Investigation of Chi-Squared Distribution Discrepancies Between CUTE Data and Simulation Data in SuperCDMS

Nick Wang, Department of Physics, University of Toronto
Supervisor: Professor Pekka Sinervo

December 5, 2025

Abstract

This report presents the findings of an investigation into chi-squared distribution discrepancies between experimental data from the Cryogenic Underground TEst facility (CUTE) and simulation data within the Super Cryogenic Dark Matter Search (SuperCDMS). The primary objective was to characterize and understand the "early rise problem", where chi-squared distributions computed from simulation data exhibit an earlier rise compared to CUTE data at energies above approximately 4 keV.

Initial visual inspection of K-shell average Phonon Total (PT) pulses in the time domain revealed differences in the decay tail between CUTE and simulation pulses, motivating a detailed standard deviation analysis in both the time and frequency domains. This analysis indicated that CUTE data exhibits greater pulse-to-pulse variation and higher baseline noise levels than simulation for K-shell events. In contrast to reasonable agreement between CUTE average pulse Power Spectral Density (PSD) and CUTE PT template PSD, discrepancies were observed between the simulation average pulse PSD and simulation PT template PSD in the low-frequency components (0–3 kHz). Consequently, CUTE and simulation raw data were reprocessed with a modified cutoff frequency of 3 kHz. The resulting 0–3 kHz chi-squared distributions demonstrated that low-frequency components dominate the total chi-squared (0–40 kHz) in energy ranges above 4 keV.

Subsequently, pulse shape variation was investigated as a function of energy, revealing energy-dependent saturation effects in the simulation model that alter pulse shapes above 4 keV: the simulation pulse peaks become rounder and the falling edges become steeper as energy increases. To quantify how these saturation effects contribute to the chi-squared discrepancy, comprehensive frequency-domain analysis was then performed across all energy bins, revealing a complex spectral structure. At low energies (below 4 keV) in the 0–3 kHz band, simulation template agreement is moderately worse than CUTE's template agreement, resulting in simulation chi-squared values being only mildly elevated above CUTE rather than a significant discrepancy. A spectral balance mechanism operates in this regime where the excess chi-squared from the 0–3 kHz band is offset by lower chi-squared contributions at higher frequencies (3–40 kHz), maintaining apparent overall agreement. However, above 4 keV, energy-dependent saturation

effects substantially worsen the template mismatch in the 0–3 kHz band. Combined with the quadratic dependence of chi-squared on amplitude, this causes the simulation chi-squared to rise much more steeply than CUTE, breaking the spectral balance.

1 Introduction

1.1 Background and Motivation

The SuperCDMS experiment is one of the leading efforts to directly detect weakly interacting massive particles (WIMPs), a primary dark matter candidate. The experiment employs cryogenic germanium and silicon detectors operated at approximately 30 mK to measure both phonon and ionization signals from particle interactions. High-fidelity simulations are crucial for the experiment's success, as they are used for background modeling, calculating signal efficiencies, and ultimately setting limits on dark matter interactions.

As a foundational step, it is essential to validate that our simulation chain (`SourceSim`, `DMC`, `DAQSim`, `CDMSbats`) can accurately reproduce experimental data from a known calibration source. Warren Perry's study focuses on comparing CUTE detector data with a detailed simulation of the Ge-71 nuclear decay process.[1] However, initial comparisons revealed a significant discrepancy in a key data quality metric—the optimal filter chi-squared at low frequency (PTOFchisqLF)—which motivated this investigation.

For this analysis, we utilized the `Ge71_Kshell_pos50V_100kEvents_V05-09` dataset from Perry's work[2], which was generated using the following software configuration:

- G4CMP: V09-06-00
- SuperSim: V16-00-00
- DAQSim: v8.0.1
- Batcommon: v6.8.0
- CDMSbats: v6.8.1
- CDMSbats config: v0.3.8

[1] See Warren's work: `https://confluence.slac.stanford.edu/spaces/CDMS/pages/601696559/Comparing+Ge+activation+rates+in+data+and+simulation`

[2] More information on DMC samples: `https://confluence.slac.stanford.edu/spaces/CDMS/pages/511346640/DMC+Samples+for+CUTE+Analysis+and+Machine+Learning`

1.2 The Problem: A Systematic Chi-Squared Discrepancy

The primary discrepancy between the CUTE data and the simulation data manifests in the behavior of the optimal filter chi-squared as a function of energy. As shown in Figure 1, the distributions show reasonable agreement below approximately 4 keV, but diverge at higher energies.

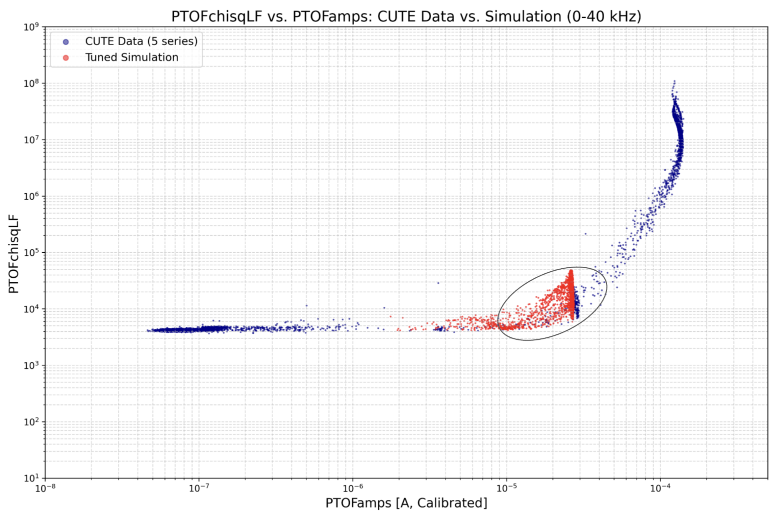


Figure 1: The PTOFchisqLF versus PTOFamps plot for CUTE data (blue) and simulation (red), illustrating the "early rise problem" in the simulation above 4 keV.

This discrepancy is critical, as it raises concerns about the validity of simulation-based predictions for data quality cuts and event selection efficiency, particularly in the energy regime most relevant for dark matter searches.

1.3 Research Objectives

The primary goal of this research was to identify the root cause of the "early rise" problem, where the simulation chi-squared values increase earlier than expected. Specifically, we aimed to investigate whether energy-dependent saturation effects are present in the simulation model above 4 keV and, if so, how these effects alter the simulation pulse shapes and translate into the observed chi-squared discrepancy through frequency-domain analysis.

