Impact of statistics and detector characteristics on data analysis

Seminario Nazionale Rivelatori Innovativi

Alberto Rotondi, Pavia University and INFN Sezione di Pavia

- statistics
- efficiency
- resolution

- **counting**
- **pile-up effects**
- **unfolding**
- **signal to background ratio**

There are 7 types of measurements

$$M\begin{pmatrix} x & 0 & 0 \\ X & \sigma & \Delta \end{pmatrix}$$

$8\text{-}M(x,0,0)=7$

There are 7 types of measurements

1 $M(x,0,\Delta) \rightarrow x \pm \frac{\Delta}{2}$ (CL=100%)

1 $M(x,\sigma,0) \rightarrow x \pm \frac{s}{\sqrt{N}}$ (CL=68%)

1 $M(x,\sigma,\Delta) \rightarrow x \pm \frac{s}{\sqrt{N}}(\mathrm{stat}) \pm \frac{\Delta}{2}(\mathrm{sys}) \Rightarrow x \pm \sqrt{\frac{s^2}{N} + \frac{\Delta^2}{12}}$ (CL~68%)

1 $M(X,0,0) \rightarrow x \pm \sqrt{x},\ \ x \pm \sqrt{x\left(1 - \frac{x}{N}\right)}$ Counting, Pile-up (CL~68%)

3 $M(X,\sigma,\Delta) \rightarrow g(y) = \int f(x)\delta(x,y)\mathrm{d}y$ Unfolding techniques -> $f(x)$

Detector statistcs

Detector Efficiency → Binomial distribution

$$B(x;n,\varepsilon) = \frac{n!}{x!(n-x)!}\varepsilon^{x}(1-\varepsilon)^{n-x}$$

Counts → Poisson distribution

$$P(x;\mu) = \frac{\mu^{x}}{x!}\,\mathrm{e}^{-\mu}$$

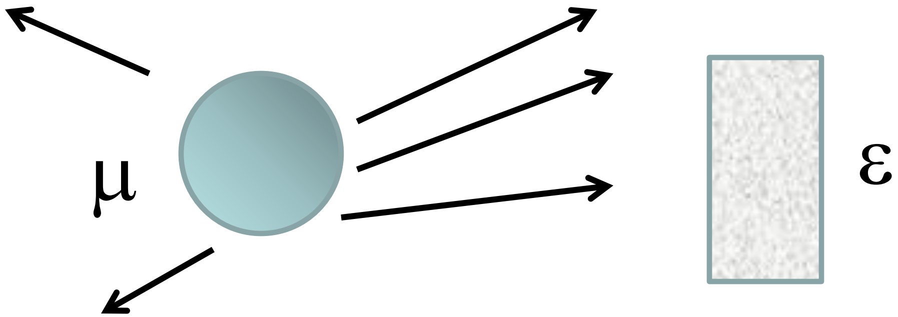
Arrival times → exponential

$$e(t;\tau) = \frac{1}{\tau}\,\mathrm{e}^{-t/\tau}, \quad \mu = \frac{\Delta t}{\tau}$$

Resolution effects → Gaussian distribution

$$G(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\,\sigma}\,e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$$

Some important facts (I)



$$B(x;\varepsilon,n)\,P(n;\mu)=\frac{n!}{x!(n-x)!}\varepsilon^{x}(1-\varepsilon)^{n-x}\frac{\mu^{n}}{n!}e^{-\mu}$$

$$y=n-x,\qquad e^{-\mu}=e^{-\mu\varepsilon}e^{-\mu(1-\varepsilon)},\qquad \mu^{n}=\mu^{n-x}\mu^{x}=\mu^{y}\mu^{x}$$

$$B(x;\varepsilon,n)\,P(n;\mu)=\frac{(\varepsilon\mu)^{x}}{x!}e^{-\varepsilon\mu}\frac{[(1-\varepsilon)\mu]^{y}}{y!}e^{-(1-\varepsilon)\mu}=P(x;\varepsilon\mu)P(n-x;(1-\varepsilon)\mu)$$

Conclusion: a binomial counter with efficiency ε that sees a Poisson source of intensity μ, counts in a poissonian way with mean $\varepsilon\mu$

Some important facts (II)

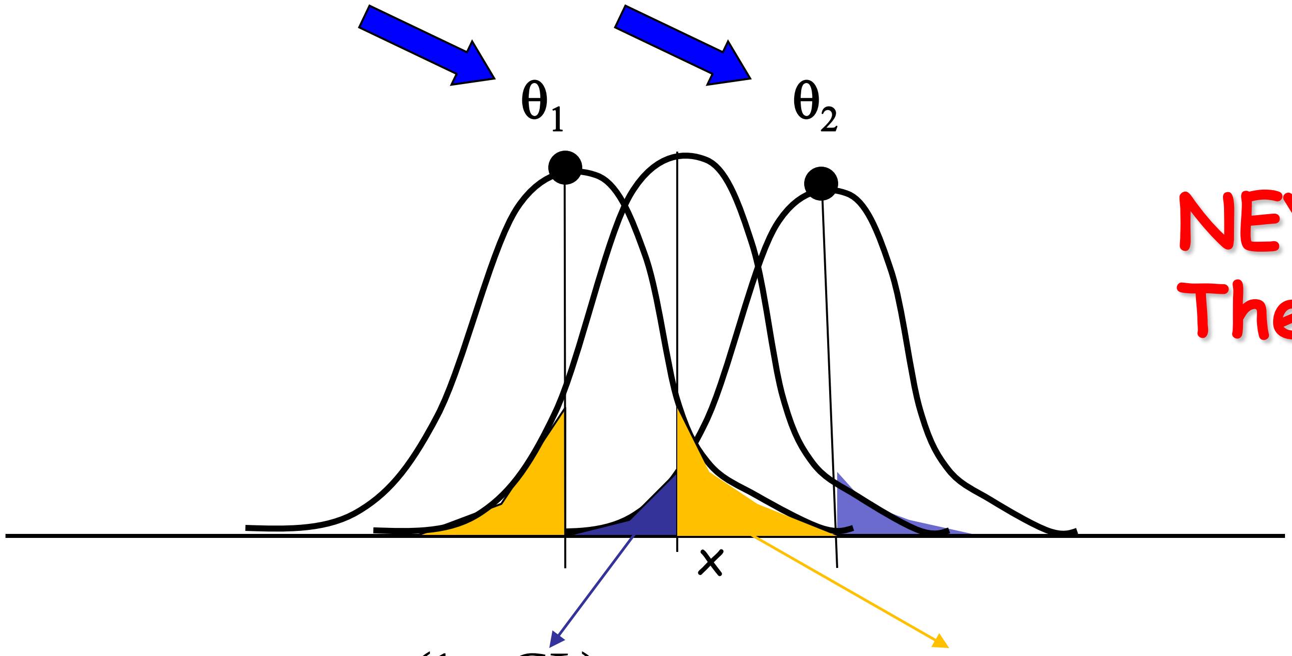
$$P(x;\mu) = \frac{\mu^x}{x!}\,\mathrm{e}^{-\mu} \quad \Longleftrightarrow \quad e(t;\tau) = \frac{1}{\tau}\,\mathrm{e}^{-t/\tau}, \quad \mu = \frac{t}{\tau}$$

Frequency domain — Time domain

$$e^{-\mu} \qquad \frac{dt}{\tau} = r\,dt$$

important: the time distribution remains the same if the clock starts at any time or if it starts at the arrival of the last event.

NEYMAN Theorem (1937)



$$\int_{-\infty}^{x} p(x;\theta_2)\,\mathrm{d}x = \frac{(1-CL)}{2},$$

$$\int_{x}^{\infty} p(x;\theta_1)\,\mathrm{d}x = \frac{(1-CL)}{2},$$

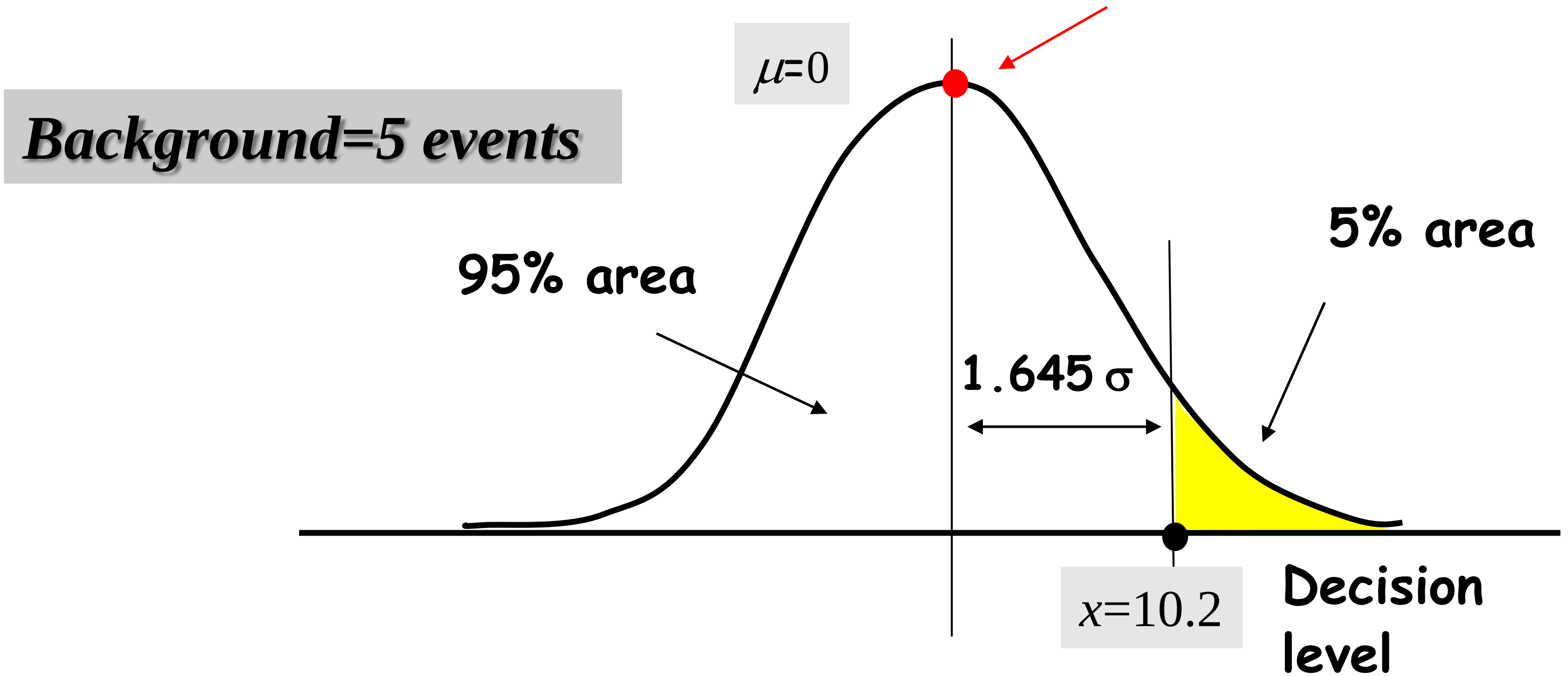
$$\sum_{k=0}^{x}\binom{n}{k}\theta_2^k(1-\theta_2)^{n-k} = \frac{(1-CL)}{2}$$

$$\sum_{k=x}^{n}\binom{n}{k}\theta_1^k(1-\theta_1)^{n-k} = \frac{(1-CL)}{2}$$

$$P\left\{\frac{|x-\mu|}{\sigma[x]} \le t_{\alpha/2}\right\} \ge CL$$

$$P\{x - t_{\alpha/2}\sigma[x] \le \mu \le x + t_{1-\alpha/2}\sigma[x]\} \ge CL$$

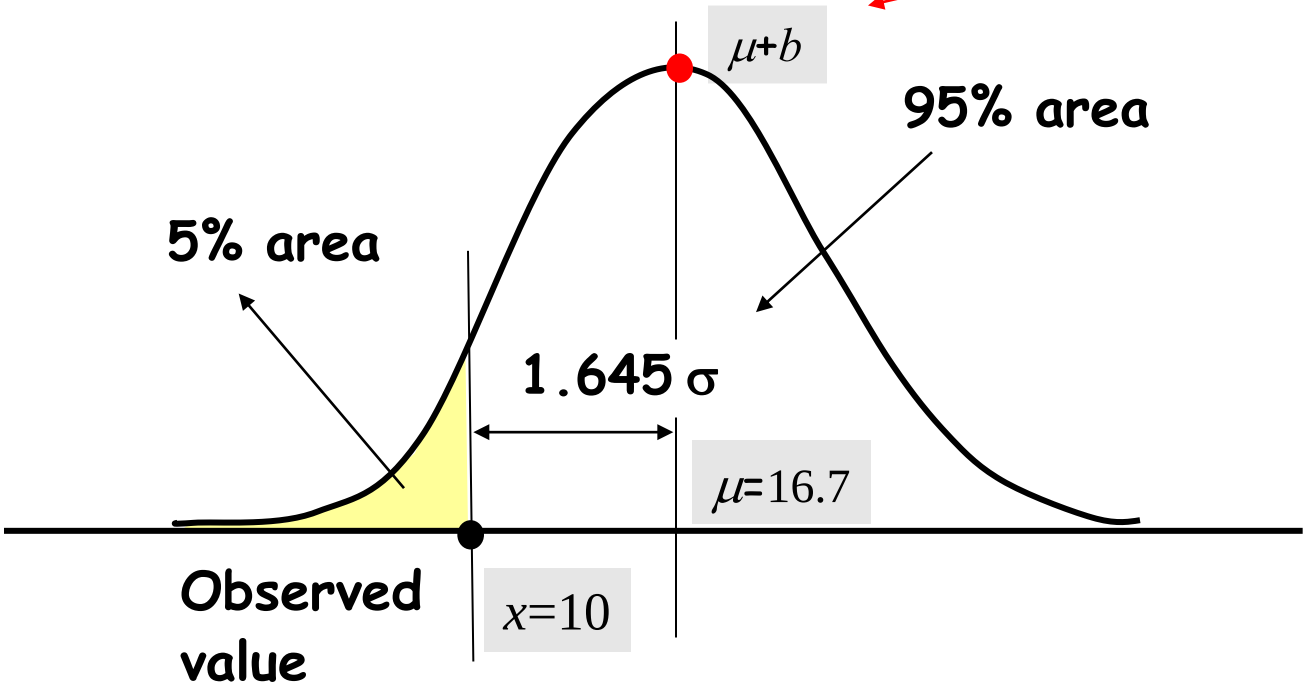
The 95% decision level...



$$5 + 1.645\sqrt{10} = 10.2$$

Meaning: if we consider the signal as detected, we will be wrong in 5% of the cases *when the signal is absent*

The 95% CL gaussian upper limit..



$$\frac{\mu+b-10}{\sqrt{\mu+2b}}=1.645 \Rightarrow \mu+b-1.645\sqrt{\mu+2b}-10=0 \xrightarrow{b=5} \mu=12.8$$

Meaning: this upper limit should give values less than the observed one in less than 5% of the experiments

Counting experiments

$$\sum_{k=x}^{\infty} \binom{n}{k} \theta_1^k (1-\theta_1)^{n-k} = \frac{(1-CL)}{2}$$

$$\sum_{k=0}^{x} \binom{n}{k} \theta_2^k (1-\theta_2)^{n-k} = \frac{(1-CL)}{2}$$

$$P\left\{ \frac{|x-\mu|}{\sigma[x]} \le t_{\alpha/2} \right\} \ge CL$$

CL=1-α is the asymptotic probability the interval will contain the true value

COVERAGE is the probability that the specific experiment does contain the true value irrespective of what the true value is

On the infinite ensemble of experiments, for a continuous variable **Coverage** and **CL** tend to coincide

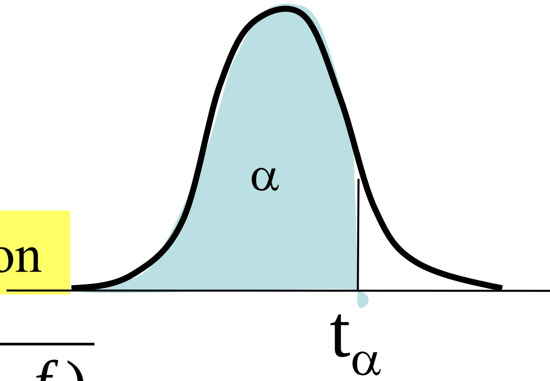
In counting experiments the variables are discrete and **CL** and **Coverage** do not coincide

What is requested is the **minimum overcoverage**

Counting experiments: Binomial case

$$P\left\{\frac{|F-p|}{\sigma[p]} \le t_\alpha\right\} = P\left\{\frac{|F-p|}{\sqrt{\dfrac{p(1-p)}{n}}} \le t_\alpha\right\} = CL$$

t=1, area 84%
Quantile α=0.84
P[|f-p|<t σ]= 68%



t is the quantile of the normal distribution

$$\frac{|f-p|}{\sqrt{\dfrac{p(1-p)}{n}}} \le |t| \longrightarrow p = \frac{f + \dfrac{t^2}{2n}}{\dfrac{t^2}{n}+1} \pm \frac{t\sqrt{\dfrac{t^2}{4n^2} + \dfrac{f(1-f)}{n}}}{\dfrac{t^2}{n}+1}$$

Wilson interval (1934)

$$\xrightarrow{n>>1} p = f \pm t_\alpha \sqrt{\frac{f(1-f)}{n}}$$

Wald (1950)
Standard in Physics

A further improvement:

The continuity correction is equivalent to
The Clopper-Pearson formula

$$\varepsilon = \frac{f_\pm + \frac{t_{\alpha/2}{}^2}{2n}}{\frac{t_{\alpha/2}{}^2}{n}+1} \pm \frac{t_{\alpha/2}\sqrt{\frac{t_{\alpha/2}{}^2}{4n^2}+\frac{f_\pm(1-f_\pm)}{n}}}{\frac{t_{\alpha/2}{}^2}{n}+1}, \quad \begin{array}{l} x = n,\ [p_1,1],\ p_1 = (1-CL)^{1/n} \\ x = 0,\ [0,p_2],\ p_2 = 1-(1-CL)^{1/n} \end{array}$$

$$f_+ = (x+0.5)/n,\ \ f_- = (x-0.5)/n,$$

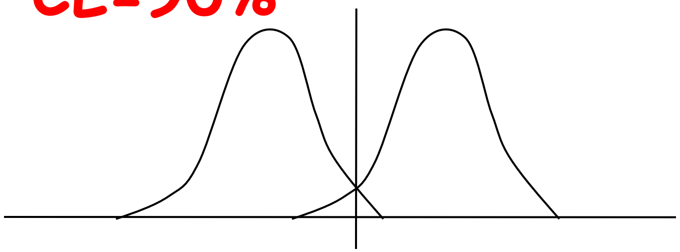
$t_{\alpha/2}$ gaussian, 1-$CL = \alpha$, $t = 1$ is 1σ

This should become the standard formula also for physicists

Elementary example

20 events have been generated and 5 passed the cut
What is the estimation of the efficiency with CL=90%?

Frequentist result: **x=5, n=20, CL=90%**



$$\sum_{k=5}^{20} \binom{n}{k} \varepsilon_1^k (1-\varepsilon_1)^{n-k} = 0.05$$

PDG

$$\sum_{k=0}^{5} \binom{n}{k} \varepsilon_2^k (1-\varepsilon_2)^{n-k} = 0.05$$

$\varepsilon_1, \varepsilon_2$ → ε =[0.104, 0.455]

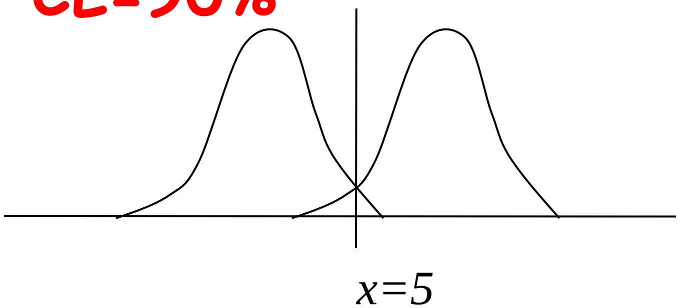
$$f \pm t_\alpha \sqrt{\frac{f(1-f)}{n}}$$

ε =[0.090, 0.410]

Elementary example

20 events have been generated and 5 passed the cut
What is the estimation of the efficiency with CL=90%?

Frequentist result: x=5, n=20, CL=90%



$$\sum_{k=5}^{20} \binom{n}{k} \varepsilon_1^k (1-\varepsilon_1)^{n-k} = 0.05$$

$$\sum_{k=0}^{5} \binom{n}{k} \varepsilon_2^k (1-\varepsilon_2)^{n-k} = 0.05$$

$\varepsilon_1, \varepsilon_2$ → ε =[0.104, 0.455]

PDG

$$\varepsilon = \frac{f_{\pm} + \frac{t_{\alpha/2}^{2}}{2n}}{\frac{t_{\alpha/2}^{2}}{n}+1} \pm \frac{t_{\alpha/2}\sqrt{\frac{t_{\alpha/2}^{2}}{4n^2} + \frac{f_{\pm}(1-f_{\pm})}{n}}}{\frac{t_{\alpha/2}^{2}}{n}+1},$$

ε =[0.145, 0.405]

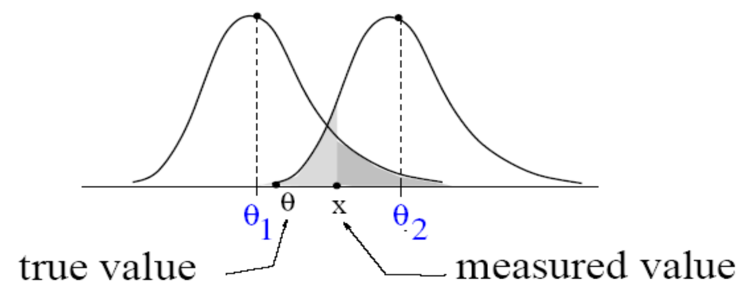
Coverage simulation

x = gRandom → Binomial(p,N) → x

$1\text{-}CL = \alpha$

$$\sum_{k=x}^{n}\binom{n}{k} p_1^k (1-p_1)^{n-k} = \alpha/2$$

$$\sum_{k=0}^{x}\binom{n}{k} p_2^k (1-p_2)^{n-k} = \alpha/2$$



$\int_x^{\infty} p(x;\theta_1)\,\mathrm{d}x = c_1 \qquad \int_{-\infty}^{x} p(x;\theta_2)\,\mathrm{d}x = c_2$

where

$$\theta \in [\theta_1, \theta_2]\ ,\quad 1-(c_1+c_2) = CL$$

MC techniques can be used: grid over θ to find the values θ_1 and θ_2 satisfying these integrals

TMath:: BinomialI(p,N,x)

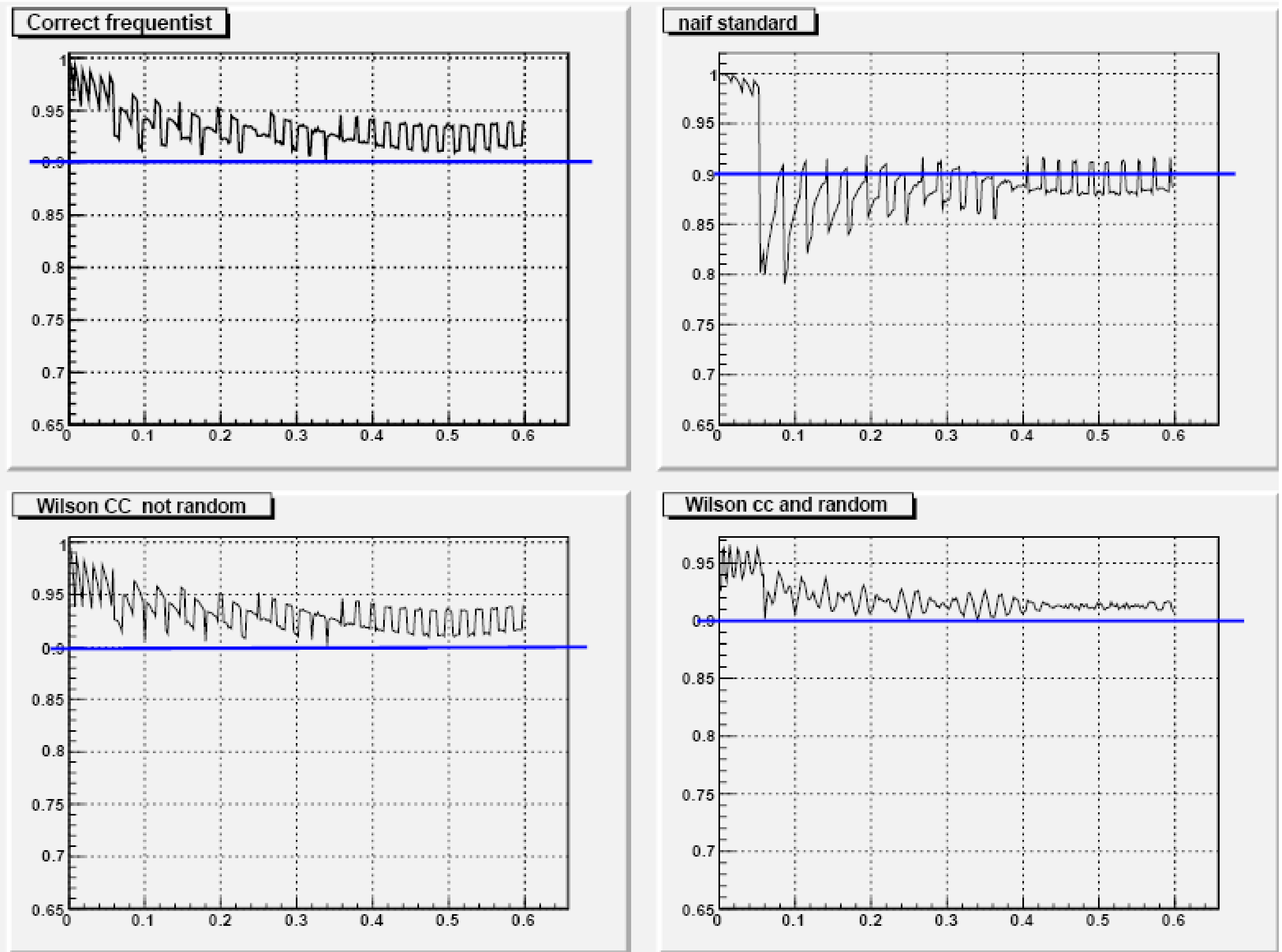
p_1 x/n p_1 p_2 x/n p_2

p

$f=k/n$

One expects $f\sim CL$

Simulate many x with a true p and check when the intervals contain the true value p. Compare this frequency with the stated CL



