Maximizing Sensitivity for the Cryogenic Dark Matter Search Low Threshold Analysis

James Reed Watson

January 17, 2018

Abstract

Dark matter has a wealth of evidence to support its existence in abundance on galactic and cosmological scales. One candidate for dark matter is the weakly-interacting massive particle (WIMP). Many experiments are attempting to detect this particle, either directly or indirectly. One such experiment is the Cryogenic Dark Matter Search (CDMS), which published results in 2014 where a mass-dependent limit was set on the WIMP-nucleon interaction cross section. Due to the rarity of these interactions, the analysis depended heavily on the optimization scheme of several selection criteria in order to adequetely filter out background events. Certain choices regarding the method of optimization were made that may have affected the sensitivity of the analysis. In this thesis, we discuss these choices and explore the effects of alterations to the optimization methods on the expected cross section limit. These alterations have the effect of improving the expected sensitivity to 5 GeV/c^2 WIMP interactions by 61.5% and 15 GeV/c^2 WIMP interactions by 8.89%, though when a more complete consideration of the analysis techniques taken into account these results are less conclusive.

Dedication

I dedicate this thesis to my parents, who supported me throughout college so that I could pursue my goals.

Acknowledgements

I would like to thank David Toback, who patiently worked with me until this thesis was as good as it could be. Jon Wilson and Jorge Morales provided the guidance necessary to write the code used in this thesis. I would also like to thank Adam Anderson, who explained the work that I would later build on for this thesis. The work done in this thesis was enabled by the Brazos cluster.

Contents

1 Introduction **6**
- 1.1 Dark Matter Searches with the SuperCDMS Experiment . . . 6
- 1.2 Overview of the SuperCDMS Low Threshold 2014 Analysis Strategy . . . 7
  - 1.2.1 Terminology . . . 8
- 1.3 Estimation of the Sensitivity of the Analysis . . . 9
  - 1.3.1 Description of the 2014 Low Threshold Analysis . . . 9
  - 1.3.2 Overview of Thesis: Analysis Optimization Methods . . . 11

2 Replication of the 2014 Low Threshold Analysis **13**
- 2.1 Event Selection Requirements . . . 13
  - 2.1.1 Preselection Requirements . . . 13
  - 2.1.2 Boosted Decision Tree (BDT) Selection Criteria . . . 13
- 2.2 WIMP Nuclear Recoil Energy Spectrum and Phonon Energy Deposits . . . 14
- 2.3 Efficiencies . . . 16
- 2.4 Spectrum Averaged Efficiencies and the Total Exposure . . . 16
- 2.5 Backgrounds . . . 17
- 2.6 Expected Sensitivities . . . 18
- 2.7 Validation of 2014 Results . . . 19

3 Reoptimizing the Analysis **21**
- 3.1 Background Subtraction and BDT Thresholds . . . 21
- 3.2 Incorporating Detector Differences into Limit Setting . . . 24
- 3.3 Sensitivity with the .OR. Cut . . . 25

4 Conclusions **29**
- 4.1 Summary . . . 29

A Spectrum, Efficiencies, and Backgrounds **32**
- A.1 Phonon Spectrum . . . 33
- A.2 Efficiencies . . . 35

B Sensitivities with and without Background Subtraction **38**

C BDT Threshold Dependence **43**
- C.1 SAEx vs. BDT . . . 44
- C.2 Background vs. *SAE*, BDT-parameterized . . . 49
- C.3 Relative Background and SAE . . . 54
- C.4 Background vs. BDT . . . 59

D 8-bin Optimization **64**
- D.1 Coarse Tuning . . . 65
- D.2 Fine Tuning . . . 70

E Sensitivity and WIMP Mass **75**

Chapter 1

Introduction

1.1 Dark Matter Searches with the SuperCDMS Experiment

There is a preponderance of indirect evidence for the existence of a particle that interacts with hadronic matter gravitationally but not electromagnetically, known as dark matter[1]. The visible light produced in stars provides an estimate of the distribution of atomic matter within a galaxy. This distribution, combined with Newtonian dynamics, provides a predicted rotational velocity distribution for the outermost stars in a galaxy. Observed galactic rotational curves disagree with the predictions from the observed luminosity and Newtonian gravity, suggesting a need for modifications to the gravitational force felt in and around a galaxy[2], either in the form of a correction to Newtonian gravity or a new form of matter. The galaxy cluster collision occuring in the Bullet Cluster provides evidence that the bulk of a galaxy cluster's matter comes in the form of a non-luminous and minimally interacting particle, and not any kind of theory of modified gravity[3]. If dark matter was produced in copious amounts in the early universe and later fell out of thermal equillibrium as the universe expanded, the observed dark matter density predicts its annihilation cross section at approximately the weak scale[4], leading to the "Weakly Interacting Massive Paricle" candidate for dark matter, which this thesis will consider. This coincidence is known as the "WIMP miracle". Perhaps the best evidence for dark matter (as opposed to some modified gravity theory) comes from the cosmic microwave background. The anisotropies in the matter distribution in the universe are difficult to explain with anything but a cold dark matter theory[5].

The Cryogenic Dark Matter Search (CDMS) is one of many experiments which looks for interactions between dark matter particles and earthbound detectors[6]. When a particle interacts with an atom in a CDMS detector (a cryogenic bolometer made of ultrapure germanium and known as an iZIP), it deposits energy. During WIMP scattering, a target nucleus within the detector will recoil, absorbing an amount of energy known as the "recoil energy" for that event. This energy is then transferred to the rest of the detector in the form of quantized sound waves, known as phonons. A portion of the interaction energy of the energy also goes into the ionizing the germanium atoms. This freed charge is then extracted out of the lattice with an applied voltage and then read out as a current signal. The quantity of charge extracted provides both a second energy measurement for the interaction and a method of discriminating WIMP-like recoils and electron-like recoils, which are a dominant background.

The experiment uses a number of criteria to select and count deposits of energy (events) and then compare those numbers to expected background contamination levels. Any excesses could serve as evidence for dark matter interactions in the detectors. These chosen selection criteria provide a way to calculate the expected background and detection efficiency of the experiment, which are in turn used to calculate an expected sensitivity to dark matter interactions.

This analysis uses data recorded between March 2012 and July 2014 (Run 133), and replicates the methods used in the 2014 Phys Rev. Lett[6]. We will refer to these as the "2014 Analysis" or the "PRL Analysis". Note that no evidence for dark matter was observed in that run, and limits were set on the WIMP-nucleon interaction cross-section as a function of WIMP mass. While the original analysis was strong, some decisions were made during the selection criteria optimization that potentially limited its sensitivity. We explore three assumptions of the 2014 analysis and consider the potential sentitivity under two new optimization methods.

1.2 Overview of the SuperCDMS Low Threshold 2014 Analysis Strategy

We begin by overviewing the analysis and defining the terms that will be used in our sensitivity and optimization techniques. The "Low Threshold" analysis refers to the fact that this is a search for WIMPs with masses in the 5-15 GeV/c^2 range, which deposit less energy into the detectors than high mass WIMPs. In this regime only events with total recoil energy above a low threshold are considered. The analysis is conducted on events that were chosen using a set of selection criteria chosen to maximize the sensitivity, or equivalently minimize the limit on the WIMP-nucleon interaction cross section [6]. These criteria are divided into two categories, known as "preselection" and "background discrimination" (we also use the term "BDT," to be defined below) requirements. The number of events recorded in the detectors is compared to an expected number of (non-WIMP) events and, depending on the outcome, a discovery or an upper limit on the cross section is reported.

For the purposes of choosing selection criteria and setting expected limits, the 2014 analysis was optimized under the no-signal hypothesis, using a "cut-and-count" strategy where the selection criteria determine the background count expectation value. This means that the analysis is tuned in order to maximize the expected sensitivity if either dark matter did not exist, or did not interact with normal matter at all (equivalent interaction cross section of $\sigma = 0$ cm^2) before looking at any experimental data. An important caveat to this description is that this method determines the selection criteria but is distinct from the method used to eventually set the limit for the final result based on experimental data, which is described in more detail below.

The first set of selection criteria, the preselection criteria, exist primarily to remove events that are unlikely to be from actual particle interactions (e.g. the pulses look nothing like what one would expect from a particle interaction), as well as to execute broad but effective selection on the events which appear to be from interactions with the nucleus. The second category of criteria exists to discriminate events which look like Standard Model interactions from events which might be WIMP-SM interactions.

We refer to two broad classes of events known as "electron recoils" and "nuclear recoils." These classes are differentiated using the "ionization yield": the ratio of the energy collected from the charge released from an interaction to the energy released via phonons (recoil energy). In order to quantify the differences in the ionization yield between gamma or electron interations and neutron-like interactions, the detectors are exposed to ^{252}Cf (neutron emitter, creating a WIMP-like signal) and ^{133}Ba (γ emitter) sources. The results of this exposure on a representitive detector are shown in Fig. 1.1. Electrons and photons will typically interact with the electrons within the lattice, which leads to almost equal ionization and recoil energies, or in other words, an ionization yield of 1 regardless of the recoil energy (see the contours surrounding the two set of events, the upper one is the "electron recoil band"). Neutrons (and WIMPs) on the other hand will preferentially scatter off of the nucleus of the atoms in the detector, and not generate as large an ionization yield (usually ~ 0.3). Thus, by selecting events with a particular relationship between ionization energy and recoil energy, one may exlude a large amount of background events right from the start.

Low mass WIMP interactions are likely to deposit only small amount of energy. However, at small energies, the nuclear recoil and electron recoil bands overlap, necessitating that a more complicated selection must be performed, for example machine learning in the form of the Boosted Decision Trees (BDT)[7]. The same calibration exposures used in e.g. Fig 1.1 are used to train a BDT to score and separate the electron and nuclear recoils. Separate BDTs were trained on each detector, and the expected energy distributions are used to further optimize for different mass WIMP candidates. In the final analysis, each event for a given detector will be run through each (mass-optimized) BDT for that detector and be given a score. Events with scores above these thresholds will be selected for analysis. The set of BDT thresholds with serve as the final, optimized "BDT" selection criteria. The results of this training are shown in Fig. 1.2.

Since these criteria can not simultaneously reject all background while retaining every good signal event we must estimate the efficiency for a WIMP interaction to actually be counted (event selection efficiency). Before we describe the analysis in full, and show the quantitative estimate of the sensitivity, we define some terms used in our estimation of the sensitivity that will be convenient for further exploration.

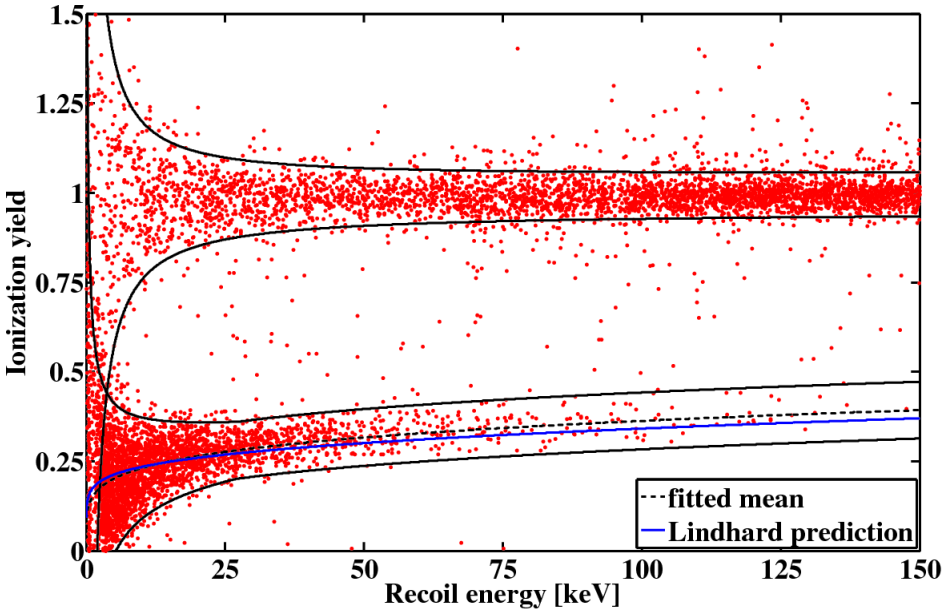

Figure 1.1: Ionization yield versus recoil energy of ^{252}Cf calibration data from a representative detector in one of the six data runs. The black/solid lines represent the chosen electron-recoil band around a yield of one and the nuclear-recoil band around 0.3. The black/dashed line denotes the mean of the latter band, while the similar but blue/solid line is the corresponding prediction from Lindhard theory [9] (screening of electric field by electrons in a solid). Taken from Ahmed *et al.* 2011[10].

1.2.1 Terminology

Our quantitative estimate of the experimental sensitivity will be described using the quantities defined below. The following sections of this note elaborate more on how each of these quantities is calculated and eventually utilized.

- N: Number of events expected or observed in the real experiment or a pseudoexperiment. In the no-signal hypothesis this is the number of background events. Units: 1.
- σ: WIMP-nucleon cross section. Units: cm^2.
- ξ: Sensitivity Parameter which corresponds to the interaction rate of a WIMP with an amount of regular matter. Units: 1/ (kg·days).
- η: Final Efficiency, the fraction of WIMP events which would pass all selection criteria, also known as the efficiency of the experiment. This efficiency is a function of BDT threshold and phonon energy. Units: (1/ keV).
- B: Background, the expected number of background events passing a set of selection criteria. Units: 1.
- T: Exposure of the detectors. This is the time the detectors were on and collecting data multiplied by the total mass of the detectors. Units: kg·days
- SAE: Spectrum-Averaged Efficiency, the final efficiency convolved with the WIMP recoil energy spectrum and multiplied by exposure. Units: 1/cm^2
- $SAEx$: Spectrum-Averaged Exposure, the SAE divided by the total WIMP spectrum over all energies. Units: kg·days.

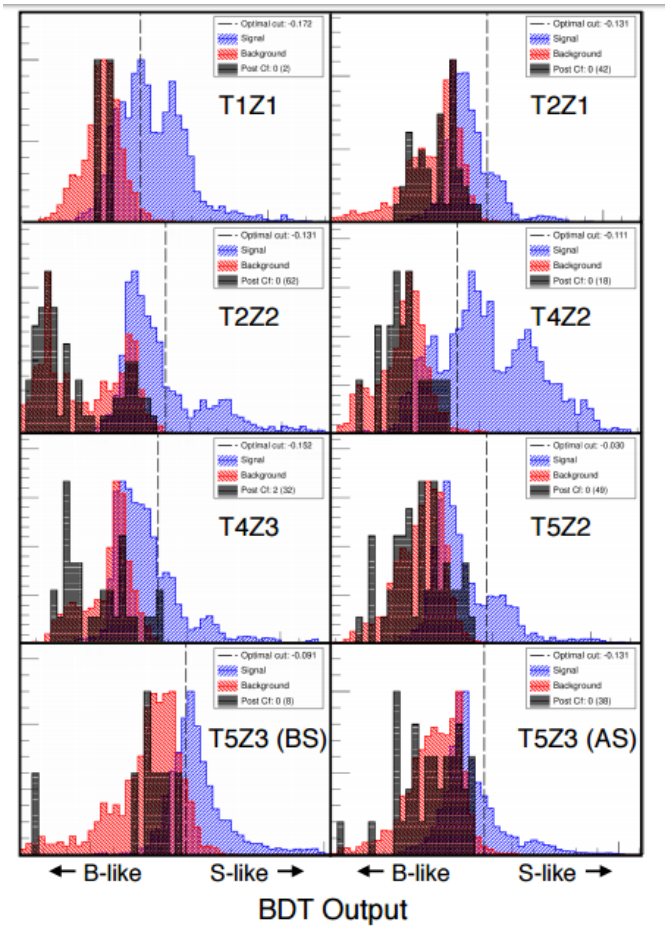

Figure 1.2: Histogram of BDT response to different input samples. The red data is a histogram of BDT responses for Ba calibration data. The blue data is a histogram of BDT responses for Ca calibration data. (Courtesy of Adam Anderson[13]). The scale on the x-axis is the BDT parameter, which is traditionally chosen to range from -1 to +1.

- R: WIMP interaction rate within the detectors. Typically we use the derivative of R with respect to either phonon or recoil energy, which is later integrated to get the total WIMP event rate. Units: 1/(kg·day).
- E_p: Phonon energy deposited during an event. Units: keV
- E_r: Recoil energy deposited during an event. Units: keV

1.3 Estimation of the Sensitivity of the Analysis

1.3.1 Description of the 2014 Low Threshold Analysis

We next describe the Low Threshold (LT) analysis as published in the 2014 paper, as well as the methods used to estimate the analysis sensitivity from the expected backgrounds and acceptance. In a given experiment, there is a direct relationship between the number of expected events seen in the detector and the interaction cross section. For the remainder of this thesis, we frequently speak in terms of pseudo-experiments. This is because we are optimizing the selection criteria in order to maximize the sensitivity, and as such are averaging over many possible realizations of the experimental results. After observing a number of events in

each detector, we can estimate the magnitude of the cross section (or sensitivity). Using the cross section, the number of expected background events, and the detector response (SAE, defined above to be the efficiency convolved with the WIMP spectrum and multiplied by exposure), we obtain an expected number of events observed in the experiment. The observed number of events is Poisson-distributed, and is related to SAE, cross section, and the number of observed background events by the following equation:

$$N = \sigma \cdot SAE + B, \tag{1.1}$$

or equivalently by:

$$N = \xi \cdot SAEx + B. \tag{1.2}$$

These equations highlight the reparameterization of SAE into $SAEx$. Note that because N will be on the order of a handful of events, $SAEx$ absorbs the order-of-magnitude scale from σ, which makes both $SAEx$ and ξ of order unity (within a few orders of magnitude, but compared to 10^{-42} this is "close"). Given the number of events observed in an experiment or pseudoexperiment, and an estimate (from other methods) of the expected backgrounds and SAE, an upper limit on the WIMP-nucleon cross section can be calculated. In this context a "pseudoexperiment" is simply a simulated potential result of the experiment that is randomly drawn from the range of possible observations. The terminology exists to emphasize that actual detector data was not used, aside from calibration information. Looking ahead to optimization, we note that the SAE is a function of the preselection criteria, the BDT selection criteria, and the trigger efficiency (efficiency for the detectors to record an event), all of which are functions themselved of the phonon energy deposited within the detector (though this is integrated over). The sensitivity of the experiment's analysis ultimately depends on only the WIMP mass and the choice of the selection criteria.

The calculation of SAE and $SAEx$ require integrations of the WIMP energy deposits and detector response over energy space. The efficiency (as a function of phonon energy) is multiplied by a theoretical WIMP energy deposit spectrum and integrated over energy. Finally, to find the spectrum-averaged efficiency (SAE), we multiply by the exposure time, which is energy independent. In equation form:

$$SAE = T \int_{E_1}^{E_2} \eta(E_p) \frac{dR}{dE_p} dE_p \tag{1.3}$$

$$SAEx = \frac{SAE}{\int_0^\infty \frac{dR}{dE_r} dE_r}, \tag{1.4}$$

where E_1 and E_2 in Eq. 1.3 are the lower and upper limits of the phonon energy considered in this experiment, which are 2 keV and 13.1 keV, and the other variables are defined in Section 1.2.

The 90% confidence level upper limits on the WIMP-nucleon interaction cross section is denoted as σ^{90}, and the limit we would expect, on average given, our backgrounds is denoted as $\rangle\sigma^{90}\langle$. Numerically, the "observed" limit in a pseudoexperiment is a function of N, SAE, and B and is given by:

$$0.9 = \int_0^{\sigma^{90}} P(N|\sigma, SAE, B) d\sigma, \tag{1.5}$$

where N is the number of events observed in an experiment or pseudo-experiment, and $P(N|\sigma, SAE, B)$ is the probability for observing N events, given an SAE, σ, and B. For all probability distributions we assume a Poisson distribution of the "observed" number of events.

The Bayesian[14] interpretation of this limit is that, given N observed events, we are 90% confident that the actual interaction cross section is less than the σ^{90}. Said another way, all interaction cross sections above that limit are excluded with 90% certainty. The limit is actually reported for both the cross section and sensitivity, so we solve Eq. 1.1 or 1.2 for σ or ξ and average over a distribution of N, which depends on B. This is discussed in more detail in Section 3.1.

We note that in Equation 1.5 and in the 2014-Low Threshold version of the analysis, the estimation of the sensitivity uses the estimate of the background to generate the number of "observed" events in the pseudoexperiments, but the limits set using those pseudoexperiments are done as if there are no backgrounds.

The limit is just the lower bound on the amount of signal that could have, on average, produced the observed number of events. Put another way, no background subtraction was performed. A possible complication is that the criteria are based on an optimization of the sensitivity, which means that a different method, e.g. one that takes the background into consideration, may yield a different set of criteria. Another difference is that the actual analysis does not use a cut-and-count method, but another method which is described in Chapter 4.1.

Separate from the considerations above is that in the 2014 analysis the detectors differed from one another due to randomness in the construction process. The effect of these differences is that each detector performs a partially independent search, the results of which are later combined. This thesis will demonstrate that the method used to combine these results can significantly effect the overall sensitivity. In the 2014 analysis, each search is optimized on four WIMP masses, 5, 7, 10, and 15 GeV/c^2. Each search uses the same preselection requirements, but their BDTs are tuned independently, and there is a different threshold for each (no attempt is made to rescale the detectors in such a way to make their BDT thresholds uniform). Eight detectors are used in this analysis because seven physical detectors were deemed suitable, one of which suffered an electrical short between data taking runs. That detector and its exposure were divided into two time periods and treated like two effective detectors [6]). These 8 effective detectors have BDT thresholds which minimize the expected 90% confidence level limit on the WIMP-nucleon cross section (without background subtraction) at each of the chosen WIMP masses. Therefore, the final BDT selection criteria for the complete analysis is characterized by 32 scalars.

Due to a variety of reasons, including correlations between the results and the desire for an experiment-wide acceptance, background estimate, and finalized candidate event list, the results of the 32 BDT selections are not considered independently in the actual experimental result. The final limit was found using selection criteria constructed from the logical .OR. of the final, tuned BDTs for the four masses on each detector. In other words, events which pass any of the 32 BDTs for any of the detectors are in the final data sample. The efficiency and backgrounds of all detectors are combined with appropriate weights and a limit is set based on the combined efficiencies and the expected number of background events across all detectors. Importantly, which detector an event occured in is not considered, only the total number which passed the cuts. This detail is the focus of our re-optimization strategy.

Though the BDT thresholds are selected based on optimizing the results for the individual WIMP masses, the final result in the 2014 analysis is calculated differently. The final analysis takes into consideration the energy distribution of the set of events that pass any of the BDTs and uses the optimal interval method (an extension of the "maximal gap method") [8] to set the final limits. The optimal interval method is not considered in this thesis, but involves finding the most likely cross section given the gaps between events in BDT-score space. The method is particularly useful in cases where an unknown but significant background is present. This includes elements of background subtraction, though the approach is different.

1.3.2 Overview of Thesis: Analysis Optimization Methods

This Thesis is broken into a number of chapters which detail the steps of our attempts at a number of new optimization methods for the analysis. In Chapter 2 the previous analysis is recreated. In Chapter 3 we focus on optimizing the cross section limit further in three main steps.

1. Incorporate background subtraction into the expected sensitivity estimation.

2. Treat the detectors as separate experiments measuring the same observable, then combine to estimate the total sensitivity.

3. Reoptimize the BDT thresholds based on the sensitivity from 2) and find new individual limits.

In Chapter 4 we compare our results to the previous results from 2), by using the .OR. selection criteria used in the original analysis.

The first assumption we consider is the method of incorporating background estimates into the optimization of selection criteria. In the 2014 analysis an estimate on the number of background events that survived the selection criteria over the course of data taking was used to calculate the expected number of events observed in the detector in the no-signal hypothesis, but not used to calculate the sensitivity for purposes of optimization (or the final analysis). This background event count estimate was only in the generation of a

distribution of possible event counts in the detectors called "pseudoexperiments." The overall expected limit was taken as an average of the upper limits set on each pseudoexperiment.

This Thesis begins by utilizing the expected background in the limit setting for each of the pseudoexperiments. The particular method of including the expected background while setting limits is called "background subtraction," and will be described in Chapter 3.1. To summarize, for a particular choice of selection criteria, an estimate of the number of background events passing all selection criteria and and estimate of efficiency are determined, then a set of pseudoexperiments (here simply a number of events N) are simulated with the assumption of no signal. A limit is set based on the results of each pseudoexperiment, which we then average to find an expected limit, a value that informs us of the sensitivity of the selection criteria.

Second, in the 2014 analysis, all detector data were pooled together in such a way that the optimization only considered the total expected background (taken as the sum of all background events across all detectors), and the final analysis only considered the total efficiency for signal to survive the selection criteria (taken as a weighted sum of the efficiencies in each detector), and the set of events observed across all detectors, irrespective of the detector in which it originated. Since the detector response to background and signal differ somewhat [6] we expect that incorporating all of the above information should lead to an improvement in the expected sensitivity of the experiment to WIMP-nucleon scattering. We will investigate this issue in Chapter 3.2.

Third, while the selection criteria used for the final result are chosen by the limit setting method used during optimization, the limit setting method in the published result differed from the method used for optimization. The optimization was performed with a detector-dependent and WIMP-mass-dependent selection criteria, but the final analysis considered the logical .OR. of all events that passes any of the criteria at any WIMP masses in order to obtain a finalized set of candidate events. In Chapter 3.3 we explore the implications of these alterations on the expected sensitivity as a function of mass.

With these considerations in mind, we proceed to explore our alternate analysis techniques in order to improve the overall sensitivity. The first is to consider a "background subtraction", where an expected limit is determined using the number of events in excess of the expected number of background events, rather than the absolute number of events. The second change takes the differing sensitivities of the detectors into consideration when combining their results into an overall expected limit. We perform a new optimization based on these changes and choose new selection criteria which are then tuned using these new assumptions. We then compare this with the expected results for the .OR. calculation in Chapter 4.

We next move on to reproducing the 2014 analysis in Chapter 2.

