Distribution function of extremely rare events from Bayesian reasoning

Kristian Piscicchia*

Museo Storico della Fisica e Centro Studi e Ricerche Enrico Fermi
INFN, Laboratori Nazionali di Frascati

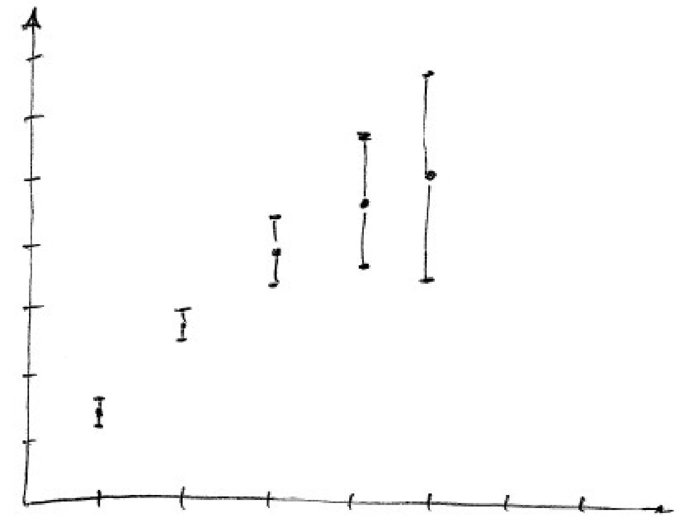
FQT2016
LNF, INFN, Frascati, Roma, Italy, 21 / 12 / 2016

*kristian.piscicchia@lnf.infn.it

MINIMUM Chi-Square @ MAXIMUM Likelihood

We want to determine the functional relation among two quantities X and Y (ex. X = energy, Y = corresponding rate)

$$y = f(x \mid a_1, \dots, a_n)$$



given N measured points it is always possible to find a polinomial of degree N-1 passing through the points, but this is not the point! Because this polinomial does not care the error bars.

If the theory suggests a function f, the problem reduces to find:

the vector $\bar{a}$ which is more probable, given the measured points (x_i, y_i) ; $i = 1, \dots, N$ within the experimental errors.

MINIMUM Chi-Square @ MAXIMUM Likelihood

Let us assume:

- x_is are not affected by errors
- y_is are Gaussian stochastic variables
- mean values and standard deviations for y_is are:

$$\mu_i = f(x_i, \bar{a}), \quad \bar{a} = (a_1, \dots, a_n) \quad ; \quad \sigma_i$$

- y_is can be considered as independent

The variable $$\chi^2 = \sum_{i=1}^{N} \frac{(y_i - \mu_i)^2}{\sigma_i^2} \qquad \chi^2 \in (0, \infty)$$

is distributed according with

$$P(\chi^2, n) = \frac{1}{2^{n/2}\,\Gamma(n/2)} \left[\chi^2\right]^{\frac{n}{2}-1} e^{-\frac{\chi^2}{2}}$$

$$E[\chi^2] = n \quad ; \quad Var[\chi^2] = 2n$$

$$\Gamma(x) = (x-1)! \quad \text{if } x \in \mathbb{N} \qquad \text{or} \qquad \Gamma(x) = \int_0^\infty e^{-t}\, t^{x-1}\, dt \quad ; \; x > 0$$

MINIMUM Chi-square

Solution to our problem: the vector $\bar{a}$ which minimises the χ^2 gives the best fit:

$$\chi^2 = \sum_{i=1}^{N} \frac{\{y_i - f(x_i | \bar{a})\}^2}{\sigma_i^2}$$

$\bar{a}$ is given by the system of n equations:

$$\begin{cases} \sum_{i=1}^{N} (y_i - f(x_i | \bar{a})) \left. \frac{\partial f}{\partial a_1} \right|_{x_i} = 0 \\ \vdots \\ \sum_{i=1}^{N} (y_i - f(x_i | \bar{a})) \left. \frac{\partial f}{\partial a_n} \right|_{x_i} = 0 \end{cases}$$

ANALYTIC SOLUTION USUALLY NOT EXISTS, BUT

MINUIT LUCKILY EXISTS

in principle the solutions are $a_i = g_i(\bar{x}, \bar{y}, \bar{\sigma})$

with errors:

$$\sigma_{a_i} = \left\{ \sum_{j=1}^{N} \left(\frac{\partial a_i}{\partial y_j} \sigma_j \right)^2 \right\}^{\frac{1}{2}}$$

HOW IS JUSTIFIED ?

