



MAX-PLANCK-GESELLSCHAFT



Max-Planck-Institut für Physik
(Werner-Heisenberg-Institut)

Signal Discovery in Sparse Spectra

F. Beaujean¹, A. Caldwell¹, D. Kollar², K. Kröniger³, S. Pashapour³

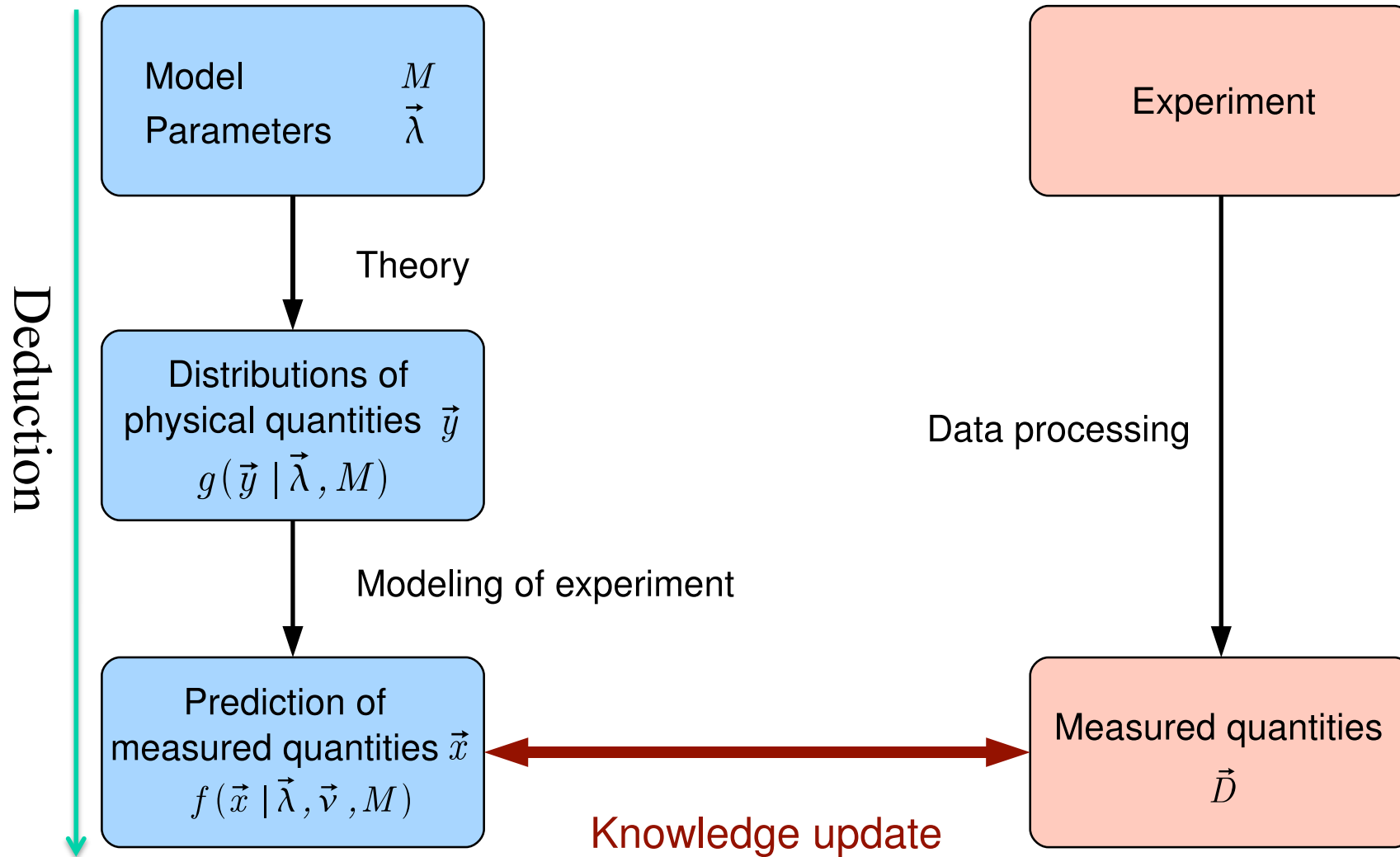
1-Max Planck Institute for Physics

2-CERN, now at AREVA

3-II. Physikalisches Institut, Universität Göttingen

1. Logical framework for data analysis
2. Discovery in sparse spectra
3. Discussion

How we learn



Logical Basis

Model building and making predictions from models follows deductive reasoning:

Given $A \rightarrow B$ (with some frequency)

Given $B \rightarrow C$ (with some frequency)

Then, given A you can conclude that C is possible with some probability (frequency)

etc.

Everything is clear, we can make frequency distributions of possible outcomes within the model, etc. **This is math**, so it is correct ...

Logical Basis

However, **in physics** what we want to know is the validity of the model given the data. i.e., logic of the form:

Given $A \rightarrow B$ (with some frequency)

Given $B \rightarrow C$ (with some frequency)

Measure C, what can we say about A? Well, can say A is a possibly correct model. What else? Need to know about other models

maybe $A_1 \rightarrow C, A_2 \rightarrow C, \dots$

We now need inductive logic to decide how much we want to believe each possible model. We can never say anything absolutely conclusive about A unless we can guarantee a complete set of alternatives A_i and only one of them can give outcome C. This does not happen in science, so **we can never say we found the true model.**

Logical basis

Purpose of science is to increase **knowledge**, where

Knowledge = **justified ~~true~~ belief**

Justification comes from the data.

Start with some knowledge or maybe plain belief

Data analysis gives updated knowledge. Need Bayes' Theorem to make coherent statements on what we believe.

nb: if don't use Bayes' theorem, belief update becomes maximally subjective.

Example: Double Beta Decay

One of the outstanding questions in Particle Physics is whether the neutrino is its own antiparticle (so-called Majorana particle).

The only practical way which has been found to search for the Majorana nature of neutrinos (particle same as antiparticle) is double beta decay (because of the light mass of neutrinos, helicity flip is very unlikely unless the neutrinos have very low energy).

For us, what is interesting is that we are looking for a peak at a well-defined energy in a sparse spectrum.

