Available online at www.sciencedirect.com
ScienceDirect
Nuclear Instruments and Methods in Physics Research A 583 (2007) 494–506
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

Unfolding of differential energy spectra in the MAGIC experiment

J. Albert[a], E. Aliu[b], H. Anderhub[c], P. Antoranz[d], A. Armada[b], M. Asensio[d], C. Baixeras[e], J.A. Barrio[d], H. Bartko[f], D. Bastieri[g], J. Becker[h], W. Bednarek[i], K. Berger[a], C. Bigongiari[g], A. Biland[c], R.K. Bock[f,g], P. Bordas[j], V. Bosch-Ramon[j], T. Bretz[a], I. Britvitch[c], M. Camara[d], E. Carmona[f], A. Chilingarian[k], S. Ciprini[l], J.A. Coarasa[f], S. Commichau[c], J.L. Contreras[d], J. Cortina[b], M.T. Costado[m,v], V. Curtef[h], V. Danielyan[k], F. Dazzi[g], A. De Angelis[n], C. Delgado[m], R. de los Reyes[d], B. De Lotto[n], E. Domingo-Santamaría[b], D. Dorner[a], M. Doro[g], M. Errando[b], M. Fagiolini[o], D. Ferenc[p], E. Fernández[b], R. Firpo[b], J. Flix[b], M.V. Fonseca[d], L. Font[e], M. Fuchs[f], N. Galante[f], R.J. García-López[m,v], M. Garczarczyk[f], M. Gaug[m], M. Giller[i], F. Goebel[f], D. Hakobyan[k], M. Hayashida[f], T. Hengstebeck[q], A. Herrero[m,v], D. Höhne[a], J. Hose[f], C.C. Hsu[f], P. Jacon[i], T. Jogler[f], R. Kosyra[f], D. Kranich[c], R. Kritzer[a], A. Laille[p], E. Lindfors[l], S. Lombardi[g], F. Longo[n], J. López[b], M. López[d], E. Lorenz[c,f], P. Majumdar[f], G. Maneva[r], K. Mannheim[a], O. Mansutti[n], M. Mariotti[g], M. Martínez[b], D. Mazin[b], C. Merck[f], M. Meucci[o], M. Meyer[a], J.M. Miranda[d], R. Mirzoyan[f], S. Mizobuchi[f], A. Moralejo[b], D. Nieto[d], K. Nilsson[l], J. Ninkovic[f], E. Oña-Wilhelmi[b], N. Otte[f,q], I. Oya[d], M. Panniello[m,✠], R. Paoletti[o], J.M. Paredes[j], M. Pasanen[l], D. Pascoli[g], F. Pauss[c], R. Pegna[o], M. Persic[n,s], L. Peruzzo[g], A. Piccioli[o], N. Puchades[b], E. Prandini[g], A. Raymers[k], W. Rhode[h], M. Ribó[j], J. Rico[b], M. Rissi[c], A. Robert[e], S. Rügamer[a], A. Saggion[g], T. Saito[f], A. Sánchez[e], P. Sartori[g], V. Scalzotto[g], V. Scapin[n], R. Schmitt[a], T. Schweizer[f], M. Shayduk[q,f], K. Shinozaki[f], S.N. Shore[t], N. Sidro[b], A. Sillanpää[l], D. Sobczynska[i], A. Stamerra[o], L.S. Stark[c], L. Takalo[l], P. Temnikov[r], D. Tescaro[b], M. Teshima[f], D.F. Torres[u], N. Turini[o], H. Vankov[r], V. Vitale[n], R.M. Wagner[f], T. Wibig[i], W. Wittek[f,*], F. Zandanel[g], R. Zanin[b], J. Zapatero[e]

[a] *Universität Würzburg, D-97074 Würzburg, Germany*
[b] *Institut de Física d'Altes Energies, Edifici Cn., E-08193 Bellaterra (Barcelona), Spain*
[c] *ETH Zurich, CH-8093, Switzerland*
[d] *Universidad Complutense, E-28040 Madrid, Spain*
[e] *Universitat Autònoma de Barcelona, E-08193 Bellaterra, Spain*
[f] *Max-Planck-Institut für Physik, D-80805 München, Germany*
[g] *Università di Padova and INFN, I-35131 Padova, Italy*
[h] *Universität Dortmund, D-44227 Dortmund, Germany*
[i] *University of Łódź, PL-90236 Lodz, Poland*
[j] *Universitat de Barcelona, E-08028 Barcelona, Spain*
[k] *Yerevan Physics Institute, AM-375036 Yerevan, Armenia*
[l] *Tuorla Observatory, FI-21500 Piikkiö, Finland*
[m] *Inst. de Astrofisica de Canarias, E-38200, La Laguna, Tenerife, Spain*
[n] *Università di Udine, and INFN Trieste, I-33100 Udine, Italy*
[o] *Università di Siena, and INFN Pisa, I-53100 Siena, Italy*
[p] *University of California, Davis, CA 95616-8677, USA*

*Corresponding author. Tel.: +49 89 32 354 291; fax: +49 89 32 26 704.
E-mail address: wittek@mppmu.mpg.de (W. Wittek).
✠Deceased.

doi:10.1016/j.nima.2007.09.048

[q]Humboldt-Universität zu Berlin, D-12489 Berlin, Germany
[r]Institute for Nuclear Research and Nuclear Energy, BG-1784 Sofia, Bulgaria
[s]INAF/Osservatorio Astronomico and INFN Trieste, I-34131 Trieste, Italy
[t]Università di Pisa, and INFN Pisa, I-56126 Pisa, Italy
[u]ICREA & Institut de Ciències de l'Espai (CSIC-IEEC), E-08193 Bellaterra, Spain
[v]Depto. de Astrofisica, Universidad, E-38206, La Laguna, Tenerife, Spain

Received 12 July 2007; received in revised form 27 September 2007; accepted 27 September 2007
Available online 10 October 2007

Abstract

The paper describes the different methods, used in the MAGIC experiment, to unfold experimental energy distributions of cosmic ray particles (γ-rays). Questions and problems related to the unfolding are discussed. Various procedures are proposed which can help to make the unfolding robust and reliable. The different methods and procedures are implemented in the MAGIC software and are used in most of the analyses.

PACS: 07.05.Kf

Keywords: Statistics techniques; Regularization; Unfolding

1. Introduction

In an Imaging-Air-Cherenkov-Telescope (IACT) experiment like MAGIC [1] the energy E of the cosmic ray particle (γ-ray) is not exactly known. It has to be estimated, the energy resolution being in the order of 20–40%. As a consequence, the experimentally measured energy spectrum is biased. The procedure to correct for the effects due to the finite energy resolution is called unfolding.

While in high-energy-physics experiments unfolding is a widely used technique, this is not the case in present day's IACT experiments. This paper deals with the unfolding procedure, which is applied as a standard tool in the MAGIC experiment. The different unfolding methods are explained in detail. Emphasis is put on the discussion of questions and problems related to the application of the unfolding to real data. It is not the aim of the paper to give a complete derivation of all formulas. For this the reader is referred to the publications [2–20]. An excellent review of unfolding methods is given in Ref. [12]. The present paper makes use of many ideas discussed in that paper.

Although here only differential energy spectra are considered, the procedures are equally well applicable to distributions of other quantities, including distributions in more than one dimension [21].

The layout of the paper is as follows. In Section 2 the notation is defined and the motivation for the unfolding procedure is given. The different unfolding methods, which means the different ways of regularisation, are presented in Section 3. The so-called Forward Unfolding, which represents an implicit unfolding under the assumption of a certain parametrisation of the unfolding result S, is explained in Section 4. In Section 5 two quantities are introduced, which are useful for an optimal choice of the regularisation strength. The criteria for this choice are collected in Section 6. Section 7 discusses various technical aspects which are important in the application of the unfolding procedure to real data. Two particular technical procedures, which ensure an unbiased and robust unfolding, are presented in Sections 8 and 9. Some unfolding results, obtained by applying the unfolding procedure to data taken in the MAGIC experiment, are discussed in Section 10. Finally, the pros and contras of the method of Correction Factors, which is an alternative way of correcting data for effects due to the finite experimental resolution, are listed in Section 11. A summary is given in Section 12.

2. The aim of the unfolding procedure

In this section the notations are defined and the motivation for the unfolding procedure is explained.