Chapter 2

Replication of the 2014 Low Threshold Analysis

Here we reproduce the results of the 2014 low threshold analysis with a focus on the optimization of event selection criteria. We present a replication of the original analysis and optimization method and show that we find thresholds consistent with that result. In Section 2.1 we outline the selection criteria for discriminating signal events from background. In Section 2.2 we discuss the the theoretical nucleon and phonon energy spectra for WIMP events. In Section 2.3 we show how the efficiencies for the detectors are calculated. In Sections 2.4 and 2.5 we explain the method for calculating the $SAEx$ and the B respectively. Section 2.6 shows the method for translating these results into the expected sensitivities, and the results. Section 2.7 serves as a cross check of the method with the results of the 2014 analysis.

2.1 Event Selection Requirements

2.1.1 Preselection Requirements

Selection requirements are chosen to discriminate between signal-like events and background-like events in a way that maximizes sensivity. Such optimization is achieved by balancing background rejection with signal acceptance. This analysis contains two primary categories of selection requirements: preselection requirements and BDT seleciton criteria. Preselection requirements are placed before the BDT selection criteria and discriminate against a number of different effects, including random noise and events that are unlikely to be result from nuclear recoils. In addition, there is a selection on the event time, where we only accept events before June 1, 2014. The preselection requirements are described in more detail in Adam Anderson's Thesis[13].

2.1.2 Boosted Decision Tree (BDT) Selection Criteria

A *Boosted Decision Tree* is a machine learning method of discriminating the probable signal events from the probable background events by, in effect, mapping the observed values of the events to a single score[7]. The "boosting" that it refers to is the fact that the algorithm is made of a series of trees, each of which are trained on the residual of the preceding tree.

Each BDT was trained on data from two sets calibration runs, where sources were placed near the detector to characterize its response. A ^{252}Cf sample, a neutron source, provides data for the signal and a ^{133}Ba sample, which is a beta emitter, provides data for the background. These samples were chosen for the responses they generate in the detector, and both sets were multiplied by an energy dependent weight function to make their recoil spectrum appear more like the WIMP energy spectrum and background model, respectively. The BDTs give each event a score which is effectively a (discontinuous) function of the spatial location (z and ρ in the detector's cylindrical coordinates), the phonon energy, and the ionization yield of an event in the detectors which outputs a number in the range [-1,1] with -1 being more "background-like"

and 1 being more "signal-like." On a more technical level, the (unboosted) decision trees takes the space of the four parameters and carves out two spaces which it assigns to background and signal, respectively. The boosting process trains more decision trees on the residuals, which are then summed together with appropriate weights. This analysis did not consider changes to the training of these BDTs.

In a cut-and-count analysis such as this one, all events with a BDT score above a chosen threshold "pass" the cut and are included in the analysis. Changing the threshold alters both the *SAEx* and the expected number of background events. The BDT thresholds used in the PRL are listed in Table 2.1. Detector T5Z3 requires additional consideration because of an electrical short which occured during the data taking period (Dec. 12, 2012), causing the the detector's runtime to be divided into two periods, one before the short(BS) and one after the short (AS). These time periods are treated in most cases as two separate detectors, the exception being the WIMP phonon and recoil spectrum as the short altered none of the detector's physics.

2.2 WIMP Nuclear Recoil Energy Spectrum and Phonon Energy Deposits

In order to determine the sensitivity to WIMP-nucleon interactions (defined in this thesis as setting a 90% confidence level upper limit on the WIMP-nucleon cross section), we first find SAE as defined in Equation 1.3. This requires an estimate of the interaction rate as a function of expected measured recoil energy, $\frac{dR}{dE_r}$, in the Germanium detectors. This estimate is detailed in Richard Schnee's paper[11], as well as Lewin and Smith[12] and it is dependent on, among other things, an estimate of the dark matter density in the Milky Way, the speed of the Solar System through the Galaxy, the mass of a Germanium nucleus, and an assumption of Maxwellian velocity distribution. We then find an event rate $\frac{dR}{dE_r}$ per kg $\cdot$ day per keV in units of 10^{-42} cm^2(cross section) as a function of WIMP Mass and recoil energy. The results are shown in Fig. 2.1 for the four masses considered in this analysis.

While kinematics determines the amount of energy deposited into a nuclear recoil, the detectors measure the energy deposited in the form of phonons, which is a related but distinct quantity. The SuperCDMS detectors measures small changes in temperature using a transition edge sensor-SQUID combination[13]. Phonons are generated when particles interact within the detector, but not all recoil energy becomes phonons, and charges drifting through the detector will produce additional phonons. Because of this, the relationship between recoil energy and phonon energy must be modelled. This detector-dependent conversion between the E_r and E_p is described in detail in the Data Release Guide[15]. Within that guide is the charge model, a fit describing how the charge carriers generate phonons while drifting through the detector material which has been biased with some voltage. The resulting coefficients are fit individually for each detector. A more

Detector	M_{WIMP} = 5 GeV/c^2 BDT Threshold	M_{WIMP} = 7 GeV/c^2 BDT Threshold	M_{WIMP} = 10 GeV/c^2 BDT Threshold	M_{WIMP} = 15 GeV/c^2 BDT Threshold
T1Z1	0.350793	0.688206	0.268291	0.253525
T2Z1	0.180069	0.231413	0.239731	0.430344
T2Z2	0.197627	0.182307	0.338651	0.259913
T4Z2	0.224329	0.306454	0.302919	0.270932
T4Z3	0.182216	0.22599	0.279441	0.364376
T5Z2	0.199065	0.175069	0.294058	0.431485
T5Z3-BS	0.258116	0.298712	0.327549	0.262227
T5Z3-AS	0.291954	0.278486	0.368071	0.340217

Table 2.1: The BDT Thresholds in the PRL analysis. Note that we provide all the digits for future replication.

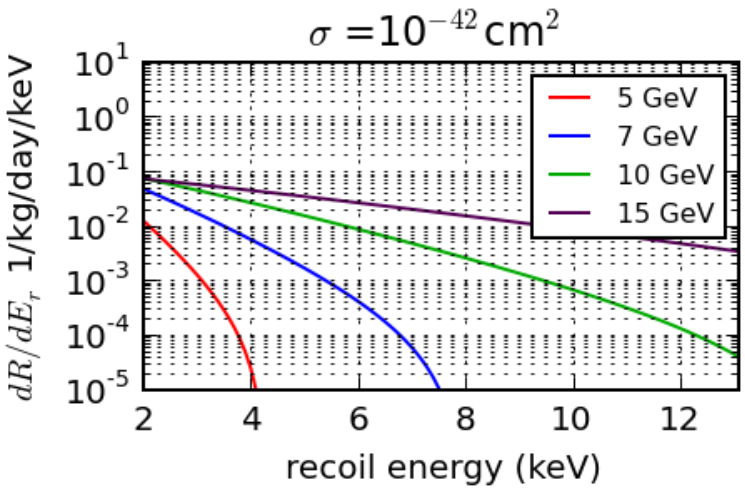

Figure 2.1: The predicted WIMP recoil energy spectrum for various WIMP masses. These curves are dectector independent as the recoil energy is a purely kinematic calculation.

explicit form of the equation for the the relationship between E_r and E_p is given by:

$$\frac{dR}{dE_p} = \frac{dR}{dE_r}(E_p - \frac{4}{3}(\alpha_1 + \alpha_2 E_p + 10^{\alpha_3}\mathrm{erf}(-\frac{E_p}{10^{\alpha_4}}))) \times(1 - \frac{4}{3}(\alpha_2 + 2 * 10^{\alpha_3-\alpha_4}\pi^{-1/2}\exp(-\frac{E_p^2}{10^{2\alpha_4}}))), \quad (2.1)$$

where α_1, α_2, α_3, and α_4 are the charge model coefficients described in the data release guide[15]. The value of $\frac{dR}{dE_p}$ is shown in Fig. 2.2 for one of the detectors. The plots for the remainder of the detectors are shown in in Appendix A.1.

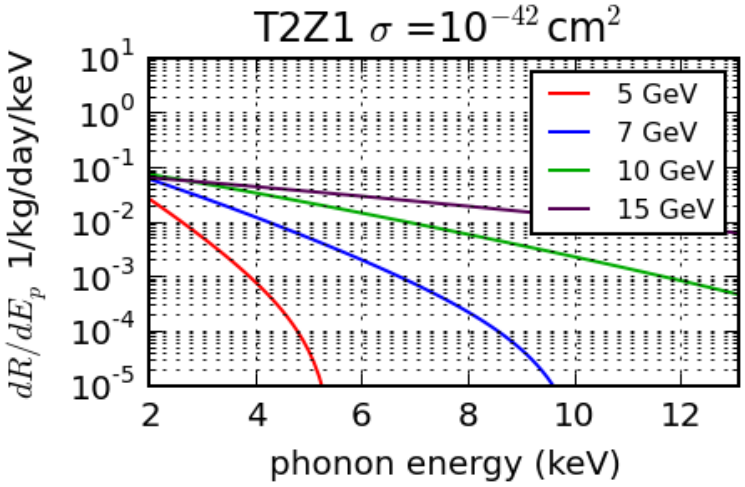

Figure 2.2: The predicted WIMP phonon energy spectrum for various WIMP masses. These curves are detector dependent and the rest are shown in Appendix A.1.

2.3 Efficiencies

The next quantity of interest in Equation 1.3 is the probability that a WIMP, having deposited energy in the detector, is actually counted by the analysis. This term in Equation 1.3 is known as the "Final Efficiency", (η) and is a product of three probabilities as shown in Equation 2.2. The first term is that of a signal event to "trigger" the detector and be written to disk. The second is the "analysis efficency,"[13] which is the probability that, given that an event passes of the first term, an event passes the preselection requirements. The final term is that the of passing the BDT selection criteria, assuming an events passes all the previous requirements. This analysis assumes that these factors comprising the efficiency are uncorrelated, enabling us to simply multiply them together.

$$\eta = (\text{Trigger Eff.}) \times (\frac{\#\text{ passing preselection crit.}}{\#\text{ events triggered}}) \times (\frac{\#\text{ passing BDT and preselection crit.}}{\#\text{ passing preselection criteria}}) \tag{2.2}$$

The methods used to calculate the trigger and analysis (preselection) efficiency are detailed in Ref [13]. Using preselection criteria 2014 result we show the recoil energy dependent efficiencies for an example detector T2Z1 in Fig. 2.3. The results for the final efficiency, η, calculation are shown for detector T2Z1 in Fig. 2.4, with the remainder being found in Appendix A.2. After multiplying the curves in Fig. 2.3 and Fig. 2.4, the final efficiency η is obtained. The result for detector T2Z1 is shown in Fig. 2.5.

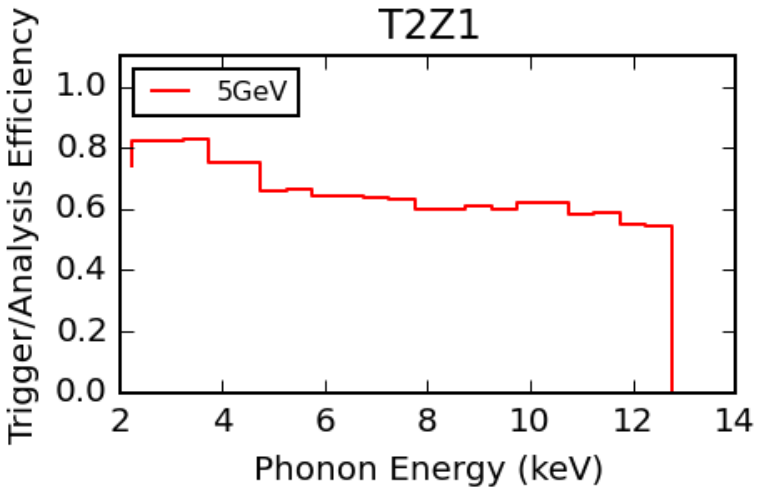

Figure 2.3: Trigger and preselection requirement efficiencies (first two terms of Equation 2.2 for Detector T2Z1. Note that while the mass is indicated, this efficiency is independent of WIMP mass.

We note that while the trigger and analysis (preselection) efficiencies are independent of mass and will not change as the BDT thresholds are altered, the BDT efficiency and the final product will change. For mathematical convenience, the efficiencies are calculated in 0.5 keV bins assuming that their values are constant within each bin. The trigger/analysis efficiencies are calculated at the center of each bin and multiplied by the BDT efficiencies.

2.4 Spectrum Averaged Efficiencies and the Total Exposure

Next, we find *SAE* and *SAEx* using Equations 1.3 and 1.4. Since both the spectrum and the efficiency are functions of phonon energy, they must be convolved in order to obtain a value of how likely a WIMP is to satisfy the selection criteria. The theoretical WIMP recoil energy spectrum is converted into a phonon energy spectrum using Equation 2.1. Next, we multiply the spectrum across the energy bins by the η in

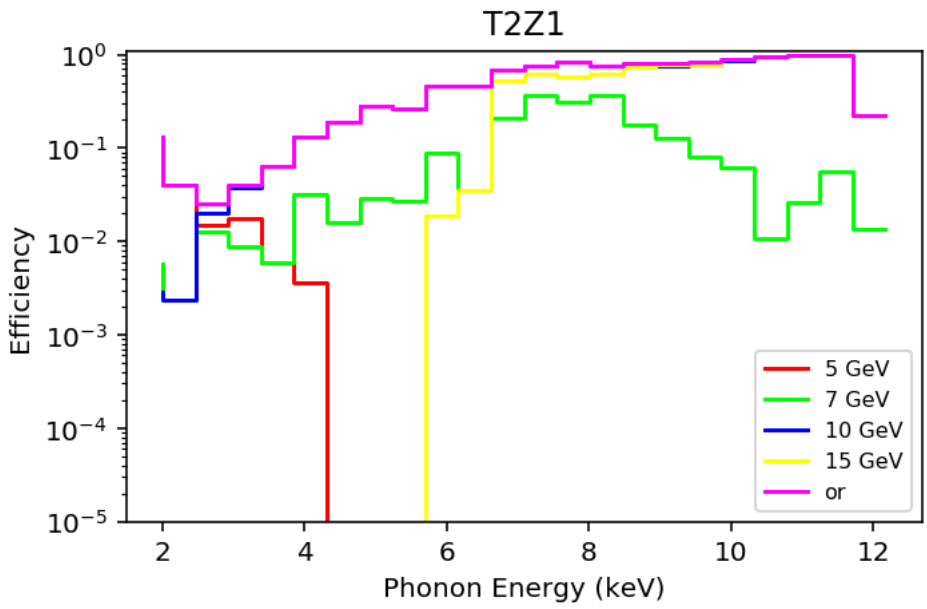

Figure 2.4: The BDT passage fraction for Cf Data for Detector T2Z1 using the BDT thresholds from the 2014 analysis. This is the third term in Equation 2.2. The BDT passage fraction for the remaining detectors is shown in Appendix A.2.

Fig. 2.5, integrate across each bin, and sum the bins across the energy range to get a "spectrum averaged efficiency". The endpoints of the LT analysis thresholds are 2 and 13.1 keV (E_1 and E_2 in Equation 1.3) and are the phonon energy in keV. We then use Equation 1.4 to obtain $SAEx$.

The final variable needed to calculate SAE and $SAEx$ are the exposures T of each of the detectors which are listed in the final column of Table 2.2. For the purposes of being useful to future users, the numbers are not rounded.

2.5 Backgrounds

The number of events from all background sources which pass all requirements is estimated using a set of possible environmental sources weighted by their contamination level and energy spectra. Each background is estimated as the expected number of events from all backgrounds sources that pass the selection criteria over the whole R133 data taking period, or "livetime". The method used to calculate estimated background is similar in some ways to the method for efficiency, but different in others. This is because we aren't simply interested in the fraction of background events that pass the selection criteria, but rather the total number expected over the R133 lifetime.

Since there are different BDT thresholds for each detector i and WIMP mass m, T_{im}, we calculate the expected background B_{im} from all samples with the following equation:

$$B_{im} = S_i \cdot \sum_{j=1}^{n} \mathrm{W}_j \cdot \Theta(\mathrm{BDT}_{jm} - T_{im}), \tag{2.3}$$

where S_i is an overall detector scaling factor, W_j is a weight assigned to each event in a given sample used to estimate the background, and $\Theta(x) = 1$ when $x >= 0$ and 0 when $x < 0$, and the sum is over events. To estimate the background we use ^{133}Ba calibration data, divided into two parts. The BDTs were trained on the former data set, then evaluated on the latter. Weights taken from [13] are applied in order to make the spectrum appear as representative pulses from various backgrounds expected in the detectors

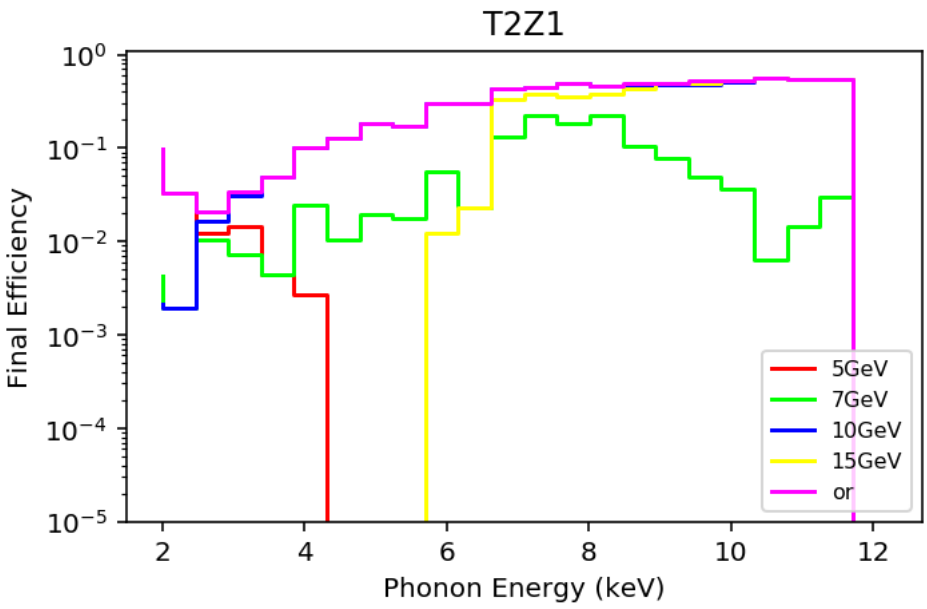

Figure 2.5: The Final Efficiency η for Cf Data for Detector T2Z1 using the thresholds used in the 2014 analysis. This is the third term in Equation 2.2. The η curve for other detectors is shown in Appendix A.2.

and normalized to the expected contamination levels within the experiment. To summarize, each event is sampled from calibration data, and if its BDT value passes a chosen threshold, a weight representing our understanding of the true backgrounds is multiplied by a scaling factor and added to our background event number expectation value.

Since the expected background for each detector is a function of the BDT threshold, and there are 32 different BDTs, 32 separate background estimates are obtained as part of the 2014 analysis. The background is estimated by summing the weights of the events that have a BDT score above the set threshold. Then the number of events used in the simulations are scaled in order to match the correct exposure for R133 (found in Table 2.2), using the variable S_i in Equation 2.3. The background estimates using the PRL BDT thresholds are shown in Table 2.3.

2.6 Expected Sensitivities

To estimate the expected σ^{90} we take the expected backgrounds and efficiencies and use the no-signal hypothesis . Ignoring systematic uncertainties, we further assume that the number of events observed in a pseudoexperiment is distributed according to Poisson statistics, where the probability to observe events N given an expectation value, λ, is:

$$P(N, \lambda) = \frac{\lambda^N e^{-\lambda}}{N!} \tag{2.4}$$

In the no signal hypothesis $\lambda = B$, and our observations are tested against it with the trial expectation value given by $\lambda' = B + \sigma \cdot SAE = B + \xi \cdot SAEx$.

This integral may be inverted with relative ease, as the inverse incomplete gamma function is found in many numerical libraries. The expected limit in the no-signal hypothesis is therefore the σ^{90} that one would expect to obtain when the expected number of observed events is the expected number of background events, $\langle N \rangle = B$. An average value is σ^{90} may be found by drawing values of N from a probability distribution given by the expected background B, or by taking the median value of N and calculating σ^{90} at that location.

Detector	*SAEx* ($M_{WIMP} = 5$ GeV/c^2) (1/kg/days)	*SAEx* ($M_{WIMP} = 7$ GeV/c^2) (1/kg/days)	*SAEx* ($M_{WIMP} = 10$ GeV/c^2) (1/kg/days)	*SAEx* ($M_{WIMP} = 15$ GeV/c^2) (1/kg/days)	Exposure (T) (kg · days)
T1Z1	0	0	1.662	6.111	80.2
T2Z1	0.333	0.146	3.611	5.41	82.9
T2Z2	0.312	2.842	6.802	12.878	80.9
T4Z2	0.011	0.258	3.312	8.716	87.4
T4Z3	0.406	1.759	4.273	10.74	83.8
T5Z2	0.273	1.764	4.523	5.621	82.7
T5Z3-BS	0.004	0.053	0.398	1.829	18.8
T5Z3-AS	0	0.168	1.205	5.627	60.37
Sum	1.339	6.99	25.786	56.932	577.07

Table 2.2: The *SAEx* values and Exposures using the 2014 BDT thresholds for each WIMP mass /detector combination. The number in the column headings corresponds to the WIMP mass at which the *SAEx* is evaluated.

Detector	B_{expected} $M_{WIMP} = 5$ GeV/c^2	B_{expected} $M_{WIMP} = 7$ GeV/c^2	B_{expected} $M_{WIMP} = 10$ GeV/c^2	B_{expected} $M_{WIMP} = 15$ GeV/c^2	B_{expected} .OR.
T1Z1	0	0	0.012	0.013	0.019
T2Z1	1.243	0.008	0.112	0.005	1.351
T2Z2	0.759	0.822	0.148	0.062	1.197
T4Z2	0.007	0.003	0.031	0.017	0.039
T4Z3	1.119	0.678	0.046	0.032	1.413
T5Z2	0.565	0.736	0.142	0.007	1.082
T5Z3-BS	0.02	0.003	0.002	0.013	0.036
T5Z3-AS	0	0.069	0	0.016	0.085
Sum	3.713	2.319	0.493	0.164	5.22

Table 2.3: Estimated number of events from all background sources using BDT thresholds in Table 2.1

Both are valid strategies, and are written as:

$$\langle \sigma^{90} \rangle_{\text{mean}} = \sum_{N=0}^{\infty} P(N|B)\sigma^{90}(N, SAE, B) \tag{2.5}$$

$$\langle \sigma^{90} \rangle_{\text{median}} = \sigma^{90}(N = B, SAE, B) \tag{2.6}$$

In this thesis we use $\langle \sigma^{90} \rangle_{\text{median}}$ as it is computationally less expensive. Note that in the median approach N can be non-integral, which is handled numerically by extending the Poisson distribution with the gamma function.

2.7 Validation of 2014 Results

Having established the method for estimating the sensitivity, we may now demonstrate that the BDT thresholds chosen for the 2014 result are consistent with maximizing sensitivity to WIMPs when no background subtraction is done. Note that while this method is similar to the optimization done in the following chapter, this is performed as a consistency check with the 2014 thresholds.

The sensitivity ξ^{90} is plotted as a function of BDT threshold for a representative detector T2Z2 for a WIMP mass of 10 GeV/c^2 in Fig. 2.6 as the green curve (the other curve will be explained in a later Chapter). The minimum is at the BDT value of 0.35, compared to 0.339 in the 2014 analysis (see Table 2.1). While these

numbers differ (to within a step size of the curve) it is clear that they are within a slowly varying region of the curve, making the results largely indifferent to small variations. The 2014 result also included systematic uncertainties, which this analysis ignored for simplicity. The published 2014 thresholds are shown as the dashed black line, while the calculated minima are shown as the other dashed lines. The 1-bin (bg) curve is the sensitivity without background subtraction, while the 1-bin(sub) curve is with background subtraction (The method of background subtraction is explored in more detail in Chapter 3). The important take-away of this plot is that the sensitivity is robust against small changes in BDT thresholds, and that the minimum is consistent with the BDT thresholds chosen for the 2014 Analysis.

Similar results are found for other mass/detector configurations and are found in Appendix . With this verification we can now test our new ideas for improving the sensitivity and be able to confidently compare with the previous results.

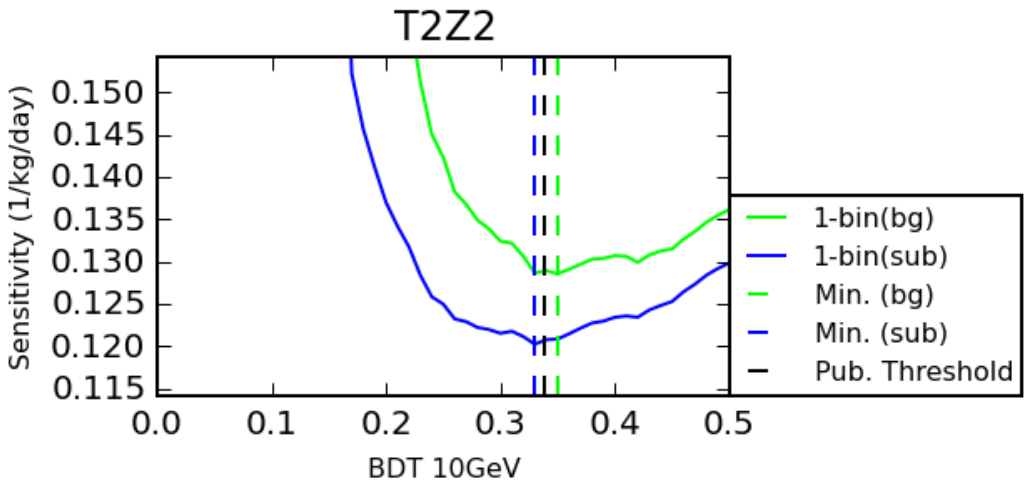

Figure 2.6: The green solid curve shows the expected sensitivity (without background subtraction) as a function of BDT threshold for detector T2Z2 assuming a 10 GeV/c^2 WIMP as is done in the PRL. We will refer to this as the 1-bin optimization in later chapers. The blue solid curve shows the 1-bin optimization for detector T2Z2 assuming a 10 GeV/c^2 WIMP with background subtraction. The vertical lines indicate the various thresholds that one could use. The middle (black) dotted line is the 2014 BDT threshold. The right (green) dotted line is the BDT threshold which minimizes the expected Poisson limit as specified in this Chapter. The left (blue) dotted line is a related limit that will be explained in Chapter 3. This plot indicates that the 2014 BDT threshold has been effectively replicated to a high degree of accuracy. The sensitivity for the remaining detectors and mass combinations is shown in Appendix B.

