EPJ Web of Conferences **55**, 03002 (2013)
DOI: 10.1051/epjconf/20135503002
© Owned by the authors, published by EDP Sciences, 2013

Unfolding in particle physics: A window on solving inverse problems

Francesco Spanò

[1] *Royal Holloway, University of London*
Egham, Surrey, TW20 0EX, United Kingdom

Abstract. Unfolding is the ensemble of techniques aimed at resolving inverse, ill-posed problems. A pedagogical introduction to the origin and main problems related to unfolding is presented and used as the the stepping stone towards the illustration of some of the most common techniques that are currently used in particle physics experiments.

1 Introduction

The problem of recovering the "true", untarnished distribution of the values for a given variable from "smeared", biased and inefficient observations is common to a variety of fields.

Figure 1 shows an example from medical imaging [1]. Positron Emission Topography (PET) aims at visualizing the blood flow and metabolic activity in an organ by introducing a positron-emitting radioactive material (tracer) and by detecting X-ray photons from electron-positron annihilation ($e^+e^- \to \gamma\gamma$) when positrons emitted by the tracer annihilate with electrons from the surrounding organic matter. The reconstruction the photon emission spatial density from the detected counts provides the organ's image and it requires inverting the process outlined in Figure 1.

Images are often blurred by detector effects and corrupted by the presence of additional random noise [2]. Inverting the process shown in figure 2 is necessary to recover the details of initial image.

A variety of particle physics measurements share the same "imaging" goal. An illustrative example is the invariant mass of the pair of top-antitop quarks produced in proton-proton (pp) collisions at a center-of-mass energy ($\sqrt{s}$) of 7 TeV at the Large Hadron Collider (LHC) [3] ($pp \to t\bar{t} + X$) . This is reconstructed by measuring a complex final state involving jets of hadrons and leptons with a multi-purpose detector (in this example the ATLAS detector [4]). The inversion of the measuring process, shown in the cartoon of figure 3, is required to correct for detector (and sometimes acceptance) effects and recover the distribution of the underlying physical property. In particle physics *unfolding* is the ensemble of statistical techniques used to solve what is defined as the *inverse problem*: infer an unknown distribution $f(y)$ for a variable y from the measured distribution $g(s)$ by using knowledge and/or assumptions on the probability distribution that links the observation to the "true" value. Other terms are used that give somewhat more emphasis to recovering one specific feature of the degraded information so the techniques are also named *unsmearing* or *deconvolving*. The proper mathematical formulation of the inverse problem is the crucial step to understand the challenges involved in the unfolding procedure.

Article available at http://www.epj-conferences.org or http://dx.doi.org/10.1051/epjconf/20135503002

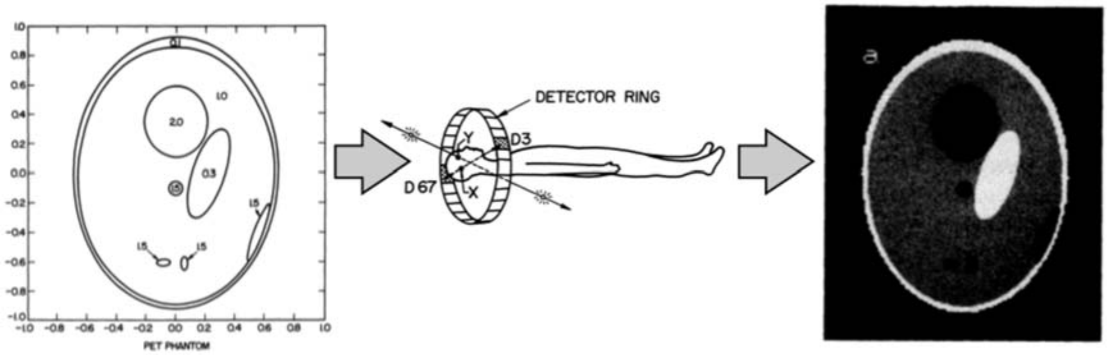


Figure 1. Scheme for the evolution of the medical imaging process using figures from Ref. [1]. The simulated photon pair emission density representing the brain (left, Figure 2) is passed to a simulation of the Positron Emission Tomography (center, Figure 1a) that produces the "observed" count distribution from the photon detector (right, Figure 3a). The names of the figures are as they appear in the reference.

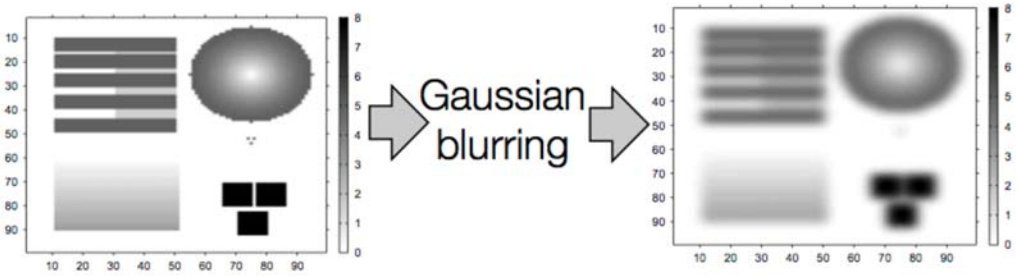


Figure 2. Scheme for the evolution of blurring and degradation of a two dimensional image using figures from Ref. [2]. The "true" simulated two-dimensional image (left, Figure 4a) is degraded by convoluting it with a Gaussian "spread" function with the addition of random Gaussian noise (see Section 4 in Ref. [2]) to produce the "observed" image (right, Figure 5A). The names of the figures are as they appear in the reference.

2 Unfolding foundations

The mathematical foundations of unfolding are intimately related to the description of the inverse problem [10] provided by the Fredholm integral equation of the first type

$$g(\mathbf{s}) = \int_{\Omega} K(\mathbf{s}, \mathbf{y}) f(\mathbf{y}) d\mathbf{y} \tag{1}$$

where the true $f(\mathbf{y})$ distribution of the variable $\mathbf{y} = (y_1,..,y_J)$ is related to the measured or observed distribution $g(\mathbf{s})$ of the variable $\mathbf{s} = (s_1,..,s_L)$ by the convolution with the *kernel* function $K(\mathbf{s}, \mathbf{y})$ [11]. In general the variables $\mathbf{y}$ and $\mathbf{s}$ belong to multidimensional spaces with different dimensions so the two integers J and L are different, in principle. The "volume" Ω represents the support of $f(\mathbf{y})$ i.e. the subspace of the multidimensional space where $\mathbf{y}$ is defined. The distribution $f(\mathbf{y})$ is transformed into the reconstructed distribution $g(\mathbf{s})$ generally because of limitations in the reconstruction of the data (biases), non-unitary and non-uniform efficiency in their collection and resolution effects.

Given the random nature of both the values of the variables to be observed and of the effects that limit their observation, retrieving $f(\mathbf{y})$ is a statistical estimation problem and the estimator needs to

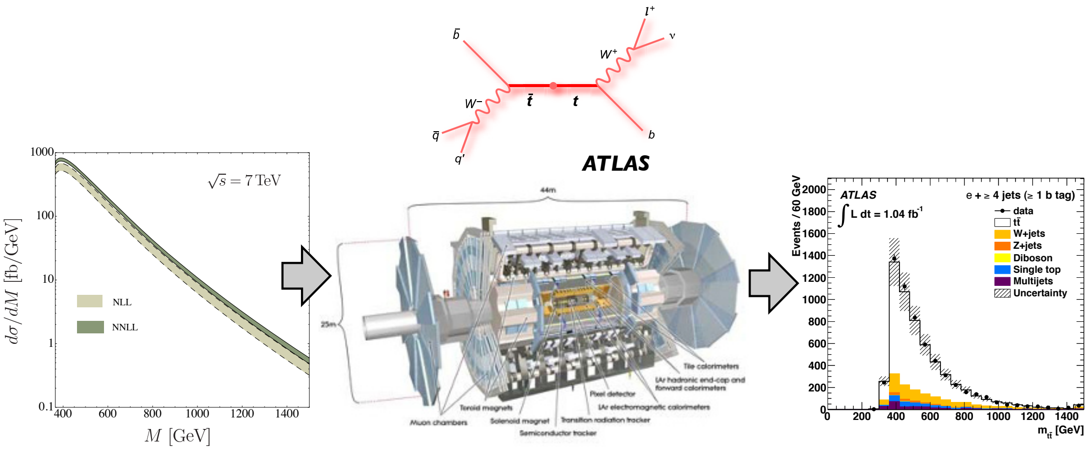


Figure 3. Scheme of evolution of the measurement of the invariant mass of the top-antitop quark system. The predicted mass distribution (left, Figure (10, right) in Ref. [5] (shown with the inclusion of theoretical uncertainties) for events featuring top-antitop quarks produced in $\sqrt{s} = 7$ TeV pp collisions at the LHC is reconstructed (right, Figure 1(a) from Ref. [6]) after the top quark decay products are measured by the ATLAS detector (middle, image from Ref. [7]). A diagram from Ref. [8] shows the final state partons from the top quark decay at leading order. The names of the figures are those by which they appear in the references.

be assessed for consistency, bias, efficiency, and robustness [11, 12]. If $f(\mathbf{y}, \mathbf{a})$ exists such that it is a prediction for $f(\mathbf{y})$ expressed as a function of a set of parameters $(a_1,..,a_P)$, the convolution of $f(\mathbf{y}, \mathbf{a})$ with $K(\mathbf{s}, \mathbf{y})$ can be compared with the observed distribution $g(\mathbf{s})$ to extract the parameter vector $\mathbf{a}$ from the data (for instance by means of a fit) and provide a complete description of $f(\mathbf{y})$. However if no parametrized prediction for $f(\mathbf{y})$ exists (as it is often the case), different techniques need to be used to estimate $f(\mathbf{y})$ from $g(\mathbf{s})$. Operatively the measurements that sample $g(\mathbf{s})$ are limited in number and affected by biases, inefficiency and imperfect resolution, so a discretized version of the integral equation 1 is used and a limited number of ingredients define the unfolding problem [13].

In the very common one dimensional case where both y and s are real variables, the measured distribution is approximated by the histogram representing the values ν_i, the expected number of counts in a given interval of s according to the definition

$$\nu_i = \int_{s_{i-1}}^{s_i} g(s)ds \tag{2}$$

where the interval of definition for s is divided in N sub-intervals by a set of $(s_1,...,s_N)$ values and any integral of $g(s)$ over a specified sub-interval provides the total number of observed events in that sub-interval.

In a similar manner the true distribution is approximated by a histogram. The range of the allowed values for y is also divided in M sub-intervals by a set of $(y_1,...,y_M)$ values and the expected number of counts in one of the sub-intervals is defined as

$$\mu_j = \int_{y_{j-1}}^{y_j} f(y)dy \tag{3}$$

The integral *kernel* $K(s, y)$ from Equation 1 is approximated by a response matrix $R(i, j)$ representing the probability that an event with a value of the y variable in bin j is observed as an event with a value of s in bin i. So Equation 1 is transformed in

$$\nu_i = \sum_{j=1}^{M} R_{i,j}\mu_j \tag{4}$$

where ν_i and μ_j are the expected number of reconstructed and "true" events in bins i and j respectively.

Consequently the first ingredient for the unfolding problem described by Equation 4 is the knowledge of the response matrix R. In general R is a rectangular matrix and by combining Equation 1 with Equation 2, it is connected to the *kernel* by the equation

$$R_{i,j} = \frac{\int_{s_{i-1}}^{s_i}\int_{y_{j-1}}^{y_j} K(s,y)f(y)dyds}{\int_{y_{j-1}}^{y_j} f(y)dy} \tag{5}$$

If the analytical formulation of the *kernel* is available, R can be determined directly from Equation 5. However most frequently R is obtained by running detailed simulation of the measuring apparatus including as many effects as possible. Monte Carlo events are generated with the best available prediction for the true distribution $f(y)$ and fully simulated with the most accurate model of the detector to produce our best guess of $g(s)$, the distribution of measured values. For some cases it is possible to measure the response to δ-like (unit-impulse) inputs that can allow to determine the *kernel* in a certain range of values, like the response of a calorimeter to a beam of particle of known energy and nature. This is equivalent to the integral $K(s,y_0) = \int_a^b K(s,y)\delta(y-y_0)dy$.

The second ingredient is the the vector of expected bin contents $\boldsymbol{\nu}$. The vector $\boldsymbol{\nu}$ is approximated by the vector $\mathbf{n} = (n_1,...,n_N)$ representing the number of observed events in each histogram bin for the variable s. By definition $\boldsymbol{\nu}$ is such that $E[n_i]= \nu_i$ where $E[n_i]$ indicates the expectation value of n_i.

Two additional ingredients are necessary to make the model built in 4 closer to reality.

First some interesting events are not observed due to inefficiencies in the detection or to the requirements imposed on the events properties. Such inefficiency is included in the estimate of the response matrix $R(i,j)$ with a proper normalization by defining

$$\sum R_{i,j} = \sum_{j=1}^{M} P(\text{observed in bin } i|\text{true value in bin } j) = P(\text{observed anywhere}|\text{true value in bin } j) = \epsilon_j \tag{6}$$

where the vector $\boldsymbol{\epsilon}= (\epsilon_1,..,\epsilon_M)$ describes the detection efficiency as a function of the histogram bin.

Secondly some of the observed events are not interesting for the measurement one wants to perform as they are due to backgrounds (events that look like the ones of interest, but have different origin) and they modify the observed distribution. Such events have their own distribution $b(s)$ in terms of the values of the observed variable s. The vector $\boldsymbol{\beta}$ of the expected number of background events in each bin of the histogram of s can be defined as

$$\beta_i = \int_{s_{i-1}}^{s_i} b(s)ds \tag{7}$$

Examples of histograms [13] featuring the vectors $\boldsymbol{\mu}$, $\boldsymbol{\epsilon}$ and the corresponding vectors $\mathbf{n}$ and $\boldsymbol{\nu}$ are shown in figure 4.