2.1. Notation

The true and measured (estimated) values of the energy of the cosmic ray particle are denoted by E_{true} and E_{est}, respectively. The data are assumed to be binned in histograms, and certain binnings are chosen independently for the distributions in E_{true} and E_{est}. Furthermore, the following definitions are introduced:

Y_i	number of events in bin i of E_{est} $(i = 1, \ldots, n_a)$
$K_{i,k}$	covariance matrix K of Y $(i = 1, \ldots, n_a;\ k = 1, \ldots, n_a)$
S_j	number of events in bin j of E_{true} $(j = 1, \ldots, n_b)$
$T_{j,l}$	covariance matrix T of S $(j = 1, \ldots, n_b;\ l = 1, \ldots, n_b)$
M_{ij}	migration matrix M $(i = 1, \ldots, n_a;\ j = 1, \ldots, n_b)$
G_{ik}	Gram's matrix $G = M \cdot M^{\mathrm{T}}$ $(i = 1, \ldots, n_a;\ k = 1, \ldots, n_a)$

with

$$\sum_{i=1}^{n_a} M_{ij} = 1 \quad \text{for all } j. \tag{1}$$

In the following it is assumed that the rank n_r of G is equal to (and not less than) the minimum of n_a and n_b, where n_a and n_b are the number of bins in E_{est} and E_{true}, respectively, which are used in the unfolding. This can always be achieved by a proper choice of the binnings in E_{true} and E_{est}.

The element M_{ij} of the migration matrix represents the fraction of events in bin j of E_{true}, which moves into bin i of E_{est}, due to the finite experimental energy resolution:

$$Y_i = \sum_{j=1}^{n_b} M_{ij} \cdot S_j \quad (i = 1, \ldots, n_a)$$

or in matrix notation

$$Y = M \cdot S. \tag{2}$$

In Eq. (2) Y, S and M are understood as quantities for a finite event sample, including the individual fluctuations. In this paper, Y and S are also called measured and true distribution, respectively. It is the aim of the unfolding procedure to determine an estimate of S, given the measurement Y and an estimate of M.

The estimate of the migration matrix M is obtained from Monte Carlo (MC) simulations, in which the development of the air shower (induced by the cosmic ray particle in the atmosphere), the emission of Cherenkov light in the air shower, the geometrical, optical and electronic properties of the telescope and the experimental procedures in the data analysis (shower reconstruction, γ/hadron separation, energy estimation, selections, cuts) are simulated [22]. M is computed from a 2-dimensional plot of the number of reconstructed MC events in the E_{est}–E_{true} plane, which was produced under the same conditions (selections, cuts) as the experimental distribution Y.

It should be noted that M only describes the migration of events. It does not describe losses of events, which will occur due to the finite acceptance of the detector, due to the trigger conditions and due to additional selections and cuts. In an IACT experiment these losses are also determined by MC simulations, and the corresponding correction factor is the effective collection area $A_{eff}(E_{true})$. $A_{eff}(E_{true})$ has to be computed again under the same conditions as the experimental distribution Y. Apart from minor effects (see Section 8), this correction can be performed quite independently of the unfolding.

The unfolding can be understood as a reshuffling of events from the bins of E_{est} into the bins of E_{true}. In this procedure the numerical values of the bin edges, both for E_{true} and E_{est}, are completely irrelevant. E_{true} and E_{est} may be even two different physical quantities, with completely different ranges of values and different units, like E_{true} = (the true energy of the cosmic ray particle) and E_{est} = (the total number of Cherenkov photons measured for the shower) [23]. Of course, unfolding makes only sense if E_{true} and E_{est} are sufficiently strongly correlated, otherwise the distribution of E_{true} cannot be inferred from a distribution of E_{est}. This is in contrast to the method of Correction Factors (see Section 11), where E_{est} has to be a good estimate of E_{true} in any case.

Because the binnings in E_{true} and E_{est} can be chosen independently and to a certain degree arbitrarily (see Section 7), the migration matrix M is in general not square. For this reason Eq. (2) can in general not be inverted to obtain S as $M^{-1} \cdot Y$. However, even when M can be inverted, $S = M^{-1} \cdot Y$ is in general not a satisfactory solution (see Section 2.2).

As the unfolding is equivalent to a reshuffling of events from the bins of the measured distribution into the bins of the true distribution the unfolding is not restricted to 1-dimensional distributions but can in the same way be applied to multi-dimensional distributions. The information necessary for the unfolding procedure is completely contained in the corresponding migration matrix, which in the case of multi-dimensional distributions describes the migration of events from the bins of the true multi-dimensional distribution into the bins of the measured multi-dimensional distribution. The dimensions of the measured and true distributions may also be different. An example for an unfolding in two dimensions is given in Ref. [21].

2.2. The direct solution of $Y = M \cdot S$

Very generally, the solution S of the system of linear equations (2) can be obtained by minimising the least-squares expression

$$\chi_0^2 = (Y - M \cdot S)^T \cdot K^{-1} \cdot (Y - M \cdot S) \tag{3}$$

where the n_b components of S are the free parameters. Minimising χ_0^2 will yield solutions for S which, after folding with M, are best compatible with the measurement Y.

Two cases have to be distinguished:

2.2.1. The underconstrained case $n_r = n_a \leqslant n_b$

Because $n_r = n_a$ the $n_a \times n_a$ matrix G can be inverted and a particular solution S_0 can be written as

$$S^0 = M^T \cdot C \quad \text{with } C = G^{-1} \cdot Y. \tag{4}$$

If $n_a < n_b$, the solutions $S = S_0 + S_T$ form a space of $(n_b - n_a)$ dimensions with $M \cdot S_T = 0$. For $n_a = n_b$ Eq. (4) reduces to

$$S^0 = M^T \cdot G^{-1} \cdot Y = M^{-1} \cdot Y \tag{5}$$

and S^0 is the only solution. In both cases $M \cdot S = M \cdot S^0 = Y$, implying $\chi_0^2 = 0$. The solutions are independent of the covariance matrix K. Moreover,

because of (1) the total number of events is not changed:

$$\sum_i Y_i = \sum_i \sum_j M_{ij} \cdot S_j^0 = \sum_j S_j^0. \quad (6)$$

2.2.2. The overconstrained case $n_a > n_b = n_r$

Minimising χ_0^2 by varying S yields

$$S^{LSQ} = H^{-1} M^T K^{-1} \cdot Y \quad (7)$$

where H is the $n_b \times n_b$ matrix

$$H = M^T K^{-1} M. \quad (8)$$

The solution S^{LSQ} now depends on K. The minimum value of χ_0^2 becomes

$$\chi_0^2 = (Y - M \cdot S^{LSQ})^T \cdot K^{-1} \cdot (Y - M \cdot S^{LSQ}). \quad (9)$$

Expression (7) is also valid for $n_r = n_a = n_b$, in which case it reduces to $S^{LSQ} = M^{-1} \cdot Y$.

It can be shown that the direct solutions (4) and (7) may lead to large errors of S_j, reflected in large absolute values of the elements of the error matrix T of S. This behaviour can be traced back to small eigenvalues of the matrix G and H, respectively [12].

3. Unfolding with regularisation—the different unfolding methods

In order to reduce the large errors of S, a procedure called regularisation is applied. By the regularisation additional constraints are imposed on S, by which some information in the measurements Y_i is discarded. Regularisation can be viewed as a smearing of the unfolded distribution with some finite resolution, which reduces the correlations between the S_i of adjacent bins at the expense that S_i is no longer an unbiased estimate of the true distribution [16].

The bias increases with increasing regularisation strength. Nevertheless, it turns out that with properly tuned regularisation (see Section 6) solutions can be obtained which are much closer to the true distribution than the direct solutions (4) or (7).

It is evident that regularisation is particularly important in the underconstrained case. However, also in the overconstrained case regularisation makes sense: even if the system of equations (2) is formally overconstrained ($n_a > n_b$), it may be effectively underconstrained. This happens for example if some of the measurements Y_i have much larger errors than the other Y_i.

In the following, three different ways of regularisation are described [12].

3.1. Adding a regularisation term in the expression for χ_0^2

In some unfolding methods regularisation is performed by adding a regularisation term Reg(S) in the expression for χ_0^2 (Eq. (3))

$$\chi^2 = \frac{w}{2} \cdot \chi_0^2 + \mathrm{Reg}(S) \quad (10)$$

w, also called regularisation parameter, is a weight which allows to steer the regularisation strength: large values of w correspond to weak regularisation, small values to strong regularisation.

3.1.1. Tikhonov's method

In Tikhonov's method [4] the regularisation term is defined as

$$\mathrm{Reg}(S) = \sum_{j=1}^{n_b} \left(\frac{d^2 S}{dx^2} \right)_j^2. \quad (11)$$

For the second derivative $(d^2S/dx^2)_j$ of S in bin j different approximations may be used. The expression used in the MAGIC software [24] is

$$\left(\frac{d^2 S}{dx^2} \right)_j = 2.0 \cdot \left(\frac{S_{j+1} - S_j}{S_{j+1} + S_j} - \frac{S_j - S_{j-1}}{S_j + S_{j-1}} \right). \quad (12)$$

This is actually an approximation for the bin-to-bin variation of $\Delta S/S$. In Ref. [7] $(d^2S/dx^2)_j$ is calculated from a spline representation of S.

