Kinematic fitting

A powerful tool of event selection and reconstruction

Sergey Yaschenko, DESY Zeuthen

University of Glasgow, 12.01.2011

Outline

> Introduction

> Techniques

> Applications

> Results of use at HERMES

> Summary

> Literature

Introduction: kinematic fitting as a part of data processing

> Raw data processing

> Detector calibrations

> Track search and reconstruction

> Momentum reconstruction

> Event selection and reconstruction

- At this stage 3-momenta of all found tracks are reconstructed
- Combine individual tracks to events
- Apply certain requirements (cuts) on track correlations to select events of interest and reject the background
- Alternatively use kinematic fitting

> Physics analysis

Introduction: kinematic fitting as a tool of event selection

- Kinematic fitting – adjustment of measured kinematic parameters under certain assumptions (conditions, constraints)

- Aims
  - Test if assumptions are true
  - Improve accuracy of measurements
  - Check if the knowledge of measurement uncertainties is correct
  - Check for possible systematic uncertainties

- Has been used for more than 50 years in particle physics, considered as a standard tool of event selection and reconstruction

- But also often considered as too complicated and not really necessary

Example 1

- Decay of a particle a to 3 particles a → 1+2+3
- Energy of the primary particle is precisely known and equal to E_a = 30 GeV
- Energies of the secondary particles are equal to
  E_1 = 5 GeV, E_2 = 10 GeV and E_3 = 15 GeV
  and measured with errors distributed by Gaussian with sigmas
  σ_1 = 1 GeV, σ_2 = 1 GeV and σ_3 = 1 GeV
- Minimization of least-squares functional

$$\chi^2 = \sum_{i=1}^{3} (E_i - E_i^{fit})^2 / \sigma_i^2$$

- under condition $$E_1^{fit} + E_2^{fit} + E_3^{fit} - E_a = 0$$

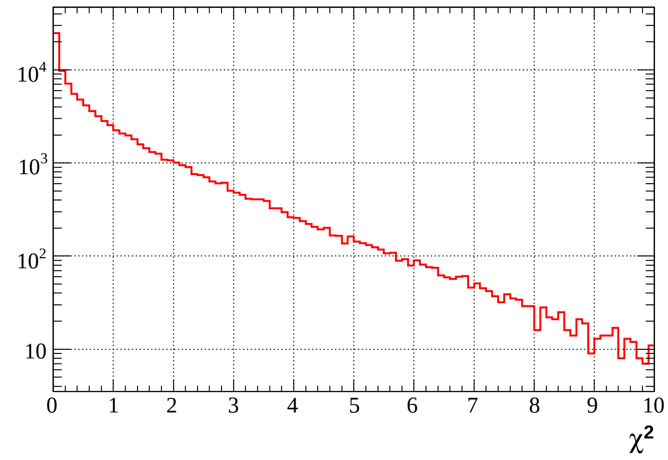
- If measurement errors are the same the result is trivial:

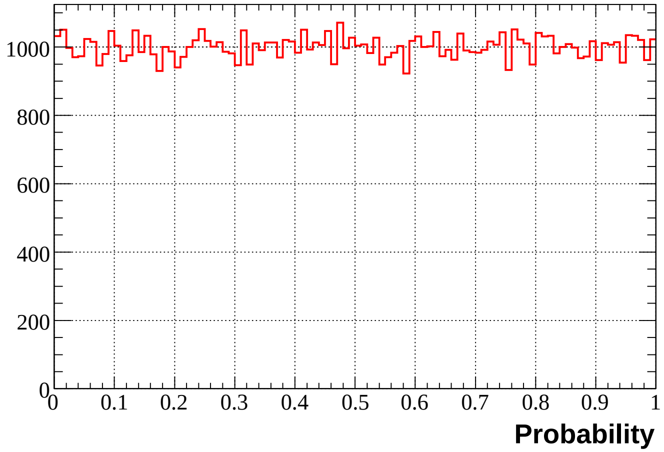
$$E_i^{fit} = E_i - \varepsilon$$

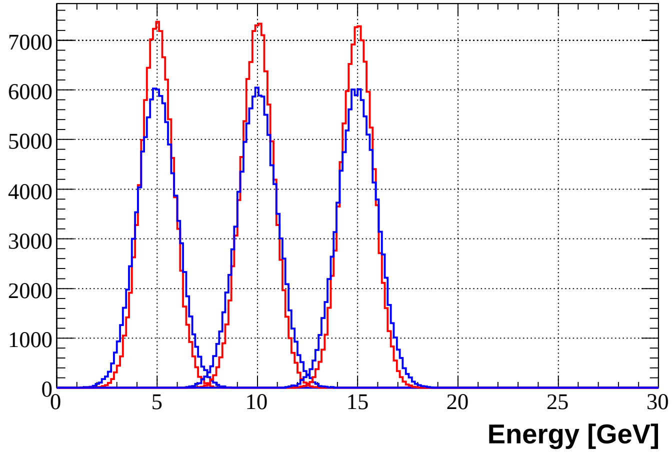
$$\varepsilon = (E_1 + E_2 + E_3 - E_a) / 3$$

Results of example 1

- Chi square distributed as chi square for one degree of freedom
- Probability distribution $\int_{\chi^2}^{\infty} f(z;n)dz$
- A measure of the probability that a chi square from the theoretical distribution is greater than chi square obtained from the fit
- Improvement of accuracies of energy measurements by factor of √(2/3)