Chapter 3

Reoptimizing the Analysis

In this Chapter we consider alterations to the optimization strategies used in the 2014 analysis. In Section 3.1 we study the effect on sensitivity and BDT thresholds of subtracting off the expected background. In Section 3.2 we analyze the effect of incorporating event-detector information into the optimization. In both cases an improvement in sensitivity is seen. These sensitivities are compared against the analysis performed in the 2014 result (here also sometimes referred to as the .OR. analysis) in Section 3.3. It happens that this improvement is more pronounced when considering lower WIMP masses for reasons that will be explained below.

3.1 Background Subtraction and BDT Thresholds

Generally, one expects that incorporating more information about an experiment's backgrounds will improve the sensitivity / limit that the experiment is able to set. Having established the method of calculating the Poisson expected sensitivities within the 2014 analysis, we now make the first alteration: considering the method of background subtraction when examining the number of events in a pseudo-experiment. While the λ parameter of the Poisson probability density function in Equation 3.1 is determined by the background estimate, for calculating limits (and integrating the probability), we can subtract out the expected number of background events in order to get an estimate of how many "excess" events could be due to signal. This method is known as "background subtraction".

The limits are set using by integrating the likelihood using Bayes' theorem[16]:

$$P(\sigma|D) = \frac{P(D|\sigma)P(\sigma)}{P(D)}, \tag{3.1}$$

where $P(A|B)$ is the probability of A given B, σ is our observable (conveniently in our case this is the cross section, also written as σ), and D is the data that depends on the value of σ. The values $P(\sigma)$ and $P(D)$ are the *prior distributions* of those values, representing our belief in their likelihood. The value $P(\sigma)$ is of particular interest, as one could choose a uniform prior distribution to reflect total ignorance, or one could incorporate some belief (such as the non-existence of dark matter). While incorporating such information may allow for a more credible limit to be set, often it is desired that the limit have certain properties, one of which is Frequentist "coverage"[14]. This is the property that, given a fixed "true" value of the parameter of interest (here the sensitivity), the data generated by the model will lead to a limit containing that true value with some frequency (e.g. 90%). For Bayesian Poisson limits, it turns out that this Frequentist property is satisfied with a uniform prior distribution.

The value $P(D)$ is the probability of observing the dataset at all, regardless of the value of σ. This is taken to be $P(D) = \int P(D|\sigma)P(\sigma)d\sigma$ taken over all possible values of σ. The left hand side, $P(\sigma|D)$ is the *posterior distribution*, which reflects our updated belief in the value of σ following the observation D.

In our case, and switching over to ξ from σ, $P(D|\sigma)$ is $Pois(N|\lambda(\xi) = SAEx\cdot\xi + B)$. Assuming a uniform

prior, we find our posterior distribution of ξ to be equal to:

$$P(\xi|N)d\xi = \frac{e^{-\lambda(\xi)}\lambda(\xi)^N}{\int_0^\infty e^{-\lambda(\xi)}\lambda(\xi)^N d\xi}d\xi = \frac{e^{-\lambda(\xi)}\lambda(\xi)^N}{\Gamma(N+1,B)}d\lambda, \tag{3.2}$$

where $\Gamma(N+1,B)$ is the upper incomplete gamma function. To find the 90% upper limit, we integrate ξ from 0 to ξ^{90} such that we are 90% confident that the $\xi < \xi^{90}$. Though similar, in Bayesian statistics this interval is not a "confidence interval" but rather the "credible interval."

$$0.9 = SAEx \cdot \frac{\int_0^{\xi^{90}} e^{-\lambda(\xi)}\lambda(\xi)^N d\xi}{\Gamma(N+1,B)} = \frac{\gamma(N+1,\lambda(\xi^{90})) - \gamma(N+1,B)}{\Gamma(N+1,B)} \tag{3.3}$$

$$0.9\Gamma(N+1,B) = \gamma(N+1,\lambda(\xi^{90})) - \gamma(N+1,B) \tag{3.4}$$

$$0.1\Gamma(N+1,B) = \Gamma(N+1,\lambda(\xi^{90})) \tag{3.5}$$

$$\xi^{90}(N,SAEx,B) = \frac{(\Gamma^{-1}(N+1,0.1\Gamma(N+1,B)) - B)}{SAEx} \tag{3.6}$$

$$\langle \xi^{90}_{subtracted} \rangle = \xi^{90}(B,SAEx,B) \tag{3.7}$$

Including information about the backgrounds improves sensitivity, or said another way, decreases the upper limit on the interaction cross section. This of course, assumes that one understands the backgrounds enough to subtract them in the first place.

Using the event selection requirements from the 2014 paper we can calculate the expected limits in our two main scenarios: with background subtraction and without. The results are shown in Table 3.1. It is clear that background subtraction yields a lower expected limit. The effect is more pronounced (better fractional increases in sensitivity) at low masses where the expected background are relatively large. At high masses, the large energy requirements allow us to discriminate more effectively (smaller background) while retaining signal strength. The marginal benefit of this technique is limited at at higher masses because the backgrounds are lower to begin with.

We now have the tools in place to create a new optimization scheme which explicitly contains background subtraction. We consider the effect of two changes: 1) conducting the analysis treating the detectors as independent, and 2) Re-tuning the BDT thresholds using this scheme.

Before making these changes, we estimate the sensitivity's dependence on BDT threshold. Changing a BDT threshold changes both *SAEx* and B, e.g. lowering the threshold increases efficiency while letting in more background, so the impact on the sensitivity requires careful tuning. Since the *SAEx* and B are calculated for a range of BDT thresholds for each detector, one may consider the relationship as being parameterized by BDT threshold. The relationsip is shown in Fig. 3.1 for one detector/WIMP mass combination, with the remainder in Appendix C.2. The 2014 threshold is shown as the black dashed line.

The slope of the curve shows the marginal "cost" of eliminating background. A steep slope indicates that removing more background reduces the *SAEx* by a large amount. A shallow slope indicates that background can be effectively discriminated. Note the sharp drop to 0 for *SAEx* past a value of the BDT threshold. This is due to the BDT threshold moving past the BDT value of the last event in the calibration dataset, which can not appear on the log scale. For some detectors, this occurs before the BDT threshold value that appears in the 2014 PRL result, the interpretation being that the detector does not contribute to the WIMP search. This is allowed as we are interested only the results of all detectors in tandem.

Fig. 3.2 shows *SAEx* and background relative to their values at BDT = -1, i.e. before the BDT selection criteria are applied, for one detector/ WIMP mass combination. The rest of the combinations are shown in Appendix C.3. This plot is convenient for showing that the thresholds should be placed roughly where the backgrounds start to flatten out, or else too much acceptance will be lost in the process.

We expect that combining the results of detectors will further improve the sensitivity. In our figures and tables "Bg" refers to the limit found without background subtraction and "sub" refers to the limit with background subtraction. The BDT thresholds are chosen to minimize the expected limit, and the method is as follows: the BDT threshold for a particular detector/WIMP mass combination is varied, and at each point a new sensitivity is calculated. The BDT value which minimizes this sensitivity parameter is the optimal threshold. If the detectors are treated as a single, combined detector for the purposes of this analysis, this optimization can be performed on each BDT (detector and mass) individually.

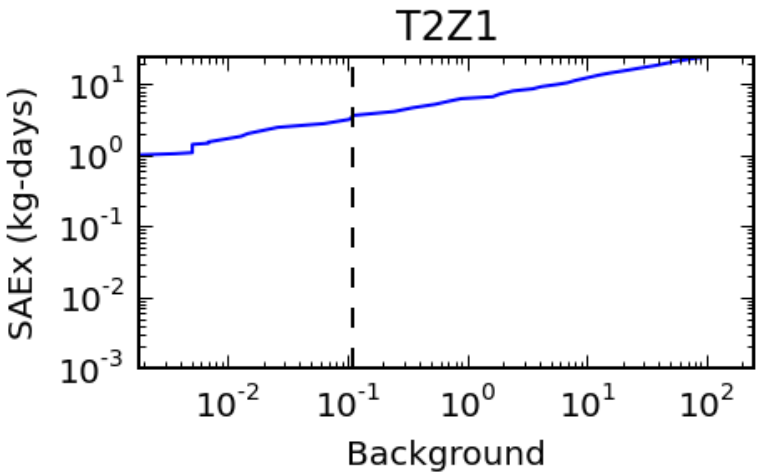

Figure 3.1: Spectrum Averaged Exposure vs. Expected Background, parameterized by BDT, for detector T2Z1 and WIMP mass 10 GeV/c^2. The other combinations appear in Appendix C.1.

We first calculate the sensitivity of the individual detector's limits both with and without background subtraction. The results of this optimization for a detector T2Z2 and a 10 GeV/c^2 were shown in Fig. 2.6. The differences in the thresholds found in the PRL and found by the minimum of the Poisson limits are also shown.

There is a clear improvement in expected sensitivity when using a background subtraction method at all thresholds of the background subtracted limit. When calculated using the PRL thresholds, the sensitivity parameters are shown in Table 3.1, showing an improvement in all cases of the background subtracted sensitivity over the non-background subtracted sensitivity.

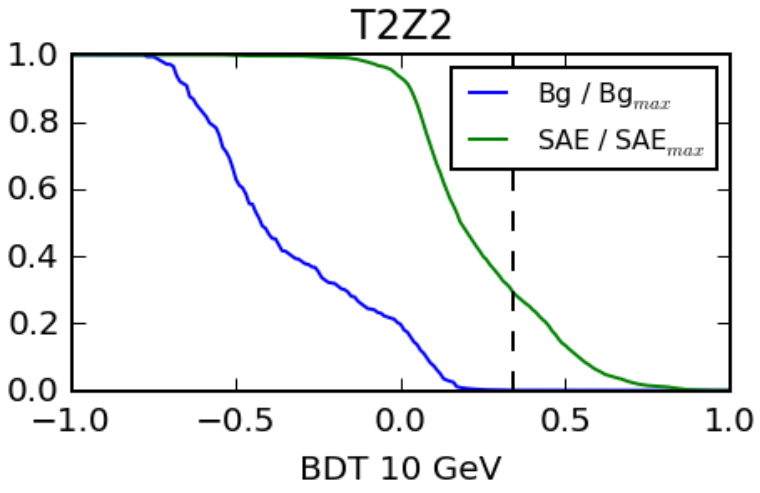

Figure 3.2: Spectrum Averaged Exposure and Expected Background, vs. BDT, relative to Pre-BDT values, for detector T2Z1 and a WIMP mass 10 GeV/c^2. The other combinations are listed in Appendix C.3.

Expected Limit Method	$\langle\xi^{90}\rangle$ (1/(kg· day)) $M_{WIMP} = 5$ Gev/c^2	$\langle\xi^{90}\rangle$ (1/(kg· day)) $M_{WIMP} = 7$ Gev/c^2	$\langle\xi^{90}\rangle$ (1/(kg· day)) $M_{WIMP} = 10$ Gev/c^2	$\langle\xi^{90}\rangle$ (1/(kg· day)) $M_{WIMP} = 15$ Gev/c^2
Background Subtraction	3.484	0.578	0.111	0.044
No Background Subtraction	5.692	0.824	0.121	0.045

Table 3.1: The expected overall sensitivity parameters $\langle\xi^{90}\rangle$ (1/(kg· day)).

3.2 Incorporating Detector Differences into Limit Setting

Next, we consider the difference between two methods: conducting the experiment as if events occured in a larger, homogeneous detector (i.e. a linear combination of the actual detectors, referred to as the 1-bin method/limit) or to treat the events occuring in the detectors as separate data channels. Because each detector has a different background and sensitivity, an event in e.g. T1Z1 is not the same as an event in e.g. T2Z2. Incorporating the target information will allow us to optimize more effectively. The effect of this choice (referred to as the 8-bin method/limit) on the sensitivity at each WIMP mass (5, 7, 10, and 15 GeV/c^2) is studied separately.

Treating the detectors as separate rather than combined sources leads us to rewrite Equation 1.5 as:

$$0.9 = \int_0^{\xi^{90}} P(\xi|\vec{N}, \vec{SAEx}, \vec{B})d\xi, \tag{3.8}$$

where the N, $SAEx$, and B scalars have been substituted by vectors representing the values for each detector. While this integral is merely one-dimensional, inverting it is difficult. Because of this, Monte Carlo integration is used, utilizing the PyMC package[18]. The model is relatively intuitive: for any BDT value each detector has a B and $SAEx$ value, and the ξ was drawn from a uniform prior distribution. The expected number of events for each detector, N_i, is taken to be equal to their respective B_i, and the incomplete gamma function is used to handle non-integral values. We multiply likelihoods to observe N_i across all detectors, and a chain of values for ξ is found through Metropolis-Hastings[17]. The expected value for ξ is taken as the 90% percentile of this chain.

To find the minimum in an 8-dimensional space (assuming a single WIMP mass) we use a gradient descent method. This optimization uses the individual (in terms of WIMP mass) selection criteria optimized simultaneously. [1]

Certain simplications were made to make the Monte Carlo integration less computationally expensive. Instead of integrating over the entire (no-signal hypothesis) probability density function, 31 of the 32 values of the BDT thresholds remained fixed at any one time, and one value was swept over a restricted range of BDT values. The value of the BDT threshold was then moved closer to the minimum value by a value proportional to the fractional improvement in the expected limit. The process was repeated for all 32 BDT thresholds, at which point the iteration was completed, and a new set of $SAEx$, B, were generated using the new thresholds. Conceptually, the thresholds were optimized "simultaneously," in that each iteration uses identical values for the thresholds that are kept static, so the order of operation does not change the results. As the values neared convergence, instead of sweeping the BDT values in increments of 0.01 (coarse), the BDT values were calculated in increments of 0.001 in a range around the current minimum. After 20 iterations, the values were considered to be adequetly converged. [2]

[1] We note that the .OR. selection criteria was used in the 2014 result and is not considered here but will be considered later. In practice, the .OR. can not be optimized neatly, since changing the threshold for one mass has effects on the signal and background at other masses, creating an ambiguity when attempting to find a convergent solution (we will get different results for different masses).

[2] An alternative method, which was not used, is the method of Gibbs sampling. In Gibbs sampling, each parameter of the model is updated in order. This means that, in our example, the 8 BDT values for a given limit, the first is updated based on the current values of the remaining 7, then the second is updated based on the new value of the first, and the old values of those following, and so on. This is typically combined with a "simulated annealing" step which gradually reduces the distance covered as the sampling continues. The annealing allows the sampler to find a global minimum quickly and stay there, which is helpful when many local minima exist.

The sensitivity near the minimum found by the last iteration of the gradient descent is shown in Fig. 3.3 for a single detector/mass combination. The plots for the remaining detector/mass combinations can be found in Appendix D.1. This indicates the location of the optimized thresholds. In Fig. 3.4, we can see the result of the optimization zoomed out over a larger domain of BDT values, and compared to the original 8-bin sensitivity (i.e. using 2014 thresholds). This shows us that the thresholds were optimized for 8-bin limits, as expected. A particular effect that may be observed is that the 8-bin limits are far less sensitive to any one particular detector than the 1-bin. One may say that this makes the result "robust" to any one detector's faults. [3]

The resulting sensitivity after reoptimization is plotted in Fig. 3.5. There it is observed that the 8-bin analysis lowers the limit marginally, but the reoptimization provides further, much larger improvements. These results of the optimization on the sensitivity parameter are shown in Table 3.2. While retaining the same BDT thresholds, there is a slight improvement in sensitivity when switching to the 8-bin method. However the biggest gains are seen following reoptimization with the 8-bin method (i.e locating new thresholds which maximize sensitivity). For 5 Gev/c^2 WIMPs, the sensitivity improves by 35%. For 15 GeV/c^2 WIMPs, the sensitivity improves by 2%.

In Table 3.3 we compare the BDT threshold values selected by our optimization to the PRL BDT thresholds, along with the corresponding (individual) $SAEx$ and B in both cases. In every case the BDT thresholds are relaxed to allow higher acceptance (and background). This analysis indicates that in this parameter range signal acceptance improves the analysis more than background harms it.

We ultimately conclude that factoring the detector that an event occurred in yields improvements in sensitivity. This improvement is most pronounced at low masses, where the backgrounds are far more significant. Next, we compare these results to those from the 2014 analysis.

M_{WIMP} (GeV/c^2)	$\langle\xi^{90}\rangle$ (bg unsubtracted) (PRL BDT thresholds) (1/kg/days)	$\langle\xi^{90}\rangle$ (bg subtracted) (PRL BDT thresholds) (1/kg/days)	$\langle\xi^{90}\rangle_{\text{8-bin}}$ (PRL BDT threhsolds) (1/kg/days)	$\langle\xi^{90}\rangle_{\text{8-bin}}$ (Reoptimized BDT thresholds) (1/kg/days)
5	5.692	3.484	3.453	2.191
7	.824	.578	0.569	0.427
10	.121	.111	0.110	0.103
15	.045	.044	0.044	0.041

Table 3.2: The 8-bin expected sensitivity parameters $\langle\xi^{90}\rangle$ before and after reoptimization.

3.3 Sensitivity with the .OR. Cut

While our optimization improves the sensitivity at each test WIMP mass, the final analysis published in the PRL uses BDT thresholds chosen based on a no-background-subtraction cut-and-count expected limit method based on results for separate WIMP masses, then sets the cross section limit σ using the full set of events which pass the .OR. of all the BDT selection requirements using the optimal interval method. This allows for setting the cross section limit as a smooth function of WIMP mass. We once more note that the final *observed* limit was found using the optimal interval method on the events which were actually measured in the detectors regardless of where they came from. This is distinct from the methods employed for optimization, which rely upon pseudo-experiments and *a priori* expectations.

Fig. 3.6 shows the effect of the options mentioned above on the expected cross section limits vs. WIMP mass. For any version of the analysis the .OR. cut, which uniquely determines the backgrounds and the acceptances, gives a unique expected limit for each WIMP mass, assuming either no background subtraction or with background subtraction. The independent variable is the WIMP mass in the spectrum calculation, and results in a set of continuous curves. The individual mass results are shown as dots with their configurations

[3]Fig. 3.4 is the result of a previous, "coarse" minimization. Due to the long timescales involved in the Monte Carlo Integration, the search window was shrunk to .025 on either side of the previous steps minimum. So Fig. 3.3 represents the actual final step, while Fig. 3.4 has approximately the same thresholds, and is used to illustrate the degree of improvement.

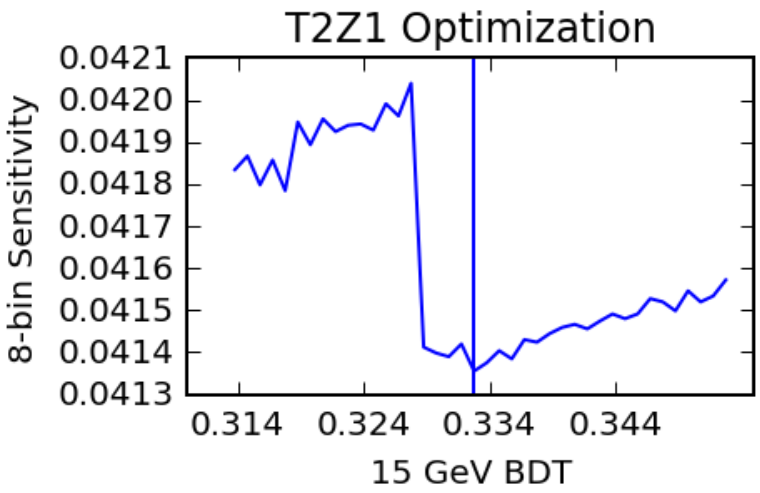

Figure 3.3: The 8-bin optimization (fine grain) shown as the expected sensitivity as a function of the 15 GeV/c^2 WIMP mass BDT score for the T2Z1 detector. The rest of the plots for the fine tuning are found at Appendix D.2.

matching their colors. One may think about the results as incremental improvements. We start with the PRL selection criteria, using a single bin, and no background subtraction (blue curve). Including background subtraction improves the expected limit (green curve), in a roughly consistent way across all WIMP masses, both in the individual and .OR. cases. Moving to 8-bin limits (while retaining background subtraction) and maintaining the PRL BDT requirements, the sensitivity improves even further, by approximately 20% at 15 GeV/c^2 and 50% at 5 GeV/c^2 (red curve). The limit is always lower using 8-bin (PRL thresholds) than the 1-bin(sub), but the effect is more pronounced at high masses where the BDT screens background more effectivly. The 8-bin limit using the 2014 BDT thresholds is always better than both 1-bin limits.

Reoptimizing on the individual 8-bin limits at the individual masses has results that are more complicated to interpret (teal curve). Consistent with our expectations, the combined limit of all detectors at each mass gets better when considered alone (i.e. before the .OR.). However, upon considering the .OR. of the mass cuts for each detector for our analysis, the optimization(of the individual masses) only improves the sensitivity at low masses, while losing sensitivity at high masses. This is due to background and signal events adding nonlinearly for the detectors when using the .OR. cut. Estimated background may be reasonable for each mass individually, but become impractically large after the .OR. assumption. For large WIMP masses the cuts are more efficient at high masses, and are more effective at screening background events. The opposite is true at low masses. Executing the .OR. takes background events from where they are not very detrimental (low masses) and puts them where they are expensive (high masses). This is why the sensitivity is improved at low masses, but not at high masses.

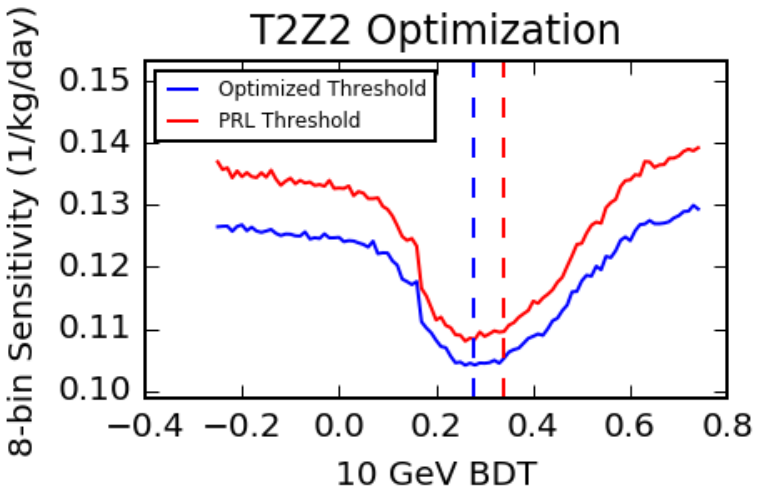

Figure 3.4: The expected 8-bin sensitivity as a function of 10 GeV BDT for detector T2Z2. The red curve shows the result of holding the other 7 detectors at the 2014 thresholds, and the blue curve shows the result of holding the other detectors at the new, reoptimized threshold. Two effects are clear: the sensitivity is improved, and the curvature is now more flat in the vicinity of the minimum. The flatness is a win, because now the analysis is more resistant to any one detector being optimized improperly. The plots for the remaining detectors and masses are found at Appendix D.1.

Detector	*SAEx* (PRL) 1/kg/days	*SAEx* (optimized) 1/kg/days	Background (PRL)	Background (optimized)	PRL BDT threshold	Optimized BDT threshold
T1Z1	1.662	2.397	0.012	0.065	0.268	0.194
T2Z1	3.611	3.658	0.112	0.150	0.240	0.239
T2Z2	6.802	8.558	0.148	0.331	0.339	0.277
T4Z2	3.312	3.391	0.031	0.031	0.303	0.288
T4Z3	4.273	6.216	0.046	0.254	0.279	0.222
T5Z2	4.523	4.919	0.142	0.157	0.294	0.281
T5Z3-BS	0.398	0.718	0.002	0.030	0.328	0.242
T5Z3-AS	1.205	3.208	0.000	0.244	0.368	0.227

Table 3.3: The analysis parameters before and after 8-bin optimization for a WIMP mass of 10 GeV/c^2. The "PRL" thresholds are those used in the 2014 PRL result and the opmizied thresholds are those chosen by gradient descent on the 8-channel Markov-Chain Monte Carlo (MCMC) method. Note that the *SAEx* is evaluated in units of $\sigma = 10^{-42}$ cm^2. Also note that in all cases the same or more background events are allowed as the BDT threshold is lowered. Allowing more background allows more signal acceptance, and this analysis indicates that this increases acceptance outweighs the increase in expected background events.

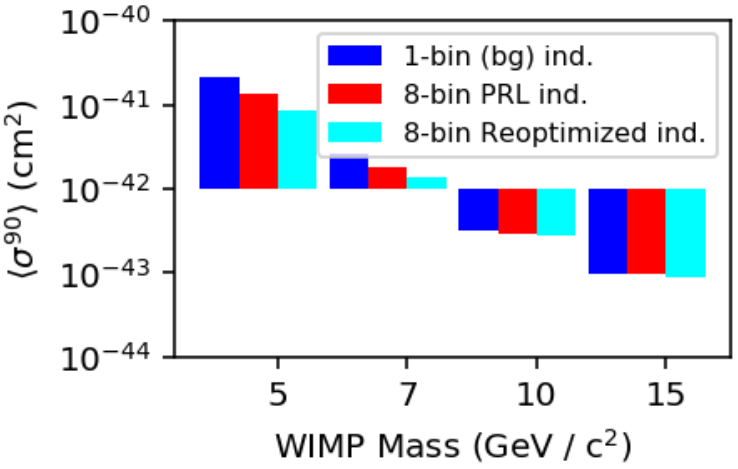

Figure 3.5: A comparison of the expected WIMP-nucleon cross section limits using the various methods specified above. The "ind." denotes that the limit is found at the specific masses denotes in the x-axis. Note the relatively large improvements at low WIMP masses, where the background is the largest. At 5 GeV/cm^2 the ξ^{90} is 5.692 (kg $\cdot$ days)$^{-1}$ in the no-background subtracted cut and count analysis with the PRL BDT selection criteria. This is improved upon in the 8-bin background subtracted reoptimization by 61.5%.