For a given value of w and after specifying Reg(S), expression (10) can be minimised numerically by varying the components of S. The minimisation also provides the error matrix T of S. The regularisation matrix R, defined in Eq. (24), can be calculated numerically by performing minimisations with modified values of Y_i.

3.1.2. Schmelling's method

In this method, which is discussed in great detail in Refs. [13,16], the regularisation term is set equal to the "cross-entropy"

$$\mathrm{Reg}(S) = \sum_{j=1}^{n_b} p_j \cdot \ln \frac{p_j}{\varepsilon_j} \quad (13)$$

p_j is the normalised distribution S

$$p_j = \frac{S_j}{\sum_{k=1}^{n_b} S_k}, \quad \sum_{j=1}^{n_b} p_j = 1 \quad (14)$$

and ε is a normalised prior distribution, which describes a prior knowledge about S. The cross-entropy Reg(S) quantifies by how much p deviates from ε. Finding S by minimising the cross-entropy Reg(S) simultaneously with the least-squares expression χ_0^2 is called the method of "Reduced Cross-Entropy".

With Reg(S) from (13), expression (10) can now be minimised to obtain the unfolded distribution S. Note that all components of S and of the prior distribution are required to be >0, because otherwise Reg(S) in Eq. (13) cannot be defined. The expressions for the error matrix T of S and for the regularisation matrix R are given in Refs. [13,16].

In the MAGIC software, the condition $\sum_i Y_i = \sum_j S_j$ is used as an additional constraint, when minimising χ^2. By this one degree of freedom is gained.

3.2. Spectral window method

In some unfolding methods regularisation is performed by suppressing small eigenvalues λ_l of G by a factor $f(\lambda_l)$ [11]. By the suppression factor $f(\lambda_l)$ the matrix G and its inverse are modified. In terms of the eigenvectors g_l of G they read

$$\widetilde{G} = \sum_{l}^{n_r} f(\lambda_l) \cdot \lambda_l \cdot g_l g_l^T \tag{15}$$

$$\widetilde{G^{-1}} = \sum_{l}^{n_r} \frac{f(\lambda_l)}{\lambda_l} \cdot g_l g_l^T \tag{16}$$

where the sums extend over all eigenvalues λ_l which are different from zero. Like G and G^{-1}, $\widetilde{G}$ and $\widetilde{G^{-1}}$ are $n_a \times n_a$ matrices. Without suppression, $f(\lambda_l) = 1$, $\widetilde{G^{-1}}$ is equal to G^{-1} in the underconstrained case $n_r = n_a \leqslant n_b$. In the overconstrained case, $n_a > n_b = n_r$, G^{-1} is undefined but $\widetilde{G^{-1}}$ can be calculated.

A similar factor $f(\kappa_l)$ can be defined to suppress small eigenvalues κ_l of H (Eq. (8)). There is considerable freedom as to the choice of the values or expressions for $f(\lambda_l)$ and $f(\kappa_l)$. One may introduce a parameter i such that in the limit $i \to \infty$ the suppression factors $f(\lambda_l, i)$ and $f(\kappa_l, i)$ tend to 1. i has a similar meaning as the weight w in Eq. (10): it determines the regularisation strength and for $i \to \infty$ the solutions tend to the direct solutions (4) and (7), respectively.

The expressions for the error matrix T of S and for the regularisation matrix R are given in Ref. [12].

3.3. Regularisation by iteration

Another way of unfolding is to calculate a solution S iteratively [3,9,11]. The regularisation is done by stopping the iteration at some point. In this case the number of iterations i plays a similar role as the weight w in Eq. (10). In the limit of an infinite number of iterations, which is equivalent to a very large weight w, the solution tends to the direct solution (4) or (7), respectively.

3.3.1. $n_r = n_a \leqslant n_b$

A simple iteration scheme [12] for solving $Y = M \cdot S$ with $G^{-1} \cdot Y = C$ is

$$C^{i+1} = C^i - \tau \cdot (GC^i - Y) \tag{17}$$

where i is the iteration number and τ is a relaxation parameter. The latter should be chosen in the range $0 < \tau < 2/\lambda_{max}$, where λ_{max} is the largest eigenvalue of G. Eq. (17) leads to

$$C^i = (1 - \tau G)^i \cdot C^0 + \tau \cdot \sum_{j=0}^{i-1} (1 - \tau G)^j \cdot Y \tag{18}$$

where C^0 is the starting value of C. The unfolded distribution is obtained as $S^i = M^T \cdot C^i$.

In terms of the suppression factor $f(\lambda_l, i)$ this regularisation can be expressed as [12]

$$f(\lambda_l, i) = [1 - (1 - \tau\lambda_l)^i] \tag{19}$$

if C^0 is set to zero, and

$$f(\lambda_l, i) = [1 - (1 - \tau\lambda_l)^i + (1 - \tau\lambda_l)^i \lambda_l] \tag{20}$$

if C^0 is chosen to be equal to Y. The solutions S^i tend to the direct solution $S^0 = M^T \cdot C = M^T \cdot G^{-1} \cdot Y$ (Eq. (4)).

This procedure can also be applied in the overconstrained case, $n_a > n_b = n_r$. In this case G^{-1} is undefined and it has to be replaced by $\widetilde{G^{-1}}$ (Eq. (16)). The solutions S^i tend to $S = M^T \cdot \widetilde{G^{-1}} \cdot Y$, with $f(\lambda_l) = 1$.

3.3.2. $n_a > n_b = n_r$

In a similar way one may define suppression factors $f(\kappa_l, i)$ for the eigenvalues of the matrix H. In this case the solution S tends to the solution S^{LSQ} (Eq. (7)).

In the representations (19) or (20) of $f(\lambda_l, i)$ or $f(\kappa_l, i)$, the integer iteration number i can be extended to real values.

In the last two methods, Spectral Window method and Regularisation by iteration, χ_0^2 does not appear explicitly. However, its value can be calculated and it is taken into account when determining the optimum regularisation strength (see Section 6).

Once the unfolded distribution S_k is determined, by any of the methods described in this section, the differential energy spectrum Φ_k of γ-rays is calculated using Eq. (22). Φ_k has the meaning of the average differential γ-ray flux in the kth bin of E_{true} (Eq. (23)).

4. Forward unfolding

An implicit unfolding can be done by representing S as a parametric function $S_k(q) = f(E_k; q)$ with parameters $q = (q_1, q_2, \ldots, q_{n_q})$ and minimising χ_0^2 in Eq. (3)

$$\chi_0^2 = \sum_{i,j=1}^{n_a} \left(Y_i - \sum_{k=1}^{n_b} M_{ik} \cdot S_k(q) \right) \cdot (K^{-1})_{ij} \cdot \left(Y_j - \sum_{l=1}^{n_b} M_{jl} \cdot S_l(q) \right) \tag{21}$$

with respect to the parameters q. The number of measurements is equal to n_a, the number of unknowns n_q. Thus, the problem is overconstrained if $n_q < n_a$, independent of the value of n_b. One degree of freedom is gained if the total number of events is required to stay constant: $\sum_j S_j = \sum_i Y_i$. If $n_q < n_b$, which is often the case, forward unfolding is better constrained than the normal unfolding, leading to more robust results.

In many cases the minimisation of χ_0^2 can be performed analytically, by solving $\partial\chi_0^2/\partial q = 0$, similarly to the procedure described in Section 2.2.

The parametrisation of S can be written in the form

$$S_k(q) = \Phi_k(q) \cdot T_{\text{eff}} \cdot \Delta(E^k_{\text{true}}) \cdot A^k_{\text{eff}} \cdot A^k_{\text{addcut}} \cdot A^k_{\text{absorption}} \tag{22}$$

with

$$\Phi_k(q) = \frac{\int_{\Delta(E^k_{\text{true}})} \Phi(E_{\text{true}}, q) \cdot \mathrm{d}E_{\text{true}}}{\Delta(E^k_{\text{true}})}. \tag{23}$$

Here $\Phi(E_{\text{true}}, q)$ is the assumed parametrisation of the differential energy spectrum of γ-rays, T_{eff} is the effective observation time, $\Delta(E^k_{\text{true}})$ denotes the kth bin in E_{true} or its width, A^k_{eff} is the effective collection area, and A^k_{addcut} is the reduction factor due to an additional cut (for example in E_{est}, see Section 7.3). A further correction $A^k_{\text{absorption}}$ can be introduced, if $\Phi(E_{\text{true}}, q)$ is supposed to represent the differential energy spectrum of γ-rays before absorption, either at the γ-source or by interaction with the extragalactic photon background. A^k_{addcut} can be determined from MC simulations, whereas $A^k_{\text{absorption}}$ can be calculated in models about the extragalactic photon background [25].