N=50
CL=0.90

n=10

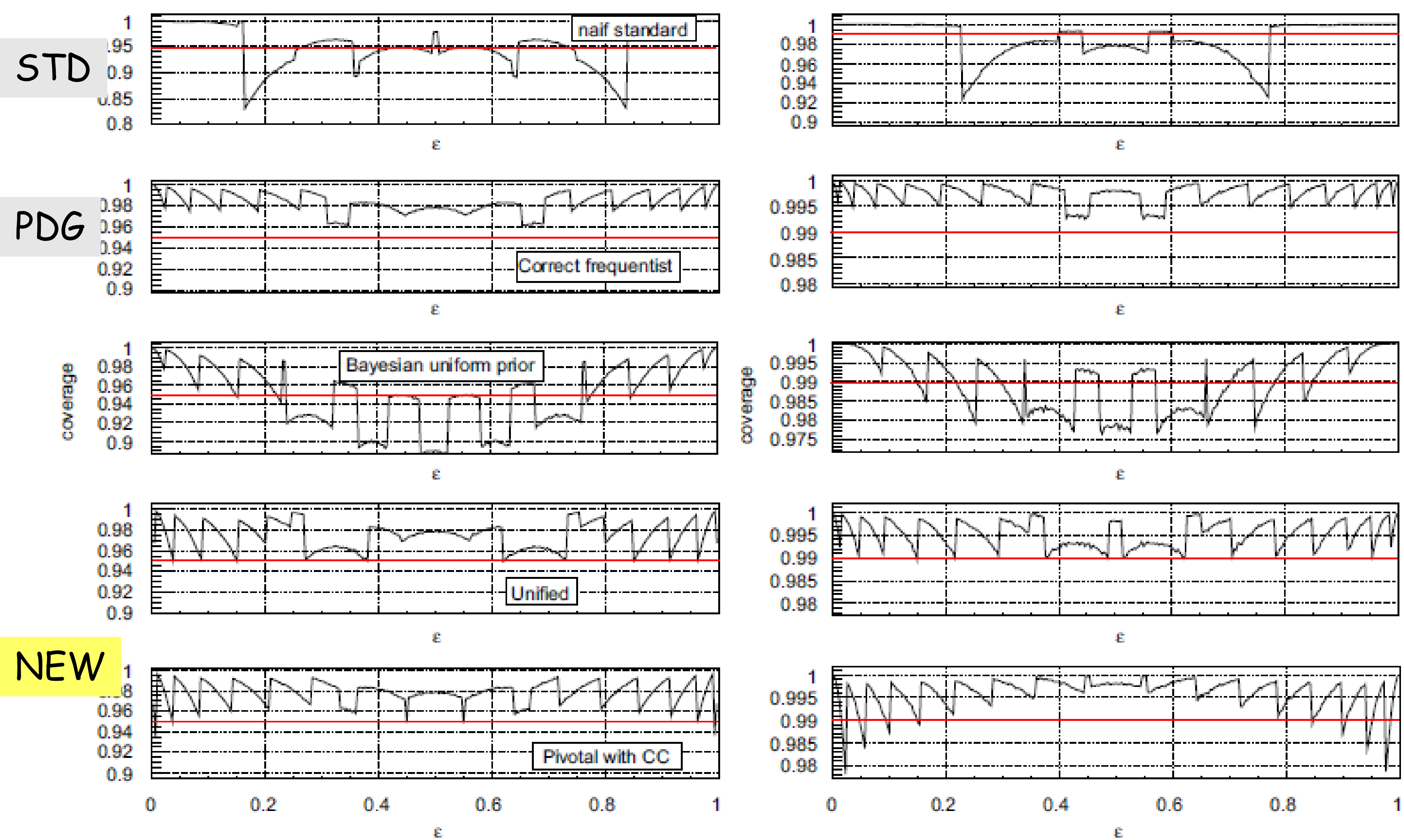


Fig. 12. Coverage of various estimation intervals for the efficiency ε in a binomial experiment with $n = 10$. The curves refer to a CL of 95% (left) and 99% (right). From top to bottom the coverages of the following intervals are reported: standard with CC of Eq. (50), classical frequentist of Eq. (39), Bayesian with uniform prior of Eq. (45), unified or likelihood ratio of Eq. (42), pivotal with CC of Eqs. (48) and (49).

n=80

A. Rotondi / Nuclear Instruments and Methods in Physics Research A 614 (2010) 105–118

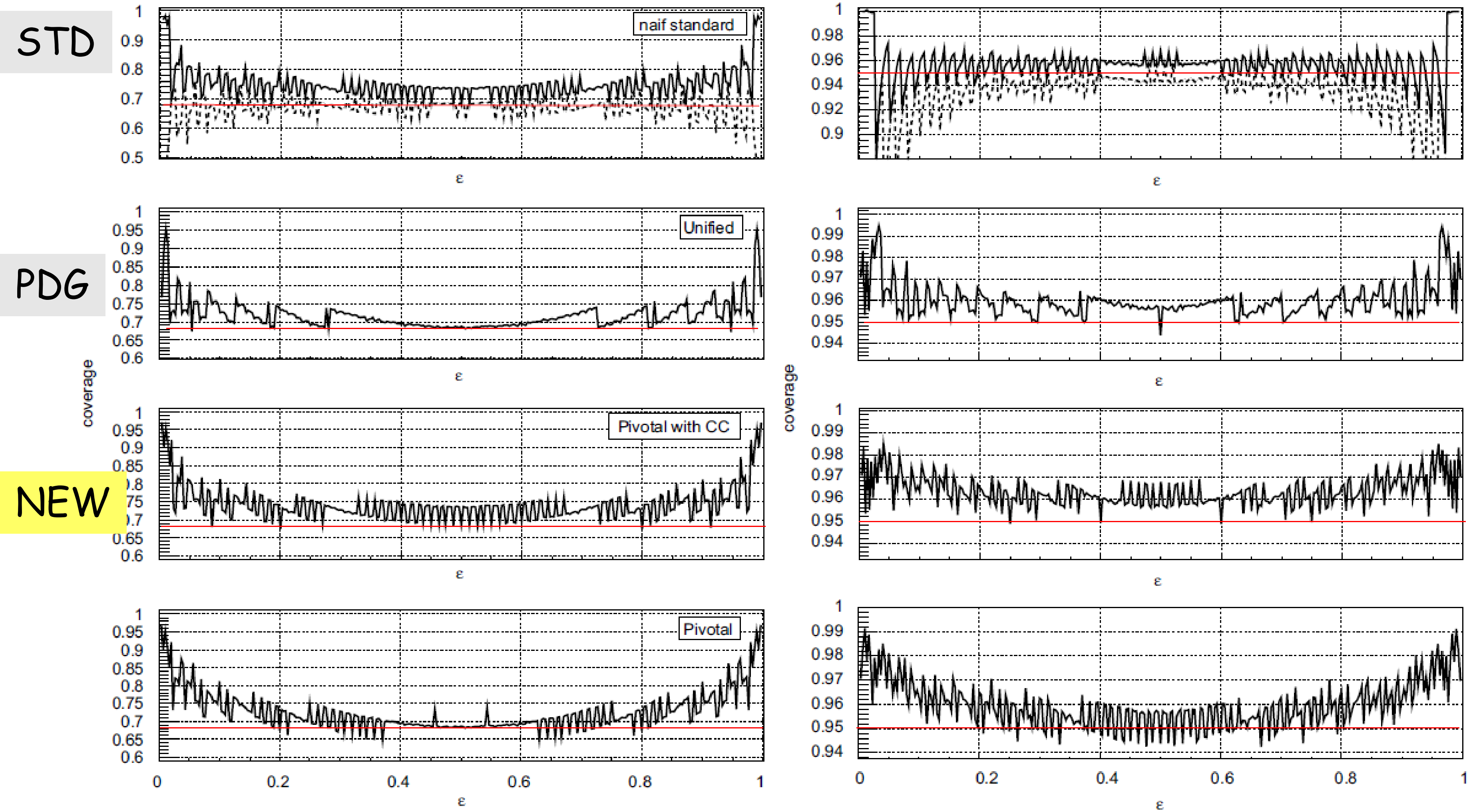
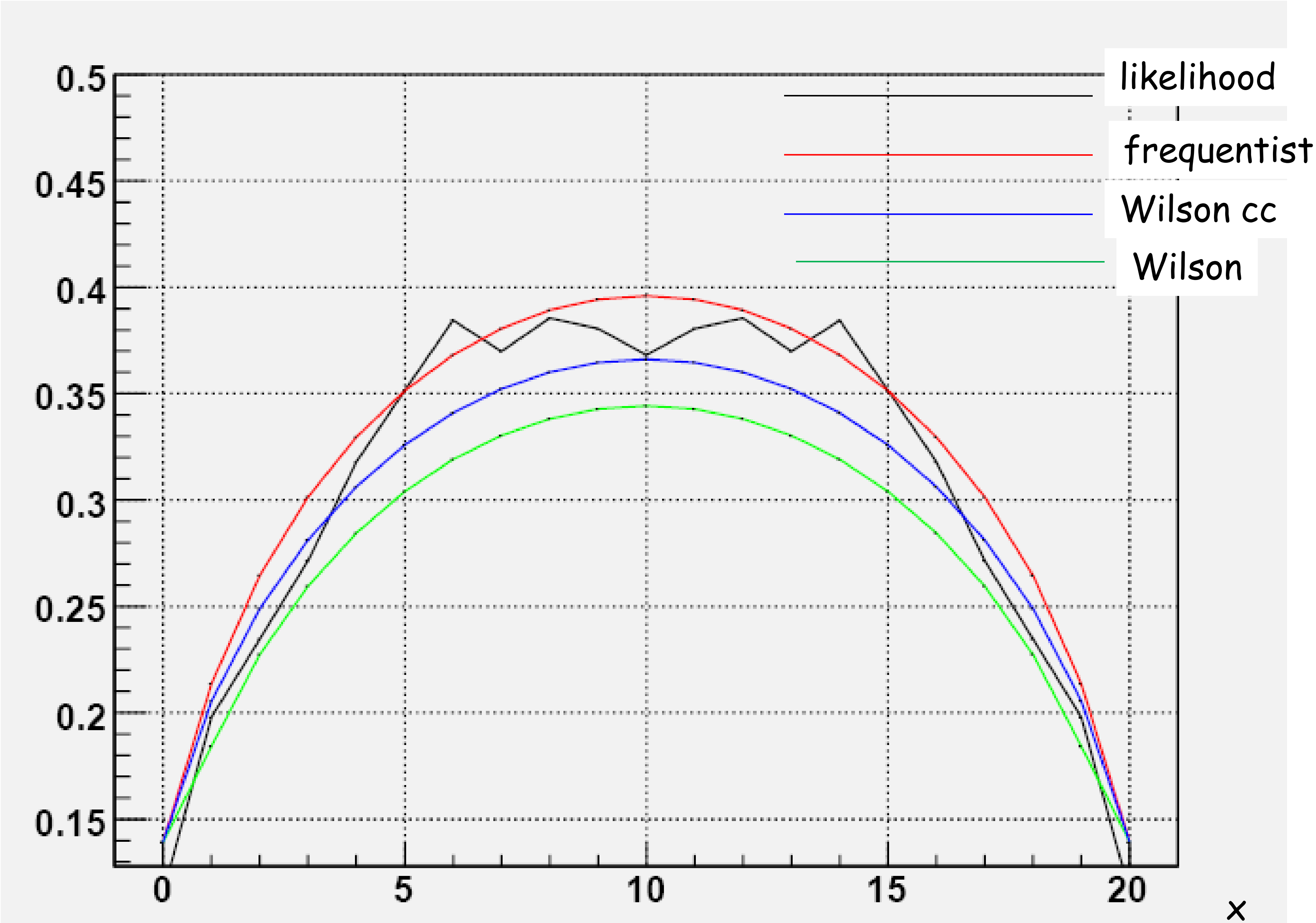


Fig. 14. As Fig. 12 for $n = 80$, $CL = 68.27\%$ (left) and $CL = 95\%$ (right). From top to bottom the coverages of the following intervals are reported: standard with CC values $c_- = c_+ = 0.5$ (full line) and without CC (short dashed line) of Eq. (50); unified or likelihood ratio of Eq. (42); pivotal from Eqs. (48) and (49) with CC values from Table 3; pivotal from Eqs. (48) and (49) without CC.

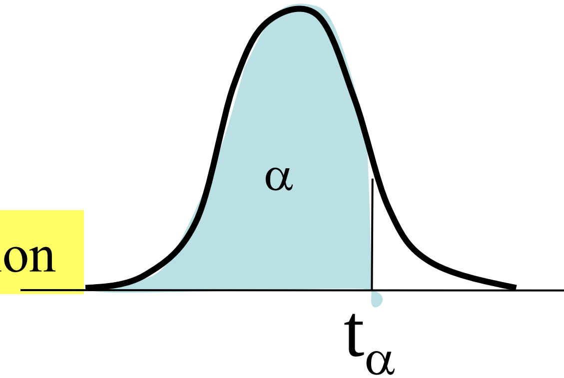
N=20 CL=0.90 Interval amplitude



Counting experiments: Poisson case

$$P\left\{\frac{|x-\mu|}{\sqrt{\mu}} \le t_\alpha \right\} \ge CL$$

t=1, area 84%
Quantile α=0.84
P[|f-p|<t σ]= 68%



t is the quantile of the normal distribution

$$\frac{(x-\mu)}{\sqrt{\mu}} = t_\alpha \;\rightarrow\; \mu = x + \frac{t_\alpha^2}{2} \pm t_\alpha \sqrt{x + \frac{t_\alpha^2}{4}}$$

Not used (why?)

$$\xrightarrow{\mu \approx x} \mu = x \pm t_\alpha \sqrt{x}$$

Standard in Physics

Counting experiments: new formula for the Poisson case

$$\frac{(x-\mu)}{\sqrt{\mu}} = t_\alpha \rightarrow \mu = x_\pm + \frac{t_\alpha^2}{2} \pm t_\alpha \sqrt{x_\pm + \frac{t_\alpha^2}{4}} \qquad x_\pm = x \pm 0.5$$

Wilson interval with Continuity correction gives the same results as …

$$\sum_{k=0}^{x} \frac{\mu_2^{\;x}}{x!} e^{-\mu_2} = \alpha / 2$$

$$\sum_{k=x}^{\infty} \frac{\mu_1^{\;x}}{x!} e^{-\mu_1} = \alpha / 2$$

Exact frequentist
Clopper Pearson (1934) (PDG)

STD

PDG

NEW

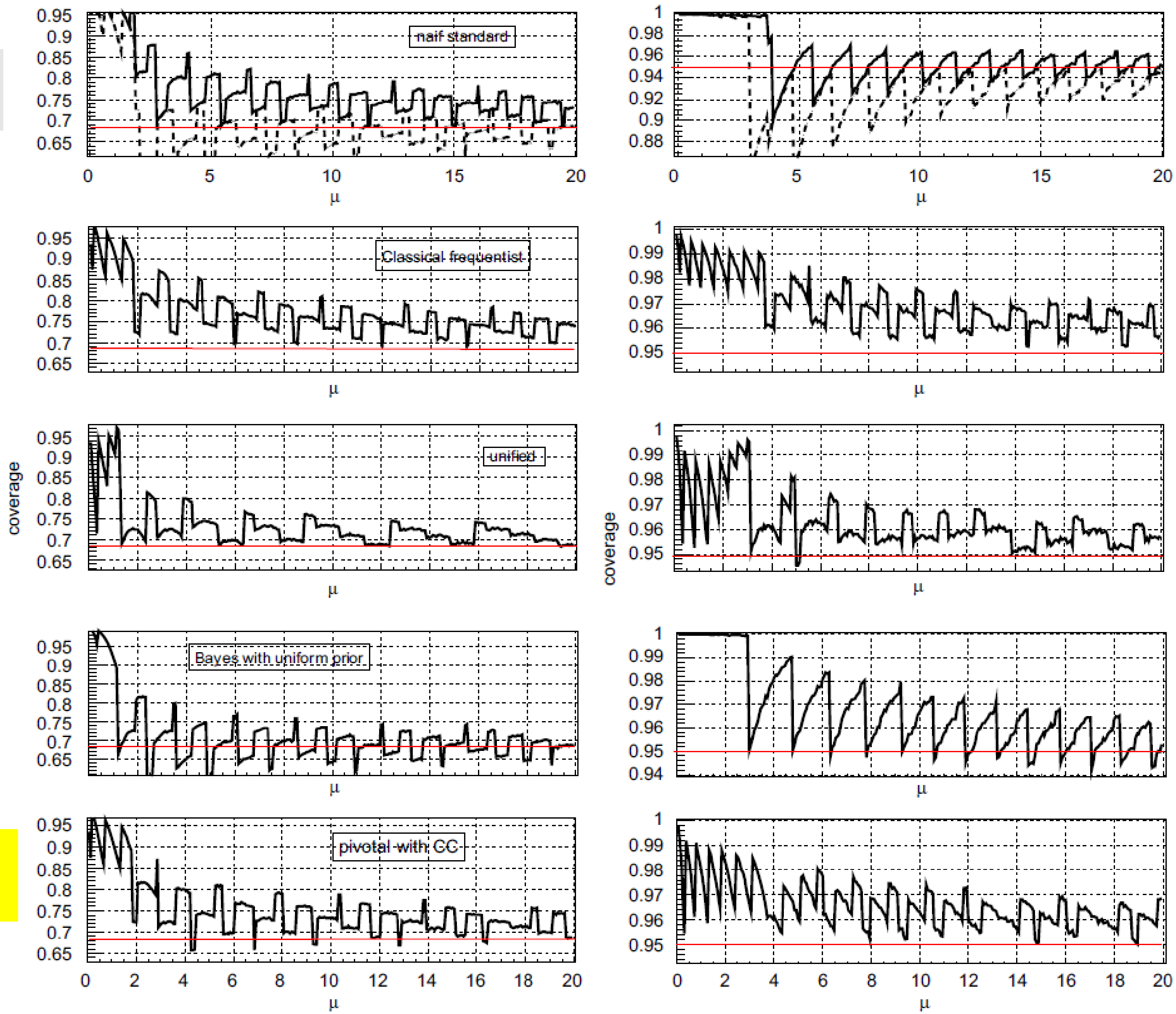
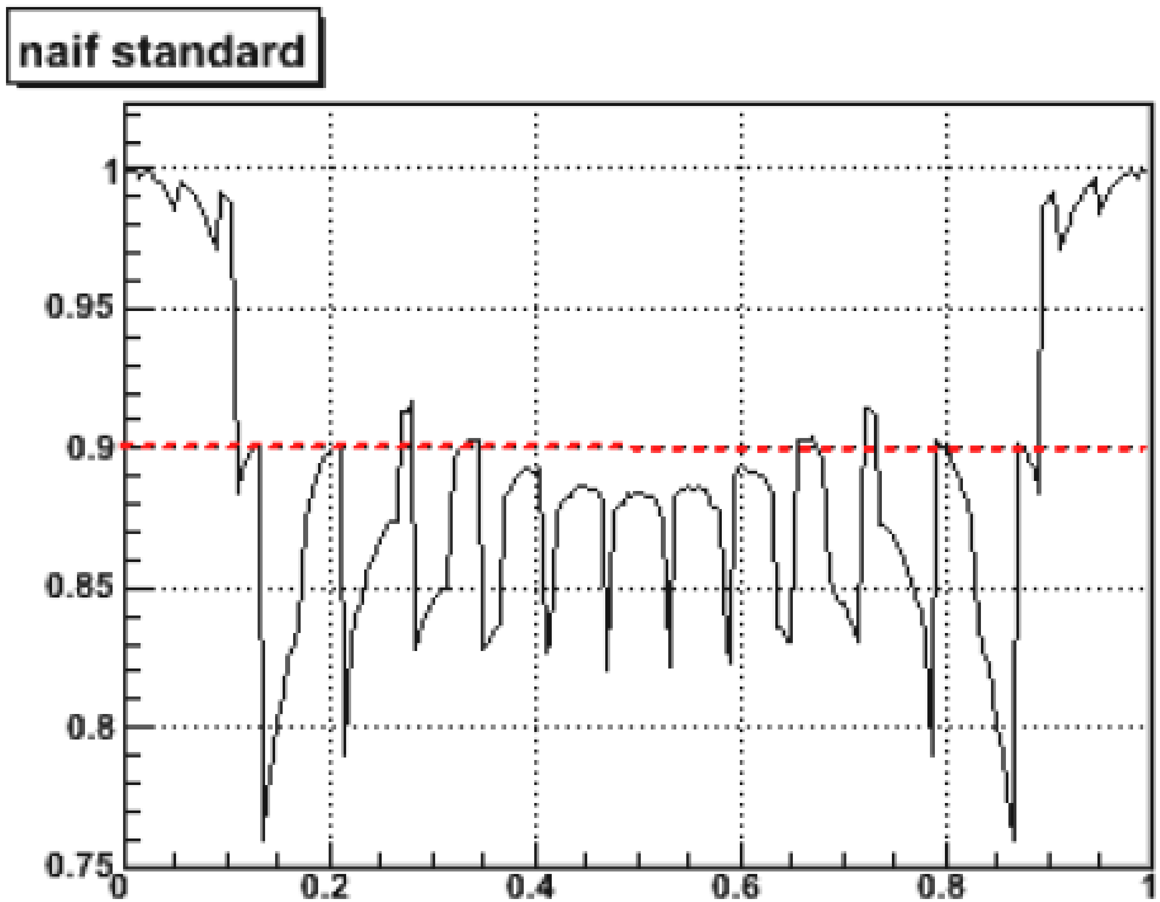


Fig. 1. Coverage of various estimation intervals in the case of Poisson distribution as a function of the true mean value μ. The curves refer to a CL of 68.27% (left) and 95% (right) marked with the horizontal full line. From top to bottom the coverages of the following intervals are reported: standard of Eq. (11) with CC (full line) and without CC (short dashed line); classical frequentist of Eq. (3); unified or likelihood ratio of Eq. (5); Bayesian with uniform prior of Eq. (8); pivotal with CC of Eq. (12).

Accurate at 2% for n>300

$$\xrightarrow{n>>1} \varepsilon = f \pm t_\alpha \sqrt{\frac{f(1-f)}{n}}$$

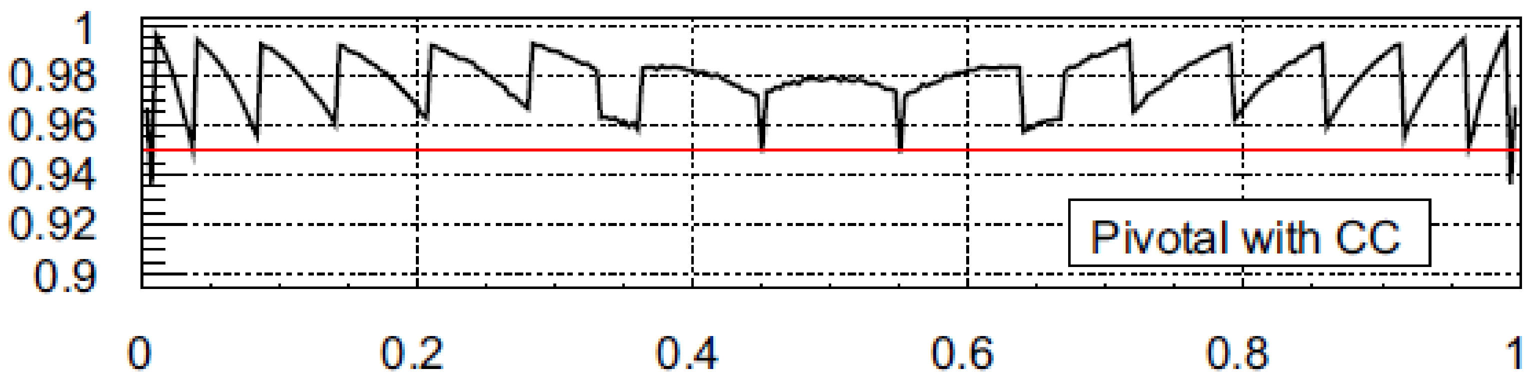


n=20

Accurate at 2% for n>10

$$\frac{f_\pm + \frac{t^2}{2n}}{\frac{t^2}{n}+1} \pm \frac{t_\alpha \sqrt{\frac{t^2}{4n^2} + \frac{f_\pm(1-f_\pm)}{n}}}{\frac{t^2}{n}+1}, \qquad f_\pm = \frac{x \pm 0.5}{n}$$

