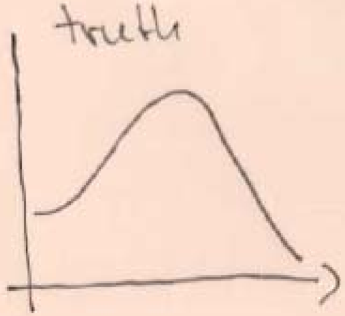
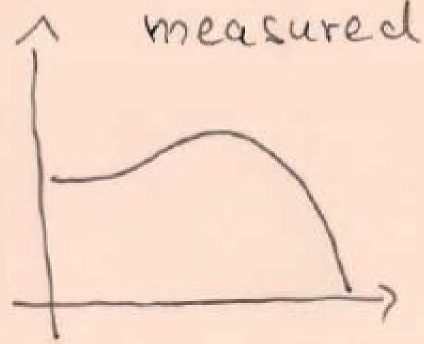
Unfolding



$$\xrightarrow[\text{+ Background}]{\text{Detector}}$$

$$\xleftarrow[\text{Unfolding}]{}$$



$$\downarrow$$

$$\begin{pmatrix} y_1 \\ \vdots \\ y_M \end{pmatrix} \qquad\qquad \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}$$

Binned data

background ("known")

$$\vec{\nu} = R\,\vec{\mu} + \vec{\beta}$$

expected true event

expected measured events

response matrix ($N \times M$) (we assume $N = M$)

("known")

"inverse problem": find estimator for $\vec{\mu}$ ($\hat{\vec{\mu}}$)

from measured data $\vec{n}$

"knowing" $\vec{\beta}$ and R

1. Easiest case: $R \approx$ diagonal

<u>correction factors</u> $C_i = \frac{\mu_i^{MC}}{\nu_i^{MC}}$ $\quad \hat{\mu}_i = C_i (n_i - \beta_i)$

2. Straightforward (but not very useful):

<u>matrix inversion</u> $\vec{\nu} = R\vec{\mu} + \vec{\beta}$

$$\Rightarrow \vec{\mu} = R^{-1}(\vec{\nu} - \vec{\beta})$$

correct! (but $\vec{\nu}$ not known!)

estimator $\hat{\vec{\mu}} = R^{-1}(\vec{n} - \vec{\beta})$

this is <u>identical</u> to maximizing Likelihood function

$$\log \mathcal{L}(\vec{\mu}) = \sum \log P(n_i; \nu_i) \qquad \nu_i = \nu_i(\mu_i)!$$

if n_i are uncorrelated

$\Rightarrow \hat{\vec{\mu}}$ is unbiased and has minimum variance!

