, T. Saab, B. Sadoulet, J. Sander, K. Schneck, R. W. Schnee, S. Scorza, B. Serfass, B. Shank, D. Speller, D. Toback, R. Underwood, S. Upadhyayula, A. N. Villano, B. Welliver, J. S. Wilson, D. H. Wright, S. Yellin, J. J. Yen, B. A. Young, and J. Zhang. New Results from the Search for Low-Mass Weakly Interacting Massive Particles with the CDMS Low Ionization Threshold Experiment. *Phys. Rev. Lett.*, 116(7):071301, 2016.

[3] Ciaran A. J. O'Hare, Anne M. Green, Julien Billard, Enectali Figueroa-Feliciano, and Louis E. Strigari. Readout strategies for directional dark matter detection beyond the neutrino background. *Phys. Rev.*, D92(6):063518, 2015.

[4] R. Agnese, A. J. Anderson, M. Asai, D. Balakishiyeva, D. Barker, R. Basu Thakur, D. A. Bauer, J. Billard, A. Borgland, M. A. Bowles, D. Brandt, P. L. Brink, R. Bunker, B. Cabrera, D. O. Caldwell, R. Calkins, D. G. Cerdeño, H. Chagani, Y. Chen, J. Cooley, B. Cornell, C. H. Crewdson, P. Cushman, M. Daal, P. C. F. Di Stefano, T. Doughty, L. Esteban, S. Fallows, E. Figueroa-Feliciano, G. L. Godfrey, S. R. Golwala, J. Hall, H. R. Harris, S. A. Hertel, T. Hofer, D. Holmgren, L. Hsu, M. E. Huber, D. Jardin, A. Jastram, O. Kamaev, B. Kara, M. H. Kelsey, A. Kennedy, M. Kiveni, K. Koch, A. Leder, B. Loer, E. Lopez Asamar, P. Lukens, R. Mahapatra, V. Mandic, K. A. McCarthy, N. Mirabolfathi, R. A. Moffatt, S. M. Oser, K. Page, W. A. Page, R. Partridge, M. Pepin, A. Phipps, K. Prasad, M. Pyle, H. Qiu, W. Rau, P. Redl, A. Reisetter, Y. Ricci, H. E. Rogers, T. Saab, B. Sadoulet, J. Sander, K. Schneck, R. W. Schnee, S. Scorza, B. Serfass, B. Shank, D. Speller, D. Toback, S. Upadhyayula, A. N.

Villano, B. Welliver, J. S. Wilson, D. H. Wright, X. Yang, S. Yellin, J. J. Yen, B. A. Young, J. Zhang, and SuperCDMS Collaboration. Improved WIMP-search reach of the CDMS II germanium data. *Physical Review D*, 92(7):072003, October 2015.

[5] K. Schneck, B. Cabrera, D. G. Cerdeño, V. Mandic, H. E. Rogers, R. Agnese, A. J. Anderson, M. Asai, D. Balakishiyeva, D. Barker, R. Basu Thakur, D. A. Bauer, J. Billard, A. Borgland, D. Brandt, P. L. Brink, R. Bunker, D. O. Caldwell, R. Calkins, H. Chagani, Y. Chen, J. Cooley, B. Cornell, C. H. Crewdson, P. Cushman, M. Daal, P. C. F. Di Stefano, T. Doughty, L. Esteban, S. Fallows, E. Figueroa-Feliciano, G. L. Godfrey, S. R. Golwala, J. Hall, H. R. Harris, T. Hofer, D. Holmgren, L. Hsu, M. E. Huber, D. M. Jardin, A. Jastram, O. Kamaev, B. Kara, M. H. Kelsey, A. Kennedy, A. Leder, B. Loer, E. Lopez Asamar, P. Lukens, R. Mahapatra, K. A. McCarthy, N. Mirabolfathi, R. A. Moffatt, J. D. Morales Mendoza, S. M. Oser, K. Page, W. A. Page, R. Partridge, M. Pepin, A. Phipps, K. Prasad, M. Pyle, H. Qiu, W. Rau, P. Redl, A. Reisetter, Y. Ricci, A. Roberts, T. Saab, B. Sadoulet, J. Sander, R. W. Schnee, S. Scorza, B. Serfass, B. Shank, D. Speller, D. Toback, S. Upadhyayula, A. N. Villano, B. Welliver, J. S. Wilson, D. H. Wright, X. Yang, S. Yellin, J. J. Yen, B. A. Young, J. Zhang, and SuperCDMS Collaboration. Dark matter effective field theory scattering in direct detection experiments. *Physical Review D*, 91(9):092004, May 2015.

