

---

# Nonlinear pile-up separation with LSTM neural networks for cryogenic particle detectors

---

**Felix Wagner**

Institute of High Energy Physics  
Austrian Academy of Sciences  
1050 Vienna, Austria  
felix.wagner@oeaw.ac.at

## Abstract

In high-background or calibration measurements with cryogenic particle detectors, a significant share of the exposure is lost due to pile-up of recoil events. We propose a method for the separation of pile-up events with an LSTM neural network and evaluate its performance on an exemplary data set. Despite a non-linear detector response function, we can reconstruct the ground truth of a severely distorted energy spectrum reasonably well.

## 1 Introduction

In modern experimental particle physics, some of the biggest open questions, such as the search for lightweight dark matter and coherent neutrino nucleus scattering, require the measurement of low energy particle scattering [1, 2]. Sensitivity to these interactions imposes certain requirements: A low energy threshold for nuclear recoils, low background, a high discrimination power against remaining background and artifacts and a sensible energy calibration down to threshold. Cryogenic detectors, as used by the CRESST, COSINUS and NUCLEUS experiments [3–5], meet these requirements, which we shall describe in Sec. 2.

In a measurement, the recoil of a particle within a cryogenic detector causes a pulse-shaped heat signal, which is then triggered and stored within a record window for further analysis. In above-ground and calibration runs, we observe high trigger rates, leading to a large amount of pile up events, i.e. heat signatures of several events within the same record window. If not treated carefully, these events distort the reconstructed recoil energy spectrum significantly. As an alternative to identifying and removing pile up events, we present in our work a method to reconstruct the individual pulse traces and corresponding recoil energies, and by that make use of the measurement’s full exposure. The separation is especially challenging if non-linearities in the detector’s response function are present, which is usually handled in an individual step of the analysis [6]. However, with an LSTM neural network [7] trained on a simulated data set, we combined these steps into one: We train the model to reconstruct the unsaturated and separated pulse traces from saturated pile-up event. Our results are described in Sec. 3, a discussion of their implications and future plans are given in Sec. 4. In this work, we used the PyTorch, PyTorch Lightning and Cait Python packages [8–10].

**Related work.** Machine learning models in rare event search experiments have been mostly applied for the rejection of artifacts and background [11, 12] and the fine-tuning of acceptance regions [13]. Specifically the rejection of pile up pulses was addressed in Ref. [14] and [15]. There are non-machine learning methods for pile up separation, e.g. in Ref. [16]. Contrary to these, we reconstruct the whole pulse trace of the individual events, with significantly faster run time than fitting methods would require, and at the same time treat the nonlinear detector response. We are not aware of any method that reports similar results for the task of nonlinear pile up separation, with comparable inference time.

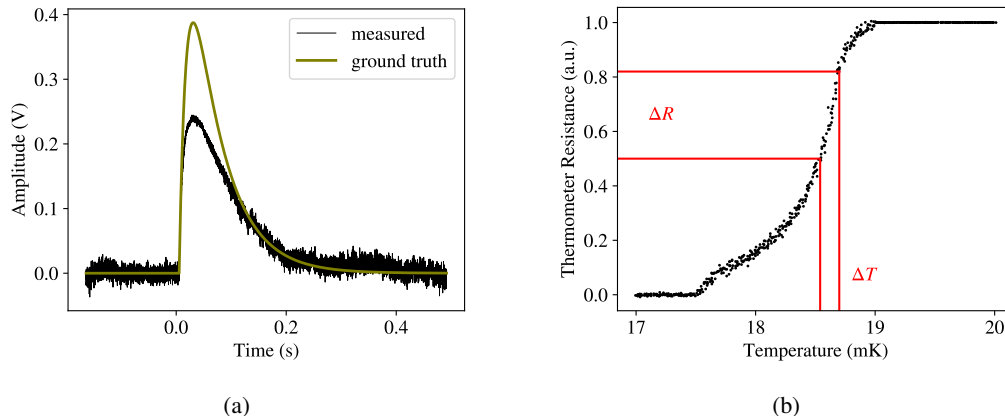


Figure 1: The measured data consists of voltage traces, representing the temperature in the superconducting thermometer. (a) During the measurement process, the pulse-shaped temperature increase in the thermometer (olive) is distorted by additive noise and a non-linear flattening (saturation) of its peak (black). (b) The transition curve of the superconducting thermometer. Operated in between the normal- and superconducting state, a small temperature increase  $\Delta T$  causes a measurable increase in resistance  $\Delta R$ .