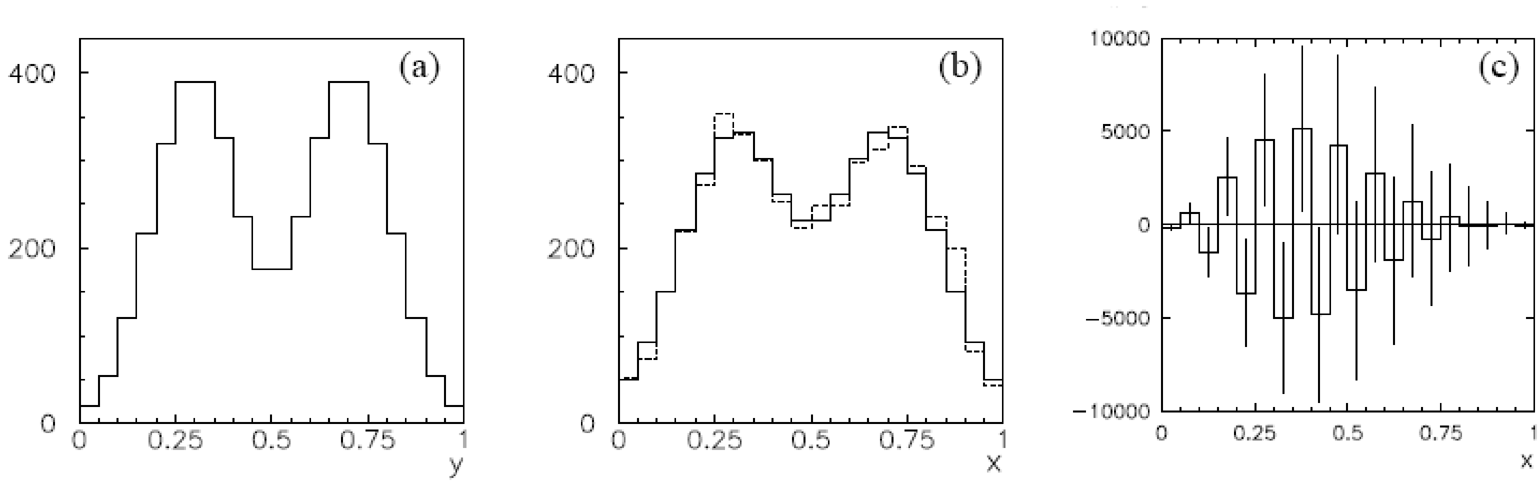


Fig. 2: Attempt to unfold using matrix inversion: (a) the 'true histogram', (b) the observed histogram n (dashed) and corresponding expectation values ν (solid), (c) the estimators $\hat{\mu}$ based on equation (8).

How badly simple matrix inversion can go „wrong" …

3. Regularization

regularization function $S(\vec{\mu})$:

sth. which is large if the solution $\hat{\vec{\mu}}$ has a desired feature,

e.g. "smoothness", "linearity", "absence of oscillations", ...

$\Rightarrow$ rather than maximizing $\log \mathcal{L}$, maximize

$$\alpha \cdot \log \mathcal{L}(\vec{\mu}) + S(\vec{\mu}) =: \Phi(\vec{\mu})$$

<u>or</u>: $\varphi(\vec{\mu}, \lambda) = \alpha \cdot \log \mathcal{L}(\vec{\mu}) + S(\vec{\mu}) + \lambda[n_{tot} - \sum_{i=1} \nu_i]$ ← normalization

choice of α: <u>trade-off</u> between best estimator and desired feature

$\hat{\vec{\mu}}_{reg}$: $\vec{\mu}$ for which $\Phi(\vec{\mu})$ = maximal

$\Rightarrow$ biased estimator!

choices for $S(\vec{\mu})$:

a) "Tikhonov"

$k = 1 \quad S(\vec{\mu}) = -\sum (\mu_i - \mu_{i+1})^2$

"flatness"

$k = 2 \quad S(\vec{\mu}) = -\sum (-\mu_i + 2\mu_{i+1} - \mu_{i+2})^2$

"curvature"

$k = 3 \quad S(\vec{\mu}) = -\sum (-\mu_i + 3\mu_{i+1} - 3\mu_{i+2} + \mu_{i+3})^2$

"skewness"

b) "max. Entropy"

$$S(\vec{\mu}) = -\sum \frac{\mu_i}{\mu_{tot}} \log \frac{\mu_i}{\mu_{tot}}$$

("abstract" (scale free) smoothness)