Parametrising S as an analytic function of E_{true}, with some free parameters q, can be understood as a kind of regularisation, because it forces the solution S and its derivatives to be continuous, leading to a suppression of the noise component of S. The Forward Unfolding does not provide an unfolded distribution S. It provides those parameter values q for an assumed parametrisation of $\Phi(E_{\text{true}})$, which minimize χ_0^2 in Eq. (21). Of course, S can then be calculated from $\Phi(E_{\text{true}})$ via (22). In contrast to the unfolding methods discussed in Section 3, in the Forward Unfolding the error of the migration matrix M can be taken into account very easily.

Under the assumption of a certain parametrisation of $\Phi(E_{\text{true}})$, the Forward Unfolding is a very robust method of determining the best parameter values q. Moreover, since there is no regularisation strength to be adjusted, the uncertainty as to its choice does not exist. Therefore Forward Unfolding represents a powerful and useful check of the unfolding results obtained by any of the methods described in Section 3. In those methods the parametrisation is only introduced after the actual unfolding of the measurements Y.

5. Useful quantities in the unfolding

In this section two quantities are explained which are useful for judging the quality of the unfolding result: the error matrix (covariance matrix) T of the unfolded distribution S and the regularisation matrix R.

5.1. The covariance matrix of the unfolded distribution S

In all unfolding methods the covariance matrix T of S can be determined. The trace of T, Trace(T), measures the noise component of S, as Trace(K) measures the noise component of Y.

In the methods where the solution S is given as $S = D \cdot Y$, like in Schmelling's method or in the Spectral Window method, the covariance matrix T of S is obtained by $T = D \cdot K \cdot D^{\mathrm{T}}$.

In those methods where the solution is determined by a numerical minimisation of χ^2, like in Tikhonov's method, T is obtained from the shape of $2.0 * \chi^2/w$ in the region around the minimum.

5.2. The regularisation matrix R

A quantity which describes how the estimates $\sum_{k=1}^{n_b} M_{jk} S_k$ of Y_j couple to the measurements Y_i is given by the $n_a \times n_a$ matrix

$$R_{ij} = \frac{\partial(\sum_{k=1}^{n_b} M_{jk} S_k)}{\partial Y_i} \quad \text{or} \quad R = \frac{\partial(M \cdot S)}{\partial Y} \tag{24}$$

also called regularisation matrix [16]. The trace of R can be interpreted as the effective number of measurements used in the unfolding procedure. The number of effectively rejected measurements is then equal to $N_{\text{rej}} = n_a -$ Trace(R).

The maximum value of Trace(R), which is equal to the rank n_{r} of G, is reached with the direct solutions (4) and (7), corresponding to the cases $n_{\text{r}} = n_a \leqslant n_b$ and $n_a > n_b = n_{\text{r}}$, respectively: For the direct solution S^0 the number of rejected measurements N_{rej} is equal to $n_a - \text{Trace}(R) = n_a - n_{\text{r}} = n_a - n_a = 0$, which means that no information is discarded. The measurements Y are completely reproduced by the unfolding: $M \cdot S^0 = Y$. For the least-squares solution S^{LSQ} the number of effectively rejected measurements is $N_{\text{rej}} = n_a - \text{Trace}(R) = n_a - n_{\text{r}} = n_a - n_b > 0$, which means that some information is discarded. The measurements Y are not exactly reproduced by the unfolding: $M \cdot S^{\text{LSQ}} \neq Y$. The fact that the system is overconstrained has a similar effect as regularisation. In both cases, with increasing regularisation strength Trace(R) is reduced and N_{rej} is increased.

6. Selecting the unfolding result

For a given unfolding method the result S depends on the regularisation strength, which is given by the weight w (or by the number of iterations i, respectively). In the literature various criteria for choosing the "best" weight are proposed [8,10,12,15,16,20]. Unfortunately, none of them provides a choice which is optimal for all cases. Reasons for this are: The optimum regularisation strength in general depends on the shape of the unknown distribution S. It also depends on the binnings in E_{true} and E_{est} and on the prior distribution (if applicable).

The effect of the regularisation is illustrated in Fig. 1, where different quantities are plotted as a function of the iteration number i. In this example an experimental energy distribution of γ-rays from the Crab Nebula [26], i.e. the

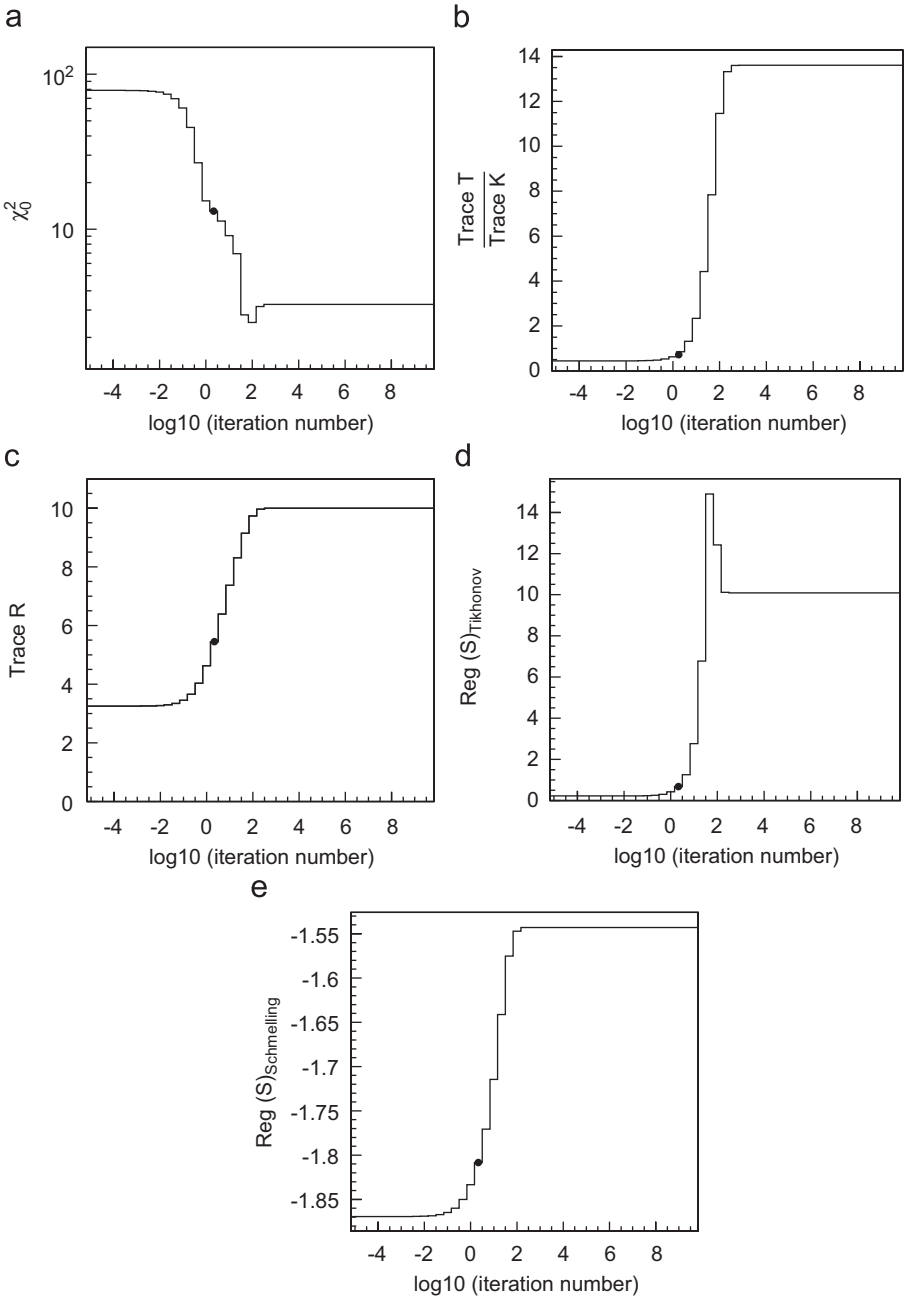

Fig. 1. Useful quantities for determining the optimum regularisation strength, plotted as a function of the iteration number i. (a) χ_0^2. (b) Trace(T)/Trace(K). (c) Trace(R). (d) Reg(S)$_{\text{Tikhonov}}$. (e) Reg(S)$_{\text{Schmelling}}$. The full circles indicate the values of the different quantities for the optimum regularisation strength.

number of excess events in bins of the estimated energy E_{est}, was unfolded. The unfolding was performed for 30 different i in the range $10^{-5}-10^{10}$, using the method of Bertero (Eq. (20)). More results from the analysis of these data are given in Section 10. With decreasing i, i.e with increasing regularisation strength, one observes an increase of χ_0^2 (Eq. (3)) and a decrease in the quantities Trace(T)/Trace(K), Trace(R), Reg(S)$_{\text{Tikhonov}}$ (Eq. (11)) and Reg(S)$_{\text{Schmelling}}$ (Eq. (13)). Very similar behaviour is found for the other unfolding methods, discussed in Section 3.