2 Methodology

To find the cause of the chi-squared discrepancy between the CUTE data and the simulation data, we developed a custom Python analysis pipeline on the Fir computing cluster. This pipeline analyzes the raw detector traces directly in both the time and frequency domains, using the `PTOFamps` and `PTOFdelay` variables from the corresponding Reduced Quantity (RQ) files.

The analysis was conducted in four distinct phases: (1) data processing and event reconstruction, (2) initial K-shell characterization and low-frequency hypothesis testing, (3) analysis of pulse shape variation as a function of amplitude, and (4) comprehensive standard deviation analysis in both the time and frequency domains.

2.1 Data Selection

For CUTE data, we selected five CUTE series located in the directory `/project/rrg-mdiamond/data/CDMS/CUTE/R37/Raw/`. The specific series analyzed were:

- 23231216_211119
- 23240109_075338
- 23240109_021236
- 23240108_203134
- 23231221_101235

For the simulation, the first (R0) and second (R1) runs were used from the directory: `/project/6049244/share/SimData/DMC\_SNOLAB\_HV/Ge71\_Kshell\_pos50V\_100kEvents\_V05-09/Raw/`.

2.1.1 Event Reconstruction and Alignment

Raw detector traces were accessed using the `rawio` library for both CUTE data and simulation data. The PT pulse was calculated as a weighted sum of the individual phonon channels (PAS1 through PFS2), using relative calibration factors derived from L-shell events:

$$PT(t) = \sum_i w_i \cdot [S_i(t) - B_i] \tag{1}$$

where i denotes the channel name, w_i is the relative calibration factor corresponding to channel i, $S_i(t)$ is the raw channel trace corresponding to channel i, and B_i is the baseline calculated as the mean of the first 5000 samples of that trace.

The relative calibration factors used for CUTE and the simulation are listed in Table 1 below.

Table 1: Relative calibration factors used for Phonon Total (PT) reconstruction.

Channel	**CUTE Factor** (w_i)	**Simulation Factor** (w_i)
PAS1	1.2861	1.140
PBS1	1.1154	1.076
PCS1	0.9602	1.002
PDS1	1.0184	0.998
PES1	0.8729	1.018
PFS1	0.7997	0.935
PAS2	1.2545	1.059
PBS2	1.0889	1.008
PCS2	0.7989	0.931
PDS2	1.0332	0.992
PES2	0.8984	0.946
PFS2	0.8735	0.895

A critical requirement for pulse shape comparison is precise temporal alignment. We utilized the Optimal Filter delay (`PTOFdelay`) values in the RQ files. By matching events via `EventTriggerID`, the `PTOFdelay` was applied to every raw trace for correct alignment. To select valid physics events from Detector 1 (Z1), we applied the following basic cuts: the event must be a physics trigger (`TriggerType == 1`) originating from the detector itself (`TriggerDetectorNum == 1`), and the reconstructed Optimal Filter amplitude must be positive (`PTOFamps > 0`). To reject low-quality events, a custom cut was also applied to the ratio of the Optimal Filter amplitude to the raw peak height. Events were required to satisfy:

$$0.8 < \frac{\texttt{PTOFamps}}{\text{Peak Height}} < 1.2 \tag{2}$$

2.2 Initial K-Shell Inspection and Low-Frequency Hypothesis

Prior to analyzing the full energy spectrum, we focused on the K-shell peak region ($9.9 - 10.7$ keV). Visual inspection of the K-shell PT pulses revealed a mismatch in the decay tail between CUTE data and simulation data. This led to a standard deviation analysis of the K-shell pulses, which indicated a significant discrepancy in the low-frequency components ($0 - 3$ kHz).

To test the hypothesis that the chi-squared discrepancy is driven primarily by these low-frequency components, we performed a validation study using the official processing software

(`CDMSbats`). We reprocessed both the CUTE and simulation data with the Optimal Filter upper cutoff frequency lowered to 3 kHz (originally 40 kHz).

2.3 Pulse Shape Variation as a Function of Amplitude

Having established the role of low frequencies in the K-shell, we extended the analysis to the full energy range to investigate whether the pulse shape mismatch is energy-dependent, specifically to test for potential saturation effects above 4 keV.

We divided the full amplitude range into small bins spanning from 0.1×10^{-5} A to 3.0×10^{-5} A with a step size of 0.1×10^{-5} A. For each bin, we collected all contributing pulses from both CUTE and simulation data. To visualize shape differences, each pulse $P(t)$ was normalized by dividing by its Optimal Filter amplitude (`PTOFamps`):

$$P_{\mathrm{norm}}(t) = \frac{P(t)}{\mathrm{PTOFamps}} \tag{3}$$

We computed the average pulse for each bin by taking the mean of all normalized pulses falling within that amplitude range. To facilitate direct visual comparison with the templates, the average pulses were re-normalized by dividing by their peak height. We then overlaid both CUTE and simulation average pulses and their respective templates (CUTE PT template and simulation PT template) across all amplitude bins. Three representative energy ranges were selected for visualization: the L-shell region (1.3 keV), an intermediate energy range (4 keV where the chi-squared divergence begins), and the K-shell region (10.3 keV where the discrepancy is fully developed). The results are shown in Figure 3, Figure 4, and Figure 5.

2.4 Comprehensive Standard Deviation and Frequency Analysis

Having observed pulse shape variations across the energy range, we extended the standard deviation analysis—previously performed only on the K-shell events—to all amplitude bins defined in Section 2.3. The objective was to characterize which frequency components contribute most significantly to the chi-squared discrepancy and how these contributions evolve with energy.

For every bin, we focused on pulse power spectral density (PSD) statistics. We computed the Fast Fourier Transform (FFT) of each individual amplitude-normalized pulse. The individual PSDs were obtained by squaring the magnitude of the FFT. From this collection of individual PSDs, we computed both the average PSD and the standard deviation of the PSDs for that bin.

We examined the 0–3 kHz band in detail, as this region was identified in the K-shell analysis as particularly relevant to the chi-squared discrepancy. Additionally, we analyzed the full 0–40 kHz frequency range to understand the complete spectral behavior across all energy bins.

3 Results

3.1 Initial K-Shell Inspection and Low-Frequency Dominance

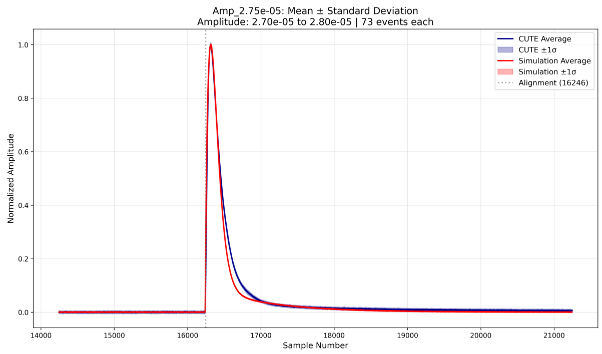
The initial investigation focused on the K-shell energy range ($9.9 - 10.7$ keV) where the chi-squared discrepancy is most pronounced.