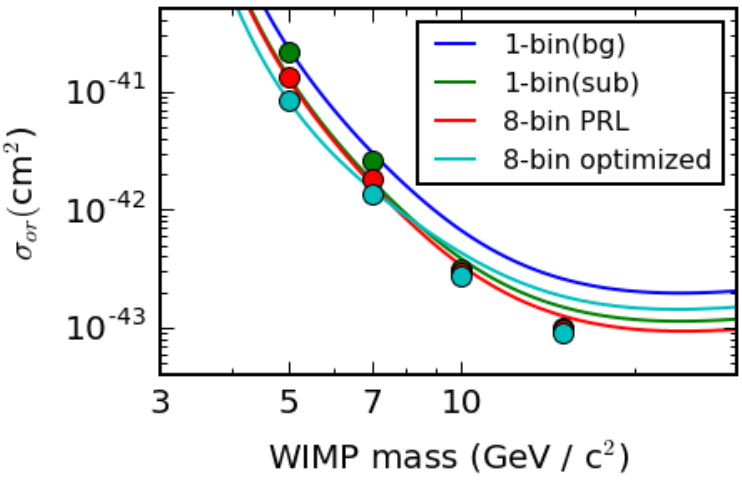

Figure 3.6: WIMP-nucleon expected cross section limits vs. mass (various methods).

Chapter 4

Conclusions

4.1 Summary

In this Thesis we have considered a number of methods for maximizing the sensitivity for the Cryogenic Dark Matter Search Low Threshold Analysis. We replicated the original expected cross section limits from the analysis based on the BDT thresholds used in the 2014 PRL. We then considered two separate improvements: adding background subtraction as well as factoring in the differences in the detectors by analyzing them as separate channels in the experiment, then combining the individual results into one overall expected sensitivity. After reoptimization, this improved the expected limit between approximately 20% at low masses and 7% at the higher masses, relative to the single-bin, background subtracted limit. Altogether, compared to the no-background subtracted 1-bin limits, the expected sensitivity to 5 GeV/c^2 WIMP interactions was improved by 61.5% and the expected sensitivity to 15 GeV/c^2 WIMP interactions were improved by approximately 8.9%, with the intermediate masses falling between those values. Thus, we urge analyzers to consider this methodology for future analyses. In all cases, we found that the sensitivity is improved by loosening the cuts slightly as the increased signal acceptance is more significant than the background leakage, especially at low masses.

While these gains manifest while using Poisson limits optimized for individual WIMP masses, their use is not as clear if one assumes the continued use of the .OR. analysis. The logical .OR. of the BDT cuts applied at different WIMP masses was used for the final result and has the benefit of having a global set of analysis requirements at the cost of reducing the effeciency as a function of mass at each individual point. This is because the reoptimization performed in this Thesis was not designed to specifically optimize the sensitivity under these requirements. Said differently, the .OR. cuts do not significantly affect the sensitivity at low masses, but do affect it significantly at high masses. This may not have a high impact as the high mass WIMP search for SuperCDMS is handled in a separate analysis [19]. It was observed that at low WIMP masses the expected limit improved, while at high masses it deteriorated. We have not studied the effect on the optimal interval method, as it does not allow for its own optimization. Its study would require comparing the results on actual detector data, which is outside the scope of this thesis. It is a method which incorporates background subtraction, so the methods shown in this thesis are expected to have an effect on that result.

In conclusion, this study has illustrated the importance of understanding the backgrounds present in SuperCDMS, as well as the benefits of considering the differences between the detectors. This must be compared against the simplicity of the .OR. cut as well as the sensitivity loss at high mass following the procedure outlined in this thesis.

Bibliography

[1] Roos M, "Dark Matter: The evidence from astronomy, astrophysics and cosmology." arXiv:1001.0316 [astro-ph.CO]

[2] Sofue Y and Rubin V., "Rotation Curves of Spiral Galaxies." *Annual Review of Astronomy and Astrophysics* 39: 137-174, 2001

[3] Clowe D. *et al.*, "A Direct Empirical Proof of the Existence of Dark Matter." *The Astrophysical Journal Letters* 648, 2006

[4] Baer H., "Dark matter production in the early Universe: beyond the thermal WIMP paradigm." *Physics Reports* 555:1-60, 2015.

[5] Skordis, C. *et al.* "Large Scale Structure in Bekenstein's theory of relativistic Modified Newtonian Dynamics." *Phys.Rev.Lett.* 96, 2006.

[6] Agnese T, *et al*, "Search for Low-Mass Weakly Interacting Massive Particles with SuperCDMS. " *Phys.Rev.Lett.* 112, 2014

[7] Roea B, Yanga H, Zhub J, "Boosted Decision Trees, A Powerful Event Classifier." Department of Physics, University of Michigan. 2005.

[8] Yellin S., "Extending the Optimal Interval Method." arXiv:0709.2701 [physics.data-an]. 2007.

[9] Lindhard, J, "On the properties of a gas of charged particles." *Danske Matematisk-fysiske Meddeleiser.* Det Kongelige Danske Videnskabernes Selskab. 1954.

[10] Ahmed, Z. *et al*, "Search for inelastic dark matter with the CDMS II experiment." *Phys.Rev. D83*, 2011

[11] Schnee, R, "Introduction to Dark Matter Experiments." *In Physics of the Large and Small: Proceedings of the 2009 Theoretical Advanced Study Institute in Elementary Particle Physics* 629-681. arXiv:1101.5205 [astro-ph.CO]. 2010.

[12] Lewin, J, Smith, F, "Review of mathematics, numerical factors, and corrections for dark matter experiments based on elastic nuclear recoil." *Astroparticle Physics* 6: 87-112, 1996.

[13] Anderson, A, "A Search for Light Weakly-Interacting Massive Particles with SuperCDMS and Applications to Neutrino Physics." MIT. 2015 .

[14] Patrignani, C, *et al*, (Particle Data Group), Chin. Phys. C, 40, 100001, Ch. 39. (2016).

[15] SuperCDMS Low Energy Data Release [online] University of California, Berkeley. http://cdms.berkeley.edu/data_releases/SuperCDMS_R133_LT/ data_release_guide.pdf

[16] Bayes, T , Price, R, "An Essay towards solving a Problem in the Doctrine of Chance. By the late Rev. Mr. Bayes, communicated by Mr. Price, in a letter to John Canton, A. M. F. R. S." *Philosophical Transactions of the Royal Society of London* 53: 370–418, 1763

[17] Robert, P, "The Metropolis Hastings Method." arXiv:1504.01896v3 [stat.CO]. 2015

[18] *PyMC*. N.A: PyMC-Devs. 2013.

[19] R Calkins and SuperCDMS collaboration 2016 J. Phys.: Conf. Ser. 718 042009

Appendix A

Spectrum, Efficiencies, and Backgrounds

A.1 Phonon Spectrum

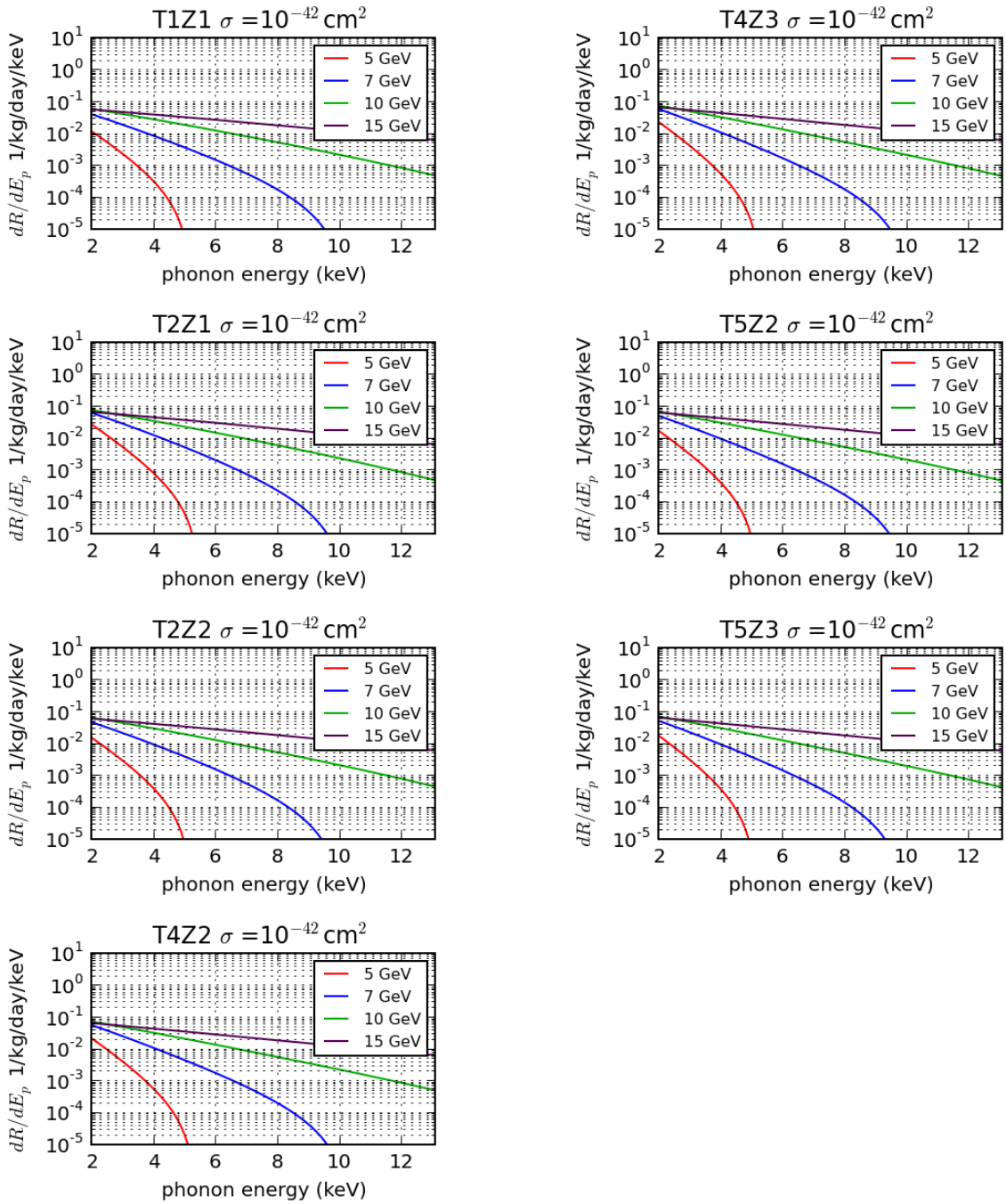

Figure A.1: The WIMP phonon energy spectrum for the various detectors.

A.2 Efficiencies

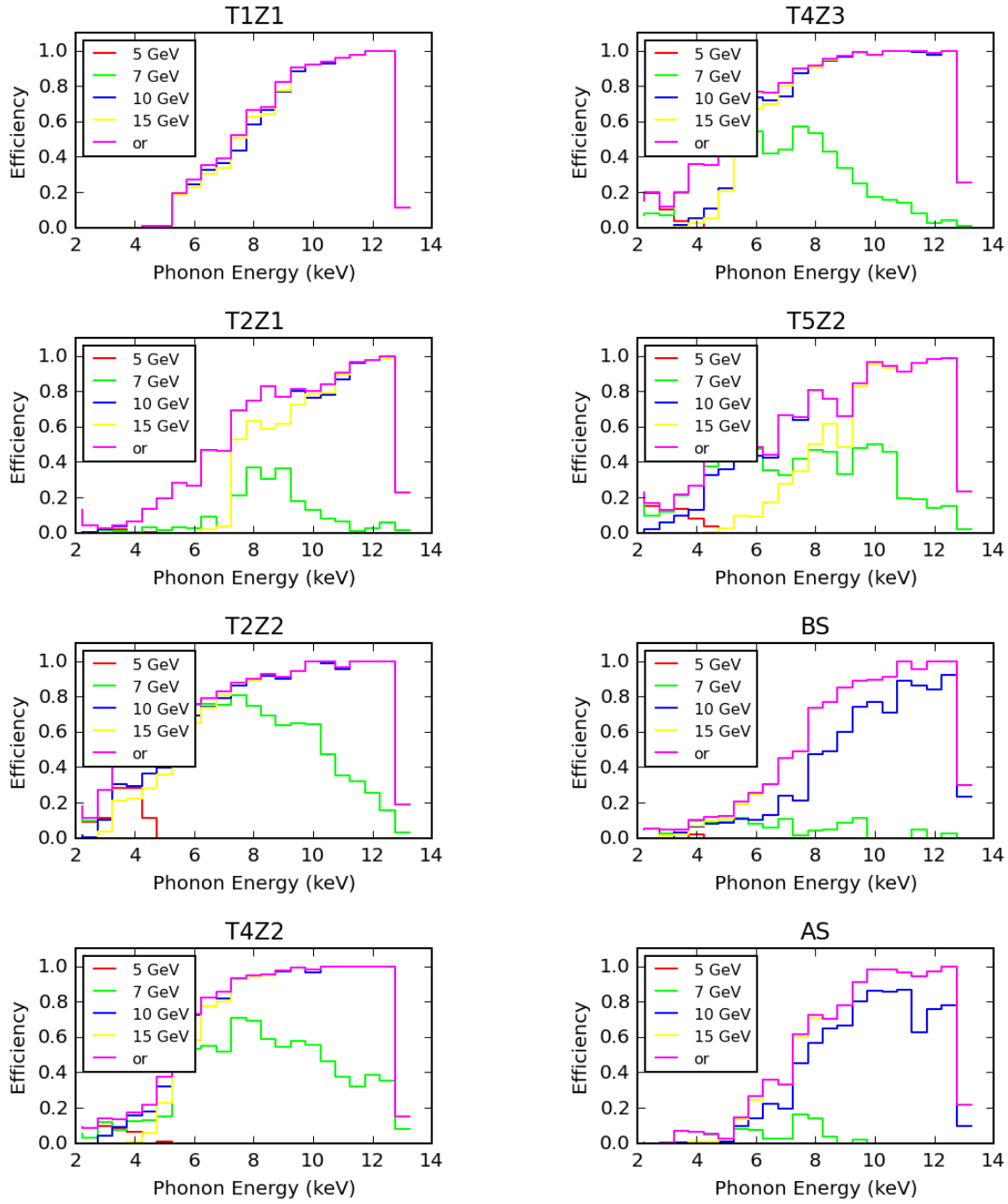

Figure A.2: The detector effiencies (BDT passage fraction) for all the detectors as a function of phonon energy.

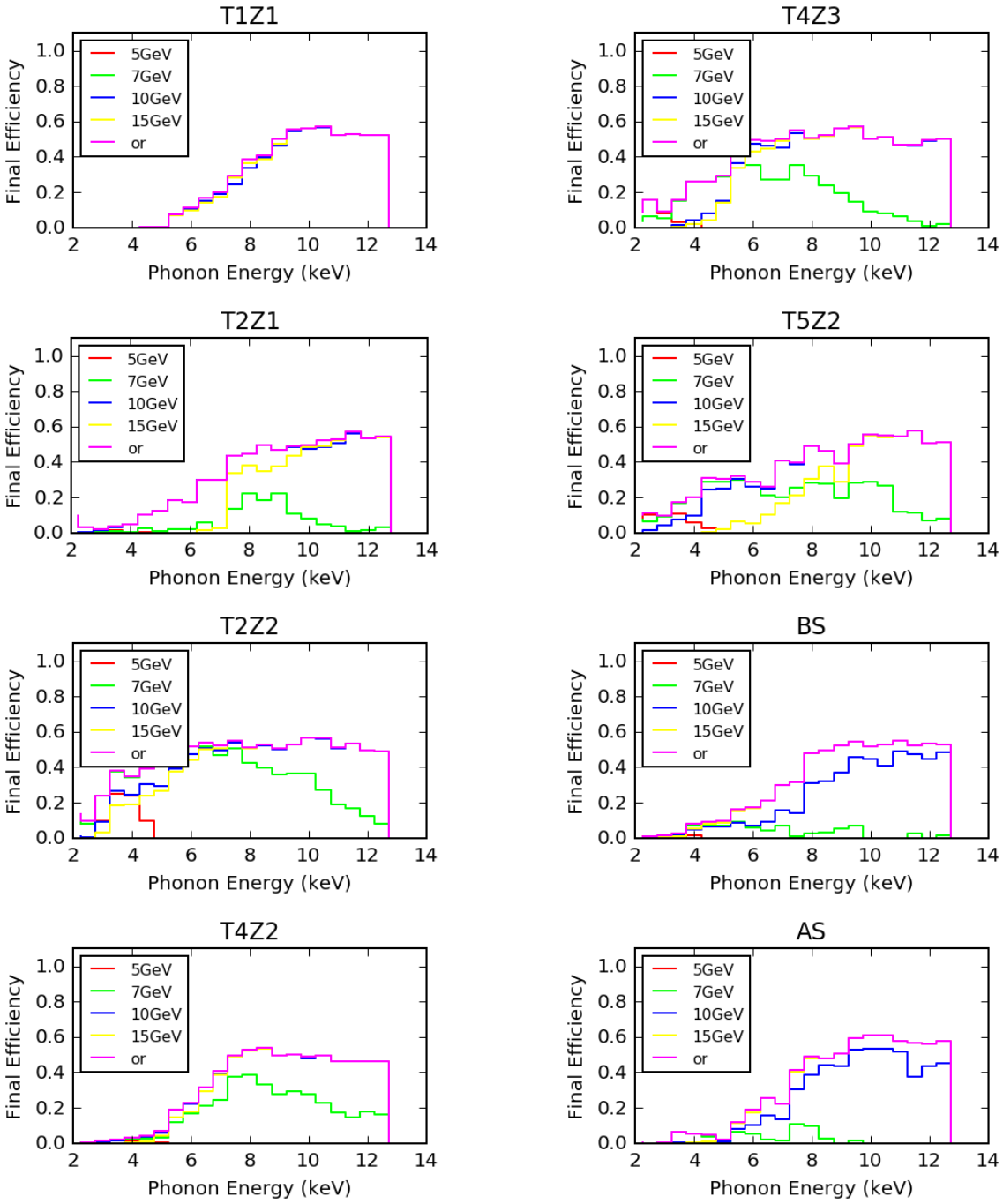

Figure A.3: The final efficiency (with analysis and trigger efficiencies) for all the detectors as a function of phonon energy.

Appendix B

Sensitivities with and without Background Subtraction

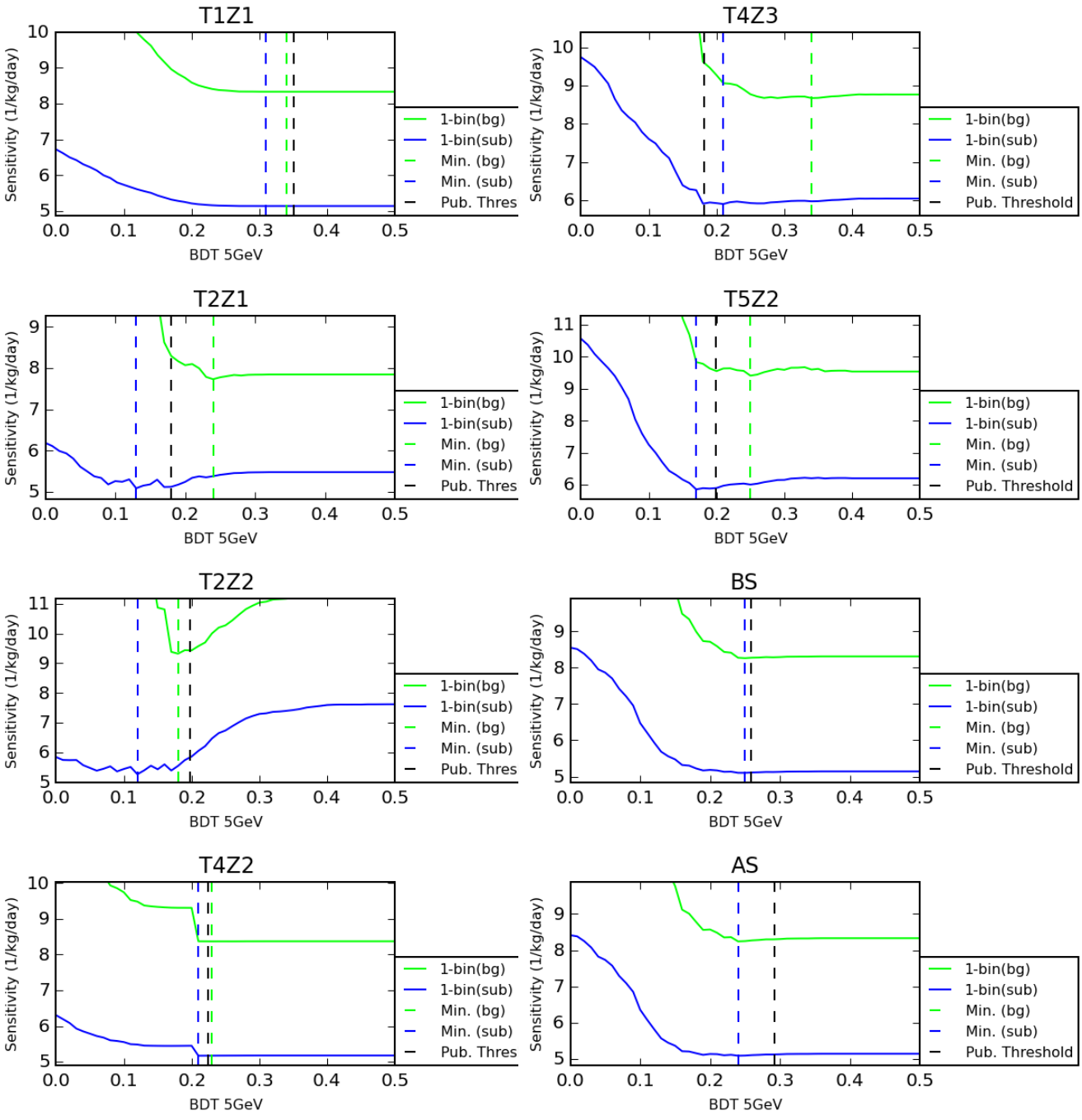

Figure B.1: The expected sensitivity parameter $\langle \xi^{90} \rangle$ (with and without background subtraction) for a 5 GeV WIMP for all detectors as a function of BDT threshold.

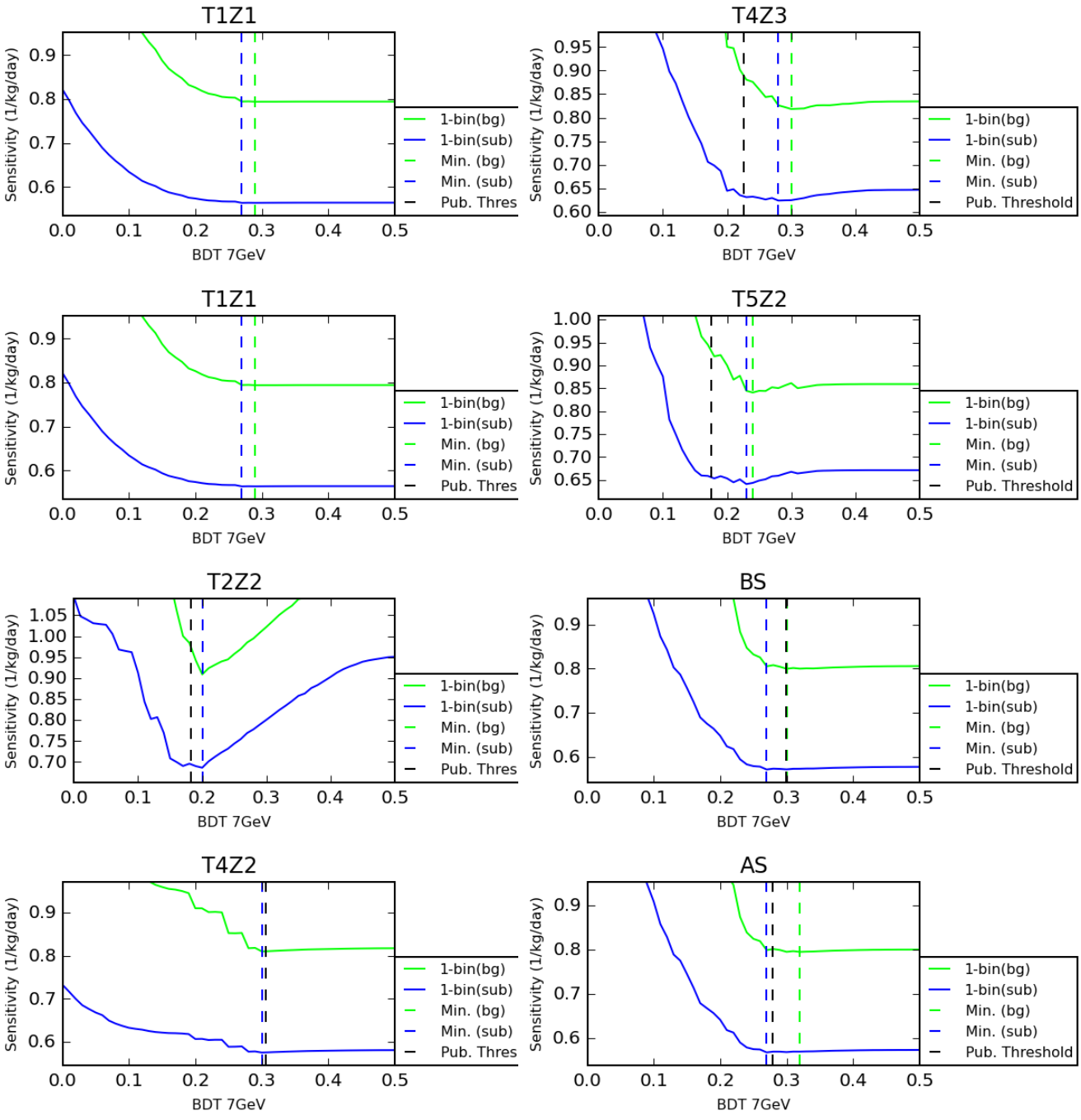

Figure B.2: The expected sensitivity parameter $\langle \xi^{90} \rangle$ (with and without background subtraction) for a 7 GeV WIMP for all detectors as a function of BDT threshold.

