Analysis notebook for improved 2vbb spectra.

The **goal** of this analysis is to compare 2 (slighty) different angular and single-electron energy distributions.

The *methodology* used to obtain the data in the analysis is as follows.

1. One spectrum was taken as a default SM model spectrum which is readily available in Falaise
2. Second spectrum was calculated based on data from Rasto Dvornicky and according to the following methodology:

- First, the angular distribution of 2vbb is given by the equation:

$$\frac{d\Gamma}{dcos(\theta)} \sim N(1 + K(\xi_{31}, \xi_{51})cos\theta)$$

equation (24) from paper by Odiviu (**O. Nitescu et al., Universe 7 (2021) 7050147**). With

$$K^{2\nu}(\xi_{31}^{2\nu}, \xi_{51}^{2\nu}) = -\frac{H_0^{2\nu} + \xi_{31} H_2^{2\nu} + \frac{5}{9}\xi_{31}^2 H_{22}^{2\nu} + \left(\frac{2}{9}\xi_{31}^2 + \xi_{51}\right) H_4^{2\nu}}{G_0^{2\nu} + \xi_{31} G_2^{2\nu} + \frac{1}{3}\xi_{31}^2 G_{22}^{2\nu} + \left(\frac{1}{3}\xi_{31}^2 + \xi_{51}\right) G_4^{2\nu}}.$$

Where H_i, G_i are the relevant phase space factors and

$$\xi_{31}^{2\nu} = \frac{M_{GT-3}^{2\nu}}{M_{GT} - \left(\frac{g_V}{g_A^{\text{eff}}}\right)^2 M_F},$$

$$\xi_{51}^{2\nu} = \frac{M_{GT-5}^{2\nu}}{M_{GT} - \left(\frac{g_V}{g_A^{\text{eff}}}\right)^2 M_F}$$

Are the NME ratios.

- Single-electron energy distribution is given by

$$\frac{d\Gamma^{2\nu}}{dE_{e_1} dE_{e_2} d(\cos\theta)} = \frac{1}{2}\frac{d\Gamma^{2\nu}}{dE_{e_1} dE_{e_2}}\left(1 + \kappa^{2\nu}(E_{e_1}, E_{e_2}, \xi_{31}^{2\nu})\cos\theta\right).$$

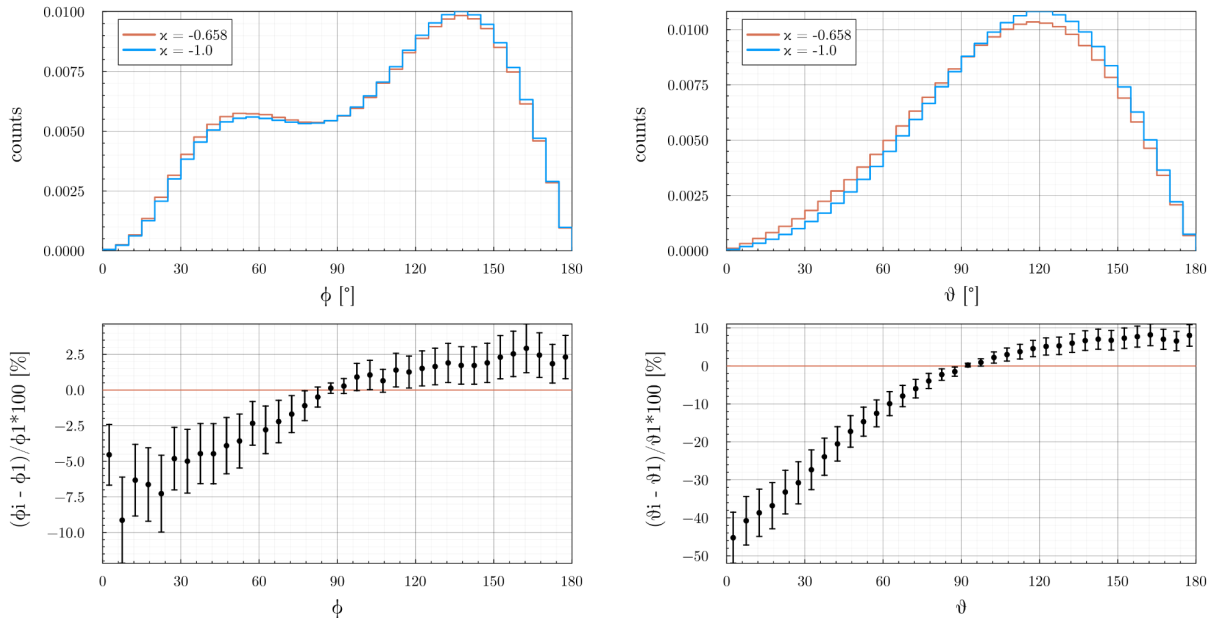
Where $\xi_{51} = 0.1397$ (assuming SSD as calculated by FS) and while ξ_{31} is a free parameter to be fit by us, we take a default value for now of ξ_{31}=0.3738 (as calculated by FS). Together using the two parameters we get $K(\xi_{31}, \xi_{51}) = -0.658.$

The simulated data were passed through *flreconstruct* and the following data cuts were applied:

1. Two negatively charged tracks reconstructed,
2. Two vertices on the source foil, within given distance from each other,
3. Sum of electron energies within the range: ``E_{sum} \in (0, 3500)~keV``,
4. Two individual Optical Module hits,
5. Two associated Optical Module hits.

Furthermore, two angular distributions were obtained from each data set. First, the so-called ϑ angular distribution, where ϑ represents the **decay** angle between the two electrons. ϑ distribution cannot be experimentally measured, it is the *truth data*. Second, the so-called φ angular distribution, where φ represents the **escape** angle - ie. the angle which can be measured with SuperNEMO, the angle between the two electrons at the moment they escape the source foil.

The two *normalized* angular distributions for each dataset are shown below.