[6] M. Pyle, E. Figueroa-Feliciano, and B. Sadoulet. Optimized Designs for Very Low Temperature Massive Calorimeters. *ArXiv e-prints*, March 2015.

[7] R. Agnese, A. J. Anderson, D. Balakishiyeva, R. Basu Thakur, D. A. Bauer, J. Billard, A. Borgland, M. A. Bowles, D. Brandt, P. L. Brink, R. Bunker, B. Cabrera, D. O. Caldwell, D. G. Cerdeno, H. Chagani, Y. Chen, J. Cooley, B. Cornell, C. H. Crewdson, P. Cushman, M. Daal, P. C. F. Di Stefano, T. Doughty, L. Esteban, S. Fallows, E. Figueroa-Feliciano, M. Fritts, G. L. Godfrey, S. R. Golwala, M. Graham, J. Hall, H. R. Harris, S. A. Hertel, T. Hofer, D. Holmgren, L. Hsu, M. E. Huber, A. Jastram, O. Kamaev, B. Kara, M. H. Kelsey, A. Kennedy, M. Kiveni, K. Koch, A. Leder, B. Loer, E. Lopez Asamar, R. Mahapatra, V. Mandic, C. Martinez, K. A. McCarthy, N. Mirabolfathi, R. A. Moffatt, D. C. Moore, R. H. Nelson, S. M. Oser, K. Page, W. A. Page, R. Partridge, M. Pepin, A. Phipps, K. Prasad, M. Pyle, H. Qiu, W. Rau, P. Redl, A. Reisetter, Y. Ricci, H. E. Rogers, T. Saab, B. Sadoulet, J. Sander, K. Schneck, R. W. Schnee, S. Scorza, B. Serfass, B. Shank, D. Speller, S. Upadhyayula, A. N. Villano, B. Welliver, D. H. Wright, S. Yellin, J. J. Yen, B. A. Young, and J. Zhang. Maximum likelihood analysis of low energy cdms ii germanium data. *Phys. Rev. D*, 91:052021, Mar 2015.

[8] R. Agnese, A. J. Anderson, D. Balakishiyeva, R. Basu Thakur, D. A. Bauer, J. Billard, A. Borgland, M. A. Bowles, D. Brandt, P. L. Brink, R. Bunker, B. Cabrera, D. O. Caldwell, D. G. Cerdeno, H. Chagani, Y. Chen, J. Cooley, B. Cornell, C. H. Crewdson, P. Cushman, M. Daal, P. C. F. Di Stefano, T. Doughty, L. Esteban, S. Fallows, E. Figueroa-Feliciano, G. L. Godfrey, S. R. Golwala, J. Hall, H. R. Harris, S. A. Hertel, T. Hofer, D. Holmgren, L. Hsu, M. E. Huber, A. Jastram, O. Kamaev, B. Kara, M. H. Kelsey, A. Kennedy, M. Kiveni, K. Koch, A. Leder, B. Loer, E. Lopez Asamar, R. Mahapatra, V. Mandic, C. Martinez, K. A. McCarthy, N. Mirabolfathi, R. A. Moffatt, D. C. Moore, H. Nelson, R. H. Nelson, R. W. Ogburn, K. Page, W. A. Page, R. Partridge, M. Pepin, A. Phipps, K. Prasad, M. Pyle, H. Qiu, W. Rau, P. Redl, A. Reisetter, Y. Ricci, H. E. Rogers, T. Saab, B. Sadoulet, J. Sander, K. Schneck, R. W. Schnee, S. Scorza, B. Serfass, B. Shank, D. Speller, S. Upadhyayula, A. N. Villano, B. Welliver, D. H. Wright, S. Yellin, J. J. Yen, B. A. Young, J. Zhang, and CDMS Collabora-

tion. First direct limits on Lightly Ionizing Particles with electric charge less than $e/6$. *Phys. Rev. Lett.*, 114(11):111302, 2015.

