Alliance Workshop on Unfolding (and Data correction)

27–28 May 2010

Volker Blobel – Universität Hamburg

Unfolding – Linear Inverse Problems

1. **Introduction**
2. **Solution of linear inverse problems**
3. **Iterative methods**
4. **Discussion – Presentation of unfolding results**

(. . . see also 57-page report: "Unfolding – Linear Inverse Problems")

---

1. Introduction

The process of the transition from the true distribution $f(x)$ to the measured distribution $g(y)$ for linear inverse problems is described by the Fredholm integral of the first kind:

$$\int_{\Omega} A(y,x) f(x)\,\mathrm{d}x = g(y)$$

The function $A(y,x)$ (called kernel in the theory of integral equations) gives the response of the measurement device to the distribution $f(x)$.

Two types of processes are based on the integral equation:

direct process:	true distribution	$f(x) \to g(y)$	measured distribution	
inverse process:	measured distribution	$g(y) \to f(x)$	true distribution .	

Numerical calculations: replace integral equation by linear equation $\boldsymbol{Ax} + \boldsymbol{\epsilon} = \boldsymbol{y}$

- measured distribution in form of a histogram, represented by a m-vector $\boldsymbol{y}$, contaminated by a statistical fluctuations represented by m-vector $\boldsymbol{\epsilon}$;
- bin-averaged values, represented by a n-vector $\boldsymbol{x}$;
- response function $A(y,x)$, represented by a m-by-n response matrix $\boldsymbol{A}$ (determined by MC methods); it does not depend on the MC input shape.

Properties of the Fredholm integral equation of 1. kind

Riemann-Lebesgue lemma: Assuming a function $f_p(x) = \sin(2\pi px)$ with $p = 1, 2, \ldots$, the lemma states that

$$\int_\Omega A(y,x)\sin(2\pi px)\,\mathrm{d}x = g_p(y) \longrightarrow 0 \qquad \text{for} \quad p \longrightarrow \infty$$

for "arbitrary" Kernel functions $A(y,x)$: high-frequency terms in $f(x)$ are damped in $g_p(y)$, and in the reconstruction of $f(x)$ from $g_p(y)$, they are amplified together with high-frequency noise.
This effect limits the possibility to reconstruct narrow structures in $f(x)$.

Null space of the Kernel: a functions $f_{\text{null}}(x)$ exists with the property

$$\int_\Omega A(y,x) f_{\text{null}}(x)\,\mathrm{d}x = 0\ .$$

Standard strategy is to set $f_{\text{null}}(x) \equiv 0$.

Literature

Per Christian Hansen.
Rank-deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion,
Volume 160 of *SIAM monographs on mathematical modeling and computation*. Society for Industrial and Applied Mathematics, Philadelphia, 1997.

Per Christian Hansen.
Discrete Inverse Problems – Insight and Algorithms,
Volume 160 of *Fundamentals of Algorithms*. Society for Industrial and Applied Mathematics, Philadelphia, 2010.

Jari Kaipio and Erkki Somersalo.
Statistical and Computational Inverse Problems,
Volume 160 of *Applied Mathematical Science*. Springer, 2004.

Jari Kaipio and Erkki Somersalo.
Statistical inverse problems: Discretization, model reduction and inverse crimes,
Journal of Computational and Applied Mathematics, 198:493 – 504, 2007.

Curtis R. Vogel.
Computational Methods for Inverse Problems,
Volume 160 of *SIAM Frontiers in Applied mathematics*. Society for Industrial and Applied Mathematics, Philadelphia, 2002.

Andreas Rieder.
Keine Probleme mit Inversen Problemen.
Vieweg, 2003.

2. Solution of linear inverse problems

In equation $\boldsymbol{Ax} + \boldsymbol{\epsilon} = \boldsymbol{y}$ with $m > n$ or even $m \gg n$ (recommended), the matrix $\boldsymbol{A}$ is a non-square matrix, and hence a direct inverse does not exist.