In general the model described in Equation 1 is then extended to

$$g(\mathbf{s}) = \int_{\Omega} K(\mathbf{s},\mathbf{y})f(\mathbf{y})d\mathbf{y} + b(\mathbf{s}) \tag{8}$$

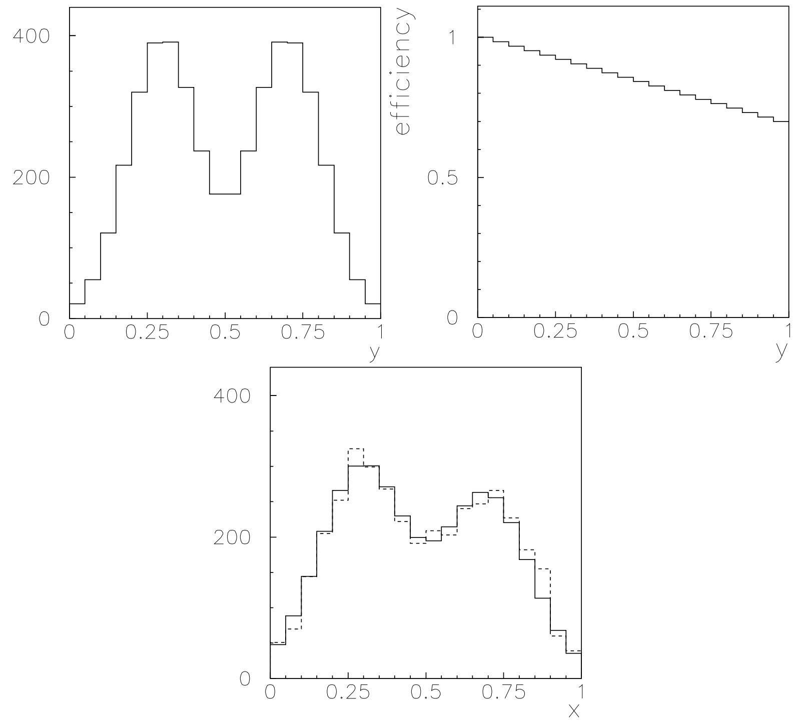


Figure 4. Examples of "true" distribution (left) ($\boldsymbol{\mu}$), a given set of efficiencies including resolution effects (center) ($\boldsymbol{\epsilon}$) and the corresponding observed (dashed, right) ($\mathbf{n}$) and expected observed distribution (solid, right)($\boldsymbol{\nu}$) [13]. The vectors $\boldsymbol{\mu}$, $\boldsymbol{\epsilon}$, $\mathbf{n}$ and $\boldsymbol{\nu}$ are defined in the text.

and its discretized one-dimensional form described in Equation 4 is consequently extended [13] to

$$E[n_i] = \nu_i = \sum_{j=1}^{M} R_{i,j}\mu_j + \beta_i \tag{9}$$

whose vectorial compact form is

$$E[\mathbf{n}] = \boldsymbol{\nu} = R\boldsymbol{\mu} + \boldsymbol{\beta} \tag{10}$$

3 The maximum likelihood solution

Given the problem described by Equation 10, the formal solution is written as

$$\boldsymbol{\mu_{est}} = R^{-1}(\boldsymbol{\nu} - \boldsymbol{\beta}) \tag{11}$$

where R^{-1} is the inverse of R. This estimate for $\boldsymbol{\mu}$ can also be derived from the principle of maximum likelihood (ML) [14]. If one assumes (fairly generally) that events are being counted in each histogram bin and that the data are consequently independent Poisson observation distributed according to

$$P(n_i|\nu_i) = \nu_i^{n_i}\frac{e^{-\nu_i}}{n_i!} \tag{12}$$

the logarithm of the global likelihood $\mathcal{L} = \prod_{i=1}^{N} P(n_i|\nu_i)$ resulting from the Poisson assumption is

$$\log\mathcal{L}(\mu) = \sum_{i=1}^{N}(n_i \log \nu_i - \nu_i - \log n_i!) \tag{13}$$

where $\boldsymbol{\nu} = \boldsymbol{\nu(\mu)}$ because of equation 10. Consequently the maximum likelihood estimator for $\boldsymbol{\nu}$ obtained by imposing $\partial\log\mathcal{L}(\mu_i)/\partial\mu_i = 0 \ \forall \ i$ is given by

$$\boldsymbol{\nu}_{ML} = \mathbf{n} \tag{14}$$

and consequently the estimate of $\boldsymbol{\mu}$ is obtained as

$$\boldsymbol{\mu}_{ML} = R^{-1}(\boldsymbol{\nu}_{ML} - \boldsymbol{\beta}) = R^{-1}(\mathbf{n} - \boldsymbol{\beta}) = \boldsymbol{\mu}_{est} \tag{15}$$

Is this solution always working ? An example shown in Ref. [13] reports a double-peaked true distribution for which the resulting ML estimate, derived according to equation 15, shows a multi-peaked shape with extremely large variances and very large anticorrelation between neighbouring bins: the estimate turns out to be very different from known input. The response matrix R for this example has sizeable non-diagonal elements and the bin size of the histogram to be "inverted" is smaller than the detector resolution encoded in the model for event migrations. Figure 5 shows the generated "true" histogram $\boldsymbol{\mu}$, the observed histogram (dashed) and the corresponding expectation values (solid) and the estimator $\boldsymbol{\mu_{est}}$.

What is happening? Insight into the reasons for the ML result can be obtained by considering an instance where the true $\boldsymbol{\mu}$ have a fine structure and the detection effects, represented by the response matrix R, dilute the true information while allowing residual structure to be present [13]. This is shown in figure 6. The application of R^{-1} aims at restoring the original histogram, according to Equation 15. If the migrations are properly modelled, the inversion returns the correct values if the input data are the expectation vector $\boldsymbol{\nu}$ of the reconstructed bin contents. However the matrix inversion is applied to one instance of the vector $\mathbf{n}$, it is not applied to its expectation value ν. As a consequence, in a suggestively descriptive way, R "assumes" that the fluctuations in $\mathbf{n}$ are the residual of a real original structure diluted by the detection effects (and not of statistical origin) and uses the given input and the available model for migrations to reconstruct μ i.e. it magnifies the fluctuations back into the result.

Independently of the large fluctuations induced by the application of the matrix inversion the maximum likelihood solution is an unbiased estimator of $\boldsymbol{\mu}$ because

$$E[\boldsymbol{\mu}_{ML}] = E[R^{-1}(\mathbf{n} - \boldsymbol{\beta})] = R^{-1}(E[\mathbf{n}] - \boldsymbol{\beta}) = R^{-1}(\boldsymbol{\nu} - \boldsymbol{\beta}) \tag{16}$$

with a covariance given by

$$U_{ML,i,j} = cov[\mu_{ML,i}, \mu_{ML,j}] = \sum_{k,l=1}^{N} R^{-1}_{i,k} R^{-1}_{j,l} cov(n_k, n_l) = \sum_{k,l=1}^{N} R^{-1}_{i,k} R^{-1}_{j,l} \delta_{k,l}\nu_k = \sum_{k=1}^{N} R^{-1}_{i,k} R^{-1}_{j,k}\nu_k \tag{17}$$

where $\delta_{k,l}$ is the Dirac δ symbol [1] and the equality $cov[n_k, n_k] = \nu$ uses the property of Poisson distributed data according to equation 12.

Under rather general condition the variance of unbiased estimators has a minimum value (effectively a lower bound) determined by the Cramér-Rao-Frechet bound [14]:

$$U^{-1}_{min,k,l} = -E[\frac{\partial^2 log\mathcal{L}}{\partial\mu_k\partial\mu_l}] = \sum_{i=1}^{N} R_{ik}R_{i,l}/\nu_i \tag{18}$$

[1] $\delta_{k,l} = 1$ only if k=l, it is zero otherwise.

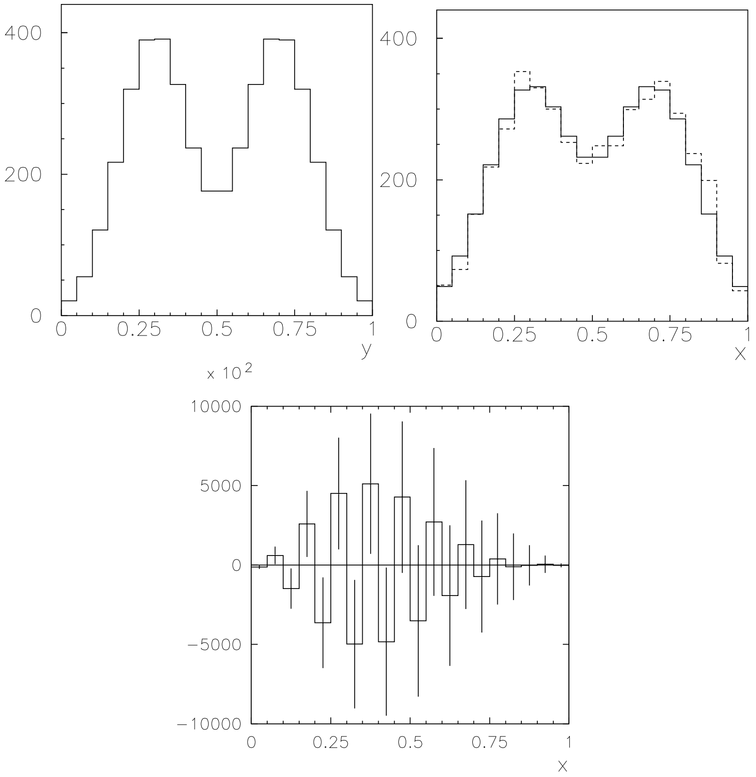


Figure 5. Examples of "true" distribution (left) ($\boldsymbol{\mu}$), the observed (dashed, middle) (**n**) and the expected observed distribution (solid, middle) ($\boldsymbol{\nu}$) assuming imperfect resolution and perfect detection efficiency, the resulting estimate for $\boldsymbol{\mu}_{est}$ using the ML solution (right) [13]. The vectors $\boldsymbol{\mu}$, $\boldsymbol{\nu}$, **n** and $\boldsymbol{\mu}_{est}$ are defined in the text.

If this equation is inverted [2] the minimum variance equals the ML variance obtained in Equation 17 i.e. $U_{i,j} = U_{min,i,j}$. Consequently the ML solution provides the unbiased estimator with the smallest variance. As a consequence estimators providing an additional reduction in variance with respect to the ML estimator will necessarily introduce a bias in the estimate of the true distribution. The balance between bias and variance is a crucial item in the unfolding procedure. Understanding the origin of the large fluctuations in the ML estimator allows to develop techniques to reduce the fluctuations (and consequently the variance of the estimator) while understanding the limitations in terms of bias of the estimator.

4 Correction factors: a "diagonal" ML

A simple step towards a small variance estimator consists in a simplification of Equation 15 derived by taking the same binning for $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ and assuming R to be diagonal (no migrations of events between bins when transforming the true distribution into the measured one). The resulting estimate for $\boldsymbol{\mu}$ is

$$\mu_{i,est} = C_i(n_i - \beta_i) \tag{19}$$

[2] One can use equations 17 and 18 to verify that $UU_{min}^{-1} = \mathbb{1}$.

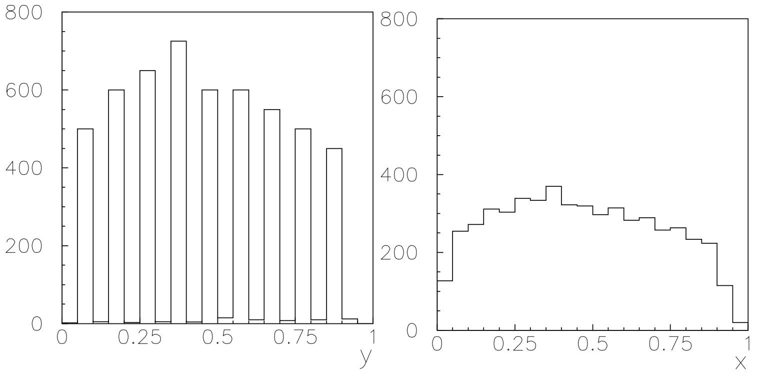


Figure 6. Examples of “true” distribution with fine structure (left) ($\boldsymbol{\mu}$) and the expected observed distribution (right) ($\boldsymbol{\nu}$) [13]. The vectors $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ are defined in the text.

where C_i are correction factors (often called “bin-by-bin corrections”) and they are usually derived from full simulation of the process under investigation. This provides an estimate for the expected number of reconstructed events μ_i^{MC} and true events ν_i^{MC} and the correction factors are simply derived as

$$C_i = \frac{\mu_i^{MC}}{\nu_i^{MC}} \tag{20}$$

The corresponding covariance matrix is estimated [13] to be

$$U_{C,i,j} = cov[\mu_i^{MC}, \mu_j^{MC}] = C_i^2 cov[n_i, n_j] \tag{21}$$

The correction factor C_i is often of order unity so the variance of the estimators is not much larger than the Poisson statistical uncertainty in the data and it is typically reduced with respect to the ML estimator uncertainty. In relation to the uncertainties in Equation 21 a simple example due to R. Cousins and reported in Ref. [15]) points out their limitations. If one assumes that, for a given bin i of the distribution to be corrected, the values are C_i= 0.1, $\beta_i = 0$ and $n_i = 100$, the estimate $\mu_{i,C}$ for the expected number of events in this bin is obtained by $C_i n_i$= 10 and the associated standard deviation is $C_i \sqrt{n_i}$=1. However this estimate maintains that only 10 of the 100 events that are observed in the bin are actually belonging to the bin, while the remaining 90 events migrated in from other bins. It is then contradictory to have a measurement with a 10% uncertainty when there are in fact only 10 events that are actually carrying information about the bin content.