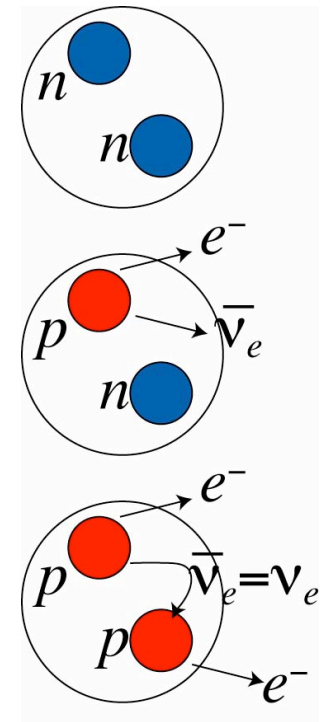
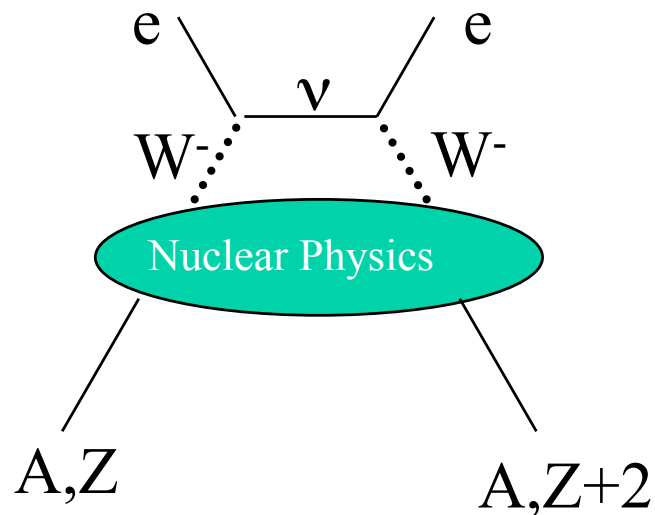
A. Caldwell, K. Kröninger, Phys. Rev. D 74 (2006) 092003

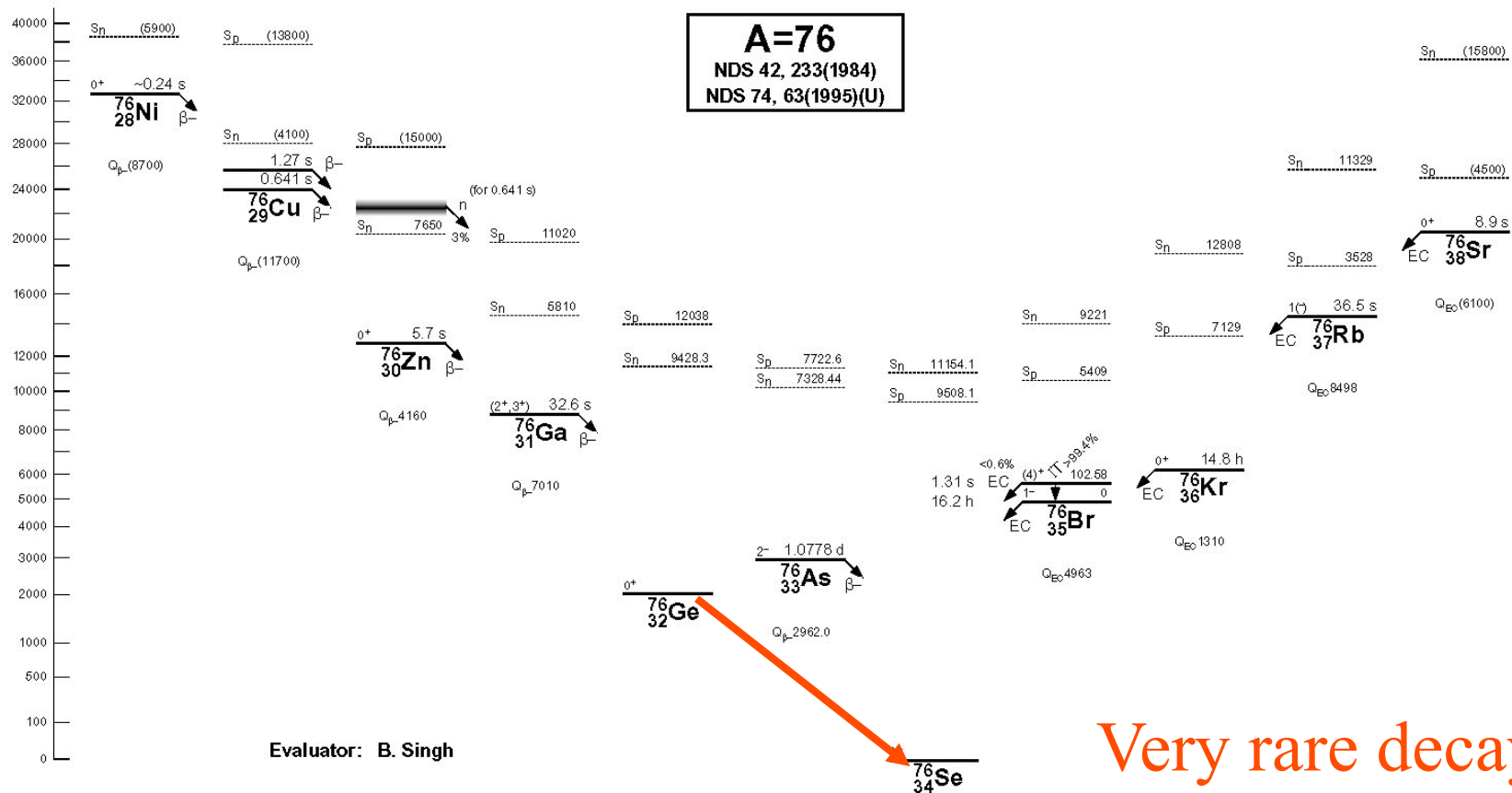
How can we test if neutrinos are Dirac or Majorana particles ?