In **linear** inverse theory the estimate $\widehat{\boldsymbol{x}}$ is a **linear** transformation of the measurement $\boldsymbol{y}$:

$$\widehat{\boldsymbol{x}} = \boldsymbol{A}^{\#}\, \boldsymbol{y}$$

where $\boldsymbol{A}^{\#}$ is called a *generalized inverse* and can include regularization. Inserting $\boldsymbol{y} = \boldsymbol{Ax} + \boldsymbol{\epsilon}$:

$$\widehat{\boldsymbol{x}} = \left(\boldsymbol{A}^{\#}\boldsymbol{A}\right)\boldsymbol{x} + \boldsymbol{A}^{\#}\boldsymbol{\epsilon} \quad = \quad \boldsymbol{x} \quad + \quad \underbrace{\left(\boldsymbol{A}^{\#}\boldsymbol{A} - \boldsymbol{I}\right)\boldsymbol{x}}_{\text{regularization error}} \quad + \quad \underbrace{\left(\boldsymbol{A}^{\#}\right)\boldsymbol{\epsilon}}_{\text{error contribution}}$$

The n-by-n matrix $\boldsymbol{\Xi} = \boldsymbol{A}^{\#}\boldsymbol{A}$ is called the *resolution matrix*.

The existence of the matrix $\boldsymbol{A}^{\#}$ is essential, it allows the standard calculation of the covariance matrix of estimate $\widehat{\boldsymbol{x}}$:

$$\boldsymbol{V}_x = \boldsymbol{A}^{\#}\, \boldsymbol{V}_y\, \boldsymbol{A}^{\#\mathrm{T}}$$

Some unfolding methods (CFM, ...) construct $\widehat{\boldsymbol{x}}$, without defining the generalized $\boldsymbol{A}^{\#}$, thus **no** standard covariance matrix calculation, inclusing the correlation between elements, is possible.

Least Squares solution

Using Least Squares as statistical model, with $F(\boldsymbol{x}) = \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|^2 = \text{minimum}$, matrix $\boldsymbol{A}^{\#}$ becomes:

$$\boldsymbol{A}^{\#} = \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{V}_y^{-1}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{V}_y^{-1} \qquad \boldsymbol{V}_x = \boldsymbol{A}^{\#}\,\boldsymbol{V}_y\boldsymbol{A}^{\#\mathrm{T}} = \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{V}_y^{-1}\boldsymbol{A}\right)^{-1}$$

For this definition the resolution matrix $\boldsymbol{\Xi}$ becomes the unit matrix:

$$\boldsymbol{\Xi} = \boldsymbol{A}^{\#} \;\times\; \boldsymbol{A} = \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{V}_y^{-1}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{V}_y^{-1}\boldsymbol{A} = \boldsymbol{I}$$

i.e. the regularization error vanishes.

Study transformation properties of symmetric matrix $\boldsymbol{C} = \boldsymbol{A}\boldsymbol{V}_y^{-1}\boldsymbol{A}^{\mathrm{T}}$ by diagonalization:

$$\boldsymbol{C} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathrm{T}} \qquad \boldsymbol{\Lambda} = \boldsymbol{U}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{U} \qquad \boldsymbol{U}^{\mathrm{T}}\boldsymbol{U} = \boldsymbol{U}\boldsymbol{U}^{\mathrm{T}} = \boldsymbol{I} \qquad \boldsymbol{C}^{-1} = \boldsymbol{U}\boldsymbol{\Lambda}^{-1}\boldsymbol{U}^{\mathrm{T}}$$

with diagonal matrix $\boldsymbol{\Lambda}$ of eigenvalues in decreasing order: $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_n \geq 0$ and columns of orthogonal matrix $\boldsymbol{U} = [\boldsymbol{u}_1, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_n]$ given by the eigenvectors $\boldsymbol{u}_j$.

