Exploring Fully Bayesian Unfolding for γ-ray Spectra

Vala Maria Valsdóttir

A thesis presented for the degree of

Master of Science

Department of Physics

University of Oslo

August 2020

Abstract

In this thesis a new method for the unfolding of γ-ray spectra using Bayesian statistics has been investigated. The method in question goes by the name Fully Bayesian Unfolding. By making use of Bayes' theorem the method samples from a likelihood function and prior distribution to achieve a posterior distribution for each energy-bin of the spectrum.

The γ-ray spectra are obtained from experiments at the Oslo Cyclotron Laboratory with OSCAR an array of $LaBr_3$:Ce scintillator detectors and SiRi a particle telescope. The reaction chosen to test the method was the ^{28}Si(p,p$'\gamma$) reaction.

The new FBU method was compared to the iterative unfolding method that was developed and is currently used for unfolding in the Oslo nuclear physics group. In addition, the iterative unfolding method is a well known unfolding method used in nuclear physics research outside of the Oslo nuclear physics group.

The motivation of the study is to extract the uncertainty from the unfolding method. These uncertainties can be propagated into further analysis such as the Oslo-Method or others. In addition the iterative unfolding method can suffer from negative fluctuations. Using Bayesian unfolding prevents this since the priors can have lower bounds set to zero, thus only the positive solutions are presented in the outcome. The Fully Bayesian Unfolding was found to deliver on these terms both giving reasonable uncertainty estimates and give a stable solutions for the γ-ray spectra tested in this work. The main source of discrepancy was found in the threshold region ($[0, 160]$ keV), which in turn seems to have an effect on the re-distribution of counts to the γ-ray peaks. The discrepancy was decreased for the FBU when including the background into the unfolding. Since the raw spectra do not contain counts in the threshold area the discrepancy can only come from the response matrix. The response matrix was made under different conditions than the experiment tested in this thesis, and is still a work in progress. The total efficiency of OSCAR has not been fully established either and was estimated using the geometric efficiency. There is uncertainty in the total efficiency calculated in this thesis and it can also have an affect on the discrepancies seen in the unfolded spectrum.

Acknowledgement

Finally, all the rehearsals in the shower for my Oscar acceptance speech comes to use! Never in my wildest dreams did I think it would be put to use in the context of me getting a degree of master of science. Luckily, I will not have to fake cry whilst writing this. Firstly, I have to thank my supervisors Ann-Cecilie Larsen and Frank Leonel Bello Garrote. Ann-Cecilie is the busiest woman I have ever met and at the same time she somehow manages to listen to my endless blabbering about unfolding, response matrices, machine learning, statistics, calibration and other quite boring topics. I thank you for always taking the time, even though you have much more important things to attend to. Frank Leonel Bello Garrote, the first thing one thinks of when hearing the name is someone that wrote a constitution. Instead he got stuck helping me with my thesis. For me he will forever be the person that took the time to read through my thesis and always giving insightful feedback, even when officially on vacation. I will forever be gratefully for that. Also, thanks to all of the nuclear physics group in general. All of you have at some point helped me, and all of you have been kind and inviting. Thank you all for that.

To my favorite lecturer Morten Hjorth-Jensen: Thank you for introducing me to all the topics that I did not know I loved and for always finding time to discuss and help me.

Thanks to Andreas Kvellestad for your insight into Bayesian Statistics, one zoom meeting with you really helped me progress with my thesis.

Special thanks has to be given to my peers Erlend, Dorthea, Maria and Per-Dimitri. Erlend, without you I would not have this particular thesis. Thank you for sitting endless hours discussing what unfolding method I should try out, before you suggested Bayesian unfolding. Dorthea, it has been so great discussing unfolding with you on zoom. Thank you for reading my thesis and giving insightful feedback right before my delivery. Maria, you have been the greatest friend during all of this. Thank you for all the support, always telling me that I am great. Per-Dimitri, you had to live with me in lock down in a small flat during the writing of my thesis, and you still seem to like me! In addition you always helped me when I got stuck. Thank you for all the years studying physics. From the first year of bachelors at NTNU to a masters at UiO, we have always worked together and it has always been a pleasure.

At last, I want to thank my family. I have a big family stretched across both Norway and Iceland, so you will have to do with me not addressing all the names. Thank you for the endless support. Even though you have no idea what I am doing, all of you have always helped and pushed me to finish. Thank you, I love you all!

Contents

Chapter 1

Introduction

> Having aspirations of becoming a scientist, what I really wanted to know was how to analyse experimental data. Thus, I eagerly looked forward to the lectures on statistics. Sadly, they were a great disappointment.
>
> *D. S. Sivia*
>
> *Data Analysis: A Bayesian Tutorial*

The probability formulations of Bayes where published posthumously in 1763 in the essay called 'An Essay Towards Solving a Problem in the Doctrines of Chances' [1]. The question that Bayes' tried to answer in the essay had been up for debate for a long time already then: how should one reason where it is not possible to argue with certainty [2]? While the answer came in Bayes' essay the ideas of Bayes, at the time, where not considered to be groundbreaking. There was nothing wrong with the ideas published in the essay, but rather the concept itself, seen in light of science. While Thomas Bayes was alive he was a fellow of the Royal Society, and was a known theologian, philosopher and mathematician. Still the ideas published in the essay where disregarded and forgot, until 1812 when Laplace wrote down the idea mathematically which is known today as Bayes' theorem. Laplace rediscovered the ideas Bayes' and the theorem by himself, and also used it to solve problems in various fields as celestial mechanics and medical statistics [2].

To introduce probabilistic thinking, one should start with two concepts: deductive logic and inductive logic. A typical question that is often encountered is: given y, find x. This is a typical example where deductive logic can be applied. For this problem the well-defined axioms of algebra can be used to find the correct answer. This kind of logic is very helpful in many mathematical problems. Many scientists today, however, have the reversed problem: given the outcome what are the causes of the outcome? For example, when throwing a dice and one gets the number 6 three out of four times, is the dice fair? In this type of problem one has to use inductive logic to find the plausible cause of the outcome. By pondering over this problem one might realize that this is

more challenging than using the axioms of algebra to find x. In a case where the outcome is known but the cause unknown, it might seem reasonable to make an inference based on the experimental data and any prior knowledge accessible, while still being open to revising the starting position if new information comes to light [2]. The above sentence is Bayes' theorem put into words,

$$\text{prob}(X \mid Y, I) = \frac{\text{prob}(Y \mid X, I) \times \text{prob}(X, I)}{\text{prob}(Y \mid I)}.$$

The term $\text{prob}(X, I)$ is called the prior, it represents the state of knowledge or ignorance about the truth of X. This is modified by the term $\text{prob}(Y \mid X, I)$ which is called the likelihood and presents the data, here called Y, that one wishes to infer knowledge about. The denominator is the marginal probability of Y, and plays a crucial role in some data analysis problems (model selection problems). The term on the left side is called the posterior and represents the state of knowledge achieved about X being true in the light of Y. I denotes the relevant background information at hand. Going back to inductive logical thinking: If there is a hypothesis that needs testing, let say, is the dice fair, it might be a good place to start with how much one believes that the dice is fair. From a probabilistic view point, if one specifies how much one believes the dice to be fair, at the same time one is specifying how much one believes that the dice is not fair. The second assertion one can make, is slightly more complicated [2]: If one specifies how much one believes that a hypothesis Y is true, and how much one believes that X is true given Y, then automatically one has stated the belief of them both being true. Making the following constraint: given the data, one should always arrive at the same conclusion even if different analysing paths are chosen. This can only be true if the rules of probability theory are used when making the guess of how much one believes something to be true:

$$\text{prob}(X \mid I) + \text{prob}(\overline{X} \mid I) = 1$$

and

$$\text{prob}(X, Y \mid I) = \text{prob}(X \mid Y, I) \times \text{prob}(Y \mid I),$$

where the notation $\overline{X}$ means when X is false, and the vertical bar $\mid$ stands for 'given'. The first equation is called the sum rule and the second equation is called the product rule. From these two equations Bayes' theorem can be derived.

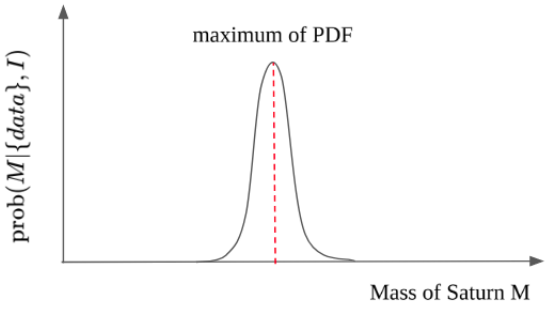

Figure 1: Illustration of the resulting posterior distribution of the mass of Saturn

Going back to the story of Laplace: After his death the Bayesian ideas where yet again disregarded because of the problematic concept of probability represented as a degree of belief, i.e how much one thinks something is true, based on the evidence at hand [2]. Instead the idea of probability being tied up to the frequency of an event occurring, became the winning approach. Since frequency can be measured, it was seen as the most objective way of dealing with random phenomena. However, in many scientific problems, the frequency of an event is not always available, and the approach thus has limits. A great example of this comes from the work of Laplace himself. He used his probability theory (Bayes' theorem) to estimate the mass of Saturn. This was before much data on the orbit of Saturn was collected, thus the frequency approach was not possible at the time. In short, he started by calculating the posterior probability distribution function (PDF) of the mass M, given the data and all the relevant background information I (i.e knowledge of classical mechanics). This gave him a PDF as illustrated in figure 1. The PDF shows where the mass of Saturn is most likely to be. The best estimate of the mass is represented by the maximum of the PDF, and the width represents the uncertainty in Laplace's estimate. Over the past 150 years, data has been collected on the orbit of Saturn and Laplace's estimate has change only by 0.63% [2]. In the 17th century, before enough observations where made, looking at the mass of Saturn in the light of the frequency definition one is prohibited from using probability theory, since the mass of Saturn is a constant and not a random variable. This means it has no frequency distribution and probability theory can not be applied [2]. If the PDF in the figure should be interpreted with a frequentist approach, one could have sampled data from a large ensemble of universes, where everything remained constant apart from the mass of Saturn. This immediately sounds like a far fetched idea. Problems similar to estimating the mass of Saturn appears in many fields of physics and the Bayesian approach has therefore been growing in terms of use the past years.

In this thesis, new methods using Bayes' theorem for unfolding of data from nuclear experiments performed at the Oslo Cyclotron Laboratory (OCL), will be investigated. However, this method is general and can be applied to other types of data. As previously mentioned, Bayes' theorem includes the relevant background information I, in this study the relevant background information is the well-known behavior of experimental setup and data processing at OCL. The setup at the laboratory consists of a $\Delta E - E$ Silicon Ring particle-detector system

(SiRi) [3] and 30 $LaBr_3$:Ce scintillator detectors referred to as the Oslo Scintillator ARray (OSCAR). In this thesis γ-ray spectra obtained from experiments with OSCAR will be used to investigate the Bayesian unfolding method. When the γ-ray hits the detector it undergoes different types of interactions with the matter, such as the photo-electric effect, Compton scattering and pair production. The aim of unfolding, is to reconstruct the spectrum to a set of full-energy peaks, i.e when the γ-ray energy is fully absorbed by the detector. To obtain such a spectrum, accurate knowledge of the detector response functions is crucial. This knowledge presents the background information I that are known beforehand.

The question of unfolding is the following: Given the data and the response function, what was the actual truth-spectrum? Then if Bayes' theorem is stated, without X and Y, and instead with the hypothesis of what is believed to be the truth and the data, it might become clear that the question above fits the theorem:

$$\text{prob}(truth \mid data, I) \propto \text{prob}(data \mid truth, I) \times \text{prob}(truth \mid I). \tag{1.1}$$

In this version of the theorem the term in the denominator is omitted and thus there is a proportionality instead of an equal sign. In many data analysis problems this is sufficient since often the posterior is not dependent on the term [2]. By making a hypothesis about what is the truth and test it against the evidence (the data), Bayes' theorem proves useful in the estimation of the unfolded spectrum. Since the true quantity of interest is not known in the case of unfolding, the data is the only measurement one has. The power of Bayes' theorem lies in the fact that it relates the quantity of interest, i.e the truth-spectrum, to the observed data and evidence at hand. In other words, it is easier to examine the data, given the hypothesis of the truth, than it is to examine our hypothesis given the data. In addition, by applying Bayesian statistics to the unfolding, one obtains uncertainty estimates of the method, which can be very useful for the further analysis.

Since Bayesian statistics is rather new in the field of nuclear physics, chapter2 of this thesis is dedicated to an introduction of the Bayesian approach. Chapter 3 will elaborate further on the background information, that is the experimental setup and data processing at OCL. Part II of the thesis includes chapter 4 of the method for unfolding. In chapter 5 the results for unfolding will be presented and in chapter 6 the model checking and results of the unfolding will be discussed. The thesis ends with a summary and future outlook in chapter 7.

Part I

Part 1: Theory and Background Knowledge

Chapter 2

The Bayesian Approach

> A decision was wise, even though it led to disastrous consequences, if the evidence at hand indicated it was the best one to make; and a decision was foolish, even though it led to the happiest possible consequences, if it was unreasonable to expect those consequences
>
> *Herodotus around 500 BC [2]*

In recent years Bayesian statistics or Bayesian inference has become more popular among scientist in many different fields. Inference meaning when a conclusion can be achieved on the basis of evidence and reasoning. The upside of using Bayes' theorem instead of the classical frequentist approach is the inherent flexibility. A well defined Bayesian model can incorporate multiple levels of randomness, and in addition has the ability to combine information from different sources, while offering uncertainty summaries of the estimates. In some cases, it gives more nuanced results which can better describe real-world problems. Another motivation for focusing on Bayesian methods is its philosophical aspects. If not being familiar with classical statistics and its frame work, a receiver of statistical information, for example the likelihood of it raining tomorrow stated by the weather forecaster on the news will often be interpret in a Bayesian 'frame of mind' (without being aware of this, of course). Meaning, they will interpret the numbers as probability statements about the likely values of unknown quantities conditional on the evidence in the data [4]. One can argue that Bayesian statistics is more naturally intuitive than the frequentist approach, but at the same time one has to be careful when making that statement. Since Bayes' theorem incorporates prior knowledge and the likelihood of the data, a good model needs to be specified to infer meaningful estimates. After specifying a good model the resulting posteriors should be evaluated. As with classical statistics, there is a need for verifying the model after sampling the posteriors. The drawback with Bayesian inference is that specifying a model might not be clear cut, there are many paths that can be taken and all models are dependent of the user understanding the data, and having enough knowledge to

initialize and defined it. The same argument goes for inferring estimates from posteriors and model checking. It is a job that requires careful consideration and can lead to conclusions that are more open to interpretation. This chapter aims to shed light on the possibilities within the Bayesian frame work. Construction of a well defined probability model will be discussed, including choosing priors and likelihood functions in section 1. The posterior distribution and how it can be interpreted together with model checking is presented in section 2 and 3. In the end the sampling method used in this thesis is introduced.

1 Probability Models

1.1 Bayes' Theorem

As stated before, Bayes' theorem follows directly by combining the definition of conditional probability with the product and sum rules [2]. The sum rule being

$$\text{prob}(X \mid I) + \text{prob}(\overline{X} \mid I) = 1 \tag{2.1}$$

which states that the probability of X being true plus the probability of $\overline{X}$ not being true equals to one. The product rule

$$\text{prob}(X, Y \mid I) = \text{prob}(X \mid Y, I) \times \text{prob}(Y \mid I), \tag{2.2}$$

states that the joint event of X and Y, given I equals the probability of X being true given that Y and I is true, times the probability of Y given I. If the above equation 2.2 is transposed, that is X and Y are interchanged

$$\text{prob}(Y, X \mid I) = \text{prob}(Y \mid X, I) \times \text{prob}(X \mid I), \tag{2.3}$$

the statement says that the joint event of X and Y given I equals the probability of Y being true given that X and I is true, times the probability of X given I. The right side of equation 2.2 thus equals the right side of equation 2.3, so

$$\text{prob}(X \mid Y, I) \times \text{prob}(Y \mid I) = \text{prob}(Y \mid X, I) \times \text{prob}(X \mid I). \tag{2.4}$$

If both sides of the equal sign is divided by $\text{prob}(Y \mid I)$, this gives Bayes' theorem for X being true given that Y and I is true

$$\text{prob}(X \mid Y, I) = \frac{\text{prob}(Y \mid X, I) \times \text{prob}(X \mid I)}{\text{prob}(Y \mid I)}, \tag{2.5}$$

where I stands for all the background information at hand, because there is no such thing as an absolute probability [2]. Example of background information is the experimental setup of the experiment where the observed values Y where obtained and the data processing done before the statistical analysis takes place. In the case of experiments in nuclear physics the signals detected comes from a nuclear reaction. This is taken into account when it comes to unfolding. The data that to be unfolded are γ-ray transitions from a nuclear decay. So the probability assigned in detecting a γ-ray for a given energy will depend on the conditions of the experiment. All this information is relevant and is included into the background information I of Bayes' theorem. It may affect how the hypothesis (prior belief) is initiated, but it is not directly included into the prior. The conditioning on I is often omitted, but it must always be remembered that it exists [2]. Therefore, chapter 3 is dedicated to the relevant background information of the data used in unfolding, and Bayes' theorem will thus be presented without it for the rest of the thesis.

1.2 Likelihood

The likelihood function is a function of the parameters that one wishes to infer given a set of observed data, and the function is defined on the parameter scale. It then gives the unnormalized probability of a particular parameter value for a fixed data set, so it is a measurement of the fit of a statistical model to sampled data [5].

In the Bayesian frame work, the likelihood is the conditional probability of the observations Y being true given the hypothesis X. So the data is the evidence at hand, and the hypothesis is the prior [2]. As more samples are drawn the evidence should influence the posterior (outcome) more than the prior beliefs. In theory, if the prior is very objective, the likelihood (evidence) should update the prior belief so that the posterior distributions includes the wanted parameter that one aims to infer [6]. If this is not the case, the model might be misspecified. In this thesis the likelihood function used is the Poisson likelihood. The Poisson distribution is often used to model numbers of events occurring in a fixed time period, i.e. count data, which are assumed to be independent of each other [7]. When using a Poisson likelihood for sampling it is assumed that the observed values $Y = [y_1, ..., y_n]$ are independent and an identically distributed sequence of Poisson random variables [7]. The probability mass function of the sequence is then

$$\text{prob}_Y(y_j) = \begin{cases} \frac{X^{y_j}}{y_j!}\exp(-X) & y_j \in R_Y \\ 0 & y_j \notin R_Y \end{cases}, \tag{2.6}$$

where R_Y is the set of realizations y_j can take, and X is the parameter of interest [7]. The likelihood function of the Poisson distribution is

$$L(X; y_1, ..., y_n) = \prod_{j=1}^{N} \frac{X^{y_j}}{y_j!}\exp(-X). \tag{2.7}$$

The resulting posterior distribution should then be centered at a point between the maximum likelihood estimate (MLE) of X and the maximum of the prior distribution, thus a weighted average of likelihood and prior

estimates [6]. If the sample size is small the posterior should be closer to the prior, and if the sample size is large it will be closer to the likelihood. The MLE for the estimator of the unfolding problem is derived in chapter 4.

1.3 Priors

When using Bayes' theorem to infer knowledge of an observation a prior is needed, that is prob($X \mid I$). A prior represents all the knowledge that the user has about the observations before inference. The prior should be a known probability distribution, like the Normal distribution, Uniform distribution and so on. These are used, with some initial values as the prior belief of how the estimated parameters are distributed. There exist two main groups of priors, **objective priors** and **subjective priors** [8]. If the user does not wish the prior to influence the likelihood of the observations an objective prior is more suitable for the model, and in the opposite case a subjective prior is better suited. An example where an subjective prior can be used is in situations where there might exist evidence from previous experiments or theory, that can be incorporated into the prior and then tested against the observed data. Meaning that the probability mass assigned to prior is distributed so it prioritises some areas more than others [8]. Objective priors can be used when the data is the only evidence and the scientist wishes the likelihood to influence the outcome more than the prior belief. That said, a fully objective prior is hard to achieve. The uniform prior is an example of a distribution often used when talking about an objective priors,

$$\text{prob}(X \mid \mathit{lower}, \mathit{upper}) = \begin{cases} \frac{1}{\mathit{upper}-\mathit{lower}} & \text{for } \mathit{lower} \leq X \leq \mathit{upper} \\ 0 & \text{for } X < \mathit{lower} \text{ or } X > \mathit{upper} \end{cases}, \tag{2.8}$$

where *lower* is the lower bound and *upper* is the upper bound of the probability distribution. The objectiveness of the uniform prior comes from that the bounds can be adjusted to include all observations, and each observation is sampled with equal probability [8]. It sounds as a good objective prior, but in some cases it may not be objective to say that every region should be sampled with equal probability. There exists as many priors as there exists probability distributions, which can be confusing for someone new to Bayesian statistics, as it seems to be an ocean of possibilities. Luckily, good knowledge of the data and having an idea of what likelihood should be chosen makes it easier to pick a suitable prior. In addition there are many ways to verify if the prior was a good choice when analysing the posteriors.

2 The Posterior Distribution

The posterior distribution is the result of Bayesian inference and thus the result of the likelihood multiplied with the prior 2.5. There are many possible ways to infer knowledge from the posterior distribution, and this section aims to give an overlook over the analysis of the posteriors done in this thesis.

2.1 Credible Interval

Instead of retrieving a point estimate when applying Bayes' theorem, one ends up with a posterior distribution $\text{prob}(X \mid Y, I)$. Credible intervals are tools often used in Bayesian statistics to estimate how certain (or uncertain) the model is. A standard measure of confidence in some quantity X is typically the width of the posterior distribution, that can be measured using a $100(1-\alpha)\%$ credible interval [9], which then include $1-\alpha$ of the posterior density mass. The α stands for the level of credibility. In addition, credible intervals gives the uncertainty related to the wanted estimated quantity.

$$C_\alpha(Y) = (lower, upper) : P(lower \leq X \leq upper \mid Y) = 1 - \alpha \tag{2.9}$$

where $C_\alpha(Y)$ is a continuous interval and *lower* and *upper* are the bounds of the interval.

When first being introduced to credible intervals it might appear similar to the frequentist confidence intervals, but they are actually quite different. The credible intervals are the intervals (simply a certain 'width') of the obtained posterior distributions, while the confidence interval is an estimation itself. Meaning that with the frequentist approach the given experiment is designed so that out of a 100 experiments there is a confidence that $100(1-\alpha)\%$ of the experiments can be expected to include the true value of the parameter [10]. Whilst credible intervals are measures of confidence in the statistical model itself. There exist three main credible intervals that are often used in Bayesian Statistics:

- The **Highest Posterior Density interval** (HPD): Finding the value with highest probability density in the posterior, and the area that covers $100(1-\alpha)\%$ on both sides of this point is the HPD credible interval. By finding the highest probability point p^* in the posterior such that

$$1 - \alpha = \int_{X:\text{prob}(X|Y)>p^*} p(X \mid Y) dX, \tag{2.10}$$

  then the HPD is defined as

$$C_\alpha(Y) = X : \text{prob}(X \mid Y) \geq p^* \tag{2.11}$$

  [9].

- The **Equal Tailed interval** (ETI): Is found by estimating the median of the distribution so that there is $(1-\alpha)/2$ mass in each tail [9]. This interval gives the percentiles with the median at 50th percentile, therefore densities are irrelevant. If the 50th percentile is not positioned around the highest probable density, the ETI might not include the values of highest probability. The HPD is typically narrower than the ETI, even though it contains the same amount of mass, since every point inside of it has higher density than every point outside of it [9].

- **Assuming a Mean of the Distribution**: Estimating the mean of the distribution and finding the standard deviation within an area $100(1-\alpha)\%$ of both sides of the mean, typically:

$$\mu \pm 2\sigma. \tag{2.12}$$

2.2 Maximum A Priori

After estimating a credible interval it is possible to find a point estimate of the distribution by choosing the mean of the interval or the median. However, when computing a posterior it involves sampling and this can be computationally expensive, at least in high dimensions. The solution can be to compute the Maximum A Priori (MAP) estimate, which is the mode of the posterior distribution. It reduces the optimization problem, so efficient algorithms can be used [9]. In addition the MAP estimation has many similarities with the well known Maximum Likelihood Estimation (MLE). In fact, by using a uniform prior for the MAP so that the prior assigns equal weights on all possible values of X the MAP is equal to MLE. When finding the MAP estimate, Bayes' theorem is used

$$\begin{aligned}\text{prob}(X \mid Y, I) &= \frac{\text{prob}(Y \mid X, I) \times \text{prob}(X \mid I)}{\text{prob}(Y \mid I)} \\ &\propto \text{prob}(Y \mid X, I) \times \text{prob}(X \mid I).\end{aligned}$$

First the maximum of the parameter X must be found by computing

$$\begin{aligned}X_{MAP} &= \arg\max_X \text{prob}(Y \mid X)\text{prob}(X) \\ &= \arg\max_X \prod_i \text{prob}(y_i \mid X)\text{prob}(X) \\ &= \arg\max_X \sum_i \log(\text{prob}(y_i \mid X)) + \log(\text{prob}(X))\end{aligned}$$

.

If the prior is uniform $\text{prob}(X) = \textit{constant}$, it vanishes under the derivation when maximizing for X and in that particular case $X_{MAP} = X_{MLE}$.

The drawback of using MAP estimation is the same as for MLE; it does not provide a measure of uncertainty [9]. The MAP estimates the mode of the posterior distribution, which is the number that appears most frequently in the posterior. The mode is a point of measure zero, while the mean and median take the volume of the distribution into account. Also, choosing the mode as a summary of a posterior distribution can be a poor choice, since it can be untypical of the posterior distribution [9]. Lets say that the posterior is a continuous distribution including various numbers, then the most likely number does not have to be the most frequent.

3 Model Checking

Bayesian prior-to-posterior inference is dependent on a well defined probability model and can give misleading inference if the model is poorly initialized. Often, the suspect of a bad initialized model is the prior distribution, but the likelihood can make the model equally sensitive, by not making a good fit to the data [4]. Therefore, model checking has been developed, with the two main analysing tools being **posterior predictive checks** and **sensitivity analyses**. Both can be incorporated into the usual prior-posterior analysis.