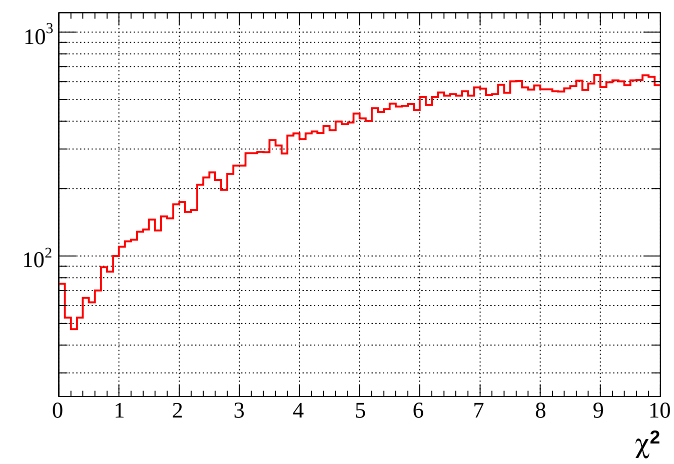


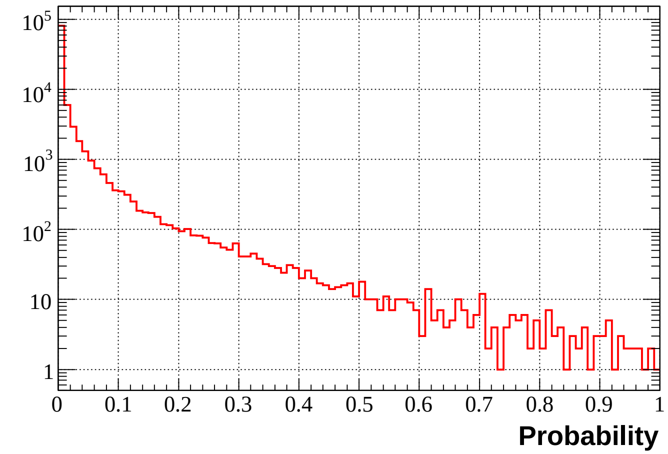


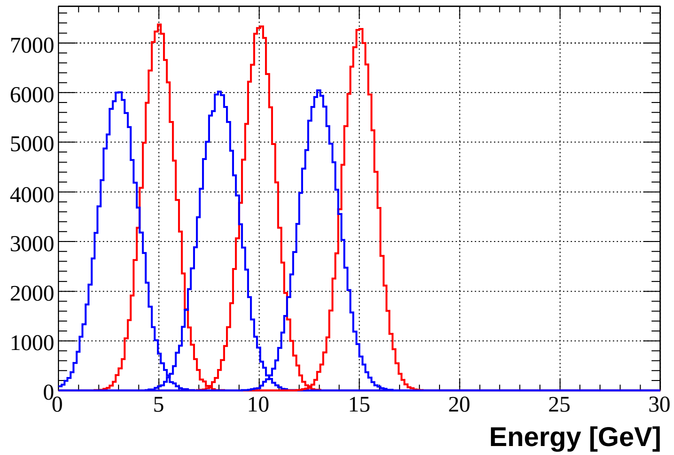


Background process

- Decay of a particle a to 4 particles a → 1+2+3+4
- Energy of the primary particle is precisely known and equal to E_a=30 GeV
- Energies of the secondary particles are equal to
  E_1 = 3 GeV, E_2 = 8 GeV and E_3 = 13 GeV
  and measured with errors distributed by Gaussian with sigma
  σ_1 = 1 GeV, σ_2 = 1 GeV and σ_3 = 1 GeV
- Particles 4 with energy of 6 GeV is unmeasured (missed particle)
- Fitting assuming a → 1+2+3 hypothesis

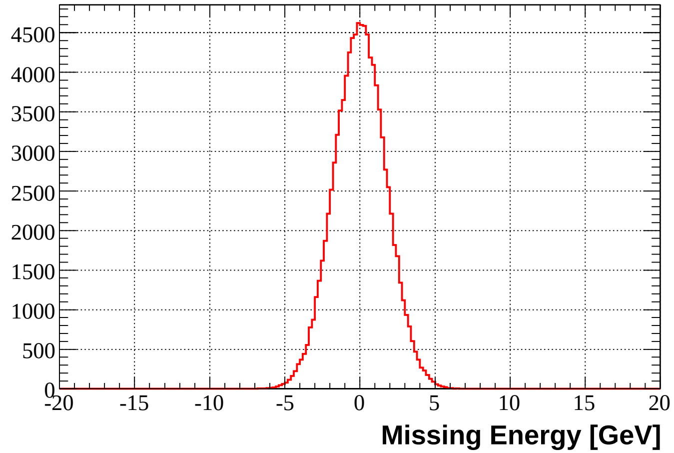


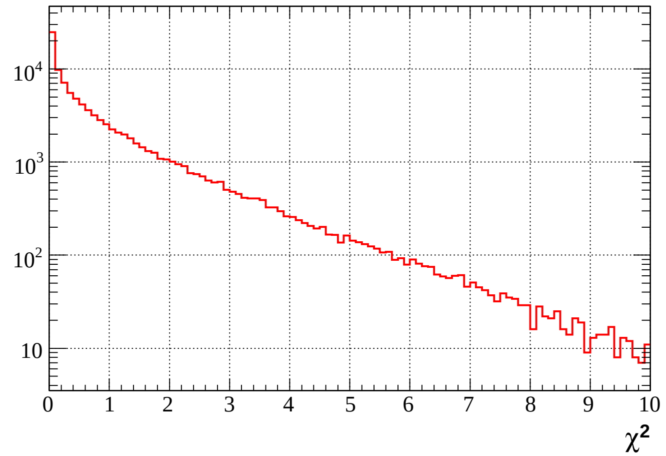




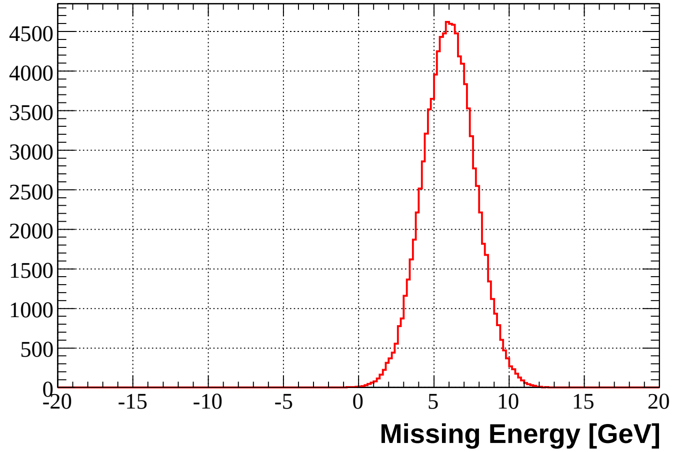
Missing energy method vs kinematic fitting

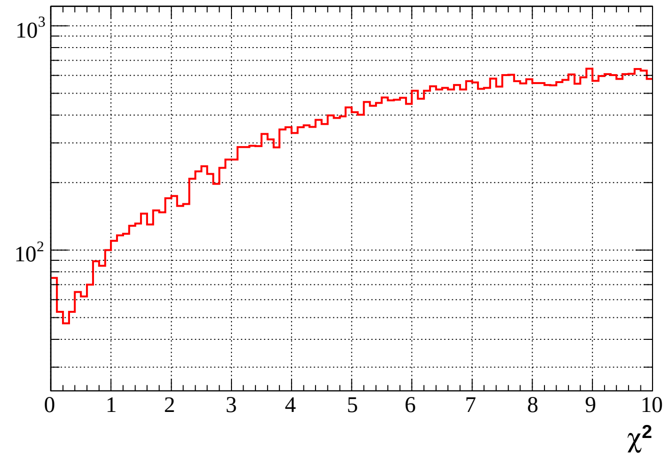
- Process of interest: missing energy is peaked near zero, chi square is distributed as chi square for 1 degree of freedom





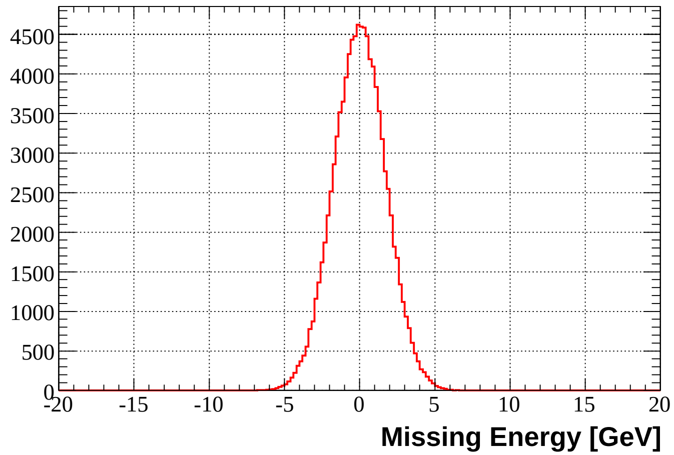
- Background process: missing energy peak is shifted, chi square distribution is completely different