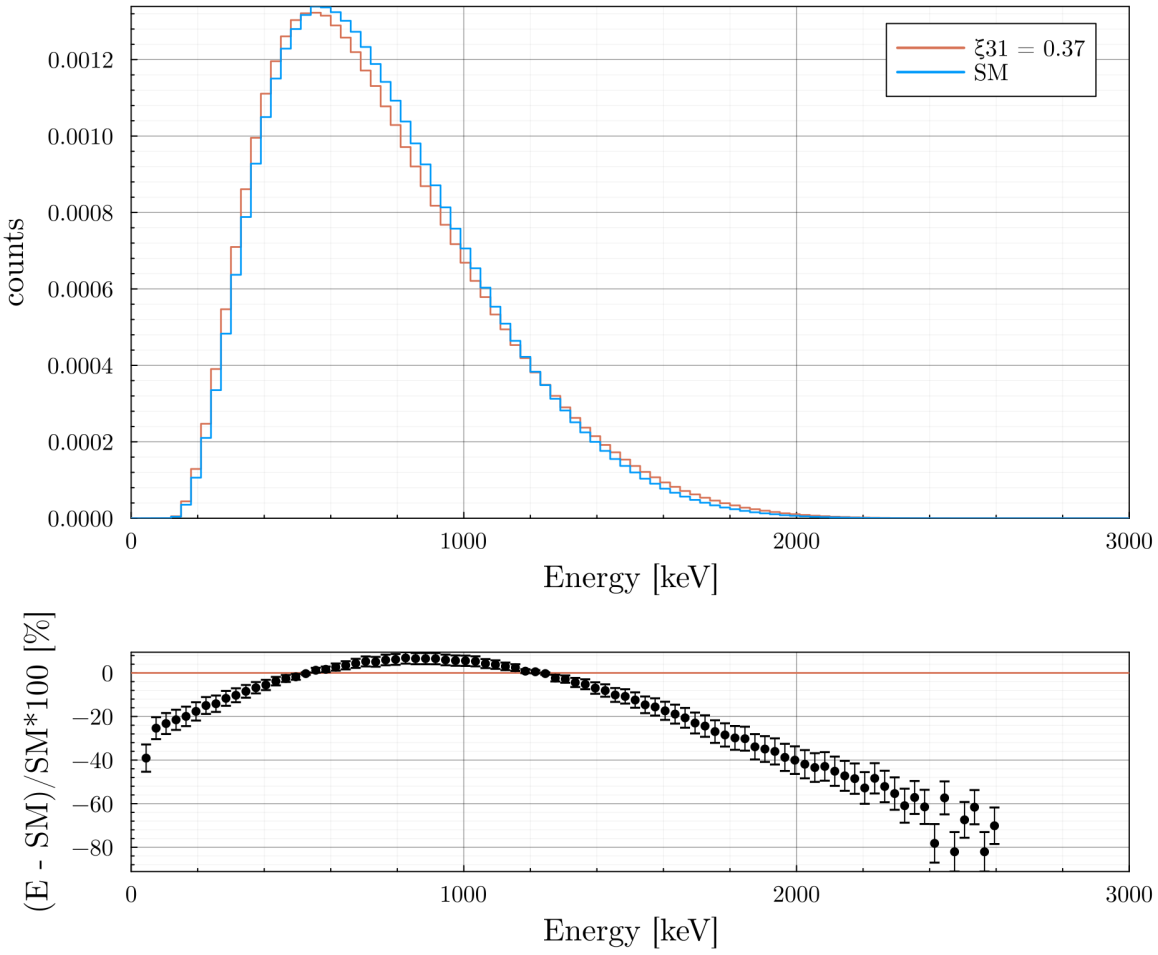


Looking at the two figures for the residuals, we can notice the following. First, for both distributions φ and ϑ the biggest difference is at low angles (though, we must take into account that the detector performance at such angles is lowered compared to higher angles).

Second, while the difference for the ϑ distribution can be up to ~ 40%, the change is nowhere near as drastic for the φ distribution. This provides a challenge, as ϑ is not experimentally measurable so we must deal with the smaller difference of the two.

From now on, we will only deal with **φ distribution**. ϑ will not be shown.

We can look at the same graphs but for the comparison of single electron spectra:



For energy distributions, we can see that the greatest difference is seen for low, or very high energies. The lower range is unfortunate for this is not the main interest region for SN. However, at higher energies we expect lower background and this could potentially be a more interesting region.

Statistical analysis of spectral comparison

We employ three methods to compare the spectra. The main goal of the analysis is to answer the question

“how many events do we need to measure to be able to distinguish two distributions from each other”.
We defined this as the variable ***S.***

The three methods used are:

1. **Bin-by-bin comparison**
2. **Kolmogorov-Smirnov Hypothesis test**
3. **ChiSquare Hypothesis test**

Bin-by-bin

We first describe the bin-by-bin method:

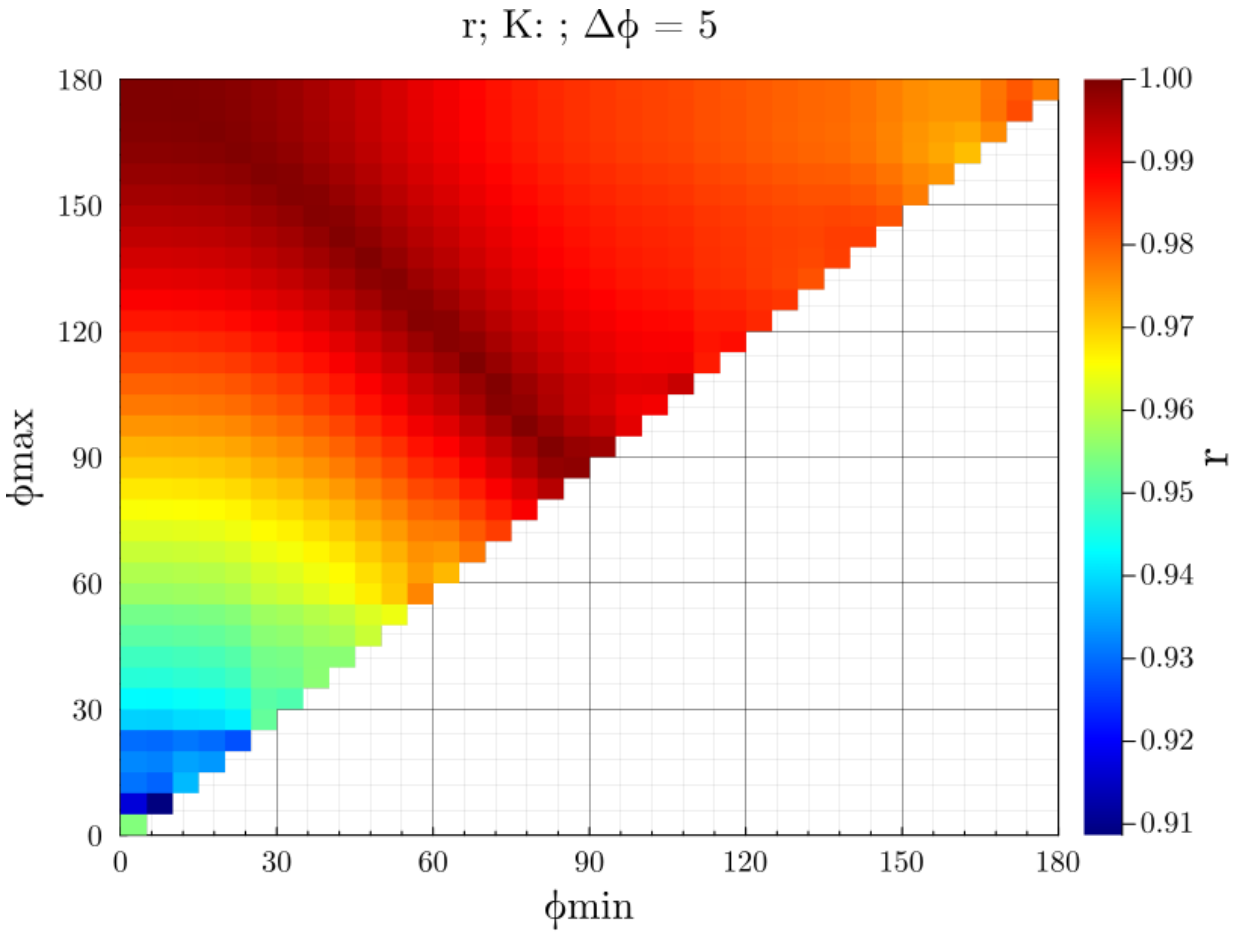
The goal is to quantify the difference between the two angular distributions looking at a single bin (the choice of bin is subject to the analysis). We wish to provide the following answers:

1. How many events (how large a statistics) must be measured to be able to distinguish two spectral shapes within given n_σ confidence?
2. Is it feasible to obtain such statistics within the 5 year data-taking period of SuperNEMO? If not, what would have to be the parameters?
2. What is the sensitivity of SuperNEMO toward the given distribution?

To fulfill the first goal, we have devised a methodology with the aim to find ROI, where the difference is statistically most significant.

The main idea arises from comparing the difference between the two distributions in terms of bin heights ($h_i^j; j \in (1,2)$). First, we introduce the ratio $r_i \equiv \frac{h_i^{j,k}}{h_i^{k,j}}$, where whether j is in the numerator or k is determined dependent on whether j < k or j > k, respectively. Furthermore, to be more *fair* in comparisons, we exchange $h_i^{j,k}$ with $\varepsilon_i^{j,k} \equiv \frac{h_i^{j,k}}{n_{TotalEvents}}$, ie. the normalized bin height. Thus, the ratio now becomes: $r_i = \frac{\varepsilon_i^{j,k}}{\varepsilon_i^{k,j}}$.

To determine ROI where the ratio r_i is most favourable for our purposes, that is the lowest number, we create *maps* of various ranges of φ.
The maps are created by taking some range $\varphi \in (\varphi_{min}, \varphi_{max})$, and calculating the respective r_i.
The results are shown in the figure below. Each square in the figure represents a certain range, to be read out by the upper-left corner of the square. (That is, the range $\varphi \in (0, \Delta\varphi)$ presented by the square in the down-left corner, the very first square.)



In the figure we can see the calculated r_i for each range of phi. The most ideal value is the lowest, ie. r = 0.91 for phi (5, 10).

This however, does not tell the whole story yet. We must consider the uncertainty of the result as well as the uncertainty of the simulation. Furthermore, since the aim was to answer **how many events are required** to distinguish the two spectra, we still have some steps to take.

We will therefore produce a few more figures and calculations.

First, to calculate the number of events required, we begin with the following:

Assume M represents the number of events in the *smaller* bin i and N represents the number of events in the *larger* bin. Then in order to distinguish the two bins from each other at the level of n_σ, their difference must be at least equal (or greater) than the sum of their respective uncertainties.

$$M - N = n_\sigma(\Delta M + \Delta N); \Delta M = \sqrt{M}, \Delta N = \sqrt{N}$$

Dividing the equation by M, substituting $r = \frac{M}{N}$ and rearranging yields:

$\tilde{M}(r) = n_\sigma^2(\frac{r+\sqrt{r}}{1-r})^2$, where $\tilde{M}$ simply represents the minimum number of events in M required to distinguish M from N by n_σ.

Now, using uncertainty propagation, we can find the uncertainty on $\tilde{M}$ as:

$$\Delta\tilde{M} = n_\sigma^2(\frac{r+\sqrt{r}}{1-r})^3 r\sqrt{1/M + 1/N}.$$

The uncertainty in $\tilde{M}$ can be improved by obtaining higher statistics.