3.1 Posterior Predictive Checks

The question one should ask after the posteriors are retrieved: is the model consistent with the data? If the model fits the observed data, then sampling S replicated data under the specified model should be consistent with the observed data. After the new posteriors (the posterior predictive distributions) are retrieved for the replicated data, these can be compared with the observed data. If there are systematic differences between the replications and the data, this indicates a potential failure of the model [4]. So the posterior predictive check is the comparison of the real tested quantities $T(Y, X^s)$ and the predicted test quantities $T(Y^{rep}, X^s)$. From there, different metrics can be applied to reveal discrepancies. One such metrics is the p-values, which is the proportion of the S simulations where the test quantities are equal or exceed the real values:

$$T(Y^{rep}, X^s) \geq T(Y, X^s) \quad s = 1, ..., S. \tag{2.13}$$

The p-values are known to many from classical statistics, but in the Bayesian case the usage is different. Extreme p-values does not lead to rejection of the model, instead it sheds light on the limits of the model. There may be need to make further improvements to the model, or the model simply cannot be expected to capture the wanted aspects of the observed data. The solution can be to expand the model in an appropriate way, or check if the data contains the information that one set out to infer. Another way to find systematical differences is to look at the relative difference between the real tested quantities and the predicted test quantities. Also, the prior distribution and posterior distributions can be compared, if the prior distribution contradict the posterior distribution another prior distribution can be chosen, that fits the problem better. This can be compared to the frequentist approach, where the model is iteratively altered to provide a better fit to the data [4]. However, in contrast to the frequentist approach, Bayesian model checking does not requite special methods to handle nuisance parameters, by drawing samples from the posteriors one automatically averages over all parameters in the model [4].

3.2 Sensitivity Analysis

Another question to ask when the model has returned a posterior distributions is: If using another reasonable approach, how much will the posterior change? How sensitive is the result to changes? As an example, the model will often overestimate the uncertainty since all the background knowledge of the data is not incorporated into

the priors. The basic method of sensitivity analysis is to fit many probability models to the same observations [4]. By replacing objective priors with subjective priors that contain substantial prior knowledge, can reveal if the objective prior is influencing the posteriors in an unwanted way. If the model is correctly initialized, the median of the posteriors should not change, since they are in general less sensitive to changes in the model compared to the mean or highest density estimates.

4 Sampling Methods

When sampling from the likelihood and prior the main algorithm used is Monte Carlo sampling. The numerical Monte Carlo methods are widely used in a variety of areas (physics, chemistry, biology,engineering etc). These methods imply statistical simulation which allows omitting writing all differential equations that describe a system and its dynamics, since the process is simulated directly. If the given system can be described by certain probability distribution functions, the random sampling based on these functions allows the passage through the number of steps and achieve a solution. However, the separate results of these samplings should be accumulated in a way that the final result is obtained in the most efficient way. That is why the Monte Carlo technique should be based on a combination of some sampling technique and a selection algorithm. Here there is a variety of algorithms that can be applied like the Metropolis algorithm or Gibbs sampling, where the aim is to govern the movement of the samples towards the most probable result. Unless this algorithm is implemented the step sizes and directions for movement of components of the system are independent and random, making the movement towards the equilibrium state much less efficient. In this thesis, the No-U-Turn sampler was used, which is a Hamiltonian Monte Carlo (HMC) that is a Markov chain Monte Carlo (MCMC) algorithm. The algorithm was developed by M. D. Hoffman and A. Gelman and published in 2014 [11], it is implemented into the PyMC3 library [12]. The HMC avoids the random walk behavior and sensitivity to correlated parameters that often plague many MCMC methods [11]. The drawback of many sampling algorithms when applying them to complex Bayesian models is that they take long time to converge. When the parameters are continuous rather than discrete, HMC suppresses the random walk behavior by transforming the sampling from a target distribution into a problem of simulating Hamiltonian dynamics [13]. This minimizes the cost of the HMC per sample, if the target distribution has dimension D the cost is around $O(D^{5/4})$, compared to the random-walk Metropolis $O(D^2)$ cost [14]. Unfortunately, the HMC is sensitive to two parameters, the step size and the number of steps. This is what the extension to HMC, called the No-U-Turn sampler (NUTs) want to eliminate. The HMC can drop dramatically in efficiency if the step size and number of steps is poorly chosen, requiring costly tuning runs. NUTs was thus developed as a MCMC algorithm that closely resembles the HMC, where the step size is automatically tuned and the number of steps does not need to be initialized [11].

Chapter 3

Background Knowledge

The experimental nuclear physics group at the OCL [15] is well known for their investigation of the γ-ray strength function and nuclear level density in the excitation energy region from ground state and up to the neutron separation energy. To be able to study this region the Oslo-method has been developed [16–19]. The γ-ray strength function measures the average nuclear electromagnetic response, which is directly related to the reduced transition probabilities and γ-transmission coefficients. The nuclear level density is the number of nuclear energy levels per energy unit (MeV^{-1}) as a function of excitation energy. For their method and work the OCL group has gained international attention, thus paving the way for important collaborations with other experimental nuclear physics groups around the world. The laboratory is in addition used to study medical isotopes (e.g. ^{18}F, ^{211}At, ^{205}Bi). In this chapter the instrumental setup of the OCL will be elaborated, as well as the processing of data obtained in a nuclear experiment, this is the background knowledge I mentioned in chapter 2, which is taken into consideration when unfolding. This knowledge is not directly implemented into the prior, but it heavily influences the users understanding of the raw data that is to be unfolded. In this thesis the ^{28}Si(p,p$'\gamma$) reaction was used to test the fully Bayesian unfolding method.

1 Experimental setup

The data used for study of the Bayesian unfolding method has been obtained from an experiment at the OCL (4.6). The cyclotron at the OCL is a MC-35 Scanditronix. The cyclotron accelerates light particles (e.g. p, d, ^{3}He, ^{4}He), with a beam energy between 2 MeV and 47 MeV. The beam is focused with quadruple analyzing magnet, then the dipole magnets bends the beam 90°, before it enters the experimental hall, where it hits the target placed in the center of a target chamber. In the experimental hall the detectors, SiRi and OSCAR are placed. SiRi is a $\Delta E - E$ silicon particle telescope and is either placed in forward or backward position relative to the beam direction. OSCAR is a $LaBr_3$:Ce scintillator array designed to detect γ-rays.

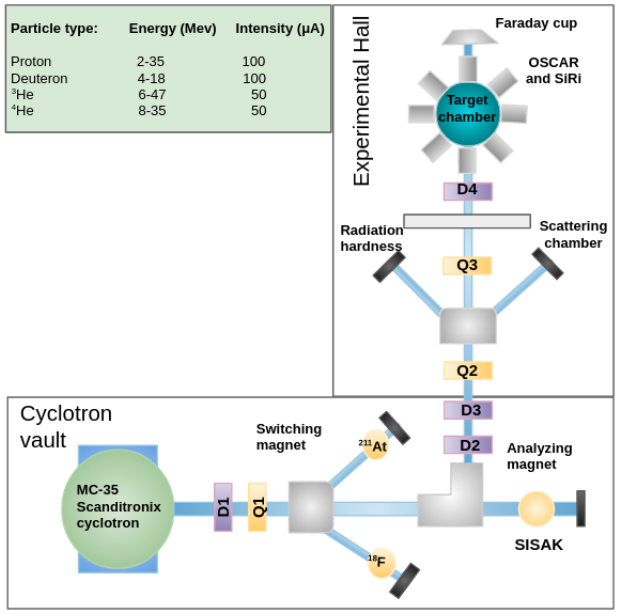

Figure 3.1: Sketch of the setup at the Oslo Cyclotron Laboratory (not to scale)

The particle telescope SiRi

The telescope system consists of two types of silicon detectors, the thin ΔE detector (130 μm) and the thick E detector (1550 μm) [3]. These detectors are designed for measuring the energy, of the charged particle and to differentiate between various particle species [3]. SiRi is designed with 8 modules of E detectors set up behind 8 ΔE detectors, which make a trapezoidal shape. A picture of the SiRi detector and an illustration of the front ΔE detector are shown in figure 3.3.

SiRi is placed 5 cm from the center of the target. The pads of the ΔE detectors correspond to 2° of the scattering angle, that means they cover the angle in forward position $\theta = 47 \pm 7°$ and in backwards position $\theta = 133 \pm 7°$.

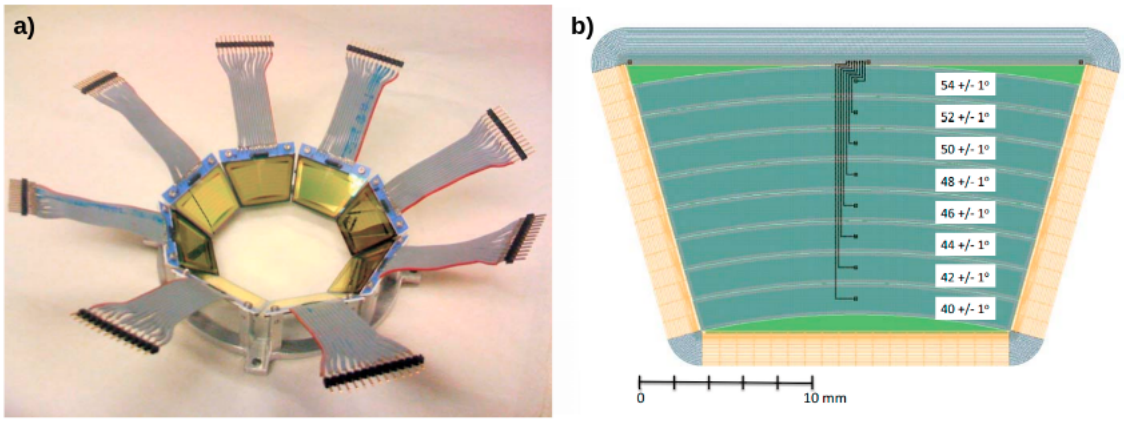

Figure 3.2: **a)** a picture of the SiRi particle telescope. **b)** An illustration of the ΔE detector with the angles in forward position. The both figures are taken from M. Guttormsen *et al.* [3]

Figure 3.3: A picture of the SiRi particle telescope set up for the ^{28}Si(p, p′) experiment.

Since the recoil corrections are important, especially for light nuclei, the measurement of θ must be with an uncertainty of typically $\pm 1^\circ$ [3]. The goal is to be able to identify the initial excitation energy E_x of the reaction product. When charged particles travel through matter, they will interact with electrons that are on its path through the electromagnetic force. This interaction makes the particles lose speed and therefore also lose energy, until stopped completely, depositing all their energy into the detector. Silicon detectors go under the category called semiconductors. When these detectors are connected to a voltage the charge is collected [20]. The signal produced will then be proportional to the energy deposited.

The $\Delta E - E$ technique is based on the energy loss of charged particles in matter, which is given by the Bethe-Bloch equation [20–22]:

$$\frac{dE}{dx} = 2\pi N_a r_e^2 m_e c^2 \rho \frac{Z}{A} \frac{z^2}{\beta^2} \left[\ln \left(\frac{2m_e \gamma^2 v^2 W_{\max}}{I^2} \right) - 2\beta + \delta - 2\frac{C}{Z} \right], \tag{3.1}$$

the parameters can be written as [20]:

A – atomic weight of absorber

Z – atomic number of absorber

N_a – Avogadro's constant

m_e – electron mass

r_e – classical electron radius

I – mean excitation potential

W_{max} – maximum energy transfer per collision

ρ – density of absorber

z – charge of an incident particle

$\beta - \frac{v}{c}$

γ – Lorentz factor$1/\sqrt{1-\beta^2}$

δ – density correction

C – shell correction

$\frac{dE}{dx}$, put in words, means the energy loss per unit length or stopping power. When the incoming particles have relatively low energy, the stopping power and the particle energy have the following relationship

$$\Delta E \sim \frac{z^2}{\beta^2} \sim \frac{mz^2}{E}, \tag{3.2}$$

where m is the mass of the detected particle. The equation above shows that the energy loss depends on the charge and the mass of given particle. When a particle interacts with the $\Delta E - E$-detectors, this principle is used to identify the particle in question. The particle may travel through the thin ΔE detector and lose some of its energy, before the rest is lost in the thick E detector. Since each particle type has different charge and mass they can be distinguished following the Bethe-Bloch formula. These can be represented in so-called 'banana'-plots, where each ΔE channel is plotted with the corresponding E channel (3.7). More information on the typical data representation and processing will be elaborated in the section below 2.

The OSCAR Array

The Cerium doped Lanthanum bromide ($LaBr_3$:Ce) is an inorganic scintillator detector. It stands out among scintillators with its excellent properties, such as an energy resolution around $2.7 - 3.3\%$ FWHM at 661.6 keV, almost perfect light yield proportionality (> 100 keV) and very good intrinsic time resolution (< 1 ns) [23]. OSCAR is fairly new and was officially applied for the first time at OCL in 2018. It consists of 30 $3.5'' \times 8''$(89mm $\times$ 203mm) $LaBr_3$:Ce scintillators. Each scintillator is connected to voltage dividers (LABRVD) and Hamamatsu's R10233-100 photomultiplier tubes [23]. A picture of OSCAR is shown in figure 3.4. Each scintillator and corresponding photomultiplier tube is wrapped in an aluminium housing and is mounted in a icosahedron-shaped frame. The frame can be adjusted to achieve the desired distance between detector and target.

A scintillator has the ability to convert γ-rays into photons in the visible range of the spectrum [20]. These γ-rays will excite atoms in the crystal, which again de-excite with visible light photons. These are collected by photocathodes that transform the photons into electrons. The photomultiplier amplifies the electric signal further. When penetrating the matter of the detector, there are various interactions the γ-ray can undergo, where the three most important are: photoelectric effect, Compton scattering and pair creation:

Figure 3.4: Picture of the detector OSCAR array

- **Photoelectric effect:** Full absorption of the γ-ray by an atomic electron. When absorbing all the energy the atom the ejects the now energetic electron, the energy of the electron being

$$E_e = E_\gamma - \text{B.E}, \tag{3.3}$$

where B.E is the binding energy of the electron [20].

- **Compton scattering** is the elastic collision between a γ-ray and a free electron [24]. The electrons in the detector are not free, however if the γ-energy is high with respect to the electron binding energy, the electrons can be considered as free [20]. The energy of the scattered γ-ray, $E_\gamma^{'}$ is thus

$$E_\gamma^{'} = \frac{E_\gamma}{1 + \frac{E_\gamma}{m_e c^2}(1 - \cos\theta)}, \tag{3.4}$$

where θ is the angle between the incident γ-ray and the electron [20]. In figure 3.5 the Compton distribution is that of the recoil electrons. The maximum recoil energy is known as the Compton edge, and is also illustrated in the figure.

- **Pair Production** occurs when the incident γ-ray has an energy at least two times larger than the mass of the electron (1022 keV) [24]. The process involves the transformation of a γ-ray into an electron-positron pair [20]. Another necessity for pair production to take place is the presence of a third body (nucleus), for the momentum conservation. The positron will annihilate with a nearby electron creating two new γ-rays of energy 511 keV in the process. These can both escape the detector, or one escapes the detector, or none of them escape. In figure 3.5 the resulting single escape ($E_{full} - 511$ keV) and double escape

(E_{full} − 1022 keV) peaks are shown. The annihilation peak in the figure is when the positron annihilates with the electron [25].

The true photon spectrum thus differs from the measured spectrum, which makes it necessary to unfold the raw spectrum. The unfolding is described in 4. In figure 3.6, the ^{28}Si(p,p$'\gamma$) reaction is illustrated. A kinematics analysis of the given experiment is carried out beforehand, and the detectors will be setup accordingly. The beam energy, target thickness, and the type of reactions expected, decides how far OSCAR is placed from the target and also if SiRi is put in backward or forward position. The application Qkinz is used to calculate how the 8 SiRi angles will register particles based on the kinematic conditions used as input [26].

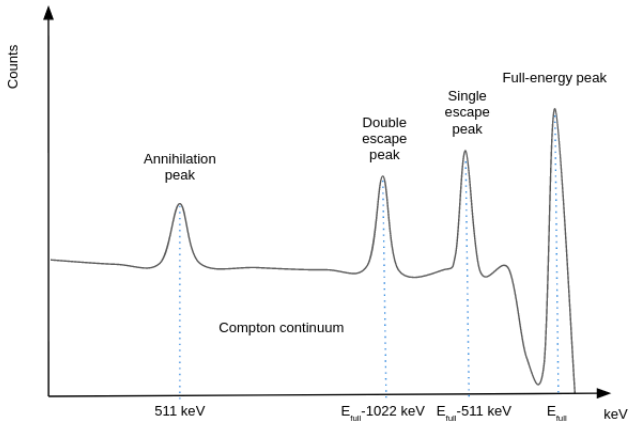

Figure 3.5: Illustration of a measured spectrum for an incoming mono-energetic γ-ray.

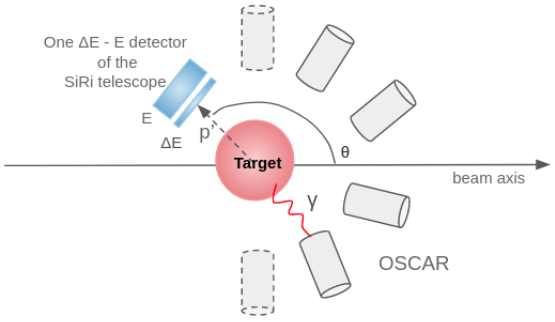

Figure 3.6: A typical reaction of an the experiment at OCL, with the particle telescope in backward position. The $LaBr_3$:Ce of OSCAR are placed around the target as indicated in the figure.

2 Data Processing

When a particle interacts with SiRi, it will free electrons and these will be collected and form an electric pulse. Similarly for OSCAR, the photons are transformed to electrons which form an electric signal. The signal travels through a digital pulse processor called Digital Gamma Finders (DGF). Each DGF has 16 channels and each channel is connected to a separate detector, 64 channels for each ΔE detector, 8 channels for each E detector and 30 channels for each $LaBr_3$:Ce detector. The DGF simultaneously measure the amplitude and analyse the pulse shape for the channels [27]. The signals travel through optic cables and are stored on a computer. These signals are time stamped, and the events in the E detectors and OSCAR arriving within ± 1 μs with respect to a reference ΔE detector, are grouped together. The raw data are then ready for further processing.

2.1 Data processing for SiRi

The total energy deposited in the telescope detector SiRi is proportional to the charge of the particle. Furthermore, the signal that is produced has a pulse shape, and there is proportionality between the integral of the pulse shape observed and the charge. The pulse height and pulse integral are again also proportional. The pulse heights are distributed over sequential channels, and they are therefore proportional to the energy deposited in the detector. However, they do not contain information on the absolute value of the energy. The first step in the data processing is to connect the channel number to their respective energies. Often, some of the peaks can be established, i.e the energy is known beforehand. By setting a dependency of that energy to the given channel number, and assuming a linear response, the proportionality between energy and channel becomes

$$E = a \times N_{ch} + b, \tag{3.5}$$

where E is the energy, N_{ch} is the channel number, a is the gain ($\frac{keV}{\text{ch}}$) and b is the energy shift. For this assumption, at least two peaks must be identified. There is a total of 64 combinations of $E - \Delta E$ detectors, since there are 8 ΔE detectors for each 8 E detector. Due to the different performance and attached electronics of each detector, all the combinations have their distinctive energy-to-channel correspondence. Figure 3.7 shows one combination of a calibrated $E - \Delta E$ spectrum, the so-called banana spectrum, for the ^{28}Si(p, p$'$) reaction. As the Bethe-Bloch formula 3.1 states, each banana, is designated to a different type of scattered particle. The lowest banana is designated to the lightest scattered particle, the proton (Z= 1). As indicated in the figure, the first blob to the far right of the (p, p$'$) channel is the ground state of ^{28}Si, and to the left of it is the first excited state of the 2^+ of ^{28}Si at energy 1779 keV. The peaks calibrated are compared to the theoretically predicted values taken from Qkinz [26].

The time stamp of the E-detectors with respect to the different ΔE detectors, also has to be accounted for. The aim is to find the true coincidences between a given E and ΔE detector. A time gate can be applied, from when the particle hits the ΔE detector until it is stopped in the E detector. Typically, a graphical cut is made to

select the events from both the $E - \Delta E$ spectrum and the time spectrum, to get the events from both detectors that originate from the same event.

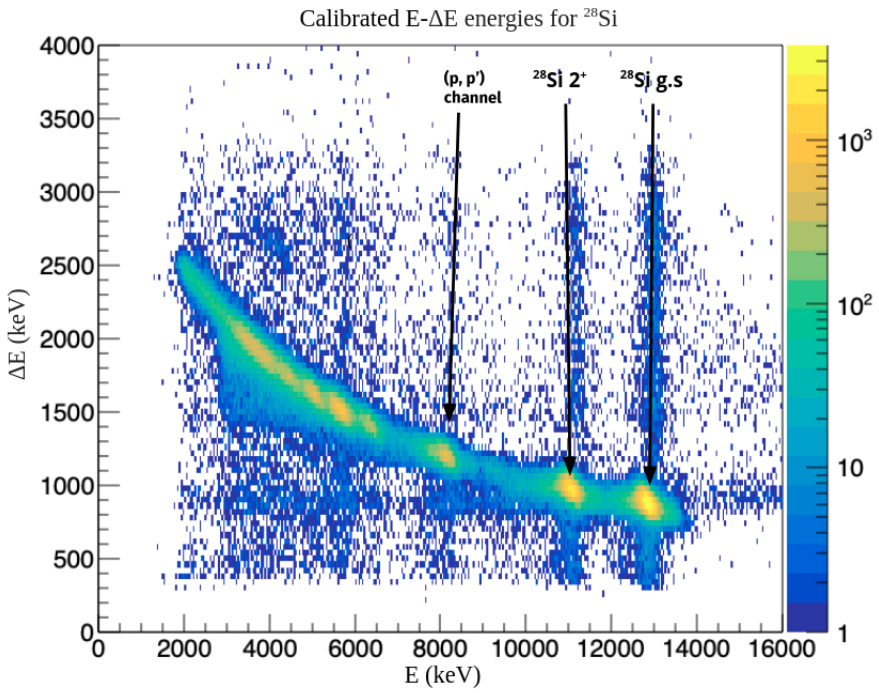

Figure 3.7: The calibrated ΔE-E spectrum for the ^{28}Si(p, p′) experiment, with SiRi in backward position. The x-axis displays the energy deposited in the E-detector, and the y-axis the energy deposited in the ΔE detector.

2.2 Data processing for OSCAR

The main difference between SiRi and OSCAR is that the channel number and energy correspondence is not always linear for OSCAR. The qualities that make the $LaBr_3$:Ce good detectors is, unfortunately, the origin of the nonlinear effects [23]. The high light yield and the short decay time induces a large current in the PMTs, which leads to saturation effects. A polynomial calibration is therefore recommended,

$$E = \beta_0 + \beta_1 N_{ch} + \beta_2 N_{ch}^2 + O(N_{ch}^2) \tag{3.6}$$

where E is the energy, N_{ch} is the channel number and $\beta_0, \beta_1, \beta_2, ...$ are the calibration fit coefficients. An in-beam nuclear reaction can be used to produce monochromatic γ-rays for the calibration of OSCAR. The $^{28}Si(p_p{}'\gamma)$ reaction is used for calibration. By calibrating $^{28}Si(p, p')$ for the particle detectors, using the process mentioned above, one uses the knowledge of the excitation energies to identify the γ-peaks in OSCAR. Achieving a good polynomial fit for these γ-transitions of $^{28}Si(p, p'\gamma)$ can give values for the coefficients to be used in the calibration for the actual nuclear reaction under investigation.

The time alignment for the OSCAR detectors need to be treated in reference to the particle detectors. When a particle hits the ΔE detector and the corresponding E detector, a γ-ray from the same reaction might be incident in one of the $LaBr_3$:Ce detectors. Therefore, the zero time reference for one scintillator should correspond with the zero time reference of the particle telescopes. The ΔE detectors have better time resolution compared to the E detector. It is thus best to consider the time difference between the signals in the ΔE and the $LaBr_3$:Ce, to optimize the resolution in the time spectrum. All the combinations of $LaBr_3$:Ce and ΔE (64×30) should be accounted for. When combining all the 30 scintillator detectors with one ΔE detector one sees high intensity lines, these are the prompt peaks for $LaBr_3$:Ce. The prompt peaks are then shifted to the 0 time value as can be seen in figure 3.8 a). In the prompt peak lies the true events that coincide with a particle detected in the particle telescopes. To get as many true events as possible a narrow time gate is often put on this peak (see figure 3.8 b)). As sorted events are taken within $\pm 1\mu$s, it will also include signals that do not come from the true coincidences. After applying a gate on the prompt peak there will still be random background present within the time gate. The solution is to apply another gate on one of the other beam pulses, of the same width as the prompt gate 3.8. This is then subtracted from the prompt peak.

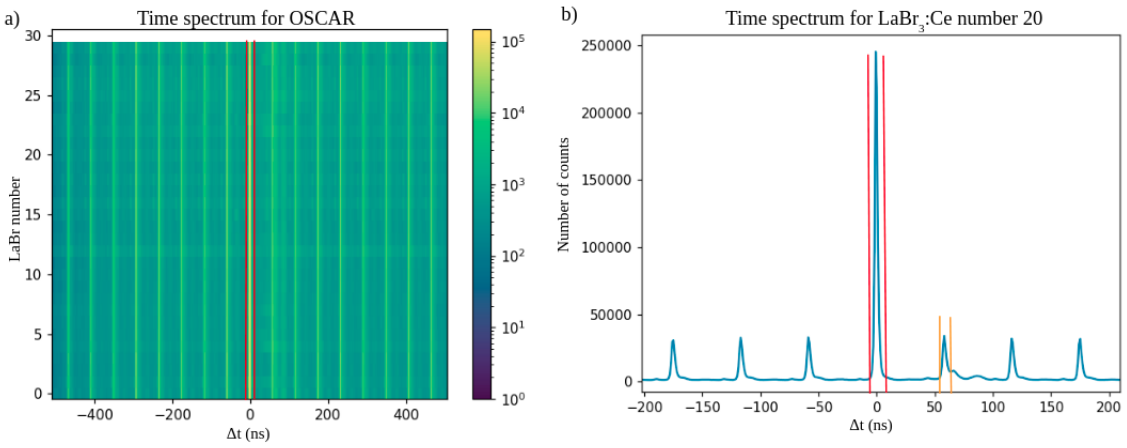

Figure 3.8: a) Time spectrum for all the 30 $LaBr_3$:Ce (y-axis) and time (x-axis). b) The projection of $LaBr_3$:Ce number 20 on the x-axis. The red line is the time gate on the prompt peak, and the orange line is the gate on one of the beam pulses that are subtracted from the prompt peak.