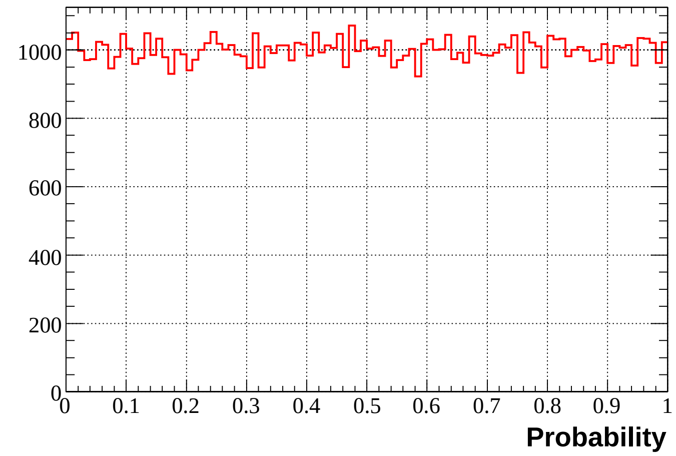


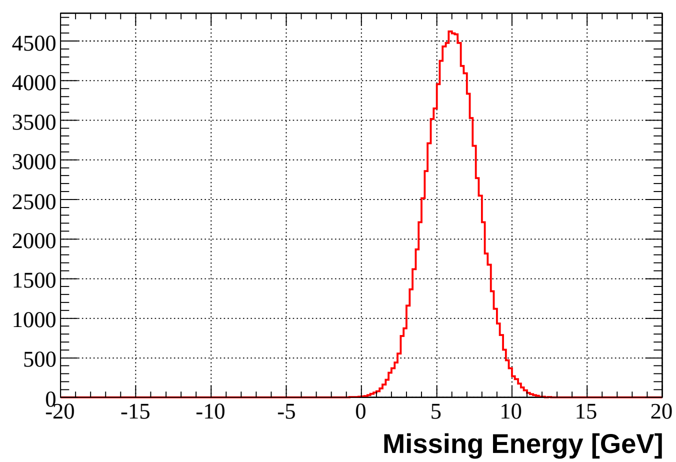


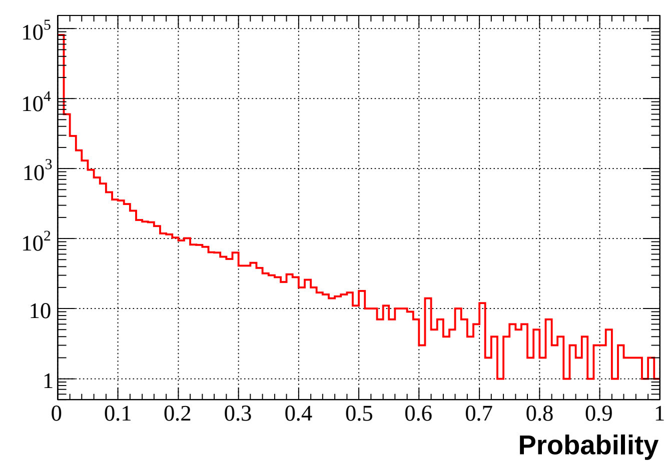
Missing energy method vs kinematic fitting

> Apply a cut on the missing energy or to chi-square or probability distribution to select the process of interest and reject the background









Kinematic fitting technique

- Minimization of least squares functional

$$\chi^2 = \sum_{i=1}^{n} (y_i - \eta_i)^2 / \sigma_i^2$$

under constraints

$$f_1(\eta_1, \eta_2, \ldots, \eta_n) = 0$$

$$f_2(\eta_1, \eta_2, \ldots, \eta_n) = 0$$

$$\ldots \ldots \ldots$$

$$f_m(\eta_1, \eta_2, \ldots, \eta_n) = 0$$

y_i – *measured kinematic parameters,* η_i – *fit parameters,*
σ_i – *measurement errors,* n – number of kinematic parameters,
m – number of constraints

- In case of correlations between kinematic parameters, covariance matrix G_y should be used

$$\chi^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} (y_i - \eta_i) G_{y(i,j)} (y_j - \eta_j)$$

Minimization of least-squares functions with constraints

> Method of elements
  - Using *q* equations of constraints eliminate *q* of the *n* kinematic parameters as functions of *n-q* parameters
  - Not always possible, procedure is not automatic

> Method of Lagrange multipliers
  - Automatic procedure, widely used
  - Exact solution if constraints depend linearly on parameters
  - Simple iterative procedure if constraints are non-linear

> Method of orthogonal transformation
  - Mentioned in one textbook, not widely used

> Method of penalty functions
  - Automatic procedure
  - Can be effectively used with constraints of inequality type

Method of Lagrange multipliers

- Measurements give instead of true quantities η_i values y_i

$$y_i = \eta_i + \varepsilon_i$$

- Measurement errors are normally distributed about zero with standard deviations σ_i

$$E(\varepsilon_i) = 0,$$

$$E(\varepsilon_i^2) = \sigma_i^2$$

- m equations of constraints

$$f_k(\vec{\eta}) = 0, k = 1,2,\ldots m$$

- If equations are linear write in the matrix form

$$B\vec{\eta} + \vec{b}_0 = 0$$

- Define

$$\vec{c} = B\vec{y} + \vec{b}_0$$

- Minimization of Lagrange function

$$L = \vec{\varepsilon}^T G_y \vec{\varepsilon} + 2\vec{\lambda}^T (\vec{c} - B\vec{\varepsilon})$$

- Total derivative should vanish

$$dL = 2\vec{\varepsilon}^T G_y d\vec{\varepsilon} - 2\vec{\lambda}^T B d\vec{\varepsilon} = 0$$

Method of Lagrange multipliers

- Solve system of linear equations

$$\begin{cases} \vec{\varepsilon}^{\,T} G_y - \vec{\lambda}^T B = 0 \\ \vec{c} - B\vec{\varepsilon} = 0 \end{cases}$$

- Obtain

$$\vec{c} - BG_y^{-1}B^T\vec{\lambda} = 0$$

- Can be easily solved

$$\tilde{\vec{\lambda}} = \left(BG_y^{-1}B^T\right)^{-1}\vec{c}$$

- Estimators of the measurement errors

$$\tilde{\vec{\varepsilon}} = G_y^{-1}B^T\left(BG_y^{-1}B^T\right)^{-1}\vec{c}$$

- Estimators for fit parameters

$$\tilde{\vec{\eta}} = \vec{y} - \tilde{\vec{\varepsilon}}$$

- Using abbreviation

$$G_B = \left(BG_y^{-1}B^T\right)^{-1}$$

- Obtain by applying error propagation

$$G_{\tilde{\eta}}^{-1} = G_y^{-1} - G_y^{-1}B^T G_B B G_y^{-1}$$

Solution of example 1

> In the example 1 $B = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$

> Vector $\vec{c} = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} E_1 \\ E_2 \\ E_3 \end{pmatrix} - E_a = E_1 + E_2 + E_3 - E_a$

$$G_B = \left(\begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right)^{-1} = \frac{1}{3}$$

> Estimation of fit parameters

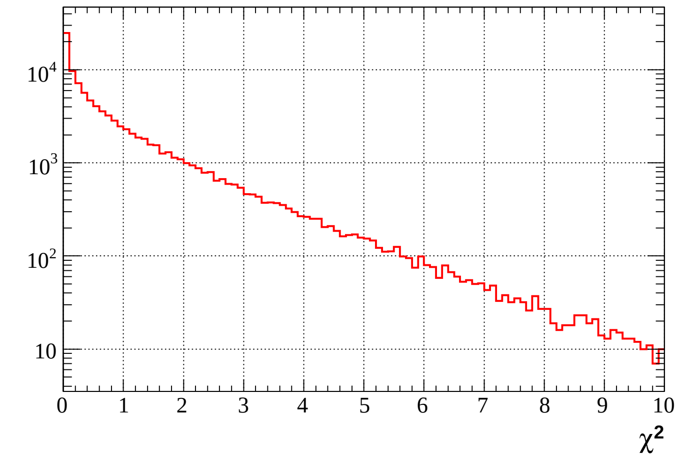
$$\tilde{\eta}_i = E_i - \frac{1}{3}(E_1 + E_2 + E_3 - E_a)$$