Experimental Problem:

$$P(\nu_L \rightarrow \nu_R) \propto \left(\frac{m_\nu}{E_\nu} \right)^2 \quad m \leq \text{eV}, E \text{ MeV or more}$$

Only known technique is **neutrinoless double beta decay**:





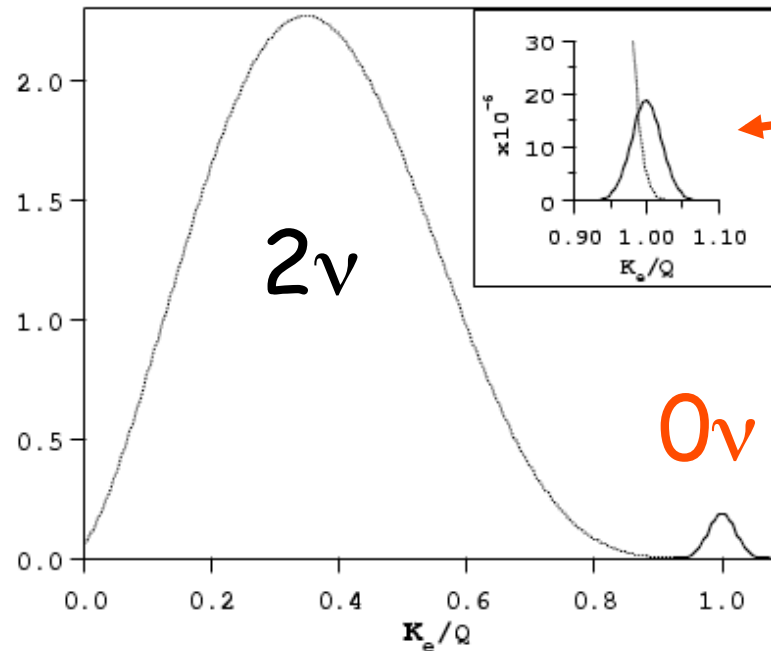
Very rare decay
lifetimes $>10^{20}$
years !

$(A,Z) \rightarrow (A,Z+1)+e+\nu$ energetically forbidden

$(A,Z) \rightarrow (A,Z+2)+2e+2\nu$ is allowed.

Then, for Majorana particle $(A,Z) \rightarrow (A,Z+2)+2e$ possible

Normalized energy spectrum



If resolution poor

If resolution good

0ν -DBD rate

Phase space $\propto Q^5$

Nuclear matrix element

Effective Majorana mass

$$1/\tau = G(Q,Z) |M_{\text{nucl}}|^2 \langle m_{ee} \rangle^2$$

$$1/\tau = G(Q,Z) |M_{\text{nucl}}|^2 \langle m_{ee} \rangle^2$$

Some numbers:

$$G(Q,Z) \sim 10^{-25} (\text{yr ev}^2)^{-1} \quad M_{\text{nucl}} \sim 2-3$$

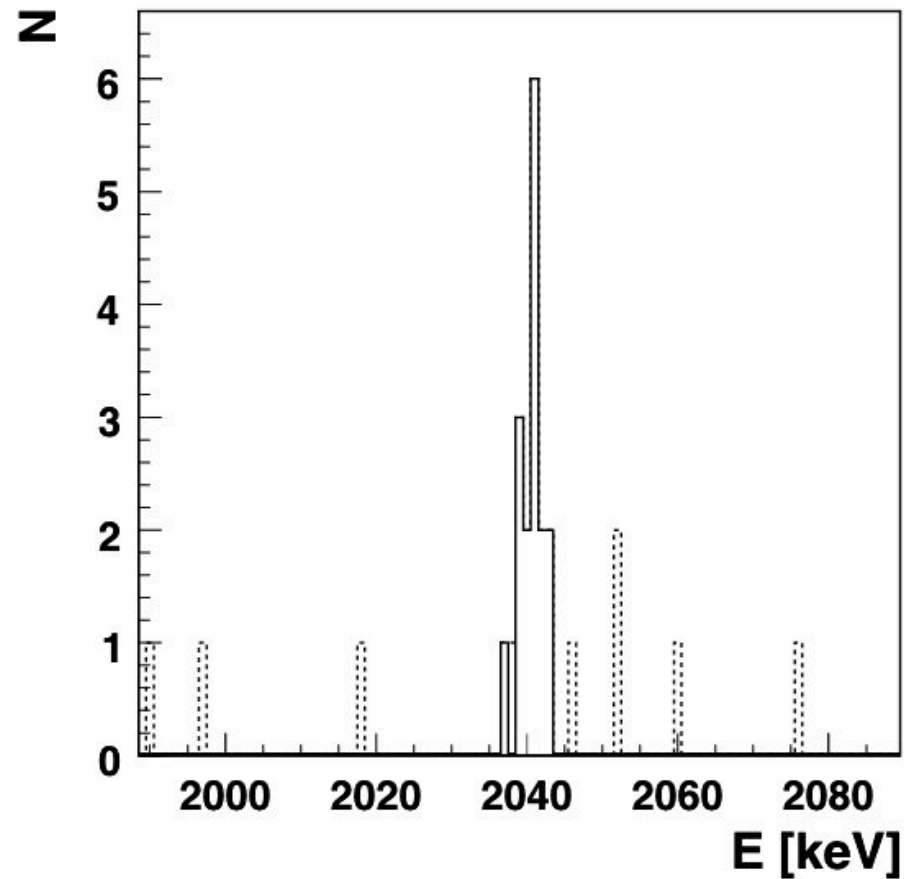
So $\tau \sim 1 \cdot 10^{25}$ yrs for $\langle m_{ee} \rangle = 100$ meV

The **chance for an atom to have decayed** via neutrinoless double beta decay **since the Big-Bang is 10^{-16}** . This is a **RARE decay**.

Conclusion: need 1000's of mole-years of exposure for sensitivity at 100 meV level. I.e., many kg of material watched over many years.

And, the backgrounds must be extremely low !!!

Discovery or not ?



Analyze energy spectrum and decide if there is evidence for a signal.
Counting experiment – Poisson statistics.

Sparse Spectra

Define the proposition:

H = The observed spectrum is due to background only

If $p(H|\text{spectrum}) < \text{cut}$, can claim ‘evidence’ for something beyond background. If we assume that **what is not background is signal**, then we claim evidence for the signal.