ε

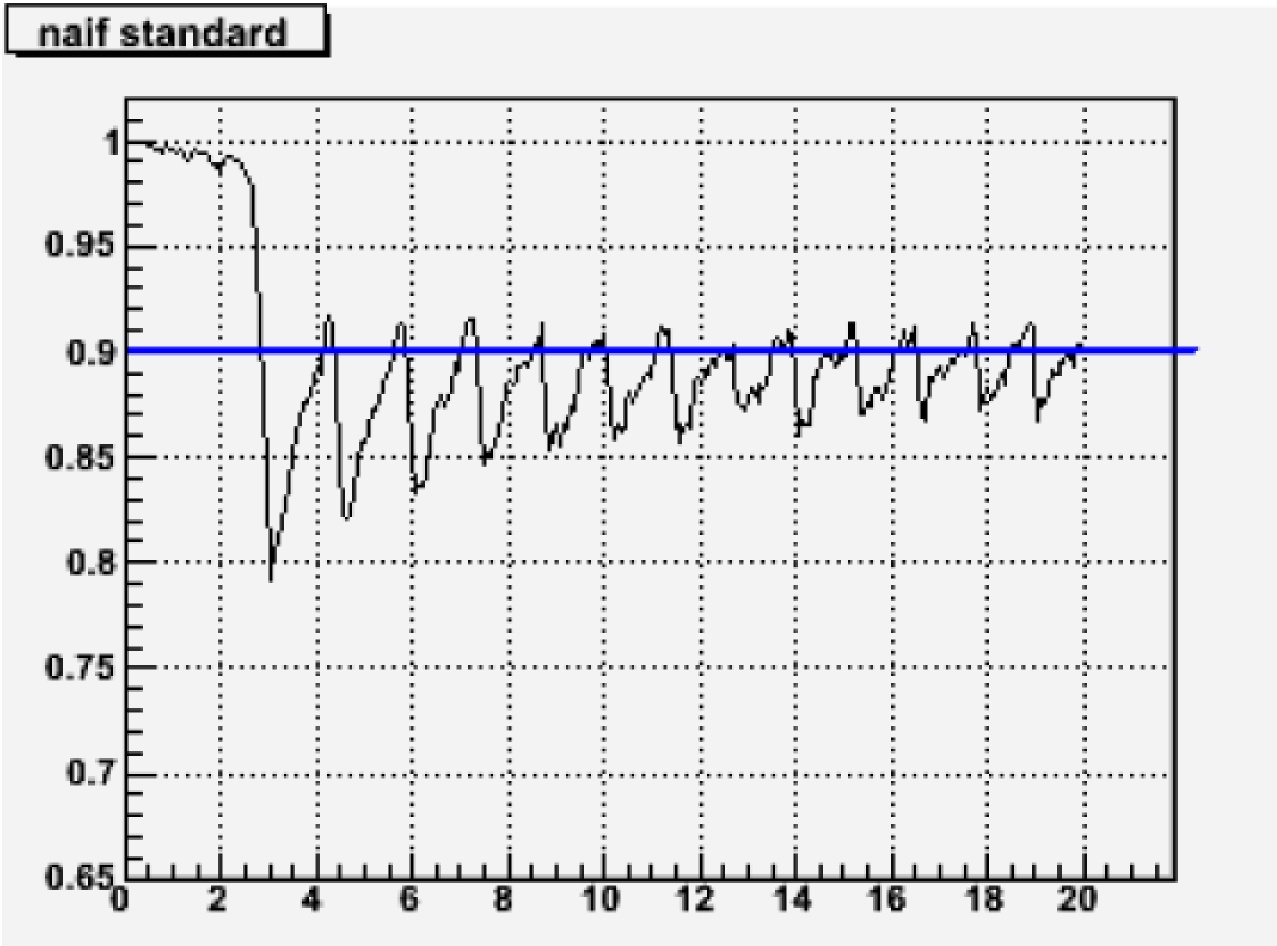


ε

n=10

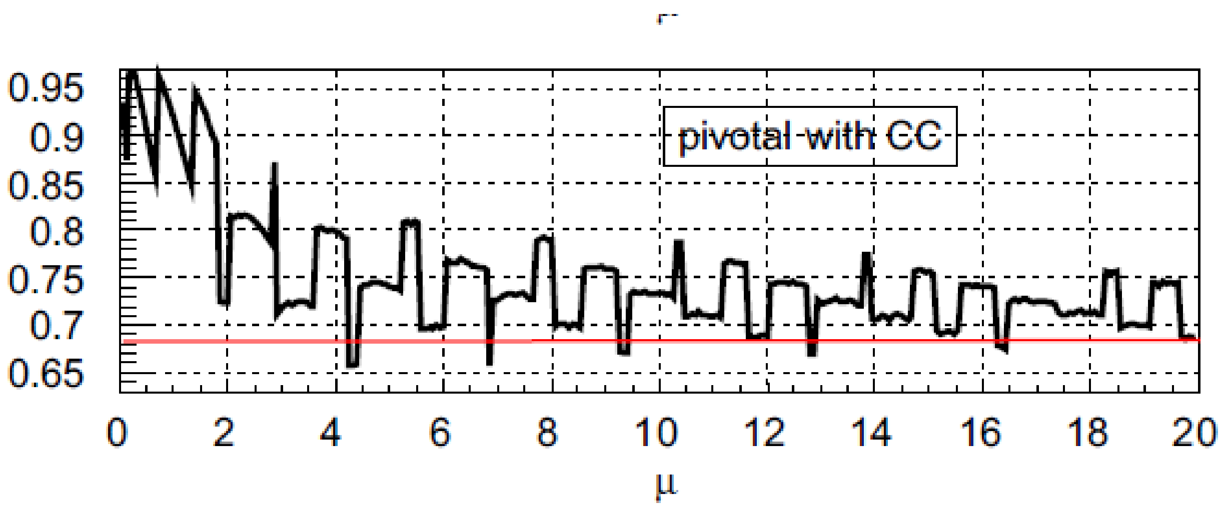
Accurate at 2% for x>80

$$\mu = x \pm t_\alpha \sqrt{x}$$



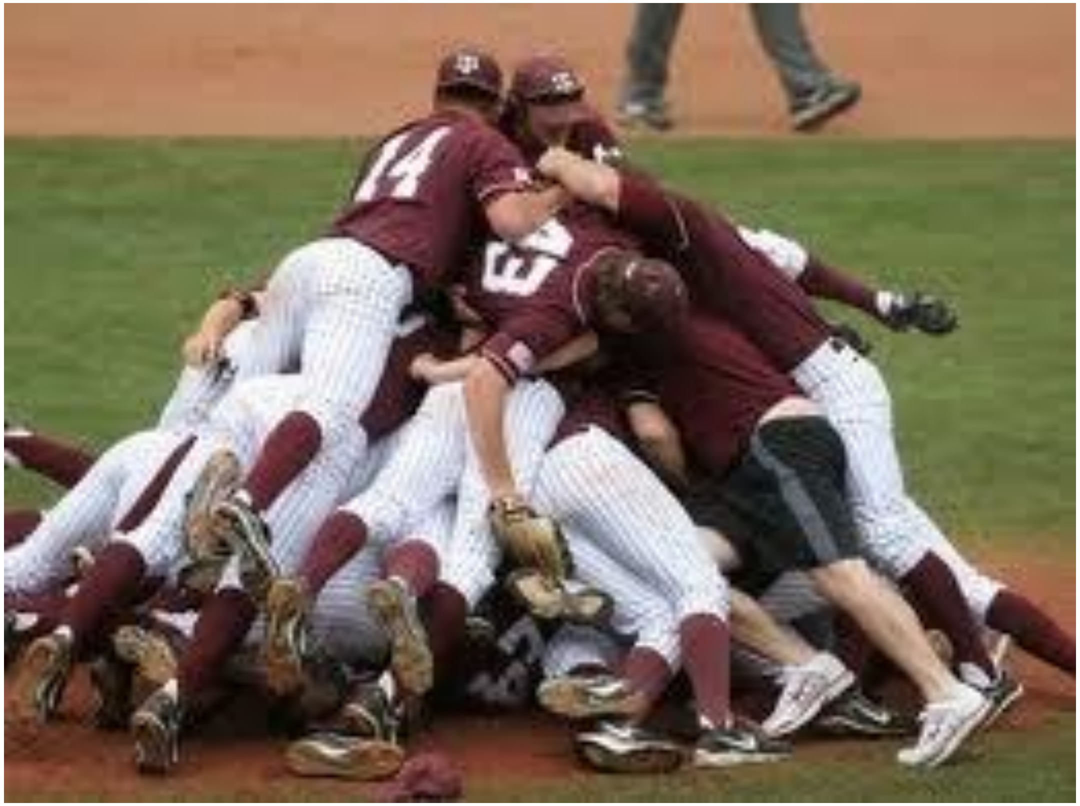
Accurate at 2% for x>0

$$\mu = x_{\pm} + \frac{t_\alpha^2}{2} \pm t_\alpha \sqrt{x_{\pm} + \frac{t_\alpha^2}{4}} \qquad x_{\pm} = x \pm 0.5$$



See A. Rotondi NIM A 614(2010)105
S. Costanza, A. Rotondi NIM A 669(2012)85

Pile-up



("dynamic" efficiency)

…..from textbooks

Dead time= minimum amount of time between two pulses so that they are recorded as separate pulses

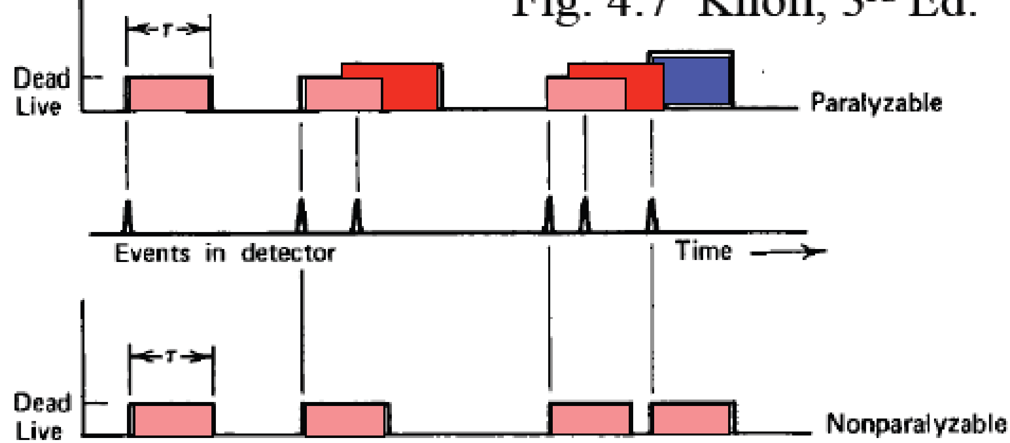
The efficiency of a system to measure and record pulses depends on the time taken up by all components of the signal processing. There are two classes of systems, those that require a fixed recovery time and those that don't.

Dead Time Models:

a) Paralyzable – detector system is affected by the radiation even if the signal is not processed. (a "slow" detector or electronics)

b) Nonparalyzable – fixed dead-time

Fig. 4.7 Knoll, 3rd Ed.

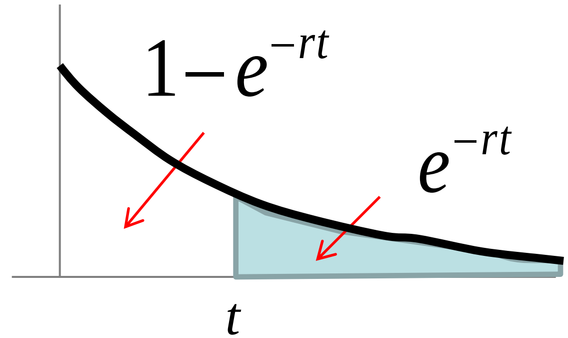


True rate: r or n (in text)

or m (in text) Dead-time: τ

….. from textbooks

a) Paralyzable (extending) dead time τ
a count is possible only after a dead time
from the last arrival



$$r_{obs} = r_{true}\, e^{-r_{true}\tau} \xrightarrow{r_{true}\,\tau<<1} r_{true} = \frac{r_{obs}}{1-r_{obs}\tau}$$

b) Non-Paralyzable dead time τ
A count is possible after a dead time
from the last count

Fraction dead = $R_{obs}\tau / T \;= r_{obs}\tau$

Loss rate = $r_{true}(r_{obs}\tau) \rightarrow r_{obs} = r_{true}(1-r_{obs}\tau) \rightarrow r_{true} = \dfrac{r_{obs}}{1-r_{obs}\tau}$

$$\rightarrow r_{true} = \frac{r_{obs}T}{T-r_{obs}\tau\, T} = r_{obs}\frac{T}{T_L}$$

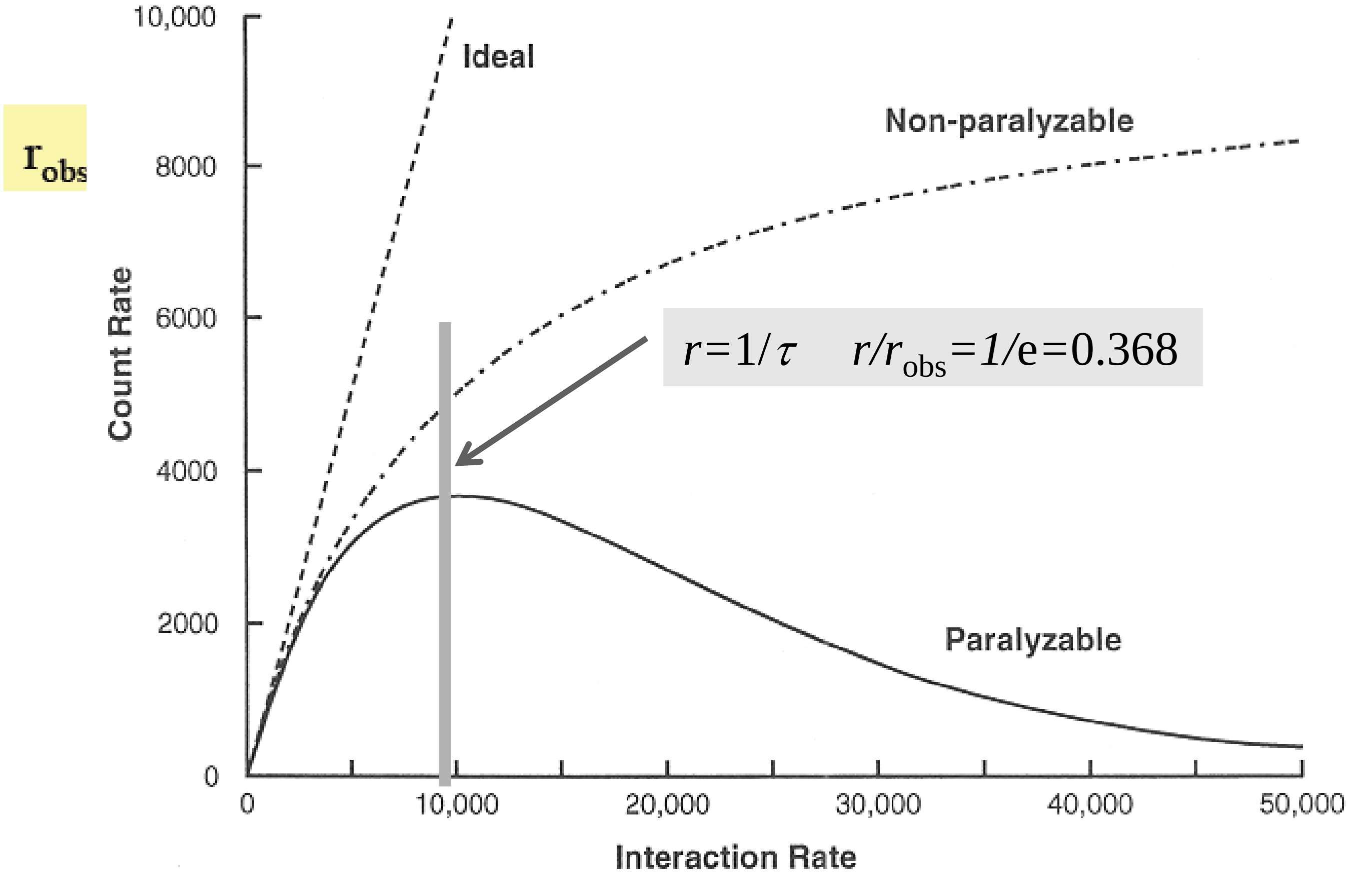
….. from textbooks

Dead Time Models:

a) Paralyzable – $r_{obs} = r\, e^{-r\tau}$

b) Nonparalyzable – $r = r_{obs} / (1 - r_{obs}\, \tau)$

Fig. 4.8 Knoll, 3rd Ed.



$$I = I_0 e^{-\tau I_0}$$

$$\xrightarrow{I_0 \tau = P t_d / t_l} \frac{I_0}{I} = e^{P t_d / t_l}$$

$$\Rightarrow \ln I = \ln I_0 - P \frac{t_d}{t_l}$$

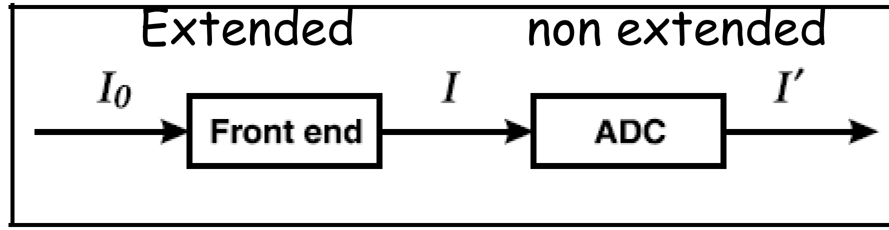
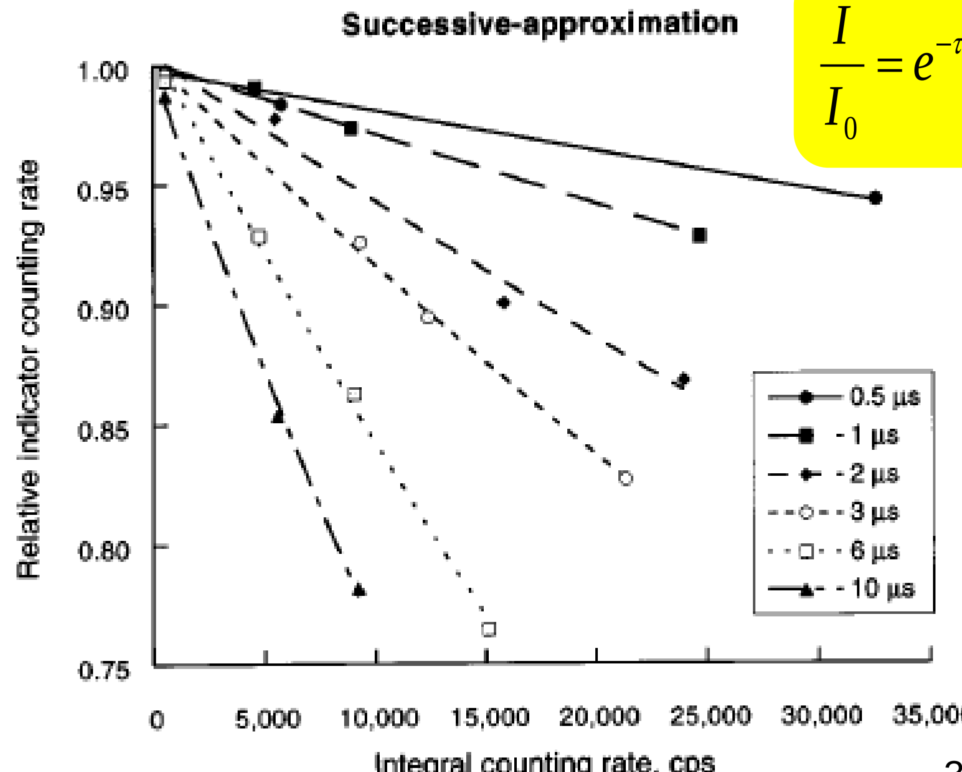


Figure 4 Mathematical model of a counting system with a decaying source feeding extending (front-end) and non-extending (ADC) dead times in series.



$$\frac{I}{I_0} = e^{-\tau I_0}$$

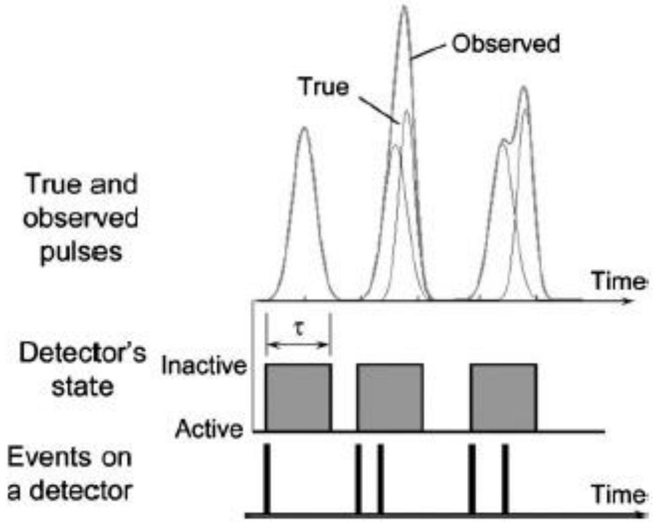
Beam (trigger) pile-up

Count loss

Pile-up

spectrum distortion

Interaction (detector) pile-up



An example of a Xilinx Spartan 6 FPGA programming/evaluation board

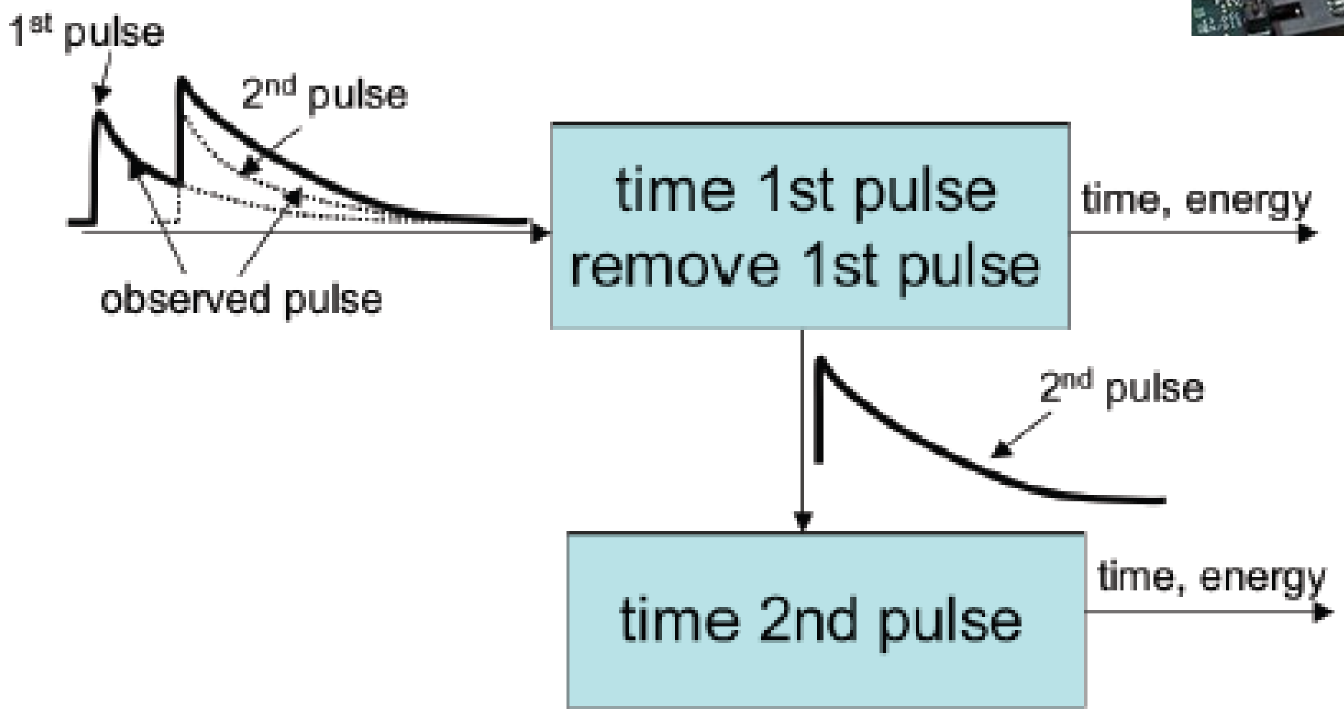


Fig. 1. Block diagram of the overall pulse pileup correction algorithm. Note that number of stages is dependent on the expected rate of pileup.

Ho to deal with pile-up?

