A story which illlustrates that

Unfolding is not unsmearing

or, looking at the math, only very little unsmearing . . .

Michael Schmelling / MPI for Nuclear Physics

source: generated by AI

Alice Wise
Experimentalist
Global Dynamics

source: generated by AI

Bob Smart
Experimentalist
SaharaTech

source: generated by AI

Eve Hibou
Theorist
Saint Marie IAS

source: generated by AI

Jack Trader
Sales Manager
ACME Instruments

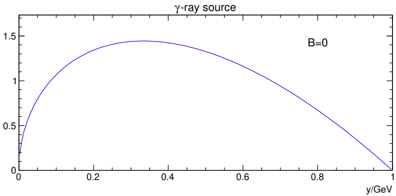
New Physics theory prediction by Eve

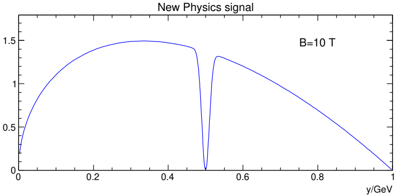
source: generated by AI

❖ “catalized ALP conversion”

In a strong B-field that contributes an axial-vector, γ-rays with $E = m_{\mathrm{ALP}}$ passing through Gd-vapour resonantly convert into ALPs. A measurement requires a broad-band γ-ray source and a good spectrometer.

experimental signature:





ACME Instruments high-resolution spectrometers

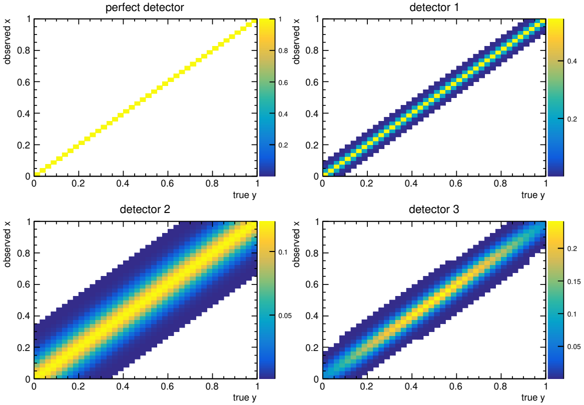

source: generated by AI

- Jack Trader:
  - all detectors by ACME Instruments are perfectly linear
  - the response is very well known
  - correction for detector effects requires only linear algebra
    - discrete unfolding problem for a counting experiment:

$$a_i = \sum_{j=1}^{n_b} R_{ij}\, b_j \qquad i = 1, \ldots, n_a \quad \text{or} \quad a = R\, b$$

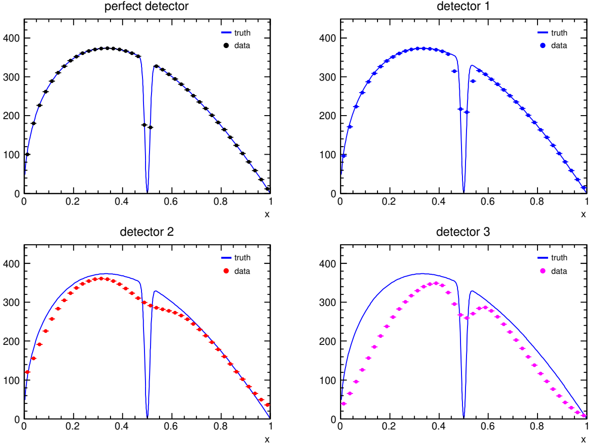
- a: vector of bin contents of the measured distribution
- R: response matrix
- b: vector of bin contents of the true distribution

❖ response matrices for ACME Instruments spectrometers

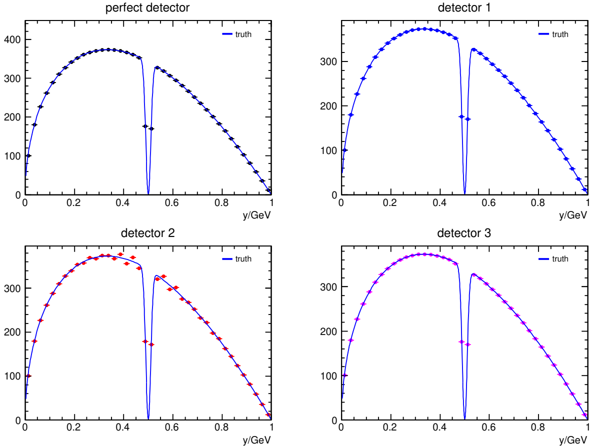


- detector 1:
  high resolution and
  uniform efficiency;
  available in a year
- detector 2:
  low resolution and
  uniform efficiency;
  on stock
- detector 3:
  good resolution but
  non-uniform efficiency;
  sold out; not produced
  anymore

❖ expected measurements $a = R\, b$



❖ corrected measurements $b = R^{-1}a$



Competing experimental groups

source: generated by AI

❖ Alice Wise

- γ-ray source with rate of 100 Hz
- prefer to wait until detector 1 becomes available
- start measurements with a delay of 1 year