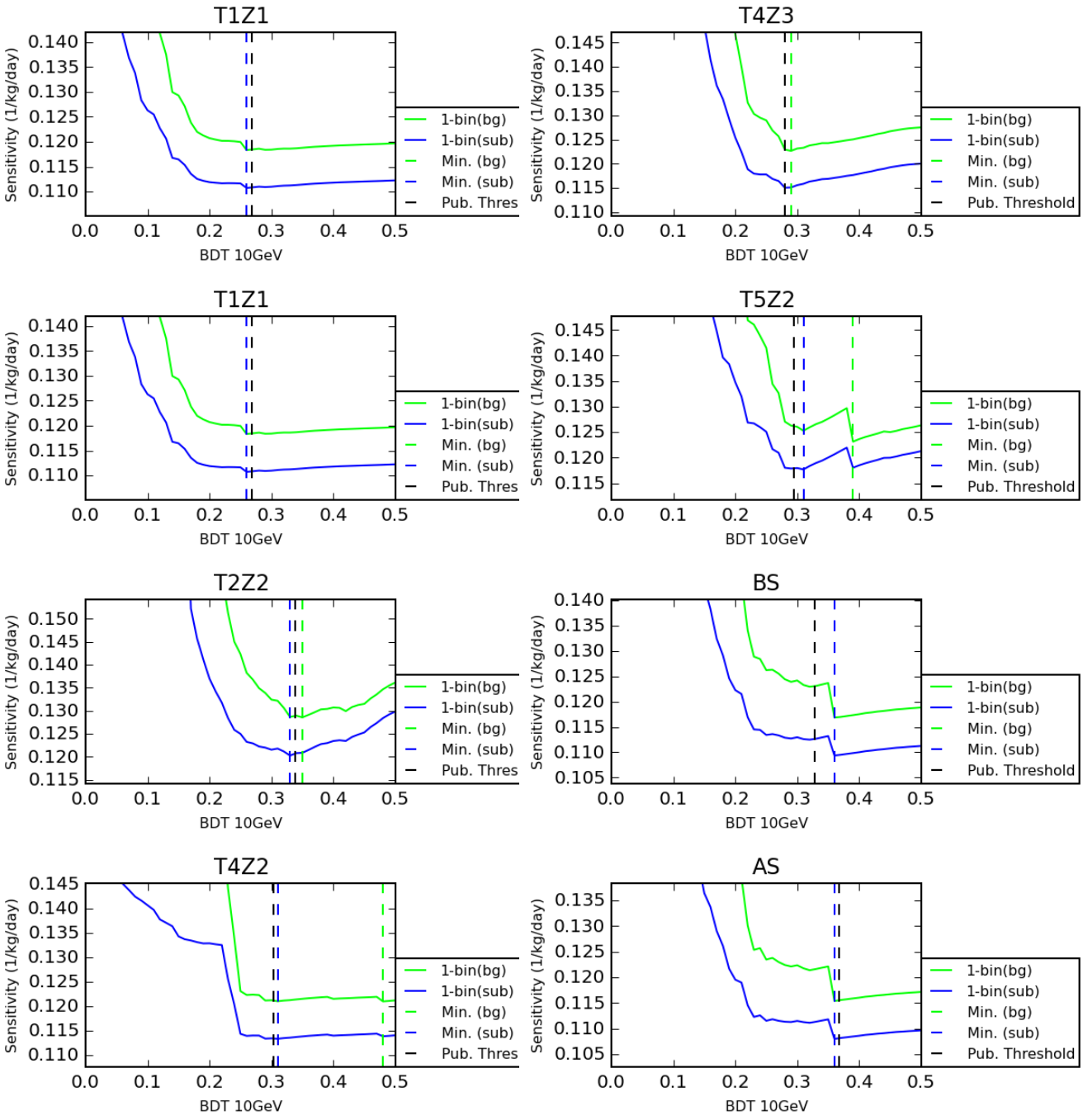

Figure B.3: The expected sensitivity parameter $\langle \xi^{90} \rangle$ (with and without background subtraction) for a 10 GeV WIMP for all detectors as a function of BDT threshold.

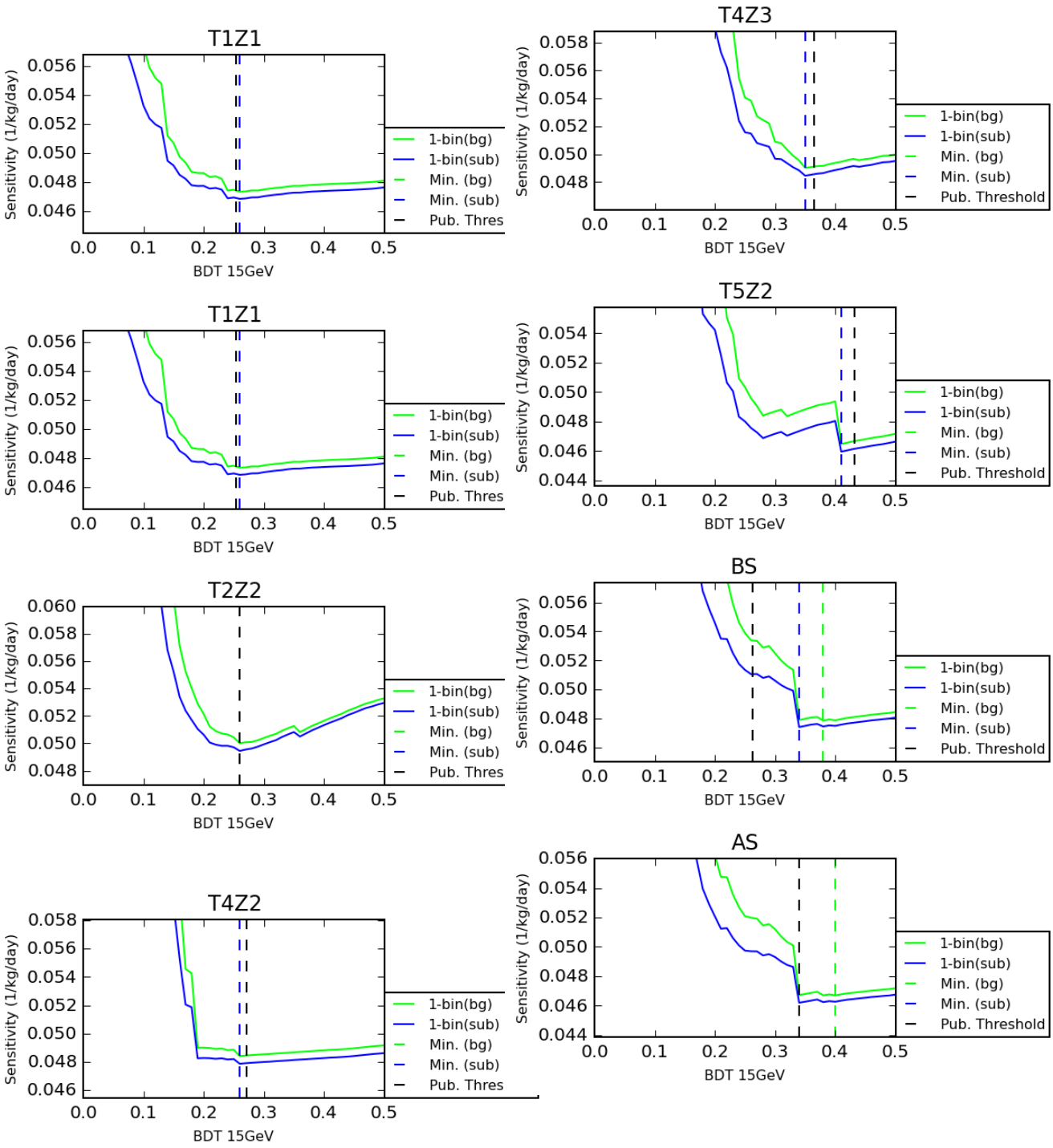

Figure B.4: The expected sensitivity parameter $\langle \xi^{90} \rangle$ (with and without background subtraction) for a 15 GeV WIMP for all detectors as a function of BDT threshold.

Appendix C

BDT Threshold Dependence

C.1 SAEx vs. BDT

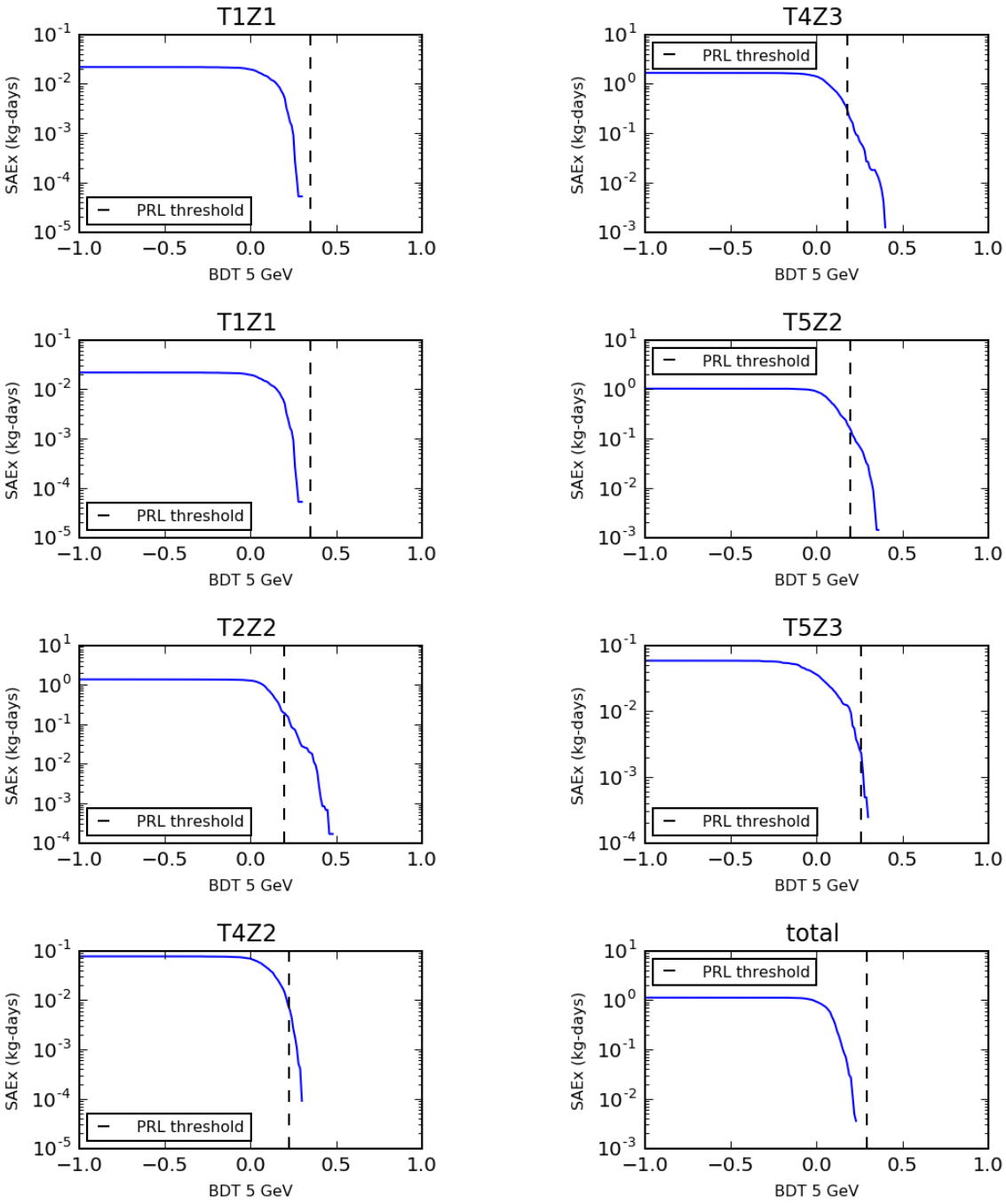

Figure C.1: The spectrum averaged exposure vs. BDT threshold for a 5 GeV WIMP for all detectors.

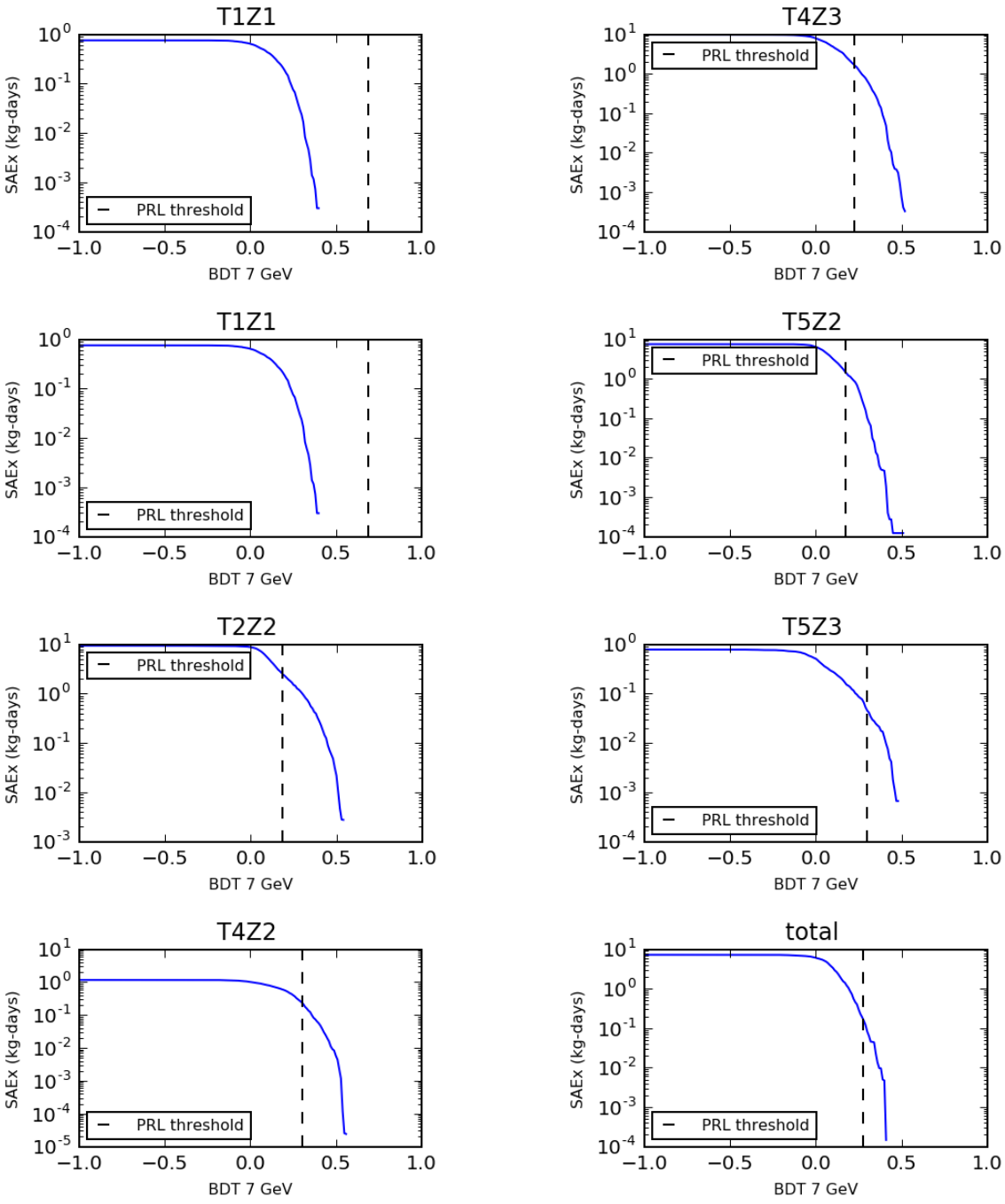

Figure C.2: The spectrum averaged exposure vs. BDT threshold for a 7 GeV WIMP for all detectors.

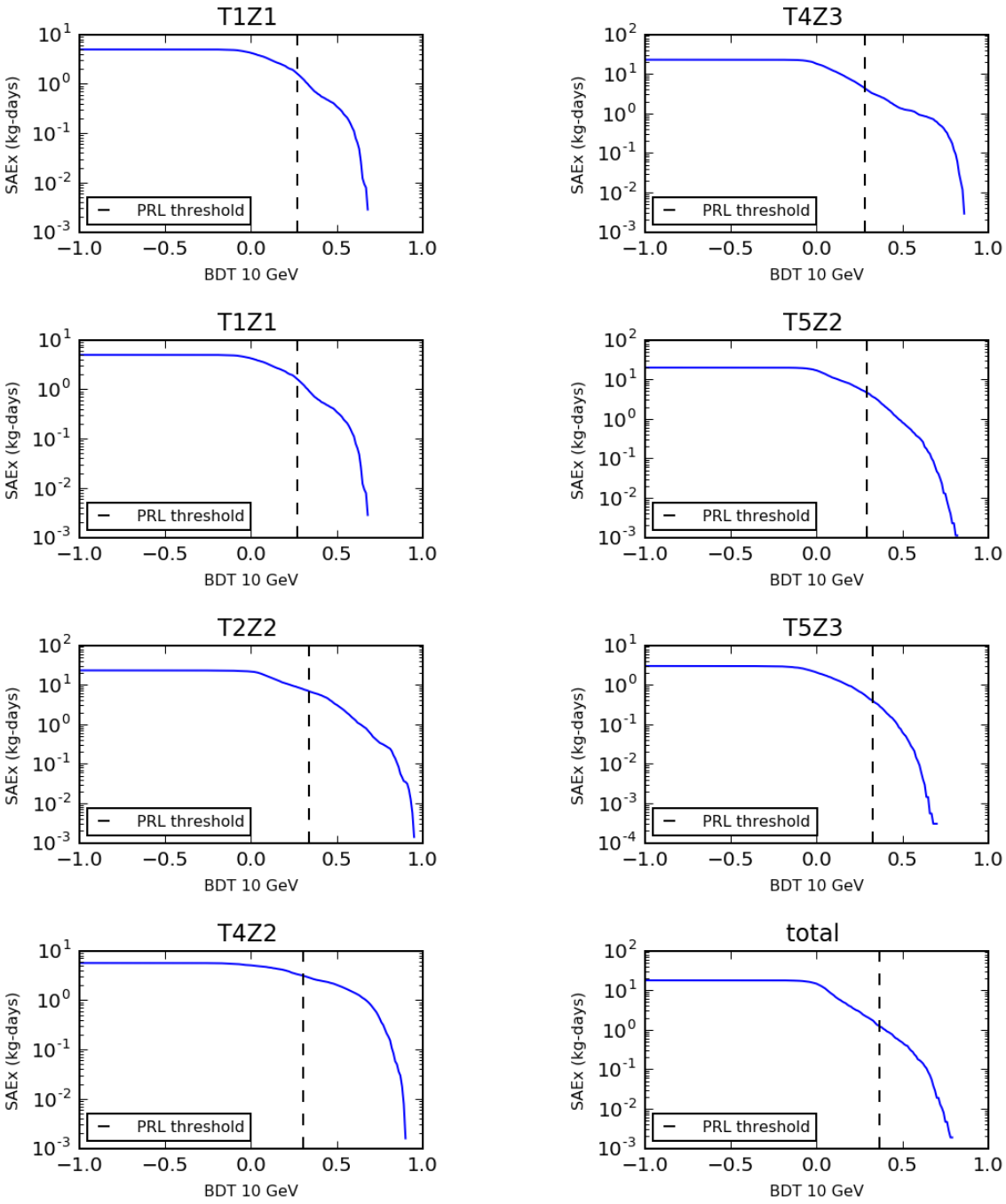

Figure C.3: The spectrum averaged exposure vs. BDT threshold for a 10 GeV WIMP for all detectors.

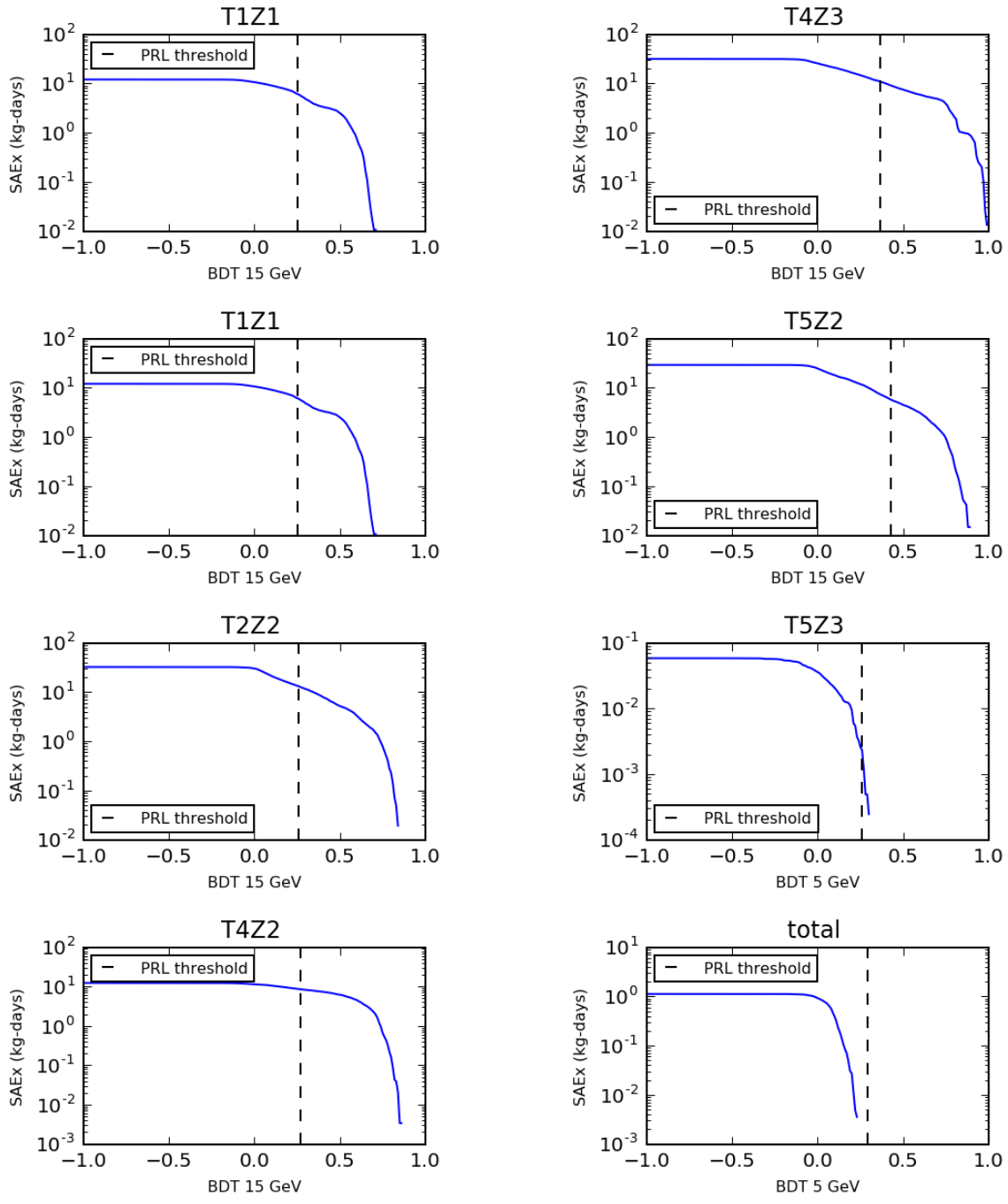

Figure C.4: The spectrum averaged exposure vs. BDT threshold for a 15 GeV WIMP for all detectors.

C.2 Background vs. SAE, BDT-parameterized

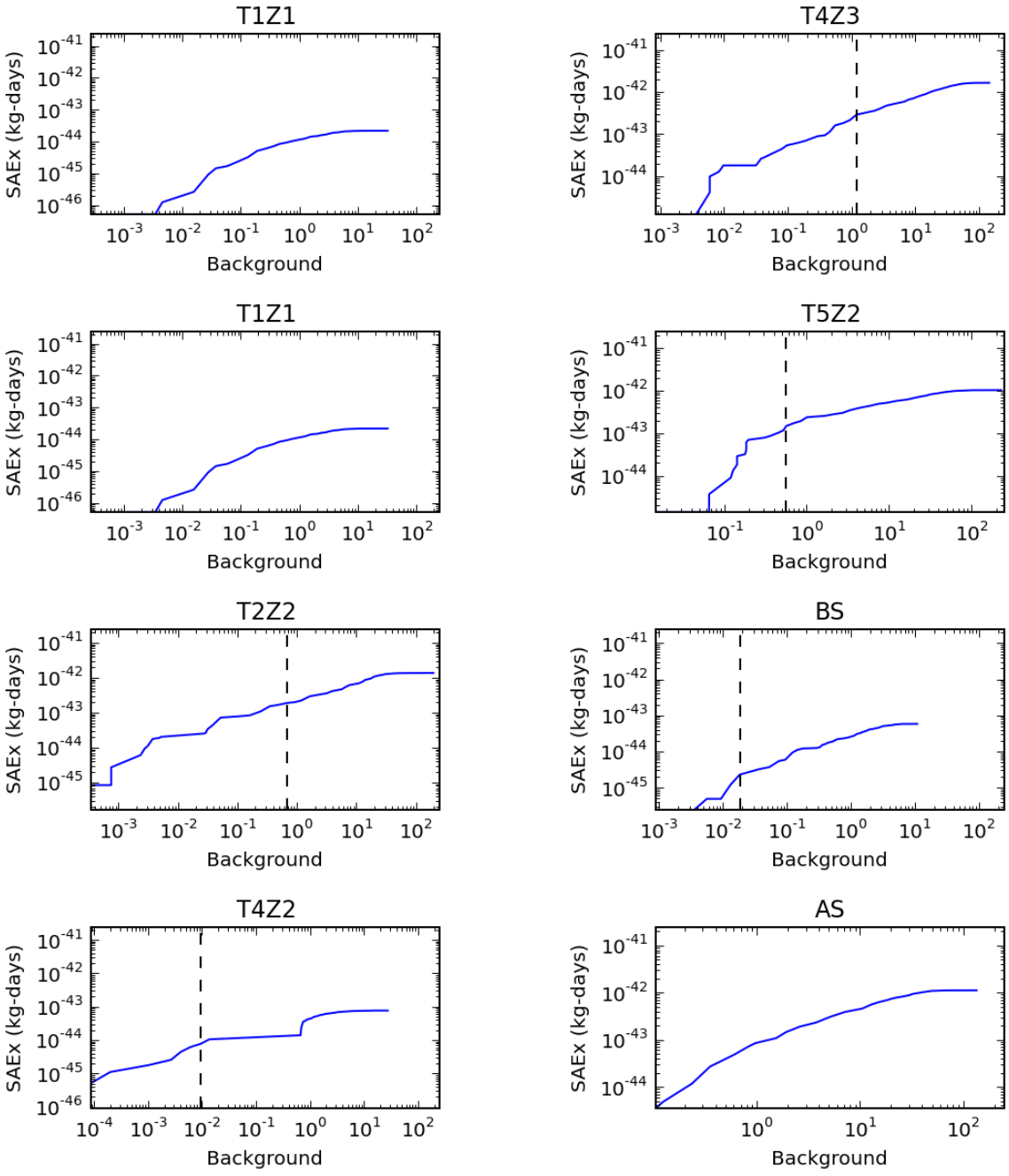

Figure C.5: The expected number of background events vs. *SAE* for a 5 GeV WIMP.