The bias corresponding to this technique, defined as E[$\mu_{i,est}$]- μ_i, is estimated [13] to be

$$b = (\frac{\mu_i^{MC}}{\nu_i^{MC}} - \frac{\mu_i}{\nu_i^{sig}})\nu_i^{sig} \tag{22}$$

where ν_i^{sig}= ν_i - β_i. The bias b is zero only if the simulation provides a proper description of the (unknown) true distribution and the bias pulls the result towards the values derived by the model that is used to determine the correction factor.

Ultimately the values of C_i depend circularly on the assumed true distribution one is trying to find. In addition the bin-to-bin correlations are completely neglected and uncertainties are only diagonal. The sum of the estimated events can be different from the sum of the observed number of events, differently from the ML estimator. The reduction in statistical uncertainty is obtained in exchange for a bias on the estimated result and the actual estimate of the bias is not simple. The bias is reduced if the migration between bins are a small fraction of the bins contents i.e. if the non-diagonal elements of the response matrix R are much much smaller than unity. Another visualization of this reduction is the requirement for the bin width to be large compared to the measurement resolution. Given its limitations in terms of possibly large biases, the technique of correction factors is a good tool for an initial approximation of the results, but it is generally advisable to avoid it for general use [3]

5 Back to basics: where to from the maximum likelihood solution?

The sensitivity to fluctuations associated with the ML solution stems from the nature of equation 15 : the original Fredholm equation 1 is an intrinsically *ill-posed* or *improper* problem [10] i.e. a problem where *"large and sometimes infinite changes in the solution could correspond to small changes in the input data"* [16] [4] In this light the stability of the solution of Equation 15 with respect to fluctuations can be quantified by how the uncertainties on the inputs are propagated to the output: a quantitative figure of merit for this propagation is the maximum ratio of relative precision of the estimated solution $\boldsymbol{\mu}_{est}$ of Equation 15 to the relative precision of the measured input vector $\mathbf{d} = \mathbf{n}$ - $\boldsymbol{\beta}$, defined as

$$c(R) = max_{\mathbf{d},\delta\mathbf{d}} \frac{\delta\boldsymbol{\mu}_{est}/\boldsymbol{\mu}_{est}}{\delta\mathbf{d}/\mathbf{d}} \tag{23}$$

The quantity $c(R)$ is called the *condition* of the R matrix and it is the upper bound on the magnification factor for the uncertainties on the input to the inversion. A large value for $c(R)$ implies instability under small fluctuations in the input i.e. a significant sensitivity to "noise" in the measurement.

A deeper analysis of equation 15 illustrates the link between fluctuations and instability and exposes the origin of instability in a quantitative manner [17] by making a connection with the *condition* of the matrix to be inverted.

The first step is to perform a transformation of variables in equation 15 such that the covariance matrix $V_{\mathbf{d}}$ of the vector $\mathbf{d}$ becomes the identity matrix. In general $V_{\mathbf{d}}$ can be non-diagonal as there can be correlations between the observations in the different bins: the Poisson-based likelihood for independent observations described by Equation 12 is consequently extended to be

$$\mathcal{L} \propto e^{-\frac{1}{2}\chi^2(\boldsymbol{\mu},\mathbf{d})} = e^{-\frac{1}{2}(R\boldsymbol{\mu}-\mathbf{d})^T V_{\mathbf{d}}^{-1}(R\boldsymbol{\mu}-\mathbf{d})} \tag{24}$$

and the estimates deriving from its maximization coincide with the least squares estimate [5]. The reduction of $V_{\mathbf{d}}$ to the identity matrix allows to write the generalized likelihood of Equation 24 in terms of significances i.e. variables normalized to their uncertainties. The transformation of variables

[3] A possible exception can be some very well behaved cases with nearly diagonal response matrices where migrations effects are minimal, the expected uncertainties are well understood and the expected bias is found to be negligible in comparison to the total final uncertainties on the unfolded results (see also Section13).

[4] A simple and powerful visualization of the ill-posed problem is also given in Ref. [10]: given that the *kernel* integration in Equation 1 tends to smooth out $f(\mathbf{y})$ and to reduce its high frequency components (edges, cusps and the like), the inversion of such a procedure will inevitably enhance the high frequency features of the input.

[5] In the limit of large expected number of events each independent Poisson variable described in Equation12 tends to a Gaussian with the same mean and variance so the resulting likelihood $\mathcal{L}$ will tend to the diagonal multivariate Gaussian distribution $\mathcal{L} \propto e^{-(R\boldsymbol{\mu}-\mathbf{d})^T D_{\mathbf{d}}^{-1}(R\boldsymbol{\mu}-\mathbf{d})}$ where $D_{\mathbf{d},i,i} = \sigma(d_i)^2$, the uncertainly on y_i, and $D_{\mathbf{d},i,j} = 0$ for $i \neq j$ (see chapter 4 of [18]). A non-diagonal multivariate Gaussian likelihood will include correlations. An example of correlated variables is given in the case where the total number of events is a fixed quantity and the bin contents of a histogram are correlated and are distributed according to a multinomial distribution. In the limit of large number of observed and expected events in each bin, the multivariate generalization is a multivariate Gaussian [18].

is a rotation in $\mathbb{R}^N$ followed by rescaling. The matrix $V_{\mathbf{d}}$ is symmetric and positive definite so there exists an $N \times N$ orthogonal matrix Q ($QQ^T = \mathbb{1}$) such that $V_{\mathbf{d}} = QV'_{\mathbf{d}}Q^T$ and $V'_{\mathbf{d}}$ is an $N \times N$ diagonal matrix such that $V'_{\mathbf{d},i,i} = v_i^2 \neq$ zero and $V'_{\mathbf{d},i,j} = 0$ for $i \neq j$. The new vector $\mathbf{d}'$ is obtained by a rotation with Q and a rescaling based on v_i as follows

$$d'_i = \frac{1}{v_i}\sum_{j=1}^{N} Q_{i,j}d_j \tag{25}$$

The new rotated and normalized $\mathbf{d}'$ vector encapsulates the statistical significance of the inputs (i.e. their size in units of their uncertainty) : it takes into account the different statistical power of the equation associated to each of the N input values (see Equation 9) . The new R' matrix is also redefined accordingly

$$R'_{i,j} = \frac{1}{v_I}\sum_{k=1}^{N} Q_{i,k}R_{k,,j} \tag{26}$$

so that equation 11 is reformulated in terms of the new variables as

$$\boldsymbol{\mu}_{est} = (R')^{-1}\mathbf{d}' \tag{27}$$

and the sum of squares to be minimized equivalent to the maximum likelihood is simplified to

$$\frac{1}{2}\chi^2(\boldsymbol{\mu}, \mathbf{d}) = (R'\boldsymbol{\mu} - \mathbf{d}')^T(R'\boldsymbol{\mu} - \mathbf{d}') \tag{28}$$

The second step is to expose the decomposition of the ML solution in terms of parameters that measure the sensitivity to fluctuations in the input [10]. Such parameters can also be related to the size of the migrations described by R' (see Section 4 of Ref. [19]) i.e. the resolution and acceptance performance of the available instruments. This is done by performing a *singular value decomposition* [20] (SVD) of R' . In general a matrix R' of dimensions $M \times N$ can be decomposed as

$$R' = U\Sigma V^T \tag{29}$$

where U and V are unitary matrices ($U^TU = V^TV = \mathbb{1}$)) respectively of dimensions $M\times M$ and $N\times N$ and $\Sigma = U^TR'V$ is a diagonal matrix of dimensions $M \times N$ i.e. such that $\Sigma_{i,j} = \sigma_i$ if and only if $i = j$ otherwise it is zero. The σ_i values are called *singular values* of the matrix R', they are non not negative and can always be arranged in non-increasing order [10]. Both matrices U and V can be written in terms of their column vectors: $U = (\mathbf{u}_1,..,\mathbf{u}_N)$ and $V = (\mathbf{v}_1,..,\mathbf{v}_N)$. If R' is replaced by its singular value decomposition in the inversion and $\sigma_j \neq 0\ \forall\ j$ the result is

$$\boldsymbol{\mu}_{est} = (R')^{-1}\mathbf{d}' = (R')^{-1}(\mathbf{n}' - \boldsymbol{\beta}') = V\Sigma^{-1}U^T\mathbf{d}' = \sum_{i=1}^{N}\frac{1}{\sigma_i}(\mathbf{u}_i^T\mathbf{d}')\mathbf{v_i} = \sum_{i=1}^{N}\frac{1}{\sigma_i}c_i\mathbf{v_i} \tag{30}$$

The singular values σ_i have important properties to characterize the unfolded result. The smoother the *kernel* corresponding to R' (i.e. the higher order continuous partial derivatives it has), the faster the decay to zero of the singular values σ_i is found to be; the smaller the value of σ_i becomes, the larger the frequency turns out to be for the component σ_i corresponds to (i.e. the more oscillations are present in the functions the corresponding *kernel* is decomposed in) [10]. The coefficients $c_i = \mathbf{u}_i^T\mathbf{d}$ can be ordered by decreasing value and they decrease rapidly with the increasing index i [21]. In addition the vector $\mathbf{c} = (c_1,..,c_N)$ has unitary covariance matrix $V_{\mathbf{c}} = \mathbb{1}$ because it is obtained by multiplying the unit-covariance $\mathbf{d}'$ by the orthogonal matrix U^T. These normalized coefficients encode the significance of the corresponding contribution to the ML result. The contribution of each c_i is weighted with the inverse of the corresponding singular value σ_i: small singular values can generate large fluctuations in the final ML result [21].

The quantitative connection between the singular value decomposition and the magnification of uncertainties in the unfolded result can be found in the condition $c(R')$: this can be re-written as

$$c(R') = ||(R')^{-1}\delta\mathbf{d}||/||(R')^{-1}\mathbf{d}||/||\delta\mathbf{d}||/||\mathbf{d}|| \tag{31}$$

and it can be shown [22] that

$$c(R') = ||R'|| \cdot ||(R')^{-1}|| = \sigma_{max}/\sigma_{min} \tag{32}$$

where $||\mathbf{d}||$ is the norm of the vector $\mathbf{d}$ resulting from the Euclidean positive definite metric in $\mathbb{R}^N$. For the matrix R', the norm $||R'||$ is induced by the Euclidean norm. If A:$\mathbb{R}^N \to \mathbb{R}^N$ is a linear application with the Euclidean norm for a vector $||\mathbf{x}|| = (\sum_i x_i^2)^{\frac{1}{2}}$ defined for both $\mathbb{R}^N$ and $\mathbb{R}^M$, the norm of the matrix A is defined as $\sqrt{\text{max eigenvalue of } A^T A}$. So the *condition* of the matrix R' can be read off from its singular value decomposition that is connected to the sensitivity to fluctuations in the unfolding problem.

The overall picture is now clearer. The singular value decomposition gives insight into the unfolding problem: ML estimators are sensitive to small effects that can lead to large changes in their values. Once the problem is described in terms of uncertainty normalized variables, the large sensitivity to small fluctuations (i.e. high frequency components, in Fourier-like language) can be derived from the high condition number $c(R)$ for the response matrix that describes the unfolding problem. In order to pose the problem more properly, it is then necessary to reduce the the impact of the low significance, highly oscillating input components while preserving the information available in the remaining high significance, more stable components. The problem is then said to have been "regularized". As the ML estimator is unbiased according to the discussion of Section 3, regularization inevitably leads to accepting a certain level of bias in exchange for a reduced variance. The bias is defined as the difference between the expected value of the unfolded result and the true unmeasured expected value. The heart of unfolding problems lies in understanding the balance between bias and uncertainty.

6 Regularized unfolding: a general view

The likelihood formulation of the unfolding problem in Equations 13 and 24 quantifies the distance between the data vector $\mathbf{n}$ and the expectation vector $\boldsymbol{\nu}$. According to that distance, in a neighbourhood of the ML solution in $\mathbb{R}^N$ the values of $\boldsymbol{\mu}$ are such that

$$\log\mathcal{L}(\boldsymbol{\mu}) \geq \log\mathcal{L}_{max} - \Delta\log\mathcal{L} \tag{33}$$

In order to filter out a certain amount of the high frequency components of the input and alleviate the sensitivity to large fluctuations, this distance definition can be modified with the goal to single out a modified solution that is still "close" to the unbiased ML estimate, but less sensitive to fluctuation. A transparent way to carry out such modification is to impose constraints on the initial likelihood by adding Lagrange multipliers and describing the regularization as a maximization procedure for a new likelihood ϕ.

The logarithm of the new likelihood to be minimized then becomes

$$\phi = \alpha\log\mathcal{L}(\boldsymbol{\mu}) + S(\boldsymbol{\mu}) \tag{34}$$

or

$$\phi = \log\mathcal{L}(\boldsymbol{\mu}) + \tau S(\boldsymbol{\mu}) \tag{35}$$

where $\mathcal{L}(\boldsymbol{\mu})$ is the initial likelihood (for instance from either Equation 13 or Eq. 24), $S(\boldsymbol{\mu})$ is called regularization function, α and τ are the regularization parameters that allow to tune the strength of the constraints (equivalent a special choice of $\Delta\log\mathcal{L}$). In addition, it is possible to add the constraint that $n_{tot} = \sum_{i=1}^{N} \nu_i$ if the solution is required to provide an unbiased estimate of the total number of events. This results in the maximization of

$$\phi = \alpha\log\mathcal{L}(\boldsymbol{\mu}) + S(\boldsymbol{\mu}) + \lambda(n_{tot} - \sum_{i=1}^{N} \nu_i) \tag{36}$$

as a function of λ and $\boldsymbol{\mu}$. It should be noted that $\sum_{i=1}^{N} \nu_i$ is a function of μ_i as $\nu_i = \sum_{i=1}^{N} R_{i,j}\nu_j + \beta_i$. The regularization function is often perceived as a measure of the level of "smoothness" required of the maximum likelihood solution. In this light, taking for instance the formulation of Equation 34, if α is set to zero, the solution is set to the smooth function encoding all the constrains (i.e. available pre-existing information): the shape of $S(\boldsymbol{\mu})$ is imposed as the correct one and the data are ignored. If α tends to infinity (i.e. α is much larger than any of the other coefficients) $S(\boldsymbol{\mu})$ carries no weight in the maximization and the ML solution is re-obtained.