However, due to the fact that we do not precisely understand the detector angular correlations, we do not know precisely the analytical value for r either. (If we knew perfectly the correlation $\theta \rightarrow \phi$, we could obtain r analytically from the input angular distributions). We must, therefore, take into account the uncertainty on r.

$$\Delta r = r\sqrt{1/M + 1/N};$$

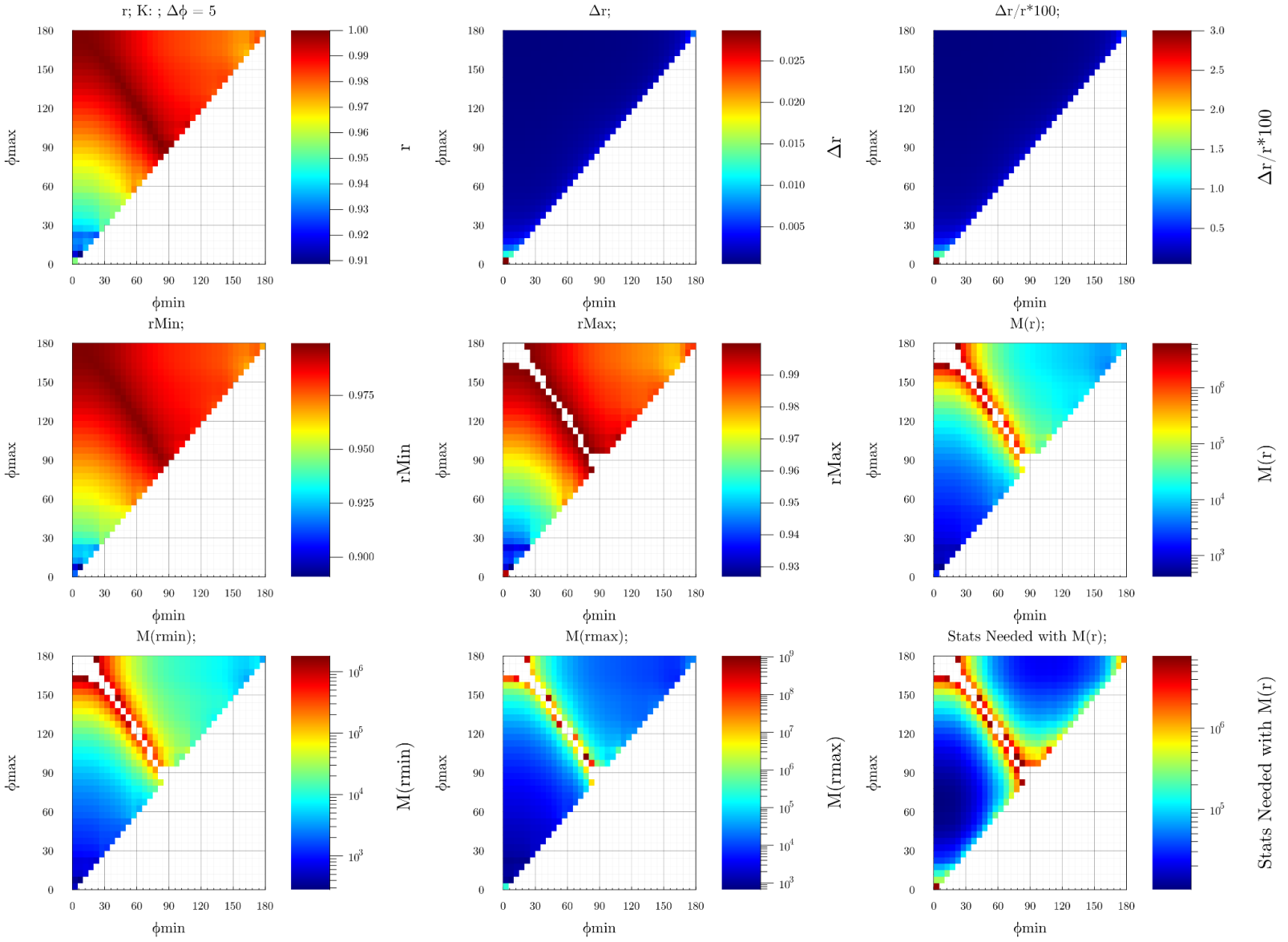
Since the behavior of r is very much non-linear ($\tilde{M}$ rapidly explodes for r close to 1.0) we consider the maximum (and minimum) uncertainties:

$$r_{max} = \frac{M+\Delta M}{N-\Delta N}\ r_{min} = \frac{M-\Delta M}{N-\Delta N}.$$

Here, the worst case scenario (r_{max}) can in some cases exceed 1.0, (ie. $r > 1.0$). Such regions will be removed from the analysis and will show in the maps as empty squares.

To answer the question *how many events SN needs to measure to distinguish* we can convert the $\tilde{M}(r)$ into the number of events needed S as follows: $S = \tilde{M}(r)/\varepsilon$. That is, we scale the number of needed events in the bin by the proportion of the total events that bin represents. S then gives the total statistics needed to obtain desired $\tilde{M}(r)$.

Finally, let us look at all of the mentioned values. In the figure below we show maps for: 1. r, 2.Δr, 3. $\Delta r/r * 100$, 4. r_{min}, 5. r_{max}, 6. $\tilde{M}(r)$, 7. $\tilde{M}(r_{min})$, and 8. $\tilde{M}(r_{max})$, 9. $S(r)$.



The optimal values are:

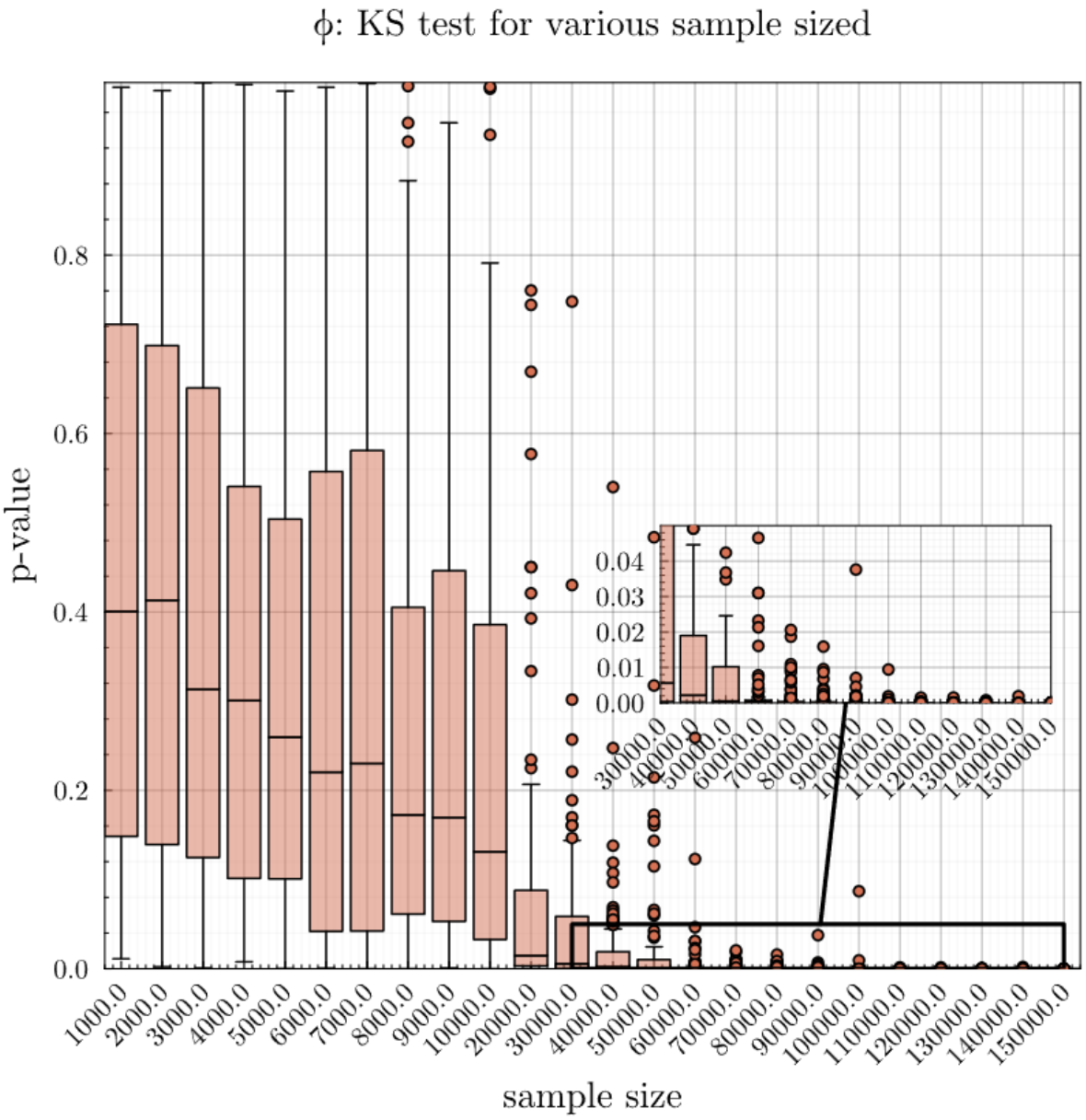
- r = 0.91 @ (5, 10)°
- δr = 0.05 @ (0, 180)°
- $M(r)$ = 415.42 @ (5, 10)°
- $M(r_{min})$ = 282.26 @ (5, 10)°
- $M(r_{max})$ = 667.48 @ (5, 10)°
- $S(r)$ = 10206.53 @ (0, 70)°

This is as far as we got so far. Conclusions are not yet made. Further analysis will be done later. Question is what combination is best for S and the uncertainties, combined with how well the detector performs for given angular region.

Kolmogorov-Smirnov (KS) and ChiSquared hypothesis tests

In the following text analysis using KS and ChiSquared hypothesis tests is performed. The methodology is as follows:
1. The spectra are split into N random subsets of size M
2. For each subset a KS and ChiSquare hypothesis test is performed and p-value is extracted.
3. For various values of CL (i.e. 90%, 95%) the corresponding S is found as the value for which M crosses the CL.

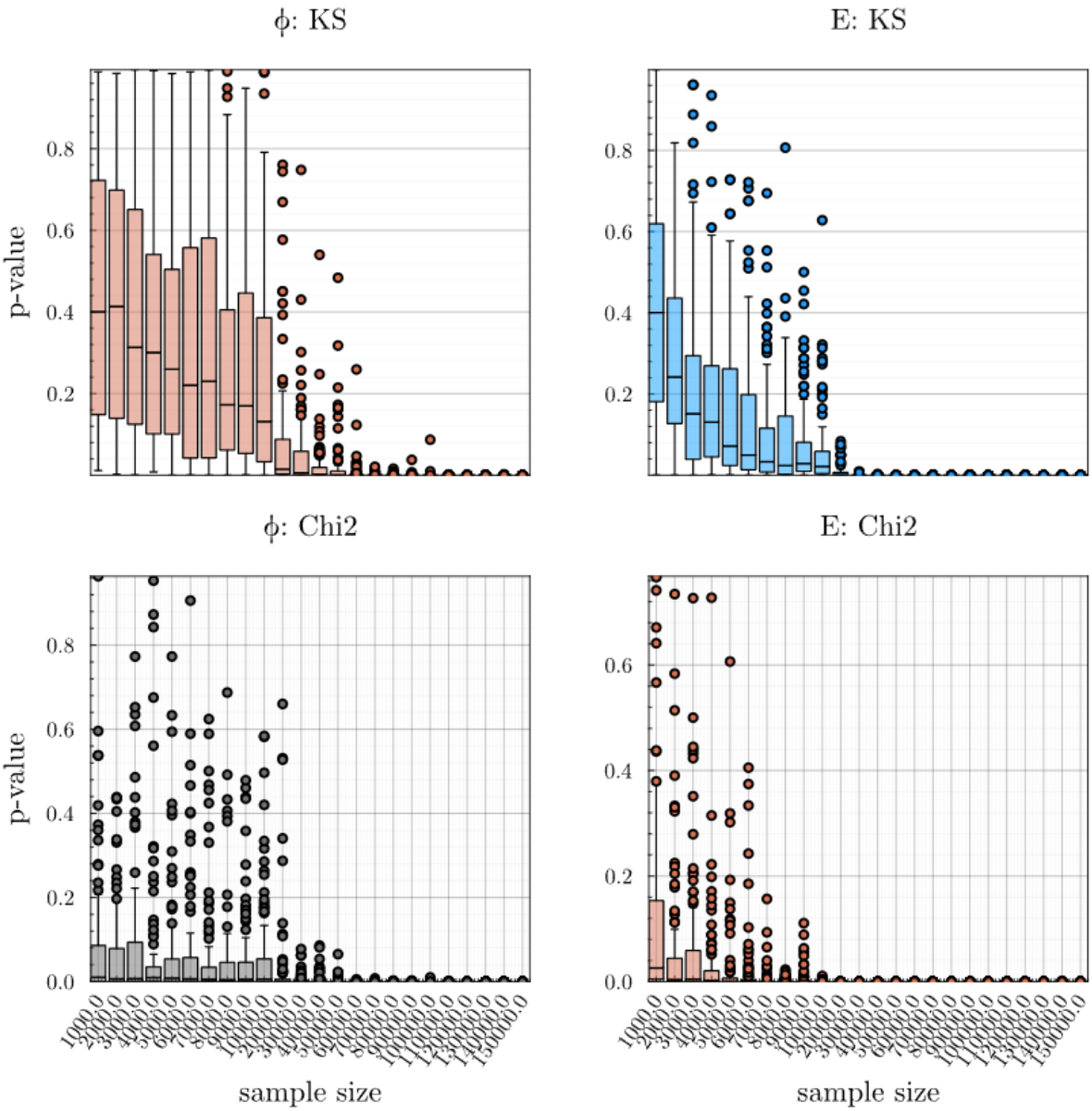
First to get a good look at how the p-values are distributed throughout the N subsets of size M we look at a boxplot of KS test for phi distribution.