MINIMUM chi-square

Each of the y_is is Gaussian

$$G\left(y_i \mid \mu_i = f(x_i, \bar{a}), \sigma_i\right) = \frac{1}{\sqrt{2\pi}\,\sigma_i} e^{-\frac{(y_i - f(x_i,\bar{a}))^2}{2\sigma_i^2}}$$

if y_is are independent the joint distribution function is

$$P(y_1, \dots, y_N) = \prod_{i=1}^{N} G(y_i) \propto e^{-\sum_{i=1}^{N} \frac{(y_i - f(x_i,\bar{a}))^2}{2\sigma_i^2}} = e^{-\frac{\chi^2}{2}}$$

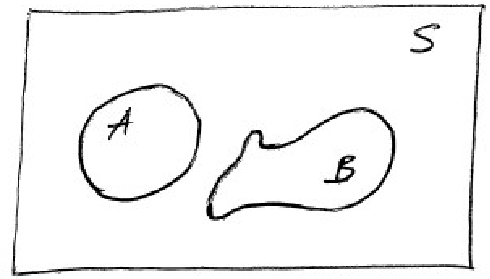
the minimum χ^2 maximizes P, <u>but</u> $G(y_i)$ are probability statements on y_i. That is if I would repeat m-times the measurement of the physical quantity related to y_i (Y_i) then

$$P\left(\mu_i - \frac{\sigma}{\sqrt{m}} \le \overline{Y}_i \le \mu_i + \frac{\sigma}{\sqrt{m}}\right) = 68\%$$

I want a probability stament on $\bar{a}$. How do reverse the probability?

BAYES THEOREM

The probability is a real function such that:



① $P(A) \geq 0$

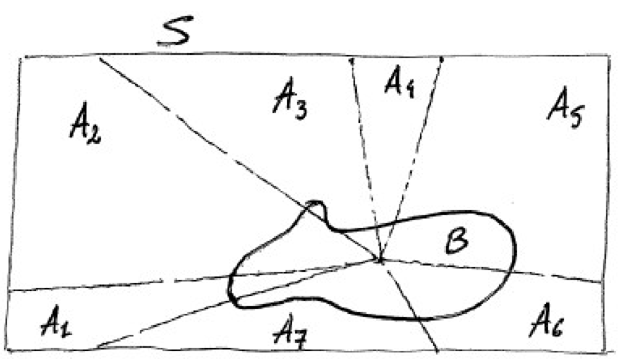
② GIVEN $(A \cap B) = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$

③ $P(S) = 1$

CONDITIONAL PROBABILITY $$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

given $A \cap B = B \cap A \Rightarrow P(A \cap B) = P(B \cap A) = P(B|A) \cdot P(A)$

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$



- if $A_i \cap A_j = \emptyset \quad \forall\, i, j$

  and $\bigcup_i A_i = S$

  then $P(B) = \sum_i P(B \cap A_i) = \sum_i P(B|A_i) \cdot P(A_i)$

BAYES THEOREM

$$\underset{\text{POSTERIOR}}{P(A|B)} = \frac{P(B|A)\cdot P(A)}{P(B)} = \frac{\overset{\text{LIKELIHOOD}}{P(B|A)}\cdot \overset{\text{PRIOR}}{P(A)}}{\sum_i P(B|A_i)\cdot P(A_i)}$$

for a continuous variable x:

$$P(x|y) = \frac{P(y|x)\cdot P(x)}{\int P(y|x')\cdot P(x')\,dx'}$$

FROM THE MINIMUM χ^2 TO THE MAXIMUM LIKELIHOOD

BAYESIAN INFERENCE OF THE TRUE VALUE μ OF y

From the likelihood

$$G(y_i \mid \mu_i, \sigma_i) = \frac{1}{\sqrt{2\pi}\,\sigma_i} e^{-\frac{(y_i - \mu_i)^2}{2\sigma_i^2}}$$

remember $\mu_i = f(x_i, \bar{a})$

we get for μ_i

$$P(\mu_i \mid y_i) = \frac{\frac{1}{\sqrt{2\pi}\,\sigma_i} e^{-\frac{(y_i-\mu_i)^2}{2\sigma_i^2}} \cdot P(\mu_i)}{\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\,\sigma_i} e^{-\frac{(y_i-\mu_i')^2}{2\sigma_i^2}} P(\mu_i')\, d\mu_i'} =$$

Considering each value of μ_i equally likely over a large interval $P(\mu_i) = \text{const}$

$$= \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(\mu_i - y_i)^2}{2\sigma_i^2}} \Rightarrow E[\mu_i] = y_i \; ; \; \sigma(\mu_i) = \sigma$$

MINIMIZING χ^2 WE FIND THE VECTOR $\bar{a}$ WHICH MAXIZES THE PROBABILITY FOR $f(x_i, \bar{a})$ to be compatible with the observables (x_i, y_i) within the errors σ_i -

PAY ATTENTION !

A noteworthy case of a prior for which the "naive" inversion is paradoxical:

$\mu \geq 0$ when y_i can fall outside

ex. signal - background when are similar,

μ is the mass of a particle ...