In the explicit formalism the ingredients for the regularization of a given likelihood $\mathcal{L}(\boldsymbol{\mu})$ are the regularization function $S(\boldsymbol{\mu})$ and a prescription for α to tune the level of filtering for the high frequency components of the input.

7 Regularized unfolding: the Tikhonov scheme

An analytic and quantitative measure of the smoothness of the unfolding solution is the mean square of the k^{th} derivative proposed by Tikhonov and Arsenin in Ref. [23]. The proposed form for the regularization function S is then

$$S[f(y)] = \int (\frac{d^k f(y)}{dy^k})^2 dy \tag{37}$$

with k in an integer number. If $k = 2$ is chosen, Equation 37 can be approximated by a sum over the numerical estimate of second derivative [24]

$$S(\boldsymbol{\mu}) = -\sum_{i=1}^{M-2} [(\mu_{i+2} - \mu_{i+1}) - (\mu_{i+1} - \mu_i)]^2 \tag{38}$$

where M is the number of values used to describe the regularization function or the number of bins used to provide its discrete description. In matrix notation it is possible to re-write $S(\boldsymbol{\mu})$ as

$$S(\boldsymbol{\mu}) = (C\boldsymbol{\mu})^T(C\boldsymbol{\mu}) \tag{39}$$

where C is the $M \times M$ matrix that encodes the definition of the second order numerical derivatives (see Section 6 in [19]) [6].

In the limit of large expected and observed number of events for the distribution of interest the logarithm of the likelihood to be maximized results from combining Equations 24, 34 and 33 into

$$\phi(\boldsymbol{\mu}, \tau) = -\frac{1}{2}\chi^2(\boldsymbol{\mu}) + \tau S(\boldsymbol{\mu}) \tag{40}$$

[6] In general a different form for C allows to use a different regularization function that is also quadratic in$\boldsymbol{\mu}$

The combined likelihood $\phi(\boldsymbol{\mu}, \tau)$ is a quadratic function of μ given the definition of χ^2 in Equation 24 and of $S(\boldsymbol{\mu})$ in Equation 38. Consequently the first partial derivatives with respect to μ and τ to be solved to minimize $\phi(\mu, \tau)$ return a system of linear equations.

Similarly to Section 5 the first step is a linear transformation such that the input variables $\mathbf{d}$ are normalized to their uncertainties and their new covariance matrix is $\mathbb{1} \in \mathbb{R}^N$ (diagonal with unitary elements). Consequently the value of $-\frac{1}{2}\chi^2(\boldsymbol{\mu})$ takes the form reported in Equation 28. The minimization of Equation 40 is then equivalent to finding the solution to the problem represented as

$$\begin{pmatrix} R'\mu \\ \sqrt{\tau}C(\mu) \end{pmatrix} = \begin{pmatrix} \mathbf{d}' \\ \mathbf{0} \end{pmatrix} \tag{41}$$

which can also be re-written as

$$\begin{pmatrix} R'C^{-1} \\ \sqrt{\tau}\mathbb{1}) \end{pmatrix} = \begin{pmatrix} \mathbf{d}' \\ \mathbf{0} \end{pmatrix} \tag{42}$$

As third step the product $R'C^{-1}$ can be expanded by a singular value decomposition like in Equation 29 and the solution of the problem can be written in terms of such expansion like in Equation 30. The major difference is the presence of the τ-dependent constraint. In order to incorporate τ in the solution of Eq. 42, a special linear transformation is performed, called Givens rotation [25] : this coordinate transformation sets the lower diagonal block proportional to τ to zero while transferring the information about the τ variable to the upper block. In this way the solution can be expressed as a function of τ and of the solution for τ=0 [19, 21]. The final result [21] is

$$\boldsymbol{\mu}_{est} = \sum_{i=1}^{N} \frac{1}{\sigma_i}(\tilde{\mathbf{u}}_i^T \mathbf{d}')\tilde{\mathbf{v}}_i = \sum_{i=1}^{N} \frac{1}{\sigma_i}\phi_i \tilde{c}_i \tilde{\mathbf{v}}_i \tag{43}$$

with $\phi_i(\tau)$= $\frac{\sigma_i^2}{\sigma_i^2+\tau}$ and $R'C^{-1}$ is SVD-decomposed as $R'C^{-1} = \tilde{U}\Sigma\tilde{V}^T$. The small values of σ_i are now regularized by the presence of τ so that they do not cause large fluctuations. The τ parameter acts like a cut off for a low pass filter to single out the highest frequencies causing the most rapid fluctuations. When σ_i is much smaller than τ the coefficient $\phi_i(\tau)$ behaves like σ_i/τ instead of behaving like $1/\sigma_i$ (see Equation 30) so ϕ_i tends to zero instead of tending to infinity and the impact of these "frequencies" is drastically reduced. It is the additional assumption on the smoothness of the solutions that reduces the importance of the solutions that result from highly oscillating solutions. While the C matrix is set by the assumption on the derivatives, the value of τ is optimized by the properties of the problem at hand. The choice of the value for the τ parameter is discussed in detail in Sections 4.3 and 7 of Ref. [19]. Here we report just the salient concepts. The reduction of the impact of higher-frequency, less statistically significant, more noise-like components is a powerful criterion for the choice of τ. The significance of each component can be read off from the coefficients of equation 43 similarly to the general discussion in Section 5. Reference [19] uses the components of the covariance-normalized, SVD-rotated input vector $\mathbf{w}$ defined as

$$\mathbf{w} = \tilde{U}\mathbf{d}' = \sum_{i=1}^{N} \tilde{\mathbf{u}}_i^T \mathbf{d}' \tag{44}$$

The components w_i of $\mathbf{w}$ that are of order unity or less are considered to represent statistically insignificant contributions and given that the impact of the components of σ_i larger than τ is suppressed according to equation 43, setting τ equal to the value of σ_k^2 for k such that $w_k \lesssim 1$ is one way to meet the chosen criterion. Additional optimization for the choice of the value of τ is possible [7], for instance by using the τ that gives that best least-squares-based comparison between a generated and the unfolded version fully simulated model for the problem under consideration. Figure 7 shows an "academic" example [19] of unfolding simulated events with this instance of Tikhonov regularization: it reports the response matrix, the superposition of the true, reconstructed and unfolded distributions, the distribution of the w_i values and the difference between the true and the unfolded result. In the technical implementation of this version for Tikhonov regularization, a correction function is used to scale the simulated truth shape of the truth distribution to obtain the unfolded result. The minimization constraint is actually imposed on the curvature of this shape correction function. As a consequence the normalization information is kept in the response matrix which is no more probability-base like in Equation 4, but it is related to the actual number of simulated events so that the statistical uncertainties on the knowledge of the matrix elements are properly kept into account in the final result.

8 Iterative unfolding

A different approach to including information about the expected true distribution to counter the enhancement of fluctuations is obtained with an iterative approach.

As pointed out in Ref. [26], the initial idea of an iterative technique for solving ill-posed problems dates back to at least 40 years ago [28]. The general description outlined by Ref. [28] provides a still valid basis to understand the core concepts of a large number of the iterative schemes used at present. The inspiration is derived from Bayes' theorem [29] where the sets involved in the formulation are named *obs* and *true* to hint respectively at the observed and true number events associated to a given property [8]. In this light Bayes' theorem can be written as

$$P(true|obs) = \frac{P(obs|true)f(true)}{\int f(true)P(obs|true)dtrue} = \frac{P(obs|true)f(true)}{g(obs)} \tag{45}$$

where $P(x|y)$ is the conditional probability that a variable x has a certain value given the value of the variable y is between y and y+dy, $f(x)$ is a probability density for x and $g(obs)$ is defined as

$$g(obs) = \int f(true)P(obs|true)dtrue. \tag{46}$$

Inverting equation 45 and using the normalization properties of P(x|y), one obtains

$$f(true) = \int g(obs)P(true|obs)dobs \tag{47}$$

Equation 47 looks like the "inverse" of 46: it should be noted that $P(true|obs)$ is actually a function of $f(true)$ itself. The proper inverse *kernel* for f(true) needs to be a function of P(obs|true) only.

Equation. 45 provides the ansatz that if an initial hypothesis is made on $f(true)$, it is possible to use $P(obs|true)$ estimated from simulation to evaluate $P(true|obs)$ by defining $g(obs)$ as the convolution of $f(true)$ and $P(obs|true)$. The estimate of $f(true)$ can be re-used as initial hypothesis for an updated estimate and the procedure can be iterated. So in the the r^{th} iterative step, using Equation 46, $g^r(obs)$ is defined as

$$g^r(obs) = \int f^r(true)P(obs|true)dtrue \tag{48}$$

[7] The current implementation of the Tikhonov scheme with n=2 used in the ROOT Unfolding framework (RooUnfold) [51] only allows to select τ= s_k^2 without any additional modification.

[8] In general *obs* and *true* can be considered names of variables.

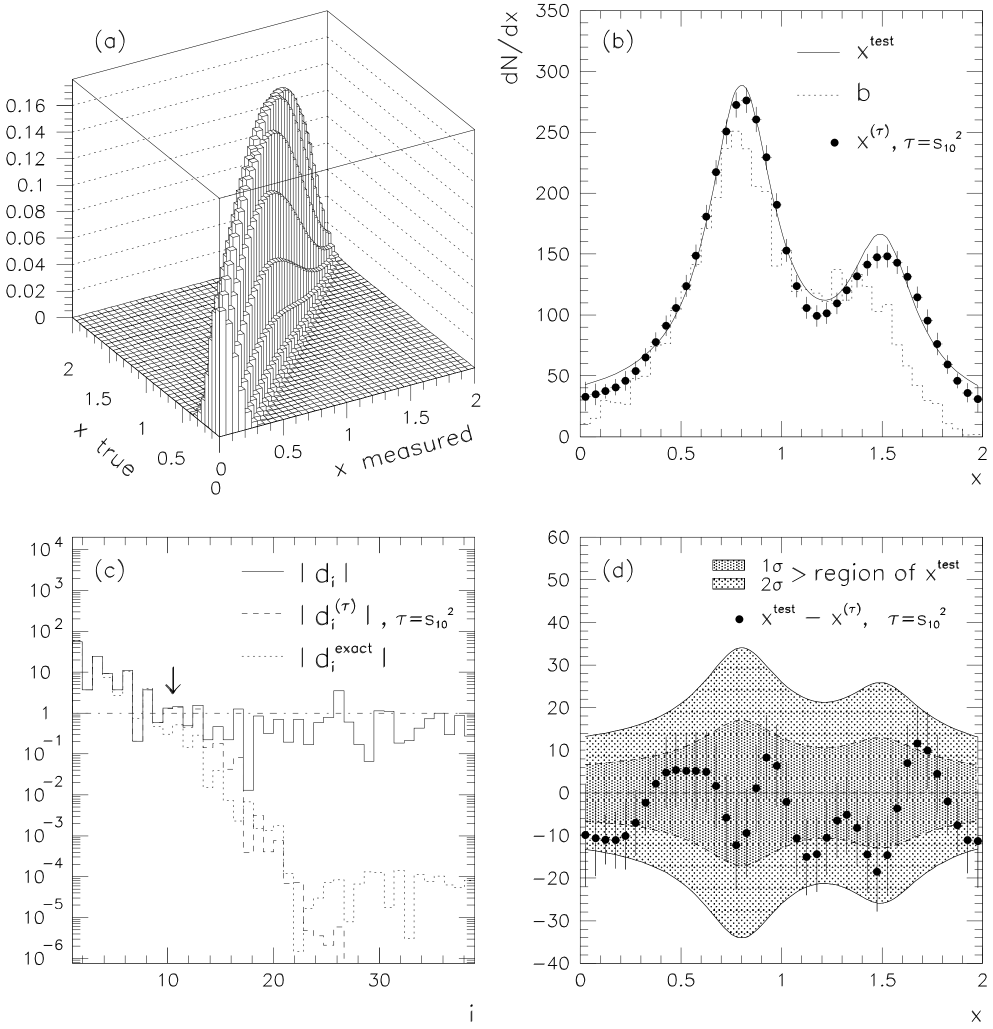


Figure 7. All the ingredients and results of an unfolding problem resolved with the n = 2 Tikhonov scheme from Ref. [19] (a) The response matrix connecting the true distribution to the measured one by encapsulating the model for the detection performance (b) The superposition of the true distribution (solid curve, labelled x^{test}), the measured distribution (dotted curve, labelled b) and the unfolded distribution (dots, labelled x^{tau} with the choice of $\tau = s_{10}^2$, the tenth singular value c) The superposition of three versions for the absolute value of the covariance normalized, SVD-rotated input vector called: the unregularized values (solid, labelled $|d_i|$), the regularized values(dashed, labelled d_i^τ, for $\tau = s_{10}^2$) corresponding to equation 44 **w**, the arrow points to the boundary between significant and non-significant components of the unfolded solution. (d) the difference between the true distribution (x^{test}) and the unfolded result (x^τ) showing the one and two standard deviation statistical bands of the true distribution.

and consequently Equation 45 returns

$$P^r(true|obs) = \frac{f^r(true)P(obs|true)}{g^r(obs)} \tag{49}$$

Using the ansatz of Equation 47, the estimate $P^r(true|obs)$ is then convoluted with the observed $g(obs)_{data}$ that is estimated from data. So a new estimate for $f(true)$, the starting point for the $(r+1)^{th}$ step, is then obtained by using Equation 49 as follows:

$$f^{r+1}(true) = \int g(obs)_{data}P^r(true|obs)dobs = \int f^r(true)\frac{g(obs)_{data}}{g^r(obs)}P(obs|true)dobs \tag{50}$$

The iteration ends at the step r for which modifications to the value of $f^{r+1}(true)$ introduced by additional steps are smaller than a given tolerance value. An equivalent statement is that the iterative scheme converges if there is a step r such that $g(obs)_{data} = g^r(obs)$ [28]. The integration in Equations 50 will tend to remove deviations from unity in $g(obs)_{data}/g^r(obs)$ that are present on a large scale compared to the support of $P(obs|true)$. On the other hand deviations on a small length scale compared to the support of $P(obs|true)$ will be tend to be averaged out in the integration. This means that the scheme is sensitive to “long wavelength” errors in $g^r(obs)$ that are usually corrected for in the initial iterations by incorporating all the useful information on the dataset. Further iterations will take more and more into account “shorter wavelength” errors more likely deriving from statistical fluctuations in $g(obs)_{data}$: the resulting corrections will tend to match $g^r(obs)$ to those fluctuations rather than to the proper $g(obs)$ value corresponding to the best estimate of $f(true)$.