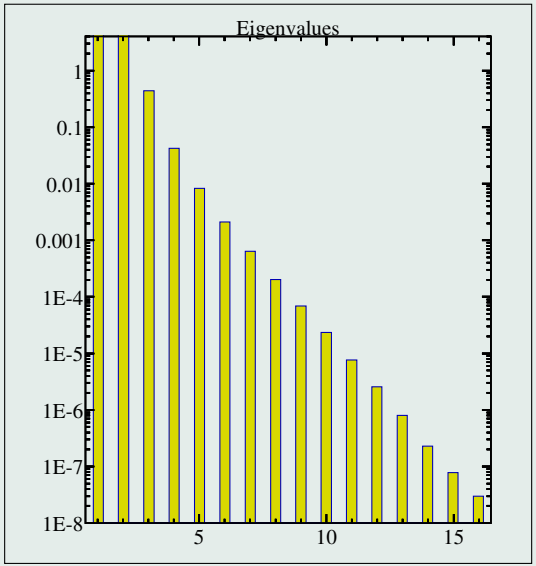
Using result from diagonalization:

$$\widehat{\boldsymbol{x}} = \boldsymbol{C}^{-1}\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{V}_y^{-1}\right)\boldsymbol{y} = \boldsymbol{U}\boldsymbol{\Lambda}^{-1/2}\;\left[\left(\boldsymbol{\Lambda}^{-1/2}\boldsymbol{U}^{\mathrm{T}}\right)\;\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{V}_y^{-1}\right)\right]\boldsymbol{y} = \boldsymbol{U}\boldsymbol{\Lambda}^{-1/2}\boldsymbol{c}$$

with vector of Fourier coefficients $\boldsymbol{c} = \left[\left(\boldsymbol{\Lambda}^{-1/2}\boldsymbol{U}^{\mathrm{T}}\right)\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{V}_y^{-1}\right)\right]\boldsymbol{y}$
i.e. the solution is $\widehat{\boldsymbol{x}}$ expressed by Fourier coefficients with $\boldsymbol{V}_c = \boldsymbol{I}$.

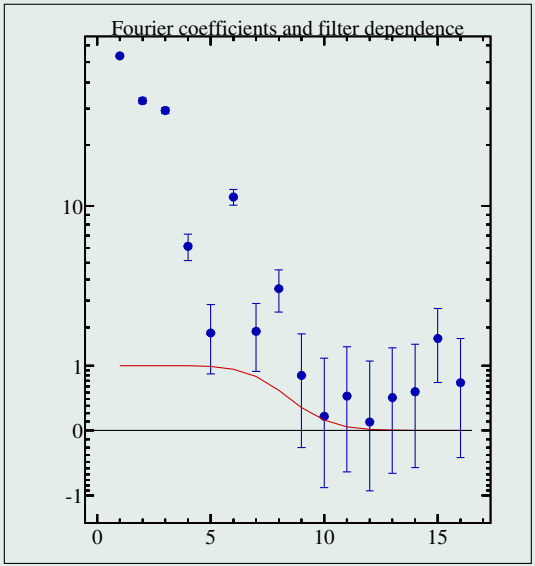
Eigenvalue spectrum and Fourier coefficients

Eigenvalues $= \lambda_j$



The eigenvalues decrease by ≈ 8 orders of magnitude, due to limited resolution.

Fourier coefficients $= c_j$



Note: the c_j are independent, and all have error $1 \Rightarrow$ only ten coefficients are significant.

The solution can be expressed by a sum in the form (with $\boldsymbol{b} = \left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{V}_y^{-1}\right)\boldsymbol{y}$)

$$\text{Solution:} \qquad \widehat{\boldsymbol{x}} = \sum_{j=1}^{n} \frac{1}{\sqrt{\lambda_j}} c_j \boldsymbol{u}_j \qquad \text{with} \quad c_j = \frac{1}{\sqrt{\lambda_j}} \left(\boldsymbol{b}^{\mathrm{T}}\boldsymbol{u}_j\right) \qquad \boldsymbol{V}_x = \sum_{j=1}^{n} \frac{1}{\lambda_j} \boldsymbol{u}_j \boldsymbol{u}_j^{\mathrm{T}}$$

Fourier coefficients c_j, which are insignificant (only noise, i.e. compatible with zero), should follow a normal distribution $N(0,1)$; this allows simple statistical tests for significance.

The insignificant coefficients can make a huge and dominating contribution to the solution $\widehat{\boldsymbol{x}}$, if the eigenvalues λ_j are small.