E.g.:

$p(H|\text{spectrum}) < 0.01$, ‘evidence’ (better >99% belief in ‘new physics’)

$p(H|\text{spectrum}) < 0.0001$, ‘discovery’ (better >99.99% belief in ‘new physics’) (very stringent, DoB contains our belief in the new physics)

Note: intended to be the real ‘degree-of-belief’.

Sparse Spectra

What we know how to calculate:

$p(\text{spectrum}|H)$ - the probability to observe the spectrum given H
(We assume Poisson statistics are valid)

How do we go from $p(\text{spectrum}|H)$ to $p(H|\text{spectrum})$?

Certainly $p(A|B) \neq p(B|A)$

(e.g., 1% probability of signal assuming Standard Model does not mean Standard Model ruled out with 99% certainty)

Need to use Bayes' theorem to reach a conclusion

$$p(H|\text{spectrum}) = \frac{p(\text{spectrum}|H)p_0(H)}{p(\text{spectrum})}$$

Double Beta Decay Example

$p_0(H)$ is the prior belief in H (before we do the experiment). It is a critical part of the Bayesian analysis. Our posterior belief in the truthfulness of H always depends on prior beliefs. E.g.,

The existing limits are $T_{1/2} > 4 \cdot 10^{25}$ yr; a positive claim for a signal exists at the level $T_{1/2} = 1.2 \cdot 10^{25}$ yr; my favorite theorist believes strongly that neutrinos are Majorana particles, but he won't tell me the neutrino mass; the theorist at a neighboring university says that he believes strongly in Leptogenesis, and in that context the neutrino is a Majorana particle but it must be very light, such that neutrinoless double beta decay is unobservable,...

What is $p(\text{spectrum})$? Expand (law of total probability)

$$p(\text{spectrum}) = p(\text{spectrum} \mid H)p(H) + p(\text{spectrum} \mid \bar{H})p(\bar{H})$$

DBD example

We need also the probability of the negation of H. In our case, we assume knowledge concerning the background, so

\bar{H} = The spectrum is due to background + signal (neutrinoless double beta decay).

I.e., we assume backgrounds are known up to normalization and some smoothly varying shape, and the only possibility other than known background is signal from neutrinoless double beta decay.

$$p(H \mid \text{spectrum}) + p(\bar{H} \mid \text{spectrum}) = 1$$

so

$$p(H \mid \text{spectrum}) = \frac{p(\text{spectrum} \mid H)p_0(H)}{p(\text{spectrum} \mid H)p_0(H) + p(\text{spectrum} \mid \bar{H})p_0(\bar{H})}$$
$$p(\bar{H} \mid \text{spectrum}) = \frac{p(\text{spectrum} \mid \bar{H})p_0(\bar{H})}{p(\text{spectrum} \mid H)p_0(H) + p(\text{spectrum} \mid \bar{H})p_0(\bar{H})}$$

DBD example

Now we know how to perform all calculations:

$$p(\text{spectrum} | H) = \int p(\text{spectrum} | B) p_0(B) dB$$

$$p(\text{spectrum} | \bar{H}) = \int p(\text{spectrum} | S, B) p_0(S) p_0(B) dB$$

Where B is the expected number of background events and S is the expected number of signal events. These quantities come with their own priors.

n_i = observed number of events in bin i

λ_i = expected number of events in bin i

$$\lambda_i = S \int_{\Delta E_i} f_S(E) dE + B \int_{\Delta E_i} f_B(E) dE$$

Where f_S and f_B are the normalized signal and background probability densities as functions of energy.

DBD example

then

$$p(\text{spectrum} \mid B) = \prod_{i=1}^N \frac{\lambda_i(0, B)^{n_i}}{n_i!} e^{-\lambda_i(0, B)}$$

$$p(\text{spectrum} \mid S, B) = \prod_{i=1}^N \frac{\lambda_i(S, B)^{n_i}}{n_i!} e^{-\lambda_i(S, B)}$$

To determine parameter values or set limits, we need

$$p(S, B \mid \text{spectrum}) = \frac{p(\text{spectrum} \mid S, B) p_0(S) p_0(B)}{\int p(\text{spectrum} \mid S, B) p_0(S) p_0(B) dS dB}$$

and then marginalize

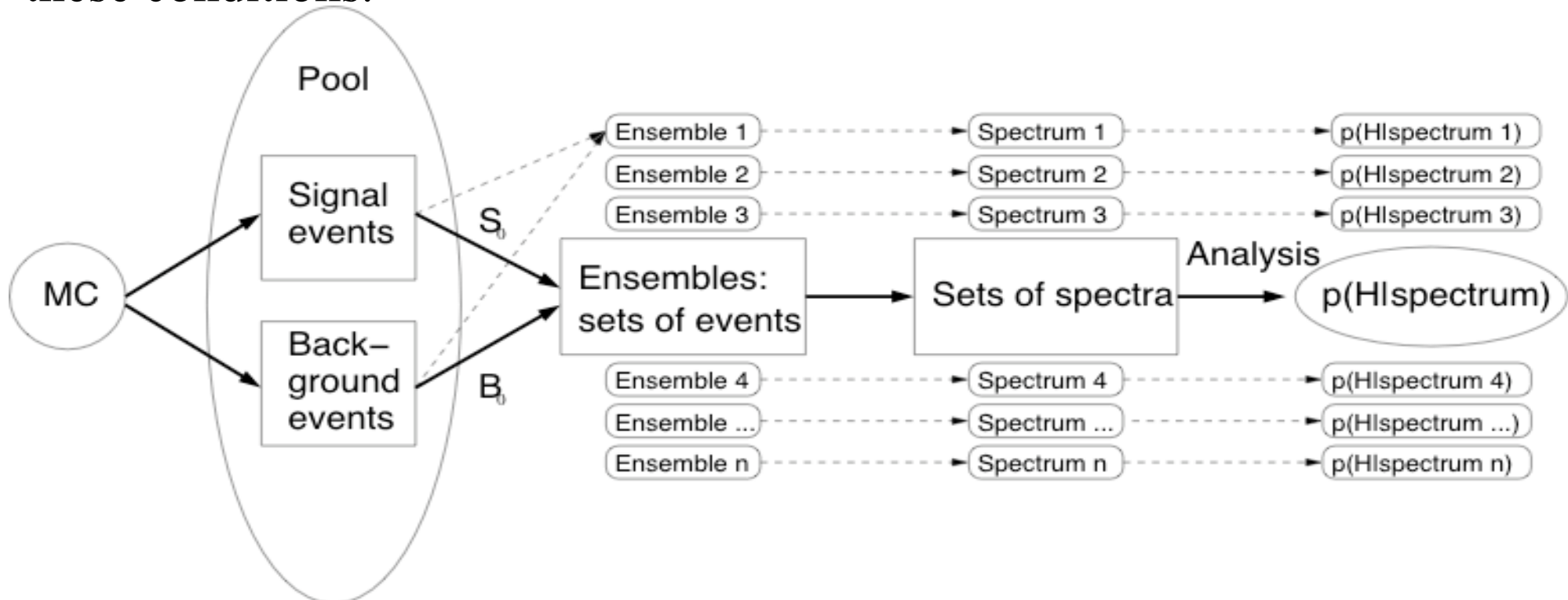
$$p(S \mid \text{spectrum}) = \int p(S, B \mid \text{spectrum}) dB$$