Obviously, an acceptable unfolding result should satisfy the following conditions:

- The χ^2-probability, calculated from the value of χ_0^2 and the number of degrees of freedom in the unfolding procedure, should be acceptable, say $>1\%$. Otherwise the unfolding result is incompatible with the measured distribution Y.
- The noise term Trace(T) of the unfolded distribution S should be comparable to the noise term Trace(K) of the measurements. The main aim of regularisation is a

suppression of the large noise term of S, which one often obtains if no regularisation is applied. A large noise term Trace(T), as well as large correlation terms of T, indicate a too fine binning in E_{true}, leading to small eigenvalues of G or H (see Section 7.1).
- Trace(R) should not be much lower than its maximum possible value, which is equal to the rank n_{r} of the matrix G. Otherwise the solution is too strongly dominated and biased by the regularisation.

For determining the "best" regularisation strength a compromise has to be found between the above requirements. It has turned out that the criterion $\text{Trace}(T) = \text{Trace}(K)$ in general leads to solutions which satisfy the above conditions reasonably well, provided the problem is not strongly overconstrained. In the latter case, where the unfolding result is better constrained, a solution with $\text{Trace}(T) < \text{Trace}(K)$ is more appropriate. In MAGIC the standard criterion for determining the optimal regularisation strength is $\text{Trace}(T) = \text{Trace}(K)$. The full circles in Fig. 1 indicate this choice. However, any other regularisation strength can be chosen by hand, if this is suggested by the behaviour of the quantities χ^2_0, Trace(T), Trace(R), $\text{Reg}(S)_{\text{Tikhonov}}$ or $\text{Reg}(S)_{\text{Schmelling}}$.

Unfolding with regularisation is a procedure which allows freedom in the choice of the regularisation method and in the choice of the regularisation strength. The above criteria for an acceptable solution strongly restrict this freedom. Nevertheless a certain degree of arbitrariness remains as to which unfolding result should be considered representative and final. In MAGIC a selected unfolding result is considered representative if all other unfolding methods yield results, which are also acceptable and statistically consistent with the selected result. In addition, it is required that also the Forward Unfolding (Section 4), using a reasonable parametrisation of $\Phi(E_{\text{true}})$, gives a consistent result.

The covariance matrix of the unfolded distribution S is obtained as explained in Section 5.1. An additional uncertainty due to the unfolding is determined from the spread of the S_j, obtained from the different unfolding methods. The error of S due to the error of the migration matrix M is usually not considered.

The uncertainty as to the choice of the regularisation strength can be expressed as $\text{Trace}(T) = \text{Trace}(K) * (1 \pm \delta)$. δ has the meaning of the relative uncertainty of Trace(T). This uncertainty corresponds to an error of the covariance matrix T rather to an error of the unfolded distribution S. No additional error is therefore assigned to S due to the uncertainty of the regularisation strength.

7. Further comments on the unfolding

In the actual application of the unfolding procedure to real data some technical details have to be considered, and they are discussed in this section.

7.1. Optimal binnings

The binning of the experimental distribution Y is often dictated by the available statistics and by the experimental errors. The binning should not be chosen too fine in order to assure significant measurements in all bins. In the case of an IACT experiment, a sufficiently large sample is required to determine the number of signal (excess) events with sufficient accuracy. The binning should not be chosen too wide either, because the binning in Y limits the reconstruction of the fine structure of the unfolded distribution S.

Another criterion for the binnings is the behaviour of Gram's matrix G. A too fine binning for S leads to strong correlations between neighbouring columns of the migration matrix M, implying small eigenvalues of G, which lead to a large noise component of S. Given a certain choice of the binning in E_{est} and thus of n_a, the bin size in E_{true} or n_b should be set such, that the system of linear equations (2) is not underconstrained. This usually leads to wider bins in E_{true} than in E_{est}. In MAGIC a typical value of $\Delta\log_{10}(E_{\text{true}})/\Delta\log_{10}(E_{\text{est}})$ is 1.4 (see Section 10).

It should also be noted that the unfolding procedure does not require equidistant bins, neither in E_{est} nor in E_{true}.

7.2. Completeness of the migration matrix

If $n_{a1} \leqslant i \leqslant n_{a2}$ specifies the range of bins of the measured distribution Y which are to be considered in the unfolding procedure, also the range in E_{true} to be considered in the unfolding has to be chosen properly: one has to make sure that all bins j of E_{true} are present, for which the column j of the migration matrix M_{ij} contributes to the selected bins of Y, i.e. for which at least one of the elements M_{ij} ($n_{a1} \leqslant i \leqslant n_{a2}$) is different from zero.

There is an exception to this rule, if for some reason certain bins j of E_{true} are not expected to contribute to the selected bins in E_{est}. This is for example the case if one of the factors A^k_{xxx} in Eq. (22) is so small that S_k can be neglected.

7.3. Additional cuts

As explained in Section 2.1, the distribution Y and the migration matrix M have to be produced under identical conditions (selections, cuts). If an additional cut is imposed when generating Y, also M has to be recalculated before doing the unfolding. If this additional cut is a cut in E_{est} one may proceed in the following way:

- Renormalise the columns j of M to the selected range in E_{est} (see Eq. (1)).
- Perform the unfolding of Y in the usual way.
- Apply a correction A^j_{addcut} to the unfolded distribution S_j, where A^j_{addcut} is the renormalisation factor for column j of M.

7.4. Starting values for the minimisation

In the cases where the unfolding procedure involves numerical minimisations, like the ones discussed in Section 3.1, the minimisation may not converge. This problem can be often solved by choosing different starting values. Another reason for non-convergence is discussed in Section 7.5. In MAGIC the standard choice of the starting distribution for S, and also of the prior distribution ε in Schmelling's method, is a distribution which is close to the measured distribution Y.

7.5. Components of S which cannot be determined in the unfolding

According to Eq. (2) those S_j for which the column M_{ij} $(i = 1, \ldots, n_a)$ is a null vector have no influence on Y and can therefore not be determined in the unfolding. These components should not be varied in the minimisation because they would unnecessarily complicate the minimisation process and may lead to non-convergence.

7.6. Dependence on the assumptions made in the MC simulation

For the unfolding the migration matrix M_{ij} is the crucial quantity. Obviously, if it does not describe the real migration of events correctly, the unfolding result will be wrong. This means that at fixed j, i.e. at fixed E_{true}, the MC simulation has to describe the migration in E correctly. This will be the case if at fixed E_{true} the shower simulation is realistic and if the detector response is simulated correctly.

On the other hand, the distribution of E_{true} in the MC need not agree with the real distribution of E: due to the normalisation of M (Eq. (1)) the bin-to-bin variation of the number of MC events in E_{true} has no influence on M_{ij} at all. This is one of the great advantages of unfolding methods like those presented in Section 3 as compared to methods based on correction factors (see Section 11).

However, there is a residual dependence of M_{ij} due to the finite binning in E_{true}: Depending on the shape of the E_{true} distribution within an E_{true} bin in the MC simulation, the calculated M_{ij} may be more representative for the lower, middle or upper part of the E_{true} bin. If the E_{true} distribution in the real data is different from that in the MC simulation the calculated M_{ij} may not be exactly the right ones.

This residual dependence of M on the shape of the E_{true} distribution in the MC can be nearly completely removed by an iteration procedure in which the M for the next iteration step is determined from a MC sample, in which the E_{true} distribution has been corrected using the unfolding result of the last iteration step (see Section 8).

8. Determining the effective collection area A and the migration matrix M for a finite bin in E_{true} and Θ

In an IACT experiment it is important that the effective collection area A, which enters in the flux calculation, is computed taking into account a realistic shape of the differential flux $\Phi(E_{\text{true}})$ and the actual distribution of effective observation times $\mathrm{d}T(\Theta)/\mathrm{d}\Theta$ in the zenith angle Θ. This is also important for the migration matrix M, which enters in the unfolding, because of the residual dependence on the flux spectrum, as discussed in Section 7.6. Recalculating A and M with the proper E_{true} and Θ spectra is the more important the bigger the $(\Delta E_{\text{true}}, \Delta\Theta)$ interval and the stronger the variations of M and A within this bin are. Often, for statistics reasons, large bin sizes in E_{est} and Θ, and thus also in E_{true}, have to be chosen.