Alternative is the Singular Value Decomposition (SVD) (used e.g. by Höcker and Kartvelishvili in GURU), which is equivalent.

Simplest method to avoid insignificant contribution: truncation!
The truncated solution can be written with a **sum up to** $k < n$ only (instead of n) . . . the noisy contributions are removed from the result, without the introduction of a bias, . . . but the rank of $\boldsymbol{V}_x$ is only k $(< n)$ ($\Rightarrow \boldsymbol{V}_x$ is singular, unless the number of data points is reduced.).

Regularization methods

Regularization: add penalty term to the χ^2-expression, in order to make a smooth cut-off.

$$\textit{Thikhonov-Phillips:} \quad F_\tau(\boldsymbol{x}) = \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|^2 + \tau \, \|\boldsymbol{L}\boldsymbol{x}\|^2 = \text{minimum}$$

with $\boldsymbol{L} = \boldsymbol{1}$ or $=$ derivative expresssion (first, second, third). The factor τ is called regularization parameter.

Fourier coefficients $\boldsymbol{c} = \boldsymbol{\Lambda}^{-1/2}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{b}$ filter factor $\varphi_j = \dfrac{\lambda_j}{\lambda_j + \tau} = 1 \ \ldots \ 0.5 \ \ldots \ 0$

$$\text{Solution:} \quad \widehat{\boldsymbol{x}} = \sum_{j=1}^{n} \frac{\varphi_j}{\sqrt{\lambda_j}} c_j \boldsymbol{u}_j \quad \text{with} \quad c_j = \frac{1}{\sqrt{\lambda_j}} \left(\boldsymbol{b}^{\mathrm{T}}\boldsymbol{u}_j\right)$$

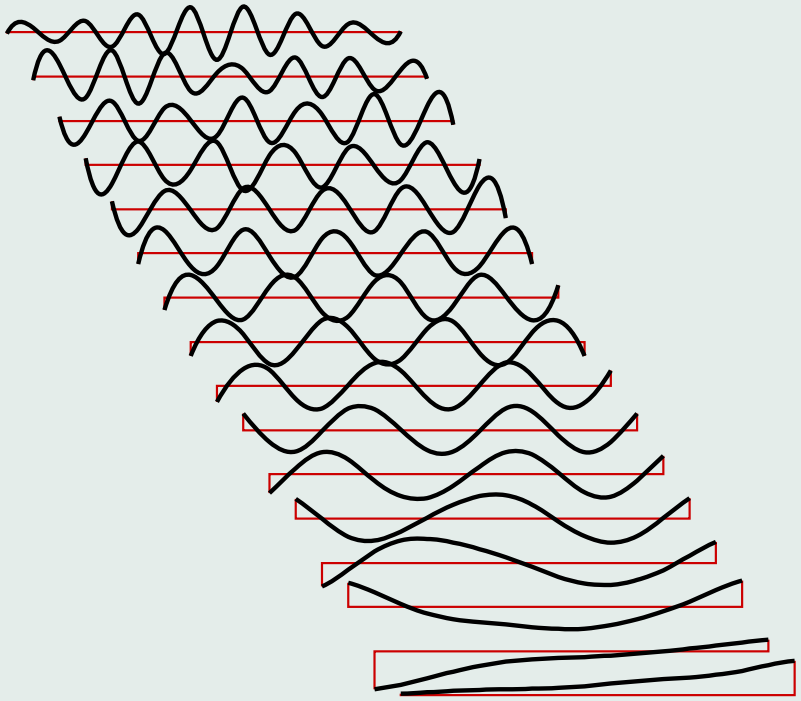
The result $\widehat{\boldsymbol{x}}$ is expressed as a superposition of eigenvectors $\boldsymbol{u}_j$, each weighted with the Fourier coefficient c_j and $1/\sqrt{\lambda_j}$ (!); the filter factor φ_j reduces the effect of insignificant contributions, without introducing a bias.