## 2 Detector concept

Cryogenic detectors are small crystals with a mass of  $\mathcal{O}(1 - 100)$  g, operated at temperatures below 50 mK. At these low temperatures, an individual particle recoil inside the detector material produces a detectable pulse-shaped heat signal (see Fig. 1 (a), olive). The crystal is equipped with a sensitive thermometer. In our work we consider a transition edge sensor (TES), i.e. a superconducting film evaporated onto the crystal. The temperature of the superconductor is fine tuned, such that the thermometer is in transition from its normal conducting to its superconducting phase (see Fig. 1 (b)). This way, the heat signal of the particle recoil causes a measurable increase in the thermometers resistance. For low energy recoils, the transition curve of the thermometer is well approximated with a linear function, i.e. the pulse-shaped heat signature is translated to a pulse-shape voltage signal, only distorted by additive noise. In this regime, the height of the pulse is directly proportional to the energy of the particle recoil. However, for particle recoils with higher energy, the normal-conducting (saturated) region of the thermometer is reached, which we observe as a non-linear flattening of the pulse's peak (see Fig. 1 (a), black). Throughout a measurement, the shape of the transition curve is continuously measured with injected test signals, making the reconstruction of saturated pulse heights possible. Typical materials used for the detector crystal are  $\text{CaWO}_4$ ,  $\text{LiO}_2$ ,  $\text{Al}_2\text{O}_3$ , Si and NaI [1].

### 2.1 Used data

For TES-based detectors, an accurate, parametric model of the response function was derived in Ref. [17]. In Ref. [18], a method for the simulation of noise traces with a given noise power spectrum (NPS) is introduced. In Ref. [6], the upper end of the detectors saturation curve was approximated with an appropriate fit function. We combine these methods into a simulation of realistic recoil events, which we use exclusively throughout this work. Based on previous work, we expect a good transfer of the results to measured data [19–21], see also Sec. 4.

We simulate noise traces with a record length of 16384 samples, and a sampling frequency of 25 kHz. The pulse shape is chosen to have a significant signal length of 150 ms and is superposed twice to the noise trace, to augment the desired pile up events (see Fig. 2). The onset of the first pulse is located in close proximity to the typical trigger position at one fourth of the record window. The second pulse is located on the tail of the first pulse, uniformly distributed between the trigger position and the end of the record window. Both pulse heights are uniformly distributed between 0 and 0.5 V. Finally, the thermometers saturation curve is applied to the voltage trace, flattening the peaks of pulses above 0.25 V.

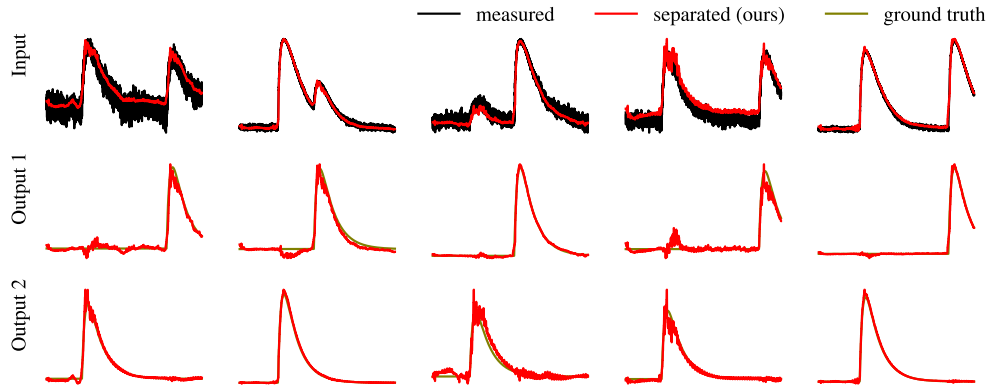


Figure 2: The record windows of the recorded particle recoils. The five columns correspond to five independent samples in the data set. The first row shows the time series which are input to the LSTM (black) and the summed outputs after application of the non-linear transition function (red). We observe a good agreement in most samples, with minor distortions for events with low signal-to-noise ratio. The second and third row show the two output traces of the LSTM (red), which are mostly in good agreement with the ground truth, plotted underneath (olive).