e.g., 90% probability upper limit, S_{90} from solving

$$\int_0^{S_{90}} p(S \mid \text{spectrum}) dS = 0.90$$

DBD example

So we know how to calculate probabilities given an experimental outcome. What do we do to check the sensitivity of the experiment ? We generate ensembles of possible experimental results, which will depend on particular choices of background and signal, B_0 and S_0 . Then we can make distributions of the probabilities which could result under these conditions.



GERDA example

Assumptions for GERDA:

$$p_0(H) = p_0(\bar{H}) = 1/2$$

$$p_0(S) = \frac{1}{S_{\max}} \quad 0 \leq S \leq S_{\max} \quad p_0(S) = 0 \text{ otherwise}$$

$$p_0(B) = \frac{e^{-\frac{(B-\mu_B)^2}{2\sigma_B^2}}}{\int_0^{\infty} e^{-\frac{(B-\mu_B)^2}{2\sigma_B^2}} dB} \quad B \geq 0; \quad p_0(B) = 0 \quad B < 0$$

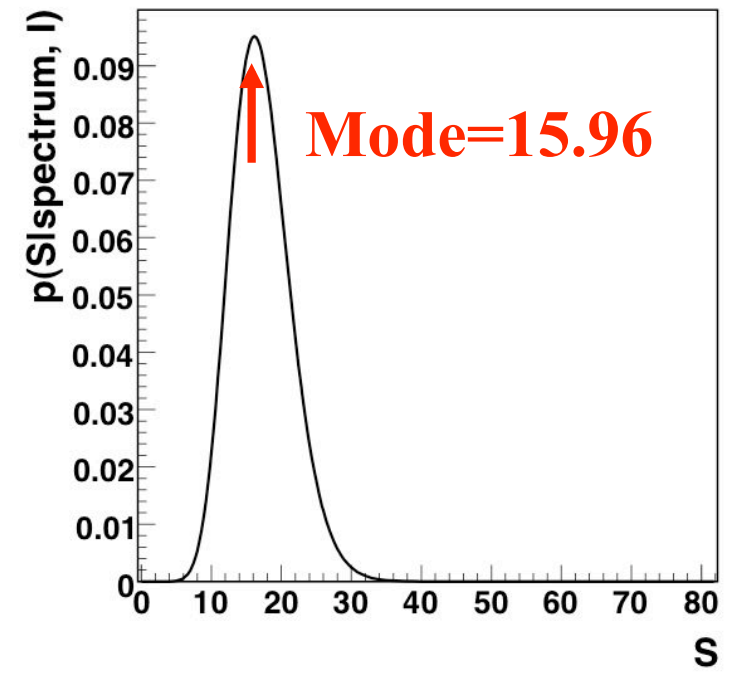
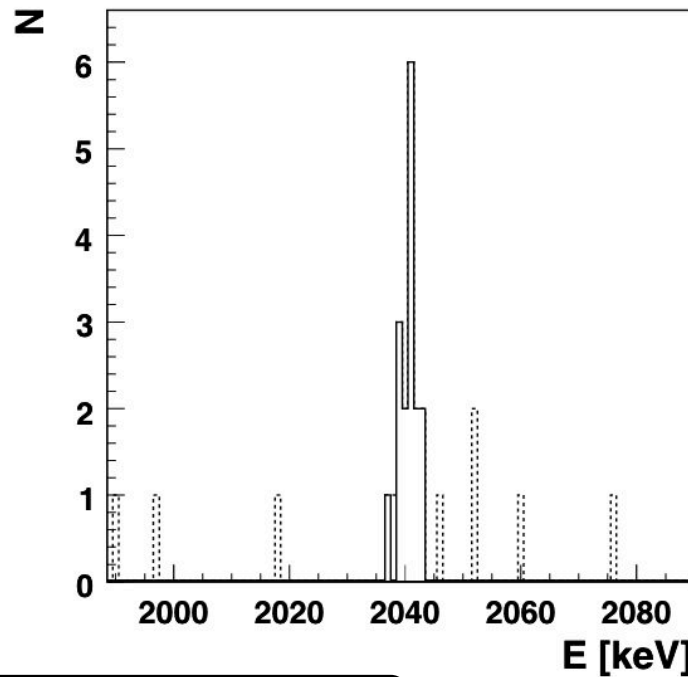
S_{\max} was calculated assuming $T_{1/2} = 0.5 \cdot 10^{25}$ yr

$$\mu_B = B_0, \quad \sigma_B = B_0/2$$

100 keV window analyzed. B_0 total background in this window.