As shown in Figure 2, the analysis proceeds from the time domain to the frequency domain and then to chi-squared attribution:

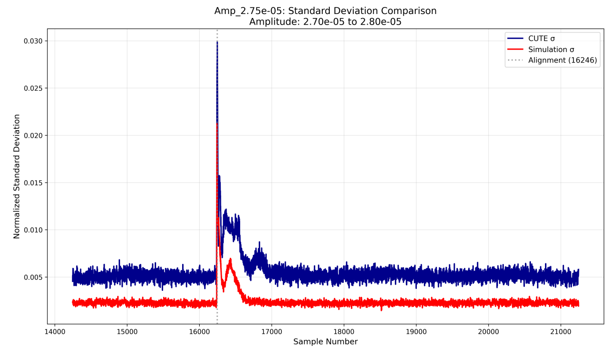
- **Time Domain (a, b):** Time-domain analysis of the K-shell pulses revealed three primary characteristics. First, the CUTE data exhibits roughly two times higher baseline noise levels compared to the simulation data ((Figure 2(b))). Second, the pulse-to-pulse variation in the CUTE data is larger than in the simulation across the entire pulse window ((Figure 2(b))). Finally, a pulse shape mismatch is evident in the decaying tail ((Figure 2(a))).

- **Frequency Domain (c, d):** The average PSDs were obtained by averaging the spectra of individual normalized pulses. Critically, both the CUTE template PSD (gold) and simulation template PSD (purple) are plotted for comparison. The CUTE average PSD aligns reasonably well with the CUTE template PSD throughout the 0–3 kHz range, except at the lowest frequencies (0.05–0.4 kHz). In contrast, the simulation average PSD shows greater deviation from the simulation template PSD. The deviation is most visible in the 0–1.5 kHz range, where the simulation average PSD exhibits higher spectral amplitude than the simulation template PSD.

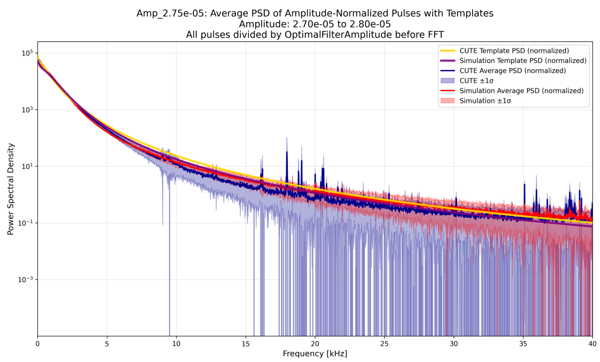
- **Chi-Squared (e, f):** Comparing the chi-squared versus amplitude distributions confirms that the "early rise" is clearly dominated by the 0–3 kHz components.



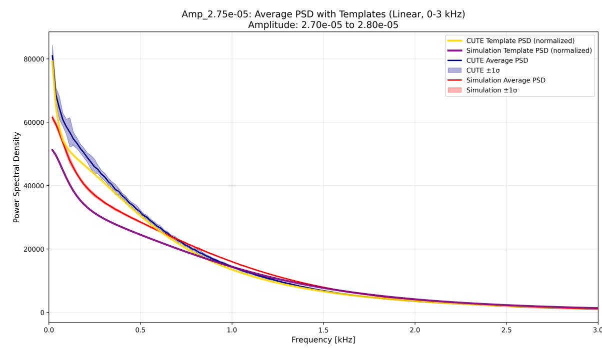
(a) Average pulses (Peak=1) with $\pm 1\sigma$



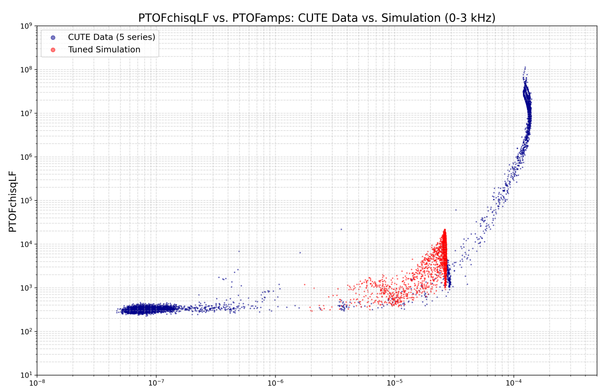
(b) Standard Deviation $\sigma(t)$ Comparison



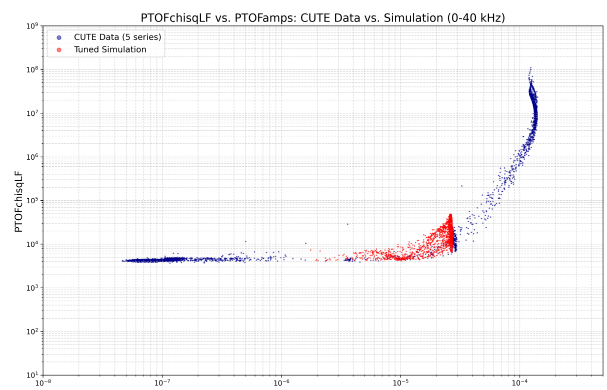
(c) average PSD (Full Range, Log Scale)



(d) average PSD (0–3 kHz, Linear Scale)



(e) χ^2 vs Amplitude (0–3 kHz)



(f) χ^2 vs Amplitude (0–40 kHz)

Figure 2: Comprehensive K-shell pulse shape analysis. (a) Overlay of average CUTE and simulation pulses with $\pm 1\sigma$ bands (renormalized to peak height 1). The CUTE data exhibits significantly wider variation bands and shows mismatch in the decay tail. (b) Time-domain standard deviation comparison showing CUTE data consistently exceeds simulation across baseline, peak, and tail regions. (c, d) Average power spectral density with both templates shown (CUTE template in gold, simulation template in purple). The low-frequency region (0–3 kHz) shows that simulation data deviates more from its template than CUTE data deviates from its template. (e) Chi-squared distribution using only 0–3 kHz bandwidth clearly exposes the "early rise" anomaly. (f) Total chi-squared distribution (0–40 kHz).

3.2 Pulse Shape Evolution As a Function of Energy

To investigate whether the chi-squared rise above 4 keV is caused by energy-dependent effects, we analyzed average pulse shapes across the whole amplitude range. We selected three representative amplitude bins to illustrate the behavior across the energy spectrum:

- **L-Shell Region:** Amplitude bin 3.0×10^{-6} A to 4.0×10^{-6} A (approximately 1.3 keV).
- **Intermediate Energy Region:** Amplitude bin 1.1×10^{-5} A to 1.2×10^{-5} A (approximately 4 keV).
- **K-Shell Region:** Amplitude bin 2.7×10^{-5} A to 2.8×10^{-5} A (approximately 10.3 keV).