The discretized version of the iteration technique described in Ref. [26] is based on the assumption that the response matrix R is positive definite and the input binned distribution y_i in non negative so that the relation $\mathbf{d} = \mathbf{n} - \boldsymbol{\beta} = \mathrm{R}\boldsymbol{\mu}$ (from Section 3) can be inverted iteratively. The starting point is an hypothesis-guess on $\boldsymbol{\mu}$ called $\boldsymbol{\mu}^0$ to produce the first estimate $\mathbf{d}^0 = \mathrm{R}\,\boldsymbol{\mu}^0$. At the r^{th} step of the iteration the estimate for the vector $\mathbf{d}$ is

$$d_i^r = \sum_{j=1}^{N} R_{i,j}\mu_j^r \tag{51}$$

which is the discrete form of 48 with the correspondence $R_{i,j} \rightarrow P(obs|true)$, $\mu_j^r \rightarrow f^r(true)$. This step can be defined "folding" and it corresponds to equation 48. Consequently the r^{th} estimate of $\boldsymbol{\mu}$ is obtained as in Equation 50 by "integrating" the response matrix $R_{i,j}$ over the updating function d_i/d_i^r

$$\mu_j^{r+1} = 1/\epsilon_j \sum_{i=1}^{N} R_{i,j}\mu_j^r(d_i/d_i^r) \tag{52}$$

where $\mathbf{d}$ is the data input vector and corresponds to $g(obs)_{data}$ in Equation 50. The values of ϵ_j represent efficiencies corrections to account for experimental acceptance losses. This step can be defined as the“unfolding” one and it corresponds to equation 50. From Equation 52, the convergence of the iteration is linked to the updating function y_i/y_i^r approaching unity: in fact, given the normalization condition $\sum_i R(i,j) = 1$, $y_i/y_i^r \rightarrow 1$ implies $\mu_j^{r+1} \rightarrow \mu_j^r$. If the uncertainties are Poisson distributed such an iteration technique is empirically found to converge to the ML solution [30].

8.1 Iterative Unfolding: a recent example

An implementation of the general iterative approach is the technique of Ref. [31]. The standard basic iterative steps outlined above are carried out. In this approach the updating function of equation 52 is obtained by defining the problem in terms of the relation of "causes" to"effects". The "causes" are defined as the content C_i of the bins of the given, unknown true distribution distribution one wants to recover; the "effects" are the contents of the bins of the observed distribution E_i. The connection between the "causes" and the "effects" is obtained by the simulation-derived response matrix $P(E_J|C_i)$, the probability that a given cause C_i results into a given effect E_j (coherently with to the definition of response matrix in Section 2). The losses due to limitations in the observations are represented by a common"trash" bin and need to be recovered by efficiency corrections, again estimated by simulation. It is emphasized that the variables the iteration estimates are the expected contents of the bins C_i, the probability that a certain fraction of the total events are found in a given bin, not the overall probability for the distributions. So instead of writing Bayes theorem for a given distribution (spectrum) in the form of

$$P(x_C|X_E, R, I) \propto P(X_E|x_C, R, I)P(x_C|I) \tag{53}$$

one restarts from the probability of "causes" (i.e. the bin contents) and so each "cell" (i.e. bin content) of the discretized distribution is considered an independent cause of an effect and the probability is written for the bins as

$$P(C_i|E_J, I) \propto P(E_j|C_i, I)P(C_i|I). \tag{54}$$

With these definitions choosing $P(C_i|I) = p_0 = 1/M$ means that the probability content of a given cause i.e. a bin is a constant over the bins. This does not mean that all spectra have the same probably (it is not a statement on the probability of the distribution), on the contrary it implies a flat, precisely determined initial spectrum and consequently a very strong prior statement that would bias the result if no additional information were used: this is the starting point and the motivation for iterations. Due to this procedure the final estimate for the unfolded distribution is not expressed in terms of a given prior.

The iteration is then such that

$$P(C_i|E_j, I) \propto P(E_j|C_i, I)P(C_i|I). \tag{55}$$

The first estimate of the number of events $n(C_i)$ in the bin i of "causes" can be written as

$$n(C_i) = \frac{1}{\epsilon_i} \sum_{j=1}^{N_E} n(E_j)P(C_i|E_j) \tag{56}$$

This can also be written as

$$n(C_i) = \sum_{j=1}^{N_E} M_{i,j} n(E_J) \tag{57}$$

with

$$M_{i,j} = \frac{P(C_i|E_J, I)P_0(C_i)}{[\sum_{l=1}^{n_E} P(E_J|C_i)][\sum_{l=1}^{n_C} P(E_J|C_i)P_0(C_l)]} \tag{58}$$

where, in connection with Equation 52, $n(C_i)$ is corresponds to μ_j^1, $n(E_j)$ corresponds to d_i, $P(C_i|E_J, I)$ corresponds to $R(i, j)$, $P_0(C_l)$ corresponds to μ_i^0 and the denominator corresponds to d_i^1. The additional iterative steps follow exactly the evolution of the discrete scheme outlined in Section 8 (see the description of Equations 51 and 52). A distinctive feature of this approach is that the estimated distribution can be (and very often is) smoothed at each iteration step before the "unfolding" step by a customizable polynomial fit to a problem-specific function. In principle any smoothing technique can be applied. This smoothing is not applied in the last step of the iteration. The iteration is continued until the solution is considered stable. The criterion for stability is dependent on the analysis; one suggested possibility is to quantify the agreement with the previous iteration by means of a least squares test.

An example from particle physics of the usage of the iterative unfolding technique can be found in the measurement of the distribution of the difference between the absolute rapidities ($\Delta|y_t|$) of the reconstructed top quark and anti-top quark in a sample enhanced in $t\bar{t}$ events obtained by LHC pp collisions at $\sqrt{s} = 7$ TeV [6]. In the standard model of particle physics this distribution is expected to show a slight asymmetry (at the sub-percent level) in the amount of events with positive and negative differences [6, 32]. The asymmetry is obtained by integrating the unfolded differential $t\bar{t}$ production cross section as $\Delta|y_t|$. The migration matrix is shown in figure 8. The measured and unfolded distribution for one specific set of events (featuring a single electron plus jets with at least one b-tagged jet) is shown in figure 8. A set of basic tests are performed to choose the number of iterations, the statistical and systematic uncertainty are propagated through the unfolding scheme and the stability and robustness of the procedure is tested. The number of iterations is such that the expected variation of the value for the asymmetry is stable within 0.1% in simulated $t\bar{t}$ events. The statistical uncertainty estimate is determined with simulated experiments and the systematic uncertainty is propagated to both the response matrix and the background. Simulated $t\bar{t}$ events are re-weighted so produce different samples with different true asymmetry. The analysis is performed on each different sample and the set input asymmetries is then plotted versus the resulting reconstructed asymmetries after unfolding to check the linearity of the unfolding procedure. The small biases observed in the reconstructed distributions and the extracted asymmetry are quantified by the largest relative deviation over all the bins and the mean uncertainty-normalized relative difference between true and unfolded values from the pull distributions, respectively. Such values are used to assign additional systematic uncertainties to the unfolded distributions and the final asymmetry.

9 Regularization unfolding and Entropy

Information theory can provide an alternative and deeply meaningful form for the regularization function of equation 34. Shannon's information entropy [33] for a given distribution is defined as:

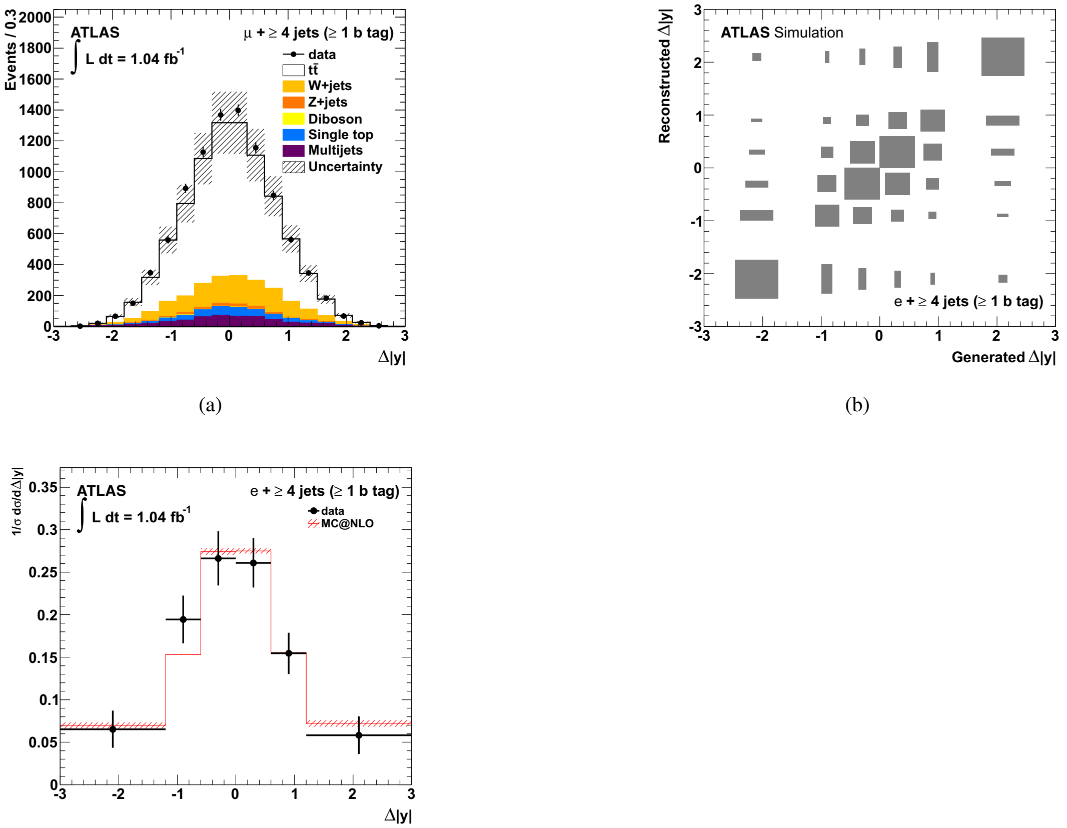
$$H = -\sum_{i=1}^{M} p_i \log p_i \tag{59}$$

where p_i is the probability for a given event/system to occur/be in a given subset i of the available phase space. The entropy H measures the amount of uncertainty represented by the probability distribution of a given variable and consequently determines the information content that any observation extracted from that population brings to the observer [9].

When new information about a variable is acquired the gain can be quantified by the change in uncertainty (information) between the initial estimate of the probability distribution for the variable and the new one. As the entropy H measures the information change, it is at the basis of the principle of minimum relative entropy (or cross-entropy) [34]: if there is not enough information to specify a probability distribution uniquely, a consistent estimator for it is obtained by minimizing

$$S(\boldsymbol{\mu}) = H(\boldsymbol{\mu}) = \sum_{i}^{M} \mu_i \log \frac{\mu_i}{\epsilon_i} \tag{60}$$

[9] An outcome from a distribution with a large Shannon entropy is more useful to the observer as it is less predictable than one with small entropy (which is actually fairly predictable): the observed outcome carries more information.



(a)

(b)

(c)

Figure 8. (a) Reconstructed distribution of the difference between the absolute rapidities of top quark and antitop quark (Δ |y|) in top quark pair events observed by the ATLAS detector in pp collisions at $\sqrt{s} = 7$ TeV at the LHC. The observed data are represented by the dots, the predicted amount of events and their breakdown in different sources are shown in the histograms in different colours and illustrated in the legend. (b) Migration matrix from simulated top quark pair events. (c) Unfolded differential cross section for the production of top quark pair events as a function of Δ |y| (dots) compared with the prediction from the standard model (red histogram). All the plots are taken from reference [6].

where $\boldsymbol{\mu}$ is the estimator vector for the unknown probability distribution, the index i goes from 1 to the number of M bins of the distribution and $\boldsymbol{\epsilon}$ is the reference probability distribution, representing the best knowledge about the true, unknown distribution. This method is used whenever the true distribution is known to be non-negative everywhere. When the only knowledge about the true distribution is its being non-negative and the reference distribution is taken to be a constant over all bins ($\epsilon_i = \epsilon_0\ \forall i$), the relative entropy of Equation 60 is reduced to the absolute entropy of Equation 59 up to a constant and the principle of minimum relative entropy is equivalent to the principle of maximum entropy [35]. The axiomatic derivation [34] for the minimum relative entropy estimator defines it as the distribution

μ_i that has the minimal distance from the reference, initial estimate ϵ_i in terms of information, but respects a given set of constraints.