In the following it is assumed that the effective collection area A and the migration matrix M are known functions of E_{true} and Θ. This can be achieved by determining them from a sample of MC γ-ray events in very fine bins of E_{true} and Θ. The aim is to calculate an average $\overline{A}$ of the effective collection area and an average $\overline{M}$ of the migration matrix, which are representative for a finite bin $(\Delta E_{\text{true}}, \Delta\Theta)$ in E_{true} and Θ.

8.1. The effective collection area

The number of observed events in a $(\Delta E_{\text{true}}, \Delta\Theta)$ bin is given by

$$S = \int_{\Delta\Theta}\int_{\Delta E_{\text{true}}} A(E_{\text{true}}, \Theta)\cdot\Phi(E_{\text{true}}) \cdot\frac{\mathrm{d}T(\Theta)}{\mathrm{d}\Theta}\cdot\mathrm{d}E_{\text{true}}\cdot\mathrm{d}\Theta. \tag{25}$$

Here $\Phi(E_{\text{true}})$ is the differential γ-ray flux to be measured, $\mathrm{d}T(\Theta)/\mathrm{d}\Theta$ is the distribution of observation times in the experimental data and $A(E_{\text{true}}, \Theta)$ is the known dependence of the effective collection area on E_{true} and Θ.

With the definitions

$$\Delta T = \int_{\Delta\Theta}\frac{\mathrm{d}T(\Theta)}{\mathrm{d}\Theta}\cdot\mathrm{d}\Theta \quad (= \text{total observation time}), \tag{26}$$

$$\overline{\Phi} = \frac{1}{\Delta E_{\text{true}}}\int_{\Delta E_{\text{true}}}\Phi(E_{\text{true}})\cdot\mathrm{d}E_{\text{true}}, \tag{27}$$

$$\overline{A} = \frac{\int_{\Delta\Theta}\int_{\Delta E_{\text{true}}} A(E_{\text{true}}, \Theta)\cdot\Phi(E_{\text{true}})\cdot(\mathrm{d}T(\Theta)/\mathrm{d}\Theta)\cdot\mathrm{d}E_{\text{true}}\cdot\mathrm{d}\Theta}{\Delta T\cdot\overline{\Phi}\cdot\Delta E_{\text{true}}}. \tag{28}$$

Eq. (25) can be rewritten as

$$\overline{\Phi} = \frac{S}{\Delta T\cdot\overline{A}\cdot\Delta E_{\text{true}}}. \tag{29}$$

This is the usual formula for converting numbers of events S into differential fluxes Φ. Because of definitions (27) and (28) the differential flux $\overline{\Phi}$ in an E_{true} bin, as determined in the unfolding, is the average differential flux in this bin. Therefore, when quoting or plotting a result for $\overline{\Phi}$ the bin edges in E_{true} should also be given or shown.

8.2. The migration matrix

The number of reconstructed MC events in a bin i of E_{est} can be written as

$$N_i = \int_{\Delta\Theta}\int_{\Delta E_{\mathrm{true}}} M_i(E_{\mathrm{true}},\Theta)\cdot A(E_{\mathrm{true}},\Theta) \cdot \Phi(E_{\mathrm{true}})\cdot\frac{\mathrm{d}T(\Theta)}{\mathrm{d}\Theta}\cdot \mathrm{d}E_{\mathrm{true}}\cdot \mathrm{d}\Theta \tag{30}$$

with

$$\sum_k M_k(E_{\mathrm{true}},\Theta)=1 \quad \text{for all } E_{\mathrm{true}} \text{ and } \Theta.$$

$M_i(E_{\mathrm{true}},\Theta)$ is the element of the normalised migration matrix for the ith bin in E_{est}, at an energy E_{true}. The index j of M_{ij} is replaced by the variable E_{true}. The dependence of $M_i(E_{\mathrm{true}},\Theta)$ on E_{true} and Θ is assumed to be known from MC simulations.

The average migration matrix $\overline{M}_i$ for the selected ΔE_{true} bin, to be used in the unfolding, is then obtained by

$$\overline{M}_i = \frac{N_i}{\sum_k N_k} = \frac{\int_{\Delta\Theta}\int_{\Delta E_{\mathrm{true}}} M_i(E_{\mathrm{true}},\Theta)\cdot A(E_{\mathrm{true}},\Theta)\cdot\Phi(E_{\mathrm{true}})\cdot(\mathrm{d}T(\Theta)/\mathrm{d}\Theta)\cdot\mathrm{d}E_{\mathrm{true}}\cdot\mathrm{d}\Theta}{\int_{\Delta\Theta}\int_{\Delta E_{\mathrm{true}}} A(E_{\mathrm{true}},\Theta)\cdot\Phi(E_{\mathrm{true}})\cdot(\mathrm{d}T(\Theta)/\mathrm{d}\Theta)\cdot\mathrm{d}E_{\mathrm{true}}\cdot\mathrm{d}\Theta}. \tag{31}$$

The averages $\overline{M}_i$ and $\overline{A}$ are calculated according to (31) and (28), respectively, using an approximation $\Phi_1(E_{\mathrm{true}})$ of the function $\Phi(E_{\mathrm{true}})$. The measured distribution Y is unfolded using $\overline{M}_i$, yielding the unfolded distribution S. A new approximation $\Phi_2(E_{\mathrm{true}})$ is then determined from S according to Eq. (29). The procedure is iterated until $\Phi(E_{\mathrm{true}})$ has converged. In practice parametric functions are used as approximations of $\Phi(E_{\mathrm{true}})$, and in general the convergence is found to be very fast.

9. Combining data before applying the unfolding procedure

One often has the situation that there exist several measured distributions Y^v of the same quantity. If the Y^v were obtained under different conditions also the migration matrices M^v and the effective collection areas A^v will be different for the different measurements.

In an IACT experiment the different conditions may be

- Different modes of observation (ON/OFF mode, wobble mode, observation in the presence of moon light, …).
- Different ranges of the zenith angle.
- Different detector conditions.
- etc.

In order to determine a final unfolded distribution S one may proceed in different ways:

- *Individual unfolding*: Unfold each Y^v using M^v to obtain S^v, and combine the S^v to obtain the final solution S.
- *Global unfolding*: Combine the Y^v and M^v to obtain a global Y and M, do an unfolding of Y using M, which will give the final solution S.

There is one important argument in favour of the second option: Each of the Y^v, or some of them, may have large statistical errors, making the unfolding of the individual Y^v unstable. One common unfolding of the global Y using the global M will be more robust, in general.

In the case of an IACT experiment, the following relations hold for each measurement v:

$$Y^v = M^v\cdot S^v,$$
$$S^v_j = T^v\cdot A^v_j\cdot\Phi_j\cdot\Delta E_j \tag{32}$$

j denotes the jth bin in E_{true}, ΔE_j is its width, Φ_j is the average flux and A_j is the effective collection area in this bin, and T is the effective observation time. Since Φ and ΔE are the same for all v the relations for the combined data read:

$$Y = M\cdot S,$$
$$S_j = T\cdot A_j\cdot\Phi_j\cdot\Delta E_j. \tag{33}$$

Inserting $Y=\sum_v Y^v$ and $T=\sum_v T^v$ in Eq. (33) and using (32) and $\sum_i M_{ij}=\sum_i M^v_{ij}=1$ one obtains

$$A_j = \frac{\sum_v A^v_j\cdot T^v}{T}, \tag{34}$$

$$M_{ij} = \frac{\sum_v M^v_{ij}\cdot(A^v_j T^v)}{\sum_\mu (A^\mu_j T^\mu)} = \frac{\sum_v M^v_{ij}\cdot(A^v_j T^v)}{A_j T}. \tag{35}$$

Relations (34) and (35) give the prescription how to combine the individual M^v to obtain the global M, and how to combine the individual A^v_j to obtain the global A_j. The measured distribution Y is unfolded using M, and the unfolded distribution S is converted into a flux Φ according to (33). As can be seen from (34) and (35) the A_j are weighted averages of the A^v_j with weights $w_v = T^v/T$, and the M_{ij} are weighted averages of the M^v_{ij} with weights $w_v = A^v_j T^v/(\sum_\mu A^\mu_j T^\mu)$. The weights can be interpreted as fractions because they add up to 1. They are correlated and their covariance matrix has to be taken into account when calculating the errors of A_j and M_{ij}.