- to measure **dead time and live time**

- with the **Time-To-Count technique**, the detector is armed at the same time a counter is started. When a strike occurs, the counter is stopped for a time longer than the supposed dead time. **The rate r is thus measured, not estimated: <t>=1/r.**

- to use a **pile-up rejection** system
- to use **digital methods** in ADC signal processing

Unfolding Methods

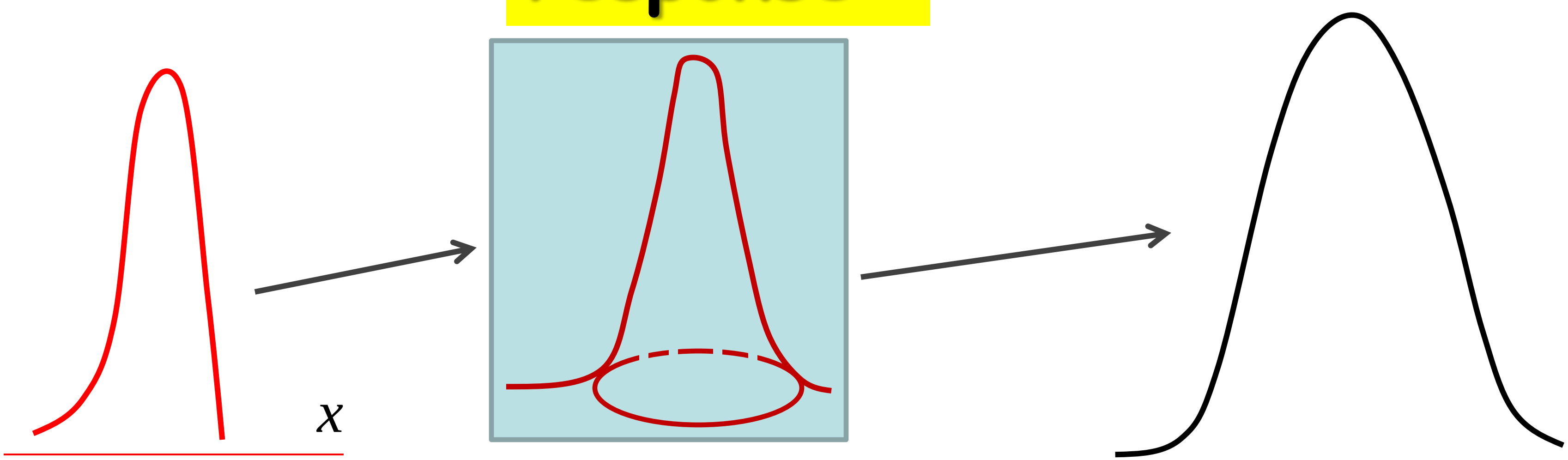
$$g=f*\delta$$

Folding is a common process in physics

signal

Apparatus response

Observed signal



$f(x)$

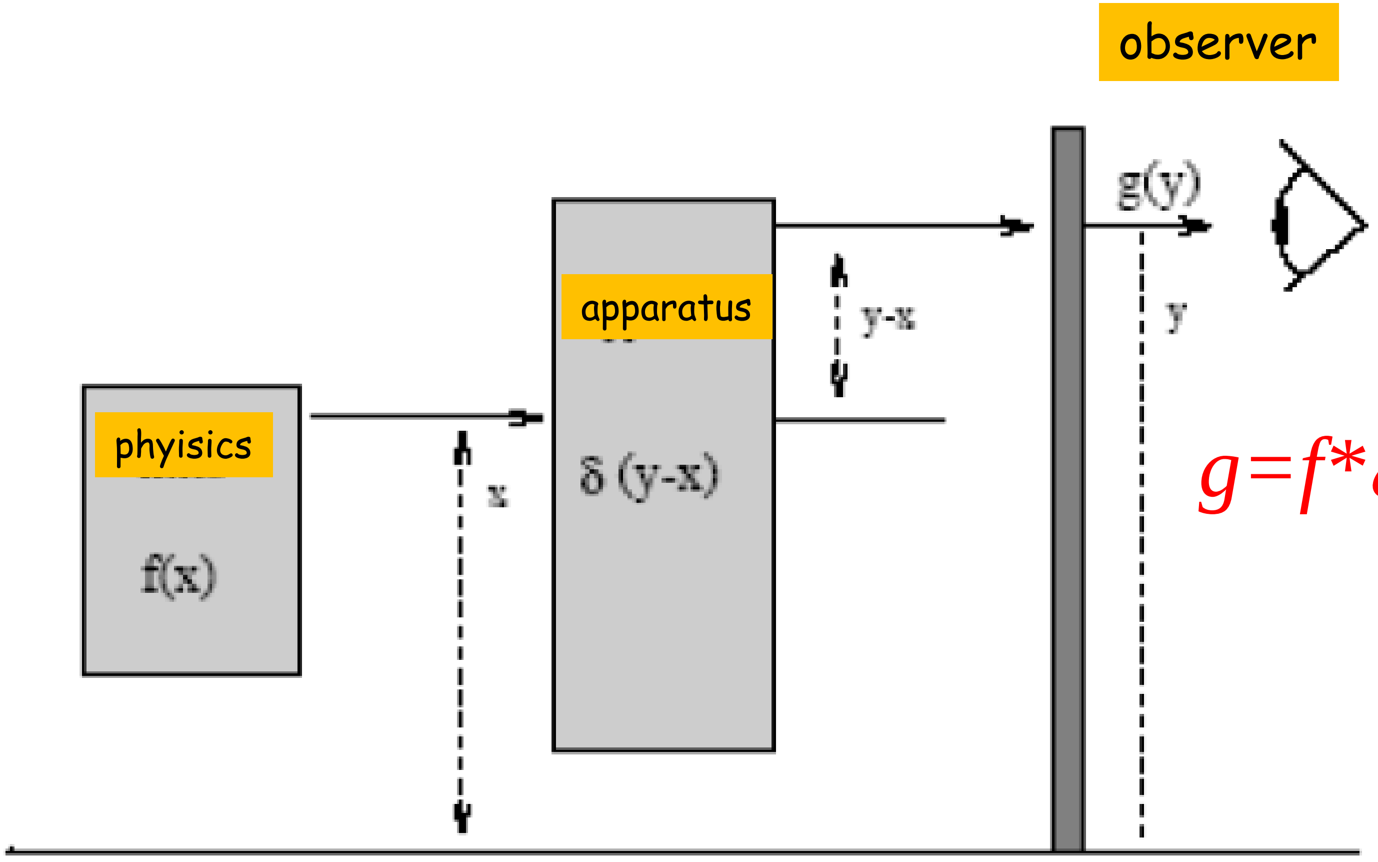
$\delta(x,y)$

$g(y)=\int f(x)\,\delta(x,y)\,\mathrm{d}x$

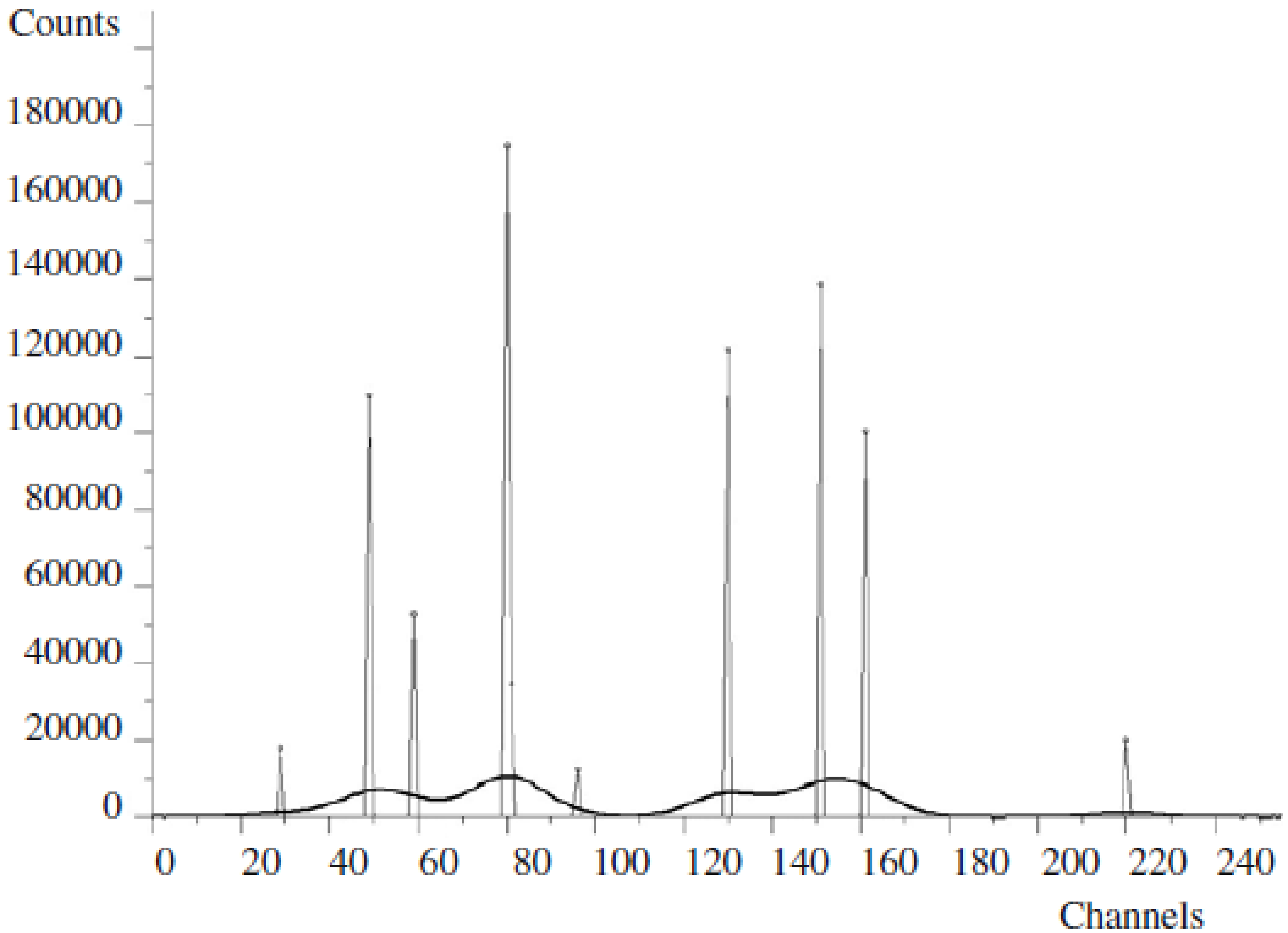
$g_i=\sum_j \Delta_{ij} f_j$

Convolution is a linear folding

$$g(y) = \int f(x)\,\delta(y-x)\,\mathrm{d}x$$



$g = f * \delta$


Counts
180000
160000
140000
120000
100000
80000
60000
40000
20000
0
0
20
40
60
80
100
120
140
160
180
200
220
240
Channels

Fourier Techniques

$$f(x) = \int F(t)\, \mathrm{e}^{2\pi i\, xt}\, \mathrm{d}t$$

Convolution:

$$f(x) = \int g(y) \delta(x - y)\, \mathrm{d}y$$

$$\int F(t)\, \mathrm{e}^{2\pi i xt}\, \mathrm{d}t = \int G(t)\, \mathrm{e}^{2\pi i\, yt}\, \Delta(t)\, \mathrm{e}^{2\pi i\,(x-y)t}\, \mathrm{d}t$$

$$\int F(t)\, \mathrm{e}^{2\pi i\, xt}\, \mathrm{d}t = \int G(t)\, \Delta(t)\, \mathrm{e}^{2\pi i\, xt}\, \mathrm{d}t \rightarrow F(t) = G(t)\, \Delta(t)$$

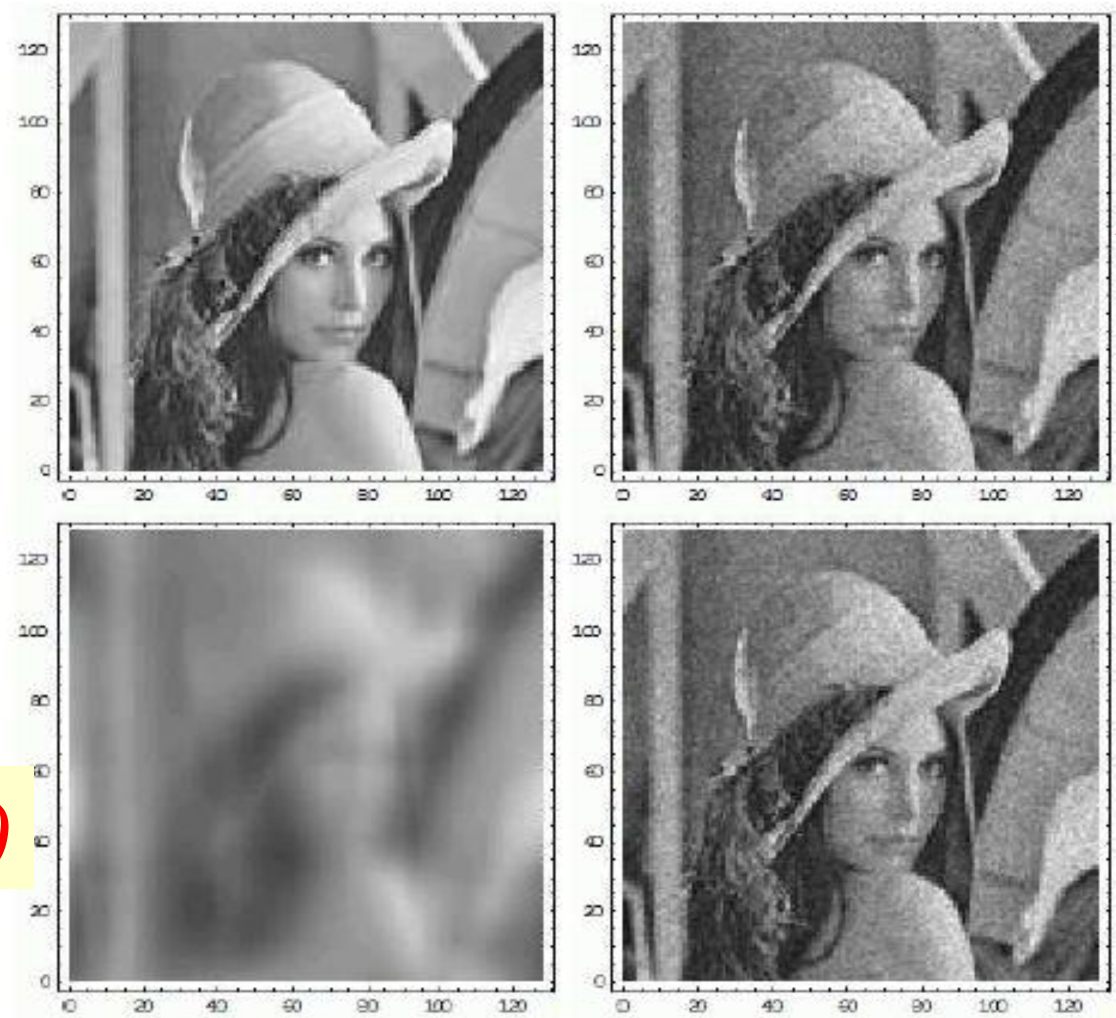
original

Poisson statistics

$f+R$

Gaussian smearing

$g=\delta*(f+R)$



Fourier restored

Figure 11: **Lena restored by FFT: The original image (top left) is sampled with Poisson statistics (top right) and smeared with a 2D 10-bins Gaussian PSF (bottom left): the Fourier restored image (bottom right) is similar to the Poisson sampled image. In this case the noise term N is neglected.**

original

Gaussian smearing

$g=f*\delta$

Poisson statistics

$g=f*\delta+R$

Fourier (un)restored

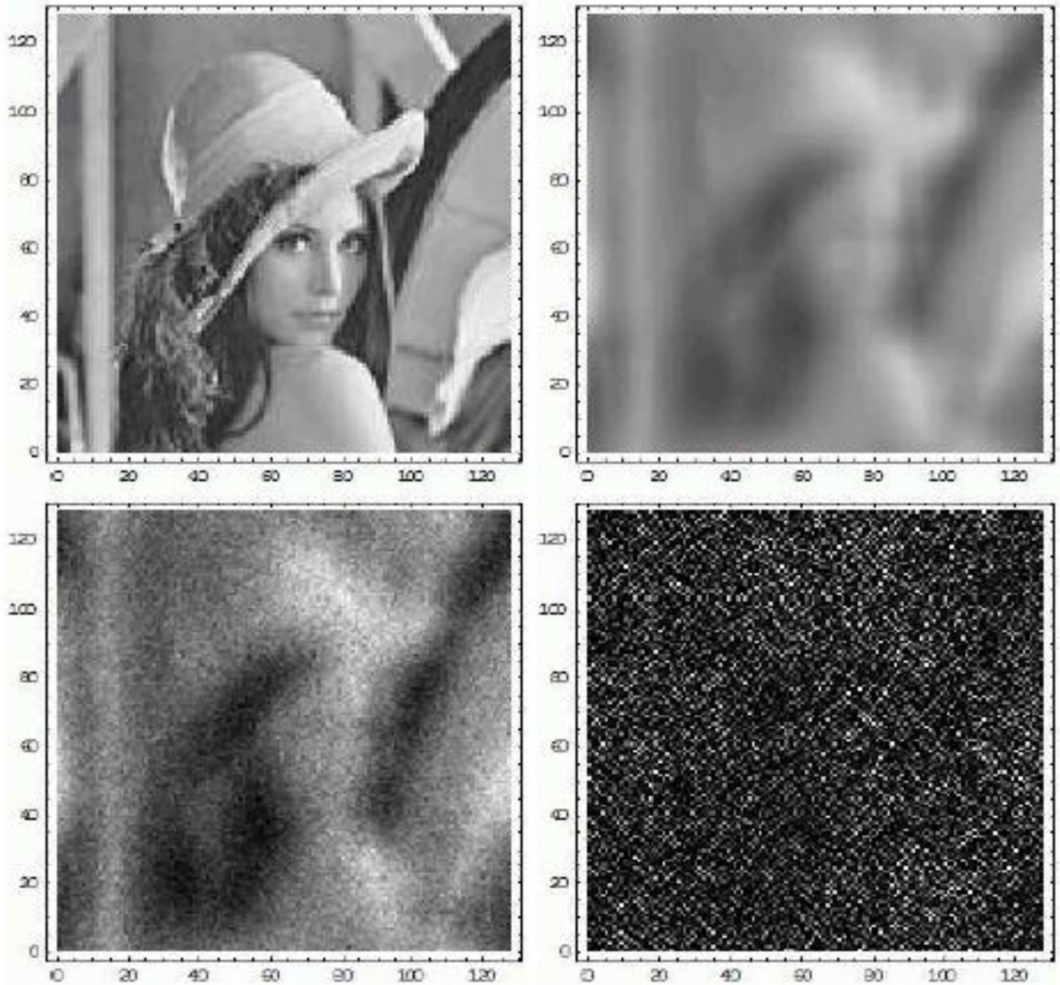


Figure 12: **Lena not restored by FFT: In this case the noise term N is not ignored: the original image (top left) is smeared with a 2D 10-bins Gaussian PSF (top right) and the result is sampled with Poisson statistics (bottom left): the Fourier restored image (bottom right) cannot recover the information lost in the noise. Another approach, statistical in nature, is required.**

The problem with fluctuations

original

Gaussian smearing

Poisson statistics

Fourier (un)restored

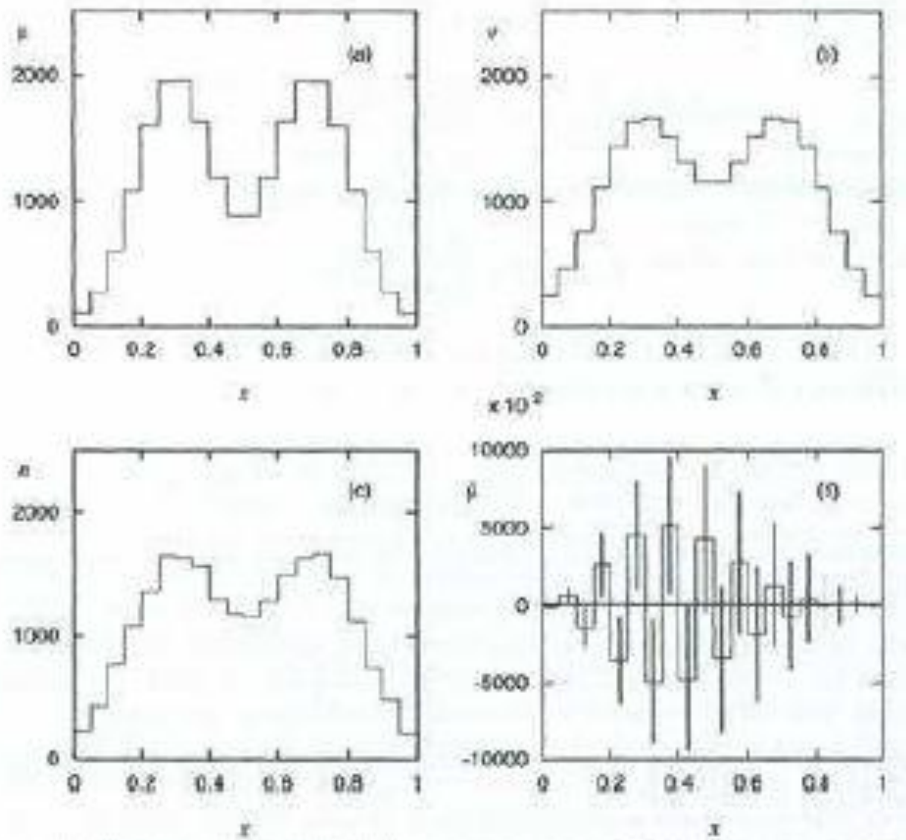


Fig. 11.1 (a) A hypothetical true histogram μ, (b) the histogram of expectation values $\nu = R\mu$, (c) the histogram of observed data n, and (d) the estimators $\hat{\mu}$ obtained from inversion of the response matrix.

A reminder…..
(Bayes theorem)

$$\left.\begin{aligned} P(A|B) &= \frac{P(A\cap B)}{P(B)} \\ P(B|A) &= \frac{P(A\cap B)}{P(A)} \end{aligned}\right\} \Rightarrow P(A|B)P(B) = P(B|A)P(A)$$

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

1D Unfolding: Bayesian iterative algorithm

Alice, Atlas,(RooFit), PHYSTAT2011

$$P(m_i)=\sum_j P(m_i\,|\,t_j)P(t_j)$$

$$P(t_i)=\sum_j P(t_i\,|\,m_j)P(m_j)$$

***t*= true, *m*=measured**

$$P(t_i\,|\,m_j)=\frac{P(m_j\,|\,t_i)P(t_i)}{\sum_k P(m_j\,|\,t_k)P(t_k)}$$

resolution

$$P^{(k+1)}(t_i)=\sum_j P(t_i\,|\,m_j)P(m_j)=\sum_j \frac{P(m_j\,|\,t_i)P^{(k)}(t_i)P(m_j)}{\sum_n P(m_j\,|\,t_n)P^{(k)}(t_n)}$$

$$P^{(k+1)}(t_i)=\sum_j \frac{P(m_j\,|\,t_i)}{\varepsilon_i}\,P^{(k)}(t_i)\frac{P(m_j)}{P^{(k)}(m_j)}$$

efficiency

1D Unfolding: Bayesian algorithm

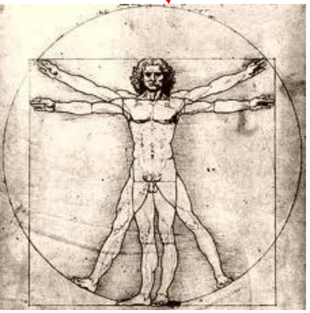
$$\sum_j P(m_j \mid t_i) = \begin{cases} 1 \\ or \\ \varepsilon_i \end{cases}$$

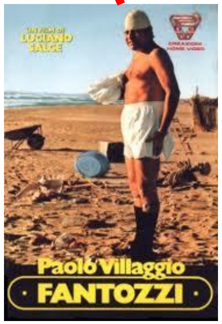
$$t_i^{(k+1)} = t_i^{(k)} \sum_j P(m_j \mid t_i) \frac{m_j}{m_j^{(k)}} \frac{1}{\varepsilon_i}$$

Unfolding matrix

$$t_i^{(k+1)} = \sum_j M_{ij} m_j ,$$

$$M_{ij} = \frac{1}{\varepsilon_i} P(m_j \mid t_i) \frac{t_i^{(k)}}{m_j^{(k)}}$$





$$m_j^{(k)} = \sum_p P(m_j \mid t_p)\, t_p^{(k)}$$

Efficiency and resolution

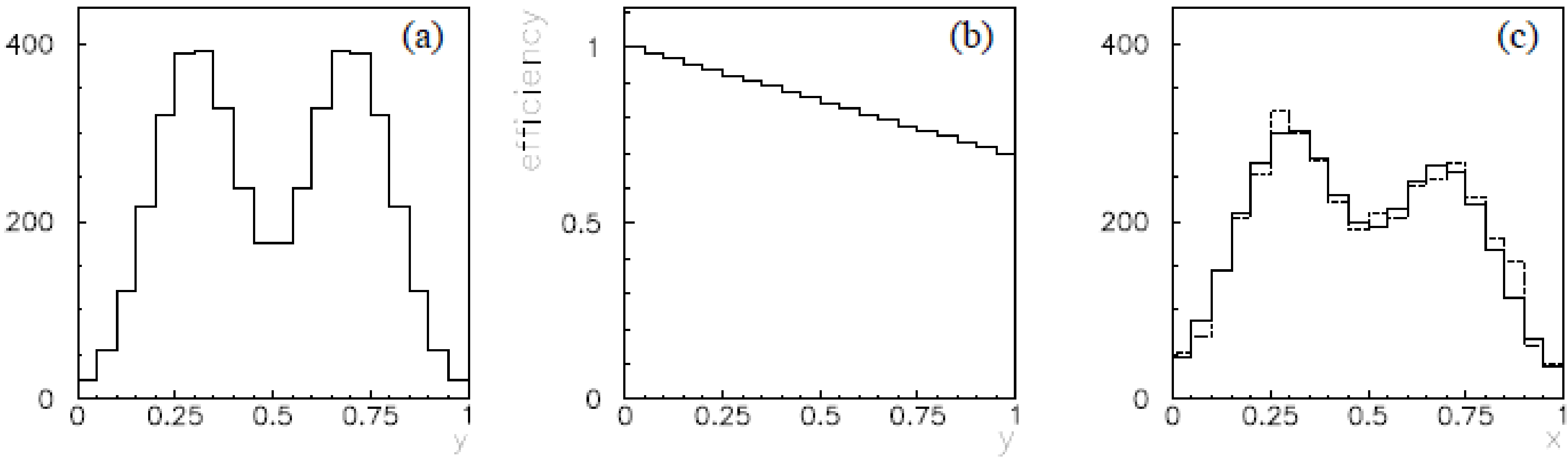
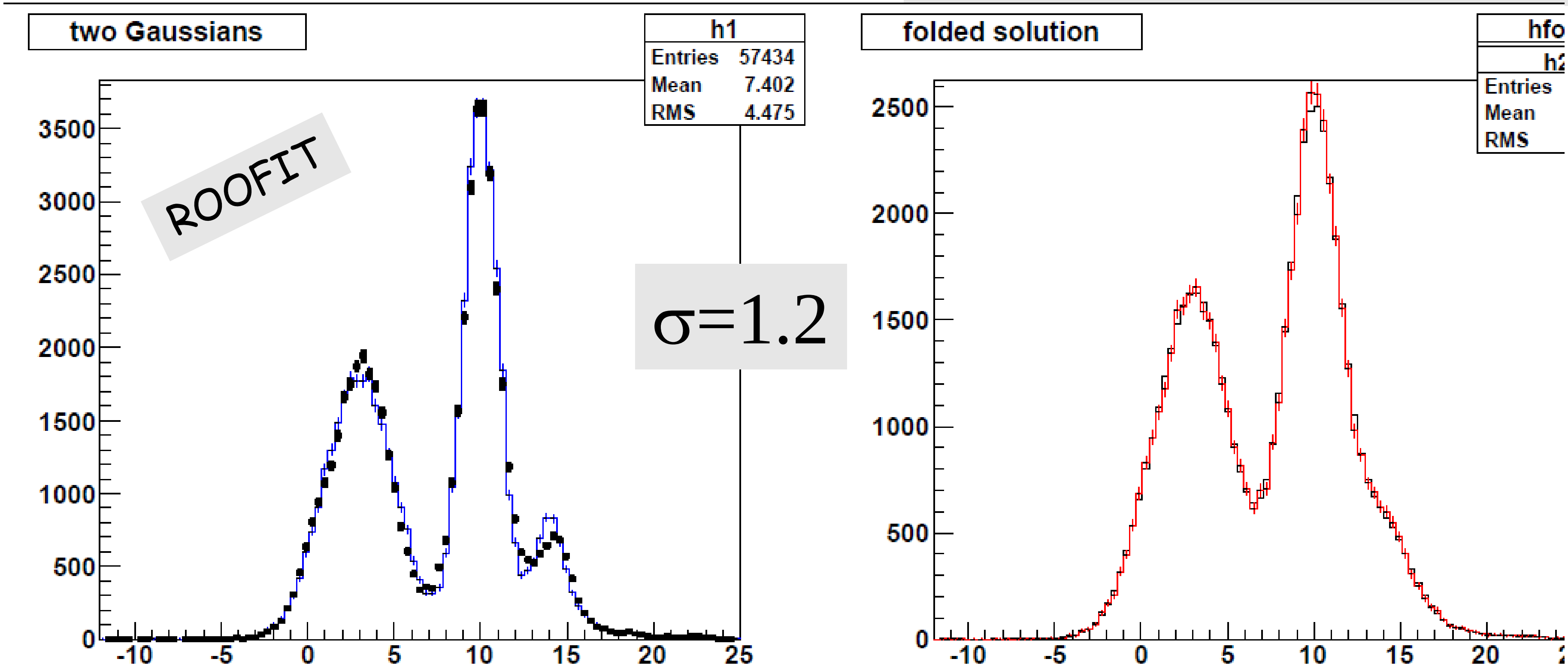