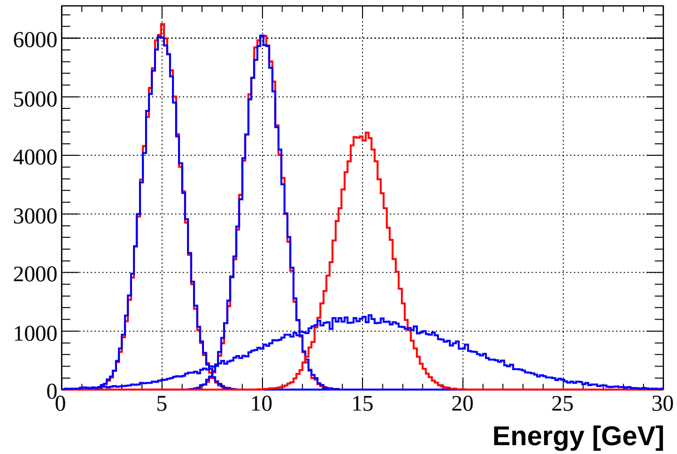
> Estimation of error matrix of fit parameters

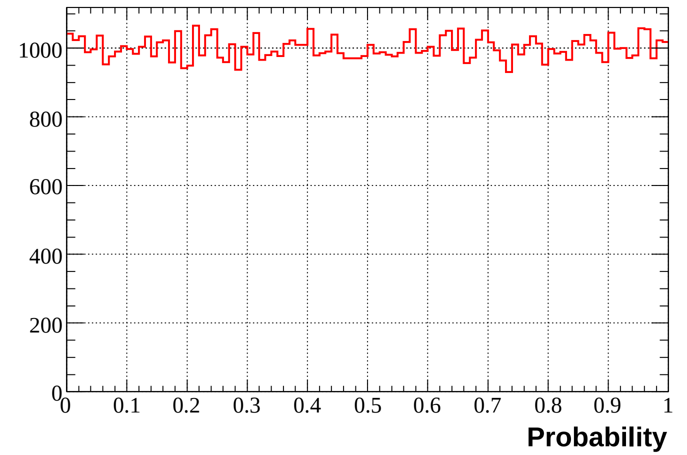
$$G_{\tilde{\eta}}^{-1} = \frac{1}{3} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$$

Example 2

- Decay of a particle a to 3 particles a → 1+2+3
- Energy of the primary particle is precisely known and equal to $E_a = 30$ GeV
- Energies of the secondary particles are equal to $E_1 = 5$ GeV, $E_2 = 10$ GeV and $E_3 = 15$ GeV and measured with errors distributed by Gaussian with sigma $\sigma_1 = 1$ GeV, $\sigma_2 = 1$ GeV and $\sigma_3 = 5$ GeV







$$E_i^{fit} = E_i - \varepsilon_i$$

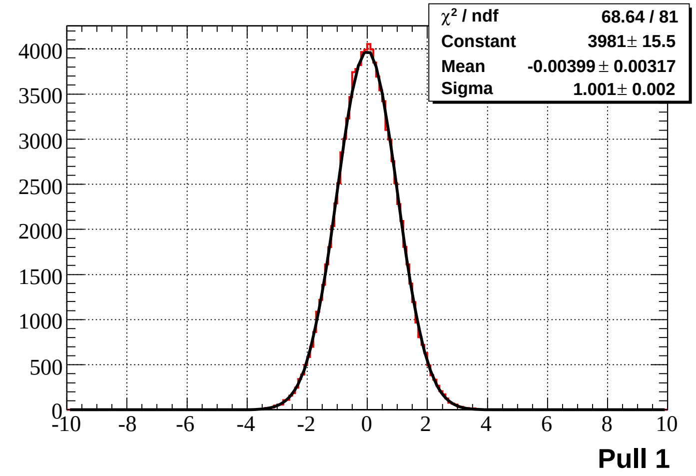
$$\varepsilon_i = (E_1 + E_2 + E_3 - E_a) \cdot \sigma_i^2 / (\sigma_1^2 + \sigma_2^2 + \sigma_3^2)$$

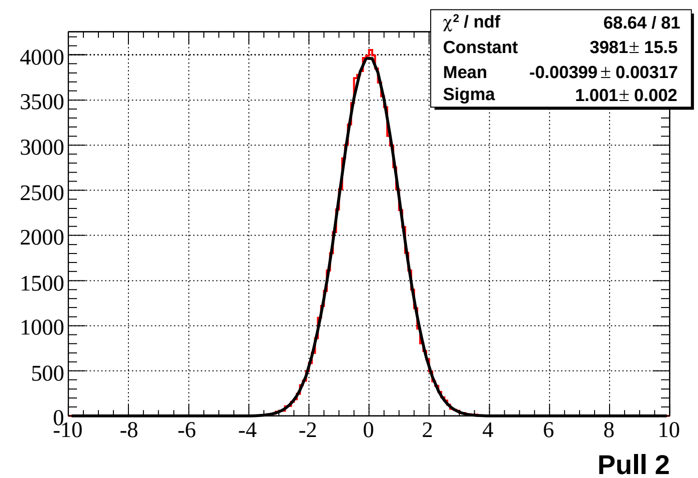
Results of example 2

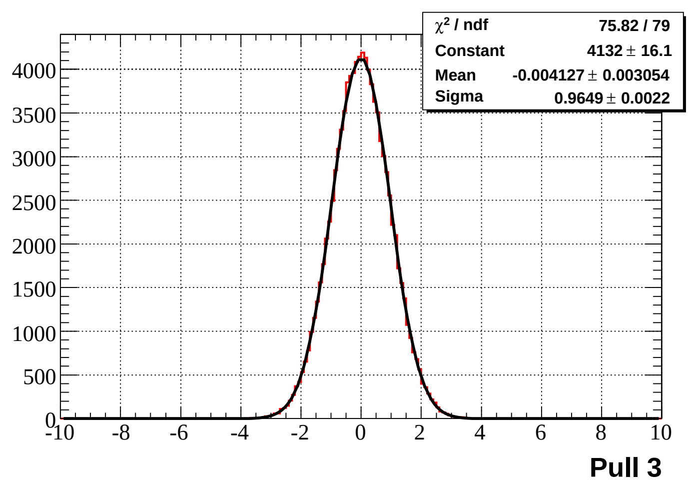
- Pull distributions defined as

$$pull_i = \frac{\varepsilon_i}{\sigma_i(\varepsilon_i)} = \frac{y_i - \eta_i}{\sqrt{\sigma^2(y_i) - \sigma^2(\eta_i)}}$$

- Should be normally distributed around zero with standard deviation equal to unity
- Help to understand if measurement errors are known correctly
- Check for possible systematic effects