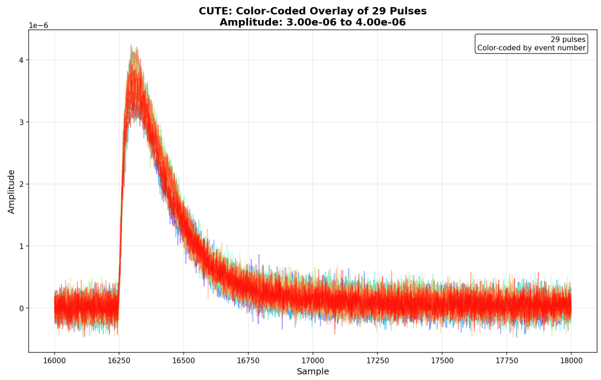
3.2.1 Evidence for Energy-Dependent Saturation Effects in Simulation Model

When comparing each dataset to its respective template, distinct patterns emerge across the energy range. Figure 3 displays the L-shell bin, Figure 4 shows the intermediate energy bin, and Figure 5 presents the K-shell bin.

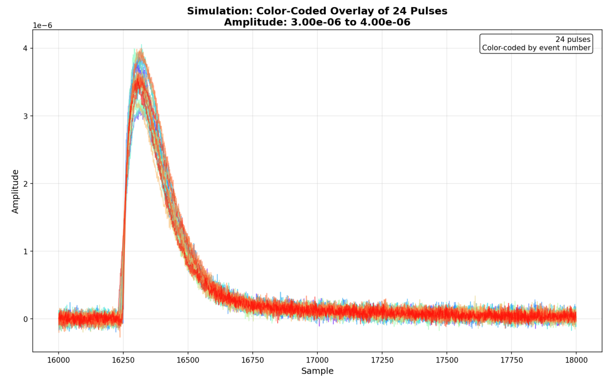
At the L-shell energy, both CUTE and simulation data show good agreement with their respective templates. At intermediate energy (4 keV), deviations begin to appear in the simulation pulses compared to the simulation template. At the K-shell energy, the simulation pulses show pronounced deviations from the simulation template:

1. **Peak Rounding:** The pulse peak becomes more rounded and broader compared to the template peak.
2. **Faster Descent:** The falling edge exhibits a more rapid return to baseline than predicted by the template.

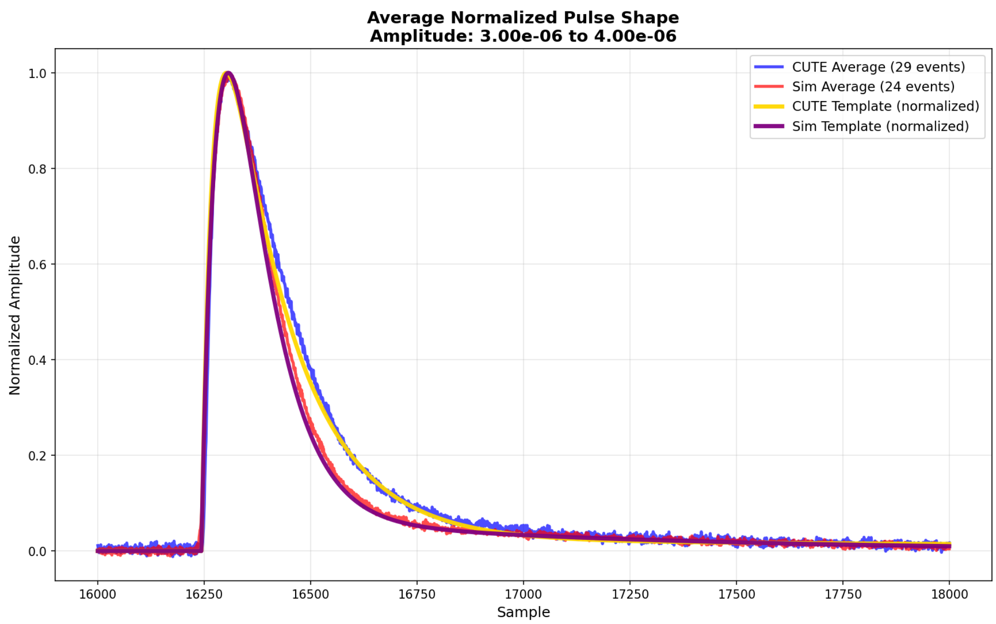
The CUTE data maintains good consistency with the CUTE template across all three energy ranges. The simulation pulse shapes should remain consistent with the simulation template across all energies, but instead exhibit changes at higher energies indicating energy-dependent saturation effects. When chi-squared is calculated for each dataset using its respective template, this energy-dependent saturation behavior in the simulation drives the observed "early rise" in the simulation chi-squared distribution above 4 keV.