Eqs. (34) and (35) also show that A_j and M_{ij} can be obtained without knowing which spectra $\Phi(E_{\mathrm{true}})$ and $\mathrm{d}T(\Theta)/\mathrm{d}\Theta$ were used to compute the A^v_j and M^v_{ij}.

In the approach discussed above the individual measurements Y^v are combined in the simple way $Y=\sum_v Y^v$. In cases where the resolutions of the individual measurements differ strongly a different combination may be more

appropriate. The generalisation to the case $Y = \sum_v v^v Y^v$, where the v^v are appropriate weights, is straightforward: $T = \sum_v T^v$ has to be replaced by $T = \sum_v v^v T^v$ and in Eqs. (34) and (35) T^v and T^μ have to be replaced by $v^v T^v$ and $v^\mu T^\mu$, respectively.

10. Application to experimental data

In this section an experimental energy distribution of γ-rays from the Crab Nebula is unfolded, which was obtained in an analysis of data taken with the MAGIC telescope [26]. The migration matrix M, as determined from a sample of γ-MC events, is plotted in Fig. 2(a) as a function of E_{est} and E_{true}. The size of the boxes is proportional to the value of M_{ij}, where i and j are the bin numbers in E_{est} and E_{true}, respectively. The experimental distribution Y of the number of γ-excess events as a function of E_{est} is displayed in Fig. 2(b). Both distributions are after all cuts and selections, except a cut in E_{est}.

The vertical and horizontal lines in Figs. 2(a) and (c) indicate the ranges in E_{est} or E_{true}, which were selected for the unfolding. The range in E_{est} is given by those E_{est} bins for which a significant number of excess events could be determined. The range in E_{true} comprises all those E_{true} bins which are expected to contribute to the selected range in E_{est}. Fig. 2(a) suggests that these are all 15 E_{true} bins. However, according to the plot of the effective collection area A_{eff} in Fig. 2(c) the contribution from $E_{\text{true}} < 60\,\text{GeV}$ is expected to be negligible. The same holds for $E_{\text{true}} > 9\,\text{TeV}$, due to the strongly decreasing γ-ray flux with increasing energy. This is confirmed by a Forward Unfolding of Y in which a differential γ-ray flux of the form $\mathrm{d}N/\mathrm{d}A \cdot \mathrm{d}t \cdot \mathrm{d}E = f_0 \cdot (E/300\,\text{GeV})^\alpha$, with $\alpha = a + b \cdot \log_{10}(E/300\,\text{GeV})$, was assumed. This leads to a number of E_{est} and E_{true} bins of $n_a = 14$ and $n_b = 10$, respectively, which are used in the unfolding procedure. The size of the $\log_{10}(E_{\text{true}})$ bins was deliberately chosen wider than the $\log_{10}(E_{\text{est}})$ bins by a factor of 1.4, in order to better constrain the unfolding. The rank of Gram's matrix G is equal to $n_r = 10$, as can be seen from Fig. 2(d), which shows the size of the eigenvalues λ_l of G as a function of l. Two to three of these eigenvalues are much smaller than the maximum eigenvalue, and they are the reason for the large values of $\text{Trace}(T)/\text{Trace}(K)$ at large iteration number (low regularisation strength) in Fig. 1(b).

The optimum regularisation strength and thus the final solution S was determined using the criterion $\text{Trace}(T)/\text{Trace}(K) = 1$. The estimates $\sum_j M_{ij} \cdot S_j$ (open circles) of Y_i are compared with Y_i (histogram) in Fig. 2(b). These data enter in the calculation of χ_0^2 (Eq. (3)). The number of degrees of freedom in the unfolding procedure is $n_a - n_b + 1 = 5$, because the number of measurements is

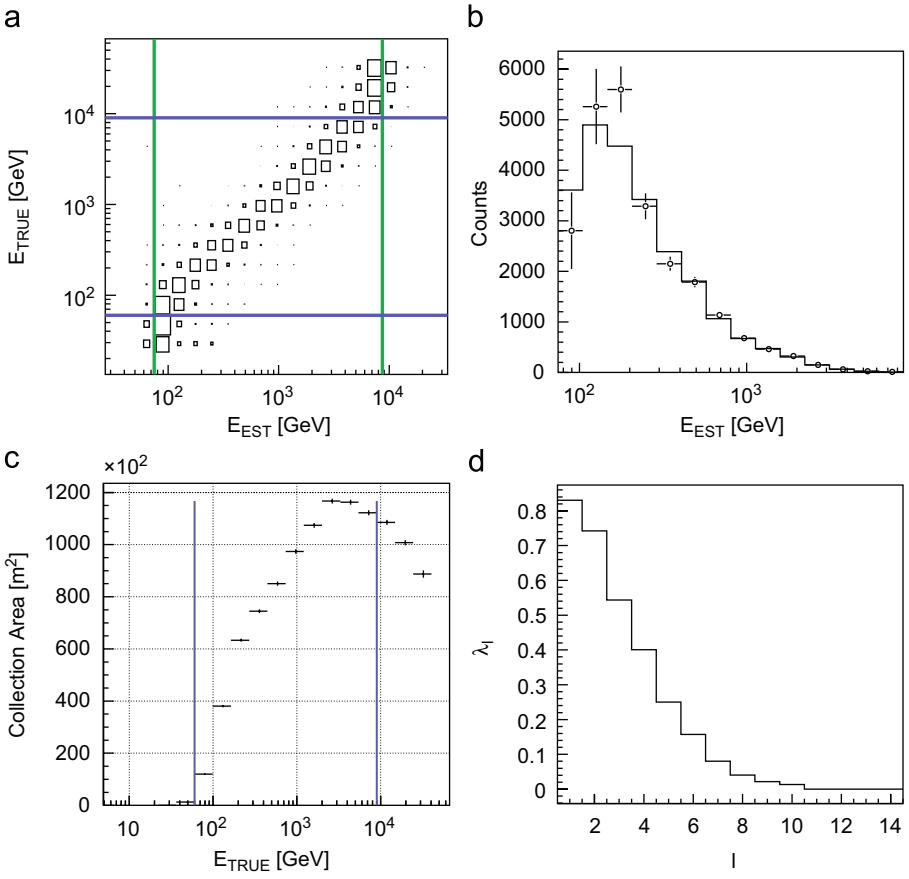

Fig. 2. Input data for the Unfolding and for the calculation of the differential γ-ray flux. (a) Original migration matrix M. (b) Distribution Y to be unfolded (open circles), and the unfolded distribution S folded with M (histogram). (c) Effective collection area A_{eff}. (d) Eigenvalues of Gram's matrix G.

n_a, the number of unknowns is n_b, and the relation $\sum_j S_j = \sum_i Y_i$ (Eq. (6)) is used as additional constraint. As can be seen from Figs. 1(d) and (e) the values of $\mathrm{Reg}(S)_{\mathrm{Tikhonov}}$ and $\mathrm{Reg}(S)_{\mathrm{Schmelling}}$ are much lower at the selected regularisation strength than without regularisation. This means that the solution S is smoothed by the regularisation. The black symbols in Fig. 3(a) represent the final solution S. The solution before correction for the cut in E_{est} is drawn with grey symbols.

The final differential γ-ray flux Φ, as computed from S according to Eq. (22), is drawn in Fig. 3(b). The solid line represents the result of a fit of the expression $f_0 \cdot (E/300\,\mathrm{GeV})^{\alpha}$, with $\alpha = a + b \cdot \log_{10}(E/300\,\mathrm{GeV})$, to the data points. The number of degrees of freedom for this fit is $n_b - 3 = 7$, because the number of data points is $n_b = 10$ and the number of free parameters is 3 (f_0, a and b). The χ^2 is 8 for 7 degrees of freedom. Setting $b = 0$ in the fit yields a χ^2 of 24 for 8 degrees of freedom. This fit is clearly disfavoured as compared to the fit in which the slope α is energy dependent. In these fits the full correlation matrix T of S has been taken into account.

The result of the latter fit for $\Phi(E_{\mathrm{true}})$ was used to recalculate the averages $\overline{M}_i$ (31) and $\overline{A}$ (28) for the individual E_{true} bins. The unfolding was repeated using the recalculated $\overline{M}_i$ and $\overline{A}$, yielding new results for S, Φ and the fit parameters. After 1 iteration this procedure converged.

Very similar results were obtained with the other unfolding methods. The spread of the S_j, obtained with the different unfolding methods, can be seen in Fig. 3(c). This spread can be regarded as an estimate of the systematic error due to the unfolding.