Plot of the eigenvectors $\boldsymbol{u}_j$

solution $\quad \widehat{\boldsymbol{x}} = \sum_{j=1}^{n} \frac{\varphi_j}{\sqrt{\lambda_j}}\, c_j\, \boldsymbol{u}_j$

The solution $\widehat{\boldsymbol{x}}$ is expressed as a superposition of normalized eigenvectors $\boldsymbol{u}_j$, each

- proportional to the Fourier coefficient c_j, and
- weigthted by $1/\sqrt{\lambda_j}$ (!);
- the filter factor φ_j reduces or removes the effect of insignificant contributions.



Regularization parameter I

Discrepancy principle (Morozov): choose τ such that $\|\boldsymbol{A}\boldsymbol{x}_{\mathrm{reg}} - \boldsymbol{y}\|_2^2 = n_{df}$

L-curve method: plot logs of $\|\boldsymbol{L}\boldsymbol{x}_{\mathrm{reg}}\|_2^2$ versus $\|\boldsymbol{A}\boldsymbol{x}_{\mathrm{reg}} - \boldsymbol{y}\|_2^2$ for a set of τ-values $\Rightarrow$ dependence has shape of an L with a distinct corner: optimal value of τ.

Effective number of degrees of freedom: ($\mathcal{RUN}$) determine n_{df} from spectrum of Fourier coefficients c_j, and determine τ such that sum of filter factors φ_j

$$\sum_{j=1}^{n} \frac{\lambda_j}{\lambda_j + \tau} = \sum_{j=1}^{n} \varphi_j = n_{df}$$

Minimum of global correlation: minimum mean value of global correlation coefficients

Definition: $\rho_j = \sqrt{1 - \left[\left(\boldsymbol{V}_x\right)_{jj} \cdot \left(\boldsymbol{V}_x^{-1}\right)_{jj} \right]^{-1}}$ with $0 \leq \rho_j \leq 1$

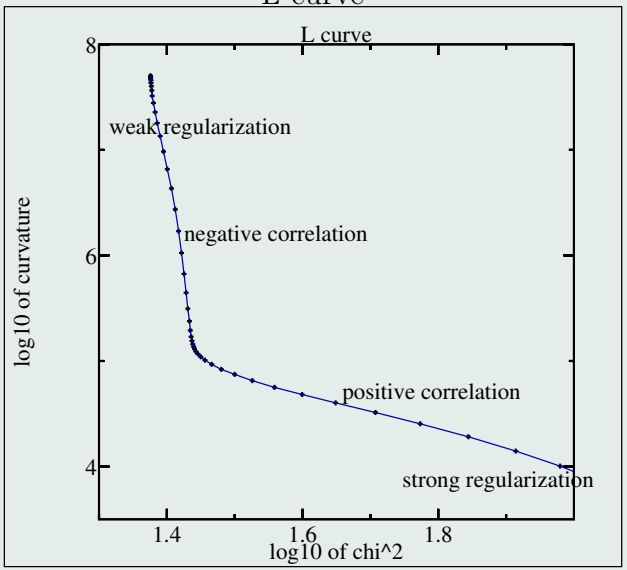
The global correlation coefficient is a measure of the total amount of correlation between element j of $\boldsymbol{x}$ and all other elements.

The arithmetic and the geometric mean of all n global correlation coefficients is determined for a large range of τ-values: the τ-value with the smallest mean value is accepted.

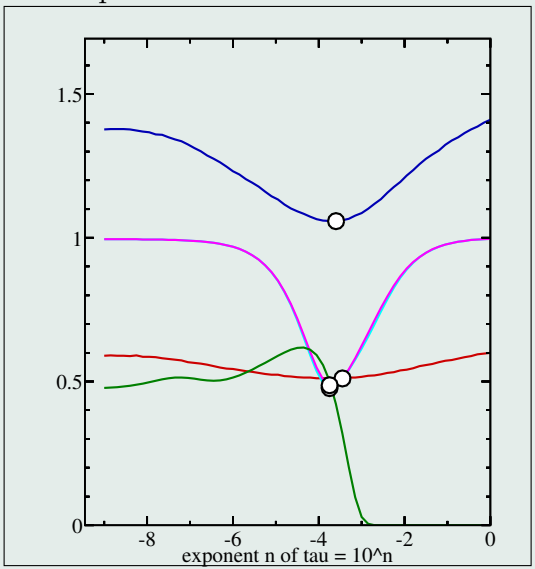
. . . seems to be the best method!