10. Unfolding ctd…

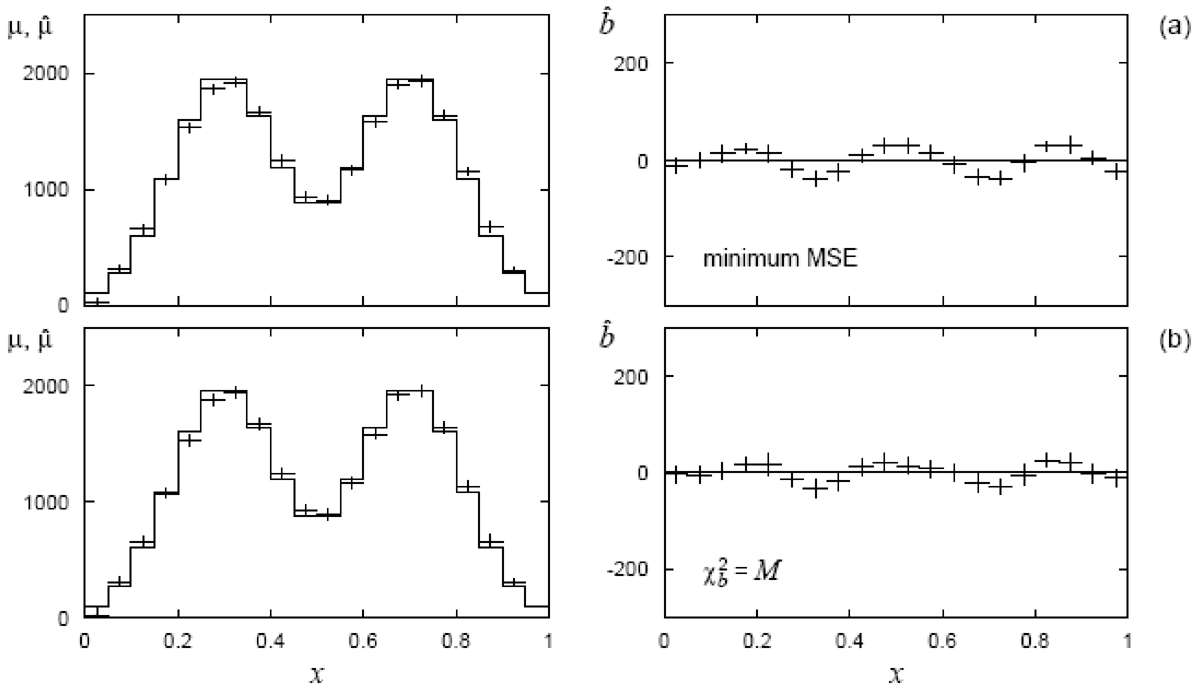


Fig. 4: Unfolded distributions using Tikhonov regularization shown as points with the true distribution shown as a histogram (left) and the estimated biases (right). The regularization parameter α was determined in (a) by the minimum mean squared error, and in (b) using $\chi_b^2 = M$ (see text).

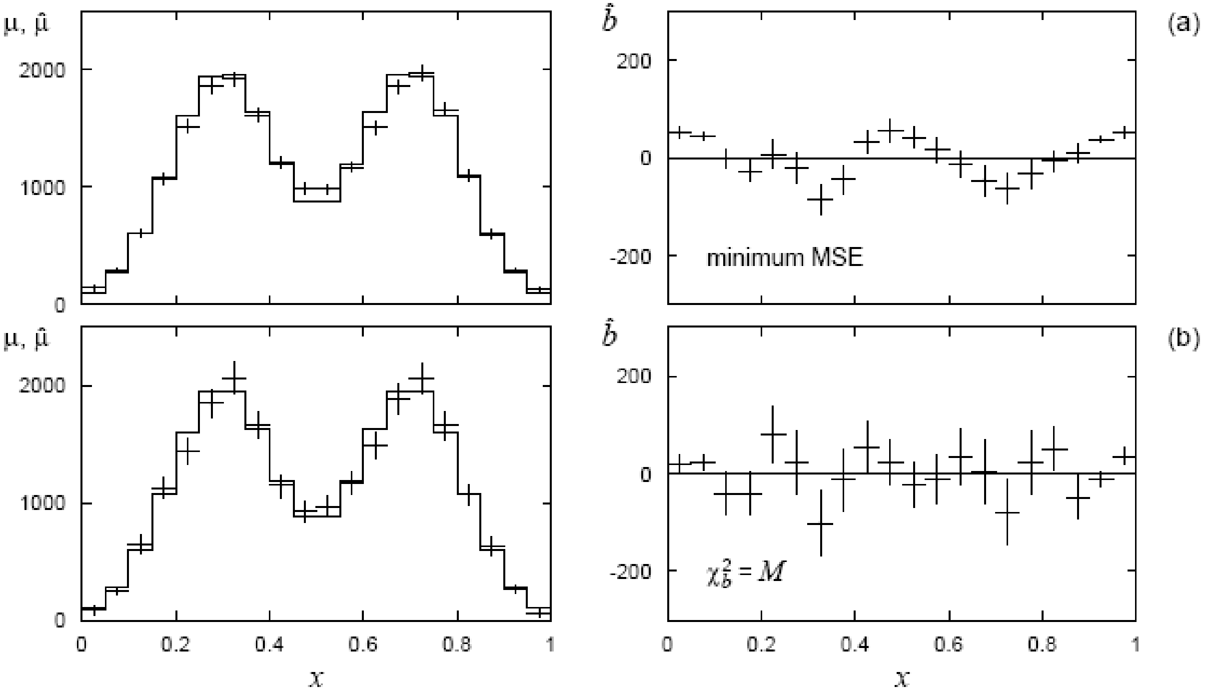


Fig. 5: Unfolded distributions using MaxEnt regularization shown as points with the true distribution shown as a histogram (left) and the estimated biases (right). The regularization parameter α was determined in (a) by the minimum mean squared error, and in (b) using $\chi_b^2 = M$ (see text).

c) "cross-entropy" $S(\vec{\mu}) = -\sum p_i \log \frac{p_i}{q_i}$

$p_i = \frac{\mu_i}{\mu_{tot}}$ q_i = "reference distribution"
(prior belief ...)

