Computational Resources—*Using Generalized Adversarial Networks to ensure trusted science results and maximize science reach within the dark matter community*

1 Computational resource estimates

1.1 CPU

The following work will require CPU cycles:

- Validating generated data
- Reproducing dark matter analyses with generated data

Validating generated data is a "light" analysis task and we estimate 5000 SUs for this task.

Analysis reproduction consists of analyzing existing code, producing new code that inserts the generated data into the data sets, and re-running the original analysis code. Of these tasks, only re-running the original analysis code consumes significant CPU. We estimate 10,000 SUs for this task.

1.2 GPU

Training the Generative Adversarial Networks requires GPU resources. We estimate a complete training cycle to consume 10,000 SUs of GPU and expect to complete many training cycles, on order of 1000, over the course of the project as we test trained networks and improve them.

We therefore estimate our GPU requirements to be 10 MSUs.

1.3 Storage

Training data drives the storage requirements for this project; training data sets consist of existing CDMS and SuperCDMS experimental data and are no larger than 10 TB. Storage for this data is already provided by the collaboration and therefore storage resources are not discussed below.

Additional storage is needed for

- Intermediate files produced by re-analysis efforts (100 GB)
- Generated data (100 GB)
- Software and documentation (1 GB)

These additional storage requirements are modest and are met by the default allocations on the systems described below.

2 Meeting project needs

2.1 XSEDE

An XSEDE allocation for CPU (100 MSUs) and GPU (20 MSUs) resources has been submitted for the upcoming year.

This request is a submission to a competitive program, is a yearly submission, and is not guaranteed. Below are two additional resource plans that will meet the project needs if this resource request is denied.

2.2 SuperCDMS collaboration resources

The work proposed focuses on SuperCDMS data and, if successful, provides direct benefit to the SuperCDMS collaboration. This work is therefore in scope for SuperCDMS science operations and is eligible to use SuperCDMS computing resources. SuperCDMS has GPU allocations at Compute Canada and has CPU allocations at SLAC, Compuate Canada, A&M, and four additional sites. In addition, SuperCDMS has a storage allocation of 140 TB through the Open Storage Network.

The CPU and storage allocations are more than sufficient to meet the needs of the proposed work. To meet

our GPU requirements we would also use local resources, detailed below.

2.3 Local resources

Resources available for immediate prototyping are (1) CPU, GPU, and Storage provided by Co-PI Banaei-Kashani's Big Data Lab and (2) Alderaan, the campus cluster.

Both these resources will allow progress on prototyping and production of the proposed work.