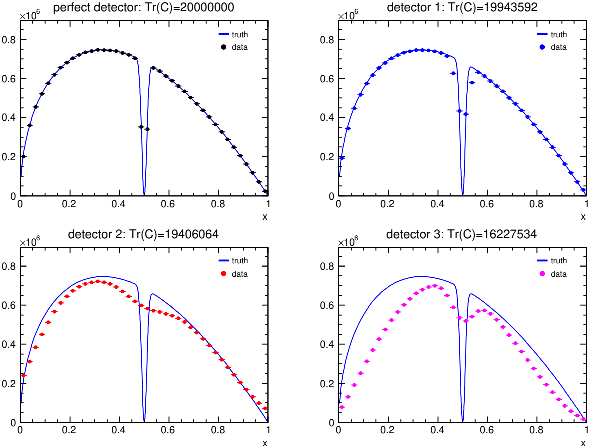
source: generated by AI

❖ Bob Smart

- γ-ray source with rate of 1 Hz
- wants to start measurements immediately
- buy detector 2 and rely on unfolding

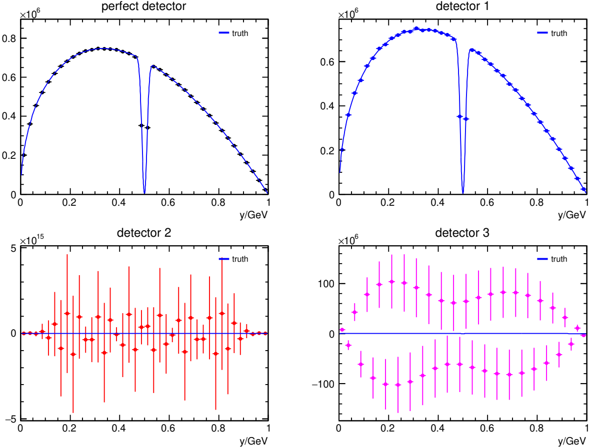
results with $N = 2 \times 10^7$ events ➔

❖ raw data



perfect detector: Tr(C)=20000000
detector 1: Tr(C)=19943592
detector 2: Tr(C)=19406064
detector 3: Tr(C)=16227534
×10^6
truth
data
x
0
0.2
0.4
0.6
0.8
1

❖ unfolded data



Lessons learned and next steps

➜ use the best possible detector!

➜ understand the limitations of unfolding!

Outline

- Diagonalization of $a = R\,b$
- Information content of the data
- Estimating the truth
- Summary

some linear algebra that shows what affects all unfolding methods ➜

1 Diagonalization of $a = R\,b$

❖ expand b into eigenvectors of a symmetric positive definite matrix I

$$I = R^T C_a^{-1} R = V\,S\,V^T \quad \text{with} \quad V\,V^T = V^T V = 1$$

- R: response matrix
- C_a: covariance matrix of the measurements a
- I: inverse of the covariance matrix C_b when estimating b by a least-squares fit;
  Fisher Information matrix for Poisson or gaussian distributed data
- V: "Fisher basis" – columns of V are the orthonormal eigenvectors of I
- S: diagonal matrix of the eigenvalues – sorted in decreasing order

results ➜

- expansion of b in the Fisher basis ➜ expansion coefficients β

$$b = V\,\beta$$

- allow for (over-)constrained problems and determine β by a least-squares fit

$$\chi^2 = \left(a - R\,V\,\beta\right)^T C_a^{-1} \left(a - R\,V\,\beta\right) \stackrel{!}{=} \text{min}$$

  - best-fit values

$$\beta = S^{-1}\alpha \quad \text{with} \quad \alpha = V^T R^T C_a^{-1}\,a$$

  - diagonal covariance matrices

$$C_\alpha = S \quad \text{and} \quad C_\beta = S^{-1}$$

  - minimum χ^2

$$\chi^2_{\text{min}} = a^T C_a^{-1}\,a - \beta^T C_\beta^{-1}\,\beta \approx n_a - n_b \ll N$$

2 Information content of the data

- number of events used to measure the expansion coefficients
  - consider a Poisson distributed variable $x = n$ with error $\sigma_x = \sqrt{n}$

$$\left(\frac{x}{\sigma_x}\right)^2 = \left(\frac{n}{\sqrt{n}}\right)^2 = n$$

    - counting experiment: $(x/\sigma_x)^2$ is the number of events in a measurement x

  - number of events n_i contributing to the measurements β_i:

$$n_i = \beta_i^2 S_{ii}$$

$$\sum n_i = \beta^T C_\beta^{-1} \beta = a^T C_a^{-1} a - \chi^2_{\text{min}} \approx \sum a_i = N$$