If we have a prior knowledge which can be modelled as Gaussian

$$P(\mu|\mu_0,\sigma_0) = \frac{1}{\sqrt{2\pi}\,\sigma_0}\, e^{-\frac{(\mu-\mu_0)^2}{2\sigma_0^2}}$$

then the posterior results:

$$P(\mu|y,\mu_0,\sigma_0) = \frac{1}{\sqrt{2\pi}\,\sigma_1}\, e^{-\frac{(\mu-\mu_1)^2}{2\sigma_1^2}}$$

with

$$\mu_1 = E[\mu] = \frac{\sigma_1^2}{\sigma^2} d + \frac{\sigma_1^2}{\sigma_0^2} \mu_0$$

$$\sigma_1^2 = Var[\mu] = \left\{ \sigma_0^{-2} + \sigma^{-2} \right\}^{-1}$$

Which reduces to uniform prior in the limit

$\sigma_0 \to \infty$; $\mu_0 = const.$

BERNULLI PROCESSS:

Consider the stochastic variable which can assume the value 1 (success) and 0 (failure). $P(1) = p$; $P(0) = 1 - p = q$

$$P(x) = p^x (1-p)^{1-x} = \text{probability of } x \text{ consecutive successes}$$

BINOMIAL DISTRIBUTION:

make n attempts of a Bernulli process, what is the probability of x successes?

$$p^x (1-p)^{n-x}$$

times the number of ways to choose n objects among n independently of the order:

$$\frac{n \cdot (n-1) \cdot (n-2) \cdots (n-x+1)}{x!} = \frac{n!}{(n-x)!\, x!} = \binom{n}{x}$$

$$E[x] = n p \quad ; \quad \sigma = \sqrt{n \cdot p \cdot q}$$

take the limit of the trials $n \to \infty$
take the probability of a single success $p \to 0$ } but $n \cdot p = \lambda$ finite

obtain:

$$P(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$ POISSON DISTRIBUTION.

Consider a process taking place in time and space (decay of an unstable nucleus) you want the counts over a period t.

① the probability of one count in Δt is $p = \tau \cdot \Delta t$
② the probability of two counts in Δt is negligible
③ the number of counts in Δt_1 is independent on those in Δt_2

In each interval Δt we observe a Bernulli process, $\Delta t = \frac{t}{n} \to 0 \Rightarrow p \to 0$ when the number of trials $n \to \infty$.
The variable number of counts in t is Poissonian with $\lambda = n \cdot p = \tau t$

$e^{-\tau t}$ τ is the rate.

An example related to collapse models

W. f. collapse reminder

$$d|\psi_t\rangle = \left[-\frac{i}{\hbar}H\,dt + \sqrt{\lambda}\int d^3x(N(\mathbf{x}) - \langle N(\mathbf{x})\rangle_t)dW_t(\mathbf{x}) - \frac{\lambda}{2}\int d^3x(N(\mathbf{x}) - \langle N(\mathbf{x})\rangle_t)^2 dt\right]|\psi_t\rangle$$

System's Hamiltonian

NEW COLLAPSE TERMS ➡ **New Physics**

$N(\mathbf{x}) = a^\dagger(\mathbf{x})a(\mathbf{x})$ particle density operator — **choice of the preferred basis**

$\langle N(\mathbf{x})\rangle_t = \langle\psi_t|N(\mathbf{x})|\psi_t\rangle$ — **nonlinearity**

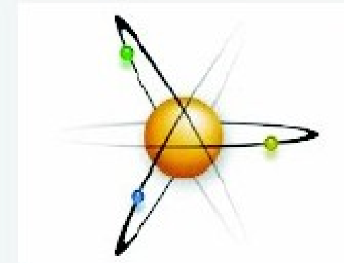
$W_t(\mathbf{x}) = \text{noise} \quad \mathbb{E}[W_t(\mathbf{x})] = 0, \quad \mathbb{E}[W_t(\mathbf{x})W_s(\mathbf{y})] = \delta(t-s)e^{-(\alpha/4)(\mathbf{x}-\mathbf{y})^2}$ — **stochasticity**

$\lambda =$ collapse strength $\qquad r_C = 1/\sqrt{\alpha} =$ correlation length — **two parameters**

the only possible modification of the Schrödinger equation, compatible with the non-faster-than-light signaling condition!

Which values for λ and r_c?

Microscopic world
(few particles)



$$\lambda \sim 10^{-8\pm2}\mathrm{s}^{-1}$$

QUANTUM – CLASSICAL TRANSITION (Adler - 2007)

Mesoscopic world
Latent image formation
+
perception in the eye
(~ 10^4 - 10^5 particles)

S.L. Adler, JPA 40, 2935 (2007)

A. Bassi, D.A. Deckert & L. Ferialdi, EPL 92, 50006 (2010)

$$\lambda \sim 10^{-17}\mathrm{s}^{-1}$$

QUANTUM – CLASSICAL TRANSITION (GRW - 1986)

Macroscopic world
(> 10^{13} particles)