(a) L-Shell: Individual CUTE Pulses

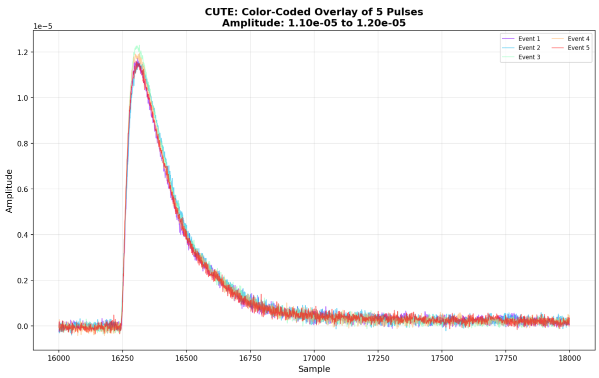


(b) L-Shell: Individual Simulation Pulses

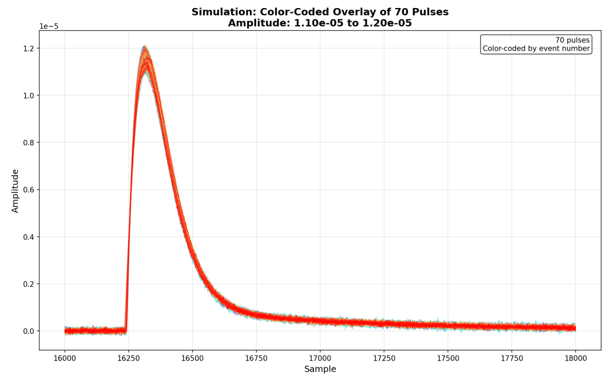


(c) L-Shell: Average Pulse Overlay with Templates

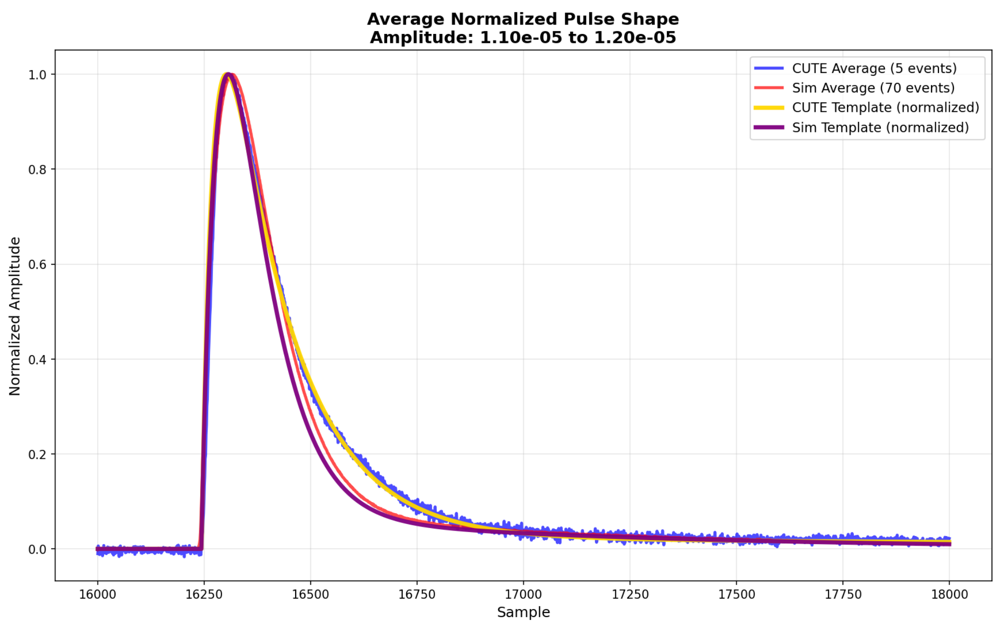
Figure 3: Pulse shape analysis for the L-shell amplitude bin ($3.0-4.0\times10^{-6}$ A, approximately 1.3 keV). (a, b) Individual CUTE and simulation pulses showing pulse-to-pulse variation. (c) Average pulse overlay showing good agreement between simulation data and simulation template (purple) at low energy, and good agreement between CUTE data and CUTE template (gold).



(a) Intermediate Energy: Individual CUTE Pulses

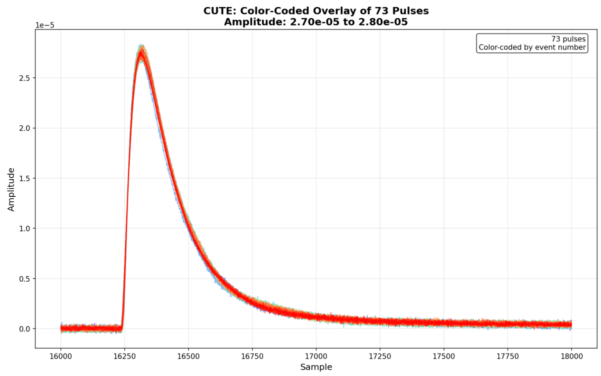


(b) Intermediate Energy: Individual Simulation Pulses

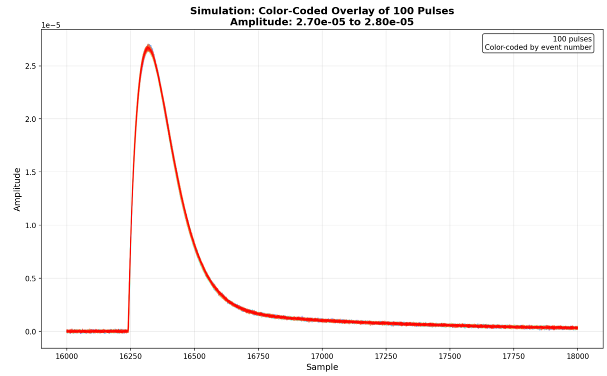


(c) Intermediate Energy: Average Pulse Overlay with Templates

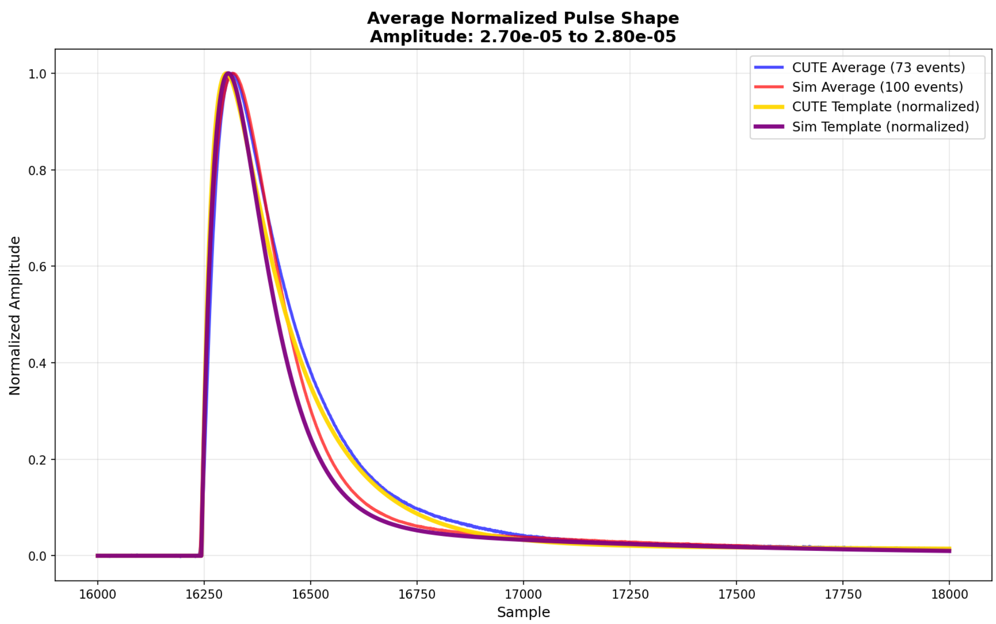
Figure 4: Pulse shape analysis for the intermediate energy amplitude bin ($1.1 - 1.2 \times 10^{-5}$ A, approximately 4.0 keV). (a, b) Individual CUTE and simulation pulses showing pulse-to-pulse variation. (c) Average pulse overlay revealing the onset of energy-dependent artifacts: simulation pulses begin showing deviations from the simulation template (purple), with emerging peak rounding and faster falling edge. The CUTE data continues to maintain good consistency with the CUTE template (gold).