Fig. 1: Illustration of ingredients for unfolding: (a) a 'true histogram' μ, (b) a possible set of efficiencies ε, and (c) the observed histogram n (dashed) and the corresponding expectation values ν (solid).

$$\sum_j P(m_j \mid t_i) = \begin{cases} 1 \\ or \\ \varepsilon_i \end{cases}$$

Bayesian algorithm- starting solution: the data

Gaussian spread σ = ± 4 channels



true

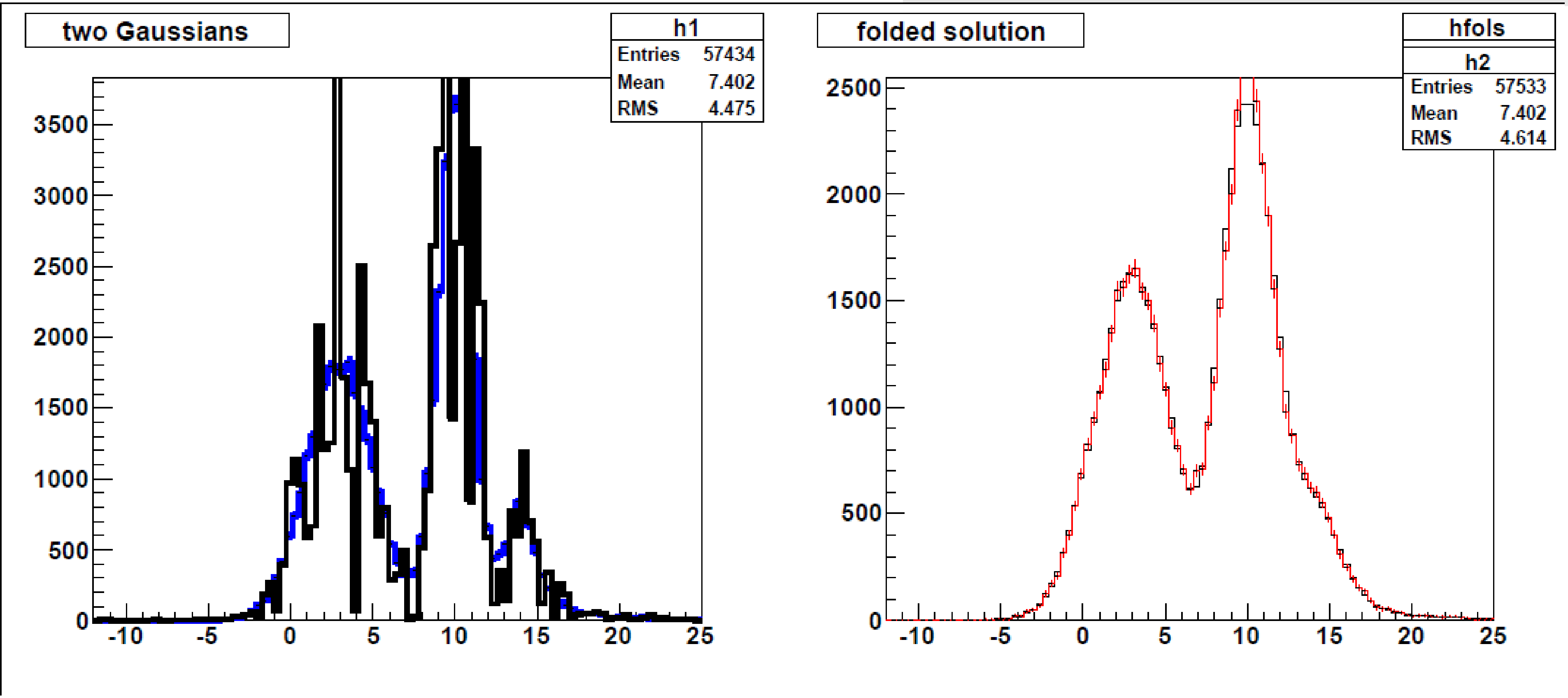
unfolded

Folded with the solution

data

Bayesian algorithm- starting solution: uniform

Gaussian spread $\sigma = \pm\, 4$ channels



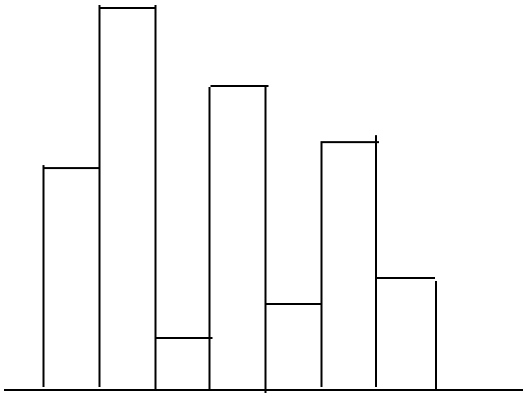
true

unfolded

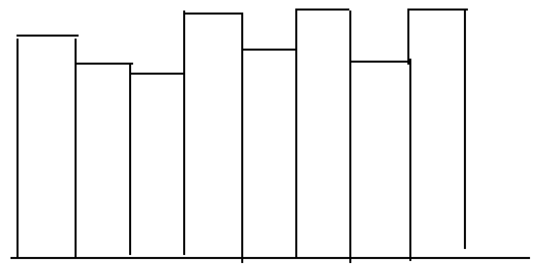
Folded with the solution

data

Why these oscillations?:



spike solution
HIGLY PROBABLE



smooth solution
UNLIKE

The smeared distributions of two input distributions cannot be distinguished if they agree on a large scale of x but differ by oscillations on a "microscopic" scale much smaller than the experimental resolution

$$F * \Delta = F * (\Delta + N) \quad \text{if} \quad F * N = 0$$

many solutions give a good χ^2

the spike ones are more probable!

Cure: to add to χ^2 an empirical regularization term $C[p]$.

$$\chi^2 \to \alpha\, \chi^2 + C[P(\text{true})]$$

or

$$\chi^2 \to \chi^2 + \alpha C[P(\text{true})]$$

or
to increase the DoF by using a parametric model

$$P(\nu \mid \mu) P(\mu) \to P(\nu \mid \mu')$$

Poisson likelihood fit with penalty regularization

for a single bin i with expectation d_i:

$$\ln P(m;\mu) = \ln \frac{\mu^m}{m!} e^{-\mu}$$

$$\ln L_i = m_i \ln \mu_i - \mu_i = m_i \ln \sum_j P(m_i \mid t_j) t_j - \sum_j P(m_i \mid t_j) t_j$$

for the histogram with penalty term R:

$$\ln L = \sum_k \left[m_k \ln \sum_j P(m_k \mid t_j) t_j - \sum_j P(m_k \mid t_j) t_j \right] - R$$

Frequently used penalty term

$$R = \alpha \sum_{i=2}^{N-1} (2t_i - t_{i-1} - t_{i+1})^2$$

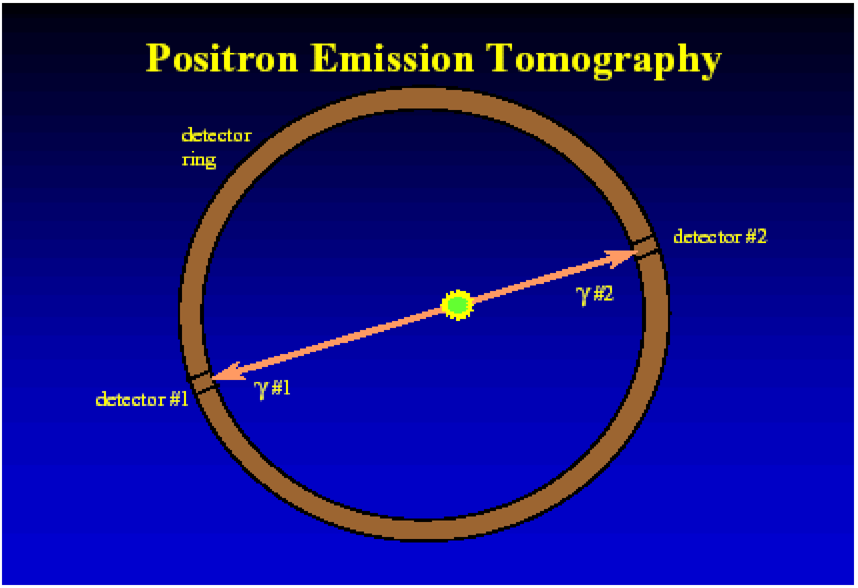
Statistical effects

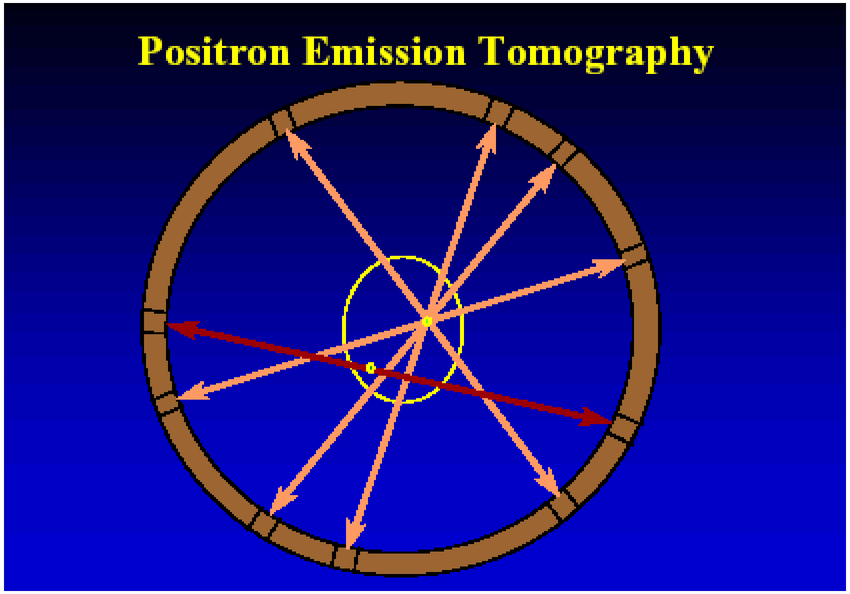
- background $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_N)$

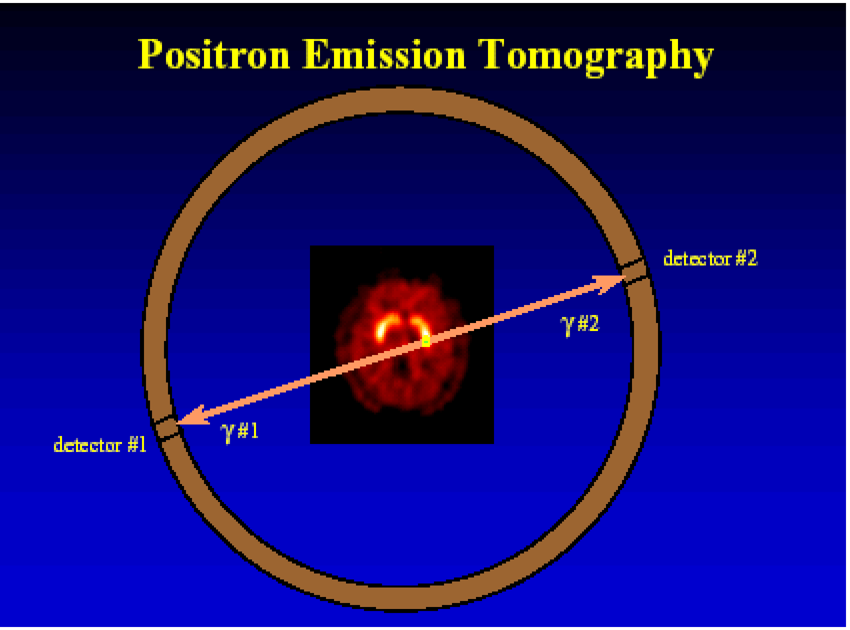
$$\nu_i = \sum_j R_{ij}\mu_j + \beta_i \longrightarrow \nu = R*\mu + \beta$$

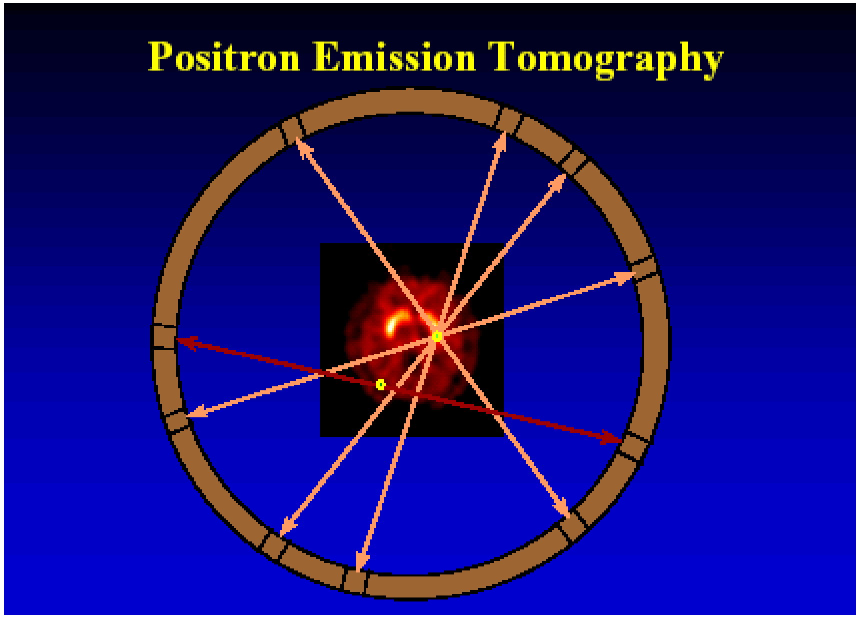
- The number of observed events in the case of random processes of detection:

$$n_i = \frac{\nu_i^{n_i}}{n_i!}\mathrm{e}^{-\nu_i}$$


Positron Emission Tomography
detector ring
detector #2
γ #2
detector #1
γ #1


Positron Emission Tomography


Positron Emission Tomography
detector #2
γ #2
detector #1
γ #1


Positron Emission Tomography

$$N_{ij}(\text{th}) = N P_{ij}(\text{obs}) = N \sum_{i'j'} P_{i'j'}(\text{true})\, P_v(\text{obs}_{ij}|\text{true}_{i'j'}) \quad (25)$$

We write this equation considering the operator R:

$$n = R * \mu$$

The iterative method (Van Cittert 1930) adds with a weight β the residual r_k to the current solution

$$\mu_{k+1} = \mu_k + \beta[n - R * \mu_k] \quad (26)$$

The method is based on the known equation

$$\sum_{i=0}^{k} q^n = \frac{1 - q^{k+1}}{1 - q} \quad (27)$$

for $k \to \infty$ the series converges if $|q < 1|$.
By applying iteratively (26)

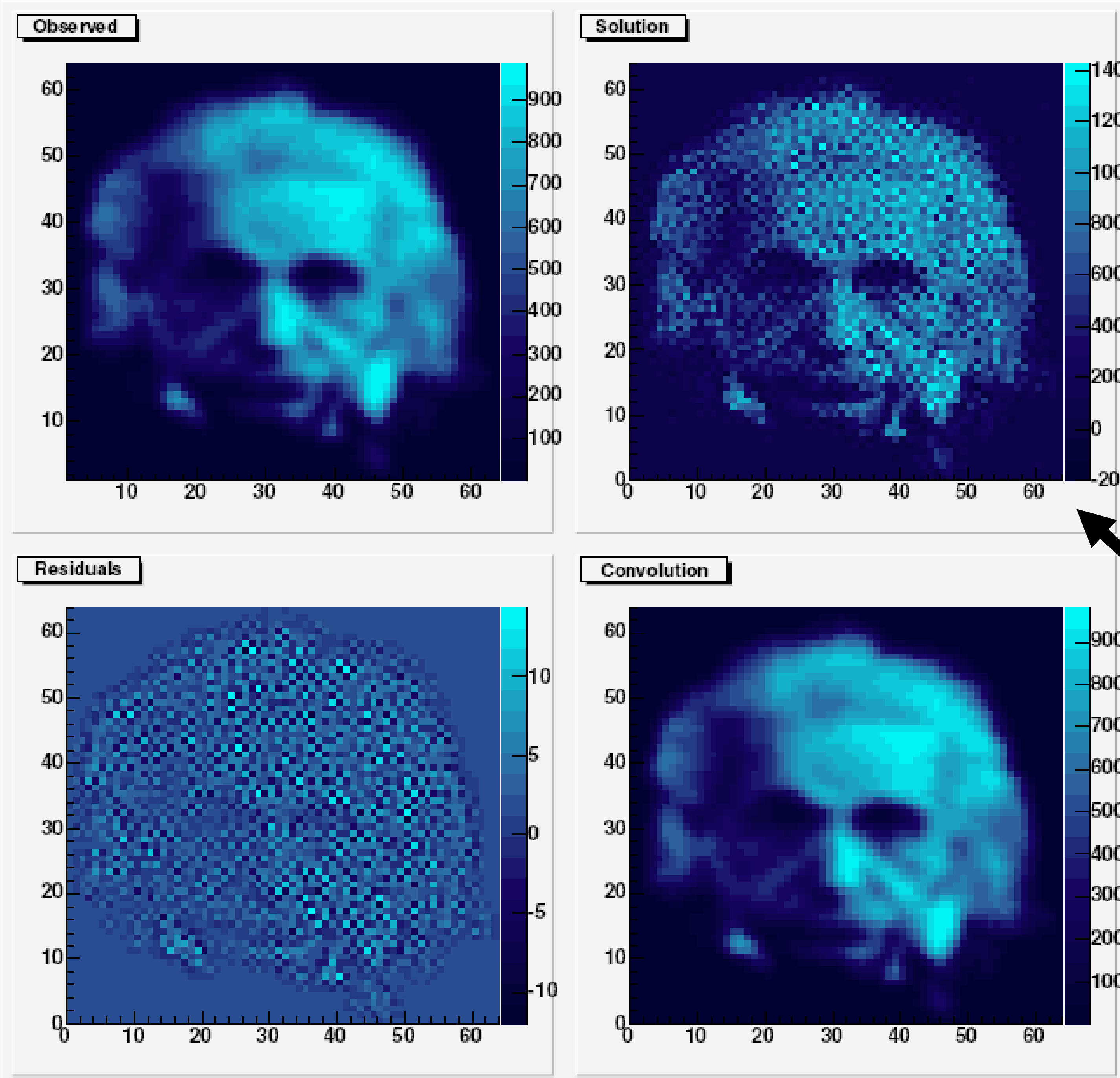
$$\begin{aligned}
\mu_{k+1} &= \beta n + (1 - \beta R)\mu_k = \beta n + (1 - \beta R)(\beta n + (1 - \beta R)\mu_{k-1}) \\
&= \beta n + \beta(1 - \beta R)n + (1 - \beta R)^2 \mu_{k-1} \\
&= \beta n + \beta(1 - \beta R)n + \beta(1 - \beta R)^2 n + (1 - \beta R)^3 \mu_{k-2} \ldots \\
&= \sum_{i=0}^{k} \beta(1 - \beta R)^i n \ .
\end{aligned}$$

From (27):

$$\mu_{k+1} = \frac{I - (I - \beta R)^{k+1}}{\beta R} \beta n \to R^{-1} n = \mu \ , \quad \text{for} \quad k \to \infty \ .$$

if $|I - \beta R| < 1$

The iterative principle

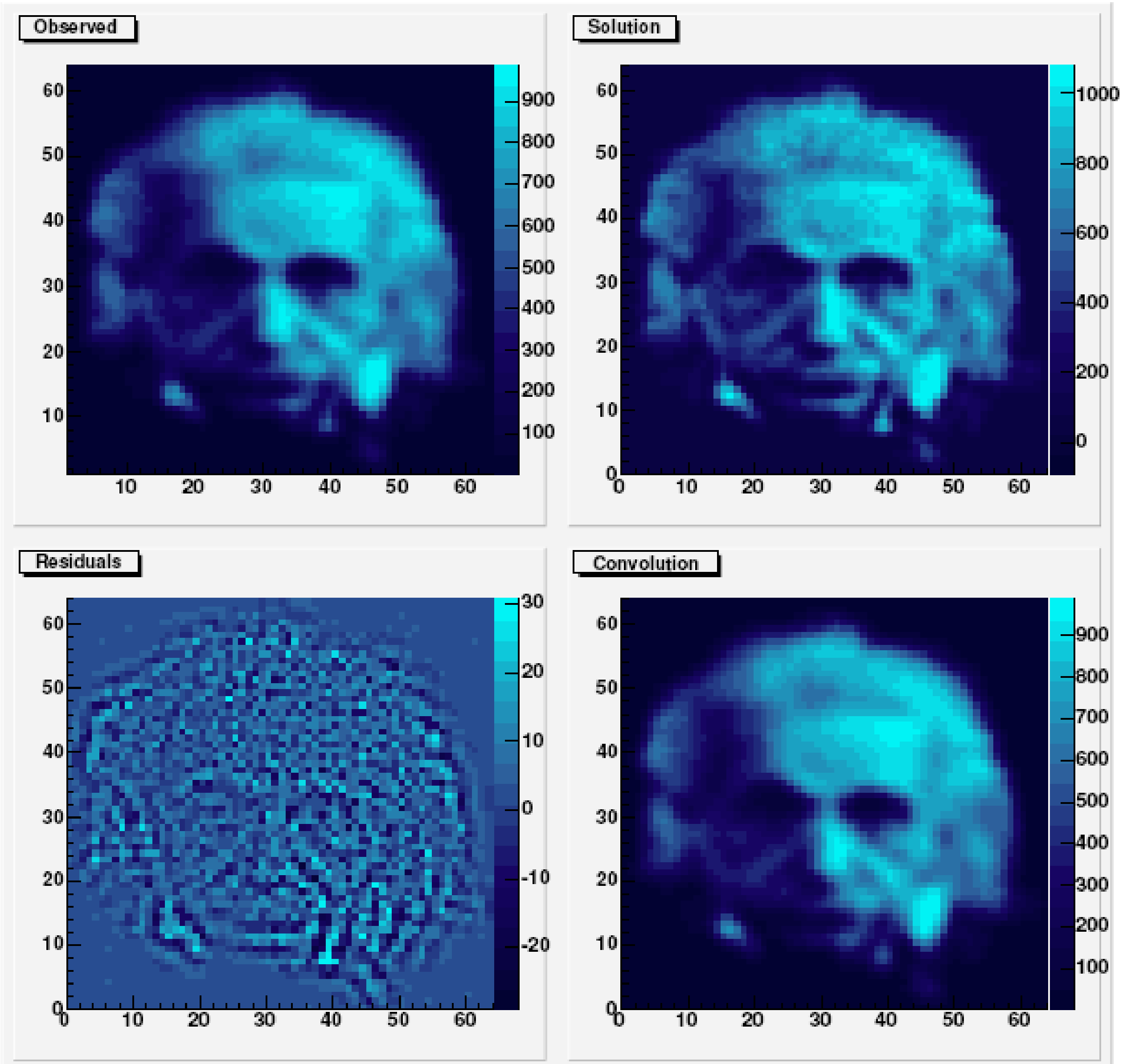


The iterative Principle without best fit

Bad!