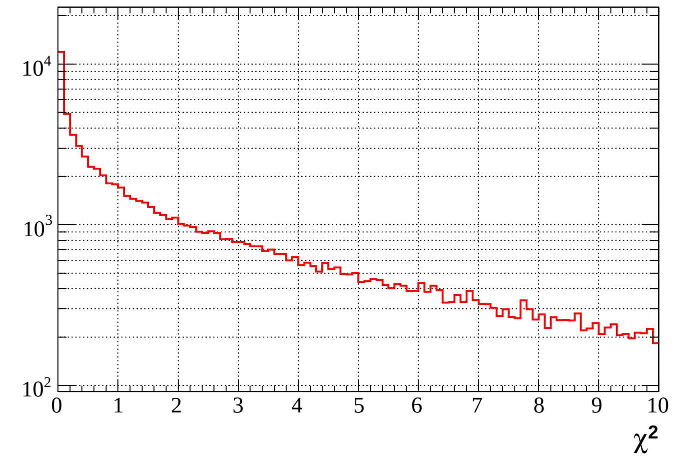


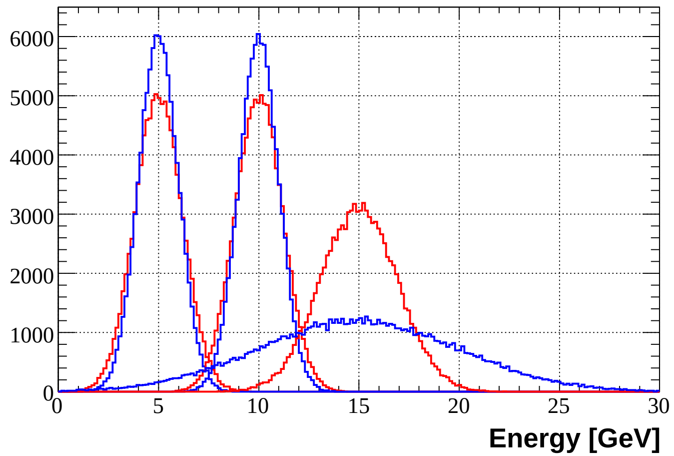


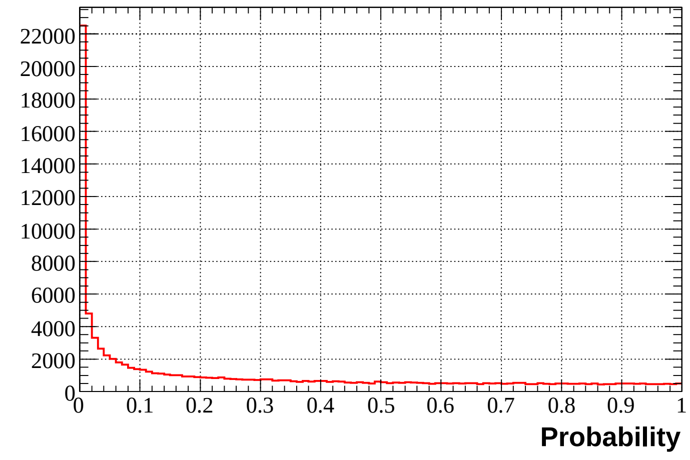


Example 3

- Decay of a particle 3 particles a → 1+2+3
- Energy of the primary particle is precisely known and equal to E_a=30 GeV
- Energies of the secondary particles are equal to E_1 = 5 GeV, E_2 = 10 GeV and E_3 = 15 GeV and measured with errors distributed by Gaussian with sigma σ_1 = 1 GeV, σ_2 = 1 GeV and σ_3 = 5 GeV
- Measurement of error of E_3 is not known and assumed to be 2 GeV

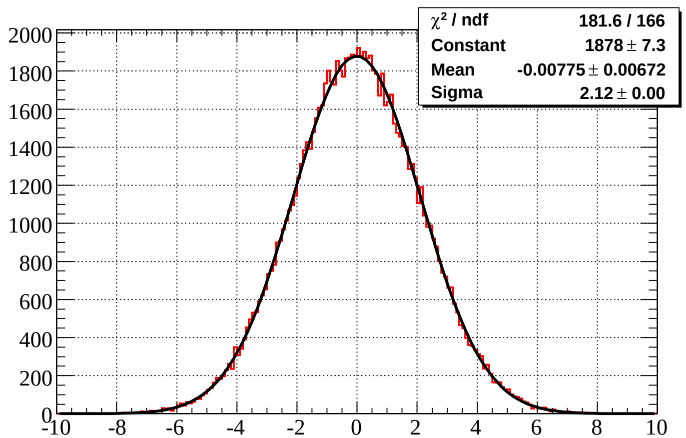




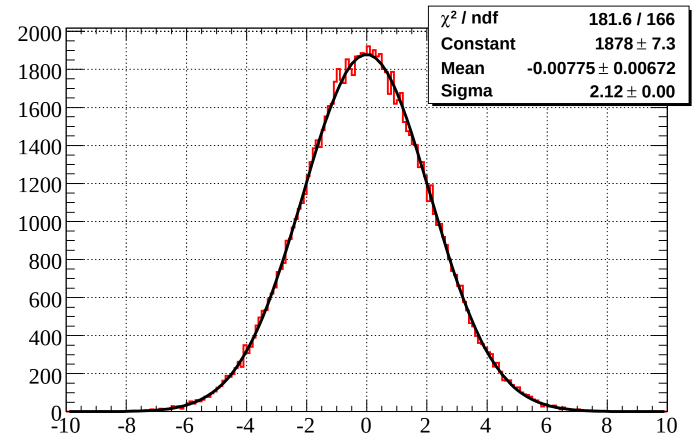


Results of example 3

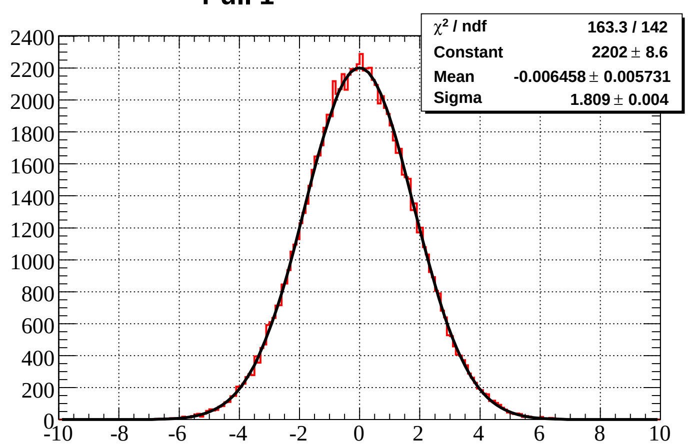
> Pull distributions are wider, indication that something is wrong with the knowledge of measurement errors