Additional insight into the use of information entropy is provided in Ref. [36] where the minimum relative entropy estimate is interpreted as a maximum likelihood estimate. The negative logarithm of the likelihood for a given set of binned observation n_i to be compatible with a prior distribution ϵ_i and to satisfy the the response matrix constraints (see Eq. 24 is considered. This likelihood is shown to be proportional to the regularization function $S(\boldsymbol{\mu})$ in equation 60 up to a constant term (see Appendix A of [36]). The likelihood for a given set of binned observation n_i deriving from a true unknown distribution μ_i to be compatible with a prior distribution ϵ_i is represented by a multinomial distribution. The negative logarithm of this likelihood is shown to be proportional to the cross-entropy $S(\boldsymbol{\mu})$ in equation 60 up to a constant term (see Appendix A of [36]). The distribution μ_i is connected to the observed data by the response matrix and the likelihood for this requirement is generally represented by the generalized $\mathcal{L}(\boldsymbol{\mu})$ of Equation 24. In the end in the estimate of $\boldsymbol{\mu}$ minimizing the cross-entropy $S(\boldsymbol{\mu}, \boldsymbol{\epsilon})$ with the response matrix constraint corresponds to maximizing the distribution $\phi(\boldsymbol{\mu})$ in Equation 34, the negative logarithm of the full likelihood for the origin and detection of the observed events, in which the cross-entropy $S(\boldsymbol{\mu})$ is interpreted as the regularization function. In addition the interpretation of $S(\boldsymbol{\mu}, \boldsymbol{\epsilon})$ as a "prior" p.d.f for μ provides the justification in a Bayesian framework [37].

9.1 Automatic Regularized Unfolding

An implementation of the minimum relative entropy principle to provide a the regularization function is present in the Automatic Regularized Unfolding (ARU) [38]. This scheme is presently used to perform unfolding for one-dimensional problems. The algorithm does not require any parameter to be tuned, differently from the τ parameter for the Tikhonov scheme described in Section 7 or the number of iterations for the iterative techniques illustrated in Section 8.

ARU is a regularized fit. The unfolded distribution to be found, $b(x)$, is parametrized as the sum of flexible and smooth piece-wise, non negative polynomial curves with finite support, $b_j(x)$, called B-splines [39] i.e. $b(x) = \sum_j c_j b_j(x)$ where the range of the index j is determined by number of non-zero derivatives and of grid points that characterize the B-splines chosen for the approximation [38]. This solution form is folded with the detector *kernel* $K(y, x)$ quantifying the miscalibrations, efficiencies and resolution effects to produce the function $f(y)$

$$f(y) = \int K(y, x) b(x) dx = \sum_j c_j \int K(y, x) b_j(x) = \sum_j c_j f_j(y) \tag{61}$$

An extended maximum likelihood fit [40] of $f(y)$ to the data is then performed by minimizing

$$L(\mathbf{c}) = L_1(\mathbf{c}) + w L_2(\mathbf{c}) \tag{62}$$

In this formula $L_1(\mathbf{c})$ corresponds to the negative logarithm of the overall extended likelihood function $L_1(\mathbf{c})$ = -log $(\mathcal{L}_{stand}\mathcal{L}_{norm})$. The value of $\mathcal{L}_{stand}$ is

$$\mathcal{L}_{stand} = K \prod_i \tilde{f}(y_i|\mathbf{c}) \tag{63}$$

where K absorbs all the normalization constants, the set of y_i are the observed values for the variable y, $\tilde{f}(y_i) = f(y_I|\mathbf{c})/v(\mathbf{c})$ and $v(\mathbf{c}) = \int dy f(y) = \sum_j c_j F_j$ with $F_j = \int dy f_j(y)$. The likelihood $\mathcal{L}_{norm}$ allows to include the variation of the normalization

$$\mathcal{L}_{norm} = v(\mathbf{c})^N \frac{e^{-v(\mathbf{c})}}{N!} \tag{64}$$

So, by using Equations 63 and 64 and the associated definitions, $L_1(\mathbf{c})$ can be written as

$$\begin{aligned} L_1(\mathbf{c}) = & -\log\mathcal{L}_{stand} - \log\mathcal{L}_{norm} \\ & = -\log(K\prod_i \tilde{f}(y_i|\mathbf{c})) - \log(\upsilon(\mathbf{c})^N \frac{e^{-\upsilon(\mathbf{c})}}{N!}) \\ & = -\log K - \sum_i \log\tilde{f}(y_i|\mathbf{c}) - N\log\upsilon(\mathbf{c}) + \upsilon(\mathbf{c}) + \log N! \\ & = C - \sum_i \log\frac{f(y_i)}{\upsilon(\mathbf{c})} - N\log\upsilon(\mathbf{c}) + \upsilon(\mathbf{c}) \\ & = C - \sum_i \log f(y_i) + N\log\upsilon(\mathbf{c}) - N\log\upsilon(\mathbf{c}) + \upsilon(\mathbf{c}) \\ & = C - \sum_i \log f(y_i) + \sum_j c_j F_j \end{aligned} \tag{65}$$

where $C = -\log K + \log N!$ includes constants that can be neglected for the purpose of minimization. $L_2(c)$ is the regularization term based on the relative entropy principle

$$L_2(c) = \int b(s)\ln\frac{b(x)}{g(x)}dx - \sum_j c_j B_j \tag{66}$$

where the normalization $B_j = \int b_j(x)dx$ is included. The reference distribution $g(x)$ is chosen to be uniform while the weight w is determined by minimizing the mean integrated squared error ($MISE$) on $f(y)$ (that includes an estimate of the bias)

$$MISE(f(y)) = \int dyE[(f(y) - f_{true}(y))^2] = \int dyV[f(y)] + (f(y) - f_{true}(y))^2 \tag{67}$$

An example of the performance of the technique is shown in figure 9. Unfolding is performed on a sample of one thousand events drawn from a distribution of two Gaussian convoluted with a Gaussian *kernel*. The regularized result is compared with the unregularized solution. Figure 10 shows the distributions obtained when performing 2000 simulated experiments with 100 or 10000 events in each experiment, respectively. The uncertainty estimate from ARU is consistent with the observed standard deviation and the average bias has the same size of the statistical uncertainty. A simulation study using 1000 pseudo-experiments of 100 and 1000 events each shows the distribution for the mean and the standard deviation of the unfolded distributions with a bias that is comparable with the statistical uncertainty of the solution.

10 Non-iterative Bayesian-inspired regularization

A non-iterative regularization scheme also inspired by Bayes' theorem was recently proposed [41]. The rationale is to find the probability for the spectrum of a variable as a whole, given the probability for the observed data spectrum and the migration model, according to Bayes' formula:

$$p(\mathbf{T}|\mathbf{D} \wedge \mathcal{P}) \propto P(\mathbf{D}|\mathbf{T} \wedge \mathcal{P})\pi(\mathbf{T} \wedge \mathcal{P}) \tag{68}$$

where $\mathbf{T} = (T_1, T_{N_t})$ is the truth level binned spectrum (event density) in an N_t-dimensional space, $\mathbf{D} = (D_1, D_{N_r})$ is the observed binned spectrum in a generally different N_r-dimensional space. D_r

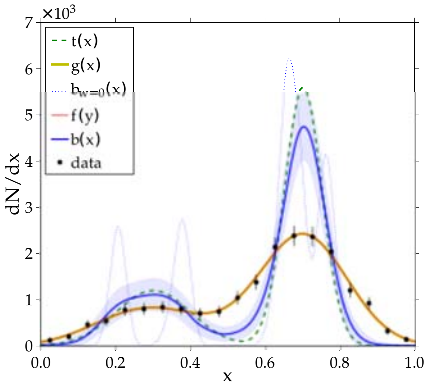


Figure 9. ARU-unfolded distribution of a one-dimensional variable x in a simulated experiment using a dataset of 1000 events [38]. The true distribution $t(x)$ is "smeared" into the histogram corresponding to the data points. In this case the folded distribution $f(y)$ used in equation 65 and the reference distribution $g(x)$ used for the regularization in Equation 66 are on top of each other. The regularized solution $b(x)$ is not showing undesired oscillations, differently from the unregularized solution $b_{w=0}(x)$.

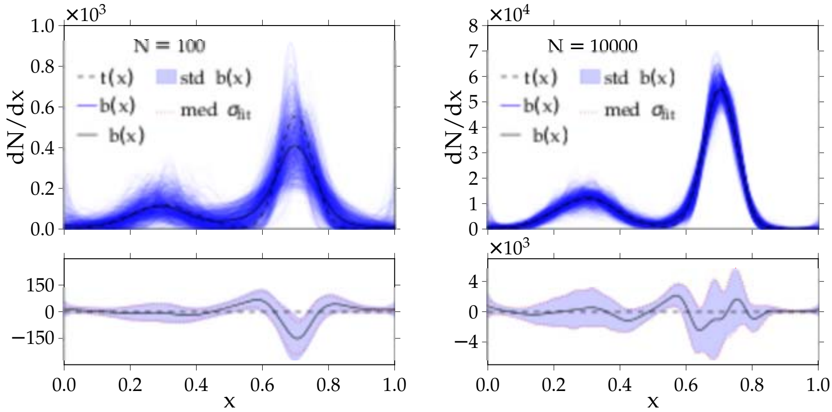


Figure 10. Superposed ARU-unfolded distributions $b(x)$ for a one-dimensional variable x resulting from unfolding 1000 pairs of simulated data sets randomly drawn according to the same true distribution with 100 events (left) and 1000 events (right) fir each pair respectively [38]. The upper plots also superpose the the true distribution $t(x)$ on top of the many unfolded solutions $b(x)$. The bottom plots show the bias of the unfolding defined as $b(x)$-$t(x)$, the standard deviation (std($b(x)$)) of the unfolded set of distribution and the median (med(σ_{fit}) of the estimated uncertainty on the solution.

is assumed to follow a Poisson distribution of mean R_r where $\mathbf{R} = (R_1, R_{N_r})$ is the N_r dimensional spectrum of the expected observations. The matrix $\mathcal{P}$ is the conditional migration matrix whose element $\mathcal{P}_{t,r}$ is the conditional probability $P(r|t)$ for an event produced in the truth level bin t to be reconstructed in the reconstructed bin r. So $P(r|t)$ is defined as

$$P(r|t) = \frac{P(t,r)}{P(t)} = \frac{\mathcal{M}_{t,r}}{\epsilon_t^{-1} \sum_{k=1}^{N_r} \mathcal{M}_{t,k}} \tag{69}$$

where $\mathcal{M}_{t,r} = P(t,r)$ is the joint probability for an event to be produced in the truth level bin t and in the reconstructed level bin r and ϵ_t is the efficiency for reconstructing events in the bins of row t defined as

$$\epsilon_t = \frac{\sum_{r=1}^{N_r} P(t,r)}{P(t)} = \frac{\sum_{r=1}^{N_r} \mathcal{M}_{tr}}{P(t)} \tag{70}$$

The interpretation of Equation 68 is that the resulting $p(\mathbf{T}|\mathbf{D} \wedge \mathcal{P})$ is the posterior probability density function (p.d.f.) for the truth level binned spectrum $\mathbf{T}$, $P(\mathbf{D}|\mathbf{T} \wedge \mathcal{P})$ is the likelihood of the observed binned spectrum $\mathbf{D}$ as a function of $\mathbf{T}$ and $\mathcal{P}$ and $\pi(\mathbf{T} \wedge \mathcal{P})$ is the prior p.d.f of $\mathbf{T}$ and $\mathcal{P}$.

If the spectrum is such that the data are counts of events, the Poisson distribution can be used

$$P(\mathbf{D}|\mathbf{T}) = \prod_{r=1}^{N_r} Poisson(D_r|\mathbf{T}) = \prod_{r=1}^{N_r} \frac{R_r^{D_r}}{D_r!} e^{-R_r} \tag{71}$$

where the mean expected number of events R_r in a bin r of $\mathbf{D}$ is

$$R_r = B_r + \sum_{t=1}^{N_t} T_t P(r|t) \tag{72}$$

and B_r is the expected number of background in bin r.

The result of the unfolding is the posterior probability distribution function $p(\mathbf{T}|\mathbf{D})$ for the whole spectrum. The form of the posterior distribution in Equation 68 is the same as the regularized likelihood that generates Equation 35 if $\mathcal{L}(\boldsymbol{\mu})$ is identified with $p(\mathbf{T}|\mathbf{D})$ and if a function $S(\mathbf{T})$ is defined such that the prior distribution is written as

$$\pi(\mathbf{T}) = e^{\alpha S(\mathbf{T})} \tag{73}$$

and $S(\mathbf{T})$ is identified with $S(\boldsymbol{\mu})$. Regularization is then interpreted as the inclusion of the prior distribution i.e. the a priori degree of belief in a specified property of the “true” spectrum $\mathbf{T}$. The posterior distribution $p(\mathbf{T}|\mathbf{D})$ is used to compute the marginal posterior distributions of the content in the bins of the spectrum $p_t(T_t|\mathbf{D})$ for $t \in [1, ...N_t]$. Such distributions are defined as

$$p_t(T_l|D) = \int .. \int p(\mathbf{T}|\mathbf{D}) dT_1..dT_{l-1} dT_{l+1}..dT_{N_t} \tag{74}$$

The estimator T_t for the content of bin t of the unfolded spectrum can then be derived by using more than one algorithm: from taking the mode or the mean of $p_t(T_l|D)$ to considering the half point of the 68% integration interval to using the mean of the Gaussian fitted to the $p_t(T_l|D)$. The uncertainty associated with the estimator is usually defined as the shorted interval in the range of T_t for which the integral of $p_t(T_l|\mathbf{D})$ amounts to 0.68. A crucial item for this technique is then the study of the convergence, stability and speed for the integration to be performed in Equation 74 [41]. Figure 11

shows an example of marginalized posterior probability distribution for the content of one bin of a given“true spectrum” resulting from a simulated experiment [41]. A steeply falling distribution of a given variable m is convoluted with an m-dependent Gaussian distribution aimed at simulating detector effects. The posterior distribution and some associated estimators are shown.