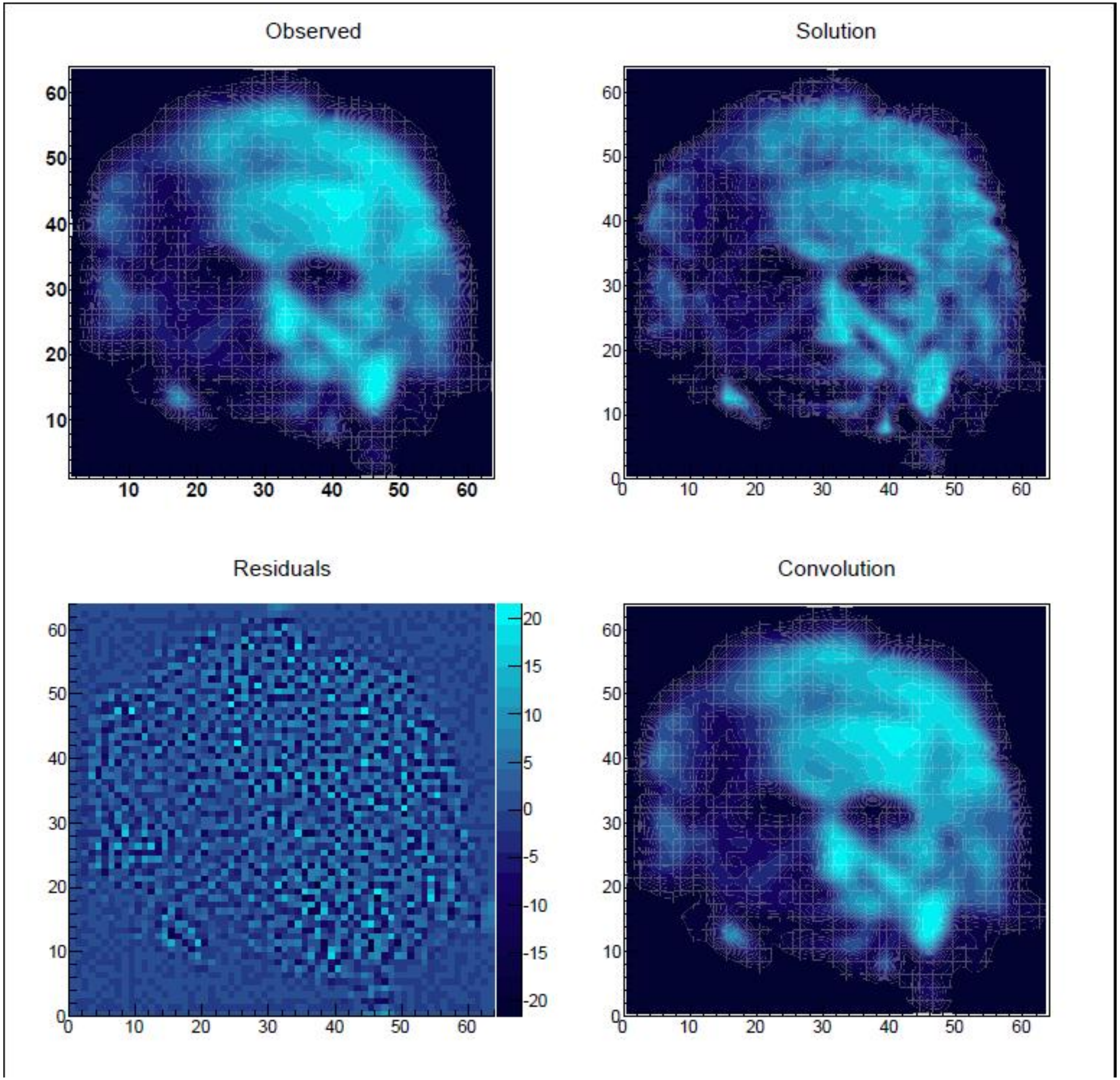
$$\mu_{k+1} = \mu_k + \beta_k[R * n - [R * R * \mu_k + \alpha(\ln \mu_k/\mu_T + I)]$$

About 40 iterations, regularized with Maximum entropy

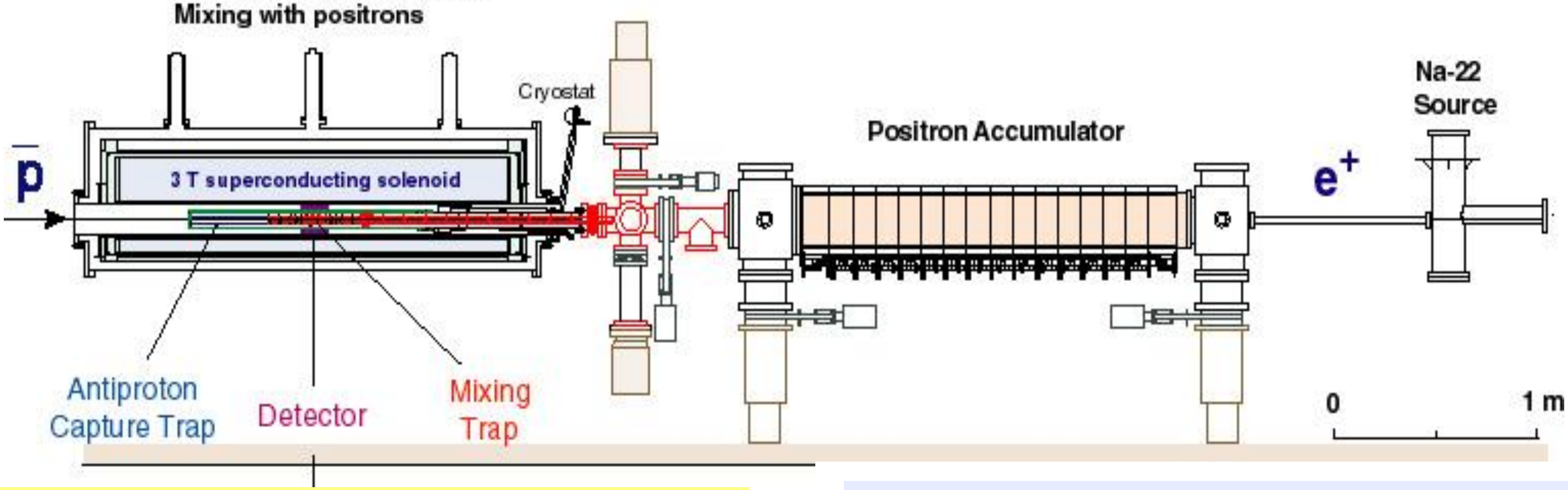


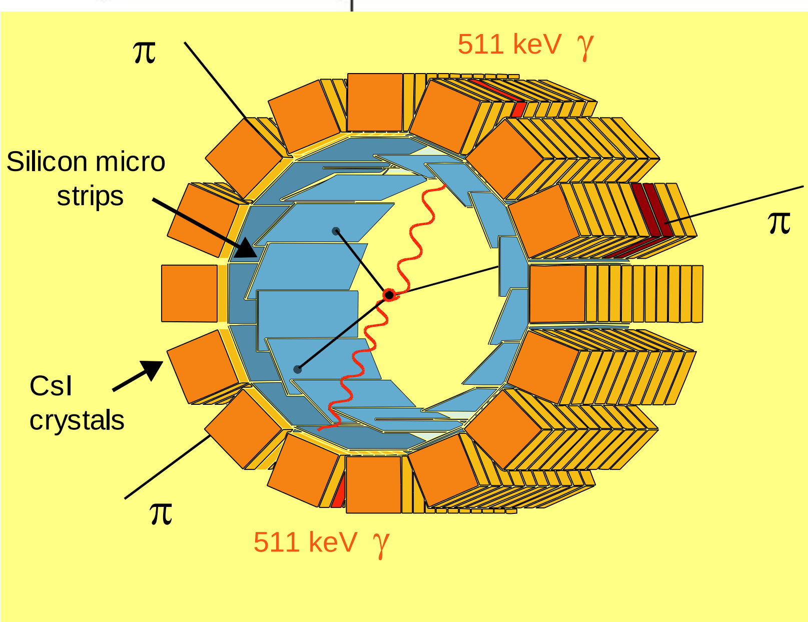
The iterative algorithm + best fit + MaxEnt regularization

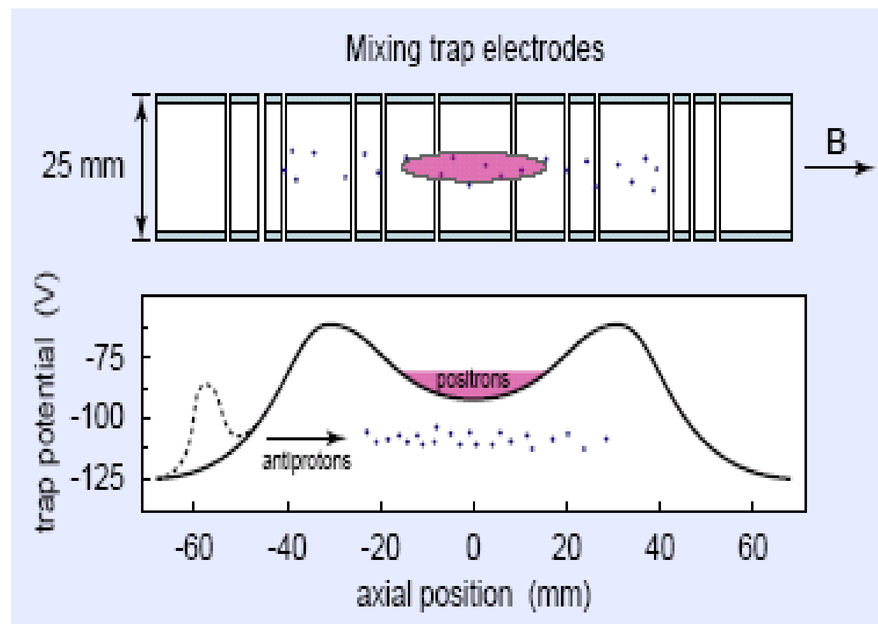
Bayesian algorithm



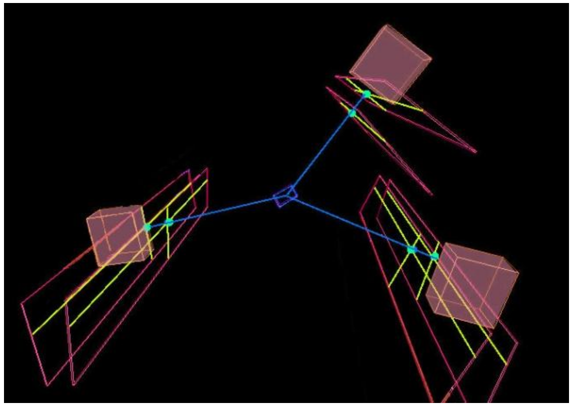
ATHENA apparatus

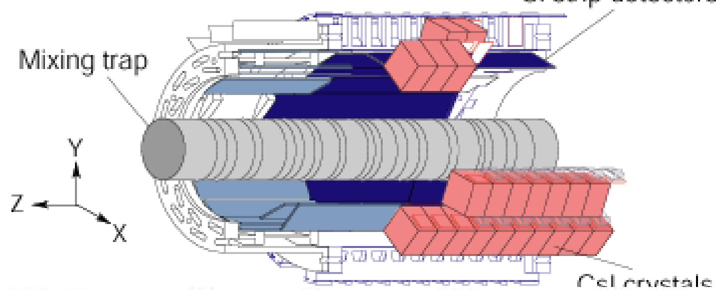






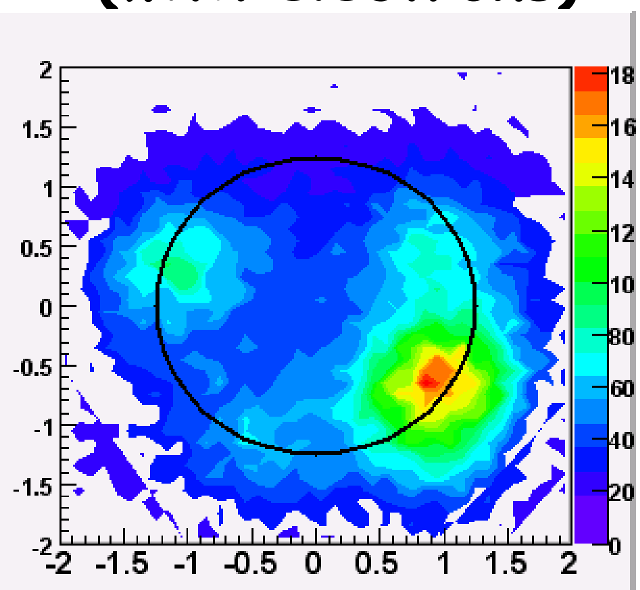
From the ATHENA detector





Distribution of annihilation vertices when antiprotons are mixed with ...

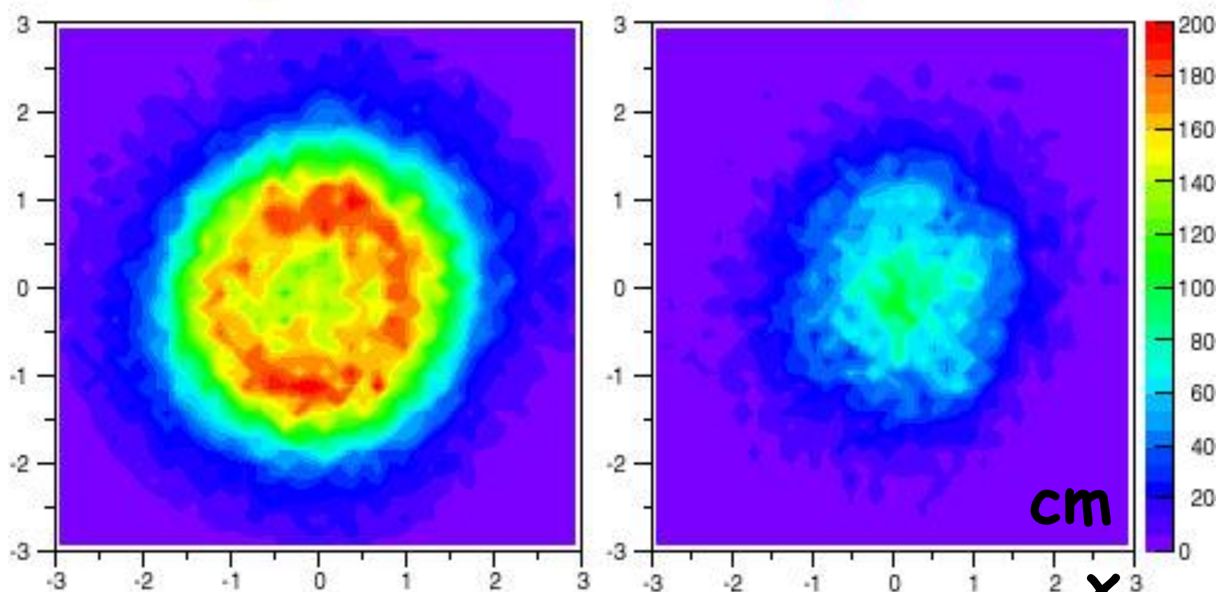
Pbar-only
(with electrons)



y

cold positrons

hot positrons

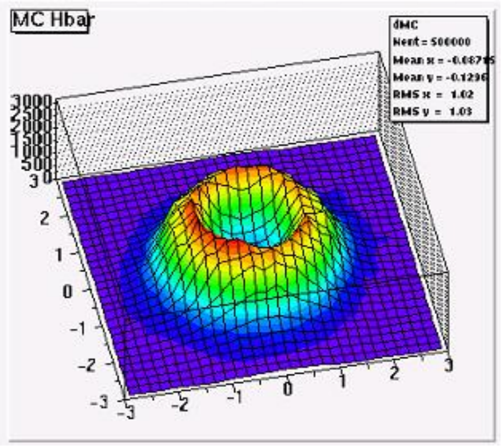


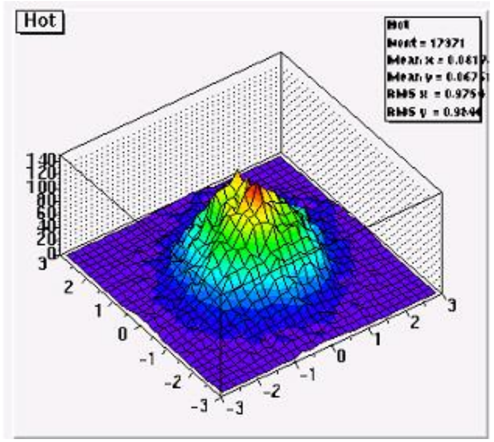
cm

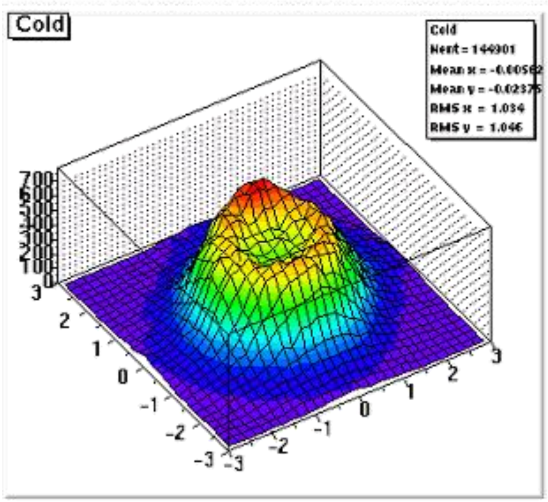
x

Annihilation vertex in the trap x-y plane

Hbar (MC) | **BCKG (HotMixData)** | **Cold Mix data**



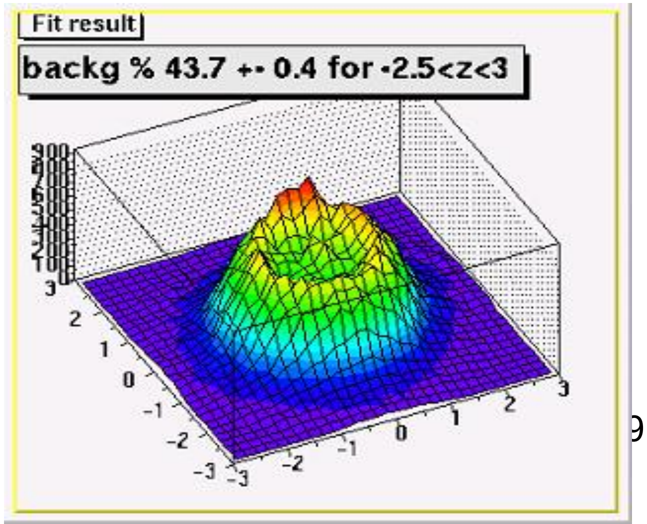




Pbar vertex XY projection (cm)

x Hbar + (1-**x**) BCKG =

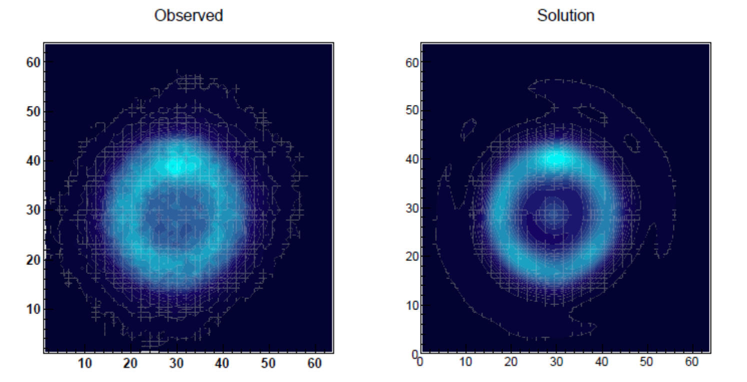
ML Fit Result

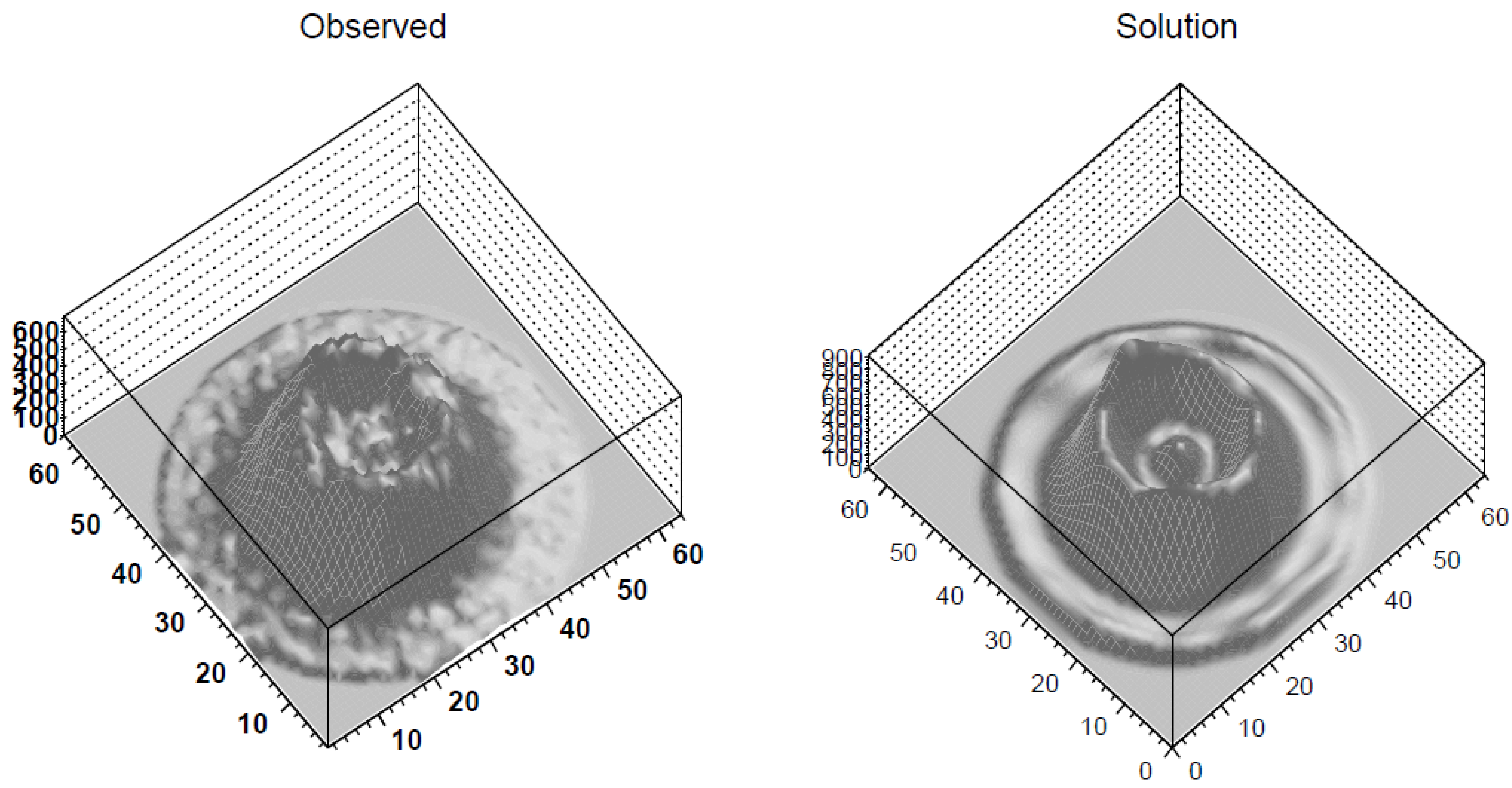


Hbar percentage

x = 0.65 ± 0.05

Iteratve best fit (Bayesian) method





Conclusions on unfolding

- iterative algorithms are used in unfolding (ill posed) problems
- sometimes they need a Bayesian regularization term
- when there are degrees of freedom, one can use a best fit of a signal+background function to the data
- to find a reliable error for the solution is still an open problem

END

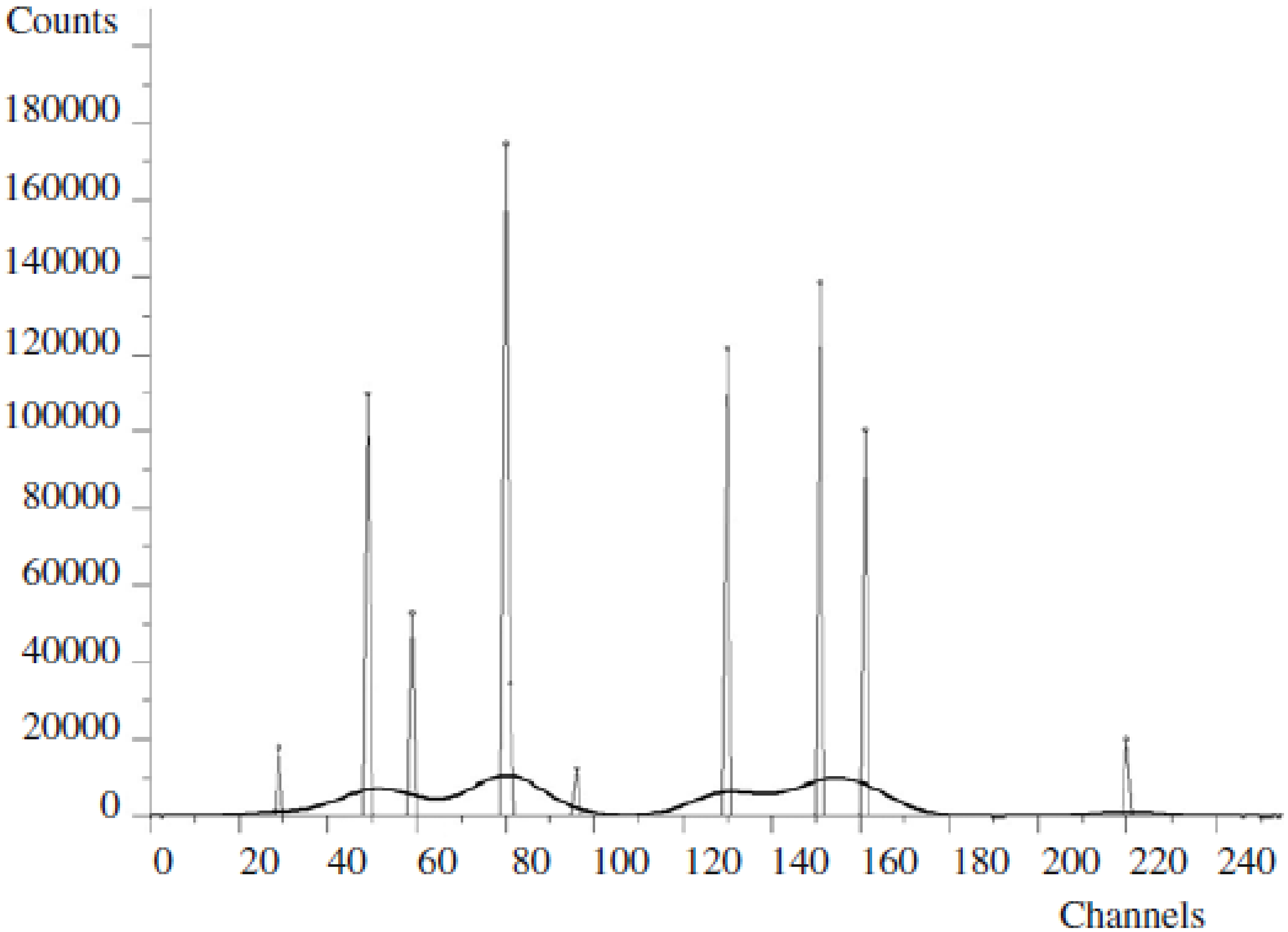
Errors

$$t_i^{(k)} = \sum_i M_{ij} m_j$$

$$\frac{\mathrm{d}t_i^{(k)}}{\mathrm{d}m_n} = M_{in} + \sum_j m_j \frac{\partial M_{ij}}{\partial m_n}$$

$$V(t_i, t_j) = \sum_l \sum_n \frac{\partial t_i^{(F)}}{\partial m_l} V_d(m_l, m_n) \frac{\partial t_j^{(F)}}{\partial m_n}$$

$$V_d(m_l, m_n) = \begin{cases} N_m(1 - N_m / N_{Tot}) & n = l \\ -N_{Tot}(N_m / N_{Tot})(N_m / N_{Tot}) & n \neq l \end{cases}$$


Counts
180000
160000
140000
120000
100000
80000
60000
40000
20000
0
0 20 40 60 80 100 120 140 160 180 200 220 240
Channels

Efficiency calculation: an OPEN PROBLEM!!

$$\frac{f+\frac{t^2}{2n}}{\frac{t^2}{n}+1} \pm \frac{t_\alpha \sqrt{\frac{t^2}{4n^2}+\frac{f(1-f)}{n}}}{\frac{t^2}{n}+1}$$