The details of the simulation process, including all used parameters, are described in App. A. The simulated data set will be made publicly available with the deanonymized version of this work.

### 3 Results

We use a data set of 60k samples, split into 42k training, 6k validation and 12k test samples.<sup>1</sup> The realistic pulse trace (Fig. 2, black) is the input to our supervised model, the two separate, unsaturated temperature pulses (Fig. 2, olive) are the training objective. As model, we use a 3-layer LSTM with an input size of 8 values per time step and 200 nodes in each hidden layer. The last layer’s hidden states of each time step are linearly mapped to two traces of 8 values each, corresponding to the time step of the two separated temperature pulses.

In a preprocessing step, we perform two transformations on the data. First, all samples are min-max normalized to a value range from 0 to 1. Second, each trace is downsampled by a factor of 32, which leaves to a record length of 512 values. A method, originally described in Ref. [22], that uses the distribution of the maxima of empty noise traces to determine the number of independent samples within the noise of a record window, leads to approximately the same number (512) of independent samples. This motivates the assumption that with our choice of downsampling, no significant information is lost. With 8 values as simultaneous input to the LSTM, one record window corresponds to 64 time steps.

We train our model with the averaged mean squared error (L2) between outputs and ground truth as loss function. The Gaussian distribution of the additive detector noise speaks in favor of a loss function sensitive to outliers (as the L2), and our decision for this loss function was experimentally confirmed: Others, e.g. the L1 loss function, were explored and brought similar or worse results. The ADAM optimizer with a learning rate of  $1e-5$  and a batch size of 32 is used, training proceeds for 100 epochs. The best validation score was reached after epoch 65, although already after 15 epochs no further significant improvement on the validation loss score is visible. The final scores were  $2.4 \cdot 10^{-4}$  on the test set, and  $1.9 \cdot 10^{-4}$  on the training set, which shows no significant overfitting on the training data. The manageable size of the trained model (3.3 MB), together with its fast inference time (Tesla P100 GPU with 12 Gig RAM / Intel(R) Xeon(R) Gold 6138 CPU @ 2.00GHz  $\sim 0.35 / 5$  minutes per training epoch) makes the model suitable for smaller machines. By choosing a reasonable inference batch size of e.g. 1000, this corresponds to the separation of  $>3000$  events / second on a single-core CPU.

<sup>1</sup>For training set sizes  $> 10k$  samples, no significant improvement of the results with larger training sets was observed.

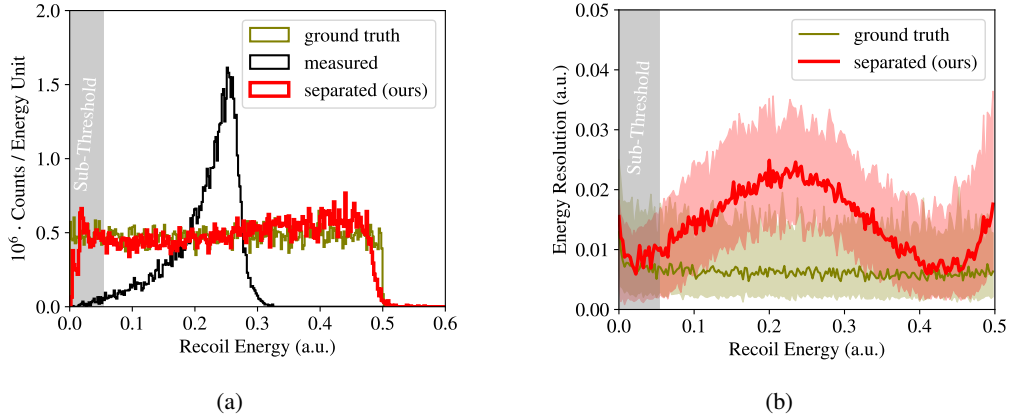


Figure 3: The results of the pulse height reconstruction on an independent test set. (a) The black histogram represents the maximal values of the piled up and saturated traces. These strongly deviate from the uniformly sampled ground truth of the appearing pulse heights (two per record window, olive). We observe a much better agreement of the maximal values of the separated pulse traces (red), with only slight deviations for strongly saturated pulse shapes, close to the upper end of the spectrum. (b) The energy resolution of pulses without pile up and saturation does not depend on the recoil energy, but only on the noise conditions (olive). The energy resolution of the reconstructed pulses does depend on the energy, with local maxima at 0.25 and 0.5, and local minima at 0 and 0.45. This peculiar shape is discussed in Sec. 3.