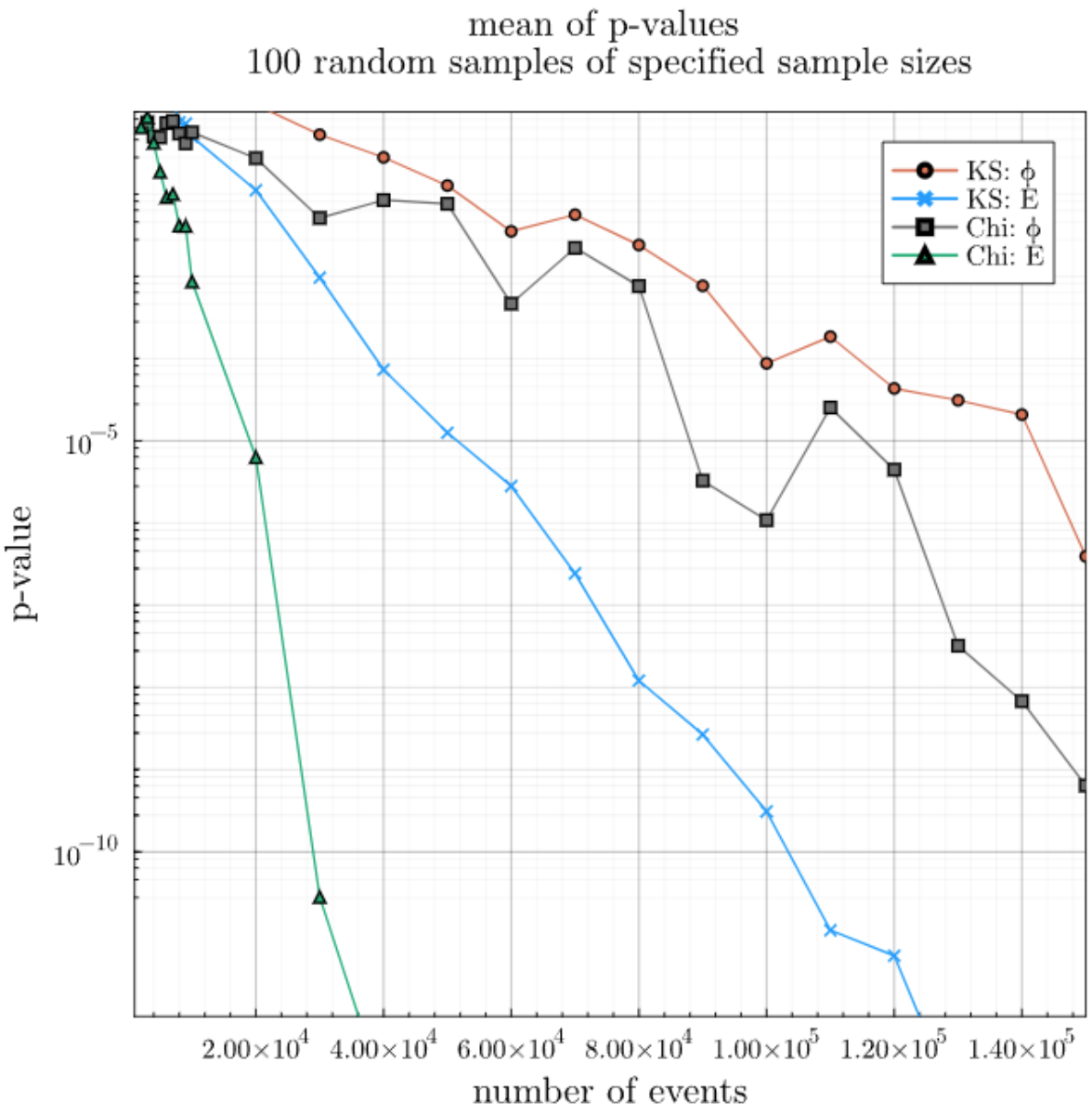
We can see that for smaller sample sizes the p-values vary greatly. Whereas for larger sample sizes, i.e. N>100k events it is very close to zero (which is desired).