2.3 The raw matrix

After the calibration and time alignments of all detectors the data can be presented in the raw coincidence matrix. The matrix contains the γ-ray spectra for each excitation energy. Figure 3.9 shows the background-subtracted raw matrix for the ^{28}Si(p, p$'\gamma$) reaction. This matrix is put together based on the data processing described in this chapter, where the data for all the detectors with the calibration and time alignment is included. This means that the matrix contain the events from the detection of a scattered proton and the γ-ray emitted in the same reaction. The background subtraction should have excluded most false coincidence events. In figure 3.9 one might notice that there are counts below the diagonal. These leftover counts should fluctuate around zero. The γ-energies emitted from a nucleus can not exceed the excitation energy of the nucleus, that would violate the energy conservation laws. In further analysis the counts exceeding the diagonal are usually cut. The coincidence matrix is the main input in the present study of the Bayesian unfolding method.

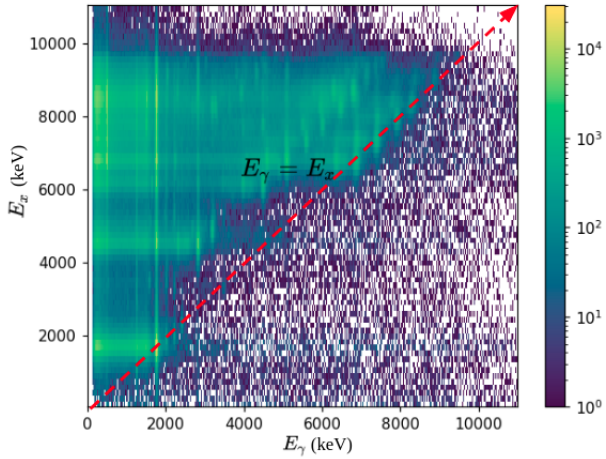

Figure 3.9: The raw coincidence matrix for the ^{28}Si(p,p$'\gamma$) reaction

Part II

Part 2: Methods and Implementation

Chapter 4

The Unfolding Methods

In the section 1 the motivations behind why unfolding methods are used on γ-ray spectra will be outlined. The response matrix used in this thesis and how it is obtained will be explained in section 2. Section 3 describes the iterative unfolding method and in section 4 the Fully Bayesian Unfolding will be explained.

1 Why Unfold?

In the introduction to statistical analysis one often starts by learning about parameter estimation. The aim is to estimate the parameters of a distribution by fitting to a given dataset. As an example, in many cases one can assume that the dataset follows a normal distribution for which one can obtain the mean and the standard deviation. Another, more complicated branch of statistical analysis, is non-parametric estimation. In this case, the data does not rely on parameters that can be estimated, but rather on ranking or order. For example, data ordered in a histogram is a non-parametric estimate of a probability distribution. In the field of nuclear physics, the problem of estimating probability distributions, where no parametric form is available, is the main part of the data analysis and is often referred to as unfolding [28].

In signal processing one learns that an input signal is registered in the receiving device which again provides an output signal. The input signal differs from the output signal since the receiving device has a certain response to the incoming signal. So, by knowing the input and output signal one can use convolution and find the response of the device. In nuclear physics, the input signal originates form a γ-ray, but it is smeared by the detector response. The output signal is folded with the response function of the detector device. Therefore, the unfolding is a deconvolutional problem where the goal is to find the true distribution of the original signal that hit the detector. Figure 4.1 helps visualize the problem: one can imagine a sharp peak with N counts, it is then folded so the counts are not longer all in the peak but are redistributed according to the detector response. This is one of the first challenges one meets when analysing γ-ray spectra. It is impossible to get back to the exact true

signal, but with an accurate knowledge of the detector response, one can retrieve a good estimation of the true signal, down to the precision of the detector resolution (shown in the middle panel of figure 4.1).

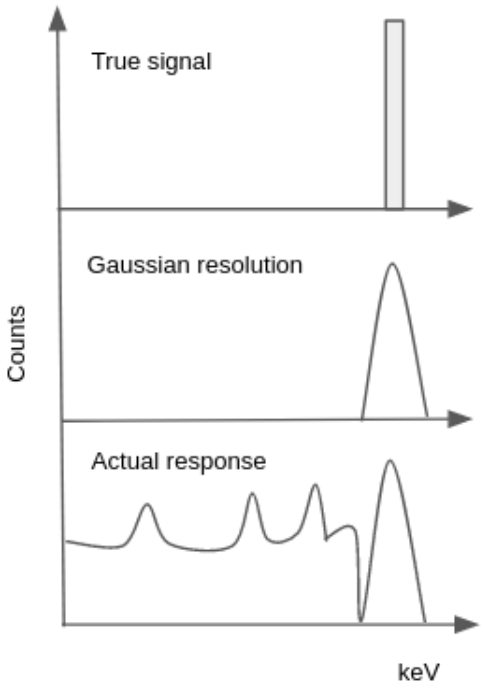

Figure 4.1: Upper panel: the true signal hitting the detector. Middle panel: folded with the detector resolution. Bottom panel: actual response of the detector.

1.1 The Bias-Variance tradeoff in Unfolding

As mentioned before, if there exists a probability density function (PDF), $g(y)$, as a function of a random variable y and the wish is to determine the PDF by sampling from data, there are two options. If the PDF has a parametric form that is known, $g(y;\boldsymbol{\phi})$, where $\boldsymbol{\phi}$ is a set of vectors of these parameters $\boldsymbol{\phi} = [\phi_1, \phi_2, ..., \phi_n]$, one can use the method of maximum likelihood to obtain the estimators $\hat{\boldsymbol{\phi}}$. If no parametrization is available, the random variable y can be represented with histograms, with D number of bins [28]. An advantage of representing the data in histograms is that the expectation value μ_i of the number of entries in a bin i can be calculated [28]. The sum of the expectation values and the probability of y to be found in bin i are: $\mu_{tot} = \sum \mu_i$ and $p_i = \frac{\mu_i}{\mu_{tot}}$.

When unfolding the data one infers the estimators for the D parameters $\mu = [\mu_1, ..., \mu_D]$, or, as another possibility, the probabilities p [28]. The true values y of a nuclear experiment will be subjected to random fluctuations due to the measuring process, which introduces a non-trivial aspect of this problem: there exists a true but unknown y for each measured value x [28]. If the detector was perfect, the measured value would be the true value but this is not the case. The measured value and the true value are therefore, as mentioned in the last

section, related through folding (convolution):

$$g_{meas}(x) = \int R(x \mid y) g_{true}(y), \tag{4.1}$$

where $R(x \mid y)$ is the response function for the detector used in the given experiment, and can be interpreted as a conditional probability

$$R_{ij} = P(\text{observed in bin } i \mid \text{true value in bin } j). \tag{4.2}$$

By summing over i, that is all observed values in the bins, the efficiency ϵ of the detector can be calculated for each j which is the bins of the expected values of y

$$\sum_{i=1}^{N} R_{i,j} = \epsilon_j. \tag{4.3}$$

The response function and how it is obtained will be elaborated further in section 2. Breaking the integral over y in equation 4.1 into a sum over bins [29], the expected number of events to be observed in bin i is

$$f_i = \sum_{j=1}^{D} R_{ij} u_j, \quad i = 1, \ldots, N \tag{4.4}$$

where

$$g_{meas}(x) \to f_i$$
$$g_{true}(y) \to u_j$$

Where $\mathbf{f} = [f_1, \ldots, f_N]$ is the expectation values of the events in the bins of the observed variable x and $\mathbf{u} = [\mu_1, \ldots, \mu_D]$ is the expectation values of the true value y. One can be tempted to think that by unfolding one obtains the original signal hitting the detector, but this shows that by unfolding one merely retrieves the expectation value of the true value. The counts in each histogram, the data, are given by the vector $\mathbf{n}= [n_1, \ldots, n_N]$. These will differ from $\mathbf{f}$, since they are the actual number of entries observed and $\mathbf{f}$ are expectation values. Suppose that there is background events for each bin i, the expectation value β_i for the number of entries that originate from these, also need to be accounted for. To relate the data with the expectation value of the measured spectrum $\mathbf{f}$, and the true spectrum $\mathbf{u}$, together with the response and the background events $\boldsymbol{\beta}$, the expected value of the data is given as

$$E[\mathbf{n}] = \mathbf{f} = R\mathbf{u} + \boldsymbol{\beta}. \tag{4.5}$$

The inversion gives

$$\mathbf{u} = R^{-1}(\mathbf{f} - \boldsymbol{\beta}). \tag{4.6}$$

In nuclear physics experiments, the data is independent and Poisson distributed [28] and the likelihood function is given as

$$L = P(n_i; f_i) = \prod_{i=1}^{N} \frac{f_i^{n_i}}{n_i!} e^{-f_i}. \tag{4.7}$$

To determine the maximum likelihood for f_i the natural logarithm of the likelihood function:

$$\ln L = [\sum_{i=1}^{N} \ln(f_i^{n_i}) + \ln\left(e^{-f_i}\right) - \ln(n_i!)] \tag{4.8}$$

$$= \sum_{i=1}^{N} [-f_i - \ln(n_i!) + n_i \ln(f_i)]. \tag{4.9}$$

$$\tag{4.10}$$

Taking the derivative of this gives

$$\frac{d \ln L}{df_i} = 0 \tag{4.11}$$

$$\Rightarrow -1 + \frac{n_i}{f_i} = 0 \tag{4.12}$$

$$\Rightarrow n_i = f_i \tag{4.13}$$

$$\Rightarrow \mathbf{f} = \mathbf{n}. \tag{4.14}$$

This shows that by taking the maximum likelihood estimation of f_i gives the expectation value of the data n_i and the inversion problem from 4.6 thus becomes

$$\hat{\boldsymbol{u}} = R^{-1}(\mathbf{n} - \boldsymbol{\beta}), \tag{4.15}$$

where $\hat{\boldsymbol{u}}$ is the estimator of $\boldsymbol{u}$. This is a non-parametric problem subjected to random fluctuations, as illustrated in figure 4.2. By performing a matrix inversion on the raw data, the resulting spectrum shows large fluctuations as is illustrated in figure 4.2 c).

The reason for these large fluctuations lies in the fact that the expectation values of the data $\boldsymbol{f}$ is not known beforehand, only the data itself $\boldsymbol{n}$. By applying R^{-1} to the expectation value $\boldsymbol{f}$ it would indeed give the true value $\boldsymbol{u}$. Since it is applied to the data, which does have statistical fluctuations, it will result in a large variance, giving oscillations in the estimators $\hat{\boldsymbol{u}}$. However, it is worth noting that by performing a matrix inversion there is no bias present:

$$E[\hat{\boldsymbol{u}}] = R^{-1}(E[\boldsymbol{n}] - \boldsymbol{\beta}) = \boldsymbol{u}. \tag{4.16}$$

Taking a look at the covariance between independent bins gives

$$\begin{aligned} U_{ij} =& \operatorname{cov}\left[\hat{u}_i, \hat{u}_j\right] \\ =& E[(\hat{u_i} - E[\hat{u_i}])(\hat{u_j} - E[\hat{u_j}])] \\ =& \sum_{k,l=1}^{N} (R^{-1})_{ik}(R^{-1})_{jl} E[(n_k - f_k)(n_l - f_l)] \\ =& \sum_{k,l=1}^{N} (R^{-1})_{ik}(R^{-1})_{jl} \operatorname{cov}[n_k, n_l] \\ =& \sum_{k=1}^{N} (R^{-1})_{ik}(R^{-1})_{jl} f_k, \end{aligned} \tag{4.17}$$

since the covariance of the data $\boldsymbol{n}$ is independent and Poisson distributed, and remembering that both k and l both sum up to N:

$$\begin{aligned} \operatorname{cov}[n_k, n_l] &= E[(n_k - E[n_k])(n_l - E[n_l])] \\ &= E[n_k n_l] - E[n_k]E[n_l] - E[n_k]E[n_l] + E[n_k]E[n_l] \\ &= E[n_k n_l] - E[n_k]E[n_l] \\ &\Rightarrow E[n_k^2] - E^2[n_k] \\ &= E[n_k^2] - f_k^2 = \operatorname{Var}(n_k) = f_k \end{aligned} \tag{4.18}$$

.

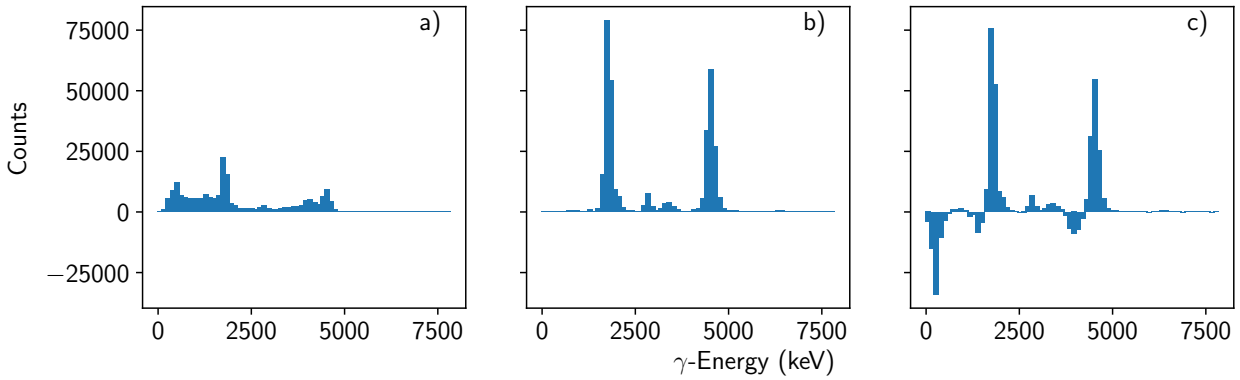

Figure 4.2: a) Example of raw data, b) Unfolded spectrum, using the iterative unfolding method, c) the spectrum obtained by matrix inversion.

In matrix form, the covariance can be denoted as $\mathrm{cov}\left[\hat{u}_i, \hat{u}_j\right] = U_{ij}$. Using the Rao-Cramér-Frechet (RCF) inequality, which gives a lower bound of an estimator's variance [29, p. 76], for the case of multiple parameters, the inverse of the covariance matrix becomes

$$\begin{aligned}\left(U^{-1}\right)_{kl} &= -E\left[\frac{\partial^2 \log L}{\partial u_k \partial u_l}\right] \\ &= \sum_{i=1}^{N} \frac{E\left[n_i\right] R_{ik} R_{il}}{f_i^2} \\ &= \sum_{i=1}^{N} \frac{R_{ik} R_{il}}{f_i}.\end{aligned} \tag{4.19}$$

If the minimal estimation of the variance is multiplied on both sides by U and twice by R^{-1}, then summed over the indices, it will give the same answer as the covariance in equation 4.17. As the RCF gives the minimal variance of the parameters, by performing a matrix inversion one actually obtains the minimum variance for all unbiased estimators, even though it has shown itself to be large (see figure 4.2). It also results in an "un-physical" solution that is more or less impossible to interpret.

By applying other unfolding methods the reduction of variance is possible, and an interpretation of the data can be achieved. All methods will introduce bias to this problem. Therefore, one has to be prepared to accept a certain bias (systematic error) in exchange for reducing the variance (statistical error), and in return attain data that can be interpreted.

2 The Response Matrix

The response function and efficiency of OSCAR is obtained by simulating the detector response with GEANT4 [30–33], a simulation toolkit, and compare these with experimentally measured mono-energetic in-beam γ-ray spectra. GEANT4 is a toolkit for simulating the passage of particles through matter. The Full-Width Half Maximum (FWHM) of the detector must be extracted from data. The FWHM for OSCAR was analyzed by fitting a Gaussian on an exponential background of the in-beam γ-ray lines from a $^{28}\mathrm{Si}(\mathrm{p}, \mathrm{p}')$ reaction. To obtain FWHM for all γ-energies, the function for FWHM was estimated to be $(\mathrm{FWHM}(E))^2 = c_0 + c_1 * E + c_2 * E^2$, with fitted parameters $c_0 = 1.13 \times 10^{-3}$, $c_1 \approx 0$ and $c_2 = 2.05(42) \times 10^{-4}$ [34]. In GEANT4, the whole geometry

of the detector setup is implemented using a class called *Constructed Solid Geometry* (CGS) solids. Then *G4EmStandardPhysics* is used, which implements standard electromagnetic interactions and neutron interactions through data-driven high-precision cross-sections [34]. The target is placed in the center of the sphere of the detector array and the energy deposited in each detector is stored in a ROOT tree [35]. The simulations are then compared to the experimental results. It was found a deviation of about $10 - 15\%$ between the simulated spectra and the experimental spectra in the number of counts below the full energy peak. The simulated spectra tended to underestimate the amount of counts. Therefore, the Compton contribution was corrected for with a factor of 1.15 for the present response function. It should be noted here that the work on the response of OSCAR is still

under completion. The response function used in this thesis was made whilst there still was analog electronics, which since has been replaced by digital electronics [34].

The aim of comparing simulations to experimental in-beam γ-rays is to ensure a correct proportion of counts belonging to the different interactions that the γ-ray undergoes within the active volume of the detector. These counts can then be normalized so that

$$p_{FE} + p_{SE} + p_{DE} + p_{CP} + p_{511} = 1, \tag{4.20}$$

p denotes probability, and where p_{FE} is the probability for the full energy peak, p_{SE} for the single escape, p_{DE} for the double escape, $p511$ for the annihilation peak and p_{CD} for the Compton distribution. In an ideal case, the response functions would be found for all incident γ-energies E_γ, but unfortunately they can only be measured for a few mono-energetic γ-lines. The normalized values from simulations, are stored in files, where the p_{FE}, p_{SE}, p_{DE}, $p511$, together with the γ-energy, FWHM and total efficiency are stored in one file, while the p_{CD} for each γ-energy, are in separate files. This simplifies the interpolation of the energies, since interpolating the Compton spectra differs from the interpolation of the peaks. The first file containing the energy peaks, can be interpolated by adding a Gaussian distribution at the interpolated peaks with the same intensity and energy resolution [19]. The Compton distribution has to be interpolated more carefully, since its shape will differ for each incident energy. Therefore, the interpolation of the Compton distributions is separated into three parts:

- Energies below the back scattering energy: linear interpolation.
- Energies between the back scattering angle and the Compton edge: interpolated using the fan method introduced by Guttormsen et al [19].
- Energies above the Compton edge: linear interpolation.

The backscattering energy is defined as

$$E_{\mathrm{b.sc}}^{meas} = E_\gamma^{true} / \left(1 + \frac{2E_\gamma^{true}}{m_\mathrm{e}c^2}\right), \tag{4.21}$$

where E_γ^{true} is the true energy of the γ-ray.

Figure 4.3 illustrates how the fan interpolation of the Compton spectra is done. As seen in the figure the angle of which the γ-ray scatters goes from 0 degree to 180 degrees. A reasonable approach is then to interpolate between channels that have the same angle for the scattered γ-ray [19]. The Compton scattering is when a γ scatters of the nucleus and transfers energy to the electron by the amount

$$E_e = E_\gamma - \frac{E_\gamma}{1 + \frac{E_\gamma}{m_e c^2}(1 - \cos\theta)} \tag{4.22}$$

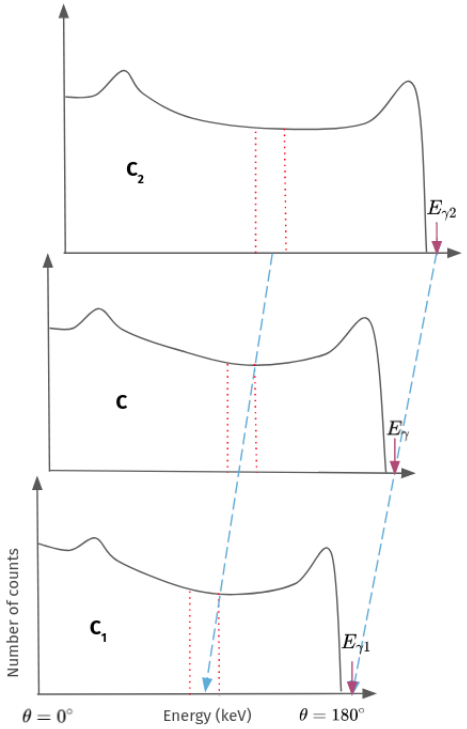

Figure 4.3: Illustration of the fan method. $\Delta\theta$ increases with the energy

where θ is the angle between the incident γ-ray and scattered γ-ray. This energy is heavily dependent on the angle and the incident γ-ray energy E_γ, so these have to be taken into account for the normalization of the interpolated Compton spectra: [19]

$$c(E) = \left(\frac{\mathrm{d}E_e}{\mathrm{d}\theta}\right)^{-1}_{E_\gamma} \left[c_1\left(E_{e1}\right)\left(\frac{\mathrm{d}E_e}{\mathrm{d}\theta}\right)_{E_{\gamma 1}} + \frac{E_\gamma - E_{\gamma 1}}{E_{\gamma 2} - E_{\gamma 1}} \times \left(c_2\left(E_{e2}\right)\left(\frac{\mathrm{d}E_e}{\mathrm{d}\theta}\right)_{E_{\gamma 2}} - c_1\left(E_{e1}\right)\left(\frac{\mathrm{d}E_e}{\mathrm{d}\theta}\right)_{E_{\gamma 1}}\right)\right]. \tag{4.23}$$

Here E_γ, $E_{\gamma 1}$ and $E_{\gamma 2}$ are the incident energies E_γ for the interpolated spectra a end the reference spectra, while E_e, E_{e1} and E_{e2} are energies of the electrons deposited in the detector due to the Compton scattering.

After the interpolation of each γ-energy the distribution are placed in a matrix with the rows representing the true γ-energies E_γ^{true}, and the columns are filled with the probability distributions for each measure γ-energy E_γ^{meas}. So for each row i the probability distribution over the columns is

$$R_i = p_{CD} + p_{FE} + p_{SE} + p_{DE} + p_{511} \tag{4.24}$$

Each bin, i.e measured energy, in the row, contains the summed probabilities of an event belonging to one of the five components for that bin, and each row should sum up to 1. Figure 4.4 shows an example of a response matrix.

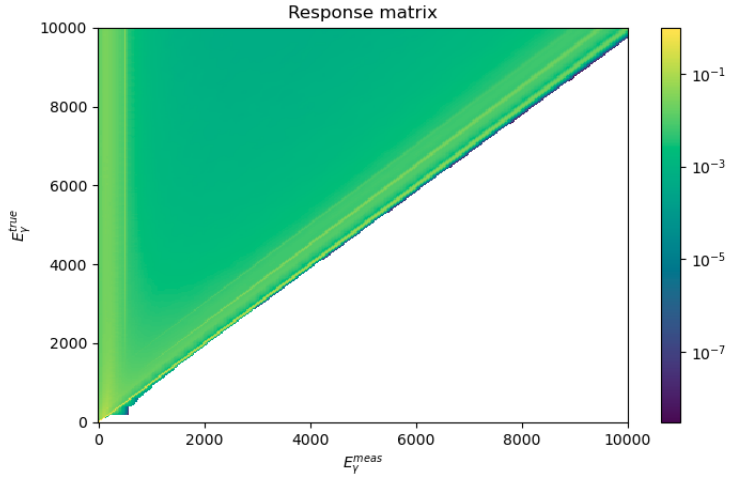

Figure 4.4: The Response Matrix of OSCAR with γ-energies from 0 to 10 MeV

The response matrix is modelled as conditional probability distributions [36]

$$R = P(E_\gamma^{meas} \mid E_\gamma^{true}), \tag{4.25}$$

where the intensities in the contour plot represent the detection probabilities for each E_γ^{true}. The probabilities together with the FWHM, total efficiency and γ-energy will be stored in a table, figure 4.5 shows an example of this.

	E	fwhm_abs	fwhm_rel_%	fwhm_rel	eff_tot	pcmp	pFE	pSE	pDE	p511
0	0.0	0.000000	19.605113	0.196051	1.18810	0.493140	0.506860	0.0	0.0	0.0
1	30.0	5.589579	18.631929	0.186319	1.18507	0.502992	0.497008	0.0	0.0	0.0
2	60.0	10.595247	17.658744	0.176587	1.18204	0.512844	0.487156	0.0	0.0	0.0
3	90.0	15.017004	16.685560	0.166856	1.17901	0.522695	0.477305	0.0	0.0	0.0
4	120.0	18.854851	15.712376	0.157124	1.17598	0.532547	0.467453	0.0	0.0	0.0
5	150.0	22.108788	14.739192	0.147392	1.17295	0.542399	0.457601	0.0	0.0	0.0

Figure 4.5: Table for the response matrix, interpolated with 30 keV energy-bins

3 The Iterative Unfolding Method

The iterative unfolding method explained here was developed by Guttormsen et al. in 1996 [19]. The method was first written in FORTRAN, but has recently been remade in the Python library package OMpy [37]. In this thesis only the OMpy package has been used for unfolding.