The separation and reconstruction of the individual pulse traces works reasonably well, as shown in Fig. 2. Minor distortions for pulses with low signal-to-noise ratio are visible, as well as residuals on the separated traces at the position of the first pulse. The impact of the distortions on the energy reconstruction of individual events depends on the choice of the pulse height estimator. While it likely impacts the maximum of the array, this effect might be reduced by commonly used fit- and matched filter methods. In this work, we use the maximum of the array as estimator, to achieve a conservative evaluation of our methods practicability. Therefore we reconstruct as a metric the pulse height spectrum, i.e. the recoil energy spectrum, of the separated pulses from the test set (LSTM predictions, Fig. 3 (a), red) and compare it with the spectrum of the simulated, individual pulses (ground truth, Fig. 3 (a), olive) and the piled-up and saturated spectrum (LSTM input, Fig. 3 (a), black). Additionally, we measure the 0.16, 0.5 and 0.84 quantiles (corresponding to the central value and a one-sigma band in the Gaussian limit) of the deviation of our reconstructed pulse heights from the ground truth value, also called the energy resolution (see Fig. 3 (b), red). We compare this value with an equivalently calculated value on a similar data set, without pile-up and saturation effects, acting as ground truth (Fig. 3 (b), olive).

For both metrics, we reach reasonable agreement of our separated pulse traces with the ground truth: The strongest relative count rate deviation in the separated recoil spectrum from the ground truth (Fig. 3 (a), red and olive) is  $\sim 20\%$  at 0.45, while it is in the measured spectrum  $> 300\%$  at 0.25. The strongest relative deviation in the energy resolution from the ground truth (Fig. 3 (b), red and olive) is a factor 3 at 0.22, corresponding to a relative error in the reconstruction of individual recoil energies of 10% (LSTM predictions) instead of 3.5% (ground truth). At threshold, the energy resolution of the LSTM predictions is roughly 25% higher than that of the ground truth, which is proportional to the necessary increase of the threshold value, to achieve the same noise trigger rate. As expected, the quality of the reconstruction decreases with rising pulse heights, due to the saturation. The observed minimum in the energy resolution at 0.45 (Fig. 3 (b), red) is a matter of ongoing investigation. One, but likely not the only, contributing effect is of statistical nature and common to all plots of prediction error w.r.t. ground truth: The negative bias of pulses at the upper end of the simulated spectrum (see Fig. 3 (a), right end of red curve) causes an overdensity of predictions around 0.45. Therefore, events with a true energy around 0.45 feature a lower error, simply because the model predicts these values more often. Overall, the achieved energy resolution is good enough to reconstruct the position of peaks (e.g. for energy calibration) in the spectrum: A Gaussian peak in the ground truth at 0.22

would lead to a peak in the reconstructed spectrum at  $0.22 \pm 0.022/\sqrt{N}$ , where  $N$  is the number of events contributing to the peak. The number of events in a calibration peak is typically high ( $\gg 1000$ ), leading to a small systematic error in the reconstruction of the peaks position. Therefore, our methods makes it possible to extract the information from physics measurements, that otherwise would need to be discarded. We lack the comparison with a state-of-the-art model, as for our application no state-of-the-art has emerged yet.

Finally we want to address the scenario, where the model is applied to a record window with only one pulse, i.e. a regular event. The static architecture of the LSTM model creates, also in this case, two output traces, which are then both interpreted as events. Because the model does not include an automatic discrimination mechanism between regular and pile-up events, one of the output traces creates a noise trigger event that does not correspond to an actual particle recoil. However, in the analysis process usually a lower threshold is imposed on the accepted events, depending on the energy resolution of the detector (Fig. 3 gray area). By accepting only traces with events above threshold, we can reject all virtual events that appeared in the separation process due to the static number or two output traces. Therefore, the model can be applied to the full data set, without prior discrimination between regular and pile-up events, and without creating additional noise trigger events.

Our trained model and code is public on GitHub [23].