We can also look at the other tests:
1. ϕ distribution using KS test
2. Single electron spectra (E) using KS test
3. ϕ distribution using Chi2 test
3. E distribution using Chi2 test

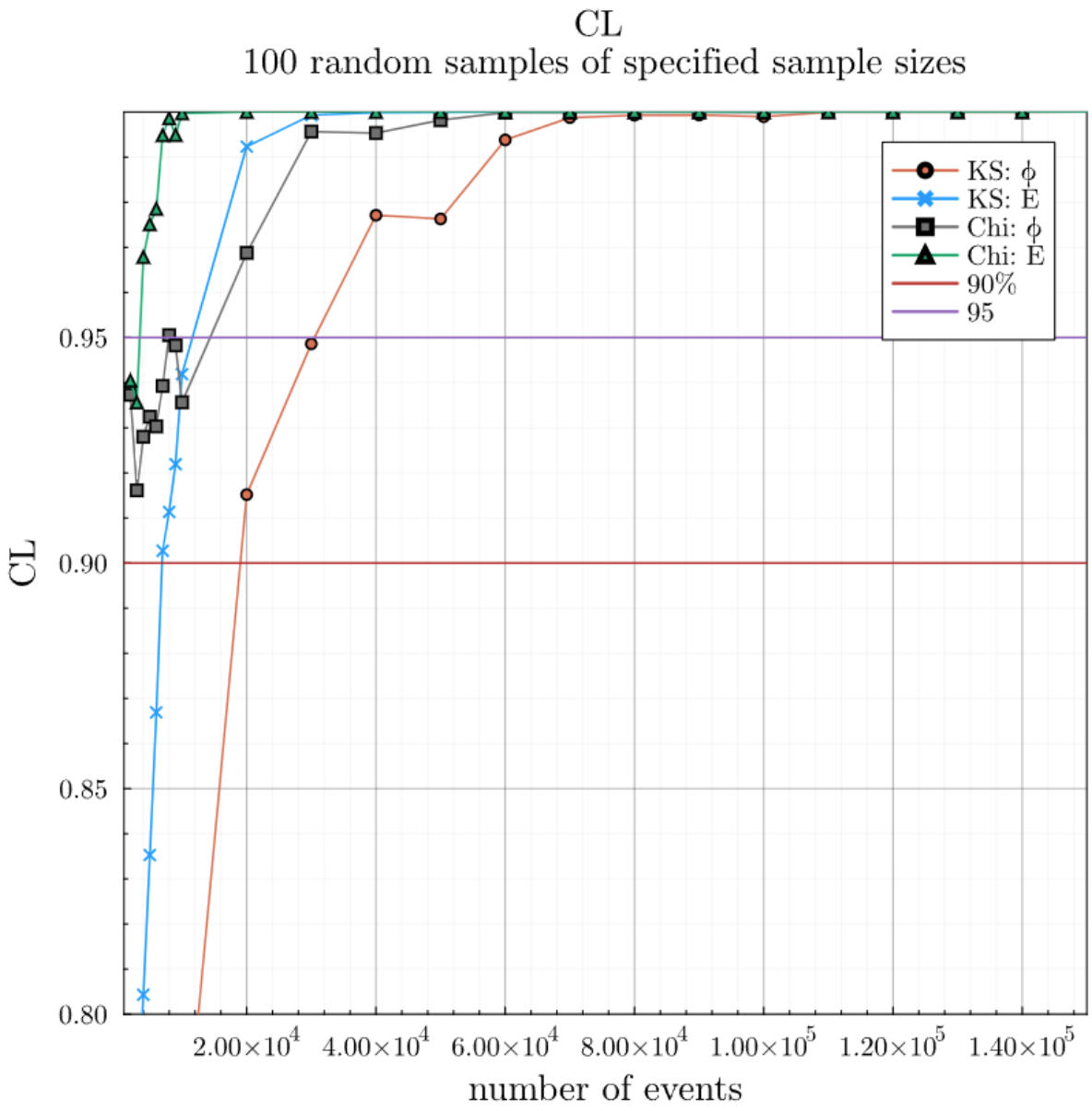
Right away we can see that KS test is more strict when it comes to rejecting H0.



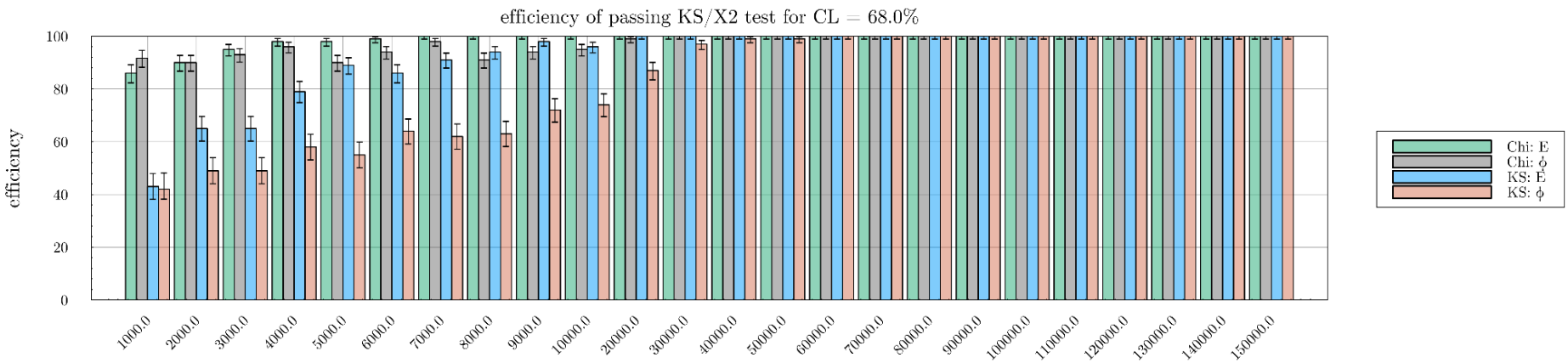
To better see the trend of the tests when compared to various sample sizes we can look at how the mean of each subset changes with changing the sample size.