Example:

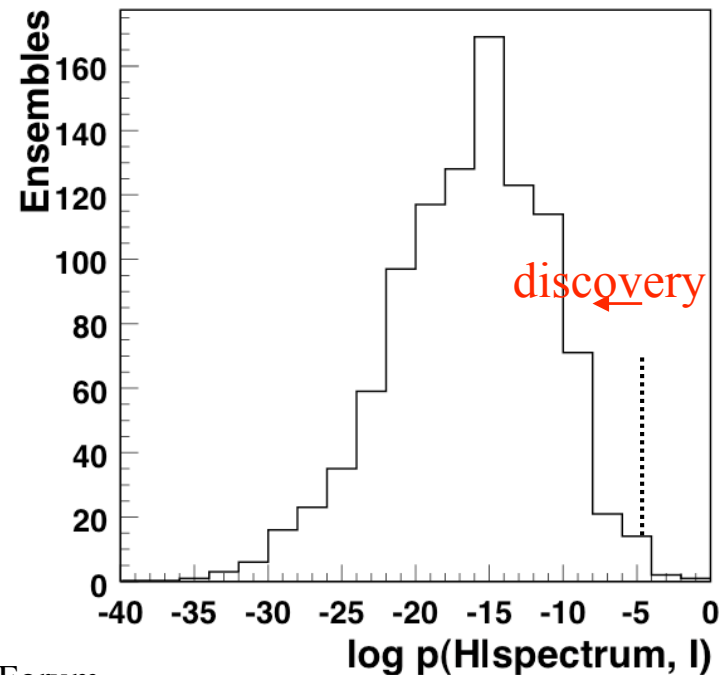
$$S_{\text{true}}=16, B_{\text{true}}=9$$



$$p(H | \text{spectrum}) = 2.2 \cdot 10^{-12}$$

1000 experiments simulated with
 $T_{1/2}=2 \cdot 10^{25}$ yr, $10^{-3}/(\text{kg keV yr})$
Exposure 100 kg-yr

About 95% chance a discovery
could be claimed

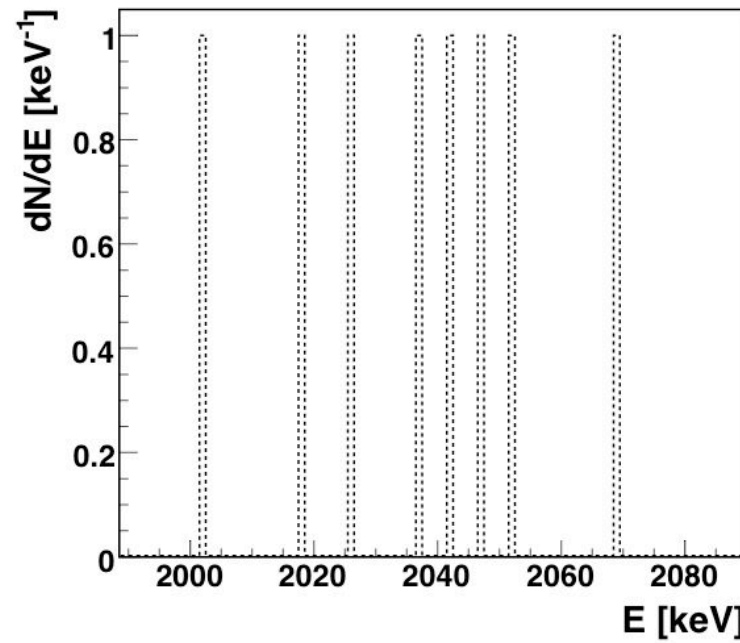


Example:

$$S_{\text{true}}=0, B_{\text{true}}=8$$

$$10^{-3}/(\text{kg keV yr})$$

Exposure 100 kg-yr

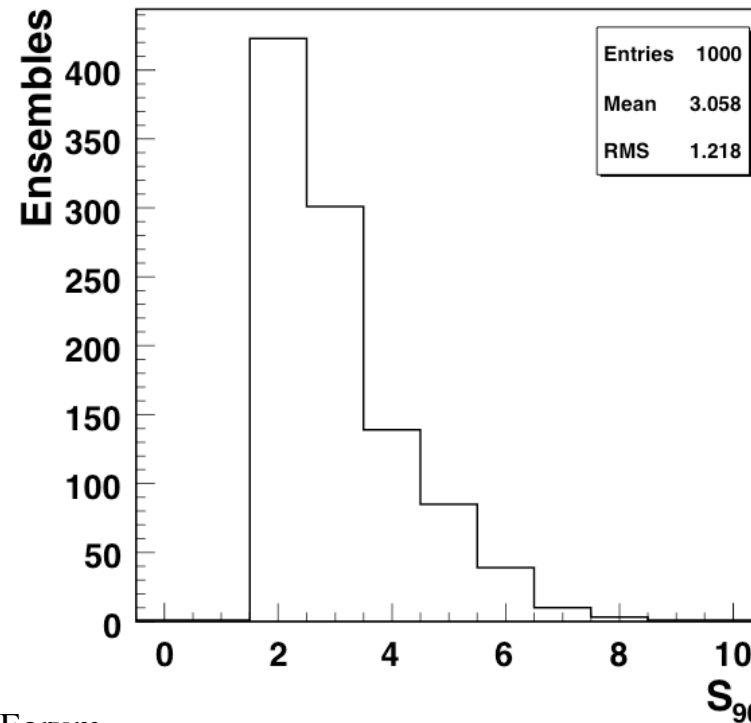


$$p(H | \text{spectrum}) = 0.93$$

$$S_{90} = 3.99$$

1000 experiments simulated

0 false claims of a discovery



GERDA example

To translate the event numbers into lifetimes, we use

$$S = \ln 2 \cdot \kappa \cdot M \cdot \varepsilon_{sig} \cdot \frac{N_A}{M_A} \cdot \frac{T}{T_{1/2}}$$

Where:

N_A is Avogadro's number

M_A is the atomic mass of ^{enr}Ge

M is the total mass of Germanium

κ is the enrichment factor (by atom, 0.86 used)

ε_{sig} is the signal efficiency (taken to be 87%)