[9] J. Billard, L. E. Strigari, and E. Figueroa-Feliciano. Solar neutrino physics with low-threshold dark matter detectors. *Phys. Rev.*, D91(9):095023, May 2015.

[10] F. Ruppin, J. Billard, E. Figueroa-Feliciano, and L. Strigari. Complementarity of dark matter detectors in light of the neutrino background. *Phys. Rev.*, D90(8):083510, 2014.

[11] R. Agnese, A. J. Anderson, M. Asai, D. Balakishiyeva, R. Basu Thakur, D. A. Bauer, J. Beaty, J. Billard, A. Borgland, M. A. Bowles, D. Brandt, P. L. Brink, R. Bunker, B. Cabrera, D. O. Caldwell, D. G. Cerdeno, H. Chagani, Y. Chen, M. Cherry, J. Cooley, B. Cornell, C. H. Crewdson, P. Cushman, M. Daal, D. DeVaney, P. C. F. Di Stefano, E. Do Couto E Silva, T. Doughty, L. Esteban, S. Fallows, E. Figueroa-Feliciano, G. L. Godfrey, S. R. Golwala, J. Hall, S. Hansen, H. R. Harris, S. A. Hertel, B. A. Hines, T. Hofer, D. Holmgren, L. Hsu, M. E. Huber, A. Jastram, O. Kamaev, B. Kara, M. H. Kelsey, S. Kenany, A. Kennedy, M. Kiveni, K. Koch, A. Leder, B. Loer, E. Lopez Asamar, R. Mahapatra, V. Mandic, C. Martinez, K. A. McCarthy, N. Mirabolfathi, R. A. Moffatt, R. H. Nelson, L. Novak, K. Page, R. Partridge, M. Pepin, A. Phipps, M. Platt, K. Prasad, M. Pyle, H. Qiu, W. Rau, P. Redl, A. Reisetter, R. W. Resch, Y. Ricci, M. Ruschman, T. Saab, B. Sadoulet, J. Sander, R. L. Schmitt, K. Schneck, R. W. Schnee, S. Scorza, D. N. Seitz, B. Serfass, B. Shank, D. Speller, A. Tomada, S. Upadhyayula, A. N. Villano, B. Welliver, D. H. Wright, S. Yellin, J. J. Yen, B. A. Young, and J. Zhang. Search for low-mass weakly interacting massive particles with supercdms. *Phys. Rev. Lett.*, 112:241302, Jun 2014.

[12] Cdms publications: http://cdms.berkeley.edu/publications.html.

[13] R Agnese, Z Ahmed, A J Anderson, S Arrenberg, D Balakishiyeva, R Basu Thakur, D A Bauer, J Billard, A Borgland, D Brandt, P L Brink, T Bruch, R Bunker, B Cabrera, D O Caldwell, D G Cerdeno, H Chagani, J Cooley, B Cornell, C H Crewdson, P Cushman, M Daal, F Dejongh, E do Couto e Silva, T Doughty, L Esteban, S Fallows, E Figueroa-Feliciano, J Filippini, J Fox, M Fritts, G L Godfrey, S R Golwala, J Hall, R H Harris, S A Hertel, T Hofer, D Holmgren, L Hsu, M E Huber, A Jastram, O Kamaev, B Kara, M H Kelsey, A Kennedy, P Kim, M Kiveni, K Koch, M Kos, S W Leman, B Loer, E Lopez Asamar, R Mahapatra, V Mandic, C Martinez, K A McCarthy, N Mirabolfathi, R A Moffatt, D C Moore, P Nadeau, R H Nelson, K Page, R Partridge, M Pepin, A Phipps, K Prasad, M Pyle, H Qiu, W Rau, P Redl, A Reisetter, Y Ricci, T Saab, B Sadoulet, J Sander, K Schneck, R W Schnee, S Scorza, B Serfass, B Shank, D Speller, K M Sundqvist, A N Villano, B Welliver, D H Wright, S Yellin, J J Yen, J Yoo, B A Young, and J Zhang. Silicon Detector Dark Matter Results from the Final Exposure of CDMS II. *Physical Review Letters*, 111(2):251301, December 2013.