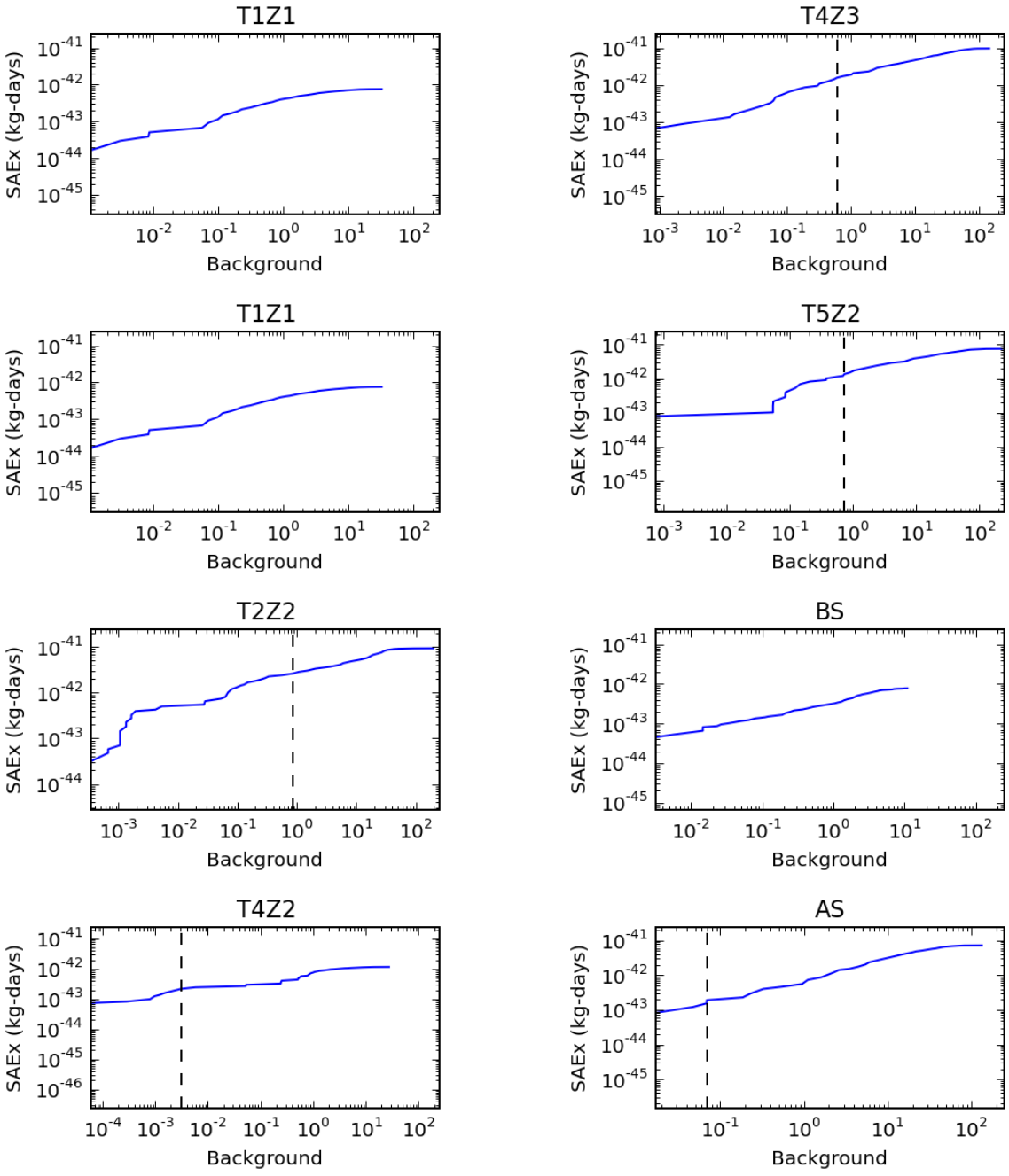

Figure C.6: The expected number of background events vs. *SAE* for a 7 GeV WIMP.

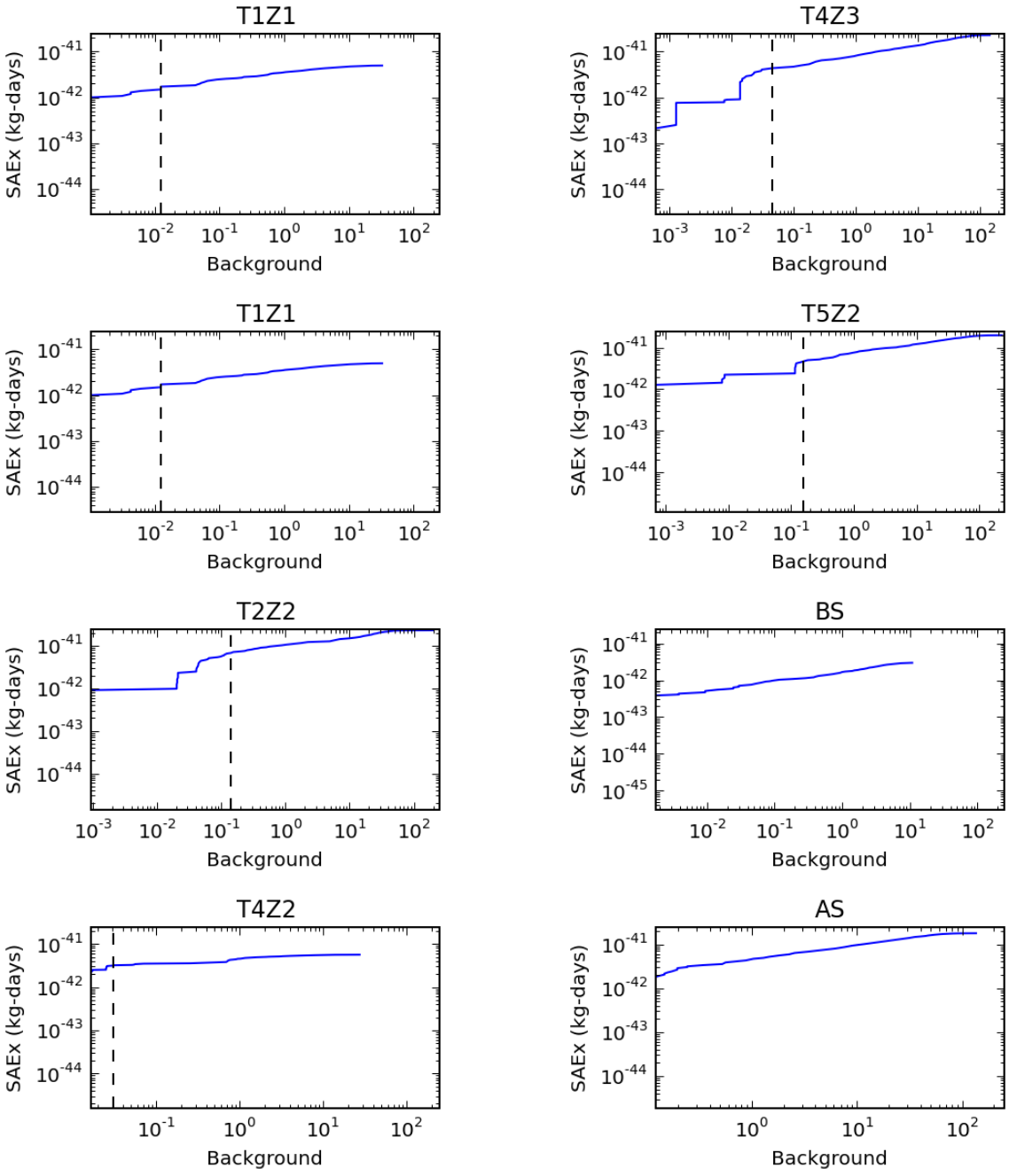

Figure C.7: The expected number of background events vs. *SAE* for a 10 GeV WIMP.

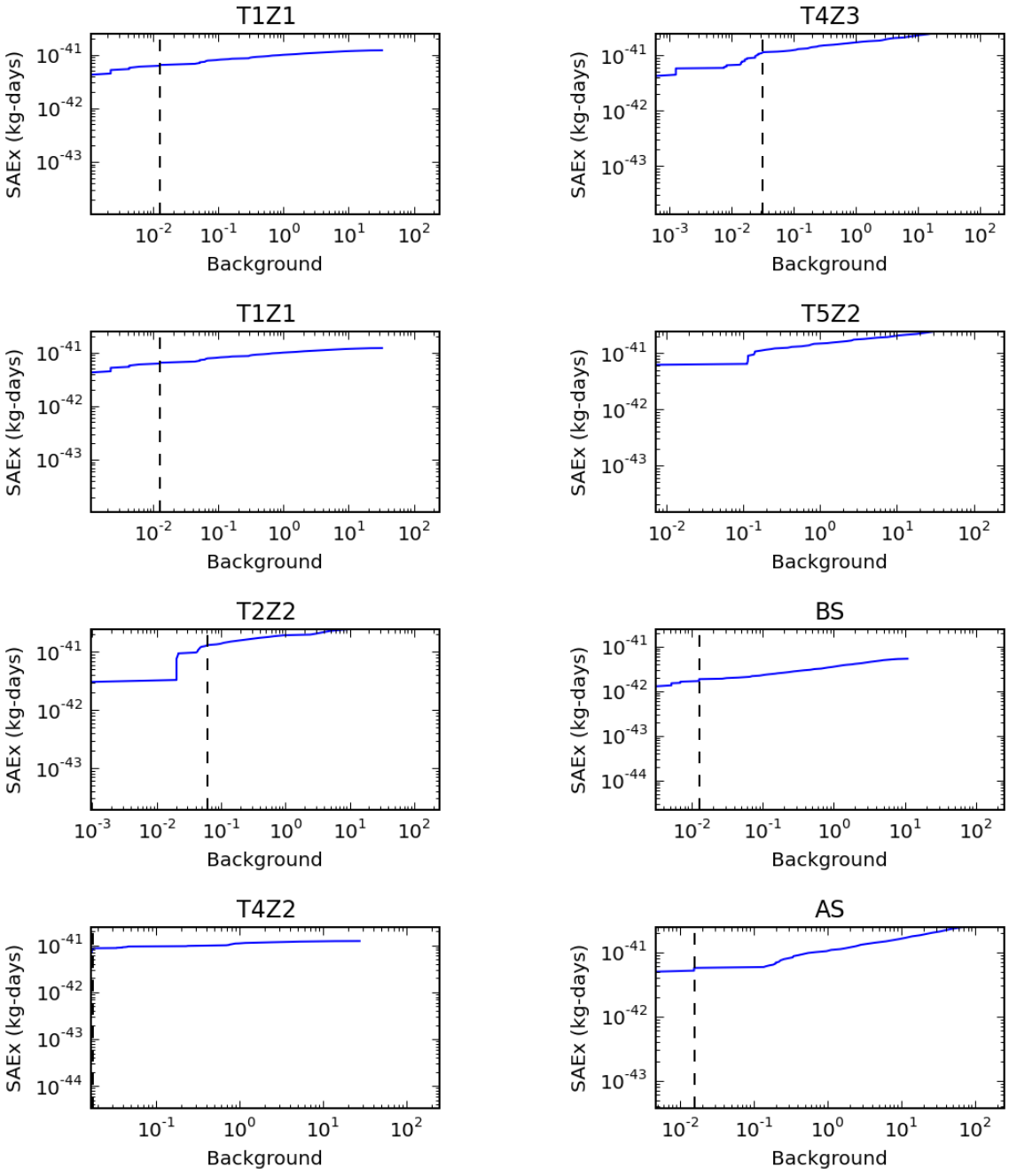

Figure C.8: The expected number of background events vs. *SAE* for a 15 GeV WIMP.

C.3 Relative Background and SAE

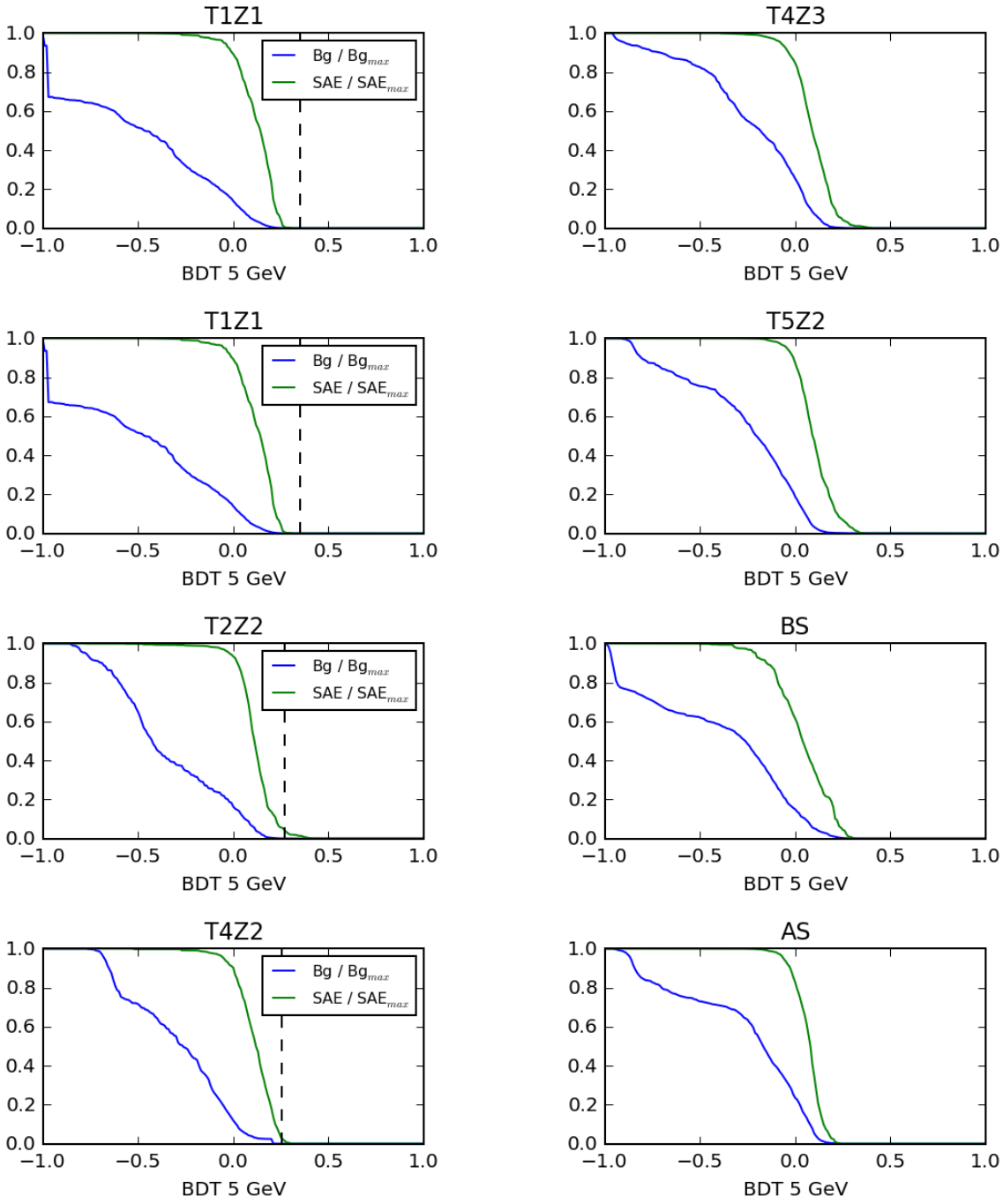

Figure C.9: The relative acceptance of expected number of background events and *SAE* vs. BDT threshold for a 5 GeV WIMP for all detectors.

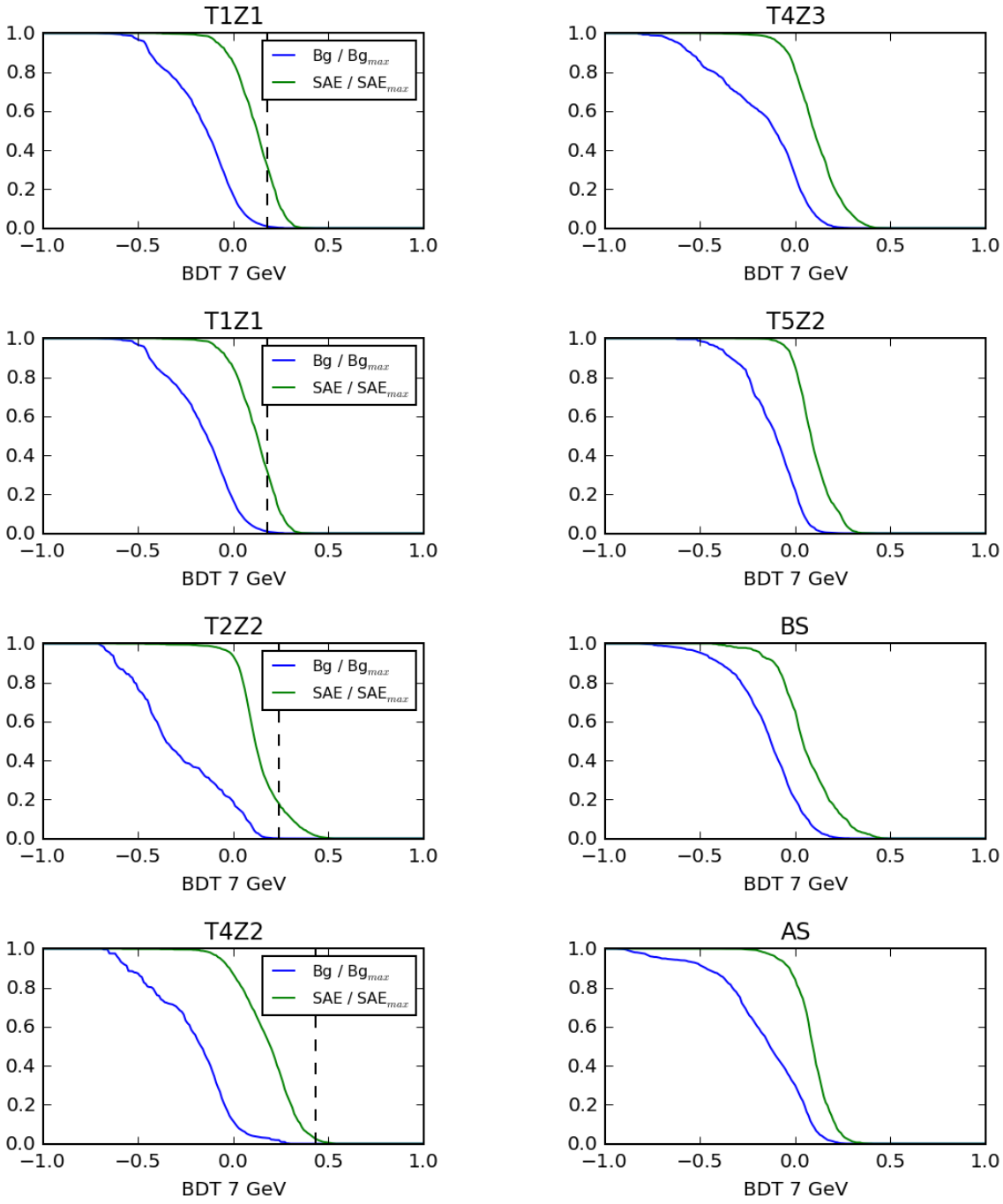

Figure C.10: The relative acceptance of expected number of background events and *SAE* vs. BDT threshold for 7 GeV WIMP for all detectors.

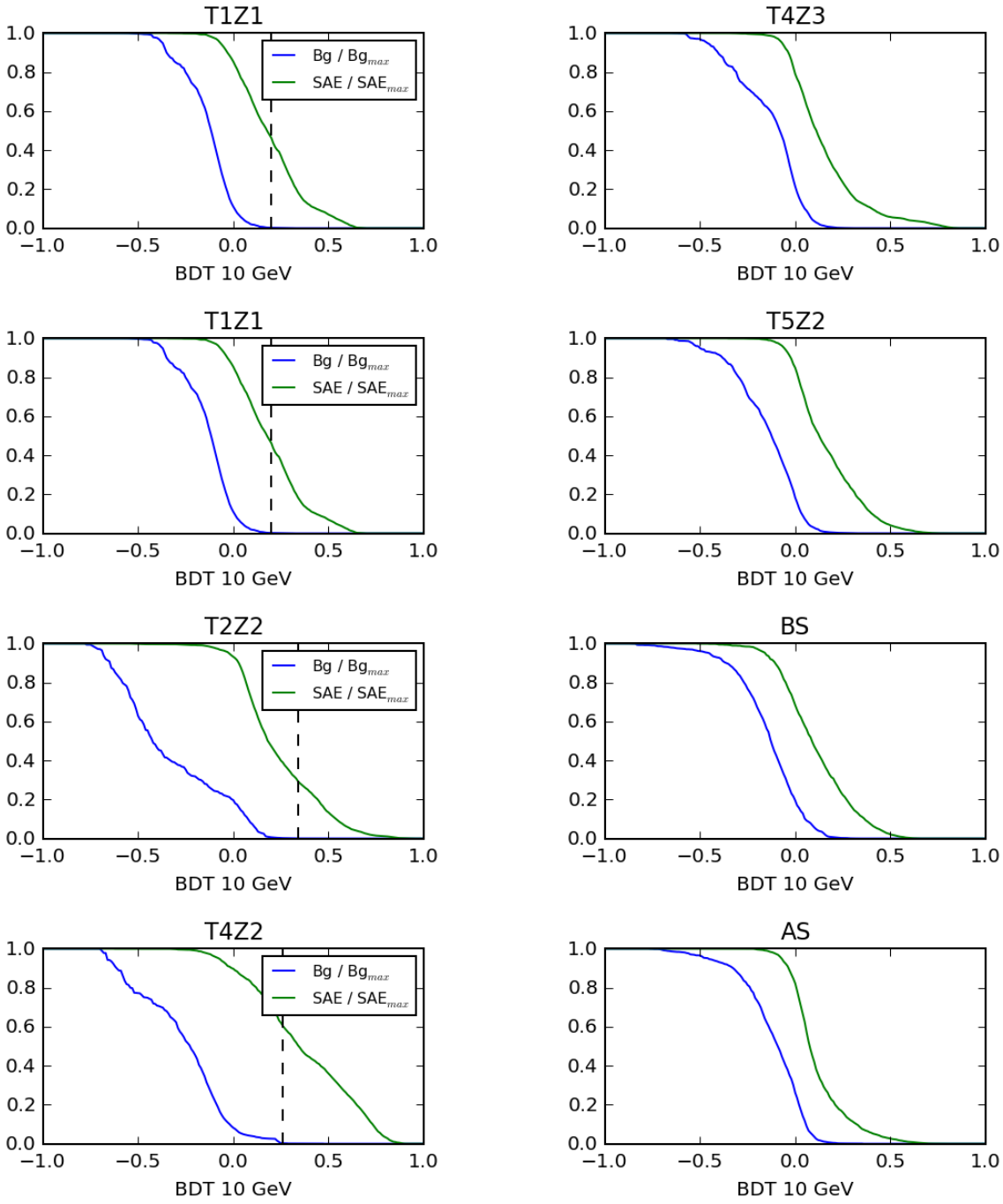

Figure C.11: The relative acceptance of expected number of background events and *SAE* vs. BDT threshold for a 10 GeV WIMP for all detectors.

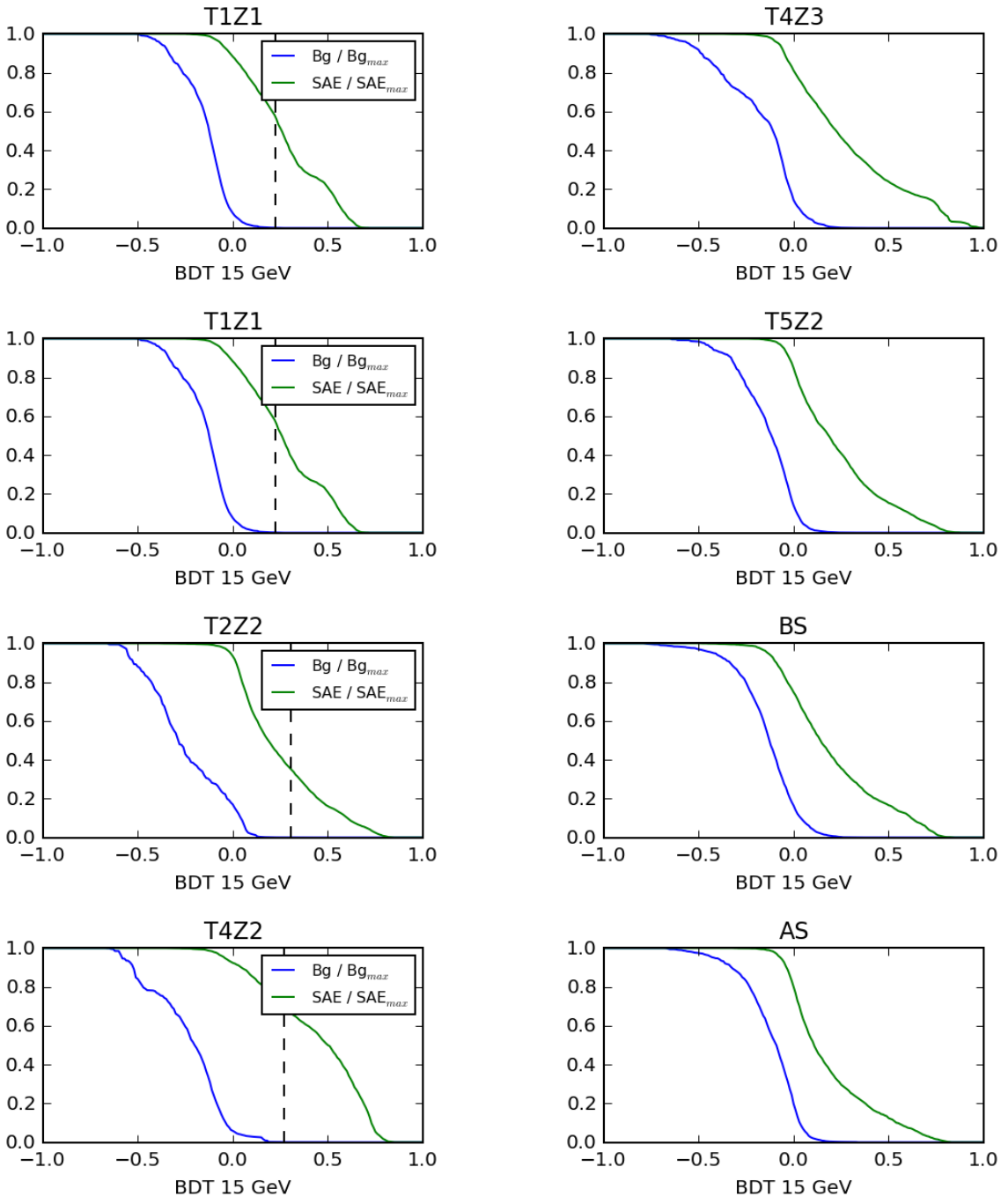

Figure C.12: The expected number of background events vs. *SAE* for a 15 GeV WIMP for all detectors.