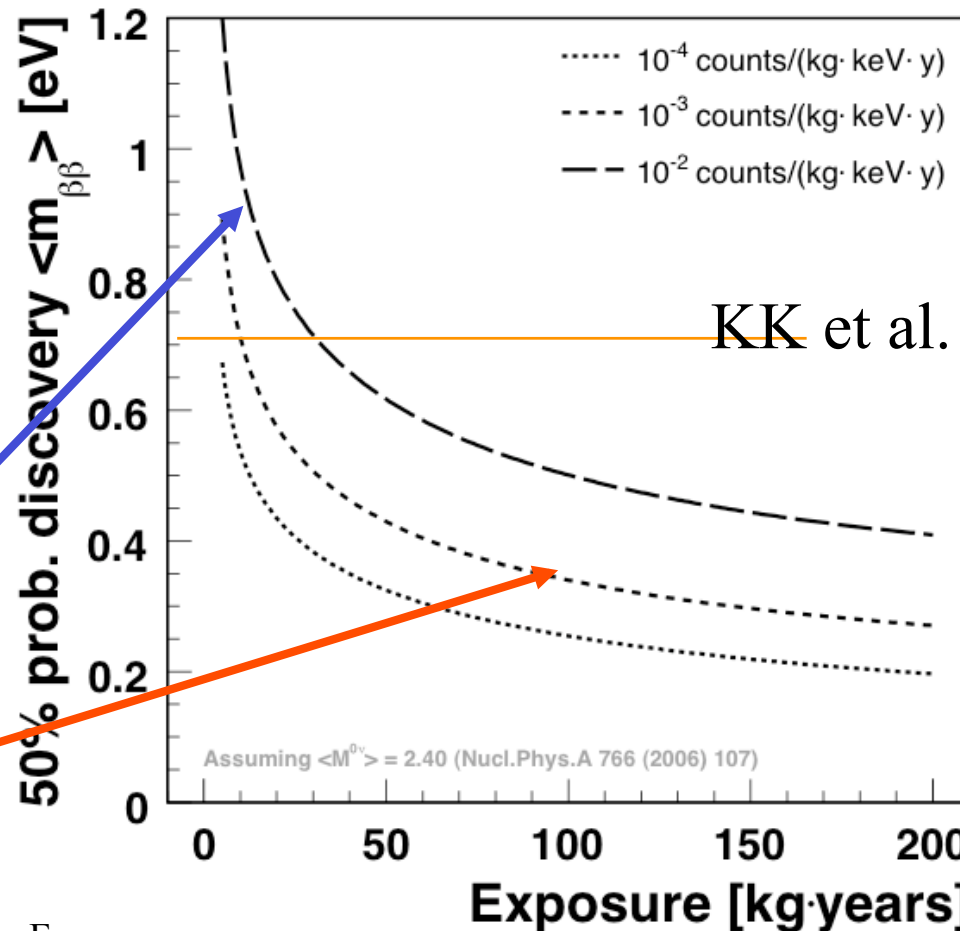
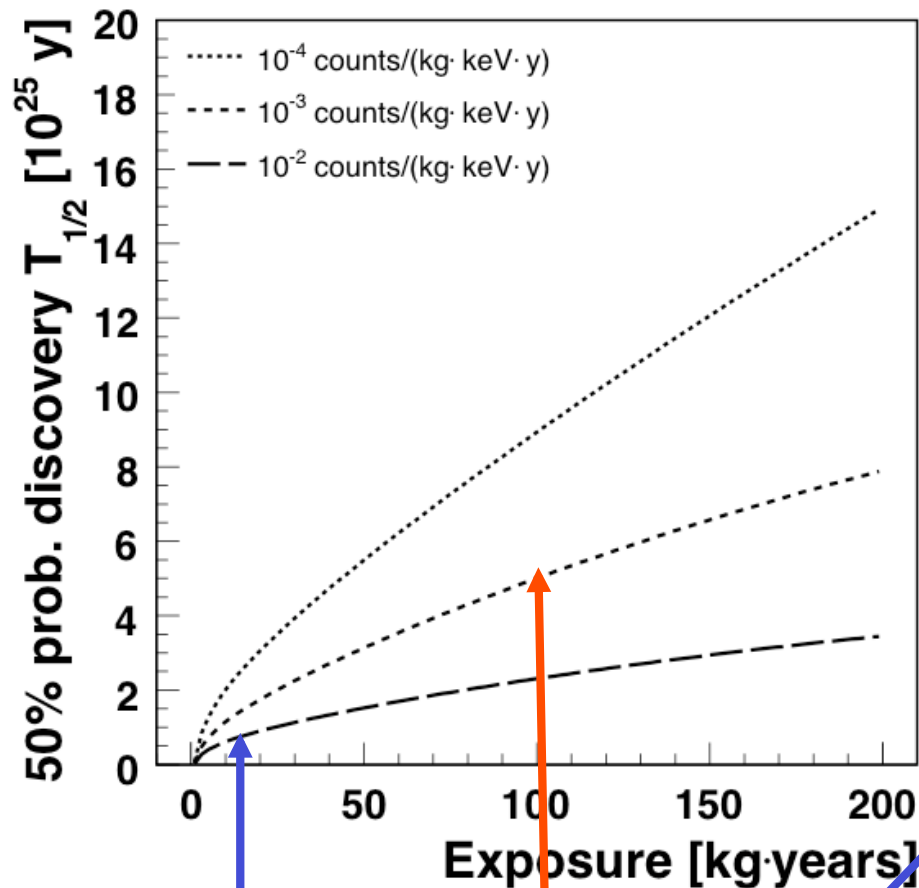
To translate $T_{1/2}$ to a mass

$$\langle m_\nu \rangle = \left(T_{1/2} G^{0\nu} \right)^{-1/2} \cdot \frac{1}{M_{0\nu}}$$

$G^{0\nu}$ and $T_{1/2}$ from Rodin, Faessler, Simkovic, Vogel nucl-th/0503063

GERDA example

Bayesian analysis: discovery defined as $P(\text{background only}|\text{spectrum}) < 0.0001$