Pull 1



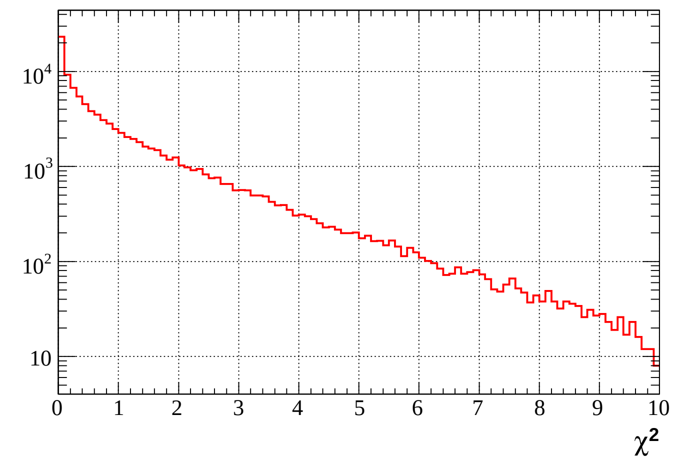
Pull 2

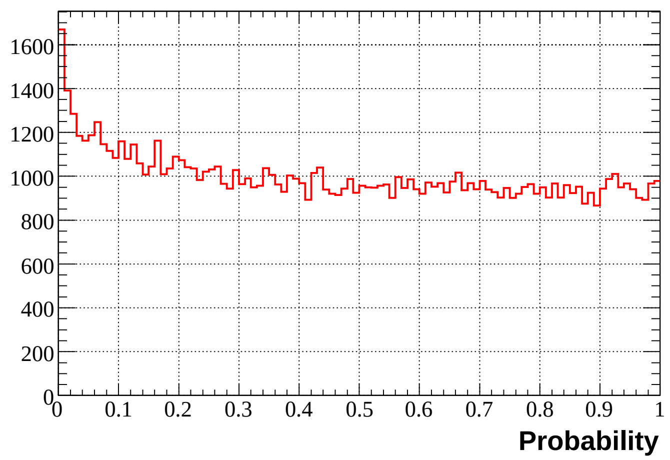


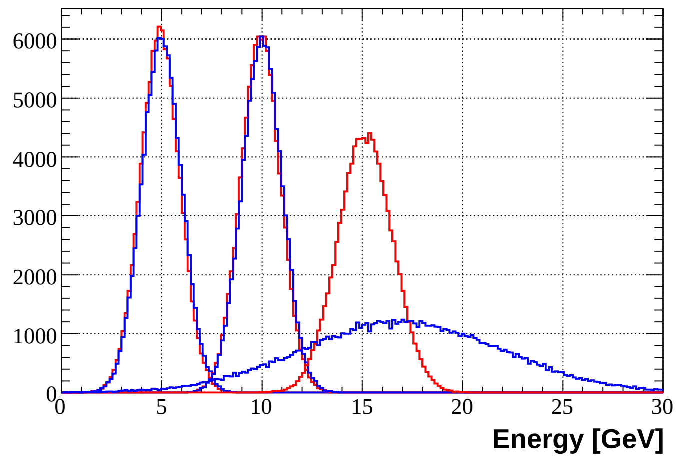
Pull 3

Example 4

- > Decay of a particle 3 particles $a \rightarrow 1+2+3$
- > Energy of the primary particle is precisely known and equal to E_a=30 GeV
- > Energies of the secondary particles are equal to E_1 = 5 GeV, E_2 = 10 GeV and E_3 = 15 GeV and measured with errors distributed by Gaussian with sigma σ_1 = 1 GeV, σ_2 = 1 GeV and σ_3 = 5 GeV
- > Bias in the measurement of E_3: 17 GeV instead of 15 GeV

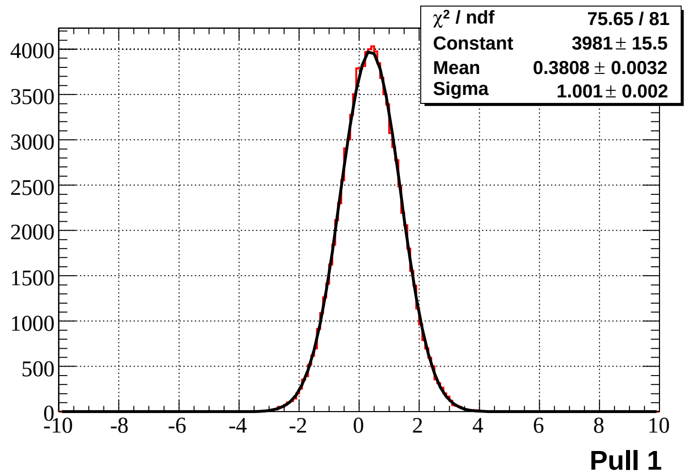


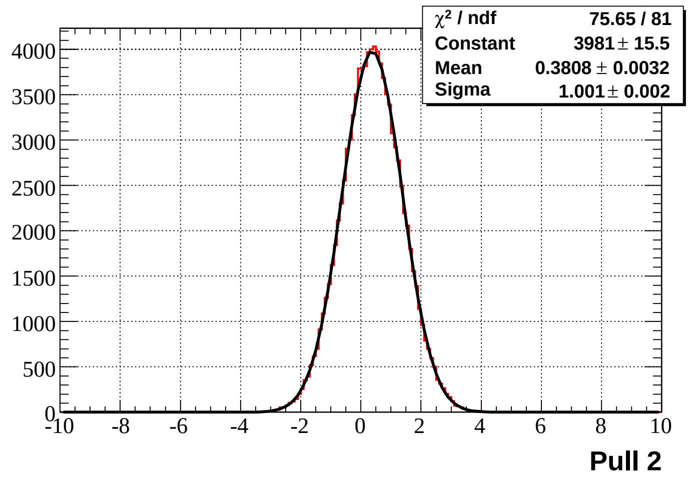


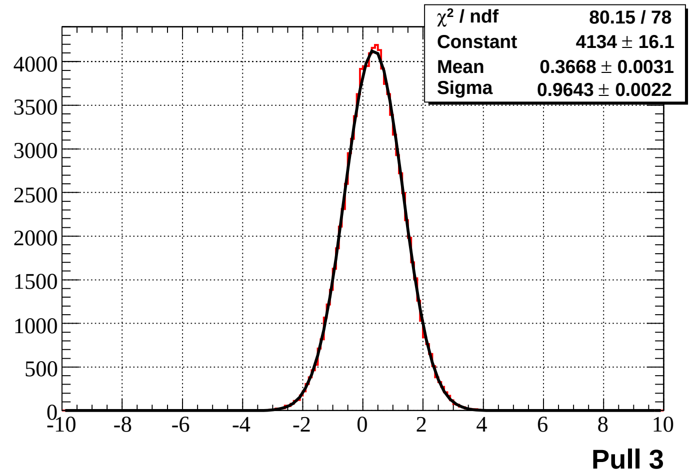


Results of example 4

> Pull distributions are shifted







Applications of kinematic fitting

- Event selection
- Background rejection
- Improvement of resolution
- Better understanding of measurement errors, search for possible sources of systematic uncertainties
  - Analysis of chi square (probability) and pull distributions
- Detector calibration
  - If number of constraints is enough, some measurements can be considered as unknown parameters, reconstructed by kinematic fitting and then used for calibration

Possible problems

> Measurement errors are not known precisely

> Systematic uncertainties

> Measurement errors are not Gaussian

> If different particles are measured in different coordinate systems – misalignment between these systems

> Multiple scattering and radiative effects usually lead to tails in chi-square and pull distributions