Variance and bias of reg. unfolding estimators

wanted : covariance of the estimators, $U_{ij} = \mathrm{cov}[\hat{\mu}_i, \hat{\mu}_j]$

$\vec{\mu}, \lambda$ are solutions to $M+1$ equations

$$F_i(\vec{\mu}, \lambda, \vec{n}) = 0 \qquad i = 1, \dots, M+1$$

$$F_i(\vec{\mu}, \lambda, \vec{n}) = \begin{cases} \frac{\partial \varphi}{\partial \mu_i} & i = 1, M \\ \frac{\partial \varphi}{\partial \lambda} & i = M+1 \end{cases}$$

expand F_i in a Taylor-series around the "measured" estimators $\vec{\tilde{\mu}}, \tilde{\lambda}, \vec{\tilde{n}}$:

$$\underbrace{F_i(\vec{\mu},\lambda,\vec{n})}_{0} \approx \underbrace{F_i(\vec{\tilde{\mu}},\tilde{\lambda},\vec{\tilde{n}})}_{0} + \sum_{j=1}^{M} \frac{\partial F_i}{\partial \mu_j}\Big|_{\tilde{\mu},\tilde{\lambda},\tilde{n}} (\mu_j - \tilde{\mu}_j)$$

$$+ \left[\frac{\partial F_i}{\partial \lambda}\right]\Big|_{\dots} (\lambda - \tilde{\lambda}) + \sum_{i=1}^{N} \frac{\partial F_i}{\partial n_j}\Big|_{\dots} (n_j - \tilde{n}_j)$$

$$\Rightarrow \hat{\vec{\mu}}(\vec{n}) \approx \vec{\tilde{\mu}} - A^{-1}B(\vec{n} - \vec{\tilde{n}}) \qquad \Rightarrow \frac{\partial \hat{\mu}_i}{\partial n_k} = -(A^{-1}B)_{ik} =: C_{ik}$$

$$\underbrace{\mathrm{cov}[\hat{\mu}_i, \hat{\mu}_j]}_{U} = \sum_{k,\ell=1}^{N} \frac{\partial \hat{\mu}_i}{\partial n_k} \frac{\partial \hat{\mu}_j}{\partial n_\ell} \underbrace{\mathrm{cov}[n_k, n_\ell]}$$

$$\boxed{U = C V C^T}$$

if $\log \mathcal{L}$ is known explicitely (e.g. for Poisson errors) U can be calculated wanted: <u>bias</u> $b_i = E[\hat{\mu}_i] - \mu_i$

$$b_i = E[\hat{\mu}_i] - \mu_i \approx \tilde{\mu}_i + \sum_{j=1}^{N} C_{ij}(\nu_j - \tilde{n}_j) - \mu_i$$

estimated bias:

$$\hat{b}_i \approx \sum_{j=1}^{N} C_{ij}(\hat{\nu}_j - n_j) = \sum_{j=1}^{N} \frac{\partial \hat{\mu}_i}{\partial n_j}(\hat{\nu}_j - n_j)$$

also covariance of $\hat{b}$, W, can be calculated:

$$W = (CR - 1)\,U\,(CR - 1)^T$$

use size of W to determine whether bias is significantly different from 0 $\Rightarrow$ get handle on reg. parameter α

<u>note</u>: reg. unfolding yields in general <u>biased</u> estimators

<u>but</u>: b can be zero, if the true function $\vec{\mu}$ has certain features, e.g. if it is

- flat for Tikhonov-reg, $k = 1$
- linear " $k = 2$ etc.
- equal to ref. distribution q for cross-entropy reg.

Choice of the regularization parameter

$$\varphi(\vec{\mu}, \lambda) = \underline{\underline{\alpha}} \log \mathcal{L}(\vec{\mu}) + S(\vec{\mu}) + \lambda \left[n_{tot} - \sum_{i=1}^{N} \nu_i \right]$$

α too large: $\rightarrow$ simple matrix inversion, oscillations

α too small: $\rightarrow$ large bias, force $\hat{\mu}$ towards the wanted feature (prior)

1. Mean squared error $\quad MSE = \frac{1}{M} \sum_{i=1}^{M} (U_{ii} + \hat{b}_i^2)$

2. bin weighted by its (error)$^{-2}$ $\quad MSE' = \frac{1}{M} \sum_{i=1}^{M} \frac{(U_{ii} + \hat{b}_i^2)}{\hat{\mu}_i}$

3. (in) significant bias $\quad \chi^2_b = \sum_{i=1}^{M} \frac{\hat{b}_i^2}{W_{ii}}$

in all case : vary α iteratively until desired value of
(MSE or MSE' or χ^2_b) is reached.

$\Rightarrow$ no unique recipe...