    - the n_i sum up to (approximately) the total number of recorded events

❖ contribution of expansion coefficients to the total variance of b

- total variance: sum of diagonal elements of C_b

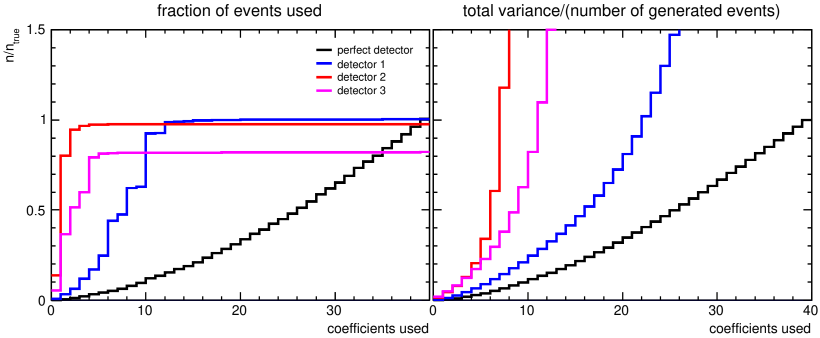
$$\mathrm{Tr}(C_b) = \mathrm{Tr}(V\, C_\beta\, V^T) = \mathrm{Tr}(S^{-1}) = \sum_i \frac{1}{S_{ii}}$$

➜ findings regarding expansion coefficients:

1. they contribute uncorrelated independent information
2. effective number of events in β_i: $\beta_i^2 S_{ii}$
3. contribution of β_i to the total variance of b: $1/S_{ii}$

β_i with small S_{ii} are not measured but dominate the variance

comparison of toy-detectors ➜



- the data contribute mainly to the first few coefficients
  - the well measured expansion coefficients allow for very efficient data reduction
- the variance explodes when using too many coefficients
  - how to construct an estimate of the true density is not obvious

3 Estimating the truth

❖ e.g. truncation: use a damping matrix D that selects the leading m coefficients

$$b = V\,\beta \;\rightarrow\; \hat{b} = V\,D\,\beta \quad \text{with} \quad D = \begin{pmatrix} 1_m & 0 \\ 0 & 0 \end{pmatrix} \quad (m\text{-dim unit matrix})$$

- relation between b and $\hat{b}$: posterior response matrix P

$$\hat{b} = V\,D\,(V^T V)\,\beta = P\,b \quad \text{with} \quad P = V\,D\,V^T \quad \text{and} \quad \mathrm{Tr}(P) = m$$

  - the truncated (regularized) estimate is a biased version of the truth
  - the bias is quantified by P in the same way as R quantifies the bias of the data
  - perfect unfolding with $P = 1$ requires $D = 1$

Estimating P when D is defined implicitly by an unfolding method

❖ exploit the relation between posterior response and covariance matrix

$$C_{\hat{b}} = P\, C_b\, P^T \quad \text{with} \quad C_b = (R^T C_a^{-1} R)^{-1}$$

- use factorisation of symmetric positive definite matrices (SPDMs) into identical SPDMs

$$C = C^{1/2} C^{1/2}$$

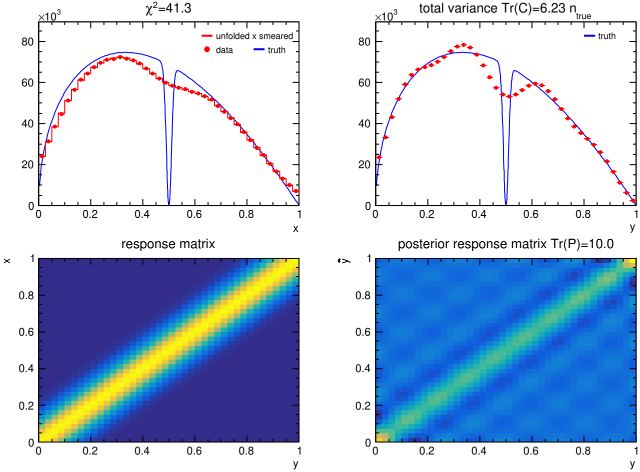
- solving for P

$$P = C_{\hat{b}}^{1/2} C_b^{-1/2} = C_{\hat{b}}^{1/2} (R^T C_a^{-1} R)^{1/2}$$

- number of coefficients actually used to construct $\hat{b}$

$$m = \mathrm{Tr}(P)$$

illustration for toy-detector 2 ➜


χ^2=41.3
×10³
unfolded x smeared
data
truth
x
total variance Tr(C)=6.23 n_{true}
×10³
truth
y
response matrix
x
y
posterior response matrix Tr(P)=10.0
ỹ
y

4 Summary

- already the discrete linear unfolding problem $a = R\,b$ is tough
  - expansion of b into eigenvectors of $R^T C_a^{-1} R$ diagonalizes the problem
    - the expansion coefficients are unbiased and statistically independent
    - the well measured coefficients allow for very efficient data reduction
  - finite resolution entails substantial loss of information
    - improvement of resolution requires LARGE statistics or extra information
  - the real problem is to construct a good estimate for b with well defined properties
    - with truncation, a fluctuation in a single coefficient may cause artefacts everywhere
  - raw data are characterised by the R, unfolded data by the posterior response P

$$P = C_{\hat{b}}^{1/2} (R^T C_a^{-1} R)^{1/2}$$

    - $\mathrm{Tr}(P)$: number of expansion coefficients effectively used