Wilson interval (1934)

$$\xrightarrow{n>>1} \varepsilon = f \pm t_\alpha \sqrt{\frac{f(1-f)}{n}}$$

Wald (1950)
Standard in Physics

$$\sum_{k=x}^{n} \binom{n}{k} \varepsilon_1^k (1-\varepsilon_1)^{n-k} = \alpha/2$$

$$\sum_{k=0}^{x} \binom{n}{k} \varepsilon_2^k (1-\varepsilon_2)^{n-k} = \alpha/2$$

Exact frequentist
Clopper Pearson (1934) (PDG)

Statistics of counting

Fixed n

$$B(x;n,\varepsilon)=\frac{n!}{x!(n-x)!}\varepsilon^{x}(1-\varepsilon)^{n-x}\qquad \mu=n\varepsilon,\quad \sigma=\sqrt{n\varepsilon(1-\varepsilon)}$$

$$n=\frac{x}{\varepsilon}\pm\frac{1}{\varepsilon}\sqrt{x(1-\varepsilon)+\frac{x^2}{\varepsilon^2}\sigma_\varepsilon^2},\qquad \varepsilon=\frac{x}{n}\pm\sqrt{\frac{x}{n^2}\left(1-\frac{x}{n}\right)}=f\pm\sqrt{\frac{f(1-f)}{n}}$$

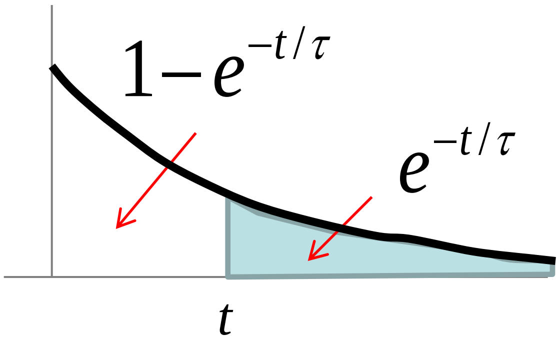
Poissonian n

$$P(x;\mu)=\frac{\mu^x}{x!}\mathrm{e}^{-\mu},\qquad \sigma=\sqrt{\mu},\qquad \mu=x\pm\sqrt{x}$$

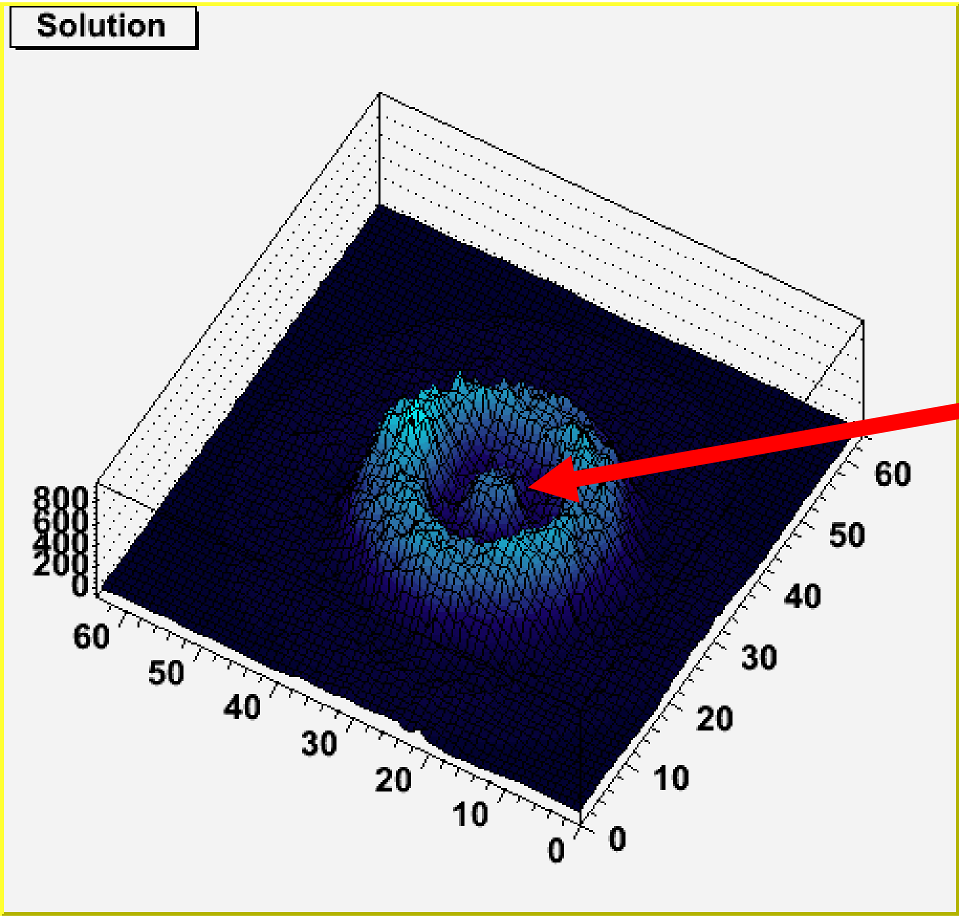
times

$$e(t;\tau)=\frac{1}{\tau}\mathrm{e}^{-t/\tau},\quad \mu=\frac{\Delta t}{\tau},\quad r=\frac{1}{\tau}$$

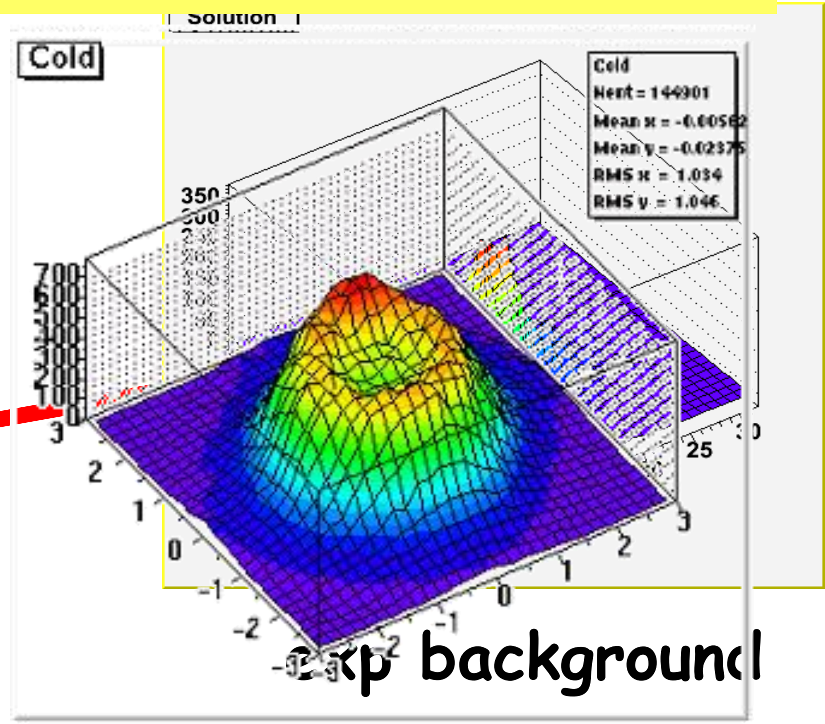
$$\langle t\rangle=\sigma=\tau=\frac{1}{r},$$



Iterative best fit (residual) method



Cold Mix

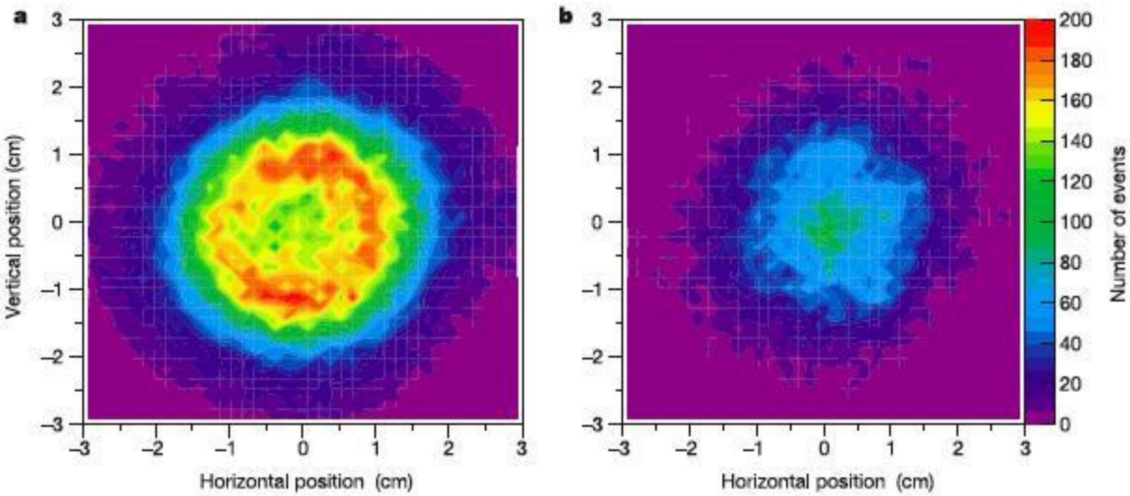


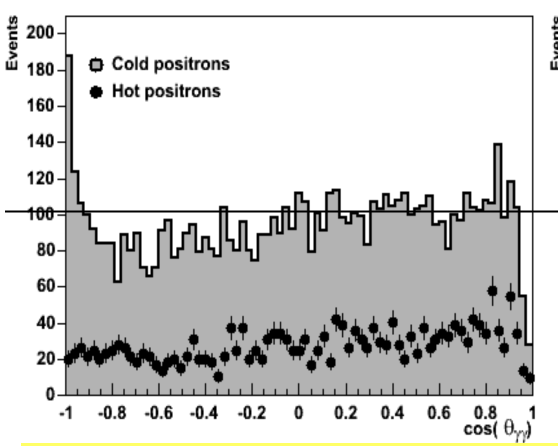
The vertex algorithm resolution function is gaussian with $\sigma \simeq 3$ mm

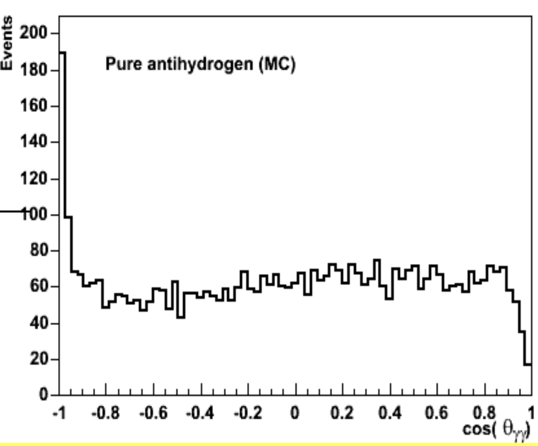
The 2D deconvolution reveals two different annihilation modes

antihydrogen !!!!!!!!!!

FIRST COLD ANTIHYDROGEN PRODUCTION & DETECTION (2002)
M. Amoretti et al., Nature 419 (2002) 456
M. Amoretti et al., Phys. Lett. B 578 (2004) 23







SIGNAL ANALYSIS:

opening angle
xy vertex distribution
radial vertex distribution

65 % +/- 10% of annihilations are due to antihydrogen

between 2002 & 2004
more than 2 millions
antihydrogen atoms
have been produced

that's about 2 x 10^{-15} mg
.. or .. 1000 Giga years for a gram

$$\frac{80}{\sqrt{190+110}} = 4.7; \quad \frac{80}{\sqrt{190}} = 6.5; \quad \frac{80}{\sqrt{110}} = 8$$

The pile-up distributions

$$e(t;\tau) = \frac{1}{\tau}\,\mathrm{e}^{-t/\tau}, \quad \mu = \frac{\Delta t}{\tau}, \; r = \frac{1}{\tau}$$

$$\langle t \rangle = \sigma = \tau = \frac{1}{r},$$

$$e(t;\tau) = \frac{r^k\, t^{k-1}}{(k-1)!}\,\mathrm{e}^{-rt}$$

$$\langle t \rangle = k\tau = \frac{k}{r}, \quad \sigma = \sqrt{k}\,r = \frac{\sqrt{k}}{\tau}$$

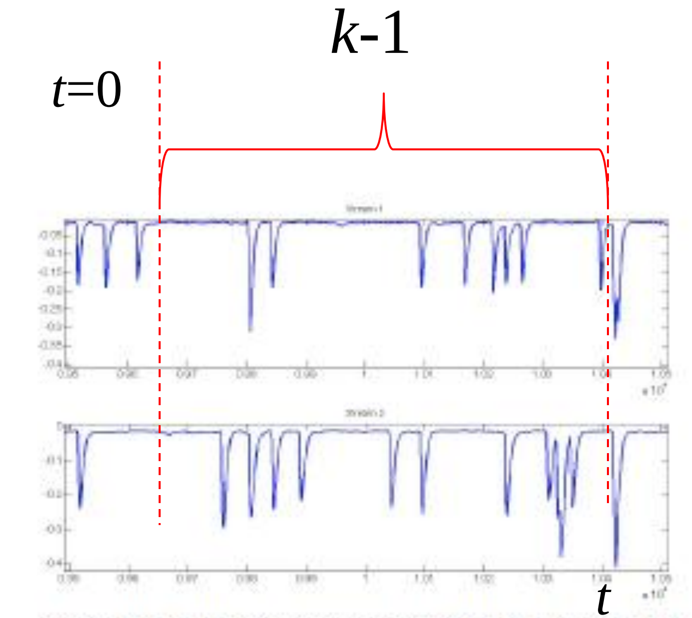


Fig. 7. Matlab plot of a pileup streams with 1000kcps from the pulse stream generator routine.

Statistical Science
2001, Vol. 16, No. 2, 101–133

Interval Estimation for a Binomial Proportion

Lawrence D. Brown, T. Tony Cai and Anirban DasGupta

Simulate many x with a true p and check when the intervals contain the true value p. Compare this frequency with the stated CL

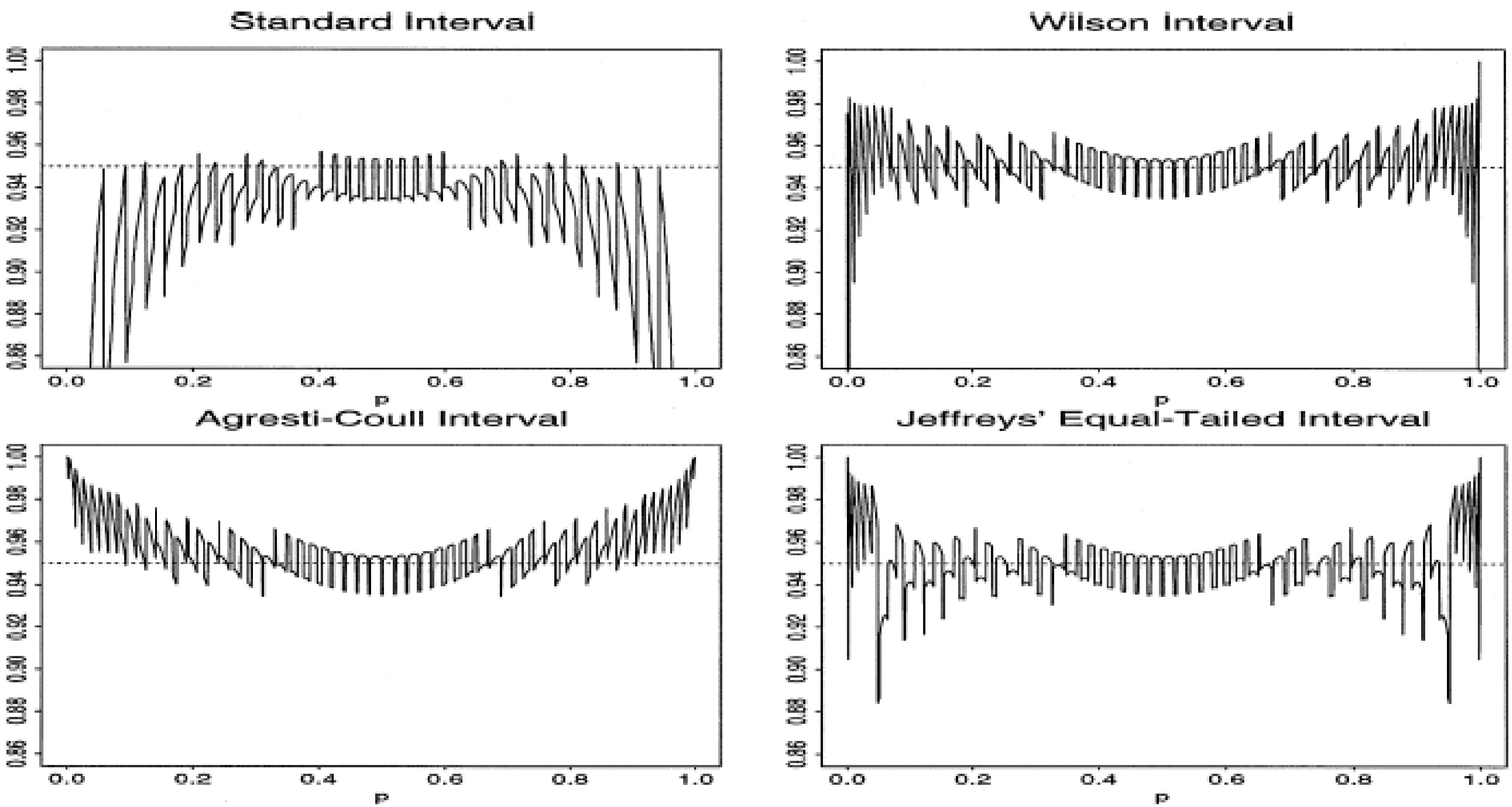


Fig. 5. *Coverage probability for* $n = 50$.

CL=0.95, n=50

Paralyzable Dead time determination

Meeks and Siegel Am.J.Phys. 76(2008)659

- With pile-up the time distribution deviates from the exponential
- the property ->

$$\frac{\int t^m r\, e^{-rt}\, \mathrm{d}t}{\left(\int t\, r\, e^{-rt}\, \mathrm{d}t\right)^m} = \frac{\langle t^m \rangle}{\langle t \rangle^m} = m!$$

in this case does not hold

- If one collects a sample of t_i , subtracs a common time T, discard the differences $(t_i\text{-}T)<0$ and calculates

$$\frac{\sum_i (t_i - T)^m / N}{\left(\sum_i (t_i - T) / N\right)^m}$$

one sees that the property above is satisfied when

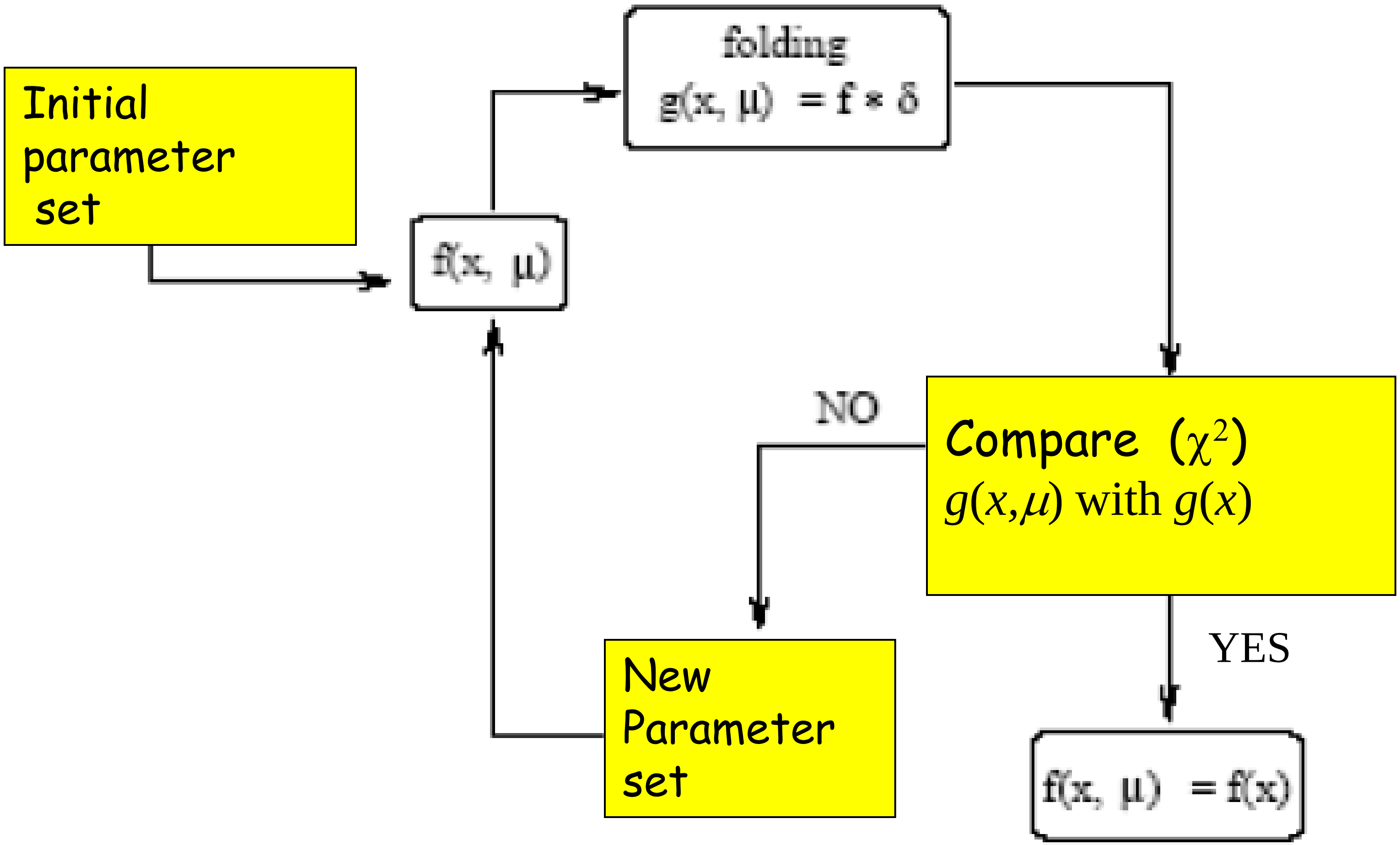
T>dead time

Table I. The moment ratios $\langle t^m \rangle / \langle t \rangle^m$ for m= 2, 3, and 4 for different delay times D for a series of 10 000 Geiger counter intervals. The statistical uncertainties in the last row are approximately the same for each number in the column.