3.1 The Iterative method

The measured spectrum $\boldsymbol{f}$ is represented by

$$\boldsymbol{f} = R\boldsymbol{u} \tag{4.26}$$

Where R is the response matrix of the discrete probabilities $R_{ij} = R\left(E^{meas}_{\gamma,j}|E^{true}_{\gamma,i}\right)\Delta E^{meas}_{\gamma}$ [36]. The folded spectrum $\boldsymbol{f}$ is the expected number of events in the bins of the actual observed variable E^{meas}_{γ} [28], and the matrix $\boldsymbol{u}$ is the expected values for the histogram of E^{true}_{γ}. As mentioned before there is background in the experimental data $\boldsymbol{B}$. The expected folded matrix, the response matrix and the expected true histogram is therefore related by

$$\boldsymbol{E}[E^{meas}_{\gamma}] = \boldsymbol{f} = R\boldsymbol{u} + \boldsymbol{B}. \tag{4.27}$$

Before the iterative method is applied the background is usually subtracted, as explained in section 2 in chapter 3. The aim of the unfolding is to obtain $\boldsymbol{u}$ given the raw data $\boldsymbol{r}$. As discussed before, the straight-forward inversion of the problem, where R^{-1} is used on the expected value of the data $\boldsymbol{E}[E^{meas}_{\gamma}] = \boldsymbol{f}$, will result in a spectrum with large oscillations between neighbouring bins [28]. Therefore, the iterative method is applied instead to approximate $\boldsymbol{u}$. The iterative method is here presented step-vise:

- At iteration $l = 0$, the initial spectrum, for the expected true spectrum $\boldsymbol{u}$ is set to be $\boldsymbol{u_0} = \boldsymbol{r}$
- Then the folded spectrum is calculated $\boldsymbol{f_1} = R\boldsymbol{u_1}$
- The initial spectrum is then updated to $\boldsymbol{u_{l+1}} = \boldsymbol{u_l} + (\boldsymbol{r} - \boldsymbol{u_l})$
- The above steps are repeated over a number of iterations chosen beforehand, until $\boldsymbol{f} \approx \boldsymbol{r}$. This is checked by taking the root-mean-square error of $\boldsymbol{f} - \boldsymbol{r}$ and the relative fluctuations in $\boldsymbol{u_l}$, which are estimated as $\frac{|\boldsymbol{u}_{l,i} - \langle \boldsymbol{u}_l \rangle|}{\langle \boldsymbol{u}_l \rangle}$, where $\langle \boldsymbol{u}_l \rangle$ is a smoothed version of the spectrum $\boldsymbol{u_l}$ [36].

The last point in the list describes the score method used to choose the best solution. In OMpy the fluctuations and the χ^2 is weighted

$$\text{score} = (1-p) \times \chi^2 + p \times \text{fluctuations}, \tag{4.28}$$

where p is the weight of the score and should range between $[0, 1]$. The weight is set by default to 0.23 in OMpy, but can be changed by the user. The iterative method can potentially lead to matrix inversion if given enough iterations, and this will result in a very good χ^2 value, but given large fluctuations as discussed before. Therefore the fluctuations are given a weight in the scoring when choosing the best solution for the unfolded spectrum. After the iterations are performed the unfolded matrix with the smallest score, is returned.

The Uncertainty Propagation of the Iterative Method

OMpy offers a method to estimate the statistical uncertainties for the unfolded matrix based on the Monte Carlo (MC) technique [36]. The method creates a number of randomly perturbated copies of the raw data. The random variables are the number of counts that are present in each bin k of the $E_\gamma - E_x$ coincidence matrix. It is assumed that the variables are independent and follow a Poisson distribution with parameter λ_i,

$$P_\lambda = p(n \mid \lambda) = \frac{\lambda^n}{n!} e^{-\lambda}, \tag{4.29}$$

where the number of counts n_k in bin k of r is an estimate for the Poisson parameter λ_k. As mentioned in previous sections the estimator λ is unbiased, since $\langle n \rangle = \lambda$. To generate an ensemble of $\boldsymbol{r}_k$ matrices, the counts in each bin k are replaced by a MC-sample from the distribution P. This creates N matrices where all are sampled from the distribution above. These matrices then represent different realizations of the experiment [36]. Defining $\tilde{\boldsymbol{r}} = [\tilde{r}_k^1, \tilde{r}_k^2, ..., \tilde{r}_i^N]^T$ as the vector for all the bins k of the N matrices. Making it possible to calculate a mean $\langle \tilde{\boldsymbol{r}}_k \rangle$,

$$\langle \tilde{\boldsymbol{r}}_k \rangle = \frac{1}{N} \sum_{m=1}^{N} \tilde{r}_k^m, \tag{4.30}$$

and also the standard deviation σ_k

$$\sigma_k = \sqrt{\frac{1}{N} \sum_{m=1}^{N} \left(\tilde{r}_k^m - \langle \tilde{\boldsymbol{r}}_k \rangle \right)^2}. \tag{4.31}$$

In the case of the raw-data the standard deviation and mean is not significant since it is given by the distribution P_λ, but this can be used to estimate the statistical uncertainties due to the counting statistics in the unfolding as well. One can also propagate these errors to the first generation method and level-densities used in the Oslo-Method.

When it comes to the bias (systematic errors) of the iterative method, there is no good way to estimate this at present. Applying any unfolding method will introduce bias, in order to reduce the variance.

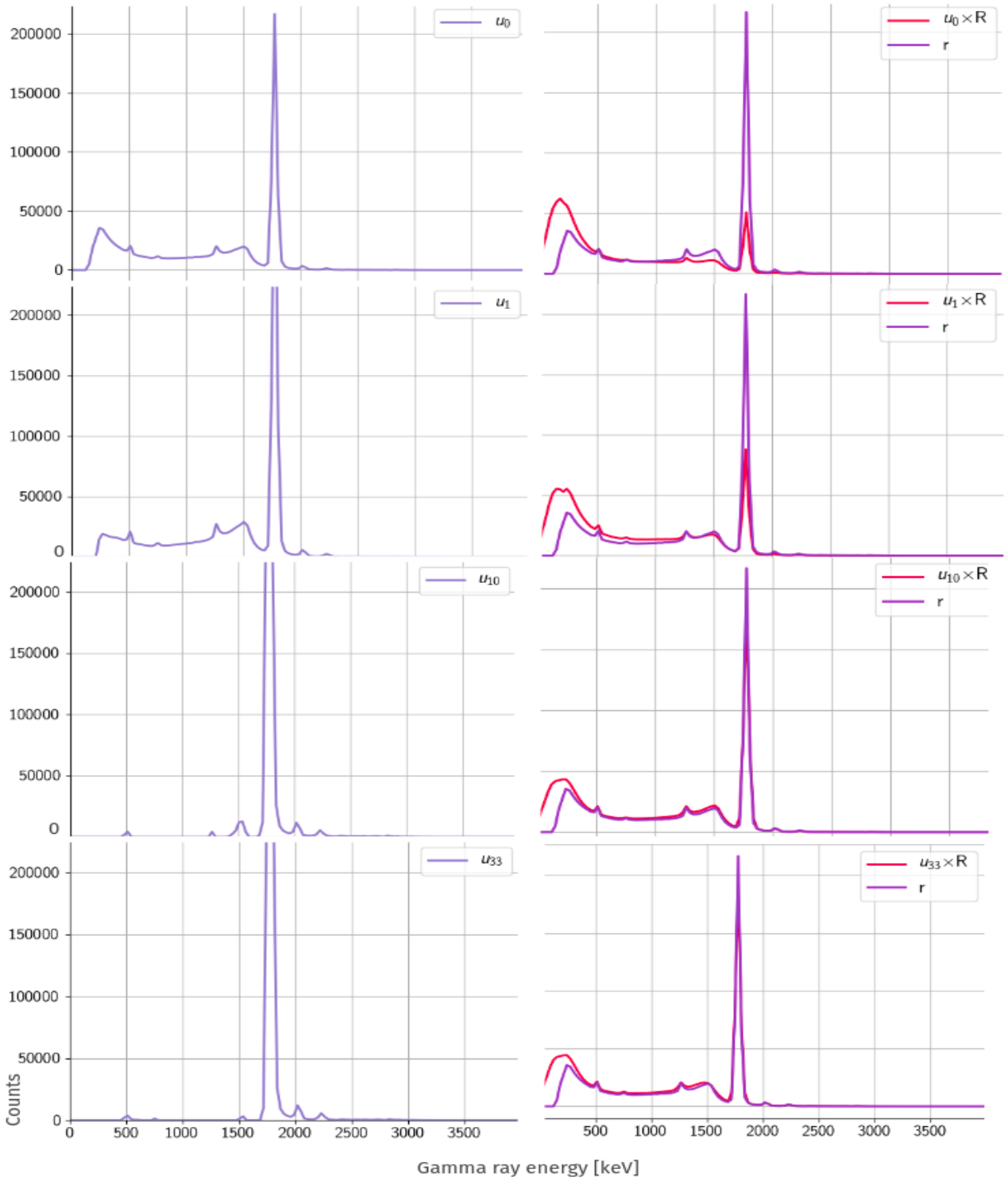

Figure 4.6: The iterative unfolding method demonstrated on ^{28}Si. On the left side is the unfolded spectrum and the on the right is the folded spectrum compared to the raw spectrum. The iterations shown are 1, 10 and 33, where $\boldsymbol{u}_0$ is the initial spectrum. Note that from iteration 1 the unfolded is outside the scale.

3.2 Compton Subtraction Method

After applying the iterative method, the usual next step is to apply the Compton subtraction method. The iterative method will reduce the oscillations in the spectrum, due to the score inducing a weight on the fluctuations. However, there will typically be large fluctuations still present, inherited from the iterative method itself. The Compton subtraction aims to smooth the spectrum further. The method is presented in the following:

- The method takes as input:
  - $\boldsymbol{r}(i)$ = the raw spectrum obtained from experiment
  - $\boldsymbol{u}_0(i)$ = the unfolded spectrum obtained from the iterative method

  Using the unfolded spectrum $\boldsymbol{u}_0(i)$ obtained from the iterative method, and and the raw spectrum $\boldsymbol{r}(i)$ to create a spectrum consisting only of the Compton spectrum $\boldsymbol{c}(i)$.
- To do this without the loss of important features, the probabilities (p_{FE}, p_{SE}, p_{DE}, p_{511}) are extracted from the response matrix table 4.5. These are all multiplied with $\boldsymbol{u}_0(i)$, except for the Compton distribution. The new spectrum obtained by this is:

$$\boldsymbol{v}(i) = \boldsymbol{u}_0(i)p_{FE}(i) + \boldsymbol{w}(i), \tag{4.32}$$

  where $\boldsymbol{w}(i)$ represents

$$\boldsymbol{w}(i) = p_{SE}(i)\boldsymbol{u}_0(i) + p_{DE}(i)\boldsymbol{u}_0(i) + \sum_i p_{511}(i)\boldsymbol{u_0}(i). \tag{4.33}$$

  To avoid getting delta functions where the peaks are located, a Gaussian smoothing function, using the resolution from given experiment (FWHM), is applied.
- To obtain the Compton distribution-spectrum, the $\boldsymbol{v}(i)$-spectrum is subtracted from the raw matrix:

$$\boldsymbol{c}(i) = \boldsymbol{r}(i) - \boldsymbol{v}(i), \tag{4.34}$$

  leaving the contributions from the Compton scattering in the spectrum $\boldsymbol{c}(i)$. The resulting Compton spectrum should, due to the nature of the Compton scattering process, be a slowly varying function with no fluctuations, so it is smoothed with a Gaussian function.
- The unfolded spectrum is obtained by subtracting the Compton spectrum and the $\boldsymbol{w}(i)$-spectrum from the raw spectrum $\boldsymbol{r}(i)$

$$\boldsymbol{u}(i) = \frac{\boldsymbol{r}(i) - \boldsymbol{c}(i) - \boldsymbol{w}(i)}{p_{FE}(i)} \tag{4.35}$$

  where the division by $p_{FE}(i)$ is the correction for the full energy probability. The resulting spectrum $\boldsymbol{u}(i)$ should inherit the statistical fluctuations that where present in the raw data, but should contain less fluctuations from the iterative unfolding method itself.

- In the end, the final unfolded spectrum is divided by the total efficiency ϵ_{tot}:

$$\boldsymbol{u}(i) = \frac{\boldsymbol{u}(i)}{\epsilon(i)_{tot}} \tag{4.36}$$

The resulting unfolded matrix has fewer fluctuations than $\boldsymbol{u}_0(i)$ from the iterative method alone. If $\boldsymbol{u}_0(i)$ was only smoothed and presented as the final unfolded, the statistical fluctuations of the original raw spectrum $\boldsymbol{r}(i)$ would be lost. Instead this method smooths the artificial fluctuations of the $\boldsymbol{u}_0(i)$ but at the same time it will preserve the statistical fluctuations from $\boldsymbol{r}(i)$.

4 Fully Bayesian Unfolding

The following section will explain the Fully Bayesian Unfolding (FBU) method used in this thesis. In the first part the formulation of the theorem will be described. Then the priors that are tested in this thesis are shown. In this method no iteration is used, instead the Monte Carlo NUTs algorithm [11] samples from the likelihood and prior which results in posterior distributions for each energy-bin of the spectrum. This method was originally developed by Georgios Choudalakis [38] for unfolding in High Energy Physics (HEP). Though, the problem of unfolding for HEP differs from unfolding in nuclear physics in many ways, both use detectors with a given response and efficiency. So the methods used for unfolding in high energy physics can be tested for unfolding γ-ray spectra as well.
The library used to develop the method is PyMC3, which is a Python package for Bayesian statistical modeling and Probabilistic Machine Learning, focusing on advanced Markov chain Monte Carlo (MCMC) and variational inference (VI) algorithms [12].

4.1 Bayes' Theorem used in Unfolding

One can ask the question; what was the actual truth spectrum $\boldsymbol{T}$ given the data $\boldsymbol{r}$? This sounds like a Bayesian question, and the aim of the method presented here is to answer it. By using Bayes theorem, matrix inversion is avoided. Instead one samples from the likelihood function and the prior. These two express the region that will be sampled from. Using Bayes' theorem the equation becomes

$$p(\boldsymbol{T} \mid \boldsymbol{r}) \propto L(\boldsymbol{r} \mid \boldsymbol{T}) p(\boldsymbol{T}), \tag{4.37}$$

where $\boldsymbol{T}$ represents the truth spectrum [38]. Since the actual truth spectrum $\boldsymbol{T}$ cannot be obtained, the estimated truth spectrum $\hat{\boldsymbol{T}}$ is the goal, and this can then be represented as the unfolded spectrum $\boldsymbol{u}$. The raw spectrum is binned $j \in \{1, 2, ...N_j\}$ where each bin represents a certain energy (keV) and contains events measured in given experiment $\boldsymbol{r} = \left(r_1, r_2, ..., r_{N_j}\right)$. Likewise, the estimated truth spectrum is also binned and energy dependent $t \in \{1, 2, ...N_t\}$ where each bin contains the number of events expected to be produced in given bin $\hat{\boldsymbol{T}} = \left\{\hat{T}_1, \hat{T}_2, ...\hat{T}_{N_t}\right\}$ [38]. Each expected event will be estimated, so each bin t needs a prior probability

distribution, hopefully containing the expected parameter $\hat{T}_{N_t}$. This results in a posterior distribution for each truth-bin t. Since the results are a set of distributions, the Bayesian unfolding does not necessarily end up with one specific solution, it rather gives a distribution that shows where the expected truth value will most likely be. In some cases, there can be two or more solutions that are equally likely, and this will show in the posterior for the given bin.

The likelihood function of the data is assumed to be independent and Poisson distributed

$$L(\boldsymbol{r} \mid \boldsymbol{T}) = \prod_{j=0}^{N_j} \frac{f_j^{r_j}}{r_j!} e^{-f_j}, \tag{4.38}$$

where $\boldsymbol{f}$ is the folded spectrum, i.e. the expected values of the data is

$$f_j = \sum_{t=1}^{N_t} T_t \cdot R, \tag{4.39}$$

where R is the response matrix [38]. In FBU the response matrix is treated differently from the response matrix explained in section 2. Instead of the rows summing up to 1, it is scaled with the efficiency of the detector [38]. The response matrix is the conditional probability of the events in bin j of the raw spectrum, given that the events are present in the truth bin

$$R(j \mid t) = \frac{P(t,j)}{P(t)} \tag{4.40}$$

and each row should sum up to the efficiency

$$\epsilon_t = \frac{\sum_{j=0}^{N_j} P(t,j)}{P(t)} \tag{4.41}$$

[38]. Since the folded is spatially dependent it can be represented in vector form

$$\boldsymbol{f} = R\boldsymbol{T}, \tag{4.42}$$

this looks similar to the section about the iterative method 3, but there the unfolded $\boldsymbol{u}$ spectrum is inferred directly. In the Bayesian method the distribution for each truth bin is inferred using the likelihood and prior belief about the truth spectrum $\boldsymbol{T}$. After sampling, a posterior distribution is achieved. How the estimated truth spectrum $\hat{\boldsymbol{T}}$ (that can then be represented as the unfolded spectrum $\boldsymbol{u}$), is inferred from the posterior distribution will be discussed later on in this section.

In all experiments there will be some background events due to different processes. These do not originate from the real truth spectrum $\boldsymbol{T}$. If the background has not been subtracted beforehand, one can include it in the Bayesian method as

$$\boldsymbol{f_j} = \boldsymbol{B}_j + \sum_{t=1}^{N_t} T_t \cdot R, \tag{4.43}$$

where $\boldsymbol{B_j}$ is the background events in the bin j [38]. The Bayesian method thus introduces a new treatment of the background.

4.2 The Posterior Distributions

The posterior is proportional to the likelihood and the prior, $p(\boldsymbol{T} \mid \boldsymbol{D}) \propto L(\boldsymbol{D} \mid \boldsymbol{T})p(\boldsymbol{T})$, and it results in a distribution for each bin. In figure 4.7, examples of posterior distributions for a given truth-bin t is shown. Inferring the estimated truth spectrum $\hat{\boldsymbol{T}}$ from these posteriors is not always straightforward. When it comes to unfolding (in the Oslo-method especially) one typically is after a point estimate of the values represented as the unfolded spectrum $\boldsymbol{u}$, to use in further calculations. This can be retrieved in many different manners with Bayesian unfolding, and it will be up to the user to decide what makes sense in terms of the raw data. In the following list are some of the possible estimations presented:

(i) Estimate the HPD interval: Find the most likely value of the posterior distribution and from that point integrate an area of 34% on each side and thus obtain an 68% HPD credible interval. Then choose the unfolded $\boldsymbol{u}$ from the middle of the HPD-interval [9].

(ii) Assuming there exists a mean and choosing the credible interval of 68% where the mean is in the center, for each posterior distribution.

(iii) Choosing a ETI (equal-tailed interval) which will have the median at its center [9].

(iv) Finding the MAP (maximum a posteriori) [9].

Choosing the highest posterior density will answer the question: where the most credible values of $\hat{\boldsymbol{T}}$ are, and what range the highest density values have. ETI gives a overview of the posterior distribution in terms of percentiles with its median at its center (50 percentile). The range of uncertainty will be in percentiles [9]. So choosing the median will not necessarily tell where the highest probability density of the parameter is located. The same argument goes for choosing a credible interval with the estimated mean at the center, it will not necessarily give the values of highest probability. The only exception is if the posteriors are normally distributed, then the mean, median and the mean of the HPD-interval will give the same estimated values. The MAP is calculated differently than the other points (see (i), (ii), (iii)). In the PyMC3 library a optimization code is used to calculate the MAP [12]. In this work the L-BFGS-B (Large-scale Bound-constrained optimization) has been used, which originates from the scikit-learn library [39]. MAP is similar to the Maximum Likelihood Estimation and will only provide a point estimate of the parameters of $\hat{\boldsymbol{T}}$ [9].

Since, in the case of unfolding, the estimated parameters have meaningful interpretations, choosing the value with highest probability is reasonable since the value with highest probability might be wanted. That said, the EIT-interval is a good choice as well, where the spectrum is then represented by the median. In chapter 5, where the results are presented this topic will be discussed and presented further.

The number of bins and the sampling will effect the resulting posterior distribution. A larger sample set, can give narrower posteriors, as the method becomes more certain. If the total amount of bins in the spectrum is large the method will require more samples.

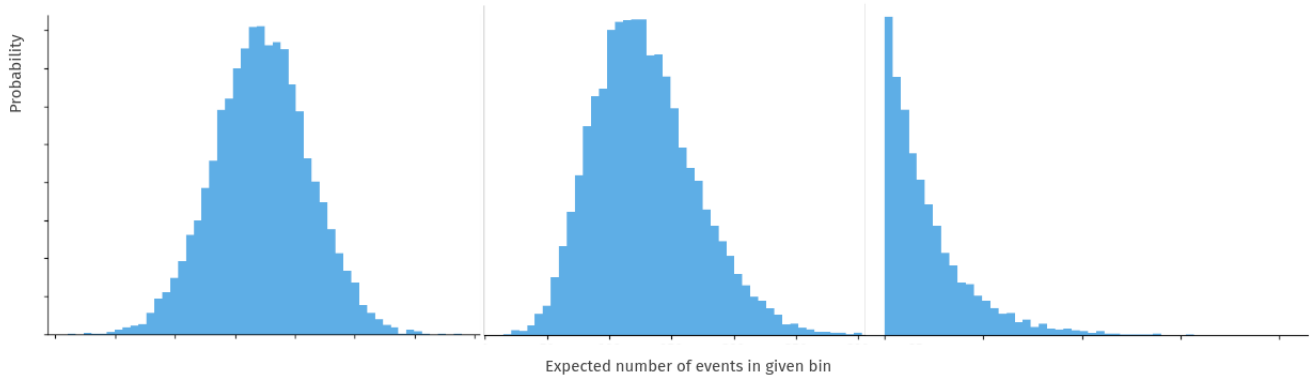

Figure 4.7: Examples of posteriors retrieved after sampling

4.3 Updating Prior Beliefs

The prior, is a very important input in this method, as the truth spectrum $\boldsymbol{T}$ is often unknown beforehand. The raw spectrum $\boldsymbol{r}$ can differ significantly from the truth spectrum, because of the detector response. In the Bayesian frame work, the prior could be chosen from the evidence at hand like theory or past experiments or one could choose what is called an objective prior. That said, there does not exist a fully objective prior, there will always be some assumptions made no matter how objective the user sets out to be. In the iterative method one uses the raw spectrum $\boldsymbol{r}$ as an initial guess for the iteration. So one can argue that the raw spectrum is a prior in that case. In the same manner, the fully Bayesian method require initial assumptions. How specific these are, is up to the user. To be as objective as possible a uniform distribution can be used

$$p\left(\boldsymbol{T}\right) = \begin{cases} \frac{1}{a-b} & T_t > 0 \\ 0 & \text{otherwise} \end{cases}, \tag{4.44}$$

where a represents the upper bound and b represents the lower. To reduce the volume sampled from, it can be a good idea to apply an efficiency correction the prior

$$p\left(\boldsymbol{T}\right) = \begin{cases} \frac{\boldsymbol{r}}{\epsilon} + \mathrm{c} & \text{if } T_t \in \left[T_t^b, T_t^a\right] \forall t \in [1, N_t] \\ 0 & \text{otherwise} \end{cases}, \tag{4.45}$$

where c is a constant and $\left[T_t^b, T_t^{\mathrm{a}}\right]$ represent the upper and lower bounds a and b. The efficiency correction to the prior is the most basic information about the spectrum: if the efficiency to detect γ-rays at 6 MeV is 0.5 it means that roughly twice as many events in that bin should be expected. That said, another important aspect is that the prior distribution for each bin is multiplied with the response matrix, in fact creating a prior for the folded spectrum: $\boldsymbol{f} = R\boldsymbol{T}$ [38]. The likelihood sampled is, as stated: $L(\boldsymbol{r} \mid \boldsymbol{T}) = \prod_{j=0}^{N_j} \frac{f_j^{r_j}}{r_j!} e^{-f_j}$. The prior constructed should resemble the raw spectrum $\boldsymbol{r}$ when multiplied with the response matrix. Since only the solutions above zero are interesting in a physical sense, the lower bound should be zero or above. This will exclude the negative solutions, since these are 'unphysical'. However, the negative solutions may be just as likely as the positive ones, thus resulting in posterior distributions like the example distribution on the far right in figure 4.7. The user of the Bayesian unfolding method is free to use other priors than the ones mentioned in this section. For example, can a normal prior be used, but one has to be careful since the normal distribution allows negative values. A good way to avoid to much fluctuations is to use regularization, one example is the max entropy (MaxEnt). This is then implemented in the uniform prior

$$p\left(\boldsymbol{T}\right) = \begin{cases} \exp(\alpha S(\boldsymbol{T})) & \text{if } T_t \in \left[T_t^{\mathrm{b}}, T_t^{\mathrm{a}}\right] \forall t \in [1, N_t] \\ 0 & \text{otherwise} \end{cases}, \tag{4.46}$$

where

$$S(\mathbf{T}) \equiv -\left(-\sum_{t=1}^{N_t} \frac{T_t}{\sum T_t} \log \frac{T_t}{\sum T_t}\right) \tag{4.47}$$

[38]. In this thesis only the prior in equation 4.45 is used.

Part III

Part 3: Results and Discussion

Chapter 5

Results

In this chapter the results for the unfolding using the Fully Bayesian Unfolding (FBU) method, will be presented. The experimental data used in this thesis was taken from the reaction ^{28}Si$(p, p'\gamma)$ has been used to verify the method. The target used has thickness $8 - 9$ mg/cm^2 and consists of 92.223% ^{28}Si, 4.685% ^{29}Si and 3.092% ^{30}Si. The detector telescope SiRi was set in backward angels and the OSCAR array was placed 16 cm from the target. The protons travel through the target and then into the SiRi detector and since the target is thick, this will affect the resolution of the experiment. The pile-up rate in OSCAR was high, which impacts the amount of low-energy γ-rays and the counts inside the 511-annihilation peak. The electronics of this experiment was digital, while the response matrix used is made for analog electronics. The experimental conditions of therefore differ from the conditions of which the response functions were made. In addition, the total efficiency of OSCAR is not fully established, which might effect the precision of the results.

First section is dedicated to the unfolding of γ-rays emitted from the 1st excited state of ^{28}Si. This is seen as a rather easy case to unfold as it has one γ-transition from the 2^+-state to the ground state. As mentioned in chapter 4, the FBU offers the inclusion of background and this function is tested in section 2. In section 3 the unfolding of the photon spectrum from all excited states is presented.