[14] R Agnese, A J Anderson, D Balakishiyeva, R Basu Thakur, D A Bauer, A Borgland, D Brandt, P L Brink, R Bunker, B Cabrera, D O Caldwell, D G Cerdeno, H Chagani, M Cherry, J Cooley, B Cornell, C H Crewdson, P Cushman, M Daal, P C F Di Stefano, E do Couto e Silva, T Doughty, L Esteban, S Fallows, E Figueroa-Feliciano, J Fox, M Fritts, G L Godfrey, S R Golwala, J Hall, H R Harris, J Hasi, S A Hertel, B A Hines, T Hofer, D Holmgren, L Hsu, M E Huber, A Jastram, O Kamaev, B Kara, M H Kelsey, S A Kenany, A Kennedy, C J Kenney, M Kiveni, K Koch, B Loer, E Lopez Asamar, R Mahapatra, V Mandic, C Martinez, K A McCarthy, N Mirabolfathi, R A Moffatt, D C Moore, P Nadeau, R H Nelson, L Novak, K Page, R Partridge, M Pepin, A Phipps, K Prasad, M Pyle, H Qiu, R Radpour, W Rau, P Redl,

A Reisetter, R W Resch, Y Ricci, T Saab, B Sadoulet, J Sander, R Schmitt, K Schneck, R W Schnee, S Scorza, D Seitz, B Serfass, B Shank, D Speller, A Tomada, A N Villano, B Welliver, D H Wright, S Yellin, J J Yen, B A Young, J Zhang, and The SuperCDMS Collaboration. Demonstration of surface electron rejection with interleaved germanium detectors for dark matter searches. *Appl. Phys. Lett.*, 103(16):164105, 2013.

[15] R Agnese, A J Anderson, M Asai, D Balakishiyeva, R Basu Thakur, D A Bauer, J Billard, A Borgland, M A Bowles, D Brandt, P L Brink, R Bunker, B Cabrera, D O Caldwell, D G Cerdeno, H Chagani, J Cooley, B Cornell, C H Crewdson, P Cushman, M Daal, P C F Di Stefano, T Doughty, L Esteban, S Fallows, E Figueroa-Feliciano, G L Godfrey, S R Golwala, J Hall, H R Harris, S A Hertel, T Hofer, D Holmgren, L Hsu, M E Huber, A Jastram, O Kamaev, B Kara, M H Kelsey, A Kennedy, M Kiveni, K Koch, B Loer, E Lopez Asamar, R Mahapatra, V Mandic, C Martinez, K A McCarthy, N Mirabolfathi, R A Moffatt, D C Moore, P Nadeau, R H Nelson, K Page, R Partridge, M Pepin, A Phipps, K Prasad, M Pyle, H Qiu, W Rau, P Redl, A Reisetter, Y Ricci, T Saab, B Sadoulet, J Sander, K Schneck, R W Schnee, S Scorza, B Serfass, B Shank, D Speller, A N Villano, B Welliver, D H Wright, S Yellin, J J Yen, B A Young, and J Zhang. CDMSlite: A Search for Low-Mass WIMPs using Voltage-Assisted Calorimetric Ionization Detection in the SuperCDMS Experiment. *Physical Review Letters*, 112(4):041302, September 2013.