G.C. Ghirardi, A. Rimini and T. Weber, PRD 34, 470 (1986)

$$r_C = 1/\sqrt{\alpha} \sim 10^{-5}\mathrm{cm}$$

Increasing size of the system

… spontaneous photon emission

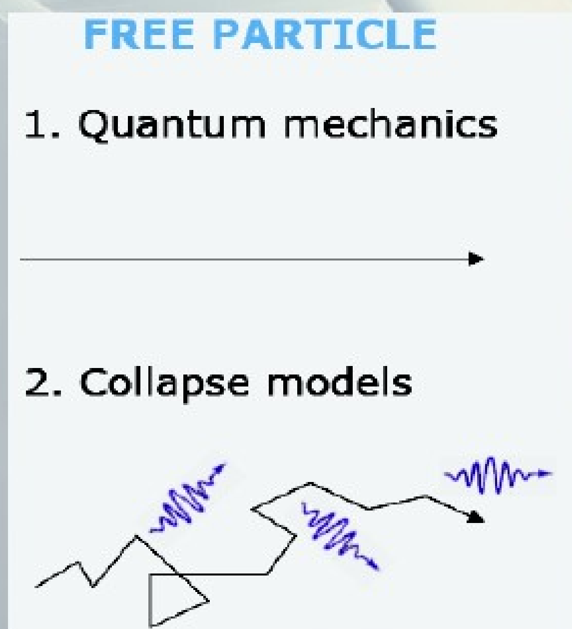
Besides collapsing the state vector to the position basis in non relativistic QM
the interaction with the stochastic field increases the expectation value of particle's energy

↓

implies for a charged particle energy radiation (not present in standard QM) !!!

↓

The comparison between theoretical prediction and experimental results provide constraints on the parameters of the CSL model



$$\frac{d\Gamma_k}{dk} = \frac{e^2\lambda}{4\pi^2 r_c^2 m^2 k}$$

Q. Fu, Phys. Rev. A 56, 1806 (1997)

S.L. Adler, A. Bassi & S. Donadi, ArXiv 1011.3941

Our limit on collapse rate - best limit in the world using IGEX data

Phys. Scr. 90 (2014) 028003, Found. Phys. (2016) 46: 263-268,
Acta Phys. Polon. B46, (2015), 147-152, J. Adv. Phys. 4, 263-266, (2015)

$\lambda < 2.5 \pm 0.2 \times 10^{-18}\ s^{-1}$ — No mass-proportional — excluded if white noise

$\lambda < 8.5 \pm 0.5 \times 10^{-12}\ s^{-1}$ — Mass-proportional — Adler value excluded if white noise

Laboratory experiments	Distance (decades) from the enhanced CSL value	Cosmological data	Distance (decades) from the enhanced CSL value
Fullerene diffraction experiments	11-12 (2-3)	Dissociation of cosmic hydrogen	18 (9)
Decay of supercurrents (SQUIDs)	15 (6)	Heating of Intergalactic medium (IGM)	9 (0)
Spontaneous X-ray emission	**6 (-3)**	Heating of protons in the universe	13 (4)
Proton decay	19 (10)	Heating of Interstellar dust grains	16 (7)

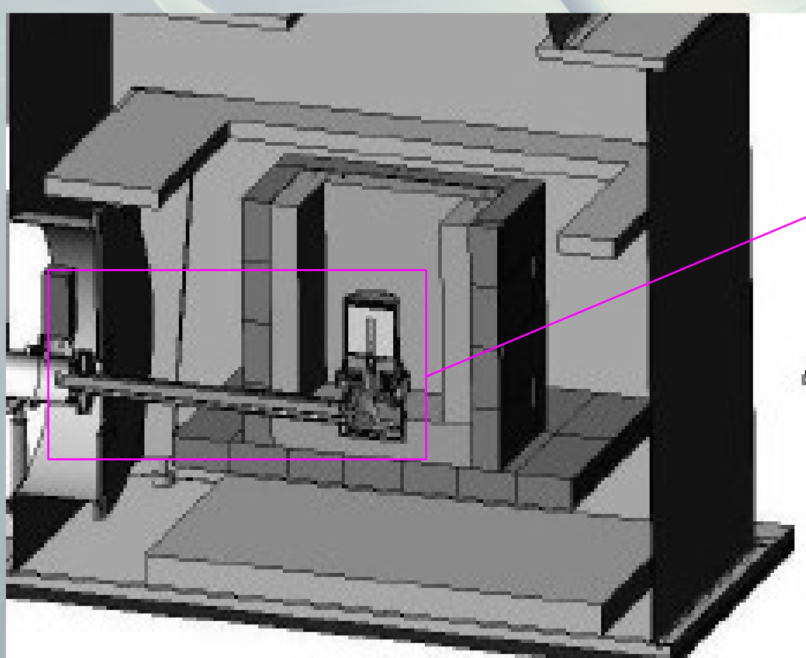
HPGe & LNGS
the most sensitive measurement of λ ever