1 Fully Bayesian Unfolding Spectrum From the First Excited state of ^{28}Si

As a first test for the FBU, the first excited state of ^{28}Si is extracted from the raw coincidence matrix and projected onto the E_γ-axis. Usually, when the raw data is intended for the Oslo-method everything below the diagonal ($> E_\gamma = E_x$) is set to zero, since the counts over the diagonal after the background subtraction should fluctuate around zero (i.e sum to zero). The matrix in 5.1 shows the raw coincidence matrix where the diagonal is set to zero. There was still negative counts in the matrix after cutting the diagonal, so the function from the OMpy library [37] called fill_and_remove_negatives was used. The function fills some of the negative counts with

counts from neighbouring bins. To prevent losing to many positive counts the rest of the negative values are set to zero. In addition, the matrix itself is cut so it ranges from $0 - 10000$ keV for both the E_x and the E_γ axis, since the original matrix contains few counts beyond these ranges. The binning was originally 5 keV energy per bin, but to speed up the sampling process used in the FBU, the matrix was re-binned to 20 keV energy-bins. For the first excited state in ^{28}Si this was considered to be fine enough binning to describe the spectrum. In figure 5.1 the red line shows the area which was projected onto the E_γ-axis, the boundaries where $E_x = 1400$ and $E_x = 2200$. This specific range was found by projecting the data onto the E_x-axis that can be seen in figure 5.18, this reveals where it is best to cut-off to get the relevant counts that belong to the first excited state. Also, the resolution in the particle detector SiRi for the 2^+ peak was found to be 360 keV, which verifies the boundaries chosen (1779 ± 360 keV).

The first excited state of ^{28}Si has energy 1779 keV, and when the nucleus de-excites to the ground state it will send out a γ-ray with an energy of 1779 keV, which is a 2^+ transition [40]. When setting the counts below the diagonal to zero, the aim is to cut everything after the 1779 keV energy. However, the target is not 100% pure ^{28}Si, but also consists of ^{29}Si and ^{30}Si [40]. The transitions of their first excited state ($(2/5)^+$ and 2^+) to their respective ground states is visible in the raw data. Since these will have counts that should be re-distributed to them when unfolding, it is important not to cut them out of the spectrum. When setting everything below the diagonal to zero one needs to be careful, the risk is the loss of valuable information. The diagonal was thus chosen to include the full energy peaks for the high energies ($E_x = E_\gamma > 7000$ keV), which resulted in the range of the first excited state being 3220 keV on the E_γ axis, as can be seen in figure 5.1. Since the counts below the diagonal are zero, these are not necessary to include in FBU. Figure 5.2 shows the projected raw data for the first excited state, a) is the raw data after setting the counts above 3220 keV to zero, and with 5 keV binning,

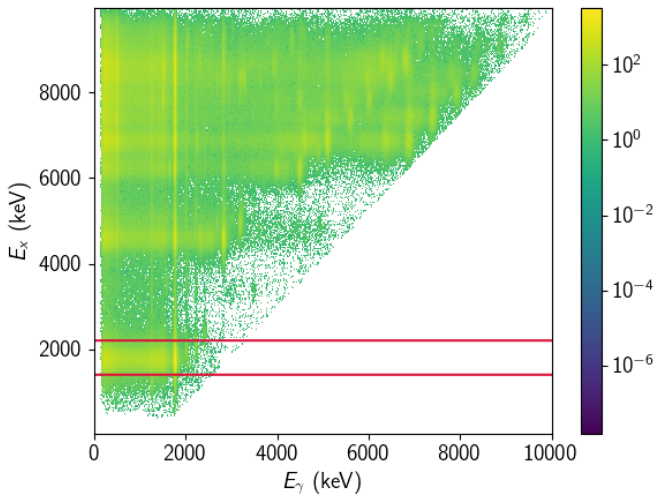

Figure 5.1: The $E_\gamma - E_x$ matrix for ^{28}Si$(p, p'\gamma)$-reaction. The red lines indicates where the cut was made for the projected $1st$ excitation energy.

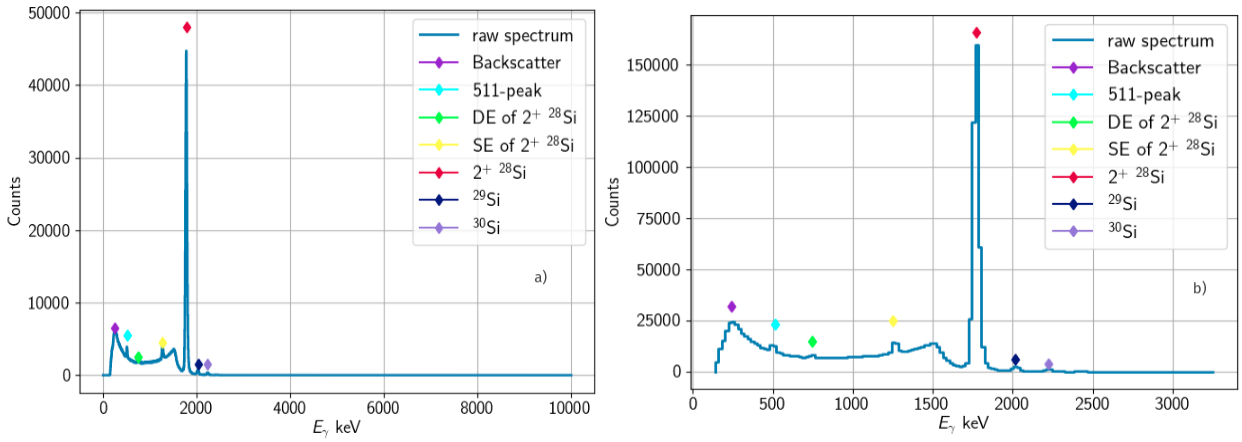

Figure 5.2: The 1st excitation energy projected on the γ-axis where all features of the spectrum are identified. a) is the projected raw data where the diagonal has been cut at $E_\gamma = 3220keV$. b) The projected raw data cut and re-binned.

b) shows the raw spectrum after cutting out the zeros and re-binning the data with 20 keV binning. All the features in the raw spectrum are identified for both plots [40].

The treatment of the threshold is important to notify. The threshold area ($[0, 160]$ keV) is the lower energy detection threshold in OSCAR, meaning that no signal is detected in this area. The response for the digital electronics has not been fully established yet. In figure 5.2 it can be seen that the energies from 0 keV and up til 160 keV are zero in the raw matrix. Since there is no treatment of this area implemented in OMpy or in the current response matrix, it was cut out before applying FBU.

As mentioned in chapter 4, FBU requires that the response matrix is scaled with the efficiency. The response matrix is made by applying the class called Response in the OMpy library, and each row sums up to 1. The efficiency is normalized to 1 for the γ-energy 1332 keV, while the actual efficiency for this energy will differ from experiment to experiment. It is assumed that the total efficiency is close to the geometric efficiency, which is calculated as

$$\epsilon_{geo} = \frac{\text{Covered Area of OSCAR}}{\text{Solid Angle}}, \tag{5.1}$$

where the covered area is the number of detectors multiplied with the detector 'nose'$= \pi r^2$. The diameter of one scintillator detector is 8.89 cm so the radius is $r = 4.45$ cm. Multiplied this with all 30 detectors, the covered area of OSCAR is equals 0.1862 m^2. The Solid Angle, when the OSCAR array is $R = 16$ cm from the target is: $4\pi R^2 = 0.3217\ m^2$. Then the geometric efficiency $\epsilon_{geo} = 0.579$ so approximately 58%. The actual total efficiency is expected be somewhat smaller than the geometric, since this was the case for the old detector CACTUS (replaced by OSCAR). The total efficiency for CACTUS was the geometric efficiency reduced by 14% for the energy 1332 keV. Assuming the same reduction for OSCAR the total efficiency for energy 1332 keV is 50%, when the distance between target and detector is 16 cm. In this present work it was decided to use this value, as a measured value has not been established yet. The efficiency was scaled from 1 to 0.5 for energy 1332 keV, as

can be seen in figure 5.3, where the scale for the full energy peak is also included. The response matrix was then scaled so the rows sum up to the efficiency instead of summing up to 1, which is the case when using the Iterative Method in OMpy.

The prior chosen for the raw data was a uniform prior, corrected with the efficiency and the scale of the full energy peak, and an added constant. The area sampled from is bound between a upper and lower bound. The sampling will take longer the bigger area that needs to be sampled. Each bin will add a dimension to the sampling, with the current 156 bins gives 156 dimensions. To make a good and efficient prior it is important to look at the prior for the folded, which is what is being sampled $\boldsymbol{f} = R\boldsymbol{T}$. In a perfect case, for the 2^+-state with γ-energy = 1779 keV, all the counts should be re-distributed to this peak. A good first try for an upper bound is to correct for the efficiency and the scale of the full energy peak. After correcting, the upper bound for the 2^+ peak did not include all the counts in the raw spectrum. A constant was therefore added to the energy-bins containing the peak, so the upper bound contained more than the sum of counts in the raw spectrum. The lower bound was set to zero, since negative solutions are 'unphysical' in this case. Figure 5.4 shows the lower and upper bound for the prior, a) shows the upper and lower bound for the actual truth spectrum, and b) shows the upper and lower bound for the folded (i.e upper and lower bounds multiplied with the response function). The priors now include the only knowledge one has about the spectrum: the prior for the folded spectrum should include all the counts of the raw spectrum, and the efficiency and full energy peak should be corrected for, and

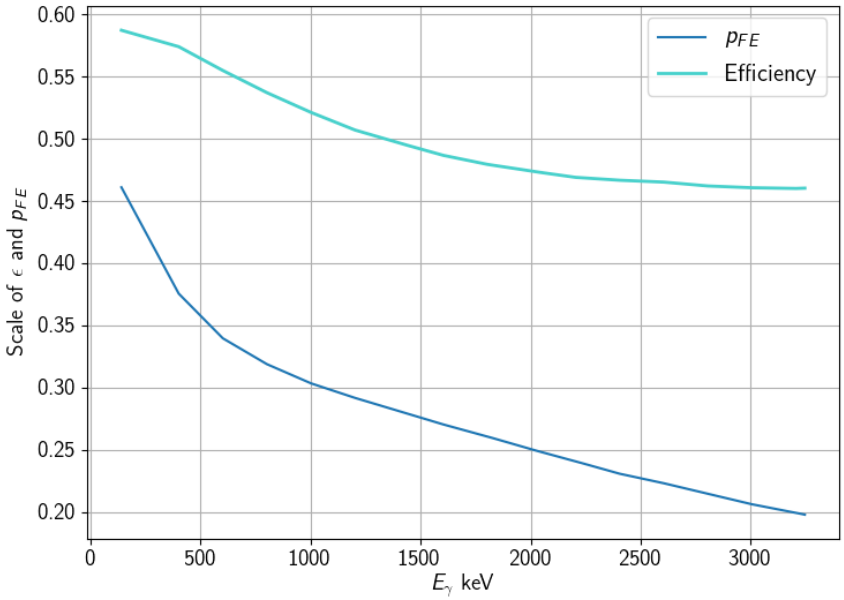

Figure 5.3: The efficiency ϵ and the scale of the full energy peak for the OSCAR array, where the actual efficiency for the 1332 keV is assumed to be 0.5.

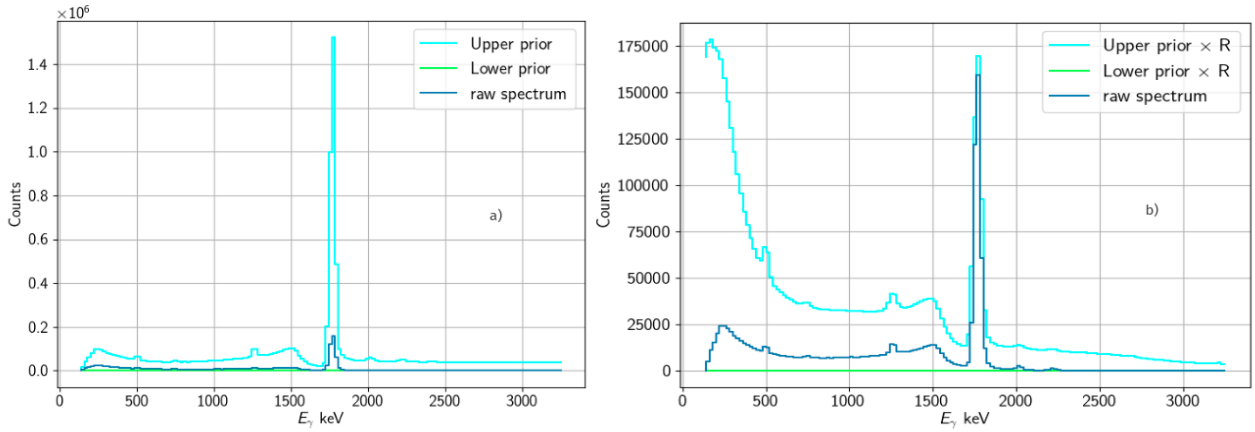

Figure 5.4: Efficiency corrected prior with an added constant. a) Upper and lower bounds for the uniform prior for the unfolded spectrum. b) Upper and lower bounds for the uniform prior multiplied with the response matrix.

lower bounds should be set to zero. In figure 5.4 b), one should note the effect from the response matrix on the upper bound of the area $0 - 200$ keV.

The uniform prior for each bin together with the response matrix and raw data are the inputs of FBU. It was sampled with 10000 samples and 2000 tuning steps for each bin, with the Monte Carlo NUTs algorithm [4]. The posterior distributions for each bin are the output of the FBU. In addition, the MAP was calculated using $L - BFGS - B$ optimization which is also included in FBU. The posterior distributions can be, among other, evaluated in histograms, which shows where the highest probability for the values are located, see figure 5.5. When using an unfolding method the goal is to find a point estimate that in turn can be presented as an unfolded spectrum. Figure 5.5 show some different approaches to finding a point estimate. The shaded area is the 68% HPD-interval which is retrieved by an inbuilt function in the PyMC3 library. The interval will always include the highest posterior density probability. The MAP which is the mode of the posterior distribution, is also included in the figure. Using the MAP will only give a point estimate, and the parameters can be biased if the mode is not representative of the posterior distribution. In figure 5.5 d) the MAP deviates the most from the sampled posteriors, which shows that the MAP is not representative for all bin-distributions. For 5.5 a), b) and c) the MAP estimates the highest probable points in the posteriors distributions. Furthermore, The MAP optimization is a faster method than using Monte Carlo sampling, it can thus be used as a good indicator for the unfolded spectrum if needed, but will not provide uncertainty estimates. The estimated mean and median are also represented in figure 5.5. A credible interval for these estimates can also be used, but as seen in the figure the HPD-interval includes the values with highest probability, while a 68% credible interval for the median and mean will not necessarily include these. In figure 5.5 a), b) and c) shows the estimated mean being positioned on the outskirts of the HPD interval, a 68% credible interval around this point would then not include the highest probability. When the posteriors are skewed towards zero the mean of the HPD will give the value closest to highest density probability. When the posteriors are closely to normally distributed the mean of the

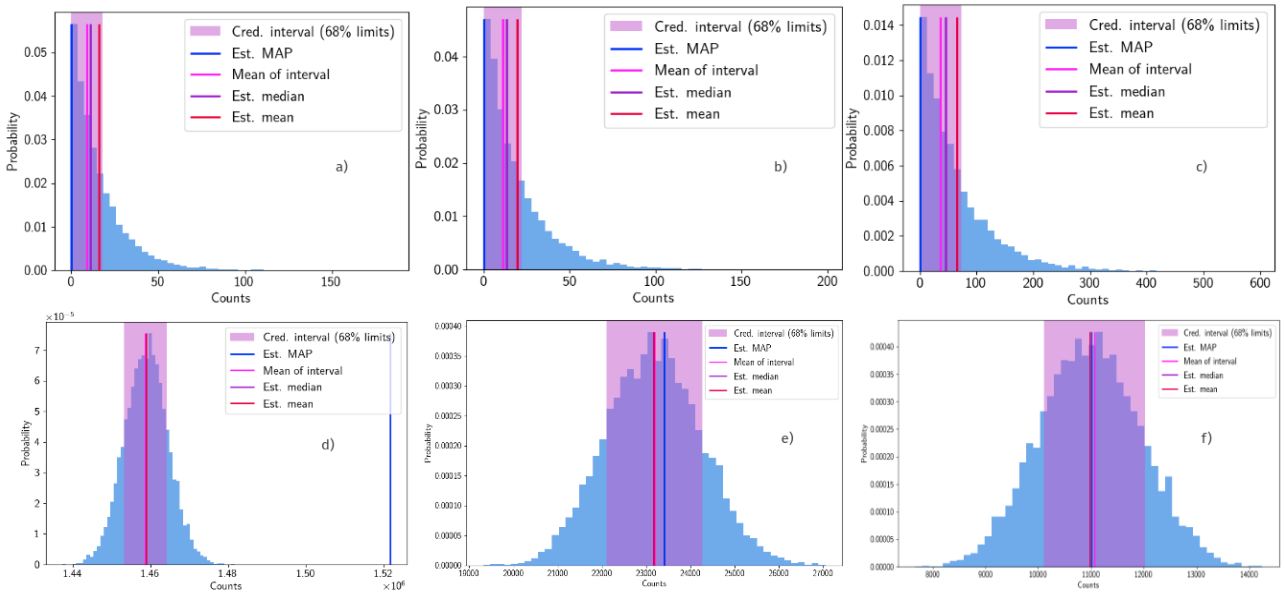

Figure 5.5: Posterior distributions showing the estimated HPD-interval with the mean of the interval as well as the estimated mean, median and MAP. a) Truth-bin for $E_\gamma = 511$ keV, where the 511-peak is located in the raw spectrum. b) Truth-bin for $E_\gamma = 754$ keV, where the double escape peak is located in the raw spectrum. c) Truth-bin for $E_\gamma = 1267$ keV, where the single escape peak is located in the raw spectrum. d) Truth-bin for $E_\gamma = 1778$ keV, where the 2^+-state of ^{28}Si is located. e) Truth-bin for $E_\gamma = 2028$ keV, where the $(\frac{2}{5})^+$-state of ^{29}Si is located. f) Truth-bin for $E_\gamma = 2235$ keV, where the 2^+-state of ^{30}Si is located.

HPD, the estimated mean and the median will all give the same values. The posteriors shown in figure 5.5 are truth-bins for the features shown in the raw spectrum 5.2. d), e), f) are the truth-bins for the 2^+-transition of ^{28}Si, $(2/5)^+$-transition of ^{29}Si, and 2^+-transition of ^{30}Si respectively. As is shown, the distributions are normally distributed. Where there are high amounts of counts, the posteriors tend to be normally distributed and it is rather straight forward to find a point estimate, since the mean, median and mean of the HPD-interval all give the same estimated value. On the other hand, a), b) and c) in figure 5.5 shows the truth-bins for the 511, double escape and single escape peaks. They are all skewed towards zero. This indicates that there might be a more probable solution that is located below zero, but those solutions are of little interest. When unfolding, the aim is to re-distribute all counts in these peaks to the full energy peaks, and these posteriors show that the highest probability is around zero.

In figure 5.6 the log plot of 500 samples picked randomly for each bin and presented together with the MAP, mean of the HPD-interval, median and the HPD-interval itself are shown. This is another way to present the posteriors for each bin in a meaningful way. Logging the MAP and HPD-interval will transform the values, as can be seen in the figure. If choosing a EIT-interval which includes the median, logging the interval does not transform the values. The median and mean of the HPD do not deviate much from each other, showing that

both can be used as estimates for the unfolded. Since a HPD-interval was chosen in this case, the mean of the HPD is picked as the estimate to represent the unfolded and folded of the FBU.

Figure 5.7 shows the log plot of the mean of the HPD plotted together with the unfolding of the Iterative Unfolding method (OMpy). The energies containing zero counts are here added to the FBU unfolded. There are fluctuations around 100 for the OMpy unfolded for lower energies. The HPD-interval is also largest in this area for the FBU, and as seen in figure 5.6 the samples are also fluctuating in this area, indicating that both methods are more uncertain here. For higher energies both methods fluctuates less and follow each other within the HPD interval.

By looking at the folded in figure 5.8 and 5.9 one can see how well the methods have unfolded the raw spectrum, as the folded spectrum should reproduce the raw spectrum. Figure 5.8 shows the folded spectrum without zeros added, for both FBU and the iterative method (OMpy) together with their relative difference to the raw spectrum. Figure 5.9 shows the same spectrum with the zero-bins added. When adding the zero-bins, the threshold is added and the effects it has on the folded spectrum can be seen. By looking at the relative difference, the adding of the threshold area does not seem to have an effect on the Iterative Method (OMpy), but it has clear effects on FBU. Before adding the threshold area the relative difference around the 2^+-transition peak of ^{28}Si is smaller than when the threshold is added. As described before, the response matrix for the threshold is

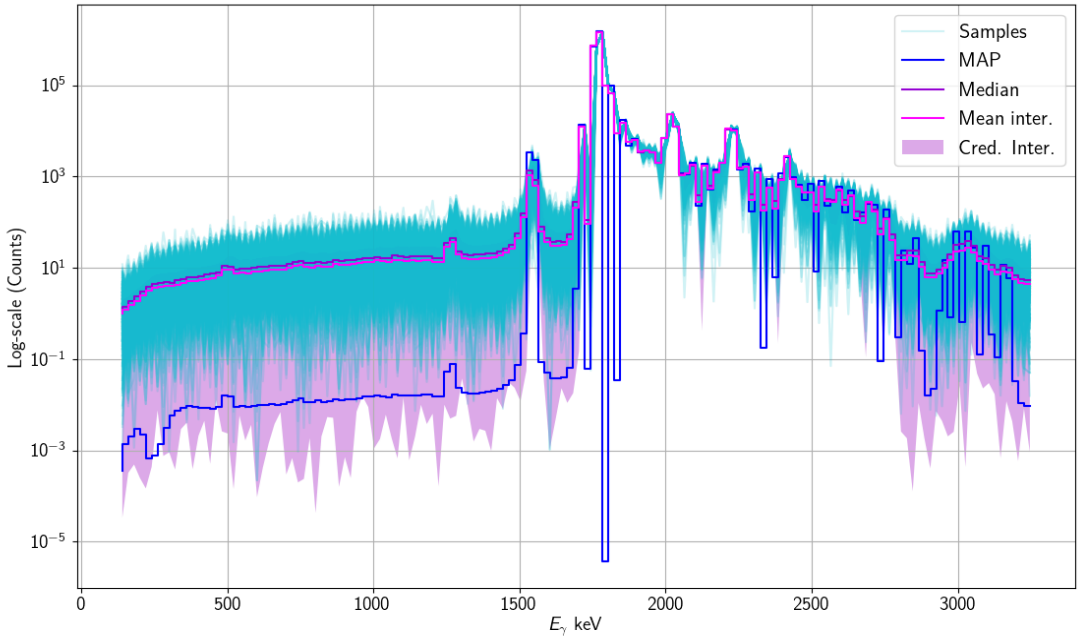

Figure 5.6: Log-plot of the estimated truth-spectrum $\hat{\boldsymbol{T}}$ for the 1st excitation energy for the ^{28}Si$(p, p'\gamma)$-reaction. 500 samples have been picked randomly for each bin and plotted together with the MAP, median, HPD-interval and the mean of the interval.

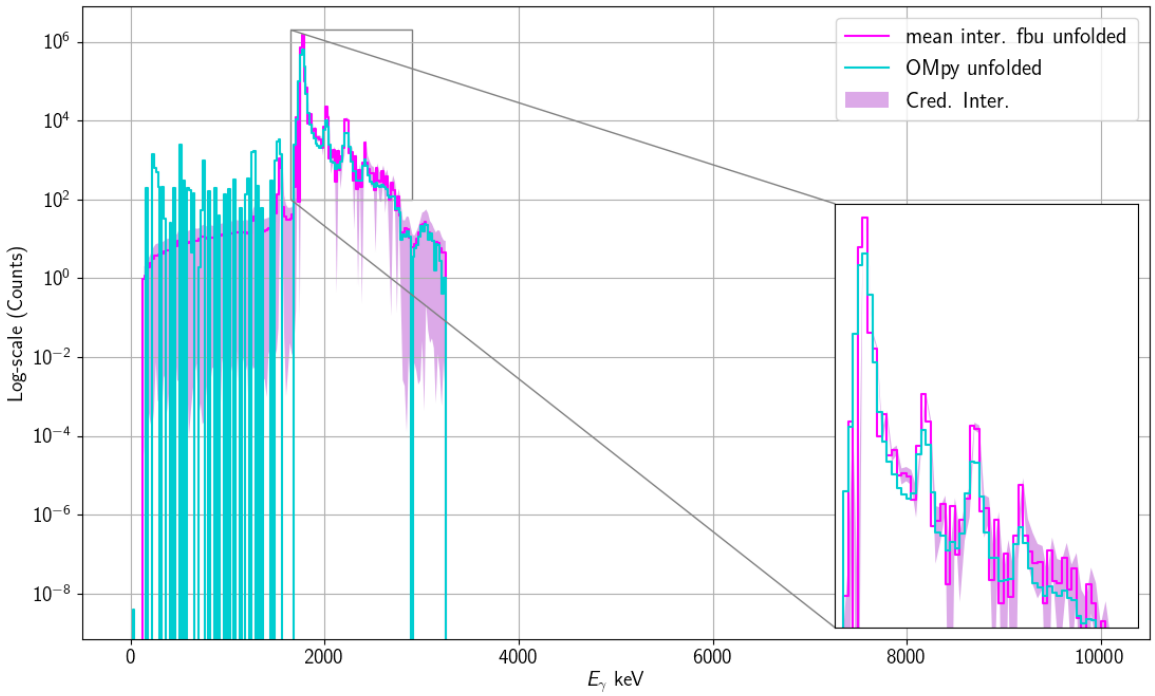

Figure 5.7: The unfolded spectrum represented with the mean of the HPD-interval and the iterative method (OMpy), for the 1st excitation energy for the ^{28}Si$(p, p'\gamma)$- reaction.

not updated for the digital electronics. Since there are no counts in the raw spectrum in this area, the likelihood function in FBU will not have any effect here. The change in relative difference can therefore not be traced to the evidence in the data. Therefore the effect might be traced to the response matrix, which is multiplied with the unfolded of FBU to produce the folded spectrum ($\boldsymbol{f} = \hat{\boldsymbol{T}}R$.

Table 5.1 shows the percentage of counts that are in the main three peaks $(2^+, (5/2)^+, 2^+)$ of the unfolded as well as the percentage of the rest-counts that have not been re-distributed. OMpy has 2% more counts in the rest, while FBU includes more counts into the 2^+-transition. The percentage redistributed to the three peaks do not

Peaks (keV):	FBU (%)	OMpy (%)
1779-peak:	91.7 %	89.8 %
2028-peak:	1.3 %	1.1 %
2235-peak:	1 %	1 %
Rest:	6 %	8 %

Table 5.1: The percentage of counts in each peak and the rest that has not been re-distributed, after unfolding.

coincide fully with the percentage of ^{29}Si and ^{30}Si that are present in the target. However, 91.7% of the counts have been re-distributed to the 2^+-transition peak of ^{28}Si for the FBU unfolding and 89.9% in OMpy.