## 4 Discussion and outlook

As shown in Sec. 3, our model performs reasonably well in the task of reconstructing recoil energy spectra from piled-up and saturated measurements, after training on simulated data. For any measurement with a cryogenic detector, the extraction of all parameters necessary for a realistic data simulation is a straight forward procedure. This renders the model useful for practical applications in the current state already. The fast inference time of our model enables real-time inference (processing  $>3000$  event/second vs typical trigger rates  $\sim 1$  Hz). However, the choice of a min/max normalization would in this case necessarily be replaced with e.g. a batch normalization layer, to account for the previously unknown min/max values in a non-static data set. There are no publicly available raw data sets of experimentally measured recoil events in cryogenic detectors. Therefore, an evaluation of our model on measured data is still work in progress, in cooperation with the CRESST and COSINUS collaborations. A publication of these results is anticipated. Also, we plan on extending the method for treatment of pile-up with more than two pulse shapes, and pile-up with detector artifacts within the same data set.

## 5 Broader impact

Apart from the previously mentioned experiments CRESST, COSINUS and NUCLEUS, many more rare event searches use cryogenic detectors with pulse-shaped time series data, e.g. the CUORE [24], CUPID [25], SuperCDMS [26], EDELWEISS [27] and RICOCHET [28] experiments. Our method can impact the reconstruction of energy spectra from above-ground and calibration runs, potentially decreasing the uncertainty in the energy scales of mentioned experiments. Furthermore, our method can be adapted to separate sensor signals in all application areas of electrical engineering.

## Acknowledgments and Disclosure of Funding

We thank Wolfgang Waltenberger and Florian Reindl for fruitful discussions and advice on cryogenic experiments and machine learning in particle physics. We thank the Institute of High Energy Physics, the Austrian Academy of Sciences and the Austrian Research Promotion Agency for their support. The computational results presented were obtained using the CLIP cluster. We also thank the reviewers of the Machine Learning and the Physical Sciences workshop at NeurIPS21 for their feedback, which helped significantly in shaping this project.

## References

- [1] Julien Billard et al. Direct detection of dark matter – appec committee report, 2021.

- [2] Gail McLaughlin. Theory and phenomenology of coherent neutrino-nucleus scattering. *AIP Conference Proceedings*, 1666(1):160001, 2015. doi: 10.1063/1.4915590. URL <https://aip.scitation.org/doi/abs/10.1063/1.4915590>.
- [3] A. H. Abdelhameed et al. First results from the cresst-iii low-mass dark matter program. *Phys. Rev. D*, 100:102002, Nov 2019. doi: 10.1103/PhysRevD.100.102002. URL <https://link.aps.org/doi/10.1103/PhysRevD.100.102002>.
- [4] F. Reindl et al. Results of the first NaI scintillating calorimeter prototypes by COSINUS. *Journal of Physics: Conference Series*, 1342:012099, jan 2020. doi: 10.1088/1742-6596/1342/1/012099. URL <https://doi.org/10.1088/1742-6596/1342/1/012099>.
- [5] J. Rothe et al. NUCLEUS: Exploring Coherent Neutrino-Nucleus Scattering with Cryogenic Detectors. *Journal of Low Temperature Physics*, 199(1):433–440, April 2020. ISSN 1573-7357. doi: 10.1007/s10909-019-02283-7. URL <https://doi.org/10.1007/s10909-019-02283-7>.
- [6] Martin Stahlberg. *Probing low-mass dark matter with CRESST-III : data analysis and first results*. PhD thesis, 2020. URL <https://repositum.tuwien.at/handle/20.500.12708/16833>.
- [7] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997.
- [8] Adam Paszke et al. Pytorch: An imperative style, high-performance deep learning library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc., 2019. URL <http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>.
- [9] William Falcon et al. Pytorch lightning. *GitHub. Note:* <https://github.com/PyTorchLightning/pytorch-lightning>, 3, 2019.
- [10] Felix et al. Wagner. fewwagner/cait: v1.0.0, July 2021. URL <https://doi.org/10.5281/zenodo.5091416>.
- [11] Artem Golovatiuk et al. Deep learning for directional dark matter search. *Journal of Physics: Conference Series*, 1525:012108, apr 2020. doi: 10.1088/1742-6596/1525/1/012108. URL <https://doi.org/10.1088/1742-6596/1525/1/012108>.
- [12] I. Coarasa et al. Machine-learning techniques applied to three-year exposure of anais-112. 17th International Conference on Topics in Astroparticle and Underground Physics (TAUP 2021), 2021. URL <https://indico.ific.uv.es/event/6178/contributions/15858/>.
- [13] E. Armengaud et al. Constraints on low-mass WIMPs from the EDELWEISS-III dark matter search. *Journal of Cosmology and Astroparticle Physics*, 2016(05):019–019, may 2016. doi: 10.1088/1475-7516/2016/05/019. URL <https://doi.org/10.1088/1475-7516/2016/05/019>.
- [14] A. Armatol et al. Novel technique for the study of pileup events in cryogenic bolometers. *Phys. Rev. C*, 104:015501, Jul 2021. doi: 10.1103/PhysRevC.104.015501. URL <https://link.aps.org/doi/10.1103/PhysRevC.104.015501>.
- [15] R. Huang et al. Pulse shape discrimination in CUPID-mo using principal component analysis. *Journal of Instrumentation*, 16(03):P03032, mar 2021. doi: 10.1088/1748-0221/16/03/p03032. URL <https://doi.org/10.1088/1748-0221/16/03/p03032>.
- [16] M. Preston et al. A feature-extraction and pile-up reconstruction algorithm for the forward-spectrometer emc of the panda experiment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1011:165601, 2021. ISSN 0168-9002. doi: <https://doi.org/10.1016/j.nima.2021.165601>. URL <https://www.sciencedirect.com/science/article/pii/S0168900221005866>.