(a) K-Shell: Individual CUTE Pulses



(b) K-Shell: Individual Simulation Pulses



(c) K-Shell: Average Pulse Overlay with Templates

Figure 5: Pulse shape analysis for the K-shell amplitude bin ($2.7 - 2.8 \times 10^{-5}$ A, approximately 10.3 keV). (a, b) Individual CUTE and simulation pulses showing pulse-to-pulse variation.(c) Clear evidence of energy-dependent artifacts in the simulation: simulation pulses show rounded peaks and faster falling edges compared to the simulation template (purple). The CUTE data maintains good consistency with the CUTE template (gold), indicating that the simulation model is incorrectly introducing saturation-like distortions at higher energies that are not present in the actual detector.

3.2.2 Global Pulse Shape Evolution

To comprehensively visualize the energy-dependent pulse shape evolution, we examined the superposition of average pulses from all amplitude bins analyzed.

Figure 6 presents this global summary with both templates shown. The left figure shows CUTE experimental data compared to the CUTE template (gold), while the right figure shows simulation data compared to the simulation template (purple). The color gradient transitions from blue (low amplitude, ~ 1 keV) to red (high amplitude, ~ 10 keV).

The key observation is that the simulation pulses (right figure) show systematic deviations from the simulation template as energy increases—the red traces at high energy deviate noticeably from the purple template, showing rounded peaks and faster descent. The CUTE pulses (left figure) show better overall agreement with the CUTE template across the energy range, though some low-energy traces show irregular shapes due to the combination of higher pulse-to-pulse variation and lower event counts in those bins.

This conclusively demonstrates that energy-dependent saturation effects are present in the simulation model. The simulation pulses should maintain consistency with the simulation template across all energies as the CUTE pulses do with their template, but instead exhibit distortions caused by energy-dependent saturation effects. These saturation effects in the simulation are the primary cause of the chi-squared discrepancy above 4 keV.

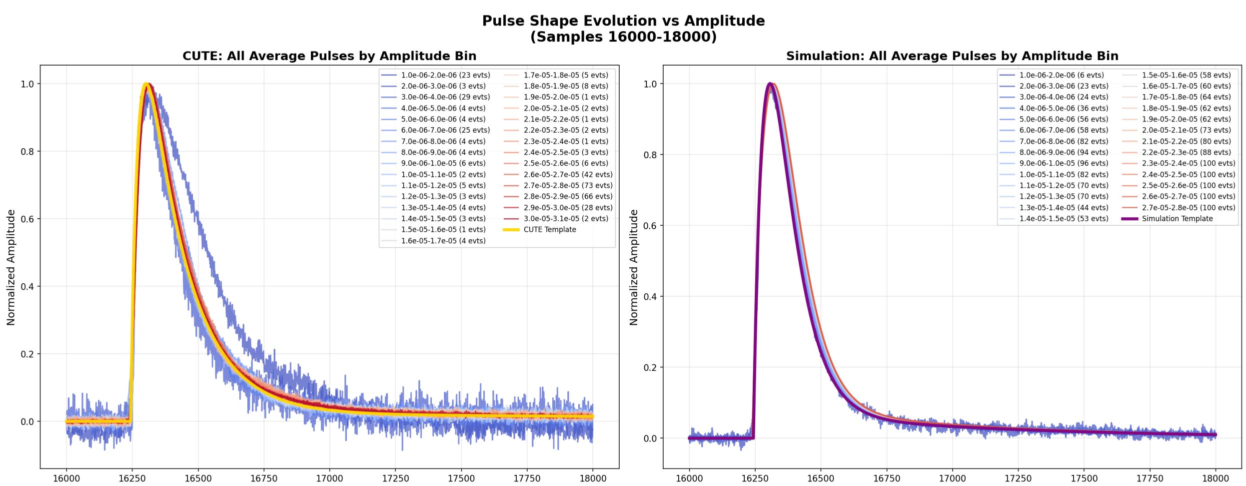


Figure 6: Global summary of pulse shape evolution across all amplitude bins ($0.1 - 3.0 \times 10^{-5}$ A). **Left:** Average CUTE PT pulses (colored by energy) compared to CUTE template (gold). The CUTE data maintains good agreement with its template across the energy range. **Right:** Average simulation pulses (colored by energy) compared to simulation template (purple). At higher energies, average simulation pulses begin to deviate from the simulation template, showing progressively rounder peaks and faster descent as energy increases, which indicates the presence of energy-dependent saturation effects in the simulation model.

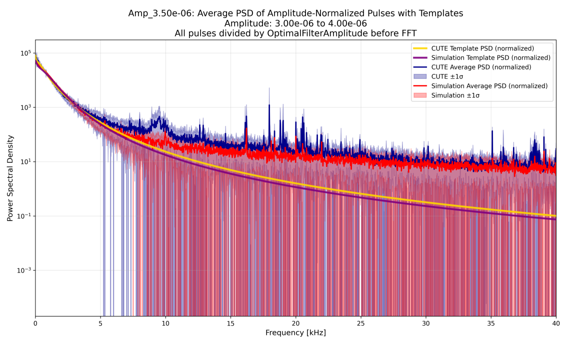
3.3 Frequency Domain Analysis Across Energy Bins

To investigate how saturation effects cause the chi-squared discrepancy, we performed frequency-domain analysis for three representative amplitude bins: L-shell (1.3 keV), intermediate energy (4 keV), and K-shell (10.3 keV).

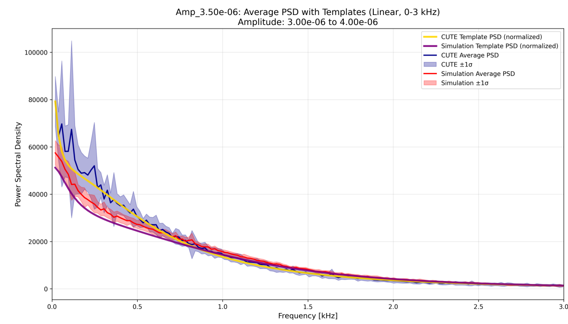
Figure 7 presents the pulse PSD comparison for these three energy bins. Each row shows two views: the full 0–40 kHz band (left) and a zoomed view of the 0–3 kHz region (right). In each figure, both the CUTE template PSD (gold) and simulation template PSD (purple) are shown as references. The average pulse PSDs for CUTE (blue) and simulation (red) are plotted with their respective $\pm 1\sigma$ variation bands shown as shaded regions.