C.4 Background vs. BDT

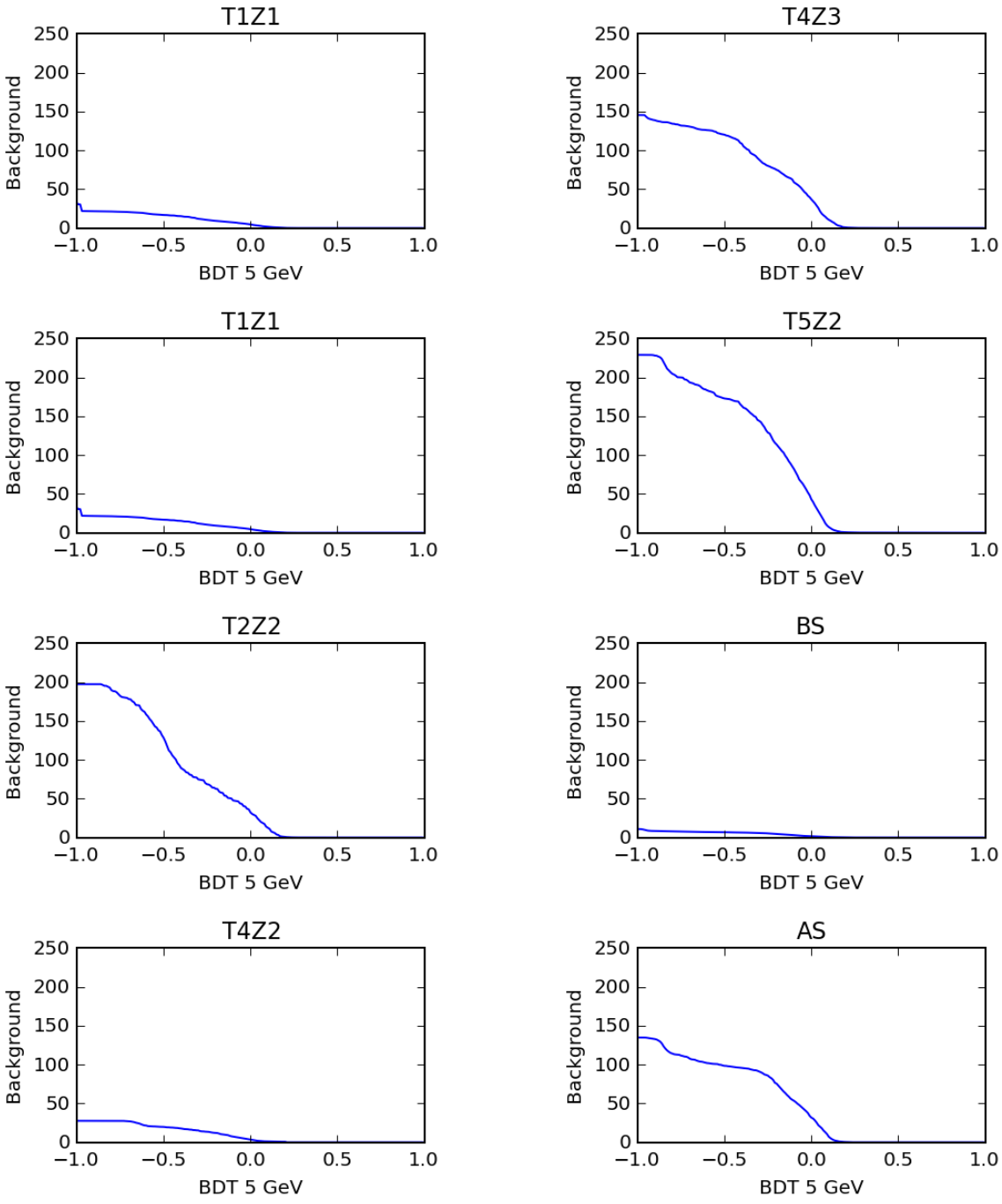

Figure C.13: The expected number of background events vs. BDT Threshold for a 5 GeV WIMP for all detectors.

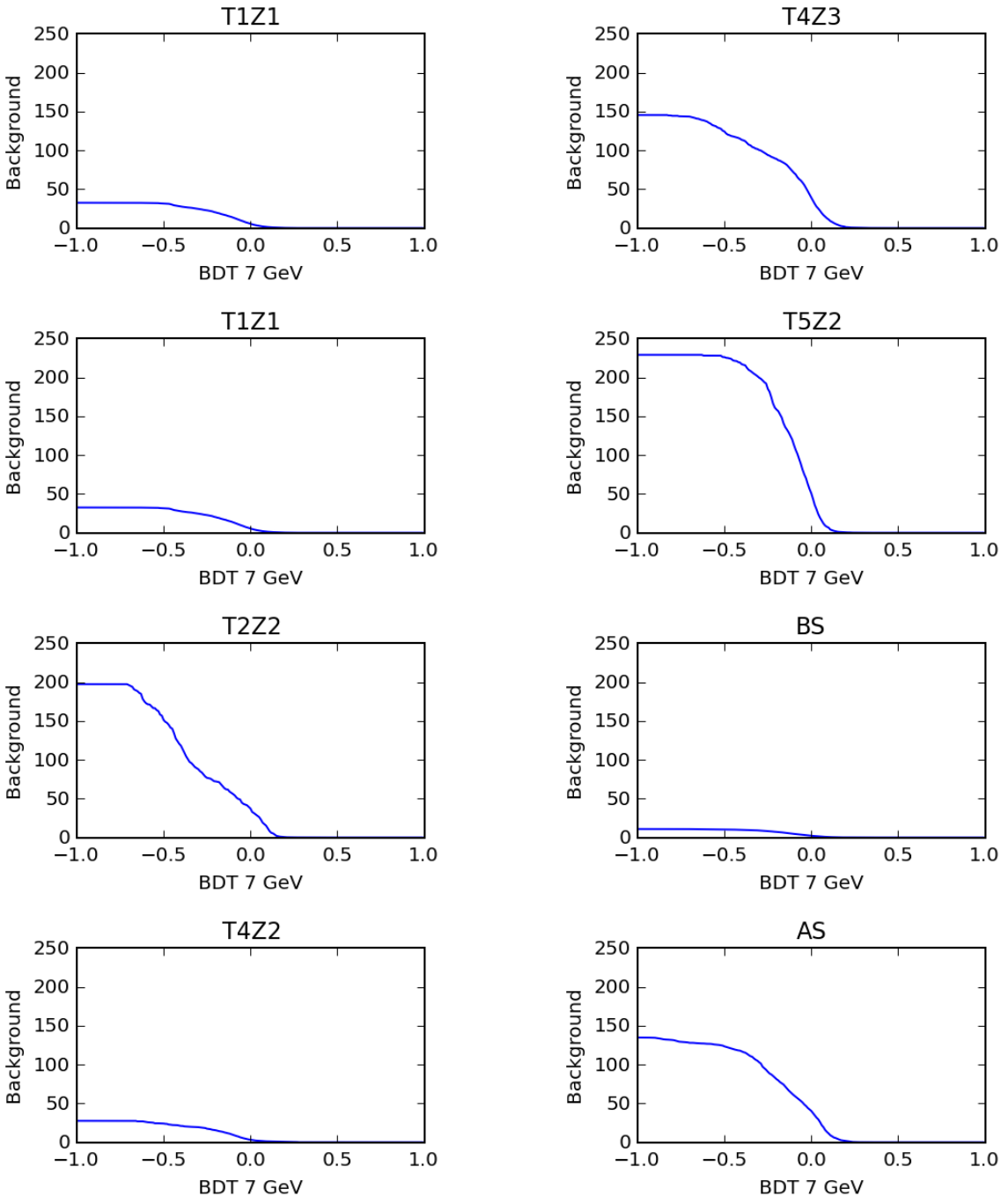

Figure C.14: The expected number of background events vs. BDT threshold for a 7 GeV WIMP for all detectors.

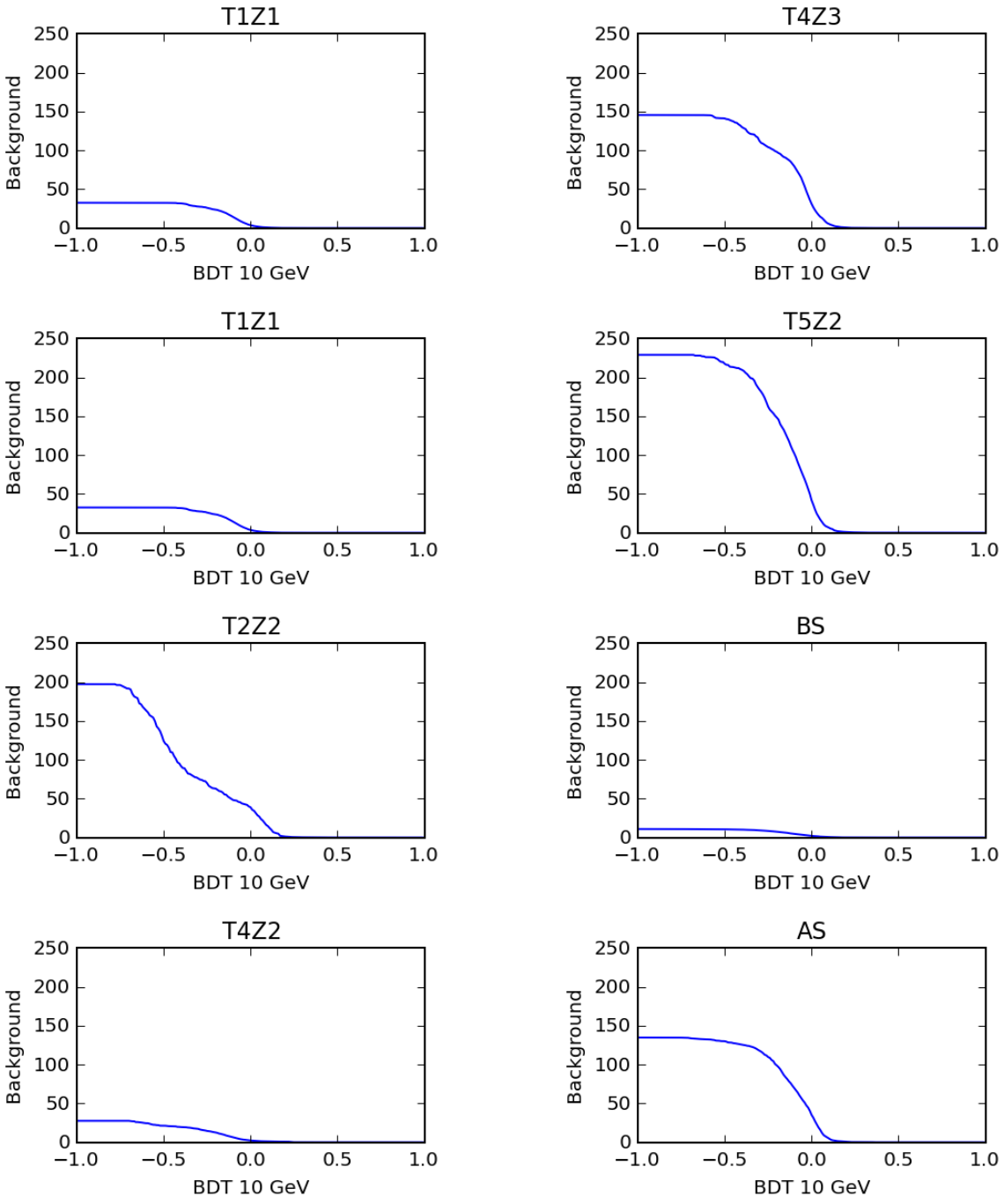

Figure C.15: The expected number of background events vs. BDT threshold for a 10 GeV WIMP for all detectors.

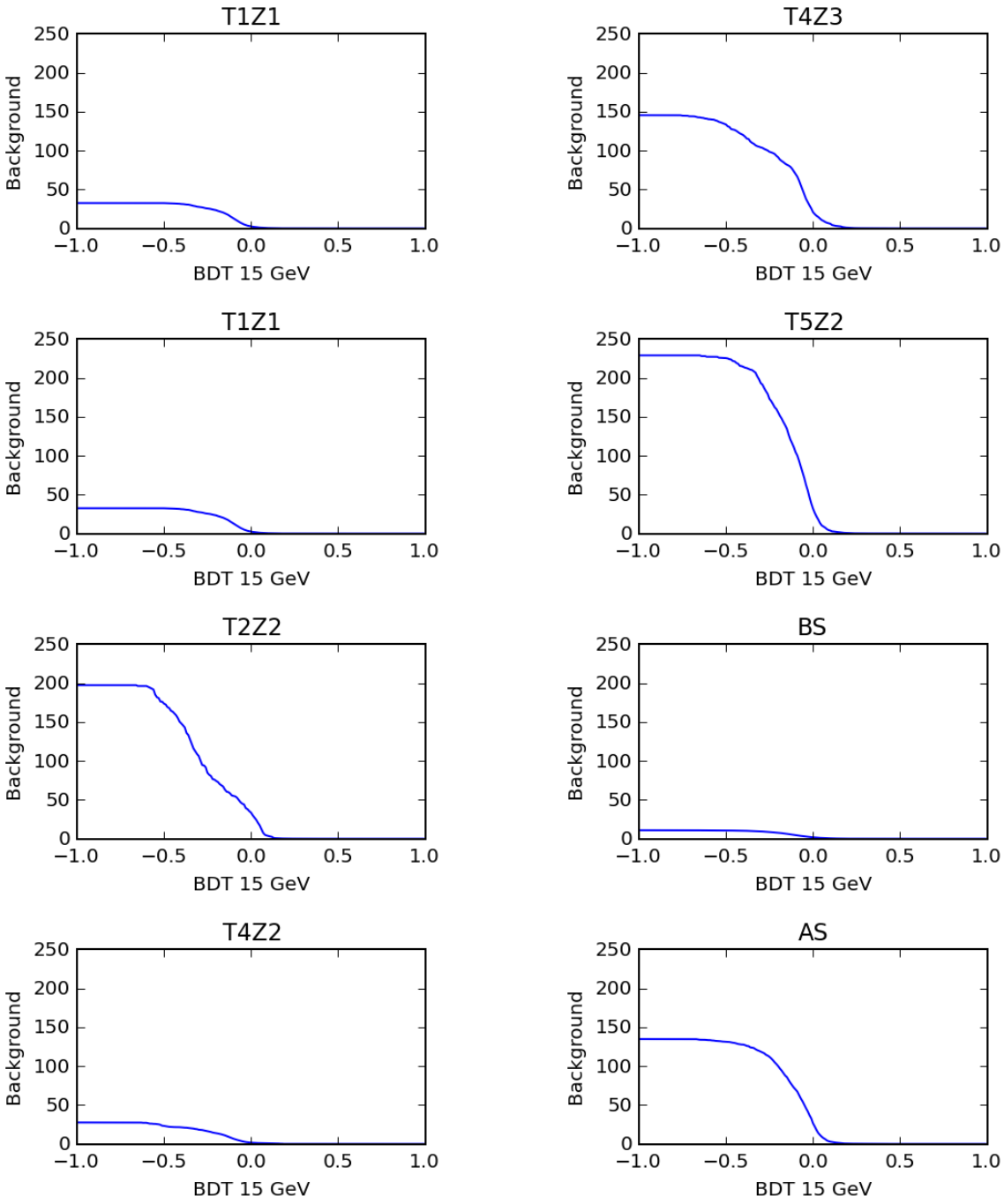

Figure C.16: The expected number of background events vs. BDT threshold for a 15 GeV WIMP for all detectors.

Appendix D

8-bin Optimization

D.1 Coarse Tuning

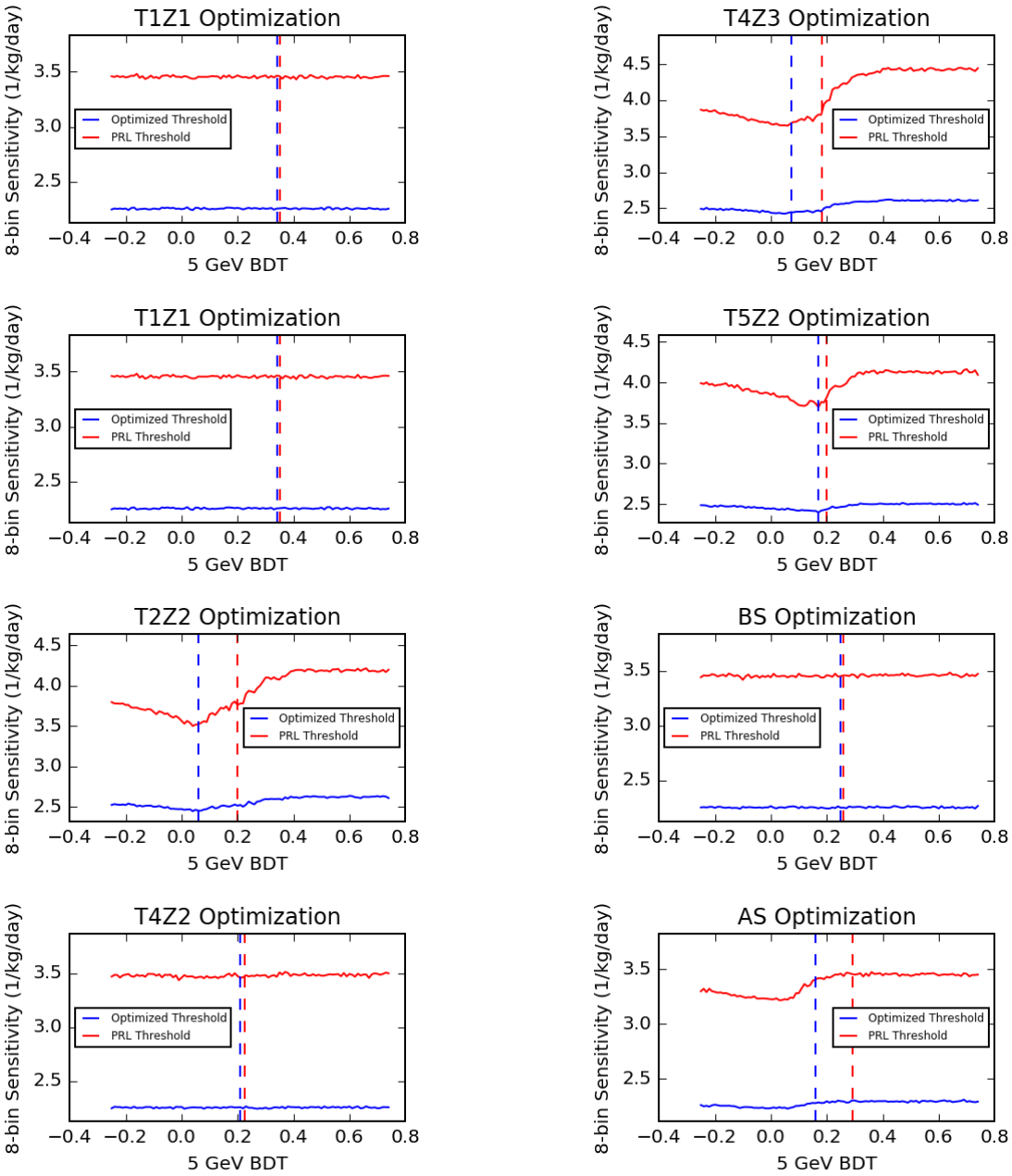

Figure D.1: 8-bin Coarse Grain BDT Optimization for a 5 GeV WIMP. Here the red curve shows 7/8 BDT threshold values being kept at 2014 values while one is altered. The blue curve shows the same situation, but 7/8 BDT values are kept at the values set by this thesis' analysis.

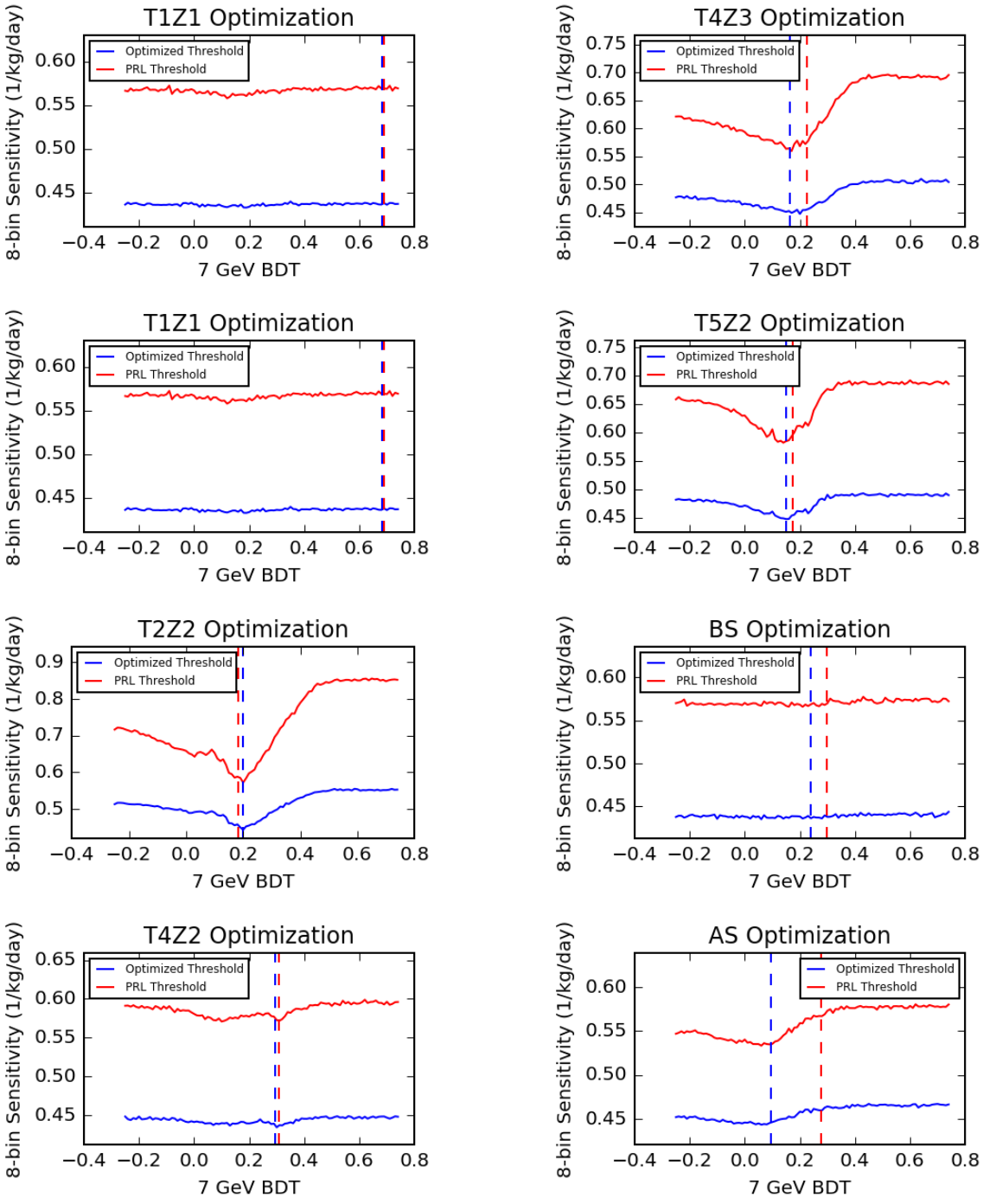

Figure D.2: 8-bin Coarse Grain BDT Optimization for a 7 GeV WIMP. Here the red curve shows 7 of 8 BDT threshold values being kept at 2014 values while one is altered. The blue curve shows the same situation, but 7 of 8 BDT values are kept at the values set by this thesis' analysis.