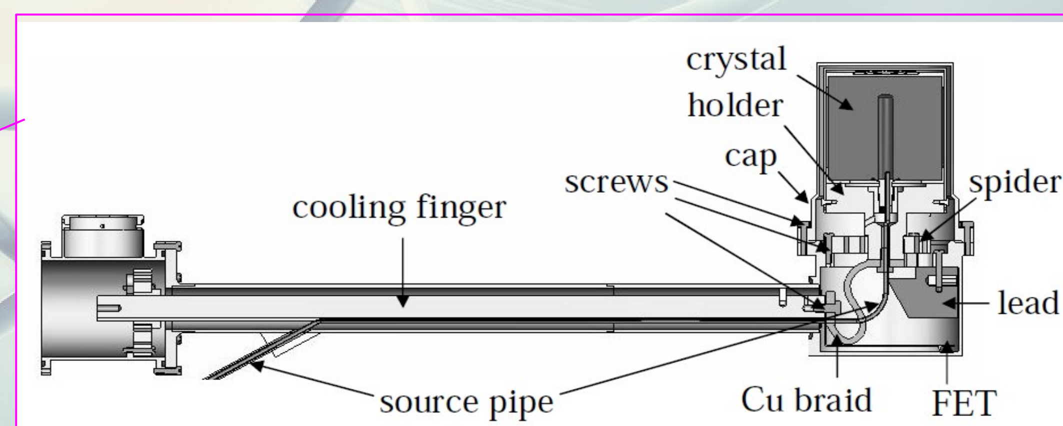
High purity Ge detector measurement:

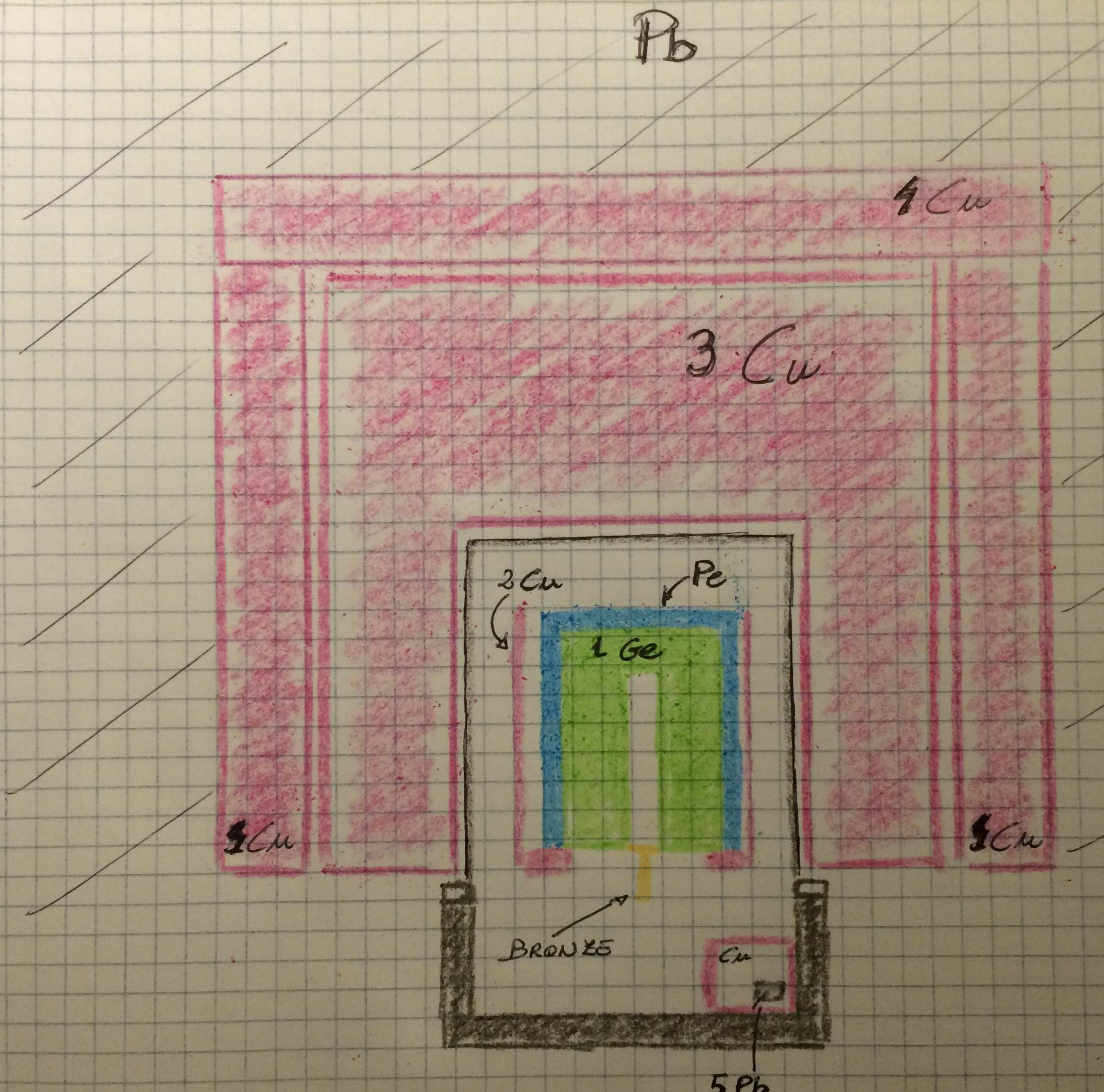
- **active Ge detector surrounded by complex electrolytic Cu + Pb shielding**