$\rightarrow$ Fig Cowan 11.2

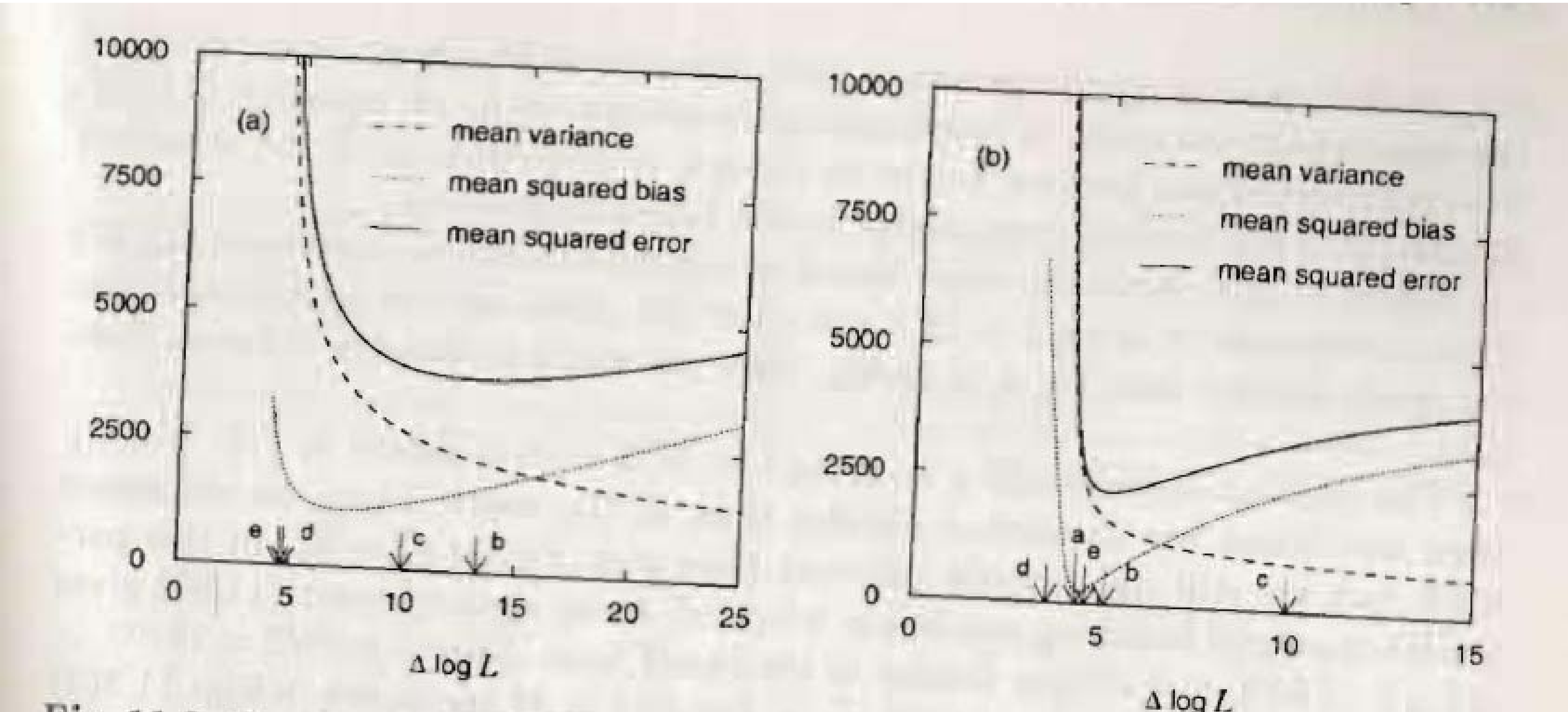


Fig. 11.2 The estimated mean variance, mean squared bias, and their sum, the mean squared error, as a function of $\Delta \log L$ for (a) MaxEnt and (b) Tikhonov regularization ($k = 2$). The arrows indicate the solutions from Figs 11.3 and 11.4: (b) is minimum MSE, (c) is $\Delta \log L = N/2$, (d) is $\Delta\chi^2_{\mathrm{eff}} = 1$, and (e) is $\chi^2 b = M$. For the MaxEnt case, the Bayesian solution $\Delta \log L = 970$ is not shown. For Tikhonov regularization, (a) gives the solution for minimum weighted MSE.

Resources

A recent workshop on Unfolding:

http://www.terascale.de/schools_and_workshops/unfolding2010/

Tikhonov regularisation (and simple correction factors)
TUnfold()   in root
http://www.desy.de/~sschmitt/

Alternative unfolding methods (iterative Bayes, single value decompostion):
http://hepunx.rl.ac.uk/~adye/software/unfold/RooUnfold.html