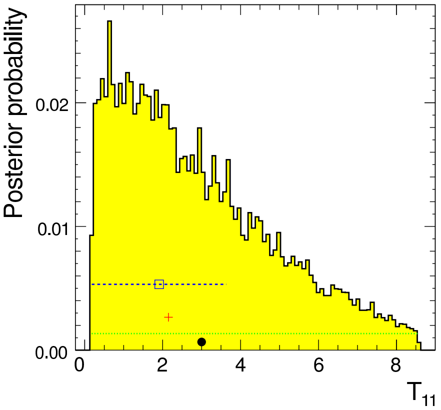


Figure 11. Example of one-dimensional marginal distribution for the content of one bin, $p_t(T_l|\mathbf{D})$, derived from a simulated model (see Fig 13 and section 6.4 of [41]) (yellow-filled distribution). In this specific case l=11 as it is the 11th bin (out of 14) of the overall distribution. The red cross marker represents the input truth bin content T_{11}.The black circle marker shows the simulated observed content D_{11}. The blue square marker represents the most likely value of T_{11} and the blue dashed line represents the shortest interval the integrates 68% of the marginalized distribution.The green dotted line shows the range in T_{11} that is used to sample the posterior $p(\mathbf{T}|\mathbf{D}$ to calculate the integral in equation 74.

11 Unfolding schemes: additional examples

Other unfolding schemes tend to build on the basic techniques outlined above. The following examples are by no means an exhaustive list: its goal is to show that the problem of unfolding can be approached in a variety of manners that involve the evolution of old ideas and their merging with new schemes. Unfolding solutions vary depending on the problem at hand and unfolding schemes are used in science beyond the realms of nuclear and particle physics . These lectures should be considered a starting point to enlarge one's knowledge about the problem and an invitation to explore the available techniques and find or develop the techniques that is more suitable for the problem the reader is interested in solving.

In the realm of iterative techniques the Iterative Dinamically Stabilized (IDS) method [42] regularizes the statistical fluctuations by iterative steps in which simultaneous corrections to the normalization of the distribution and its shape are derived. Each iteration involves three stages. Initially a correction to the normalization is provided by weighting the difference between the data and the simulation (Δd) with a monotonic regularization function $f(\Delta d)$ that depends on the statistical significance of the absolute Δd. The data-simulation difference for the content in each bin k, Δd_k, is weighted with $f(\Delta d_k)$ and the "true-to-reco" migration probability estimated from simulation. This difference is then used to correct the content of bin k in the simulation of the true distribution. The second step is an improved estimate of the background subtraction and the third step is an improved estimate of the response (migration) matrix. In both steps the bin content difference between the updated unfolded result and the true simulated distribution is weighted with $f(\Delta d_k)$ (and additionally with the simulation-derived"reco-to-true" migration probability in the third step) to provide corrections to improve background treatment and migration modelling. The goal is to derive unfolding corrections that tend to preserve real new structure present in the data (but not in the simulation) while reducing the statistical fluctuations and biases due to background subtraction. An example of application of the IDS technique in particle physics is the measurement of inclusive jet and dijet production in LHC

pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector [43] where the steeply falling distribution of the transverse momentum (p_T) for jets of colourless particles is extracted by correcting the p_T distributions for jets formed by the energy depositions in the ATLAS detector.

Iterative schemes are also provided in variants that do not require binning events into histograms. Binning free methods avoid the problem of bin size optimization. They allow to transform from one variable into another and apply selection criteria, after unfolding; small size samples can be used in arbitrarily high dimensions where this becomes a problem for techniques relying on histograms. Finally the unfolded samples automatically provide the estimate of the statistical uncertainty associated to the measurement.

A first example is the Binning Free Deconvolution proposed in Ref.[44]. For each iteration two stages are proposed. The first stage is to correct the weight of each single simulated event in the distribution of the variable of interest with the ratio of the experimental local density of events to the simulated one, both estimated by counting the events corresponding to an interval around the event value for the variable under consideration. The size of the interval of values is of the order of the experimental resolution. The second stage reassigns a weight to each event by averaging, for each event, the weights of the nearest neighbours in a given region. The size of the averaging region is variable and it governs the level of smoothing in the unfolding. Events are regrouped in bins after each iteration step. The iteration stops at the step after the minimum is reached for the regularized sum of squares defined as follows: a sum of squared differences between the bin content of the data and each intermediate corrected simulated distribution is added to a Tikhonov regularization term based on a second derivative estimate (like in equation 40).

The second example of binning free unfolding scheme is the Satellite Method [45], a technique aiming at obtaining the deconvoluted distribution associated to a certain data set in (generally) a multi-variable space. Given the observed data sample a simulated sample with exactly the same number of the events is generated, representing the unfolded, true “positions” in the multi-dimensional space. Each true “position” is associated to a set of “observed” positions representing the detector effects on the “true” position (de facto sampling the response *kernel*): the relative position of these “satellites” with respect to the associated “true” position is kept fixed. The number of “satellite” realizations is a tunable variable of the method. In each step of the iteration a test variable is computed to assess the quantitative compatibility of the experimental distribution and the expected density for simulated events resulting from the ”satellites”. The advised quantity is the statistical energy defined with the same formalism of the potential energy of two opposite charge distributions: the distributions are derived from both the experimental data and the expected simulated density and the $1/r$-law to be integrated is replaced by a general positive definite function (an exponential or a logarithm) of the distance between any two elements of the two samples. The value of the test variable is calculated at the beginning of the step and at the end, after having randomly chosen one of the “true” positions and after having moved it in the multi-variate space in a random direction, together with its satellites. The new expected density is recomputed and the test variable is recalculated: the migration is kept if the test variable has achieved a smaller value with respect to the beginning of the step, otherwise it is rejected. A new iteration starts if the test variable has not reached its minimum value. Regularization is implemented by either stopping the iteration before the results start to oscillate significantly or by varying the number of “satellite” points corresponding to a given “true” point: a larger number of satellite points implies a stronger regularization as the simulated information plays a more important role.

Finally ${}_s$*Plot* [46] is a statistical technique aimed at reconstructing the unknown true distribution of a variable (control variable) for a given data set by knowing the distribution of other variables associated to the various sources of events that compose the sample (discriminating variables). The distribution of the control variable might be known for some sources of the events, but not for all. A crucial assumption is that the control variable is uncorrelated with the discriminating variables. The first step is to perform an extended maximum likelihood fit of the probability distributions for the discriminating variables to the various sources of events in the data. This fit returns the yields of the sources and determines the parameters of of the distributions. The distribution for the control variable in a given data subset (called ${}_{s}Plot$ for the given data subset) is then obtained by weighting all the events in the subset with a function of the distribution of the discriminating variables and their estimated covariance matrix (called sWeight). The sWeight is derived by solving a matrix equation in which the average of the control variable ${}_{s}Plot$ for the given data subset is expressed as a linear combination of the control variable special distributions (${}_{in}Plot$s) for all the subsets of the data and the coefficients are provided by the covariance matrix of the discriminating variables. The ${}_{in}Plot$ for a given data subset is the control variable distribution obtained by expressing the control variable as a function of the discriminating variables. The ${}_{in}Plot$ for a given data subset is obtained by weighting all the data events in a given control variable bin with the probability that each event belongs to the data subset of interest: the value of the discriminating variable for each event is fed to the known discriminating variable distribution to obtain such probability.

12 Unfolding software tools

Some of the most recent unfolding schemes used in particle physics are now available as updated and maintained public software tools. These make it much simpler to use the proposed techniques and to provide feed-back that can be included in new versions, with general benefit to the users' community.

A sizeable collection of the available public software (and documentation) to perform unfolding in particle physics is reported under the Unfolding Framework Project [50].

Amongst the available tools, the ROOT Unfolding framework (RooUnfold) [51] usefully collects five unfolding schemes, including a basic implementation of unregularized ML unfolding through matrix inversion (see Section 3), the correction factor scheme outlined in Section 4, the regularized Tikhonov approach using the second order derivative described in Section 7 and a version of the iterative Bayesian-inspired scheme illustrated in Section 8.1. RooUnfold is a coherent C++ framework that can be linked against the ROOT libraries [52] that are widely used in particle physics analysis.

The most updated version of the iterative scheme of Section 8.1 is available at [53].

A precursor for the implementation of the Tikhonov scheme outlined in Section 7, coupling the second derivative regularization with a B-spline-based parametrization, is available in the FORTRAN-based program RUN [55]. The actual FORTRAN implementation of the Tikhonov scheme of Section 7 is available in the program GURU [56] and a C++ wrapper is also available [57] outside of ROOT.

The ARU scheme described in Section 9.1 is available in C++ format with a Python-based interface [58].

13 Optimization and choice of unfolding technique: the last two cents

Given an unfolding problem, the optimization of the balance between the bias and the total expected uncertainty, including full statistical and systematic effects, is a powerful criterion against which to scan the available degrees of freedom: unfolding scheme, binning of data, regularization parameter. Such optimization is usefully developed on large samples of simulated data so as to avoid biases induced by the specific features of the data under consideration. Quantitative figures of merit can be derived from the available information to drive the optimization [13, 47]. In the very common case of binned data, a basic figure of merit related to the total uncertainty is the mean squared averaged uncertainty over the bins of the distribution, defined as

$$MSE = \frac{1}{M}\sum_{i=1}^{M}(U_{i,i} + b_i^2) \tag{75}$$

where $U_{i,i}$ is the diagonal element of the covariance matrix of the unfolded estimates for the ith bin content and b_i is the estimated bias on each bin content as defined at the end of in section 5 and M is the number of bins in the distribution. In general different bins have different uncertainties. A modified MSE can take this into account by weighting the contribution of each bin with the inverse of the bin variance and by remembering that the mean and the variance of the Poisson-distributed content of bin i ,μ_i, are equal:

$$MSE' = \frac{1}{M}\sum_{i=1}^{M}\frac{U_{ii} + b_i^2}{\mu_i} \tag{76}$$

Another possible driving criterion is to require that the estimated bias squared, b_i^2, be smaller than its own variance $V_{ii} = cov[b_i, b_i]$ so that there is no statistically significant bias. In this way the prescription can be :

$$\sum_{i=1}^{M} b_i^2/V_{ii} = M \tag{77}$$

If the bias were statistically significant a correction for it could be considered and the uncertainty on the bias would then have to be included as systematic uncertainty on the final result.

The inclusion of systematic uncertainties in the measurement needs to take into account the correlations that the unfolding introduces on a bin-by-bin basis and to derive their combined impact on the measurement. As an example, a very powerful way to achieve this is to use pseudo-experiments [10] or toy experiments [11]: each experimental variable derived in a given pseudo- or toy experiment is the result of the sum of the effects of modifying a set of (assumed) independent modelling parameters [12]. The distributions of the modelling parameters are derived from some ancillary measurement or from an explicitly mentioned assumed distribution (usually a Gaussian or a Log-Normal distribution) and the effects are obtained by sampling such independent distributions. By calculating the observable(s) of interest in each experiment, its (their) distribution(s) can be calculated and, for instance, an estimator(estimators) for the mean(s) and the variance(s) can be obtained from the calculated distribution(s). The correlation between the various bin content can also be estimated by taking averages and variances over the pseudo- or toy experiments. This is a essentially Bayesan-inspired way of marginalizing the modelling parameters in the likelihood through Monte Carlo integration [59, 60]. An example of the formalism for the inclusion of the modelling parameters can be found in Ref [61] (see section 2 and Appendix A). The unfolding ingredients where the variations of the parameters need to be implemented are easily read off from the general likelihood solution and its elements shown in Equations 11, 13 and 36: the response matrix, the estimate of non interesting events (backgrounds), the acceptance corrections all need to be varied when input to the unfolding procedure according to their

[10] A set of simulated events including a break down of “signal” and “background’ events” with a mixture and a total size that is aimed at reproducing the available dataset. A large number of statistically independent mixtures are realized so as to simulate many instances of the experiment.

[11] A set of simulated distributions derived from assuming a probability distribution and its parameter and generating sampled distributions corresponding to the p.d.f model.

[12] For instance the a modification in the jet energy scale as a function of jet transverse momentum will cause a variation in the reconstructed mass of the top-antitop system decaying to a lepton, missing transverse energy and jets.

relation with the independent modelling parameters. In this way the all the correlations introduced by the unfolding are automatically kept into account.

In relation to the optimization, the variables that provide crucial insight in the origin of the unfolding problem represent sensitive tools to understand the role of the balance between statistical and systematic uncertainties. An important example is the condition number $c(R)$ defined in Equation 23: this quantifies the general sensitivity of the analysis to incoming fluctuations, independently of their origin. When exploring the unregularized maximum likelihood solution, the condition number will change depending on the binning of the events and on the inclusion of certain systematic uncertainties. In the regularized realm an important example is the choice of τ which is based on the distribution of the input vector w_i defined in Equation 44 Such distribution is driven by the combination of information of the response matrix and the statistical covariance matrix of the inputs linked by the specific curvature-based regularization (as illustrated in Section 7): in general systematic uncertainties are not included in the input covariance matrix, so the optimal values for the τ parameter chosen according to the criteria mentioned in Section 7 need to be extended to reflect those effects, by considering the expected variations of w_i due to the different systematic effects.

The choice for the algorithm and the optimization of its parameters is definitely dependent on the data analysis that is being carried out. Whatever the technique chosen for unfolding, it is useful to produce (and possibly report) the unregularized maximum likelihood solution using the matrix inversion solution of Equation 11: the unfolding technique does not add any bias to the result (see section 3) and it is equivalent to using the uncorrected data to test a theory or estimating its parameters, for instance by minimizing the sum of squares [47]

$$\chi^2(\boldsymbol{\theta}) = (\boldsymbol{\mu}(\boldsymbol{\theta}) - \boldsymbol{\mu}_{ML})^T U^{-1}(\boldsymbol{\mu}(\boldsymbol{\theta}) - \boldsymbol{\mu}_{ML}) \tag{78}$$

where U^{-1} is the covariance matrix on the unfolded measurement and $\boldsymbol{\theta}$ is the vector of parameters to be estimated by the minimization. The important point is that the full covariance matrix for the unfolded result needs to be used.