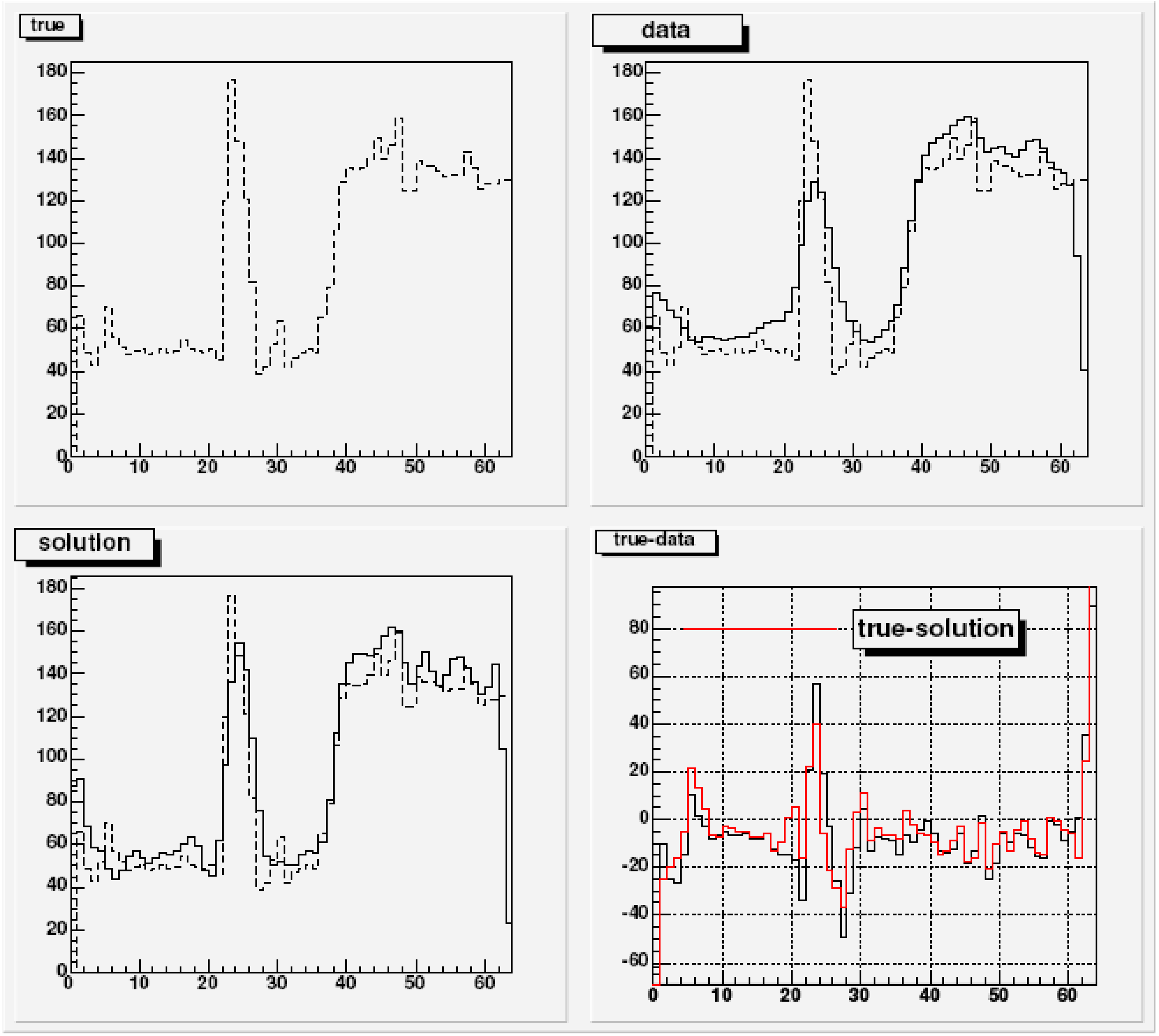
$D(\mu s)$	$\langle t^2 \rangle / \langle t \rangle^2$	$\langle t^3 \rangle / \langle t \rangle^3$	$\langle t^4 \rangle / \langle t \rangle^4$	Count rate (counts/s)	N
0	1.87	5.34	19.6	171.7	10 000
100	1.90	5.42	20.6	174.7	10 000
200	1.93	5.61	21.7	177.9	10 000
300	1.96	5.81	22.9	181.1	10 000
400	2.00	6.02	24.2	184.3	9996
500	2.01	6.07	24.4	185.0	9849
600	2.00	6.06	24.4	184.9	9660
800	2.00	6.05	24.3	184.8	9303
1000	2.01 ± 0.02	6.06 ± 0.19	24.4 ± 1.7	185 ± 2	8975

To fold …

$$g(y) = \int f(x)\,\delta(y-x)\,\mathrm{d}x$$

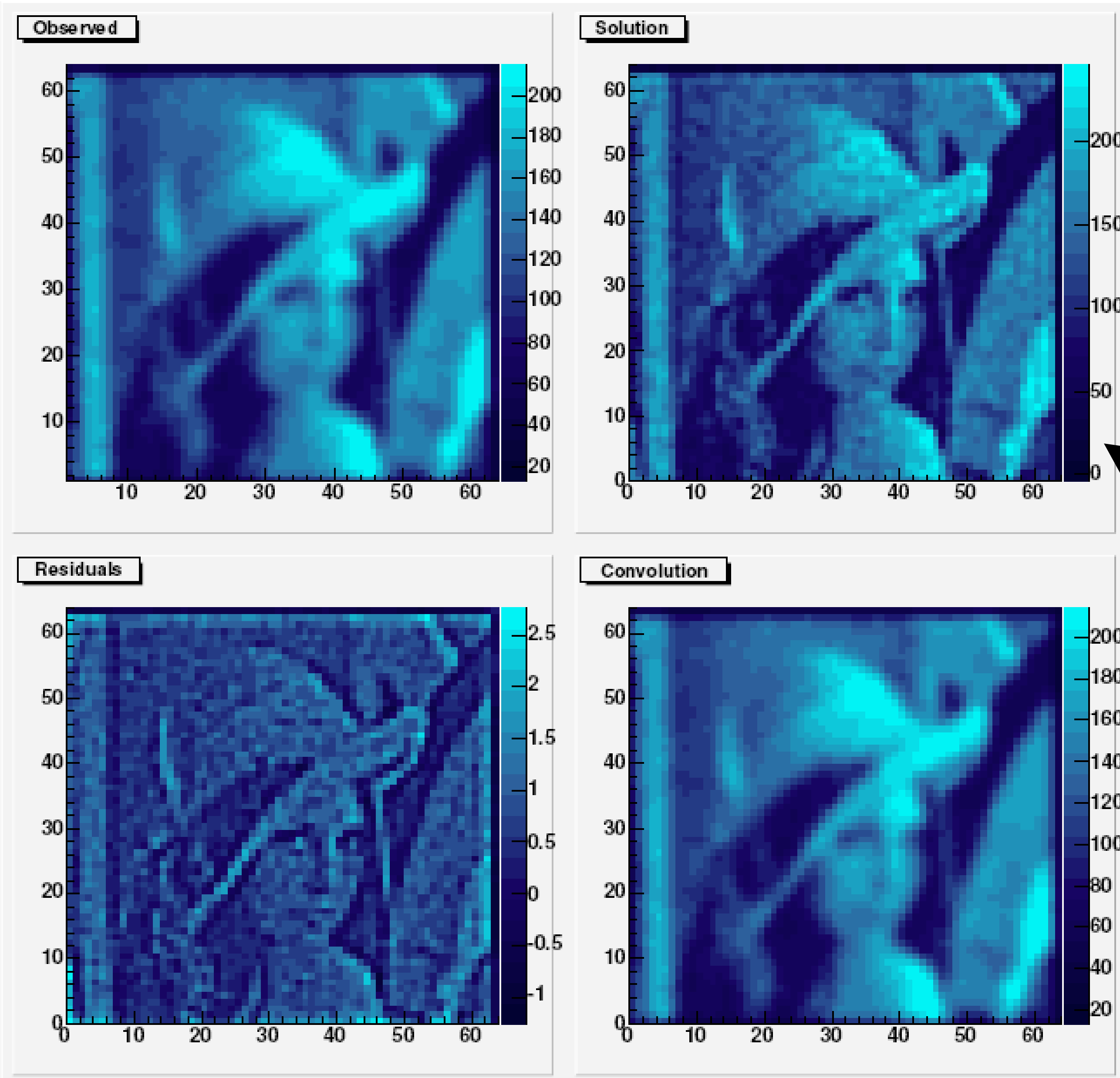


$$\mu_{k+1} = \mu_k + \beta_k[R * n - (R * R + \alpha I) * \mu_k]$$



The iterative algorithm + best fit + Tichonov regularization

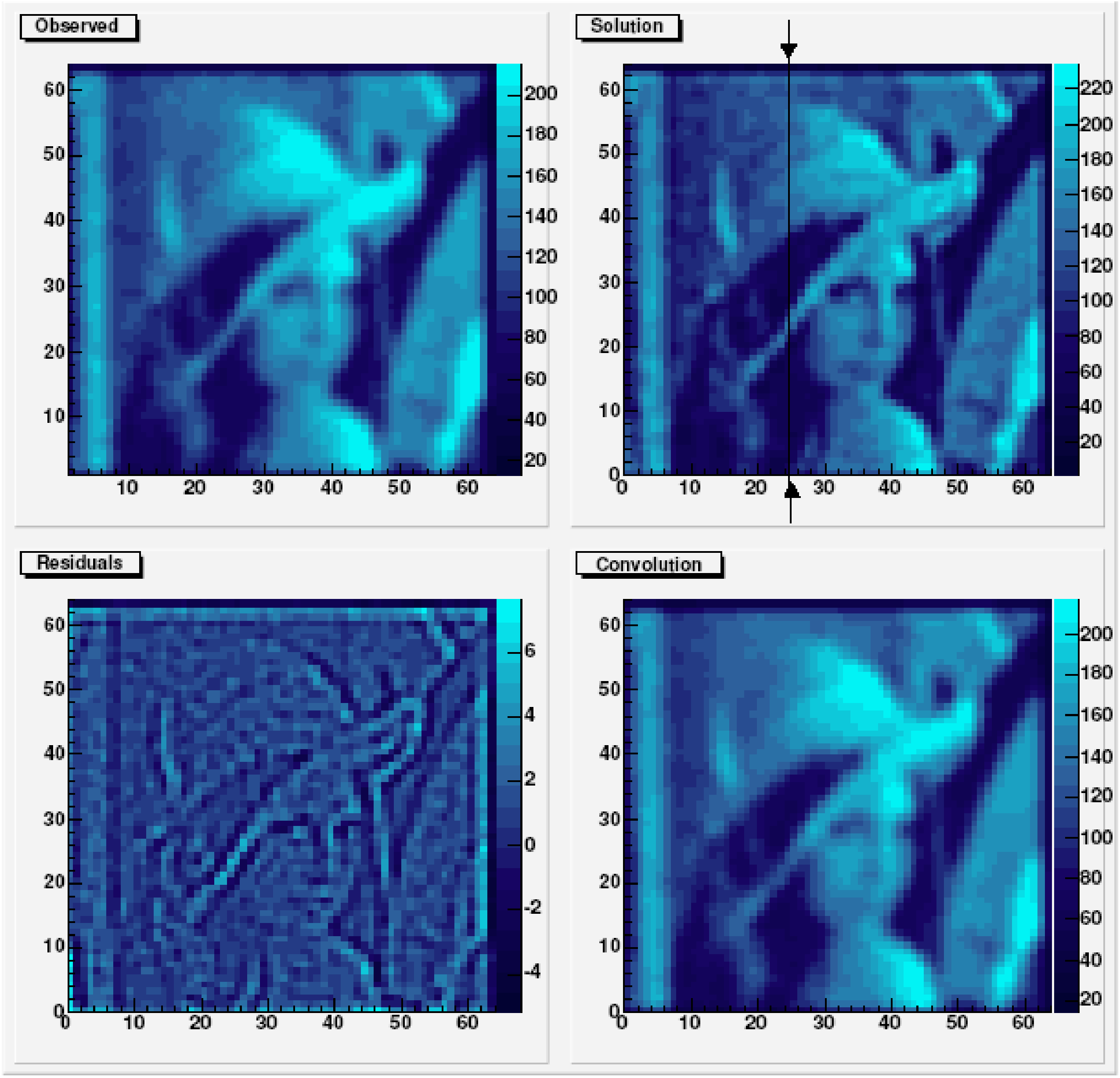
Without β $\quad \mu_{k+1} = \mu_k + [n - R * \mu_k]$



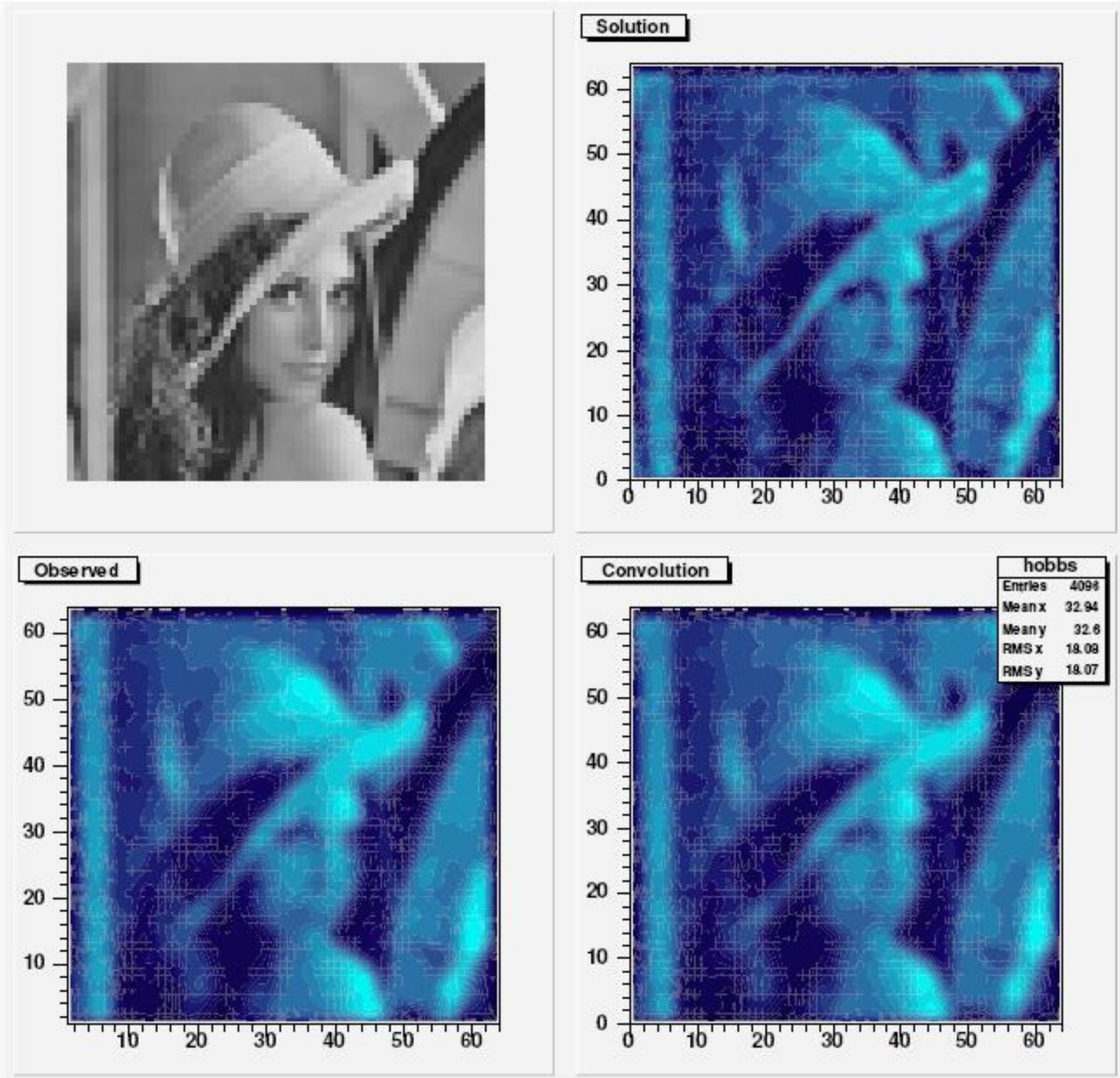
The iterative Principle without best fit

Good!

$$\mu_{k+1} = \mu_k + \beta_k[R * n - (R * R + \alpha I) * \mu_k]$$



The iterative algorithm + best fit + Tichonov regularization



The iterative Principle without best fit + smoothing

original

Poisson statistics

Gaussian smearing

Fourier restored

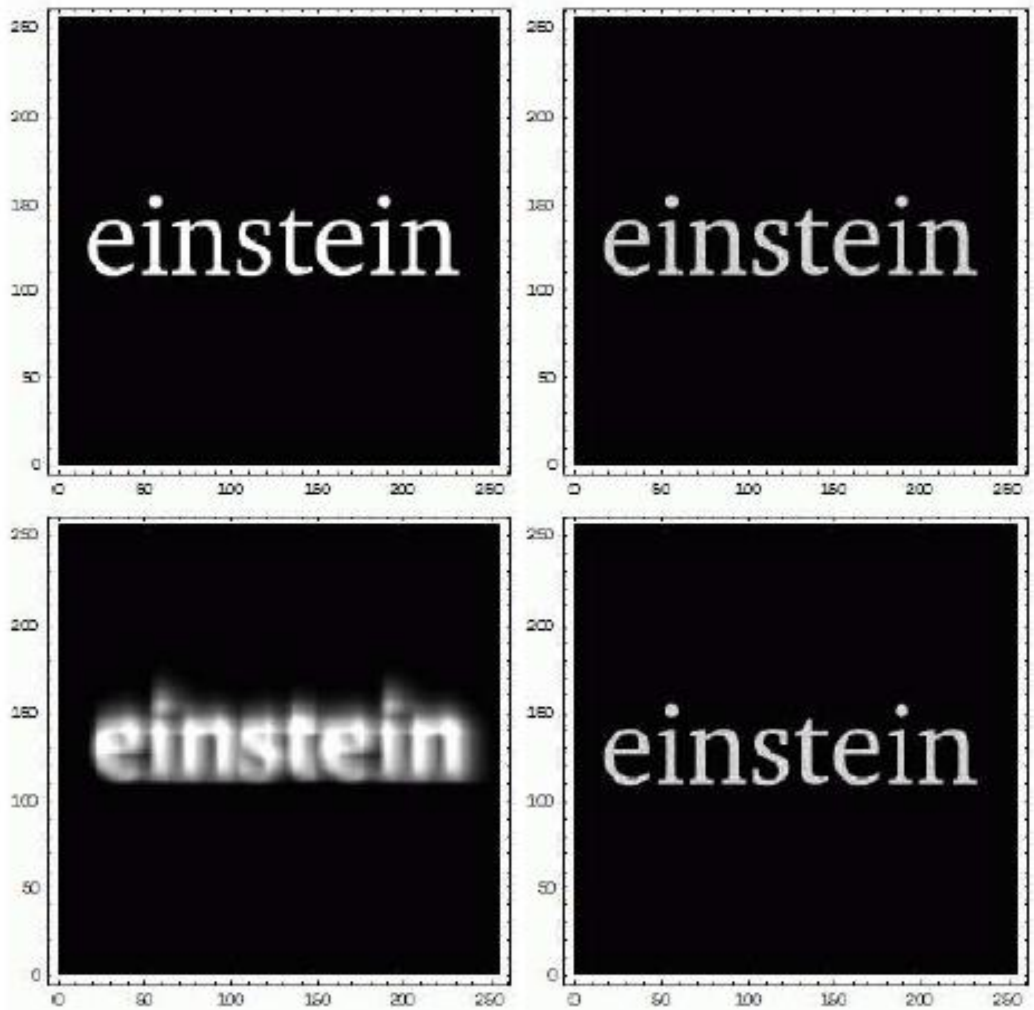


Figure 13: **Einstein restored by FFT: explanation as in Figure 1.**

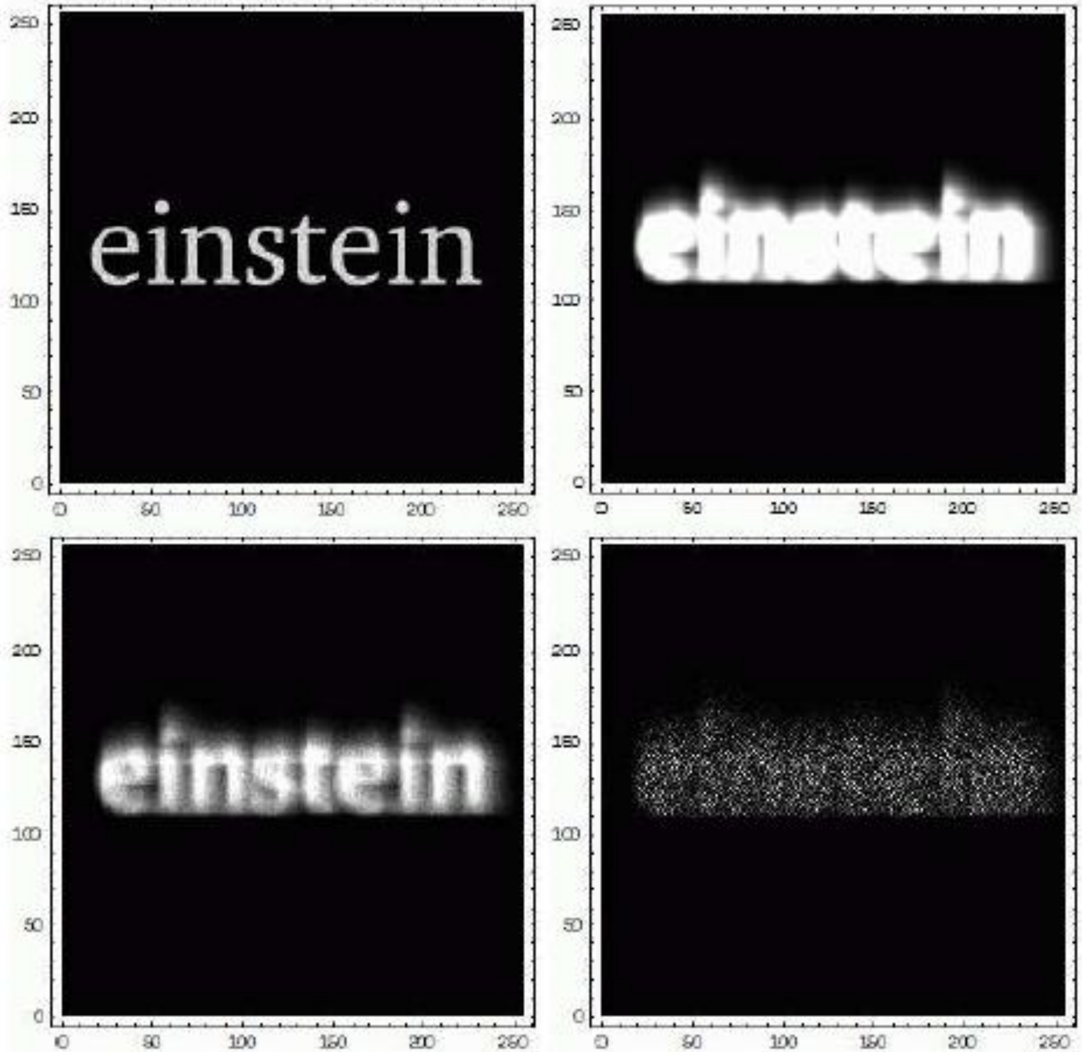


Figure 14: **Einstein not restored by FFT: explanation as in Figure 2.**

The objective function to be minimized is

$$-F(\boldsymbol{\mu}) = -2\ln L(\boldsymbol{n}|\boldsymbol{\mu}) - \alpha\, C(\boldsymbol{\mu}) + \lambda(n_T - \sum_i \nu_i) \quad (23)$$

$$\mu_j = \mu_{\rm tot} p_j = \mu_{\rm tot} \int_{\rm bin\, j} f_t(y)\, {\rm d}y$$

where $\alpha > 0$. Some regularization terms:

Regularization terms

- minimum second derivative (Tichonov)

$$C(\boldsymbol{\mu}) = -\int [f_t''(y)]^2\, {\rm d}y \simeq -\sum_{i=1}^{M-2} [-\mu_i + 2\mu_{i+1} - \mu_{i+2}]^2$$

- minimum variance:

$$C(\boldsymbol{\mu}) = -{\rm Var}[\boldsymbol{\mu}] \equiv ||C\boldsymbol{\mu}||^2 = -\sum_i \mu_i^2$$

- maximum entropy (MaxEnt)

$$C(\boldsymbol{\mu}) = -\sum_i p_i \ln p_i = -\sum_i \frac{\mu_i}{\mu_T} \ln \frac{\mu_i}{\mu_T}$$

- cross-entropy

$$C(\boldsymbol{\mu}) = -\sum_i p_i \ln \frac{p_i}{q_i} = -\sum_i \frac{\mu_i}{\mu_T} \ln \frac{\mu_i}{\mu_T\, q_i}$$

where $\boldsymbol{q} = (q_1, q_2, \ldots, q_n)$ is the most likely a priori shape for the true distribution μ_i.

Deterministic algorithms

The objective function to be minimized is

$$-F(\boldsymbol{\mu}) = -2\ln L(\boldsymbol{n}|\boldsymbol{\mu}) - \alpha\, C(\boldsymbol{\mu}) + \lambda(n_T - \sum_i \nu_i) \quad (23)$$

$$\mu_j = \mu_{\rm tot} p_j = \mu_{\rm tot} \int_{\rm bin\, j} f_t(y)\, {\rm d}y$$

where $\alpha > 0$. Some regularization terms:

- minimum second derivative (Tichonov)

$$C(\boldsymbol{\mu}) = -\int [f_t''(y)]^2\, {\rm d}y \simeq -\sum_{i=1}^{M-2} [-\mu_i + 2\mu_{i+1} - \mu_{i+2}]^2$$

- minimum variance:

$$C(\boldsymbol{\mu}) = -{\rm Var}[\boldsymbol{\mu}] \equiv ||C\boldsymbol{\mu}||^2 = -\sum_i \mu_i^2$$

- maximum entropy (MaxEnt)

$$C(\boldsymbol{\mu}) = -\sum_i p_i \ln p_i = -\sum_i \frac{\mu_i}{\mu_T} \ln \frac{\mu_i}{\mu_T}$$

- cross-entropy

$$C(\boldsymbol{\mu}) = -\sum_i p_i \ln \frac{p_i}{q_i} = -\sum_i \frac{\mu_i}{\mu_T} \ln \frac{\mu_i}{\mu_T\, q_i}$$

  where $\boldsymbol{q} = (q_1, q_2, \ldots, q_n)$ is the most likely a priori shape for the true distribution μ_i.

Image Deconvolution

$$D(\boldsymbol{x}) = \int \mathrm{d}\boldsymbol{y} I(\boldsymbol{y})\, \delta(|\boldsymbol{x} - \boldsymbol{y}|)$$

In the absence of noise

$$I = F^{-1}\left[\frac{F(D)}{F(\delta)}\right]$$

where F is the Fourier transform.

For a real image $I(n_1, n_2)$ the Fourier transform is:

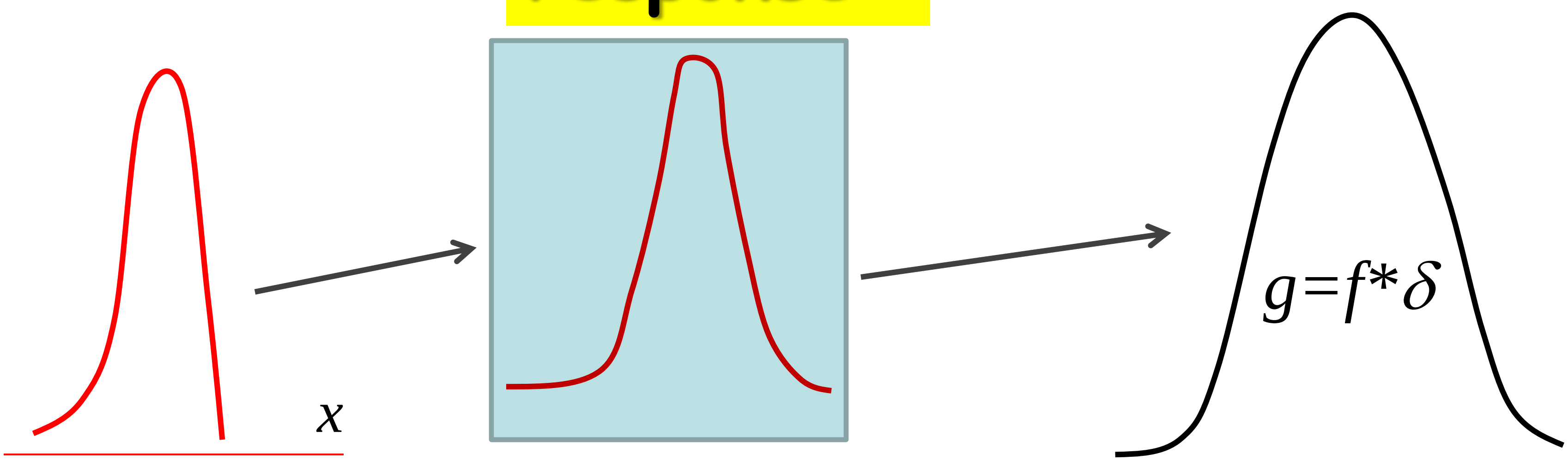
$$F(k_1, k_2) = \sum_{n_2=0}^{N_2-1} \sum_{n_1=0}^{N_1-1} \mathrm{e}^{2\pi i k_2 n_2/N_2}\, \mathrm{e}^{2\pi i k_1 n_1/N_1}\, I(n_1, n_2)$$

$$F(k_1, k_2) = FFT_2[FFT_1[I(n_1, n_2)]]$$

For the routines see for example ***Numerical Recipes*** signal

Apparatus response

Observed signal



$z = y - x$

$y = z + x$

$f(x)$

$\delta(y - x)$

$$g(y) = \int f(x)\,\delta(y - x)\,\mathrm{d}x$$