

Thesis abstract: Online processing of the large area SiPM detector signals for the DarkSide20k experiment

Giacomo Petrillo

Supervisors: Eugenio Paoloni, Simone Stracka
University of Pisa

June 1, 2021

DarkSide20k is a planned dual-phase liquid argon (LAr) time projection chamber (TPC) designed to detect dark matter, the successor to DarkSide-50. It will be the largest detector of its kind, with 20 metric tons of argon in the fiducial volume. The predicted resulting upper bound on the spin-independent WIMP-nucleon scattering cross-section, in case of no discovery, is $\approx 1 \times 10^{-47} \text{ cm}^2$ at $1 \text{ TeV}/c^2$ WIMP mass, to be compared with the current best limit $\approx 1 \times 10^{-45} \text{ cm}^2$ by XENON1T.

In this thesis we present reconstruction and characterization studies on the photodetector modules (PDMs) that will be used in the TPC. These studies are primarily meant as a support to the definition of the first stages of the online processing chain.

Each PDM has a 25 cm^2 matrix of silicon photomultipliers (SiPMs), instead of the usual photomultiplier tube (PMT). The SiPM has Geiger-mode single photon response, i.e., each detected photon produces one fixed amplitude pulse. Compared to a PMT, the photodetection efficiency is expected to be higher, reaching 50 % at room temperature, with also a better filling factor, which should be greater than 85 % in DarkSide20k, compared to 80 % in DarkSide-50, and a much better single photon resolution. The pulse looks like a sharp peak followed by a rather long exponential tail, which is a disadvantage because pulses can pile-up leading to saturation.

SiPMs have three kinds of noise: 1) stationary electric noise, which scales with the square root of the area; 2) a “dark count rate” (DCR) of pulses independent of incident light that scales with the area; 3) “correlated noise” produced by primary pulses, which contributes a factor proportional to the DCR and photon pulses.

The first two stages in the readout chain will be the digitizers and the front end processors (FEP). The digitizers find candidate pulses, and for each one send a slice of waveform to the FEP, where the final identification of pulses is decided. The performance of these stages is mainly determined by the electric noise, characterized with the signal to noise ratio (SNR), which is the ratio of the amplitude of pulses over the noise standard deviation. It influences the fake rate, i.e., the rate of random oscillations high enough to be mistakenly identified as pulses, and the temporal resolution of pulse detection.

By applying linear filters to digitized waveforms acquired from the PDMs illuminated by a pulsed laser, both in a testing setup at Laboratori Nazionali del Gran Sasso (LNGS) and in the small prototype TPC “Proto0”, we study the noise parameters of single pulse detection: SNR, temporal resolution, fake rate.

We consider 1) an autoregressive filter, which uses the least possible computational resources, 2) a matched filter without spectrum correction, which gives almost optimal performance, 3) a moving average, which is a compromise between simplicity and performance. Simple filters are needed on the digitizers, which must process all the incoming data, while the FEP will probably use the optimal filter. We also study the baseline computation and the filter length.

Then using a custom peak finder algorithm we measure the DCR and study the correlated noise, which consists in additional pulses produced recursively by each pulse, divided in two main categories: afterpulses (AP), which arrive with some delay from the parent pulse, and have smaller amplitude as the delay goes to zero, and direct cross talk (DiCT), which manifests as a integer multiplication of the amplitude of pulses because the children pulses are overlapped with the parent.

The results follow.

While with an ideal filtering procedure the post-filter SNR reaches 20, it may realistically be 13 with the resources of the digitizers, i.e., using a moving average $1\text{ }\mu\text{s}$ long for both the pulse and the baseline.

With the moving average as just described, the fake rate is 10 cps when the threshold is set to 5 standard deviations of the filtered noise, where the filter includes the subtraction of the baseline. Since the filtering procedure is not actually decided yet, we describe and test a general procedure to measure low fake rates without actually counting the threshold crossings with only 1 ms of recorded data. Of the 25 PDMs we looked at, all are within specifications but a single one which has an anomalous fake rate whose origin is still under investigation.

The temporal resolution mostly matters on the FEP, so we summarize the results with the matched filter: 1) upsampling is not necessary; 2) at low SNR the resolution diverges and how fast heavily depends on the noise spectrum, with the Proto0 noise the maximum allowed by specs, 10 ns, is reached at pre-filter SNR 2.6; 3) it is sufficient to have $1\text{ }\mu\text{s}$ of waveform per pulse sent to the FEP; 4) it is possible to lower the sampling frequency from 125 MSa/s to 62.5 MSa/s. In the thesis we show detailed curves for all the filter, filter length, SNR and sampling frequency choices.

We give upper bounds for the DCR of a 25 cm^2 SiPMs Tile at three overvoltages, 5.5 V, 7.5 V, 9.5 V (the overvoltage is the difference between the bias put on the SiPM and the breakdown voltage of the junction), which are respectively 50 cps, 170 cps, and 120 cps, to be compared with the DarkSide20k requirement of 250 cps. 5.5 V is a somewhat usual operating overvoltage while 9.5 V is considered high. Increasing the overvoltage increases both the SNR, the DCR and the correlated noise.

The analysis of correlated noise, done on the same data, gives upper bounds for AP probabilities of 2.5 %, 3.5 % and 6.5 %, and DiCT probabilities 20 %, 30 % and 50 %. These are the probabilities of said noises being generated by any given single pulse, i.e., the stacked pulses produced by DiCT count separately. The DarkSide20k specifications require the sum of correlated noise probabilities to be less than 60 % in order to avoid performance degradation, e.g., dynamic range reduction on ionization signals, where there is a lot of pile-up. For this analysis we employ the common procedures used to fit histograms with least squares, which we rederive in the Bayesian framework in an appendix.

Measuring AP and DiCT requires models. We try the models we find in the literature and in the DarkSide20k simulation and bring them into question, but we do not search for better alternatives, since their level of

accuracy should be enough for the simulation requirements. We find that the AP temporal distribution is well described by two exponential decays with constants 200 ns and 1 μ s, but not by a single one.

Finally, based on the peak finding algorithm we used in the correlated noise analysis, we suggest but do not test the following procedure to better resolve multiple pulses: do a first pass with a short filter, pick candidate peaks, do a second pass with a long filter, eventually pick additional candidates, compute the amplitude of pulses solving the linear system for the superposition of pulses using only the long filter peak amplitudes, curb peaks with low amplitude and compute it again. We sketch an argument to justify this procedure.