- [17] F. Pröbst et al. Model for cryogenic particle detectors with superconducting phase transition thermometers. *Journal of Low Temperature Physics*, 100(1):69–104, July 1995. ISSN 1573-7357. doi: 10.1007/BF00753837. URL <https://doi.org/10.1007/BF00753837>.
- [18] M. Carrettoni and O. Cremonesi. Generation of noise time series with arbitrary power spectrum. *Computer Physics Communications*, 181(12):1982–1985, 2010. ISSN 0010-4655. doi: <https://doi.org/10.1016/j.cpc.2010.09.003>. URL <https://www.sciencedirect.com/science/article/pii/S0010465510003486>.
- [19] Andreas Josef Zöller. *Artificial Neural Network Based Pulse-Shape Analysis for Cryogenic Detectors Operated in CRESST-II*. PhD thesis, Munich, Tech. U., 2016. URL <http://mediatum.ub.tum.de/?id=1303343>.
- [20] Christoph Mühlmann. Pulse-shape discrimination with deep learning in cressst, 2019. URL <https://repositum.tuwien.at/handle/20.500.12708/14865>.
- [21] Felix Wagner. Machine learning methods for the raw data analysis of cryogenic dark matter experiments, 2020. URL <https://repositum.tuwien.at/handle/20.500.12708/15023>.
- [22] M. Mancuso et al. A method to define the energy threshold depending on noise level for rare event searches. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 940:492–496, 2019. ISSN 0168-9002. doi: <https://doi.org/10.1016/j.nima.2019.06.030>. URL <https://www.sciencedirect.com/science/article/pii/S0168900219308708>.
- [23] Felix Wagner. Nonlinear pile-up separation with lstm neural networks for cryogenic particle detectors. *GitHub. Note: <https://github.com/fewagner/pileupsep>*, 2021.
- [24] D. Q. Adams et al. Improved limit on neutrinoless double-beta decay in  $^{130}\text{Te}$  with cuore. *Phys. Rev. Lett.*, 124:122501, Mar 2020. doi: 10.1103/PhysRevLett.124.122501. URL <https://link.aps.org/doi/10.1103/PhysRevLett.124.122501>.
- [25] E. Armengaud et al. The CUPID-Mo experiment for neutrinoless double-beta decay: performance and prospects. *The European Physical Journal C*, 80(1):44, January 2020. ISSN 1434-6052. doi: 10.1140/epjc/s10052-019-7578-6. URL <https://doi.org/10.1140/epjc/s10052-019-7578-6>.
- [26] I. Alkhatib et al. Light dark matter search with a high-resolution athermal phonon detector operated above ground. *Phys. Rev. Lett.*, 127:061801, Aug 2021. doi: 10.1103/PhysRevLett.127.061801. URL <https://link.aps.org/doi/10.1103/PhysRevLett.127.061801>.
- [27] E. Armengaud et al. Searching for low-mass dark matter particles with a massive ge bolometer operated above ground. *Phys. Rev. D*, 99:082003, Apr 2019. doi: 10.1103/PhysRevD.99.082003. URL <https://link.aps.org/doi/10.1103/PhysRevD.99.082003>.
- [28] J. Billard et al. Coherent neutrino scattering with low temperature bolometers at chooz reactor complex. *Journal of Physics G: Nuclear and Particle Physics*, 44(10):105101, aug 2017. doi: 10.1088/1361-6471/aa83d0. URL <https://doi.org/10.1088/1361-6471/aa83d0>.
- [29] Vanessa Zema. *Unveiling the nature of dark matter with direct detection experiments*. PhD thesis, Altra Università o Università estera, Chalmers U. Tech., 2020.