- **10B-polyethylene plates reduce the neutron flux towards the detector**

- **shield + cryostat enclosed in air tight steel housing flushed with nitrogen to avoid contact with external air (and thus radon).**






Pb
4 Cu
3 Cu
2 Cu
Pe
1 Ge
Cu
Cu
BRONZE
Cu
5 Pb

Probability distribution function of λ
theoretical information

Goal: obtain the probability distribution function PDF(λ) of the collapse rate parameter given:

- the theoretical information

Rate of spontaneously emitted photons as a consequence of *p* and *e* interaction with the stochastic field,

$$\frac{d\Gamma}{dE} = \left\{ \left(N_p^2 + N_e\right) \cdot \left(m\, n\, T\right) \right\} \frac{\lambda \hbar e^2}{4\pi^2 \epsilon_0 c^3 m_N^2 r_c^2 E}$$

(depending on λ)

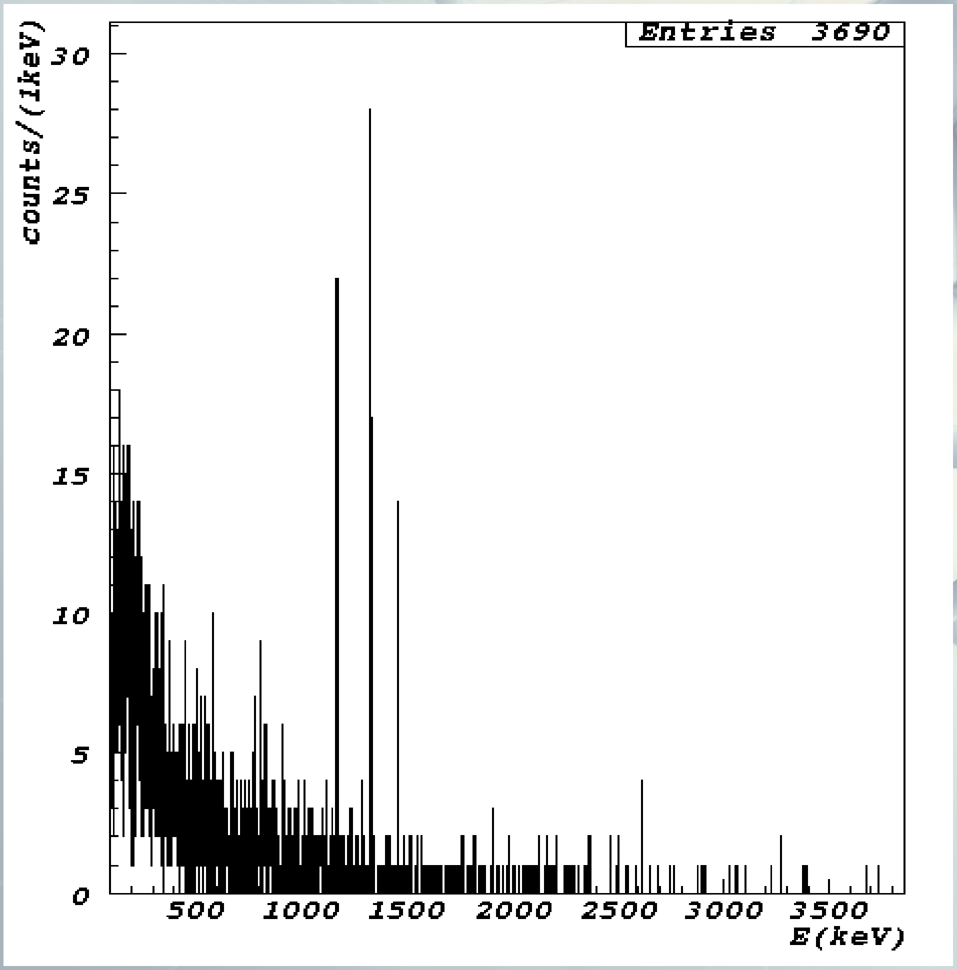
as a function of *E*

(mass of the emitting material · number of atoms per unit mass · total acquisition time)

Probability distribution function of λ
experimental information

Goal: obtain the probability distribution function PDF(λ) of the collapse rate parameter given:

- the experimental information



X-ray measurements performed in the very low background environment of the LNGS (INFN)

High Purity Germanium based detectors.
(three months data taking with 2kg germanium active mass)

According with theory constrains we consider the protons emission in the range ΔE=(1000-3800)keV.