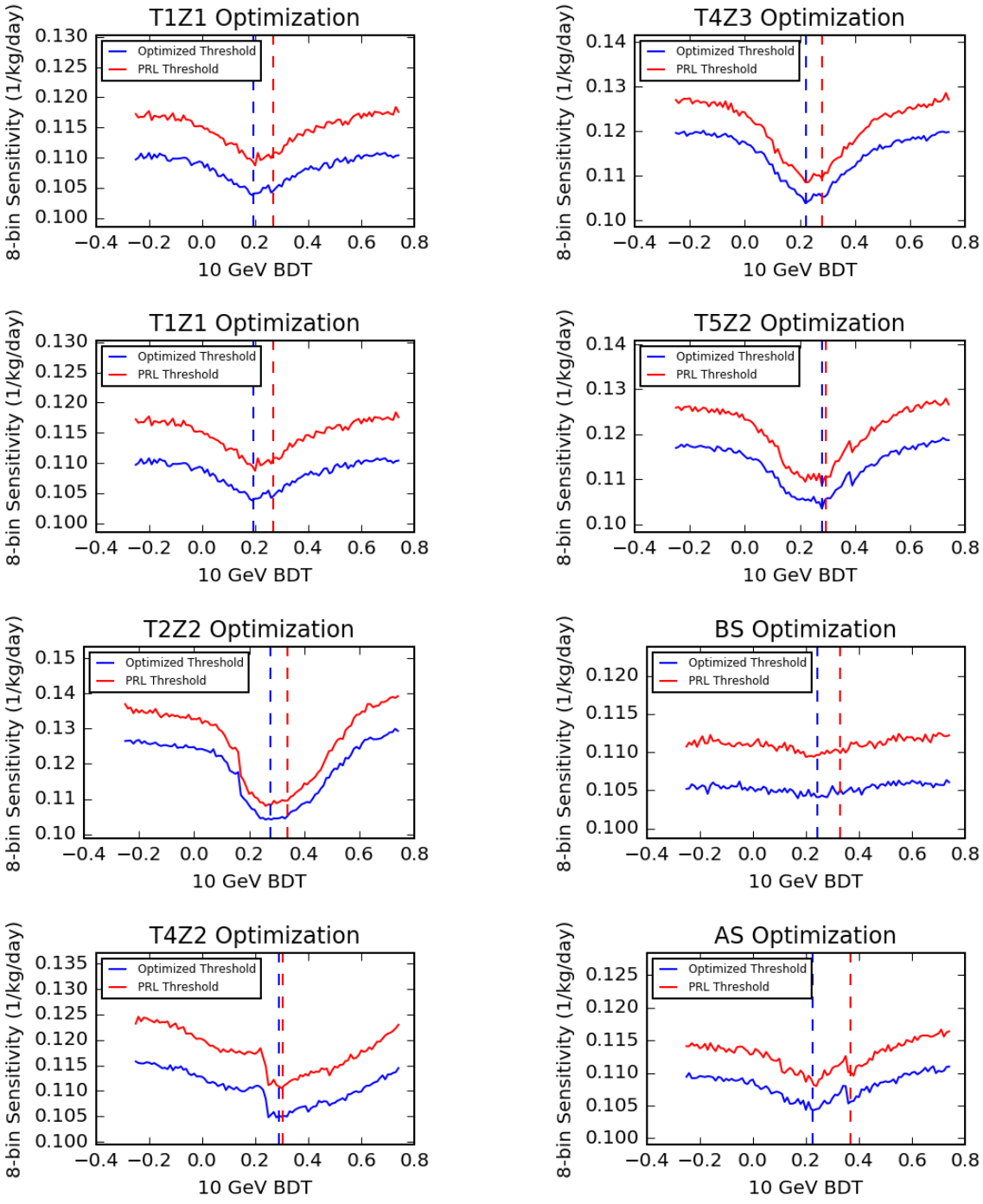

Figure D.3: 8-bin Coarse Grain BDT Optimization for a 10 GeV WIMP. Here the red curve shows 7 of 8 BDT threshold values being kept at 2014 values while one is altered. The blue curve shows the same situation, but 7 of 8 BDT values are kept at the values set by this thesis' analysis.

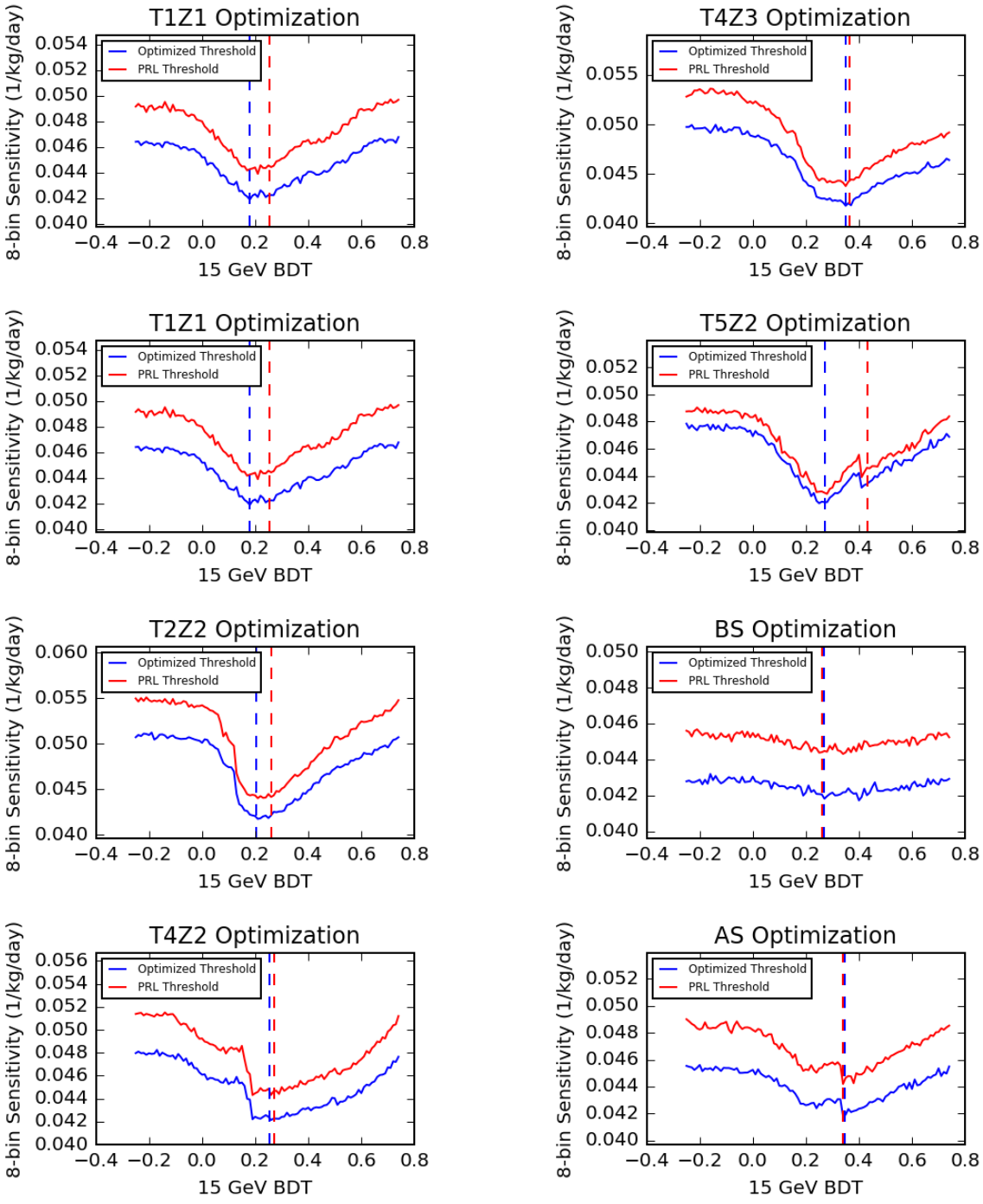

Figure D.4: 8-bin Coarse Grain BDT Optimization for a 15 GeV WIMP. Here the red curve shows 7 of 8 BDT threshold values being kept at 2014 values while one is altered. The blue curve shows the same situation, but 7 of 8 BDT values are kept at the values set by this thesis' analysis.

D.2 Fine Tuning

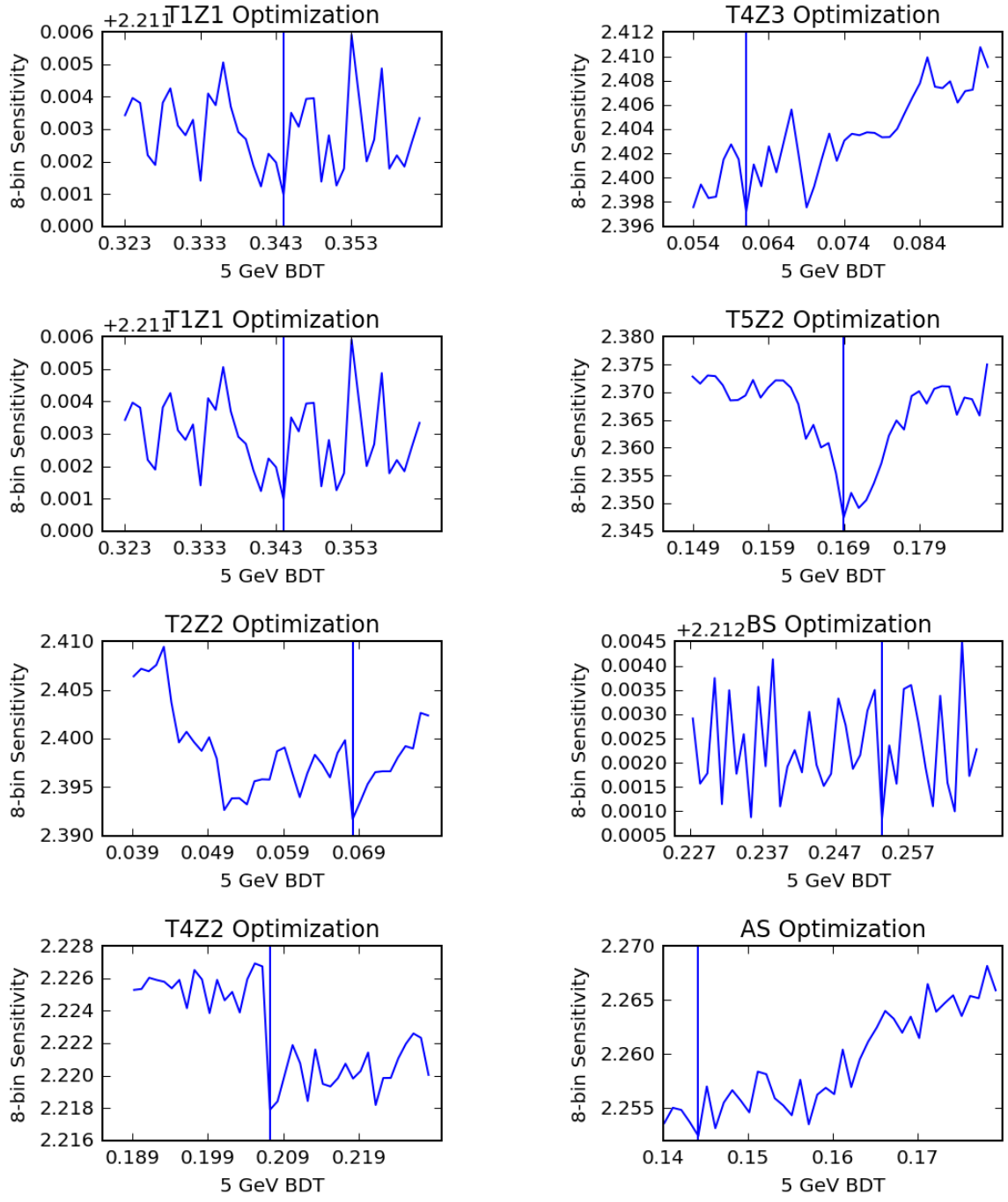

Figure D.5: The 8-bin fine grain optimization of BDT thresholds with regard to $\langle \xi^{90} \rangle$ for a 5 GeV WIMP for all detectors.

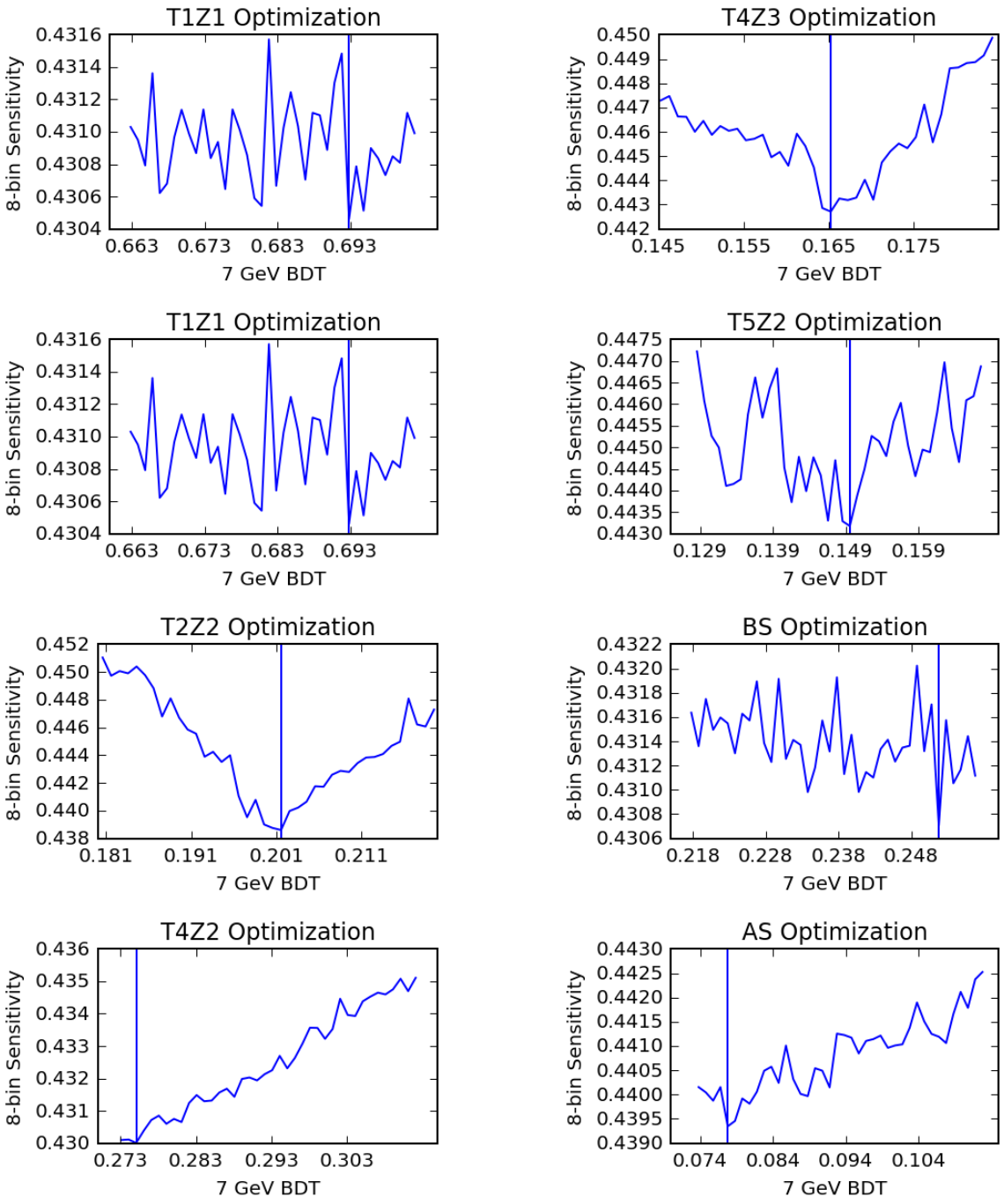

Figure D.6: The 8-bin fine grain optimization of BDT thresholds with regard to $\langle \xi^{90} \rangle$ for a 7 GeV WIMP for all detectors

.

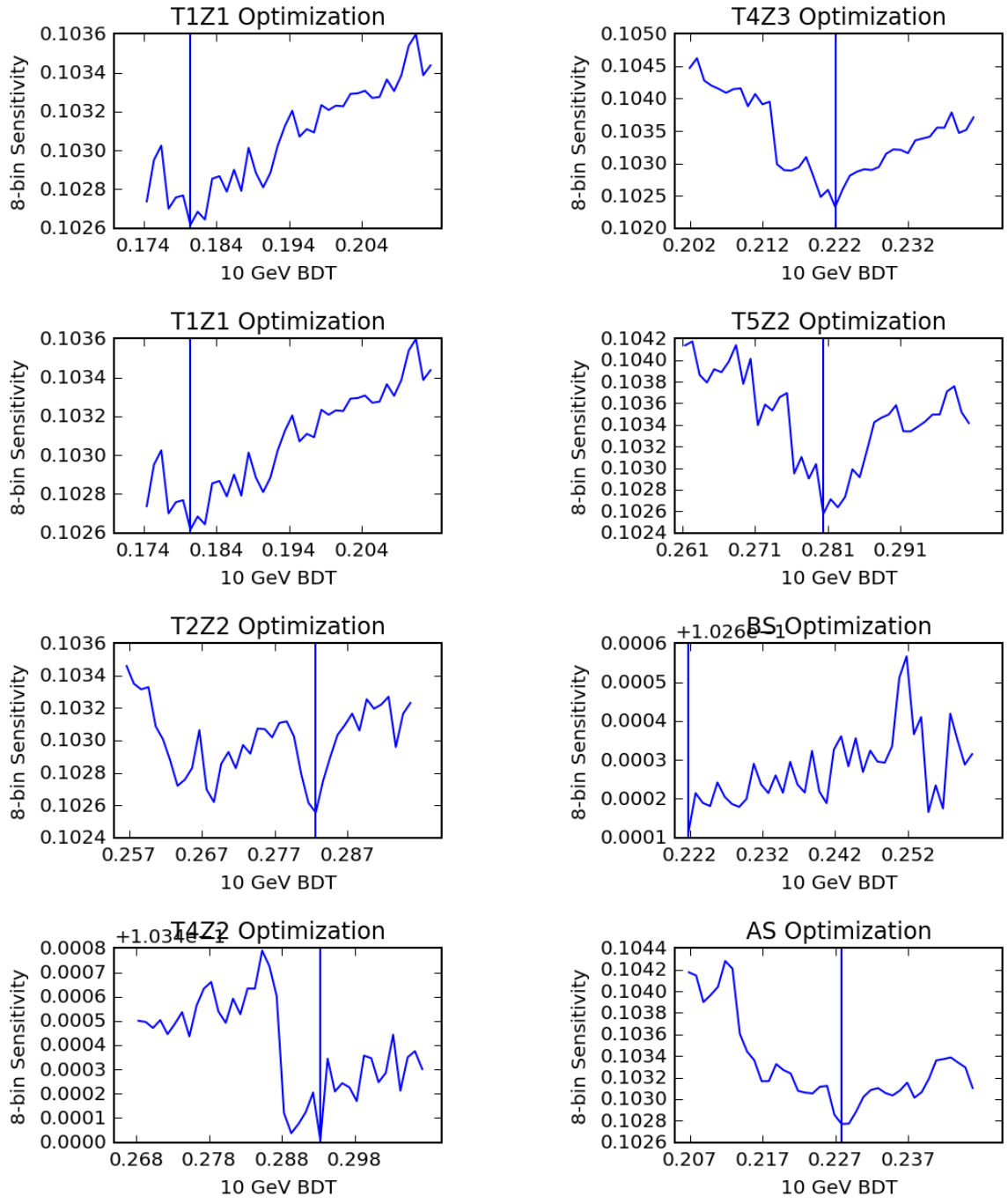

Figure D.7: The 8-bin fine grain optimization of BDT thresholds with regard to $\langle \xi^{90} \rangle$ for a 10 GeV WIMP for all detectors.

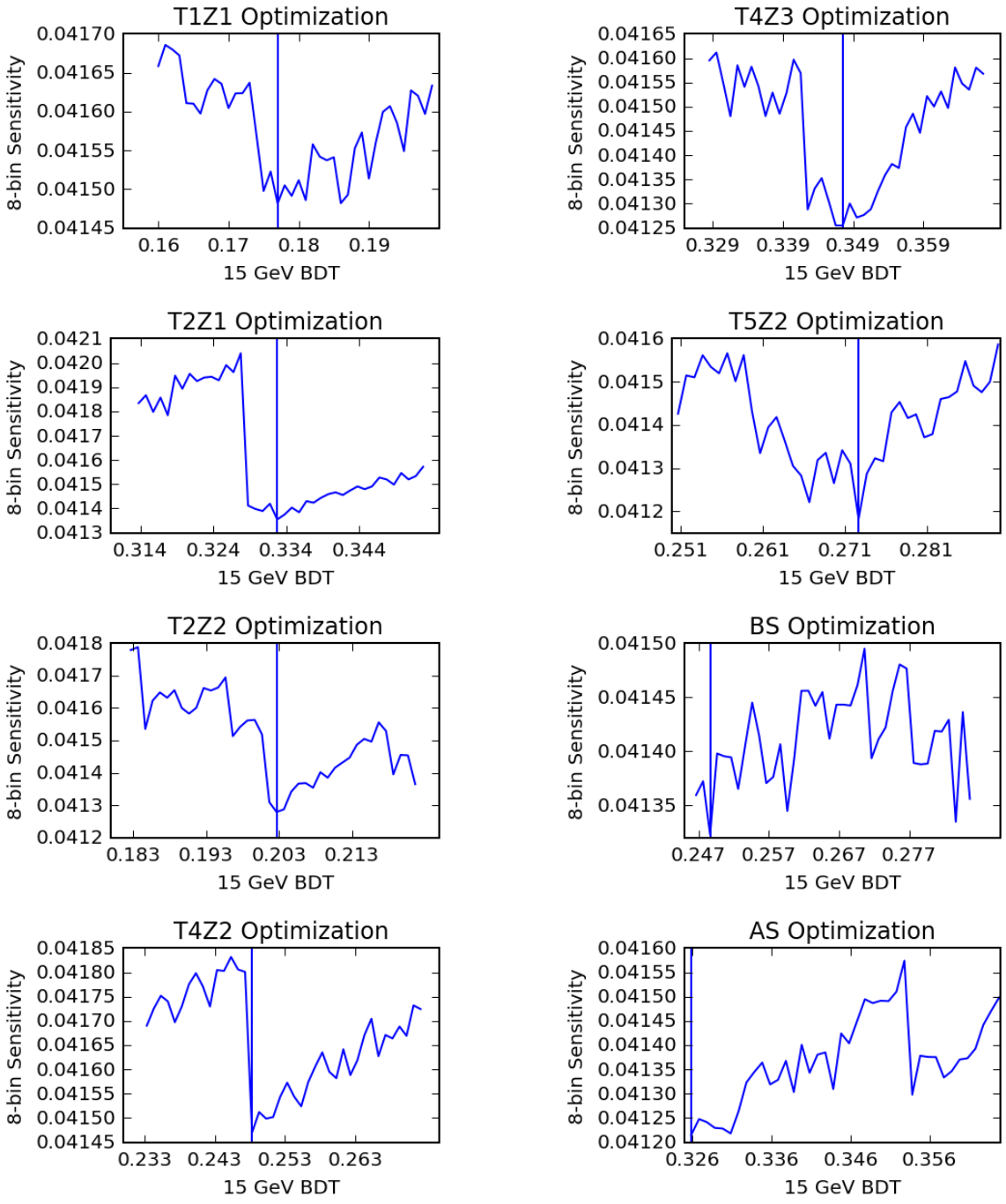

Figure D.8: The 8-bin fine grain optimization of BDT thresholds with regard to $\langle \xi^{90} \rangle$ for a 15 GeV WIMP for all detectors.

Appendix E

Sensitivity and WIMP Mass

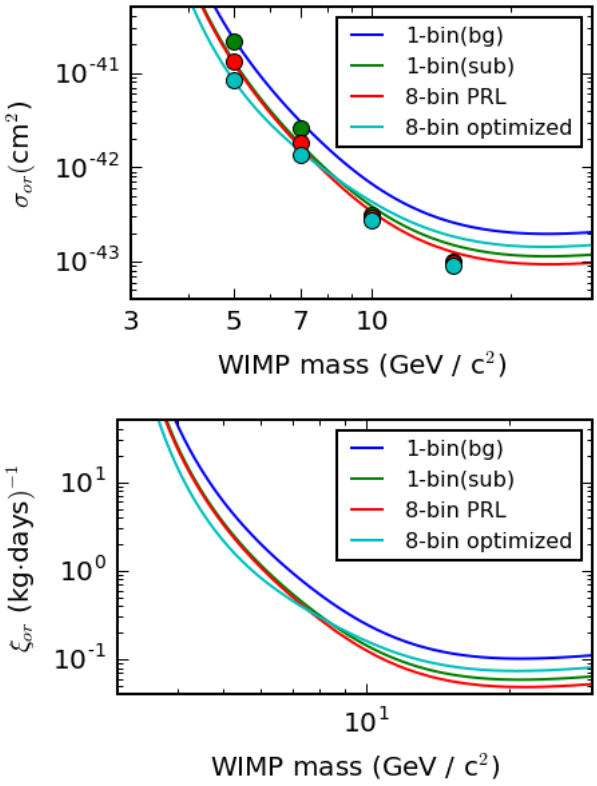

Figure E.1: The expected sensitivity cross section upper limits vs. WIMP mass (using .OR. selection).

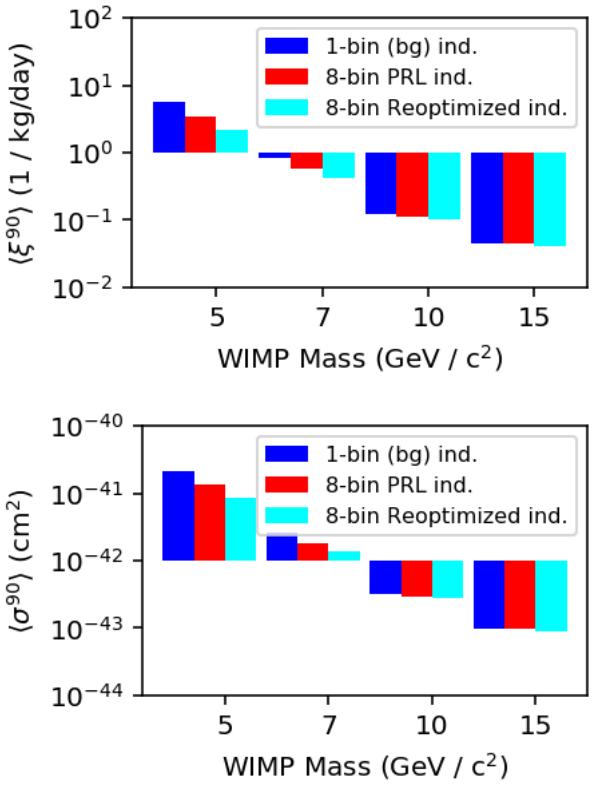

Figure E.2: Sensitivity and Cross Section Limits vs. WIMP Mass (using individual BDT selection). The bottom figure is featured in chapter three. Despite a notational difference in the legends, the same selection criteria is featured in both plots.