Regularization parameter II

L-curve



Dependence as a function of τ



optimal regularization in region of largest curvature

different τ-values are used along the curve

circle=minimum

magenta/cyan=mean global correlation

green=χ^2 probability

red=average probability

blue=relative sigma

Unfolding program $\mathcal{RUN}$

Development started 1979/1980 for neutral current neutrino experiment CHARM:

- First problem was reconstruction of cross section $\mathrm{d}\sigma/\mathrm{d}y$ from the measured values E_{had} and radius $r_{\mathrm{interaction}}$ (only these 2 quantities were measureable).
- Input are n-tuples, and 1-dim. or 2-dim. or 3-dim. measured histograms.
- Instead of Least Squares the ML method with Poisson statistic was used (sometimes only few entries/bin for $>$ 1-dim. histograms), with diagonalization of Hessian.
- For the intermediate result cubic B-splines were used to avoid discontinuities.
- A special option allows to check the consistency of MC simulation.
- Used in other experiments:
  Neutrino physics, 2-photon-physics at e^+e^- colliders, in astrophysics, and still used in 2010 (LHC, D0 at FNAL).
- Conversion to C++ in preparation (or ready? Talk by Natalie Milke).

3. Iterative methods

In case of unfolding with extremely large dimensions (e.g. 10^6 parameters in picture deblurring) iterative methods are preferred, because they avoid the n^3 time dependence of the direct solution methods, also the n^2 space dependence. The iterations require only a product of the – often sparse – response matrix with vectors. Because of the small dimension parameters for HEP unfolding iterative methods are not necessary.

Iterative methods have a kind of built-in regularization. In the literature iterative methods like Landweber iteration are explained. This and other methods are slowly or very slowly convergent algorithms. After an extremely large number of iterations the same result as in the direct methods without regularization is obtained.

$$\boldsymbol{x}^{(k)} := \boldsymbol{x}^{(k-1)} + \omega \boldsymbol{A}^{\mathrm{T}} \left(\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}^{(k-1)} \right) \qquad k = 1, 2, \ldots \qquad 0 < \omega < 2 \left\| \boldsymbol{A}^{\mathrm{T}} \boldsymbol{A} \right\|_2^{-1}$$

in one iteration.
Implicit filter factor $\varphi_i^{(k)}$ in iteration k for the i-th eigen-contribution:

$$\varphi_i^{(k)} = 1 - (1 - \omega\lambda_i)^k \qquad\qquad i = 1, 2, \ldots n$$

with the asymptotic behaviour for very large and very small eigenvalues:

$$\varphi_i^{(k)} \approx \begin{cases} 1 & \text{for large } \lambda_i \\ k\,(\omega\lambda_i) & \text{for } \lambda_i \ll 1/\omega \end{cases}$$

The number of iterations is equivalent to a regularization parameter.

4. Discussion – Presentation of unfolding results

- Is the unfolding result allowed to depend on the MC input dependence, or is independecne required?
- Is a large number of data points preferred, which means a rank-deficient covariance matrix and large positive correlations between data points.
- Or: Should the unfolded data points have a full-rank covariance matrix, which allows to use the inverse weight matrix in parameter fits? The full-rank requirement means a limitation of the number of unfolded data points – and a small correlation between the points.

  (The same result is easily converted to the previous case with a large number of data points, without increasing the information content.)

Is the unfolding result allowed to dependend on the MC input dependence?

My opinion: **NO**.
Do "blind analysis", whenever possible. Avoid any bias w.r.t. an expected result.

From a paper on the CFM: *"The correction of the detector acceptance using Monte Carlo modelling requires the cross section model used in the simulation to be sufficiently close to the data, such that migration between the bins are well reproduced. ... In practice this is achieved using an iterative MC event reweighting procedure which converges after one iteration for the measurement region."*

In an iterative method: *" ... it gives the best results (in terms of its ability to reproduce the true distribution) if one makes a realistic guess about the distribution that the true values follow ....*

What happens in the case of a completely insensitive detector?

Regularization methods will not be able to get any result!