Probability distribution function of λ
experimental information

Goal: obtain the probability distribution function PDF(λ) of the collapse rate parameter given:

- **the experimental information**

total number of counts in the selected energy range:

$$f(z_c) = \frac{\Lambda_c^{z_c} e^{-\Lambda_c}}{z_c!}$$

besides the background from standard processes let's turn on the spontaneous emission contribution ...

- z_b = number of counts due to background,
- z_s = number of counts due to signal,
- $z_c = z_b + z_s$; $z_s \sim P_{\Lambda_s}$; $z_b \sim P_{\Lambda_b}$,

$$f(z_c|P_{\lambda s}, P_{\lambda b}) = \sum_{z_s, z_b} \delta_{z_c, z_s+z_b} \, f(z_s|P_{\lambda s}) \, f(z_b|P_{\lambda b}) = \frac{(\Lambda_s + \Lambda_b)^{z_s+z_b} \, e^{(\Lambda_s+\Lambda_b)}}{z_c!}$$

Probability distribution function for λ

According with the Bayes theorem:

$$f(\lambda|\mathrm{ex},\mathrm{th}) = f(\mathrm{ex}|\lambda)\cdot f(\lambda|\mathrm{th})$$

let us assume a conservative prior [S. L. Adler, JPA 40, 2935 (2007)]

$$f(\lambda|\mathrm{th}) = 1 \qquad \lambda < 10^{-6}\mathrm{s}^{-1}$$
$$f(\lambda|\mathrm{th}) = 0 \qquad \lambda > 10^{-6}\mathrm{s}^{-1}$$

PDF(λ) is:

$$f(\lambda|\mathrm{ex},\mathrm{th}) = \frac{(\Lambda_s(\lambda)+\Lambda_b)^{z_c}\cdot e^{-(\Lambda_s(\lambda)+\Lambda_b)}}{z_c!} \qquad \lambda < 10^{-6}\mathrm{s}^{-1}$$
$$f(\lambda|\mathrm{ex},\mathrm{th}) = 0 \qquad \lambda < 10^{-6}\mathrm{s}^{-1}$$

Advantages .. - possibility to extract unambiguous limits corresponding to the probability level you prefer,

- *f*(λ) can be updated with all the experimental information at your disposal by updating the likelihood,

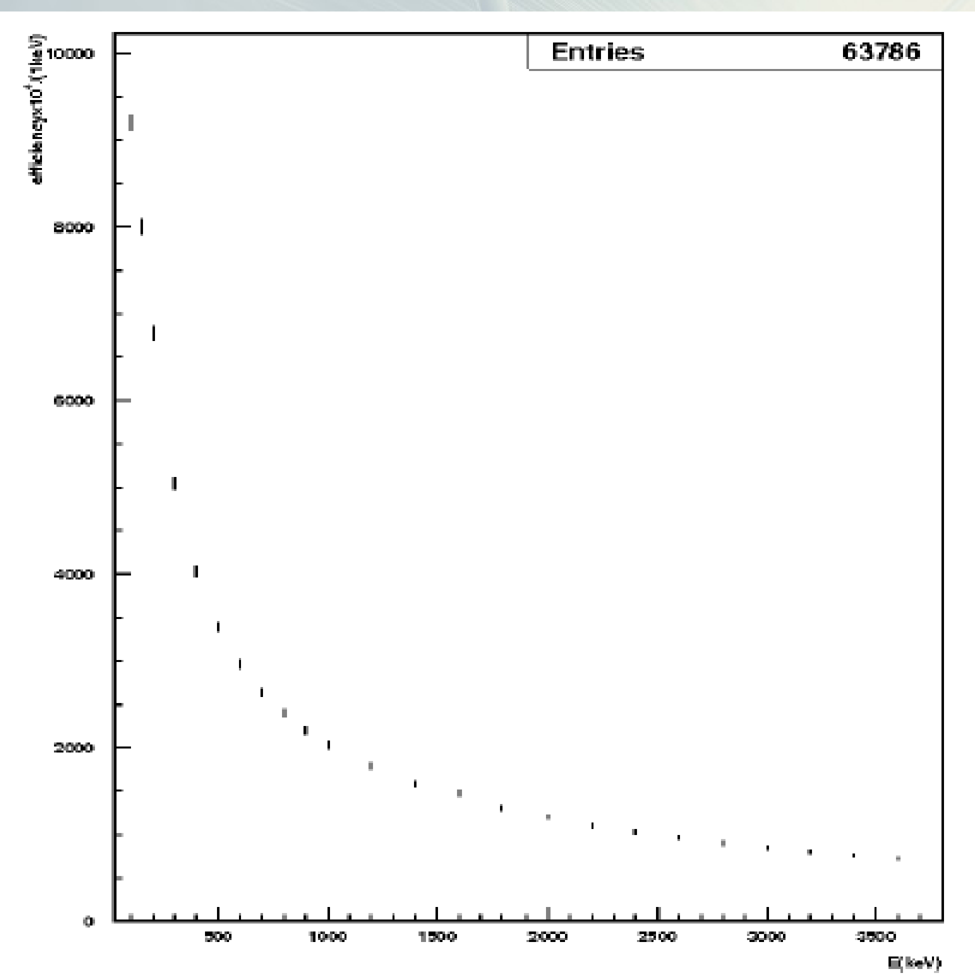
- competing or future models can be simply implemented

The expected signal .. $\Lambda_s(\lambda)$

Each material of the detector contributes to the signal rate with different:

***m*, *n* and $\varepsilon(E)$**

$\varepsilon(E)$ depends on the material and the geometry of the detector.