2 Fully Bayesian Unfolding on first excited state of ^{28}Si including background

The FBU library offers a new method for treating the background. Instead of subtracting it from the raw matrix the user can choose to use the raw matrix with background included as an input into the method. Figure 5.10

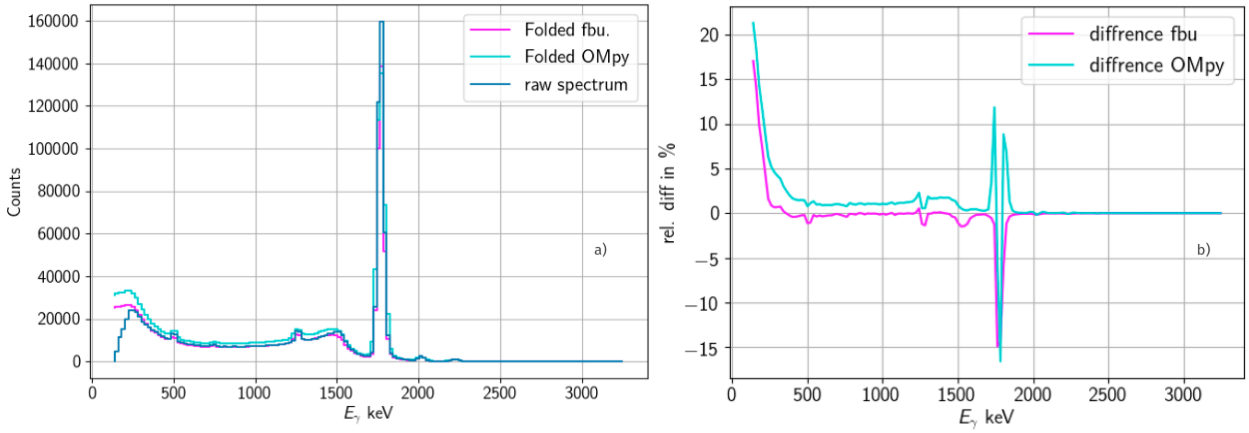

Figure 5.8: a) The folded spectrum represented with the mean of the HPD-interval and the iterative method (OMpy) compared with the raw data before adding the zero-bins, for the first excited state of the ^{28}Si$(p, p'\gamma)$-reaction.

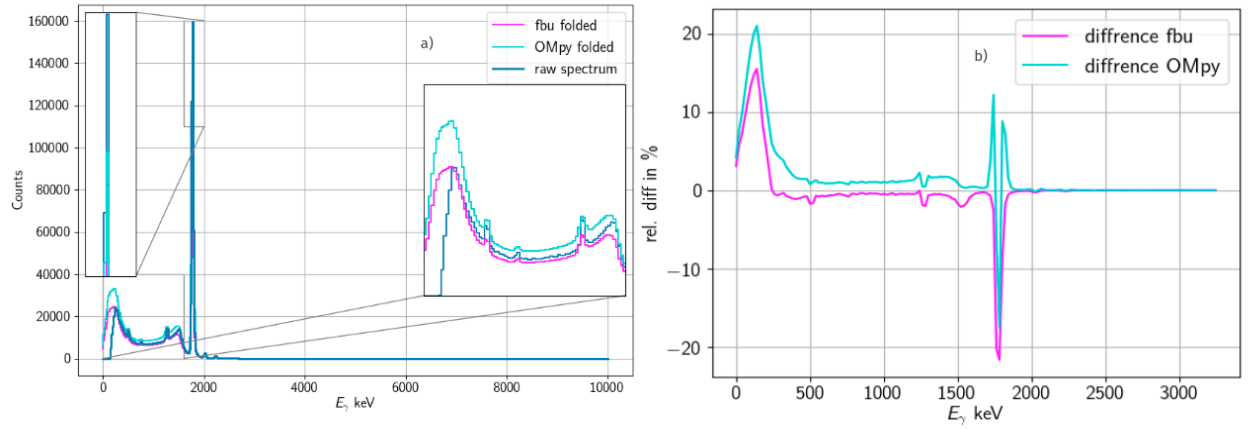

Figure 5.9: The folded spectrum represented with the mean of the HPD-interval and the iterative method (OMpy) compared with the raw data with zero-bins included, for the first excited state of the ^{28}Si$(p, p'\gamma)$-reaction.

shows the raw matrix with background included together with lines indicating the energy projected onto the E_γ-axis. The projection is the same described in last section 1.

In addition to the inclusion of background in the raw spectrum, a spectrum of the background is needed as input. Figure 5.11 shows the raw spectrum including background and the background spectrum, together with the important features of the spectrum. The FBU offers a background normalization estimate when there is uncertainty in the prompt peak versus background ratio. In chapter 3 the background used was shown in figure 3.8, the time gate for both the prompt peak and the background where chosen to be a width of 10 ns. The prompt peak time cut was 0 ± 5 ns and the background 55 ± 5 ns. Since the background and prompt peak where measured in the same experiment and the time ratio between them is 1, the background normalization uncertainty was set to zero.

The response matrix, prior, number of samples and tuning steps where initiated in the same way as previous section. Only difference is the treatment of the threshold, as the background includes counts in this area it is included. Also, the re-binning of the matrix was set to 30 keV per bin. Since no counts are set to zero under the diagonal in this case, the dimensions increase significantly making the sampling process slower.

Figures 5.12 and 5.13 show the posteriors of the truth-bins represented in histograms, with the HPD-interval, the mean of the interval and the median included. Figure 5.12 shows the truth-bins for the 511 keV peak, double escape and single escape peaks, while figure 5.13 shows the truth-bins for the 2^+-transition of ^{28}Si, $(\frac{2}{5})^+$-transition of ^{29}Si and 2^+-transition of ^{30}Si. By comparing the mean of the interval of the 2^+-state of ^{28}Si in figure 5.13 with the same state in the figure 5.5 from previous section one sees that in this case more counts are re-distributed to this peak.

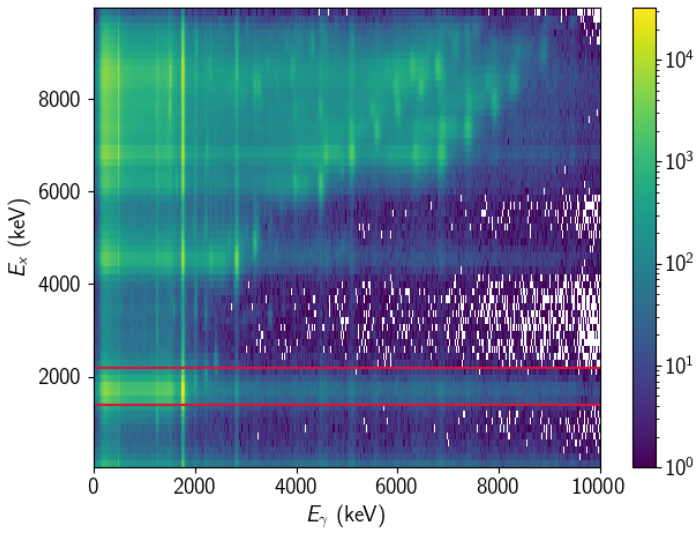

Figure 5.10: The $E_\gamma - E_x$ matrix for ^{28}Si$(p, p'\gamma)$-reaction including background. The red lines indicates where the cut was made for the projected first excited state.

In figure 5.14 the log-plot of the unfolded spectrum is presented. The HPD-interval with its mean and the median together with 500 samples that have been randomly picked for each bin are shown. It should be noted that there are some peaks present for energies higher than the 2^+-state of ^{30}Si, most notably the 2425 keV peak from $(\frac{3}{2})^+$-state of ^{29}Si [40] and also others over the $E_\gamma = E_x$ diagonal. This might indicate that a better

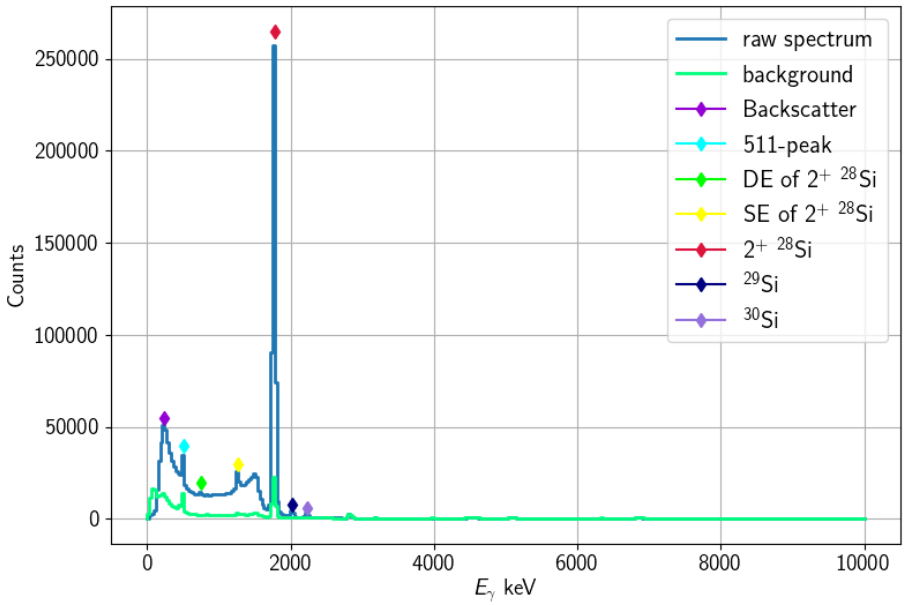

Figure 5.11: The first excited state projected on the γ-axis including background together with the background plotted separately. The features of the spectrum are also described.

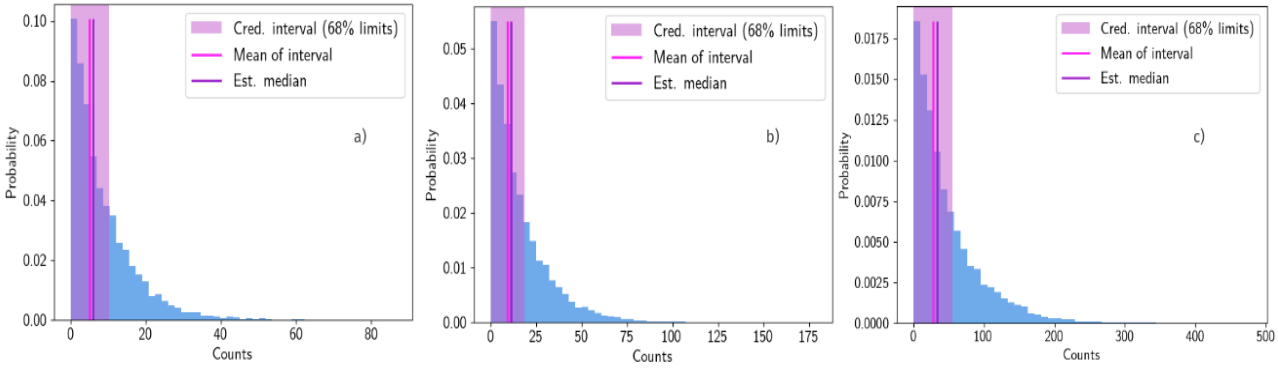

Figure 5.12: Posterior distributions for the truth-bins when background is included, showing the estimated HPD-interval with the mean of the interval as well as the estimated median. a) Truth-bin for E_γ= 511 keV, where the 511-peak is located in the raw spectrum. b) Truth-bin for E_γ= 754 keV, where the double escape peak is located in the raw spectrum. c) Truth-bin for E_γ= 1267 keV, where the single escape peak is located in the raw spectrum

background subtraction is needed, but on the other hand that can result in the loss of important information. There are some rest counts left in the single escape peak of $(\frac{5}{2})^+$-state of ^{29}Si, located at 1517 keV and the double escape peak of 2^+-state of ^{30}Si located at 1213 keV [40], meaning they have not been re-distributed back

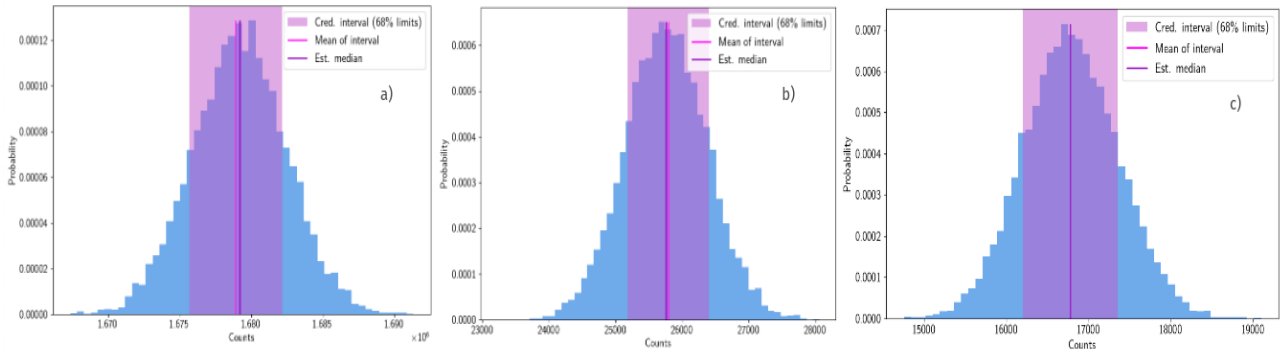

Figure 5.13: Posterior distributions for the truth-bins when background is included, showing the estimated HPD-interval with the mean of the interval as well as the estimated median. d) Truth-bin for $E_\gamma = 1778$ keV, where the 2^+-state of ^{28}Si is located. e) Truth-bin for $E_\gamma = 2028$ keV, where the $(\frac{2}{5})^+$-state of ^{29}Si is located. f) Truth-bin for $E_\gamma = 2235$ keV, where the 2^+-state of ^{30}Si is located.

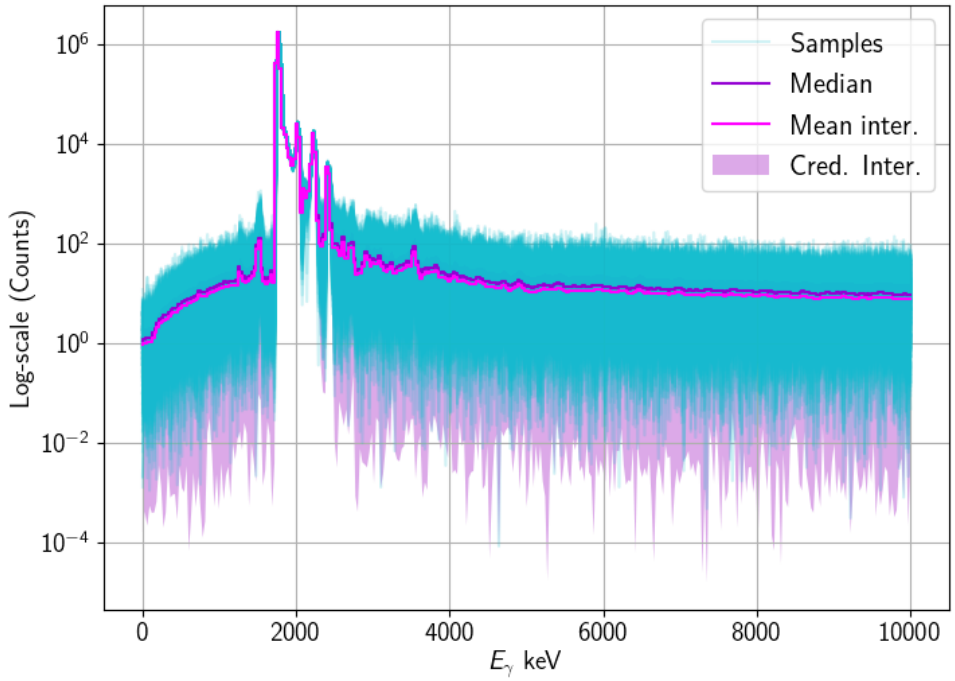

Figure 5.14: Log-plot of the estimated truth-spectrum $\hat{\boldsymbol{T}}$ for the first excited state for the ^{28}Si$(p, p'\gamma)$- reaction including background. 500 samples have been picked randomly for each bin and plotted together with median, HPD-interval and the mean of the interval.

to the peaks. These features are not show significant in the raw spectrum, probably because of the Compton contribution in this area.

Using FBU should result in a unfolded and folded spectrum without background present. Therefore the folded spectrum of FBU should be compared with the background subtracted raw-spectrum, instead of the raw spectrum containing background. This comparison is shown in figure 5.15 between both raw spectrum with and without background present. The raw-spectrum without background is obtained by subtracting the background from the raw-spectrum with background. Since there are counts present in the background spectrum in the threshold area and not in the raw-spectrum this leads to negative counts. This is also the case in the area above the 2^+-peak of ^{30}Si. The function *fill_and_remove_negative* is thus used, explained in previous section 1. The folded of FBU closely resembles the raw spectrum without background and is consistently lower than the raw-spectrum with background included. In figure 5.16 the relative difference between the folded of FBU and the raw-spectrum without background is shown. The relative difference is greatest for the 2^+-peak of ^{28}Si and in the threshold area, which is expected at this point. The factor between the sum of the raw-spectrum without background and the folded of fbu is only 1.015.

Table 5.2 includes the percentage of counts in each prominent peak in the spectrum and the rest counts. As mentioned before, the counts re-distributed to the 2^+-peak of ^{28}Si is higher than in the previous section where the background was not included. The $(\frac{2}{5})^+$-peak of ^{29}Si and 2^+-peak of ^{30}Si are approximately the same as

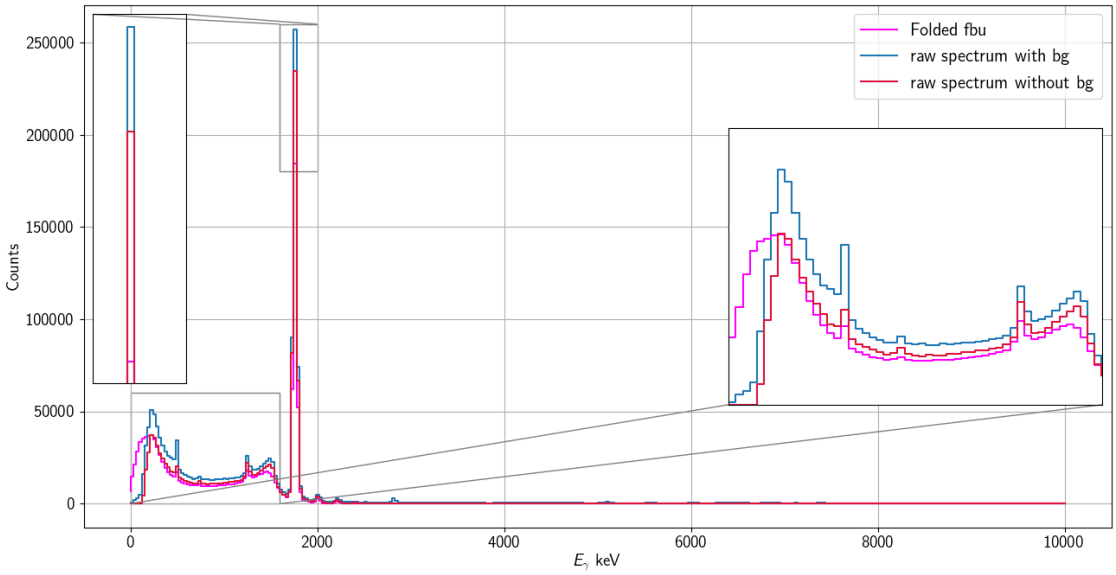

Figure 5.15: The folded spectrum of FBU represented with the mean of the HPD-interval plotted with the raw spectrum with and without background.

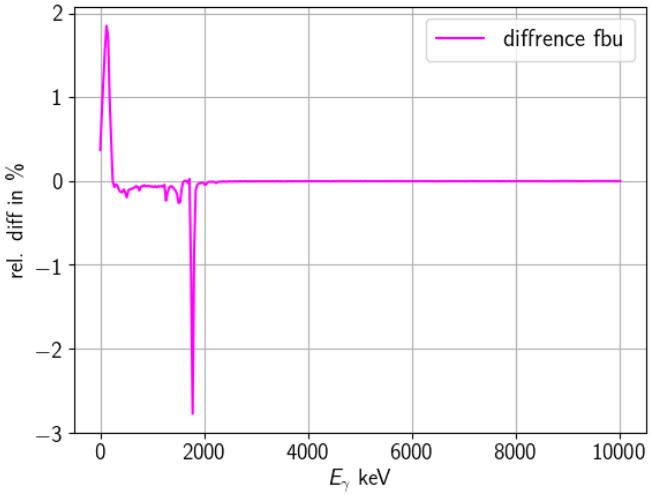

Figure 5.16: The relative difference between the folded spectrum of FBU and the raw spectrum without background.

Peaks (keV):	FBU (%)
1779-peak:	94.5 %
2028-peak:	1.3 %
2235-peak:	0.9 %
Rest:	3.3 %

Table 5.2: The percentage of counts in each peak and the rest that has not been re-distributed, after unfolding.

in the case of no background included. The rest counts are lower than in section 1, since more counts are re-distributed to the 1779-peak.

3 Fully Bayesian Unfolding for all Excited States

The first test to unfold a raw matrix using FBU will now be presented. The test took 7 days and includes the excitation energy range $[0, 8000]$ keV. As can be seen in figure 5.18 showing the raw matrix projected onto the E_x-axis, there is approximately no counts below 647 keV, this can also be verified by looking at figure 5.1 where there are little counts below this energy in the raw matrix. The unfolding with FBU was started at energy 647 keV, therefore the matrix in figure 5.17 has a defined cut there. Because of long computation time it was concluded that the upper bound of 8000 keV for the E_x-axis was sufficient to show as a first test of unfolding a matrix using FBU. To make the computation time as fast as possible it was decided to unfold the raw matrix without background. The number of bins can then be reduced without the loss of structure in the spectrum, since all bins set to zero are cut out before unfolding. For each excitation energy the spectrum was projected

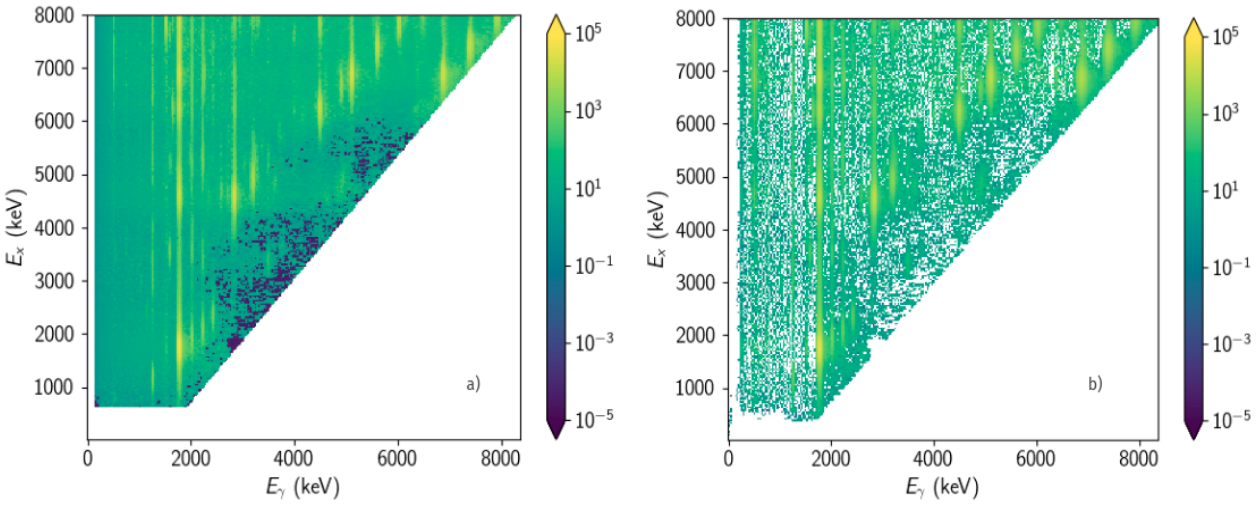

Figure 5.17: The unfolded matrix for a) FBU and b) Iterative unfolding method, in the range $[0, 8000]$ keV for the E_x-axis

onto the E_γ axis and then unfolded. The priors, efficiency and response matrix where all initiated in the same way as in previous sections. The bin size for both the E_x and the E_γ axis where set to 30 keV. The sampling was set to 6000 with 2000 tuning steps, these where reduced compared to previous sections.

Figure 5.17 a) shows the unfolded matrix of FBU and b) shows the unfolded matrix of iterative method (OMpy) for the $[0, 8000]$ keV for the E_x-axis. The main difference is the fluctuations in the unfolding method. For the

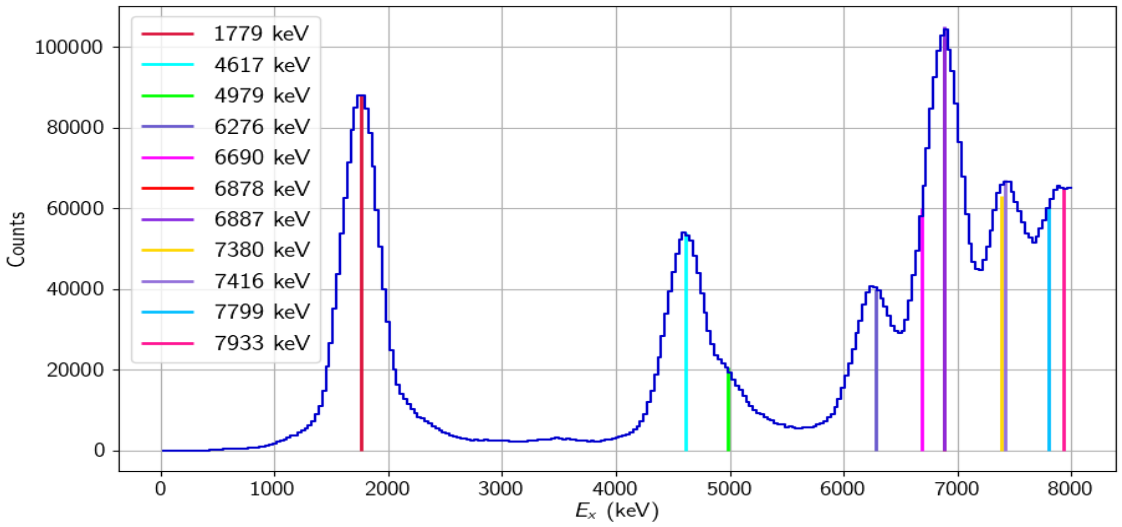

Figure 5.18: The raw matrix values projected on the E_x axis with all excitation energies in the range $[0, 8000]$ keV.