Phase I: 15 kg-yr, existing ^{76}Ge crystals

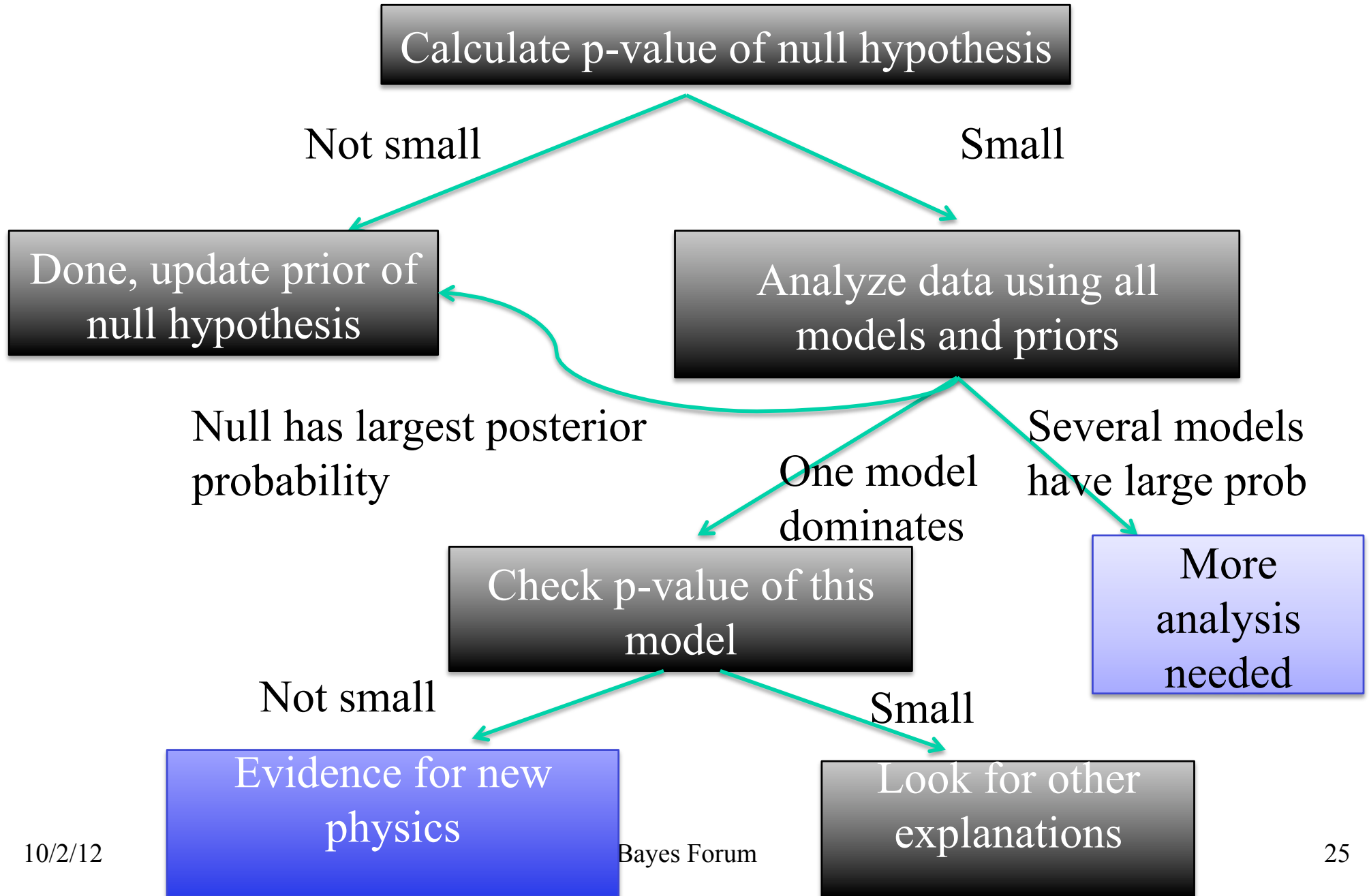
Phase II: 100 kg-yr, new segmented ^{76}Ge crystals

Discussion

The example chosen was special in the sense that we could define a complete set of models, and therefore produce a coherent probability analysis. Prior beliefs could be defined, and the knowledge update from an experiment was clear. There will always be some discussion on which prior beliefs to choose.

What if we don't have a complete set of models? No real probability analysis possible, but we still update our degree-of-belief.

Incomplete set of models



Consensus Priors

Should define informative priors whenever possible. In cases where range important, can define 2 or more consensus priors to show us what we learn from the data.

Different new physics obviously have different priors (e.g., compare Higgs at 120 GeV vs new extra dimensions, or, historically, top quark at 175 GeV vs leptoquarks at HERA).

Suggestion: form of a particle physics committee which defines consensus priors. Could be more than one prior per process (pessimistic, optimistic). Could cover all particle physics topics (dark matter, neutrino mass, Majorana, Higgs, SUSY, ...)

Having consensus priors would allow for a transparent and coherent knowledge update. Consensus priors should be updated as new data becomes available.

Look Elsewhere Effect

In the double beta example, we know where to look in energy. If we scanned the energy spectrum and looked for an excess, would we get an enhancement of false discoveries (Look Elsewhere Effect) ?

the look elsewhere effect is suppressed in Bayesian analysis by a factor:

$$\sigma / \Delta$$

Where σ is the width of the expected signal and Δ is the range over which we search.

LEE

Example: 1D spectrum, search for signal anywhere in spectrum. Assume amplitude and width of new signal fixed.

$$p(H_2|D) = \frac{\int P(D|H_2, \mu)P_0(H_2, \mu)d\mu}{\int P(D|H_2, \mu)P_0(H_2, \mu)d\mu + P(D|H_1)P_0(H_1)}$$

Assume we can take: $P_0(H_2, \mu) = P_0(H_2)P_0(\mu|H_2)$


$$P_0(H_2) = P_0(H_1) = 1/2$$

$$P_0(\mu|H_2) = \frac{1}{L_\mu}$$

$$p(H_2|D) = \frac{\int P(D|H_2, \mu)d\mu}{\int P(D|H_2, \mu)d\mu + L_\mu P(D|H_1)}$$

LEE

Can write: $\int P(D|H_2, \mu) d\mu = P(D|H_2, \mu^*) \delta_\mu$

 Mode value
for posterior

so
$$p(H_2|D) = \frac{P(D|H_2, \mu^*) \delta_\mu}{P(D|H_2, \mu^*) \delta_\mu + L_\mu P(D|H_1)}$$

Degree of belief in new physics hypothesis limited by $\frac{\delta_\mu}{L_\mu} \propto \frac{\sigma}{L_\mu}$

Summary

1. Bayesian framework is natural for quantifying scientific knowledge
2. In some cases, a complete analysis possible (e.g., double beta decay). Should aim for this whenever possible.
3. Often, cannot propose a complete set of models is not available, and a hierarchical analysis is needed. Still need to compare new model to null to say you have found something better.