11. Correction factors

A widely used method of correcting experimental distributions is the application of correction factors: Using MC data, both the "true" distribution S_k^{MC} $(k = 1, \ldots, n_c)$ and the "reconstructed" distribution Y_k^{MC} $(k = 1, \ldots, n_c)$ of some quantity are produced under certain conditions (selections, cuts). Correction factors are determined according to

$$c_k = S_k^{\mathrm{MC}} / Y_k^{\mathrm{MC}} \quad (k = 1, \ldots, n_c). \tag{36}$$

An experimental distribution Y_k, obtained under the same conditions as Y_k^{MC}, is then corrected by

$$S_k = Y_k \cdot c_k \quad (k = 1, \ldots, n_c) \tag{37}$$

to obtain the corrected distribution S_k.

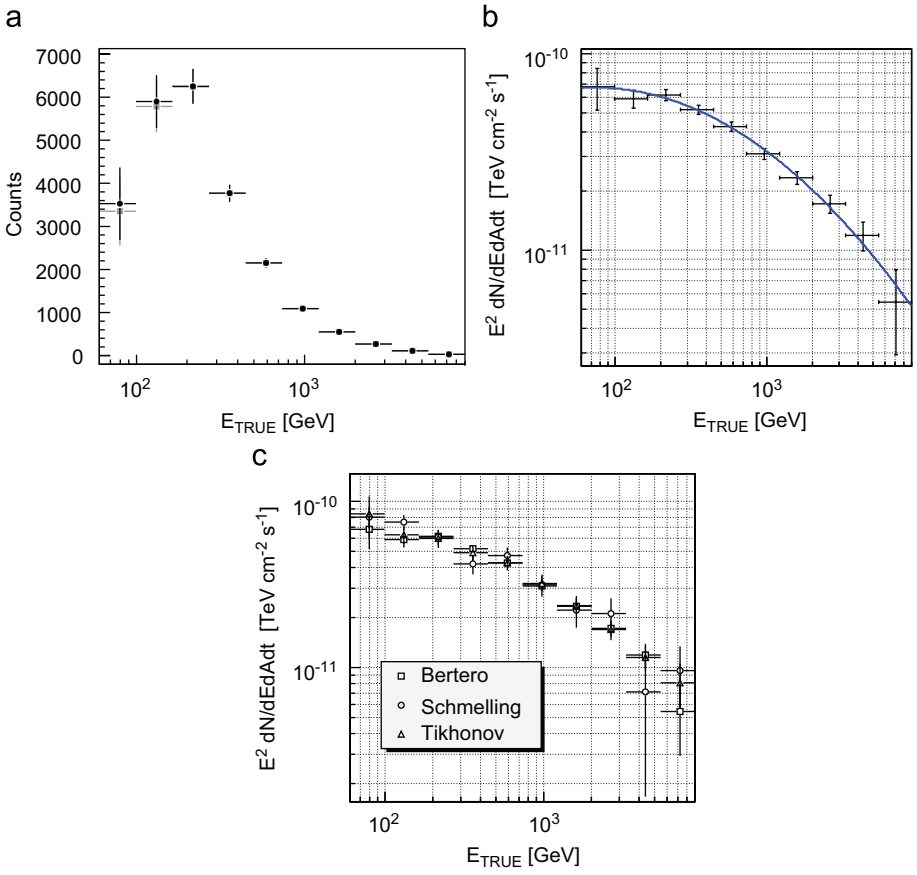

Fig. 3. Results from the Unfolding. (a) Unfolded distribution S before (grey symbols) and after the correction for the cut in E_{est} (black symbols). (b) Result for the differential γ-ray flux $\Phi(E_{\mathrm{true}})$ multiplied with E^2.] (c) Comparison of the results for $E^2 \cdot \Phi(E_{\mathrm{true}})$ from different unfolding methods.

The following properties of this procedure can be stated [16]:

- c_k is undefined if $Y_k^{\mathrm{MC}} = 0$.
- If $S_k^{\mathrm{MC}} = 0$ also $c_k = 0$ and $S_k = 0$, which means that Y_k is ignored.
- c_k depends on the shape of the MC distribution S_k^{MC}; the corrected distribution S_k is always biased towards S_k^{MC}.
- If $Y_k = 0$ also S_k is zero.
- The standard linear error propagation often yields too small errors of S_k.

The correction factors are only right if S_k^{MC} is identical to the true S_k distribution. If this is not the case one may iterate S_k^{MC}, setting S_k^{MC} equal to the last corrected experimental distribution S_k. However, this often leads to instabilities. The reason for the instabilities appears to be similar to that causing a large noise component of the direct solution (4).

In contrast to the unfolding methods presented in Section 3, there is very little freedom in choosing the binnings for S and Y. By definition, the range of values and the binnings for the true and reconstructed quantity are identical.

Advantages of the method of correction factors are that it is simple and stable. The drawbacks have been listed above, the severest one being the strong dependence of the correction factors on the assumptions made in the MC about S.

12. Summary

In this paper the procedures to unfold experimental energy distributions of γ-rays, as applied in the MAGIC experiment, are described. It is explained, how the uncertainties, which are inherent in any unfolding process, can be handled successfully. Possible problems in the unfolding are discussed and suggestions are given which can help to avoid them. Various techniques are presented, which allow to reconstruct the energy spectrum in a rather unbiased way. All algorithms are implemented in the MAGIC software, which is based on the C^{++} language and ROOT [27]. Their application to real data has shown to provide robust and reliable results. The methods and procedures are applied in most of the MAGIC analyses.

Acknowledgement

We thank Michael Schmelling for fruitful discussions and critical comments.

References

[1] C. Baixeras, et al., Nucl. Instr. and Meth. A 518 (2004) 188.
[2] R. Gold, ANL-6984, 1964, unpublished.
[3] G.I. Marchuk, Methods of Numerical Mathematics, Springer, Berlin, 1975.
[4] A.N. Tikhonov, V.Ya. Arsenin, Methods of Solution of Ill-posed Problems, Nauka, Moscow, 1979.
[5] S.W. Provencher, Comput. Phys. Commun. 27 (1982) 213, 229.
[6] V. Blobel, Unfolding methods in high-energy physics experiments, DESY 84-118, 1984.
[7] V. Blobel, 1984 CERN School of Computing, Ajguablava, Spain, CERN 85-09, 1984, p. 88.
[8] E.A. Belogorlov, et al., Nucl. Instr. and Meth. A 235 (1985) 146.
[9] S.F. Giljazov, Methods of Solution of Linear Ill-posed Problems, MSU, Moscow, 1987.
[10] V.P. Zhigunov, et al., Nucl. Instr. and Meth. A 273 (1988) 362.
[11] M. Bertero, Advances in Electronics and Electron Physics, vol. 75, Academic Press Inc., New York, 1989.
[12] V.B. Anykeyev, et al., Nucl. Instr. and Meth. A 303 (1991) 350.
[13] M. Schmelling, Nucl. Instr. and Meth. A 340 (1994) 400.
[14] V. Blobel, The RUN manual, OPAL Technical Note TN361, 1996.
[15] A. Höcker, V. Kartvelishvili, Nucl. Instr. and Meth. A 372 (1996) 469.
[16] M. Schmelling, Numerische Methoden der Datenanalyse, MPI-K Heidelberg, 1998, ⟨www.mpi-hd.mpg.de/personalhomes/michaelt/public/1998-01⟩.
[17] G. Cowan, Statistical Data Analysis, Clarendon, Oxford, 1998.
[18] G. Cowan, A survey of unfolding methods for particle physics, IPPP Workshop on Advanced Statistics Techniques in Particle Physics, Durham 2002, p. 248.
[19] W.H. Press, et al., Numerical Recipes in C++, second ed., Cambridge University Press, Cambridge, 2002.
[20] V. Blobel, An Unfolding Method for High Energy Physics, IPPP Workshop on Advanced Statistics Techniques in Particle Physics, Durham, 2002.
[21] W. Wittek, Correlations between parameters of extended air showers and their proper use in analyses, 26th International Cosmic Ray Conference, Salt Lake City, Utah, USA, 1999, astro-ph/9908029.
[22] P. Majumdar, et al. (MAGIC Collab.), 2005, Proc. of the 29th ICRC, Pune, India, 5-203, astro-ph/0508274.
[23] S. Mizobuchi, et al. (MAGIC Collab.), in: Proceedings of the 29th ICRC, Pune, India, 2005, p. 101, astro-ph/0508274.
[24] T. Bretz, R. Wagner, (MAGIC Collab.), in: Proceedings of the 28th ICRC, Tsukuba, Japan, 2003, p. 2947.
[25] T.M. Kneiske, et al., Astron. Astrophys. 413 (2004) 807.
[26] J. Albert, et al., Astrophys. J. (2007), submitted for publication, arXiv:0705.3244.
[27] R. Brun, F. Rademakers, ⟨http://root.cern.ch/⟩.