1st excited state in section 1, the iterative unfolding method fluctuated from zero to a few hundred counts in the low energy region (< 1700 keV). This fluctuation can be seen in the unfolded matrix as well. FBU does not fluctuate as much and does not give values equal to zero. Other than this difference the methods both show clear peaks where the full energy peaks are located. Both showing vertical yellow lines for the 2^+ transition and a diagonal line for the full energy peaks of each excited state, except for the states that cascade (e.g. the 4617 keV state).

In figure 5.19 a) the unfolded spectrum of FBU and the iterative method for the second excited state (4617 keV) is projected onto the E_γ axis. The γ-rays cascades, first to the first excited state (1779 keV) then to the ground state. Therefore there is no well defined peak at energy 4617 keV, but rather at energy 2838 keV ($4617 - 1779$ keV), and energy 1779 keV. The vertical line of 2838 keV peak can also be seen in the unfolded matrices in figure 5.17. The unfolded of FBU is presented with the HPD-interval showing the uncertainty in the estimation. FBU estimates higher values for the γ-ray peaks at 1779 keV and 2838 keV, and the iterative method fluctuates more for lower energies (< 2500 keV), otherwise the methods coincide more or less. In figure 5.19 b) shows the unfolded spectrum of FBU and the iterative method for the tenth excited state (7416 keV) projected onto the E_γ axis. Looking at figure 5.18 where the counts have been projected onto the E_x-axis, it can be seen that tenth excited state is close to the ninth excited state (7380 keV). Because of the resolution of the particle detector SiRi these peaks are hard to separate. The projection seen in 5.19 b) will therefore include the counts from both the ninth (7380 keV) and the tenth excited state (7416 keV). In this plot the FBU and the iterative unfolding (OMpy) coincide better compared to the other figures shown in previous section 1, since the iterative method does not fluctuate much here. The tenth excited state of energy 7416 keV can cascade, first to the first excited state (1779 keV) with an energy of 5636 keV, or transition directly to the ground state. The ninth excited state of energy 7416 keV can cascade through the third excited state (4980 keV) with energy 2400 keV or the first excited state (1779 keV) with energy 5600, or transition directly to the ground state. These γ-transitions are all

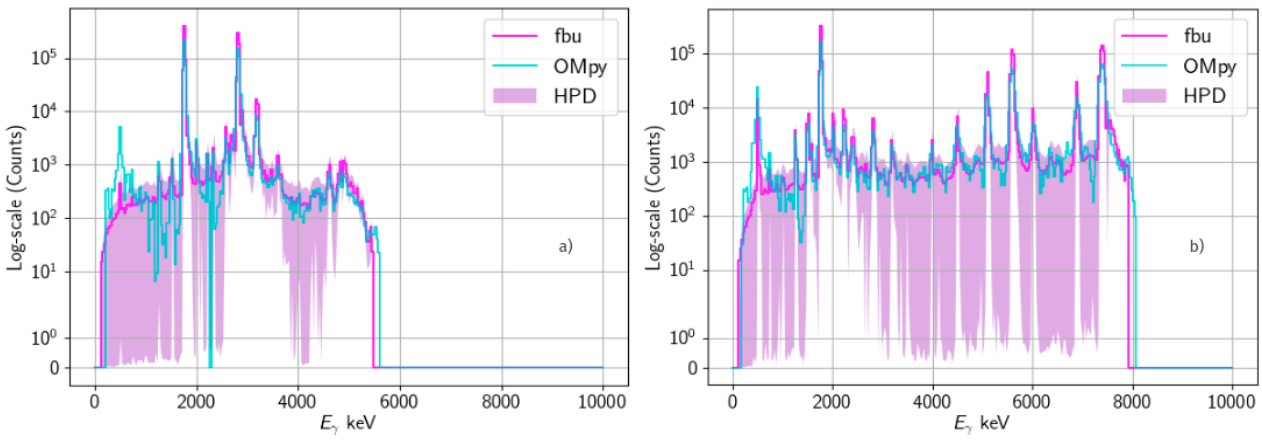

Figure 5.19: The unfolded spectrum of FBU and OMpy for a) the 4^+ excitation state (4617 keV) projected onto the E_γ axis, and b) the 2^+ excitation state (7416 keV) projected onto the E_γ axis,

visible in the figure. In addition, there are single and double escape peaks that have not been fully re-distributed to their respective full energy peaks, the same is the case for the 511-peak. The $(\frac{2}{5})^+$-transition of ^{29}Si and 2^+-transition of ^{30}Si are also present. In the end, the third excited state (4980 keV) cascades through the first excited state with an energy of 3200 keV, and this peak can also be seen in the figure. The case of tenth excited state of energy 7416 keV is an example of a spectrum where there are many cascades and it is not clear cut to identify all features of the spectrum. In most cases where one is investigating the behaviour of a nucleus, it is for γ-ray transitions in continuum, so to identify all the features of the spectrum is highly difficult. The result for the tenth excited state of energy 7416 keV, shows that FBU is able to unfold a complex spectrum to the same degree as the iterative method with the inclusion of uncertainty estimates.

Chapter 6

Discussion

The following chapter aims to shed more light on the results presented in the previous chapter. In section 1, the background subtraction versus the inclusion of background in the FBU will be discussed. The similarities and differences between the FBU and the Iterative Method will be addressed in section 2. A model checking of the FBU will be presented in section 3. In the end section 4 will address the uncertainty estimates that the OMpy library offers, are shown for the $^{28}\mathrm{Si}(p, p'\gamma)$ reaction.

1 Background Subtraction

The raw matrix without background subtraction is obtained by

$$\text{raw} = (\text{prompt} + \text{bg}) - \text{bg\_ratio} \times \text{bg}, \tag{6.1}$$

where bg is the background that is found in the time spectra in figure 3.8, prompt is the prompt peak and bg_ratio is the ratio of the background chosen [37]. The background subtracted spectrum is known, but the actual background spectrum inside the prompt peak is unknown, and therefore bg_ratio is used to subtract more or less background. When enough background is subtracted the counts above the $E_x = E_\gamma$-line should sum up to zero. If the background is greater than the counts in the raw data the subtraction will lead to negative counts. Since negative counts is 'unphysical' the function *fill_and_remove_negative* from the OMpy library [37] is used, and it takes as input a window size. The function includes two functions called *fill_negative_gauss* and *remove_negative*. The first function takes as input the values of the raw matrix, the E_γ axis and window size. It aims to fill the negative counts with positive counts from neighbouring bins that are within the window size of the neighbouring bins. The basic idea is that the negative counts are connected to the resolution and can therefore be filled from a channel within the resolution [37]. The window size should therefore reflect the detector resolution. By using a truncated Gaussian distribution, weights are made and used to fill the channels that contain negative counts. The truncation is solely for speeding up the code. Since the function is based on

weighting there might still be negative counts left in the matrix, meaning there are not enough bins with positive counts available within the window size. Then the *remove_negative* function is used, which simply sets the negative counts left to zero. Both codes from the OMpy library source code, can be found in Appendix A. Ideally, after subtracting the background, if enough background is subtracted the counts should approximately fluctuate around zero. If that is the case, filling and removing negatives to 'zero-out' the negatives is a reasonable solution. Under the study of the first excited state of ^{28}Si the pile-up rate was high which again impacts the counts inside the 511-annihilation peak. The background ratio was thus set to 1.2, so enough background would be subtracted in the area. This resulted in the sum of the counts for higher energies (> 2235 keV) being -182. It is not zero, but usually the filling and removing negatives can be justified for a few hundred counts if the statistics of the experiment is good. Meaning, that there is considerable amount of counts which in this case equals 1621107 for the 1st excited state. The negative counts then amounts to 0.011% of the total amount of counts. Figure 6.1 shows the raw spectrum including background, the background itself and the background subtracted raw spectrum without filling and removing negatives of the 1st excited state of ^{28}Si, on a symmetric logarithmic scale. In the threshold area the background includes counts while the raw spectrum including background does not. It is important to remember that the background shown here is taken from a pulse beam entering the detector at another time than the prompt beam. The actual background of the prompt beam is not known. What one can say from this plot is that the background of the prompt peak does not have counts in the threshold area while the background used for subtraction does. In this area the justification of removing negatives holds.

When using FBU the subtraction of the background is avoided. Instead it is included into the likelihood by adding it to the folded: $\boldsymbol{f} = \boldsymbol{B} + R\boldsymbol{T}$. The likelihood measures the fit between the statistical model and the evidence sampled from the data. By including the background in the likelihood the method knows the background to prompt ratio, and can thus deduce the probability of counts belonging to the background or the raw spectrum. As seen in figure 5.16 in section 2 from chapter 5, the FBU give better results when including the background when compared to the subtracted raw spectrum, when looking at the relative difference. Also, the factor the raw specturm and the folded spectrum of FBU is 1.015. This might indicate that the filling and removing negatives counts works well in this case, since the folded of FBU coincides with the background subtracted raw spectrum. In those cases when the background subtraction does not fluctuate around zero, FBU offers a solution which can be explored further.

2 The Unfolding Methods

When using unfolding methods on the raw data the goal is to deduce information that can be explained physically, and then used in further analysis. What the methods are doing is avoiding matrix inversion. Even so, many unfolding methods suffer from huge fluctuations and can sometime result in negative counts. The iterative method has proved itself to be valuable and has been tested against unfolding methods such as matrix inversion with and without regularization in the paper of P.-A. Söderström et al. Where the conclusions was that 'the iterative method gives the best qualitative reproduction of the emitted γ-ray spectrum' [41]. It is the only

method used for unfolding in the Oslo Method, which have produced many highly regarded scientific papers using the Iterative approach. Unfortunately, the method can give negative fluctuations and offers no uncertainty estimates. In figure 5.7 from chapter 5, it can be seen that the method fluctuates heavily for energies lower than the 1779 peak, though these are not negative. The FBU on the other hand is stable in this area. In cases where there are great negative fluctuations, the FBU offers a solution. Since the uniform prior has lower bounds set to zero the method will find the most likely positive solution which is desired, this is a strength compared to the iterative method. In addition, the uncertainty estimates and can be propagated into further analysis. The unfolding has been the main suspect when it comes to inconsistencies in the Oslo Method [42] [36], so with the usage of FBU this can finally be investigated. The main problem with using Bayesian inference on heavy problems like unfolding is the time it takes to sample. Over the years the methods have been greatly improved but there is still a way to go in terms of computation time. Most often there is a raw matrix that the user wishes to unfold, and if using the FBU library (built on the PyMC3 library [12]) this can take days of running time, even when parallelizing the sampling. The computer used in this thesis was a Lenovo Yoga with Intel® Core™ i7-7500U CPU @ 2.70GHz × 4 and 7.5 GiB memory. This means that the process can be faster using a computer with more than two cores and more memory, but it can never be expected to be as fast in terms of the Iterative method computation time. Another way to reduce sampling time is to re-bin the spectrum, since FBU

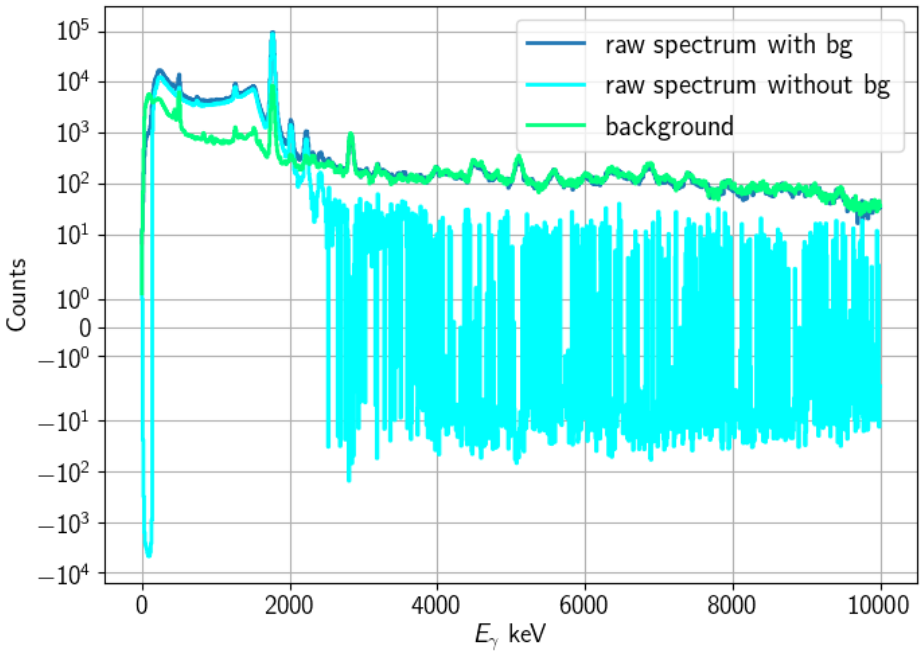

Figure 6.1: The first excited state of ^{28}Si including the raw spectrum including background, the background itself and the background subtracted raw spectrum before filling and removing negatives. The y-axis is symmetric logarithmic scale.

makes a prior for each bin the computation time extends for each bin. However, this has to be done without the loss of important features in the spectrum. In cases where fine binning is necessary it might become to complex for a typical office computer to handle. This again can be solved with varying bin size. The FBU method offers a intuitive way of analysing the raw data and unfolding them and also including uncertainty estimates. In addition the model itself can be evaluated in greater detail than in the case of other unfolding methods with various model checking techniques, some of them presented in the next section.

3 Model Checking of FBU

To make sure the model used in this study gives consistent answers, Bayesian statistics offers tools to evaluate it which are described below. The model under evaluation is

$$\begin{aligned} \boldsymbol{T} &\sim \text{Uniform}(\text{lower}, \text{upper}) \\ \boldsymbol{f} &= \boldsymbol{B} + R\boldsymbol{T} \\ \hat{\boldsymbol{T}} &\sim \text{Pois}(\boldsymbol{f}), \end{aligned}$$

where the background isincluded, but can be omitted if using background subtraction. The first check of the model is a posterior predictive check, which was explained in chapter 2. The basic idea of posterior predictive checks is that it simulates data raw^{rep} by drawing samples under the fitted model, and then compare these to the actual observed data raw [6]. Doing this might reveal systematic discrepancies between the real data and the simulated data raw^{rep} under the model. An important note on posterior predictive checks is that they do not reveal if the model is better than other models, it is just a check to verify that the model makes sense in respect to the data. In theory this check is re-using data, thus violating the likelihood principle. However, since it is used as a check to see if the model fits the data and nothing else, it has been largely accepted as a model checking method and it is implemented into the PyMC3 library [12].

Figure 6.2 a) shows the replicated sampled data raw^{rep} and their mean compared to the raw spectrum of the first excited state of ^{28}Si. The background has been subtracted, and the diagonal and threshold have been cut out of the spectrum, since they do not contain counts. Figure 6.2 b) shows the relative difference between the replicated data and the raw spectrum. Here one sees the main discrepancies between the model and the actual data. There are huge differences in the threshold area and where the 1779 keV-peak is centered. In figure 6.3 the same is shown but for the background-included spectrum. The same discrepancies are present but to a much smaller degree. It is worth noting that a relative difference in the threshold area seem to have an effect on the 1779 keV-peak. In chapter 5 the relative difference changed for the folded of the FBU when adding the zeros to the threshold area as was shown by comparing figure 5.8 with 5.9. Looking at the projected response for the same cut in excitation energy as was made in the raw matrix ($[1400, 2200]$ keV) in figure 6.4, shows the threshold

area. The prior of the unfolded is multiplied with the response matrix, and the effects from the response thus has an influence on the posteriors. Since the data is zero in this area the effects seen in the threshold area stem from the prior and response function. This might be solved by updating the response matrix according to the digital electronics.

Another interesting observation is that the model seems to perform better when the background is added. It can mean that information is lost when setting the counts below the diagonal to zero, at least in the case of unfolding with FBU. The more information fed to the method the better it performs it seems.

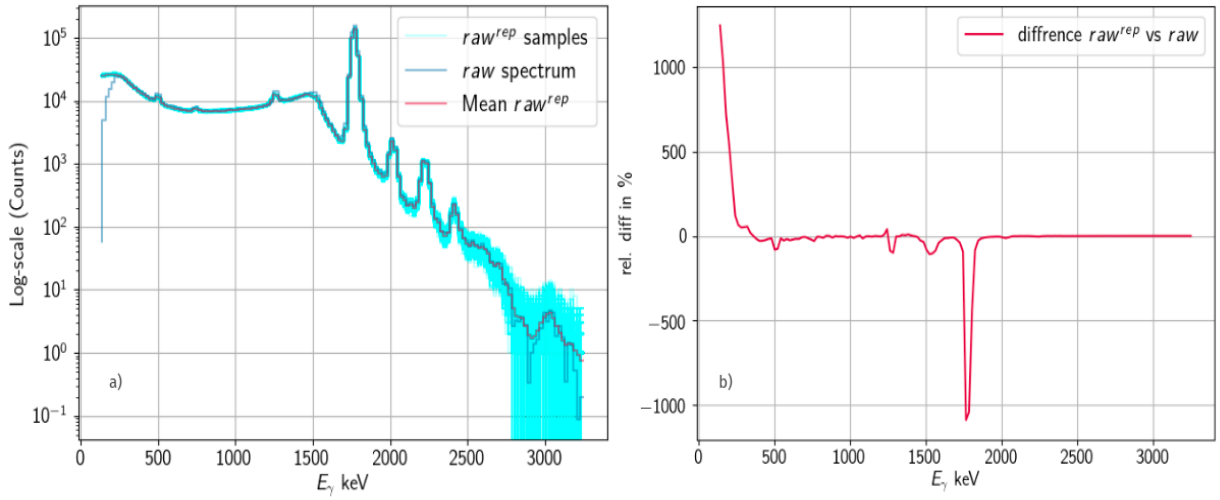

Figure 6.2: a) Replicated data raw^{rep} drawn from samples versus the actual raw spectrum where the diagonal and threshold is cut. b) The relative difference for the replicated data raw^{rep} relative to the raw spectrum

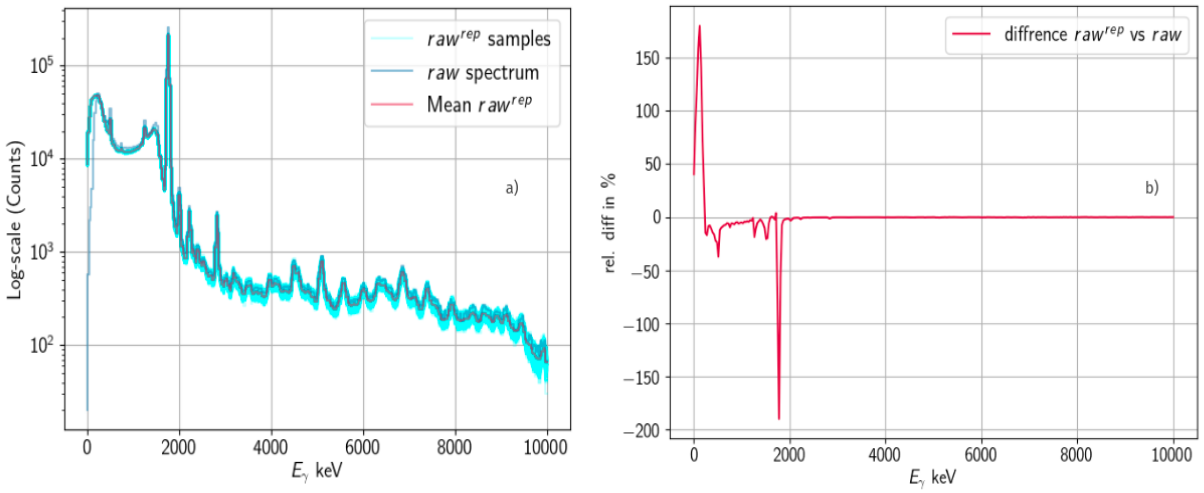

Figure 6.3: a) Replicated data raw^{rep} drawn from samples versus the actual raw spectrum with the inclusion of background. b) The relative difference for the replicated data raw^{rep} relative to the raw spectrum

Another technique used to check the model is to compare the prior probability to the posterior distribution as mentioned in chapter 2. Is the uniform prior used the best for describing the data? From looking at the posteriors for each truth-bin in figure 5.5 they vary from bin to bin, but none of them are uniformly distributed. By inspecting the posteriors a varying prior for each bin can be suggested, for example: A normal distribution where there should be peaks and a exponential distribution in the bins that are skewed towards zero. The problem with doing so is the lack of knowledge of the spectra in question. In the case of the first excited state of ^{28}Si the features of the spectrum can be identified, but this is far from the case with many other nuclei. The prior for the bin containing the counts from the 1779 keV peak ranged from 0 to 1.522e6, after correcting for efficiency and the scale of the full energy peak. Before running the sampling it was clear that the counts in this peak should not be around 0, since it is a full energy peak. Then again, making a informed prior is challenging when taking the spectrum as a whole into consideration. Features can be hidden inside the Compton distribution and therefore be easy to miss. The uniform prior does not seem to have much influence on the posterior distribution, meaning the evidence of the data is altering the prior belief. This is desired in the case of unfolding γ-ray spectra, since the outcome is often unknown. A well defined objective prior that allows the likelihood of the data to alter the posteriors so values for the unfolding can be estimated is highly desired. The only informed decision for the prior in this case, is that the lower bound is set to zero and the upper bound is corrected for the efficiency and the scale of the full energy peak.

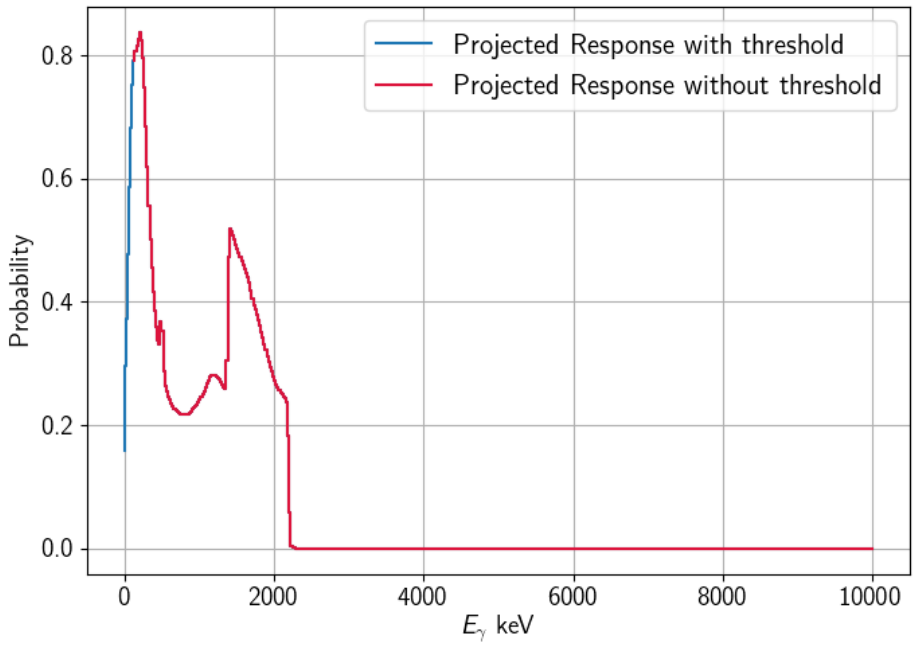

Figure 6.4: The Response projected onto the E_γ axis for excitation energy $E_x = [1400, 2200]$ keV.

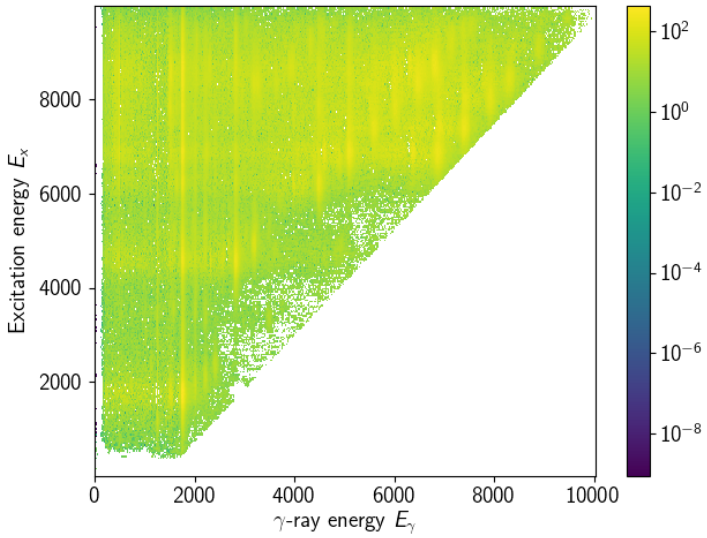

Figure 6.5: The standard deviation of the iterative unfolding method.

4 The Statistical Variance of the Iterative Method

Using the class called Ensemble from the OMpy library, the statistical variance from the raw data can be estimated. This enables the estimation of the variance for the iterative unfolding method and also for later stages in the Oslo Method. The method was explained in chapter 4. In figure 6.5 the matrix containing the standard deviation (only) after unfolding with the iterative method is shown. The method takes as input the raw matrix shown in figure 5.1, then 100 perturbed raw matrices where generated by randomly drawing from a Poisson distribution, as was explained in chapter 4. This is an unbiased estimation since the Poisson estimator λ equals the estimated value of the number of counts, thus giving the statistical variance (statistical errors) of the unfolded matrix.

In figure 6.6 the iterative method with its standard deviation and FBU with uncertainty estimates for the first excited state of ^{28}Si is presented, with background subtracted. The statistical variance of the iterative unfolding is wide in the areas where the method itself fluctuates more. Both the statistical variance and the HPD-interval are narrower in the area with high amount of counts. This is expected since in the areas with more information (i.e higher statistics) both variance for the iterative method and the uncertainty of FBU is smaller. In chapter 4 the difference between statistical error (variance) and systematic errors (bias) was explained. Figure 6.6 shows the variance of the iterative unfolding method and the FBU with its uncertainty in the estimations. The FBU offers an evaluation of itself by giving uncertainty in the estimated unfolded spectrum. Since the actual truth spectrum is not always achievable one cannot know the exact bias of the model, but rather the uncertainty of the method. The ensemble method together with FBU gives a wider picture of the error estimates of the unfolding, and both can be propagated into further analysis.