In the 0–3 kHz region, the simulation average PSD shows greater deviation from the simulation template PSD compared to the deviation between CUTE average PSD and CUTE template PSD at all three energies. However, the magnitude of this simulation template mismatch evolves with energy: at L-shell, the deviation is relatively modest, while at intermediate energy and K-shell, the deviations become more pronounced.

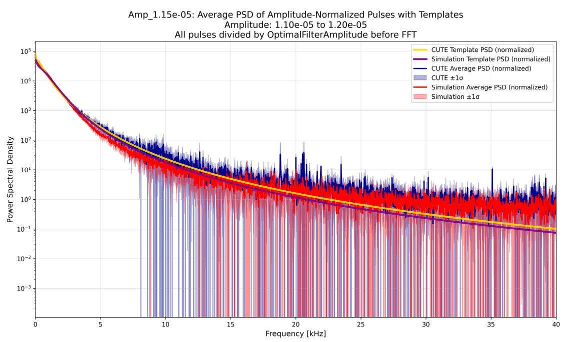
The full-band spectra (0–40 kHz) reveal additional energy-dependent patterns. At L-shell, the CUTE average PSD is generally higher than the simulation average PSD in the 3–40 kHz range, with both datasets deviating from their respective templates after approximately 7 kHz. At intermediate energy, the CUTE average PSD remains higher than simulation in the 3–40 kHz range, but this difference is reduced compared to L-shell, and both datasets show improved alignment with their respective templates. At K-shell, a reversal occurs where the simulation average PSD becomes higher than the CUTE average PSD. Notably, in the approximate range of 7–25 kHz, the simulation shows closer agreement with its template than CUTE does with its template.



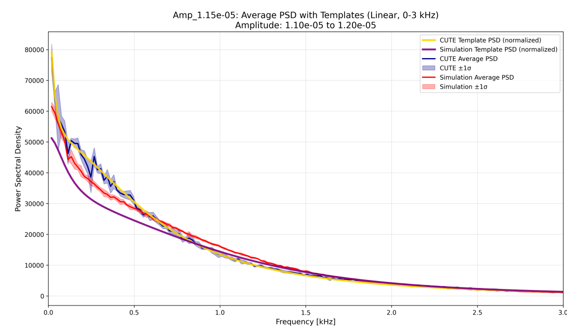
(a) L-Shell: Full Band (0-40 kHz)



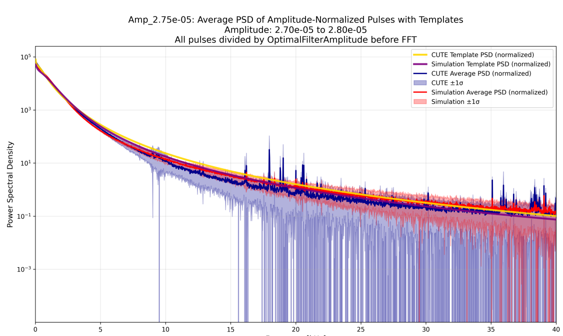
(b) L-Shell: Low Freq (0-3 kHz)



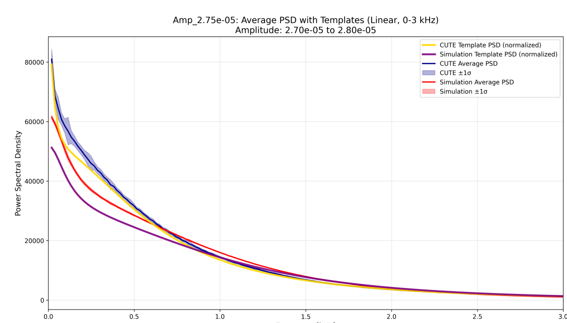
(c) 4 keV: Full Band (0-40 kHz)



(d) 4 keV: Low Freq (0-3 kHz)



(e) K-Shell: Full Band (0-40 kHz)



(f) K-Shell: Low Freq (0-3 kHz)

Figure 7: Evolution of pulse power spectral density across three energy scales: L-shell (1.3 keV), intermediate energy (4 keV), and K-shell (10.3 keV). **Left column:** Full bandwidth (0–40 kHz, log scale). **Right column:** Low-frequency region (0–3 kHz, linear scale). Both CUTE template PSD (gold) and simulation template PSD (purple) are shown for reference. Average PSDs are plotted for CUTE (blue) and simulation (red) with $\pm 1\sigma$ variation bands (shaded regions). Comparing across rows demonstrates how template agreement and spectral characteristics evolve with energy.

4 Analysis and Discussion

4.1 Analysis of Low-Frequency Mismatch

The frequency-domain analysis across energy bins reveals that the simulation consistently exhibits larger deviation from the simulation template in the low-frequency region (0–3 kHz) compared to how CUTE data deviates from the CUTE template. This pattern persists at L-shell, intermediate, and K-shell energies, though the magnitude of the simulation's template mismatch increases with energy.

Reprocessing of the K-shell data with a 3 kHz cutoff confirmed this analysis. When the bandwidth is restricted to include only these low frequencies, the chi-squared discrepancy becomes even more pronounced, demonstrating that the 0–3 kHz band is the primary contributor to the total chi-squared (0–40 kHz) discrepancy.

At low energies, this low-frequency mismatch in the simulation is relatively mild. However, at higher energies, the deviation becomes substantially more severe due to energy-dependent saturation effects that alter the pulse shape. The 0-3 kHz band is particularly sensitive to these saturation-induced pulse shape distortions, resulting in the dominant contribution to the total chi-squared discrepancy.

4.2 Evidence for Energy-Dependent Saturation Effects

The comparison of average pulse shapes across the energy range (from L-shell through intermediate energy to K-shell) demonstrates clear evidence of energy-dependent saturation effects in the simulation model.

At L-shell (1.3 keV), both CUTE and simulation pulses show good agreement with their respective templates. However, at intermediate energy (4 keV), deviations begin to appear in the simulation pulses compared to the simulation template. At K-shell (10.3 keV), pronounced saturation signs emerge in the simulation:

- **Peak Rounding:** The pulse peak becomes more rounded and broader compared to the template peak.
- **Faster Descent:** The falling edge exhibits a more rapid return to baseline than predicted by the template.

The global pulse shape summary (Figure 6) clearly visualizes this progression: simulation pulses (colored blue to red) progressively deviate from the simulation template as energy increases, while CUTE pulses maintain good consistency with the CUTE template across the whole energy range.

These energy-dependent saturation effects in the simulation indicate a deficiency in how the simulation model handles pulse generation at higher energies.

4.3 Complex Behavior for Frequencies Above 3 kHz and Noise Weighting

While the 0–3 kHz band shows consistent simulation template mismatch across all energies (with magnitude increasing above 4 keV due to saturation), the 3–40 kHz range exhibits complex energy-dependent patterns. These patterns play a crucial role in explaining the apparent total chi-squared agreement at low energies.