## A Simulation details

In Ref. [17], the detector is described as a thermal circuit, identifying the crystal phonons and the thermometer electrons as the main components (see Fig. 4 (a)), described by their temperature  $T_c, T_e$  and heat capacities  $C_c, C_e$ . Other parts are absorbed in the thermal resistances  $G_{cb}, G_{ce}, G_{eb}$ , the system is connected to a heat bath regulating the base temperature  $T_b$ . The temperature of the individual components can be calculated with a system of coupled, ordinary differential equations, depending on the power acting on the crystal and thermometer  $P_e, P_c$ :

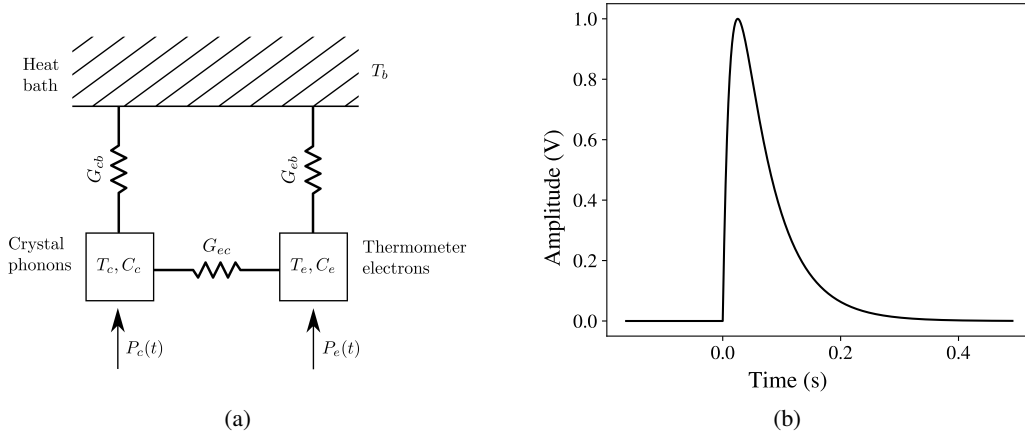


Figure 4: (a) Thermal resistance circuit of the detector: The temperature of the thermometer electrons  $T_e$  is the measured quantity, while the temperature of the crystal phonons  $T_c$  is a hidden state of the model. The temperature of the heat bath  $T_b$  is fixed. The component's heat capacities  $C_c$ ,  $C_e$  and the thermal conductivities  $G_{cb}$ ,  $G_{eb}$ ,  $G_{ec}$  govern the signal's pulse shape (b).

$$C_i \frac{dT_i}{dt} + G_{ib} (T_i - T_b) + \sum_j G_{ji} (T_j - T_i) = P_i \quad (1)$$

The indices  $i$  and  $j$  correspond to the components  $e$  and  $c$  respectively. A particle recoil inside the crystal produces a phonon population, a large share of which thermalizes through scattering on the crystals surface, while a small share thermalizes directly inside the thermometer. The waiting time for the thermalization of individual phonons is exponentially distributed, which leads to exponentially decaying power inputs in both the crystal and the thermometer. The subsequent heat transport from the crystal to the thermometer causes a second pulse component, introducing three time constants of the detectors response signal in total:  $\tau_{in}$ , the time until thermalization inside the thermometer,  $\tau_n$  the time needed for the heat transport from the crystal to the thermometer and  $\tau_t$ , the overall relaxation time of the system. Solving Eq. 1 with an Ansatz of an exponential decay for the power inputs leads to a solution for the heat increase in the thermometer, which describes the detectors response signal very well:

$$\Delta T_e(t) = \Theta(t) \left[ A_n \left( e^{-t/\tau_n} - e^{-t/\tau_{in}} \right) + A_t \left( e^{-t/\tau_t} - e^{-t/\tau_n} \right) \right] \quad (2)$$