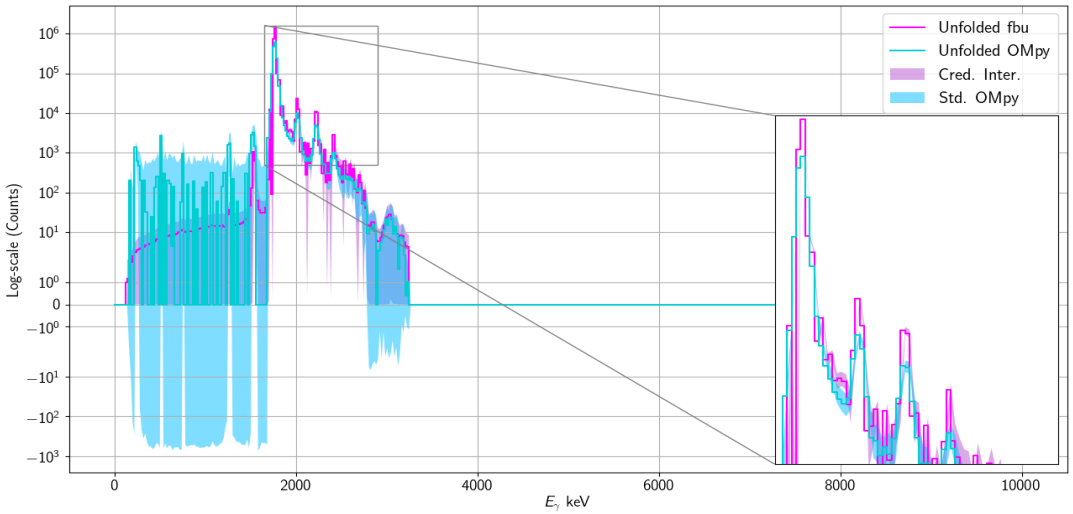

Figure 6.6: The iterative unfolding (OMpy) with standard deviation and FBU with HPD-interval, where the y-axis is symmetric logarithmic scale.

Chapter 7

Summary and Future Outlook

1 Summary

In this thesis the unfolding of γ-ray spectra attained from OSCAR, an array of $LaBr_3$:Ce scintillator detectors, was investigated using a new method called Fully Bayesian Unfolding. The FBU has never been investigated fully with nuclear physics data but was developed for high energy physics data [38]. The library is built on another library called PyMC3 [12] which is a package for Bayesian statistical modeling in python.

The FBU method uses Bayes' theorem and thus samples from a likelihood and prior resulting in posterior distributions for each bin. The posterior distributions can be evaluated in binned histograms where the probability of each values can be seen. To extract the uncertainties of the model a credible interval was chosen, in this thesis the highest probability density interval. This will give the interval including the highest probable values of the posterior distribution. There exist other credible intervals such as the Equal Tailed Interval (ETI), then including the median of the distribution. Both are often used within Bayesian statistics and it is advised that the user of the method evaluates what best describe the posterior distributions. In addition, the Maximum A Priori (MAP) was also considered during this study. To calculate the MAP a optimization code was used taken from the PyMC3 library [12]. The MAP is efficient and estimates the mode of the posterior distribution. The drawback of the method is that the mode does not always represent the posterior distribution and it does not provide uncertainty estimates. Since the unfolding requires estimated values with uncertainty the mean of the HPD-interval was chosen to describe the unfolded spectrum.

The aim of the thesis was to use Bayesian statistics to unfold raw data and thus attain uncertainty estimates that can be propagated into further analysis like the Oslo-Method. That said, the unfolding is general and can be used on data intended for other means. The data used to evaluate the FBU was taken from an experiment on the ^{28}Si(p,p$'\gamma$) reaction. The first excited state of this reaction has a clear full energy peak and features that are

easily identified. Thus making it a good candidate to evaluate the method. In addition the first test of unfolding a matrix with the unfolding method was done.

The experimental setup consisted of the particle detector SiRi in backward angles and OSCAR-array placed 16 cm away from the target. The target consists of 92.223% pure ^{28}Si, 4.685% ^{29}Si and 3.093% ^{30}Si [40] and was $8-9$ mg/cm^2 thick. The thick target and SiRi in backward angles resulted in a poorer resolution than can be expected in forward angles. With the resolution being 360 keV for the full energy peak of the first excited state. The pile up rate was high making the contribution of the 511 keV peak high.

When it comes to unfolding it is assumed that the response matrix is known. The response matrix used is made by simulating γ-ray transition with the GEANT4 toolkit [30–33]. For the iterative unfolding method each row of the matrix is normalized and sums up to zero. In FBU however, the response is scaled by the total efficiency. The total efficiency of OSCAR has not been fully established, but a value of 50% was assumed in this study for the 1332 keV energy. This value was found by calculating the geometric efficiency and then assuming that the total efficiency was reduced by 14% of the geometric efficiency (as was the case for the previous detector setup CACTUS). The high pile up rate makes the condition for which the response matrix was made quite different to the condition of the experiment under study. In addition, the response matrix was made for analog electronics, the system has later been upgraded to digital electronics. The treatment of the threshold area for the analogue electronics and digital electronics do not seem to coincide, at least not according to the results obtained in this work.

The iterative unfolding method was compared to FBU. The main difference between the models is that the iterative method does not include background. Rather it requires background subtraction, and the user has to specify a diagonal where all the counts below it are set to zero. This treatment of the raw data has its physical motivation in the fact that counts over the $E_x = E_\gamma$ line cannot stem from γ-ray transitions. In the case of the ^{28}Si(pp$'\gamma$) reaction, the background subtraction seems reasonable. Nonetheless, the FBU performs better when the background is included. The iterative unfolding method fluctuates more than the FBU for lower energies (¡ 1700 keV).

Using FBU offers a model checking to a greater extent than just evaluating the uncertainty of the credible interval. In this thesis a posterior predictive check was done for the first excited state of ^{28}Si with and without background subtraction. By replicating data under the model, the models fit to the data can be considered. For the background subtracted case, the discrepancies in the threshold area and for the 1779 peak where huge. Indicating that an evaluation of the treatment of the threshold area is needed. For the background-included spectrum the discrepancies in these areas where less but still present. In addition to the posterior predictive checks the prior chosen was evaluated. In the case of unfolding a uniform prior was used and it had the desired effect: it did not seem to influence the posterior distributions in an unwanted way.

The OMpy library [37] offers a method to estimate the statistical variance of the iterative unfolded method by using the class ensemble. This together with the FBU gives a broader knowledge of the variance and uncertainty of the unfolding methods used in this thesis.

2 Outlook

In the present work a new method for unfolding has been investigated. Since FBU has never been explored to this extent for γ-ray spectra, there is still much work to be done. The method has shown itself to be intuitive and helped shed light on some possible parts of the unfolding process that may need to be revised, like the treatment of the threshold area. The response matrix has been exploited as a possible source of the discrepancies seen in the unfolding. That said, one can never expect a perfect response function for each experiment. The condition under which the response matrix was made does not always coincide with the condition of the experiment. Also, the response matrix of OSCAR is a work in progress.

The FBU as a method for unfolding γ-ray spectra is stable and shows great potential to further use. As shown in the results for both the first excited state and unfolded matrix, the FBU re-distributes more counts to the full energy peaks, fluctuates less than the iterative unfolding method and offers uncertainty estimates. The problem is the sampling and the computation time. This is often the down side of using Monte Carlo sampling on high dimensional problems. The computation time gets better when reducing the bin size of the spectrum, but this is not always desired. This can be solved by using varying bin size, which should be explored further. Nonetheless, the user must be prepared for more computation time than the iterative unfolding method.

The FBU library is built on the PyMC3 library and the developers are currently working on PyMC4. The main difference between the two is that PyMC4 [43] is built on top of TensorFlow and the TensorFlow Probability library [44], instead of the theano library [45] which PyMC3 is built upon. It might be worth exploring the new library when it is published to full extent. The FBU library is developed for high energy physics unfolding and the code was made in 2013 [38]. Other probabilistic libraries might be worth looking into. When a good review of the possible libraries has been done a code can be implemented into the OMpy library. Made specifically for the unfolding of γ-ray spectra, not for high energy physics. Bayesian statistics is being used more extensively now than ever before since better sampling algorithms have been developed. Many computational scientist are working on bettering the computation time, and it is therefore important to keep up to date on the topic and follow the developments.

When it comes to the unfolding itself there are more possibilities to explore outside the scope of this thesis. For example, the prior used in this work was the uniform distribution which was corrected for efficiency and the scale of the full energy peak. Other priors can easily be tested and can work better for some cases than the uniform prior. In the original paper of FBU written by Georgios Choudalakis [38], priors including regularization showed promising results. Conjugate priors can also be explored, the conjugate prior distribution of a Poisson being a Gamma distribution. Choosing more specified priors that do not distribute equal probability to each draw

(like the uniform prior will) can make the computation time for FBU faster since the sampling algorithm faster reaches convergence.

Model checking is a topic that was briefly touched in this thesis and can be done to much greater extent. A sensitivity analysis could be done: then comparing the currently used model to a simpler model or a complex model. If the currently used model is sensitive to these changes it might indicate problems. If doing a thorough check of the model the outcome can be achieving more knowledge of model and also potentially bettering it for future cases.

The motivation of this thesis was to test a new unfolding method for γ-ray spectra which could provide uncertainty estimates, as well as explore the possibilities for less fluctuations and negative counts in the resulting unfolding spectra. Bayesian statistics is a powerful tool and the FBU method has shown itself to be very useful in the unfolding of γ-ray spectra. Providing priors so a lower bound can be set on the outcome, preventing it from giving negative values. The uncertainty estimates will prove useful when propagating into further analysis and can also be helpful when it comes to the unfolding itself. The user can test the model with various model checking tools and refine it for their own cases. It deserves more revising and to be tested for other reaction.

Appendices

Appendix A

```
import numpy as np
from scipy.stats import truncnorm
from typing import Optional, Tuple, Iterator, Any, Union

def fill_negative_gauss(array: np.ndarray, Eg: np.ndarray,
                        window_size: Union[int, float, np.array],
                        n_trunc: float = 3) -> np.ndarray:

    if isinstance(window_size, (int, float)):
            window_size = np.full(array.shape[1], window_size)
            sigma = window_size/2.355  # convert FWHM to sigma
        else:
            assert len(window_size) == array.shape[1], "Array length incompatible"
            sigma = window_size/2.355  # convert FWHM to sigma

        # generate truncated gauss for each Eg bin, format [Eg-bin, gauss-values]
        lower, upper = Eg - n_trunc*sigma, Eg + n_trunc*sigma
        a = (lower - Eg) / sigma
        b = (upper - Eg) / sigma
        gauss = [truncnorm(a=a, b=b, loc=p, scale=sig).pdf(Eg)
                 for p, sig in zip(Eg, sigma)]
        gauss = np.array(gauss)

        array = np.copy(array)
        N_Ex = array.shape[0]
        for i_Ex in range(N_Ex):
            row = array[i_Ex, :]
```

```
            for i_Eg in np.nonzero(row < 0)[0]:
                positives = np.where(row < 0, 0, row)
                weights = positives * gauss[i_Eg, :]

                for i_from in np.argsort(weights):
                    if row[i_from] < 0:
                        break
                    shuffle_counts(row, i_from, i_Eg)
                    if row[i_Eg] >= 0:
                        break

    return array

def remove_negative(self):
    """ Entries with negative values are set to 0 """
    self.values = np.where(self.values > 0, self.values, 0)
```

[37], [46], [47], [48].

Appendix B

Link to Github where all the codes and files used in this thesis are available:

Github Link

Bibliography

[1] T. Bayes, "An essay towards solving a problem in the doctrine of chances," *Resonance*, 2003. [Online]. Available: https://doi.org/10.1007/BF02883540

[2] D. S. Sivia and J. Skilling, *Data Analysis A Bayesian Tutorial Second Edition.* Great Clarendon Street, Oxford OX2 6DP: Oxford University Press, 2006.

[3] M. Guttormsen, A. Brüger, T. Hansen, and N. Lietaer, "The siri particle-telescope system," *Nuclear Instruments and Methods in Physics Research*, 2011. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168900211010205?via%3Dihub

[4] A. Gelman, C. J. B., H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis 2nd Edition.* Corporate Blvd., Boca Raton, Florida 33431: Chapman and Hall/CRC, 2003.

[5] J. Myung, "Tutorial on maximum likelihood estimation," *Journal of Mathematical Psycholohy*, 2003.

[6] A. Gelman and H. J., *Data Analysis Using Regression and Multilevel/Hierarchical pp 387-341.* Cambridge University, 2006.

[7] M. Taboga, *Lectures on Probability Theory and Mathematical Statistics - 3rd Edition.* CreateSpace Independent Publishing Platform, 2017.

[8] C. Davidson-Pilon *et al.*, *Bayesian Methods for Hackers.* Cameron Davidson-Pilon, 2013.

[9] K. P. Murphy, *Machine Learning: A Probabilistic Perspective.* The MIT Press, 2012.

[10] R. E. Walpole, R. H. Myers, S. L. Myers, and K. E. Ye, *Probability and Statistics for Engineers and Scientists, 9th Edition.* Pearson, 2012.

[11] M. Hoffman and G. Andrew, "The no-u-turn sampler: Adaptively setting path lengths in hamiltonian monte carlo," *Journal of Machine Learning Research*, 2014. [Online]. Available: http://jmlr.org/papers/volume15/hoffman14a/hoffman14a.pdf

[12] J. Salvatier, T. V. Wiecki, and C. Fonnesbeck, "Probabilistic programming in python using PyMC3," *PeerJ Computer Science*, vol. 2, p. e55, apr 2016. [Online]. Available: https://doi.org/10.7717/peerj-cs.55

[13] R. Neal, *Handbook of Markov Chain Monte Carlo, chapter 5: MCMC Using Hamiltonian Dynamics.* CRC Press, 2011.

[14] M. Creutz, "Global monte carlo algorithms for many-fermion systems." . *Physical Review D*, 1988.

[15] The homepage of the oslo cyclotron laboratory (ocl). [Online]. Available: https://www.mn.uio.no/fysikk/english/research/about/infrastructure/ocl/

[16] M. Guttormsen, T. Ramsoey, and J. Rekstad, "The first generation of -rays from hot nuclei," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 18, pp. 518–523, 1987.

[17] A. Schiller, L. Berghold, M. Guttormsen, E. Melby, J. Rekstad, and S. Siem, "Extraction of level density and γ strength function from primary γ spectra," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 447, no. 3, 2000.

[18] L. Henden, L. Bergholt, M. Guttormsen, J. Rekstad, and T. Tveter, "On the relation between the statistical γ-decay and the level density in 162 Dy," *Nucl. Phys. A*, vol. 589, pp. 249–266, 1995.

[19] M. Guttormsen, "The unfolding of continuum γ-ray spectra," *Nuclear Instruments and Methods in Physics Research A*, 1996. [Online]. Available: https://www.researchgate.net/publication/223322971_The_unfolding_of_continuum_g-ray_spectra

[20] W. R. Leo, *Techniques for Nuclear and Particle Physics Experiments.* CH-1814 La Tour de Peilz, Switzerland: Springer, 1993.

[21] H. Bethe, "Bremsformel für elektronen relativistischer geschwindigkeit," *Zeitschrift für Physik*, 1932. [Online]. Available: https://doi.org/10.1007/BF01342532

[22] F. Bloch, "Zur bremsung rasch bewegter teilchen beim durchgang durch materie," *Annals of Physics*, 1933. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.19334080303

[23] A. Giaz, L. Pellegri, S. Riboldi, F. Camera, N. Blasi, C. Boiano, A. Bracco, S. Brambilla, S. Ceruti, S. Coelli, F. Crespi, M. Csatlós, G. J. Frega, S., A. Krasznahorkay, S. Lodetti, B. Million, A. Owens, F. Quarati, L. Stuhl, and O. Wieland, "Characterization of large volume 3.5×8 $labr_3 : ce$ detectors," *Nuclear Instruments and Methods in Physics Research*, 2013. [Online]. Available: https://www-sciencedirect-com.ezproxy.uio.no/science/article/pii/S016890021301111X

[24] S. Tavernier, *Experimental Techniques in Nuclear and Particle Physics.* Pleinlaan2, 1050 Bruxelles: Springer, 2010.

[25] K. Buchtela, *Radiochemical Methods— Gamma-Ray Spectrometry - Encyclopedia of Analytical Science.* Vienna University of Technology, Vienna, Austria: Elsevier Ltd., 2005.

[26] V. W. Ingeberg, F. Zeiser, and E. Lima, "Qkinz kinematics calculator: oslocyclotronlab/qkinz." [Online]. Available: https://zenodo.org/record/1154143#.XufnLjczais

[27] XIA. (2019) Xia, pixie-16 user manual version 3.06. [Online]. Available: https://www.xia.com/dgf_pixie-16.html

[28] G. Cowan, "A survey of unfolding methods for particle physics," *Conf.Proc.C*, 2002. [Online]. Available: http://www.ippp.dur.ac.uk/old/statistics/proceedings/cowan.pdf

[29] ——, *Statistical Data Analysis*. Great Clarendon Street, Oxford: Oxford University Press, 1998.

[30] The geant4 homepage. [Online]. Available: https://geant4.web.cern.ch/

[31] J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Arce Dubois, M. Asai, G. Barrand, R. Capra, S. Chauvie, R. Chytracek, G. A. P. Cirrone, G. Cooperman, G. Cosmo, G. Cuttone, G. G. Daquino, M. Donszelmann, M. Dressel, G. Folger, F. Foppiano, J. Generowicz, V. Grichine, S. Guatelli, P. Gumplinger, A. Heikkinen, I. Hrivnacova, A. Howard, S. Incerti, V. Ivanchenko, T. Johnson, F. Jones, T. Koi, R. Kokoulin, M. Kossov, H. Kurashige, V. Lara, S. Larsson, F. Lei, O. Link, F. Longo, M. Maire, A. Mantero, B. Mascialino, I. McLaren, P. Mendez Lorenzo, K. Minamimoto, K. Murakami, P. Nieminen, L. Pandola, S. Parlati, L. Peralta, J. Perl, A. Pfeiffer, M. G. Pia, A. Ribon, P. Rodrigues, G. Russo, S. Sadilov, G. Santin, T. Sasaki, D. Smith, N. Starkov, S. Tanaka, E. Tcherniaev, B. Tome, A. Trindade, P. Truscott, L. Urban, M. Verderi, A. Walkden, J. P. Wellisch, D. C. Williams, D. Wright, and H. Yoshida, "Geant4 developments and applications," *IEEE Transactions on Nuclear Science*, vol. 53, no. 1, pp. 270–278, 2006.

[32] S. Agostinelli, J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Arce, M. Asai, D. Axen, S. Banerjee, G. Barrand, F. Behner, L. Bellagamba, J. Boudreau, L. Broglia, A. Brunengo, H. Burkhardt, S. Chauvie, J. Chuma, R. Chytracek, G. Cooperman, G. Cosmo, P. Degtyarenko, A. Dell'Acqua, G. Depaola, D. Dietrich, R. Enami, A. Feliciello, C. Ferguson, H. Fesefeldt, G. Folger, F. Foppiano, A. Forti, S. Garelli, S. Giani, R. Giannitrapani, D. Gibin, J. Gómez Cadenas, I. González, G. Gracia Abril, G. Greeniaus, W. Greiner, V. Grichine, A. Grossheim, S. Guatelli, P. Gumplinger, R. Hamatsu, K. Hashimoto, H. Hasui, A. Heikkinen, A. Howard, V. Ivanchenko, A. Johnson, F. Jones, J. Kallenbach, N. Kanaya, M. Kawabata, Y. Kawabata, M. Kawaguti, S. Kelner, P. Kent, A. Kimura, T. Kodama, R. Kokoulin, M. Kossov, H. Kurashige, E. Lamanna, T. Lampén, V. Lara, V. Lefebure, F. Lei, M. Liendl, W. Lockman, F. Longo, S. Magni, M. Maire, E. Medernach, K. Minamimoto, P. Mora de Freitas, Y. Morita, K. Murakami, M. Nagamatu, R. Nartallo, P. Nieminen, T. Nishimura, K. Ohtsubo, M. Okamura, S. O'Neale, Y. Oohata, K. Paech, J. Perl, A. Pfeiffer, M. Pia, F. Ranjard, A. Rybin, S. Sadilov, E. Di Salvo, G. Santin, T. Sasaki, N. Savvas, Y. Sawada, S. Scherer, S. Sei, V. Sirotenko, D. Smith, N. Starkov, H. Stoecker, J. Sulkimo, M. Takahata, S. Tanaka, E. Tcherniaev, E. Safai Tehrani, M. Tropeano, P. Truscott, H. Uno, L. Urban, P. Urban, M. Verderi, A. Walkden, W. Wander, H. Weber, J. Wellisch, T. Wenaus, D. Williams, D. Wright, T. Yamada, H. Yoshida, and D. Zschiesche, "Geant4—a simulation toolkit," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 506, no. 3, pp. 250 – 303, 2003. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0168900203013688

[33] J. Allison, K. Amako, J. Apostolakis, P. Arce, M. Asai, T. Aso, E. Bagli, A. Bagulya, S. Banerjee, G. Barrand, B. Beck, A. Bogdanov, D. Brandt, J. Brown, H. Burkhardt, P. Canal, D. Cano-Ott, S. Chauvie, K. Cho, G. Cirrone, G. Cooperman, M. Cortés-Giraldo, G. Cosmo, G. Cuttone, G. Depaola, L. Desorgher, X. Dong, A. Dotti, V. Elvira, G. Folger, Z. Francis, A. Galoyan, L. Garnier, M. Gayer, K. Genser, V. Grichine, S. Guatelli, P. Guèye, P. Gumplinger, A. Howard, I. Hřivnáčová, S. Hwang, S. Incerti, A. Ivanchenko, V. Ivanchenko, F. Jones, S. Jun, P. Kaitaniemi, N. Karakatsanis, M. Karamitros, M. Kelsey, A. Kimura, T. Koi, H. Kurashige, A. Lechner, S. Lee, F. Longo, M. Maire, D. Mancusi, A. Mantero, E. Mendoza, B. Morgan, K. Murakami, T. Nikitina, L. Pandola, P. Paprocki, J. Perl, I. Petrović, M. Pia, W. Pokorski, J. Quesada, M. Raine, M. Reis, A. Ribon, A. Ristić Fira, F. Romano, G. Russo, G. Santin, T. Sasaki, D. Sawkey, J. Shin, I. Strakovsky, A. Taborda, S. Tanaka, B. Tomé, T. Toshito, H. Tran, P. Truscott, L. Urban, V. Uzhinsky, J. Verbeke, M. Verderi, B. Wendt, H. Wenzel, D. Wright, D. Wright, T. Yamashita, J. Yarba, and H. Yoshida, "Recent developments in geant4," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 835, pp. 186 – 225, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0168900216306957

[34] F. Zeiser *et al.*, "The energy response of the oslo scintillator array oscar."

[35] R. Brun and F. Rademakers, "Root - an object oriented data analysis framework," *Nuclear Instruments and Methods in Physics Research*, 1997. [Online]. Available: http://root.cern.ch/

[36] J. E. Midtbo *et al.*, "A new software implementation of the oslo method with complete uncertainty propagation," *arXiv*, 2020. [Online]. Available: https://arxiv.org/pdf/1904.13248.pdf?fbclid=IwAR2IkI2Ewi9yGnAB36rA-BMuitQjwXzkokSq5ljg_Sdeoa0E_eUvdU3AwSA

[37] J. E. Midtbø, F. Zeiser, E. Lima, and V. W. Ingeberg. (2020) oslocyclotronlab/ompy: Ompy v1.1.0. [Online]. Available: https://zenodo.org/record/3898281#.XvE7OXUzZhE

[38] G. Choudalakis, "Fully bayesian unfolding," *arXiv*, 2012. [Online]. Available: https://arxiv.org/abs/1201.4612

[39] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[40] B. N. L. National Nuclear Data Center, "Nudat (nuclear structure and decay data)," March 18, 2008 2008.

[41] P.-A. Söderström, L. Capponi, V. Iancu, D. Lattuada, A. Pappalardo, G. V. Turturică, E. Açıksöz, D. L. Balabanski, P. Constantin, G. G. L., M. Ilie, S. Ilie, C. Matei, D. Nichita, T. Petruse, and A. Spataru, "Unfolding of sparse high-energy -ray spectra from labr3:ce detectors," *J. Instrum.*, vol. 14, 2019.

[42] A. C. Larsen, M. Guttormsen, M. Krticka, E. Betak, A. Bürger, A. Görgen, H. T. Nyhus, J. Rekstad, A. Schiller, S. Siem, H. K. Toft, G. M. Tveten, A. V. Voinov, and K. Wikan, "Analysis of possible systematic errors in the Oslo method," *Phys. Rev. C*, vol. 83, p. 23, 2011.

[43] M. Kochurov, C. Carrol, T. Wiecki, and J. Lao, "PyMC4: Exploiting Coroutines for Implementing a Probabilistic Programming Framework," *Preprint. Under review*, p. 23, 2011.

[44] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015, software available from tensorflow.org. [Online]. Available: http://tensorflow.org/

[45] The, Theano, Development, Team, R. Al-Rfou, and A. G. et al., "Theano: A Python framework for fast computation of mathematical expressions," *arXiv e-prints*, vol. abs/1605.0, p. 23, 2016.

[46] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. Jarrod Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. Carey, İ. Polat, Y. Feng, E. W. Moore, J. Vand erPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and S. . . Contributors, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods*, vol. 17, pp. 261–272, 2020.

[47] S. Van Der Walt, S. C. Colbert, and G. Varoquaux, "The numpy array: a structure for efficient numerical computation," *Computing in Science & Engineering*, vol. 13, no. 2, p. 22, 2011.

[48] T. E. Oliphant, *A guide to NumPy*. Trelgol Publishing USA, 2006, vol. 1.