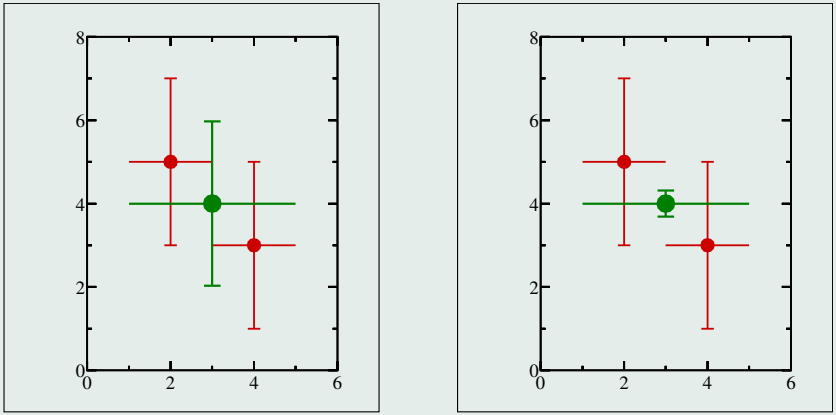
In the iterative method: *" One finds then that the final probabilities are equal to the initial ones ....*

Averaging correlated data

The Figures show two adjacent data points d_1 and d_2 with a large positive (left) and negative (right) correlation coefficient, assuming the same standard deviations: $\sigma = \sigma_1 = \sigma_2$.

$$\boldsymbol{V} = \begin{pmatrix} \sigma^2 & \rho_{12}\sigma^2 \\ \rho_{12}\sigma^2 & \sigma^2 \end{pmatrix} \qquad \text{with} \quad \rho_{12} = \pm 0.95$$

Average: the average value $\overline{\boldsymbol{d}} = \frac{1}{2}(d_1 + d_2)$ has a variance of $\boldsymbol{V}_{\overline{d}} = \frac{1}{2}(1 + \rho_{12})\,\sigma^2$ (see middle point in figures).

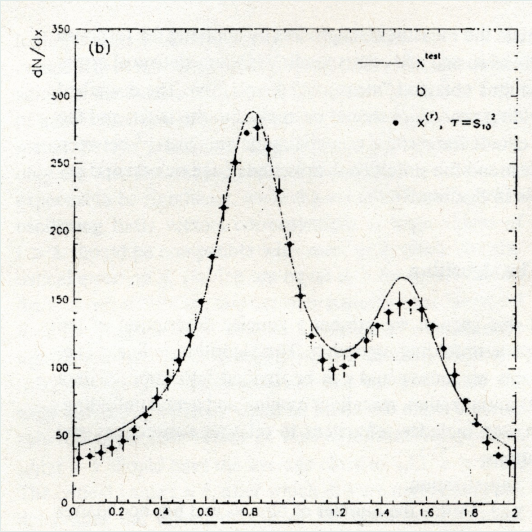


For highly correlated data the properties for averaging and χ^2-comparison with predicted values are not intuitive, but have to be performed with an explicit calculation, based on the inverse covariance matrix $\boldsymbol{V}_x^{-1} = \boldsymbol{W}_x$ (weight matrix).

Full or reduced rank of covariance matrix?

In truncation methods with $k(< n)$ elements the covariance matrix has only rank k for n data points, i.e. is singular; similar for regularization methods.
The different opinions about the rank of the covariance matrix are discussed by Hoecker and Kartvelishvili (Guru program):

"Obviously, as the number of statistically independent data points is usually smaller (and sometimes much *smaller) than the number of bins in the unfolded histogram, the latter will probably have significant bin-to-bin correlations. In our approach full propagation of errors from the measured distribution to the unfolded one is implemented, and both the covariance matrix of the unfolded solution and its inverse are easily calculated. This allows one to perform further error propagation and parameter fitting without any problem, so, contrary to the viewpoint expressed in Ref. [Blobel/*$\mathcal{RUN}$*, requires full rank], we do not think that one should use fewer bins and custom bins boundary for the unfolded hsitogram, in order to make the covariance matrix diagonal".*



40 data points and rank 10

Contents