4.3.1 Energy-Dependent Spectral Patterns Above 3 kHz

The relationship between simulation pulse PSD, CUTE pulse PSD, and their respective templates varies with energy:

- **At L-Shell:** After approximately 7 kHz, both datasets deviate from their respective templates. However, simulation shows closer template alignment in the 3–40 kHz range, with CUTE exhibiting larger upward deviations from its template.
- **At Intermediate Energy:** The CUTE average PSD remains higher than simulation in the 3–40 kHz range, but the difference is reduced compared to L-shell. Both datasets show improved alignment with their respective templates.
- **At K-Shell:** A reversal occurs where the simulation average PSD becomes higher than CUTE average PSD. In the 7–25 kHz range, simulation still shows closer agreement with its template than CUTE does, but this pattern changes at the highest frequencies.

Consequently, there is no single frequency band that cleanly separates regimes of good and poor agreement. The spectral alignment changes across the frequency range and varies significantly with energy.

4.3.2 The Role of Noise Weighting

Crucially, the chi-squared calculation depends not only on the deviation from the template (the numerator) but also on the noise power (the denominator). Our time-domain analysis confirmed that the simulation data exhibit significantly lower noise than the CUTE data.

This creates a competing effect in the chi-squared calculation:

1. **Closer Spectral Alignment (Lowers χ^2):** At low energies, in the 3–40 kHz frequency range, the simulation average pulse spectrum aligns more closely with the simulation template than CUTE data aligns with the CUTE template. This reduced mismatch tends to lower the χ^2 contribution in these regions.

2. **Stricter Standard (Raises χ^2):** Because the simulation noise is smaller, the weighting factor ($1/\text{Noise}^2$) is larger. This acts as a magnifying glass, amplifying mismatches from the template that would be negligible in the noisier experimental data.

Therefore, these observations lead us to hypothesize that the apparent agreement observed at low energies arises from a complex balance. Although the simulation consistently yields a higher chi-squared contribution in the 0–3 kHz band, the complex interplay of fit quality and noise weighting above 3 kHz appears to result in a net deficit relative to the CUTE data. At low energies, this high-frequency deficit roughly offsets the low-frequency excess, leading to a total chi-squared that matches the experimental data.

4.4 The Energy-Dependent Breakdown

This delicate spectral balance breaks down as energy increases above 4 keV, driven by the emergence of energy-dependent saturation effects in the simulation model.

- **Low Energy ($<$ 4 keV):** The simulation shows modest template mismatch in the 0–3 kHz band, contributing excess chi-squared. However, this is offset by chi-squared contributions from the 3–40 kHz range. The spectral balance mechanism maintains the apparent agreement in the total chi-squared.

- **Intermediate Energy (4 keV):** Energy-dependent saturation effects begin to emerge, causing the 0–3 kHz template mismatch to worsen. The spectral balance begins to destabilize as the low-frequency excess grows.

- **High Energy ($>$ 4 keV):** Energy-dependent saturation effects become pronounced, substantially worsening the template mismatch in the 0–3 kHz band. Because chi-squared scales with the square of amplitude, the low-frequency chi-squared contribution grows rapidly with energy. At high energies, this quadratic growth in the 0–3 kHz band overwhelms the compensating contributions from frequencies above 3 kHz.

5 Conclusion and Future Work

5.1 Summary of Findings

This investigation successfully identified the root cause of the chi-squared distribution discrepancy between CUTE experimental data and simulation data in the SuperCDMS. The "early rise problem," where simulation chi-squared values increase earlier than CUTE values above approximately 4 keV, is driven by energy-dependent saturation effects in the simulation model.

Initial K-shell analysis identified discrepancies between the simulation template PSD and simulation average pulse PSD in the 0–3 kHz frequency band, establishing this region as the primary contributor to the chi-squared discrepancy. Reprocessing with a 3 kHz cutoff validated that low-frequency components dominate the total chi-squared behavior.

Pulse shape analysis across the full energy range revealed clear evidence of energy-dependent saturation effects in the simulation. At low energies (L-shell, 1.3 keV), both CUTE and simulation pulses show good agreement with their respective templates. However, at intermediate energies (4 keV), subtle deviations begin to emerge in simulation pulses, and at high energies (K-shell, 10.3 keV), pronounced saturation signs appear: rounded peaks and faster falling edges compared to the simulation template. The CUTE experimental data maintains good consistency with its template across the whole energy range, demonstrating that the simulation's saturation effects are more severe and emerge at lower energies than in the actual detector.

Comprehensive frequency-domain analysis revealed a complex spectral balance mechanism operating at low energies. Below 4 keV, the excess chi-squared from the simulation's 0–3 kHz template mismatch is approximately offset by lower chi-squared contributions from the 3–40 kHz range. Above 4 keV, energy-dependent saturation effects break this delicate balance by substantially worsening the 0–3 kHz template mismatch, causing the observed early rise in simulation chi-squared.

These findings have important implications for SuperCDMS. Accurate simulation of detector response is essential for reliable background modeling and signal efficiency calculations in the dark matter search. The observed energy-dependent discrepancies indicate that the simulation framework requires further refinement to better match experimental behavior at higher energies.

5.2 Recommended Next Steps

To address the identified simulation deficiencies and validate the spectral balance mechanism, the following investigations are recommended:

1. Binned Frequency Chi-Squared Analysis: Develop a modified optimal filter algorithm to calculate chi-squared contributions within discrete, narrow frequency bands (e.g., 0–1 kHz, 1–2 kHz, . . . , 39–40 kHz). This would quantify precisely where simulation contributions exceed or fall below experimental data across the entire energy spectrum, providing definitive validation of the spectral balance hypothesis.

2. Investigation of Simulation Model Deficiencies: Conduct a systematic investigation of the simulation chain to identify the source of energy-dependent saturation effects.

3. Energy-Dependent Calibration Factors: Re-evaluate whether energy-independent relative calibration factors (derived solely from L-shell events) are sufficient across the full energy range. While this may not address the fundamental saturation effects in the simulation model, energy-dependent calibration factors could provide insight into systematic effects and potentially improve data-simulation agreement.

References

[1] W. Perry. (2025). *Comparing Ge activation rates in data and simulation*. SuperCDMS Confluence Page. `https://confluence.slac.stanford.edu/spaces/CDMS/pages/601696559`

[2] W. Perry. (2025). *DMC Samples for CUTE Analysis and Machine Learning*. SuperCDMS Confluence Page. `https://confluence.slac.stanford.edu/spaces/CDMS/pages/511346640`