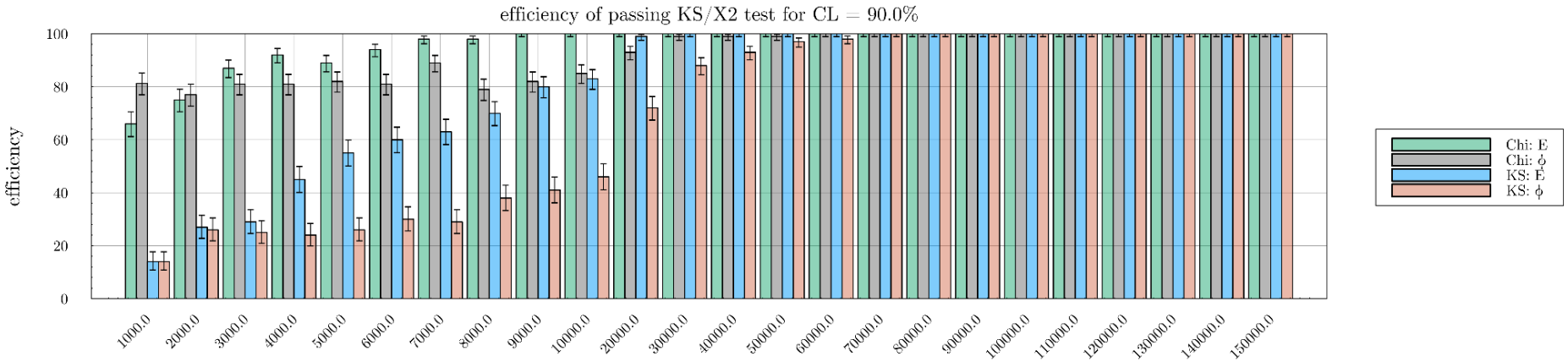


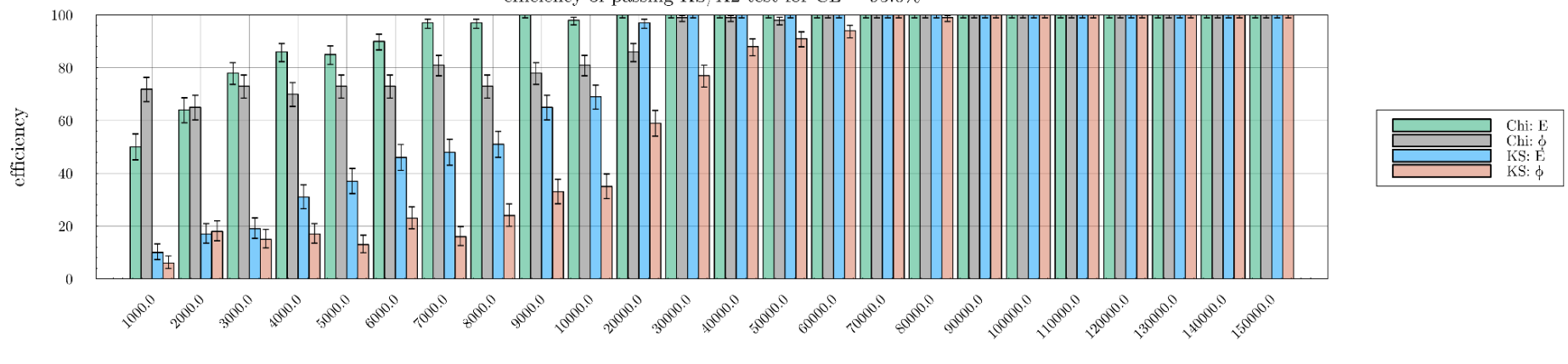
Lastly, to obtain our desired *S*, we plot values for (1-p) along with lines for 90% and 95% confidence level. In order to determine ***S*** we simply look for the number M at which the value (1-p)>CL.



However, since it can be seen that the p-values vary greatly, we can perform an additional analysis by looking at the efficiency of rejecting H0 for various sample sizes. That is, for each subset of size M we compute efficiency as $\varepsilon = \frac{M_{reject}}{M_{tot}}$.


efficiency of passing KS/X2 test for CL = 68.0%
efficiency
0
20
40
60
80
100
1000.0
2000.0
3000.0
4000.0
5000.0
6000.0
7000.0
8000.0
9000.0
10000.0
20000.0
30000.0
40000.0
50000.0
60000.0
70000.0
80000.0
90000.0
100000.0
110000.0
120000.0
130000.0
140000.0
150000.0
Chi: E
Chi: ϕ
KS: E
KS: ϕ


efficiency of passing KS/X2 test for CL = 90.0%
efficiency
0
20
40
60
80
100
1000.0
2000.0
3000.0
4000.0
5000.0
6000.0
7000.0
8000.0
9000.0
10000.0
20000.0
30000.0
40000.0
50000.0
60000.0
70000.0
80000.0
90000.0
100000.0
110000.0
120000.0
130000.0
140000.0
150000.0
Chi: E
Chi: ϕ
KS: E
KS: ϕ


efficiency of passing KS/X2 test for CL = 95.0%
efficiency
0
20
40
60
80
100
1000.0
2000.0
3000.0
4000.0
5000.0
6000.0
7000.0
8000.0
9000.0
10000.0
20000.0
30000.0
40000.0
50000.0
60000.0
70000.0
80000.0
90000.0
100000.0
110000.0
120000.0
130000.0
140000.0
150000.0
Chi: E
Chi: ϕ
KS: E
KS: ϕ

(The uncertainties were calculated according to the efficiency uncertainty rules described in M. **Paterno, "Calculating efficiencies and their uncertainties", Tech. Rep. FERMILAB-TM-2286-CD, 2004**)

From the efficiency figure we can **estimate** the desired number ***S*** by looking at each subset/test combination where it goes to 100% efficiency (eg. arbitrary 3+ times in a row).

For example for the phi distribution we could take the following numbers as rough estimates of ***S***.

CL\method	KS	X^2	bin-by-bin
68%	40k	20k	10k
90%	80k	40k	28k
95%	80k	70k	39k
99%	100k	80k	68k