Finally it is crucial to devise a test for stability and bias to check whether the unfolding scheme is robust with respect to possible fluctuations and what level of bias can be recovered by the unfolding. This is particularly important to build confidence that the unfolding is not diluting or cancelling important unexpected features of the data that are not foreseen by the models underlying the unfolding scheme. Typical tests are performed by unfolding simulated events whose true shapes are distorted with respect to the best knowledge encoded in the theoretical models included in the simulations. The induced distortions should be comparable with the total statistical and systematic uncertainty of the measurement. The distorted simulated test distributions are then unfolded and the deviation with respect to the input can be quantified, for instance, by using a least squares test between the input model and the unfolded result and exploiting the full covariance matrix including both statistical and systematic uncertainties. This provides a quantitative assessment of the capacity of the unfolding scheme to recover the input distortion and its ability to maintain “unforeseen” features present in the data.

Extremely useful insight into the unfolding problem is given in the lectures of Ref. [47] and Ref. [11], the chapter on introduction to statistics in Ref. [48] and the slides and proceedings of the PHYSTAT 2011 conference dedicated to unfolding [49].

14 Acknowledgements

The author would like to thank the organizers of SOS2012 for the invitation to contribute to the school and for the organization of a very enjoyable and enriching event. It is a pleasure to thank Glen Cowan for various enlightening discussions about the crucial aspects of unfolding and Jorgen

Sjölin for many eye-opening comments about on the inclusion of systematic uncertainties in both regularized and unregularized unfolding applications in particle physics analysis. Kyle Cranmer and Alex Read deserve thanks for a clarifying discussion on inclusion of systematic uncertainties through Monte Carlo integration.

References

[1] Y. Vardi, L. A. Shepp, L. Kaufman, *"A Statistical Model for Positron Emission Tomography"*, Journal of the American Statistical Association, Vol. 80, No. 389 (Mar., 1985), pp. 8-20 [http://www.jstor.org/stable/2288030]

[2] O. Helene, V. R. Vanin, Z. O. Guimaraes-Filho, C. Takiya, *"Variances, covariances and artifacts in image deconvolution"*, Nucl. Instr. Meth. **A**, 580 (2007) pp.1466-1473

[3] L. Evans, P. Bryant (editors), *"LHC Machine"*, 2008 JINST **3**, S08001, doi:10.1088/1748-0221/3/08/S08001 [http://iopscience.iop.org/1748-0221/3/08/S08001]

[4] the ATLAS Collaboration, *"The ATLAS Experiment at the CERN Large Hadron Collider"*, JINST **3**, 2008, S08003, doi:10.1088/1748-0221/3/08/S08003 [http://iopscience.iop.org/1748-0221/3/08/S08003]

[5] V. Ahrens, A. Ferroglia, M. Neubert, B. D. Pecjak and L. L. Yang, "Renormalization-Group Improved Predictions for Top-Quark Pair Production at Hadron Colliders", JHEP **1009**, 097 (2010) [arXiv:1003.5827 [hep-ph]].

[6] the ATLAS collaboration, *"Measurement of the charge asymmetry in top quark pair production in pp collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector"*, Eur. Phys. J. C **72**, 2039 (2012) [arXiv:1203.4211 [hep-ex]].

[7] Computer generated image of the whole ATLAS detector, CERN-GE-0803012, Photograph: Joao Pequenao, [http://cds.cern.ch/record/1095924/]

[8] Figures produced by the D0 collaboration [9] available at http://www-d0.fnal.gov/Run2Physics/top
/top_public_web_pages/top_feynman_diagrams.html

[9] The D0 experiment, http://www-d0.fnal.gov

[10] P C Hansen, *"Numerical tools for analysis and solution of Fredholm integral equations of the first kind"*, Inverse Problems **8** (1992) 849 doi:10.1088/0266-5611/8/6/005 [http://m.iopscience.iop.org/0266-5611/8/6/005?rel=sem&relno=7]

[11] Volker Blobel, *"Unfolding methods in high energy physics experiments"*, Report DESY 84-118, 1984 (also in Proceedings of the 1984 CERN School of Computing, CERN 85-09, pp. 88-127; see also http://www.desy.de/ blobel/).

[12] J. Beringer *et al.* (Particle Data Group), *"The Review of Particle Physics"*, Phys. Rev. **D**86, 010001 (2012) [http://pdg.lbl.gov/]

[13] G. Cowan, "A survey of unfolding methods for particle physics", Conf. Proc. C **0203181**, 248 (2002).

[14] *"Statistics"* (Revised by G. Cowan) in [12]

[15] L. Lyons, *"Unfolding: Introduction"*, in Proceedings of the PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN, Geneva Switzerland, 17-20 January 2011, edited by H.B.Prosper, L.Lyons, CERN-2011-006, pp. 225-228 [http://cds.cern.ch/record/1306523] and references to unfolding therein.

[16] I. P. Nedelkov, *"Improper problems in Computation Physics"*, Com. Phys. Comm. **4** (1972) 157

[17] S. Leach, *" Singular Value Decomposition. A Primer"*, [http://people.csail.mit.edu/hasinoff/320/SingularValueDecomposition.pdf], material from CSC320S: Introduction to Visual Computing course at MIT and references therein.

[18] A. G. Frodesen, O. Skjeggestad, H. Tofte, *"Probability and Statistics in particle physics"*, Hardcover: 501 pages, Publisher: Universitetsforlaget (September 1979), ISBN-10:8200019063, ISBN-13: 978-8200019060

[19] A. Hoecker, V. Kartvelishvili, *"SVD Approach to data unfolding"*, Nucl. Instr. Meth. **A 372**, 1996 (469)

[20] A. Björck, *"Least squares methods"*, Handbook of Numerical Analysis, voI I. (1990) 465-652, ed P. G. Ciarlet and J. L. Lions (Amsterdam: Elsevier)

[21] V. Blobel, *"Unfolding methods in Particle Physics"*, in Proceedings of the PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN, Geneva Switzerland, 17-20 January 2011, edited by H.B.Prosper, L.Lyons, CERN-2011-006, pp. 240-251 [http://cds.cern.ch/record/1306523] and references to unfolding therein.

[22] See for instance H. M. Antia, *"Numerical methods for scientists and Engineers"*, Birkhäuser, 2nd edition, (2002)

[23] A. N. Tikhonov ,V. Y. Arsenin *"Solutions of ill-Posed Problem"* Wiley, New York, (1977)

[24] See for instance T. M. Apostol (June 1967), *"Calculus, Vol. 1: One-Variable Calculus with an Introduction to Linear Algebra 1"* (2nd ed.), Wiley, ISBN 978-0-471-00005-1

[25] C. E. Lawson and R. J. Hanson,*"Solving Least Square Problems"*, Prentice-Hall Inc., Englewood Cliffs, 1974.

[26] G. Zech *"Regularization and error assignment to unfolded distributions"*, in Proceedings of the PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN, Geneva Switzerland, 17-20 January 2011, edited by H.B.Prosper, L.Lyons, CERN-2011-006, pp. 252-259 [http://cds.cern.ch/record/1306523] and references to unfolding therein.

[27] H. N. M̀ulthei, B. Schorr, |em " On an iterative method for the unfolding of spectra", Nucl. Instr. and Meth. A257 (1987) 371-377

[28] L. B. Lucy *"An iterative technique for the rectification of observed distributions"*, Astronomical Journal 79(6) (1974) 745

[29] See section 36.1.4. in Ref [14] and references therein.

[30] See the articles in reference 1 of [26], particularly [1] and [27].

[31] G. D'Agostini, *"A multidimensional unfolding method based on Bayes' theorem"* , Nucl. Instr. Meth. **A 362** 1995 (487)
G. D'Agostini, *" Improved Iterative Bayesian unfolding"*, http://arxiv.org/abs/1010.0632

[32] J. H. Kuhn and G. Rodrigo, *"Charge asymmetries of top quarks at hadron colliders revisited"*, JHEP **1201**, 063 (2012) [arXiv:1109.6830 [hep-ph]].

[33] C. Shannon *"A mathematical Theory of Communication"* Bell System Technical Journal **27** (3) 379-423

[34] J. E. Shore , *"Relative Entropy, Probabilistic Inference and AI"* , contribution to Proceedings of the First Conference Annual Conference on Uncertainty in Artificial Intelligence (UAI-85), Corvallis, Oregon, 1985, pp 43-47, AUAI Press [http://uai.sis.pitt.edu/papers/85/p43-shore.pdf]

[35] E. T. Jaynes, *"Information Theory and Statistical Mechanics"*, Phys. Rev. 106 (1957) 620

[36] M. Schmelling, *"The method of reduced cross-entropy. A general approach to unfold probability distributions"*, Nucl. Instr. Meth. **A 340** (1994) 400-412

[37] J. Skilling, *"Quantified Maximum Entropy"*, in Maximum Entropy and Bayesian Methods, Fundamental Theories of Physics, vol. 39, 1990, pp 341-350 and ed. P.F. Fougère (Kluwer, Dordrecht, Holland, 1990).

[38] H. P. Dembinski, M. Roth, *"ARU - towards automatic unfolding of detector effects"* in Proceedings of the PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN, Geneva Switzerland, 17-20 January 2011, edited by H.B.Prosper, L.Lyons, CERN-2011-006, pp. 285-291 [http://cds.cern.ch/record/1306523] and [http://aru.hepforge.org]

[39] C. de Boor, *"A Practical Guide to Splines"*, Springer Verlag (New York, Heidelberg, Berlin) (1978).

[40] R. Barlow, *"Extended maximum Likelihood"*, Nucl. Instrum. Meth. **A297**, 496 (1990) and references therein.

[41] G. Choudalakis, *"Fully Bayesian Unfolding"*, [arXiv:1201.4612[physics.data-an]]

[42] B. Malaescu, *"An iterative, dynamically stabilized method of data unfolding"*, [arXiv:0907.3791 [phy-
sics.data-an]]

[43] G. Aad *et al.* the ATLAS Collaboration,*"Measurement of inclusive jet and dijet production in pp collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector"*, Phys. Rev. D **86**, 014022 (2012) [arXiv:1112.6297 [hep-ex]].

[44] L. Lindemann, G. Zech, *"Unfolding by Weighting Monte Carlo Events"* , Nucl. Instr. Meth **A** 354 (1995) 516-521

[45] B. Aslan and G. Zech, *"Statistical energy as a tool for binning-free, multivariate goodness- of-fit tests, two-sample comparison and unfolding"*, Nucl. Instr. and Meth. **A** 537 (2005) 626

[46] M. Pivk, F. R. Le Diberder, *" $_s$Plot a statistical tool to unfold data distributions"* , Nucl. Inst. Meth. **A** 555:356-369, (2005)

[47] G. Cowan, *"Statistics for HEP. Lecture 4:Unfolding"*, CERN Academic Training Lectures, CERN, Geneva, Switzerland, 5th April 2012, [http://indico.cern.ch/conferenceDisplay.py?confId=173729]

[48] G. Bohm and G. Zech, *"Introduction to Statistics and Data Analysis for Physicists"*, Verlag Deutsches Elektronen-Synchrotron (2010) [http://www-library.desy.de/elbook.html]

[49] Proceedings of the PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN, Geneva Switzerland, 17-20 January 2011, edited by H.B.Prosper, L.Lyons, CERN-2011-006 [http://cds.cern.ch/record/1306523] and [http://indico.cern.ch/conferenceOtherViews.py?view=standard&confId=107747]

[50] The Unfolding Framework Project, [https://www.wiki.terascale.de/index.php/Unfolding_Framework_Project] (accessed on 9th May 2013) with software and references therein.

[51] T. Adye, *"Unfolding algorithms and tests using RooUnfold"* in Proceedings of the PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN, Geneva Switzerland, 17-20 January 2011, edited by H.B.Prosper, L.Lyons, CERN-2011-006, pp. 313-318 [http://cds.cern.ch/record/1306523] and [http://hepunx.rl.ac.uk/ adye/software/unfold/RooUnfold.html]

[52] R. Brun and F. Rademakers, *"ROOT - An Object Oriented Data Analysis Framework"*, Proceedings AIHENP'96 Workshop, Lausanne, Sep. 1996, Nucl. Inst. & Meth. in Phys. Res. A 389 (1997) 81-86. See also http://root.cern.ch/.

[53] G. D'Agostini, *"Probabillity and Statistics - Improved iterative Bayesian unfolding"* [http://www.

roma1.infn.it/~dagos/unf2_R.tgz] written using the R Framework [54]

[54] R Development Core Team (2009), x*"R: A language and environment for statistical computing"*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, [http://www.r-project.
org].

[55] V. Blobel, *"Unfolding"*, RUN program source files and Manual, [http://www.desy.de/~blobel/unfold.html]

[56] V. Kartvelishvili, *"GURU"*, [http://www.hep.lancs.ac.uk/guru.tar.gz]

[57] G. Hesketh, *"Unfolding"*, [http://www-d0.fnal.gov/ ghesketh/unfolding/]

[58] H. P. Dembinski, M. Roth, *"ARU development page"*, [http://aru.hepforge.org]

[59] M. H. Kalos, P. A. Whitlock, *"Monte Carlo Methods, Volume 1"*, Wiley-VCH Publisher (John Wiley & Sons, Inc.), 2nd Edition, (2008)

[60] R. D. Cousins, V. L. Highland, *"Incorporating systematic uncertainties into an upper limit"* , Nucl. Instr. Meth. **A.320** (1992) 331-335

[61] the ATLAS Collaboration, *"Procedure for the LHC Higgs boson search combination in summer 2011"*, ATL-PHYS-PUB-2011-011 [https://cds.cern.ch/record/1375842] and references therein.