The amplitudes of the individual signal components  $A_n$ ,  $A_t$ , as well as the time constants, depend on the thermal resistances and heat capacities. An exemplary pulse shape is shown in Fig. 4 (b). For a measurement done with cryogenic detectors, these parameters can be identified by a fit of Eq. 2 to a standard event, i.e. the superposition of several typical recoil events. Most recoil events follow this characteristic pulse shape, only for high energy events a nonlinear saturation effect occurs. Extensions of the pulse shape model, for detectors with more than two thermal components, are derived in Ref. [29].

The saturation depends on the superconductor's transition curve and was in Ref. [6] modeled with a generalised logistic function (see Fig. 5 (a)):

$$Y(t) = A + \frac{K - A}{(C + Qe^{-Bt})^{1/\nu}}, \quad (3)$$

where  $K$ ,  $C$ ,  $Q$ ,  $B$  and  $\nu$  are free fit parameters and  $A$  is derived from the condition  $Y(0) = 0$ . The variable  $t$  describes the true, unsaturated height of a signal pulse,  $Y$  describes the saturated pulse height.



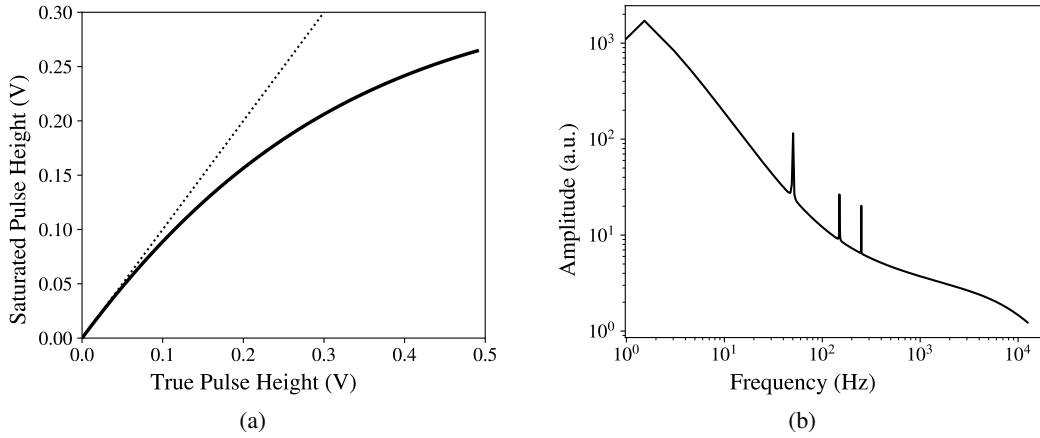


Figure 5: (a) The detector’s saturation causes a flattening of the peaks of signal pulses. This saturation can be described with a generalized logistics function (Eq. 3). The variable  $Y$  corresponds to the saturated pulse height,  $t$  to the true pulse height. (b) A simulated noise power spectrum, built from a superposition of two  $1/f^\alpha$  functions, with different parameters  $\alpha$  that describe the pink and white noise components. Disturbances from main voltage at 50, 100 and 150 Hz are added, the constant component at 0 Hz is set to zero. Frequencies above 10 kHz are cut with an anti-aliasing low pass filter.

For the simulated data we used throughout this work, we superpose simulated noise baselines and the pulse shape model, and apply the generalized logistic function point-wise.

The simulation of noise baselines follows the method proposed in Ref. [18], with an additionally added polynomial drift structure that often appears in measured noise traces. A noise power spectrum is built from a parametric model, reflecting the typical  $1/f^\alpha$  shape, the interference from mains voltage and an anti-aliasing low pass filter (see Fig. 5 (b)).

For the pulse shape model we used the parameters:

$$\tau_t = 0.05815608, \tau_{in} = 0.02059209, \tau_n = 0.01395427, A_n = 0.35, A_t = 1.97.$$

For the generalised logistic function we used the parameters:

$$K = 10, C = 3, Q = 0.5, B = 5, \nu = 0.5$$