Simulated detection eciency for γs produced inside the Germanium detector, multiplied by 10^4

Photon detection efficiencies obtained by means of MC simulations, ganerating γs in the range (E1 – E2)

$$\varepsilon(E) = N_{\text{measured events}}(E) / N_{\text{gen. events}}(E)$$

In practice .. $\Lambda_s(\lambda)$

Each material of the detector contributes to the signal rate with different:

***m*, *n* and ε(*E*)**

ε(*E*)   depends on the material and the geometry of the detector.

efficiency distributions fitted to obtain the efficiency functions:

$$\epsilon_i(E) = \sum_{j=0}^{ci} \xi_{ij} E^j$$

to obtain the signal predicted by theory & processed by the detector

$$z_s(\lambda) = \sum_i \int_{E_1}^{E_2} \left.\frac{d\Gamma}{dE}\right|_i \epsilon_i(E)\, dE =$$

$$= \sum_i \int_{E_1}^{E_2} N_{pi}^2\, \alpha_i\, \beta\, \frac{\lambda}{E} \sum_{j=0}^{ci} \xi_{ij} E^j\, dE$$

with:

$$\alpha_i = m_i n_i T,$$

$$\beta = \frac{\hbar e^2}{4\pi^2 \epsilon_0 c^3 m_N^2 r_c^2}$$

In practice .. Λ_b

Evaluation of the background:
simulation of the radionuclides decay for which materials of the setup contribute taking into account for: emission probabilities, decay scheme of each radionuclide, photons propagation and interactions inside the materials of the detector, detection eciency. The following contributions were considered:

- **Co60 from the inner Copper**
- **Co60 from the Copper block + plate**
- **Co58 from the Copper block + plate**
- **K40 from Bronze**
- **Ra226 from Bronze**
- **Bi214 from Bronze**
- **Pb214 from Bronze**
- **Bi212 from Bronze**
- **Pb212 from Bronze**
- **Tl208 from Bronze**
- **Ra226 from Poliethylene**
- **Bi214 from Poliethylene**
- **Pb214 from Poliethylene**

The expected background .. Λ_b

Evaluation of the background:
simulation of the radionuclides decay for which materials of the setup contribute taking into account for the emission probabilities and the decay scheme of each radionuclide:

- Co60 from the inner Copper
- Co60 from the Copper block + plate
- Co58 from the Copper block + plate
- K40 from Bronze
- Ra226 from Bronze
- Bi214 from Bronze
- Pb214 from Bronze
- Bi212 from Bronze
- Pb212 from Bronze
- Tl208 from Bronze
- Ra226 from Poliethylene
- Bi214 from Poliethylene
- Pb214 from Poliethylene

measured activities

$$z_{b,ij} = \frac{m_i\, A_{ij}\, T\, N_{rec,ij}}{N_{ik}}$$

detected MC γs

 # simulated events

Expected number of background counts $\Lambda_b = z_b + 1$

Presently we can describe 88% of the measured spectrum

Upper limit for the collapse rate parameter λ

- From the p.d.f we obtain the cumulative distribution function:

$$F(\lambda) = \frac{\int_0^\lambda f(\lambda|\mathrm{ex},\mathrm{th})\mathrm{d}\lambda}{\int_0^\infty f(\lambda|\mathrm{ex},\mathrm{th})\mathrm{d}\lambda} = \frac{\int_0^\lambda \frac{1}{z_c!}(a\lambda+\Lambda_b+1)^{z_c} e^{-(a\lambda+\Lambda_b+1)}\mathrm{d}\lambda}{\int_0^\infty \frac{1}{z_c!}(a\lambda+\Lambda_b+1)^{z_c} e^{-(a\lambda+\Lambda_b+1)}\mathrm{d}\lambda}$$

which we express in terms of cumulative gamma functions

$$F(\lambda) = 1 - \frac{\Gamma(z_c+1, a\lambda+1+\Lambda_b)}{\Gamma(z_c+1, 1+\Lambda_b)}$$

- put the measured z_c and the calculated $\Lambda_s(\lambda) = a\lambda + 1$, Λ_b in the cumulative distribution function

extract the limit at the desired probability level ...

$\lambda < 5 \cdot 10^{-13}$ with a probability of 95%

Gain factor ~ 20