Outline

> Introduction

> Techniques

> Applications

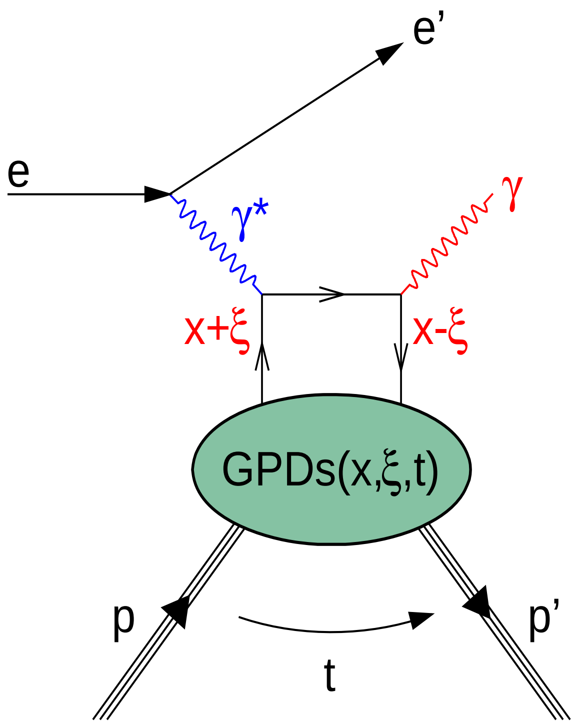
> **Results of use at HERMES**

> Literature

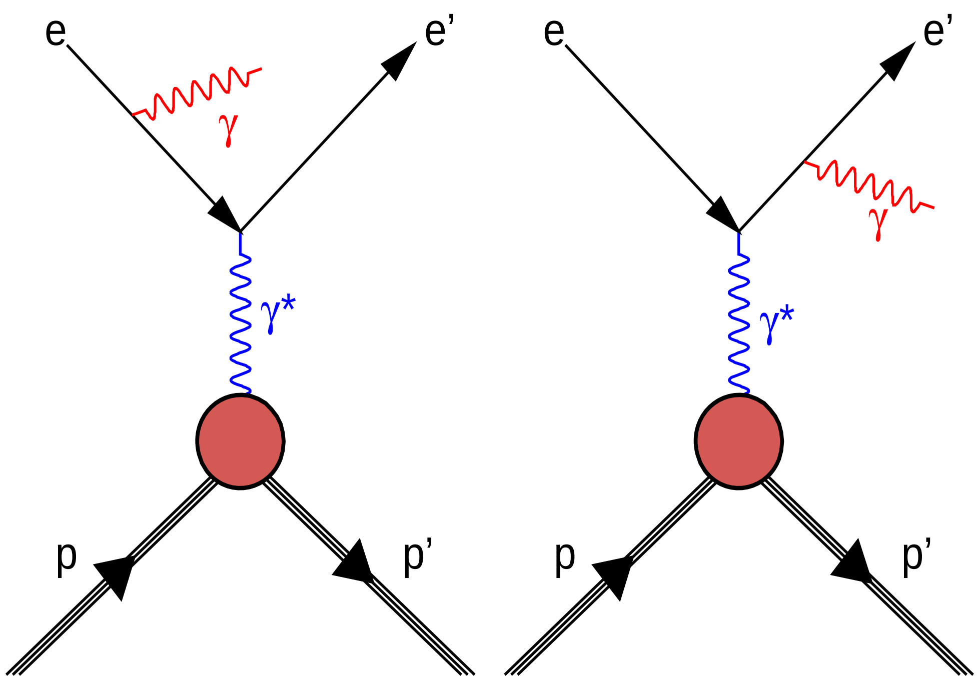
> Summary

Deeply virtual Compton scattering (DVCS) at HERMES

DVCS

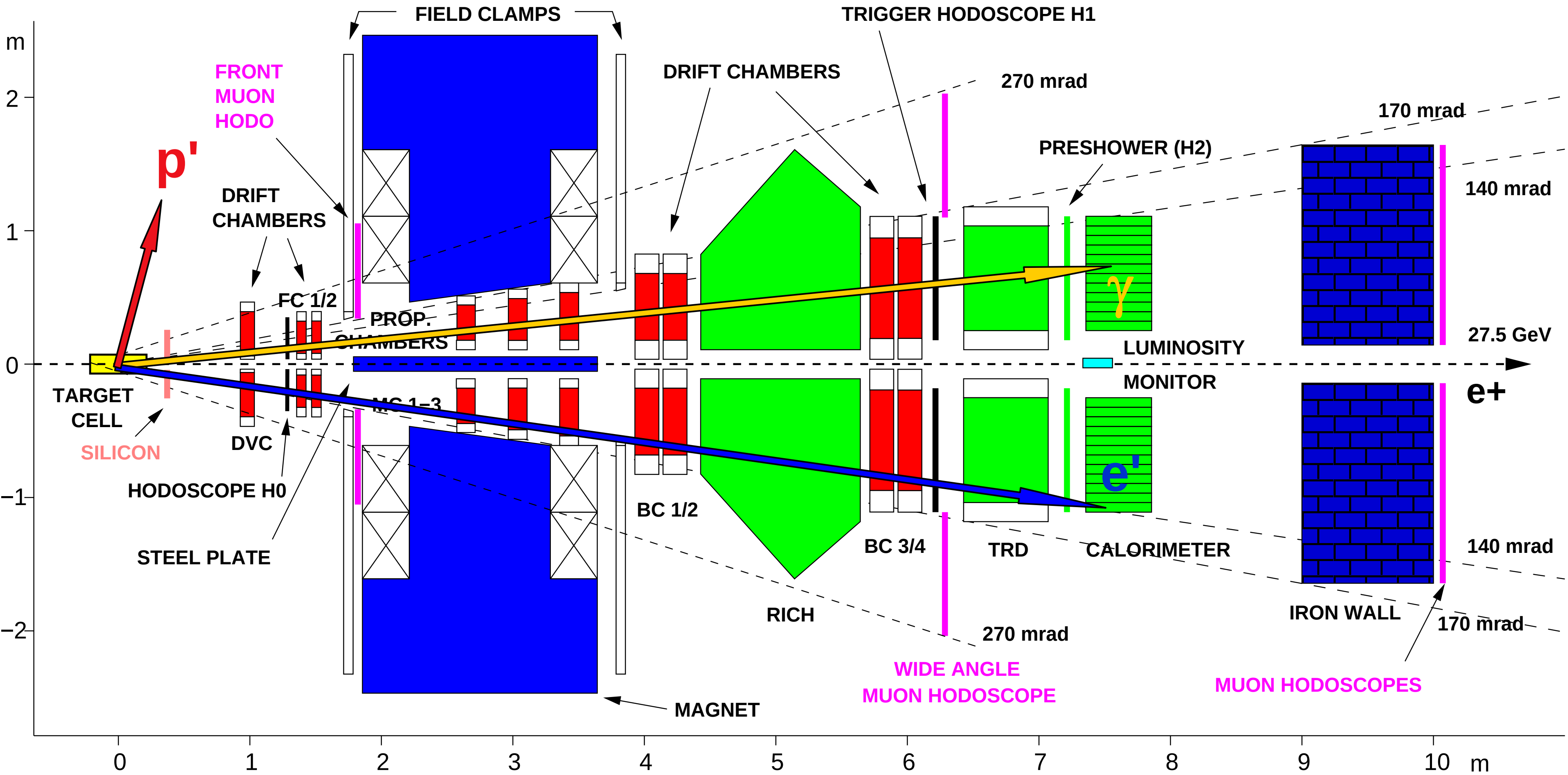


Bethe-Heitler



- DVCS and Bethe-Heitler: the same initial and final state
- Bethe-Heitler dominates at HERMES kinematics
- Generalized Parton Distributions (GPDs) accessible through cross section differences and azimuthal asymmetries via interference term

The HERMES experiment



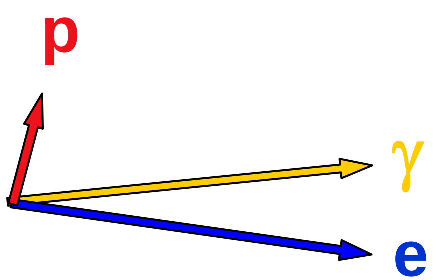
Gas targets:

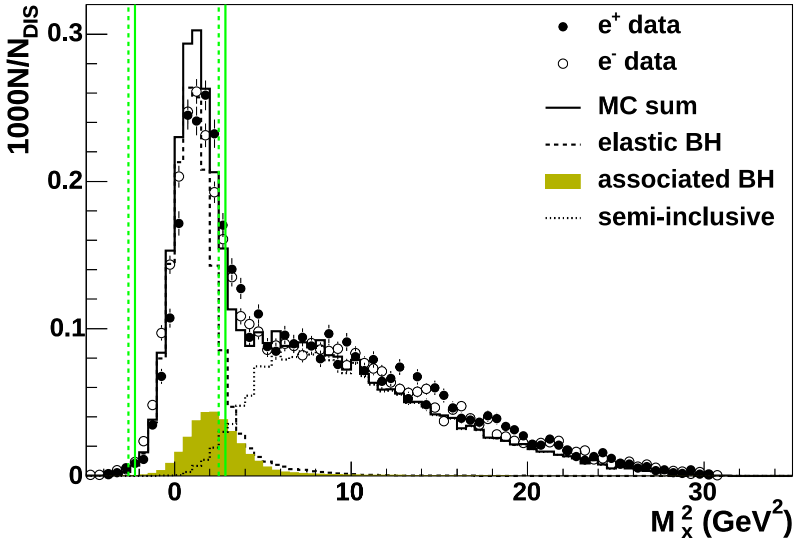
- Longitudinally polarized H, D
- Unpolarized H, D, ^{4}He, N, Ne, Kr, Xe
- Transversely polarized H

Beam:

- Longitudinally polarized e^+ and e^- with both helicities
- Energy 27.6 GeV

DVCS/BH measurement with the Recoil Detector



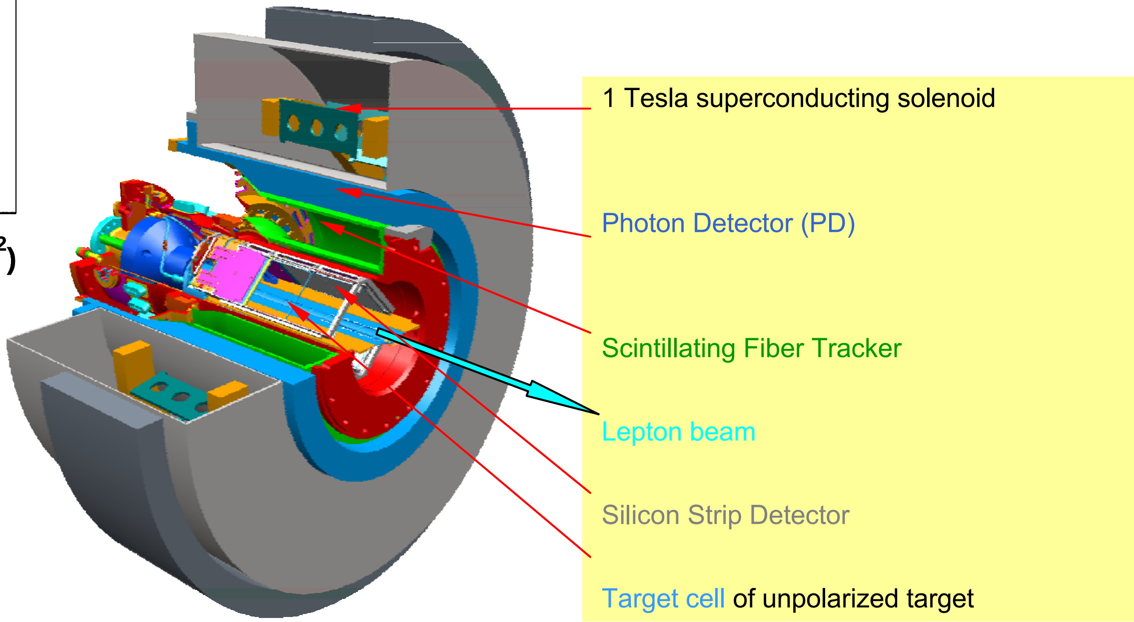


> Pre-Recoil data

- Scattered lepton and photon were detected in the forward spectrometer
- Recoil proton was not detected
- Exclusivity achieved via missing mass technique
- Associated processes (ep→eΔγ) were not resolved (12% contribution in the signal)

> Recoil data

- Detection of recoil proton
- Suppression of the background to <1% level



Elastic DVCS/BH and associated background

- Elastic DVCS/BH: $ep \rightarrow ep\gamma$
  - Beam 3-momentum is known precisely (in comparison with 3-momenta of secondary particles)
  - Target proton is at rest
  - For *e, p,* and *γ,* 3-momenta are measured with certain precision

- Background from associated process $ep \rightarrow e\Delta^+\gamma$ with $\Delta^+ \rightarrow p\pi^0$

- Not possible to separate by missing mass only

- Low efficiency of selection if one-dimensional cuts (coplanarity, momentum-momentum correlations) are applied

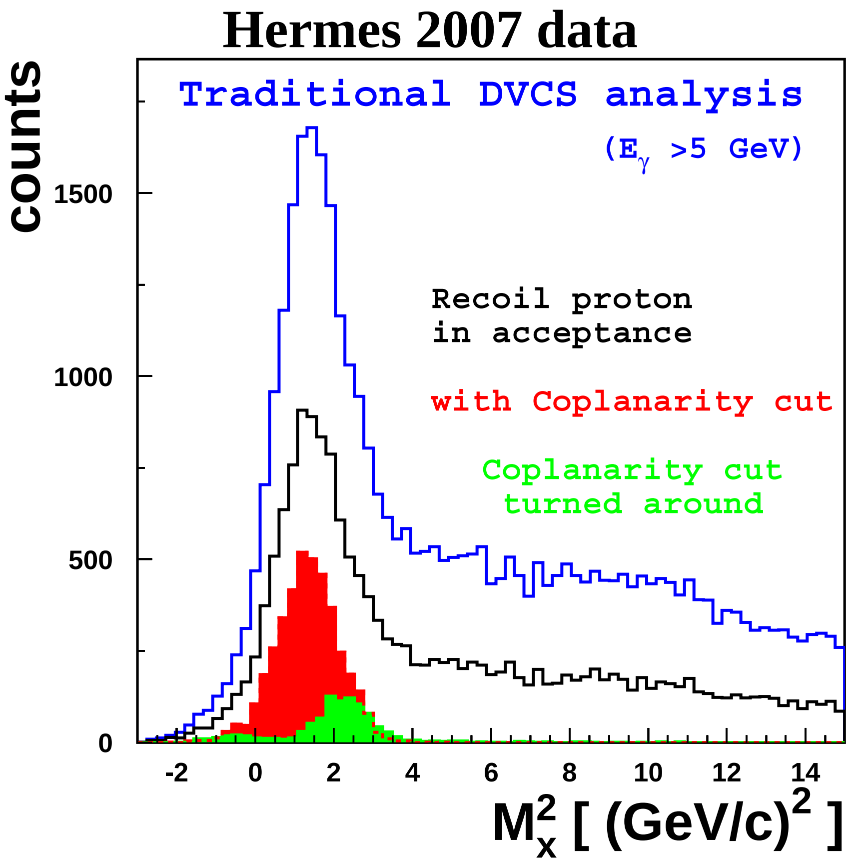
Selection of elastic DVCS/BH using simple cuts

> One-dimensional cuts
  - Momentum-momentum correlations
  - Coplanarity condition

> Efficiency of 70% at background contamination of 5%

> background contamination of 2% can be achieved but efficiency drops below 50%

> No improvement of resolutions



Kinematic fitting for elastic DVCS/BH

> **Nine** kinematic parameters:

Scattered electron	Photon	Proton
$y_1 = \tan(p_{x1} / p_{z1})$	$y_4 = \tan(p_{x2} / p_{z2})$	$y_7 = \varphi_3$
$y_2 = \tan(p_{y1} / p_{z1})$	$y_5 = \tan(p_{y2} / p_{z2})$	$y_8 = \theta_3$
$y_3 = 1/ p_1$	$y_5 = e_2$	$y_9 = 1/(p_3 \sin\theta_3)$

> **Four** equations of constraints

$$f_1 = p_{x1} + p_{x2} + p_{x3} = 0$$

$$f_2 = p_{y1} + p_{y2} + p_{y3} = 0$$

$$f_3 = p_{z1} + p_{z2} + p_{z3} - p_{beam} = 0$$

$$f_4 = e_1 + e_2 + e_3 - e_{beam} - m_p = 0$$

Constraints as functions of parameters

> Constraints are non-linear

$$f_1 = y_1 / y_3 / \sqrt{1 + y_1^2 + y_2^2} + y_4 \cdot y_6 / \sqrt{1 + y_4^2 + y_5^2} + \cos(y_7) / y_9$$

$$f_2 = y_2 / y_3 / \sqrt{1 + y_1^2 + y_2^2} + y_5 \cdot y_6 / \sqrt{1 + y_4^2 + y_5^2} + \sin(y_7) / y_9$$

$$f_3 = 1 / y_3 / \sqrt{1 + y_1^2 + y_2^2} + y_6 / \sqrt{1 + y_4^2 + y_5^2} + 1/(y_9 \tan(y_8)) - p_{beam}$$

$$f_4 = \sqrt{1 / y_3^2 + m_e^2} + y_6 + \sqrt{1/(y_9^2 \cdot \sin^2(y_8)) + m_p} - e_{beam} - m_p$$

> Derivatives of constraints can be calculated analytically or numerically

Fitting procedure

> Minimization of chi square with constraints using penalty term

$$\chi^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} (y_i - \eta_i) G_{y(i,j)} (y_j - \eta_j) + T \cdot \sum_{k=0}^{m} f_k^2 / \sigma_{ck}^2$$

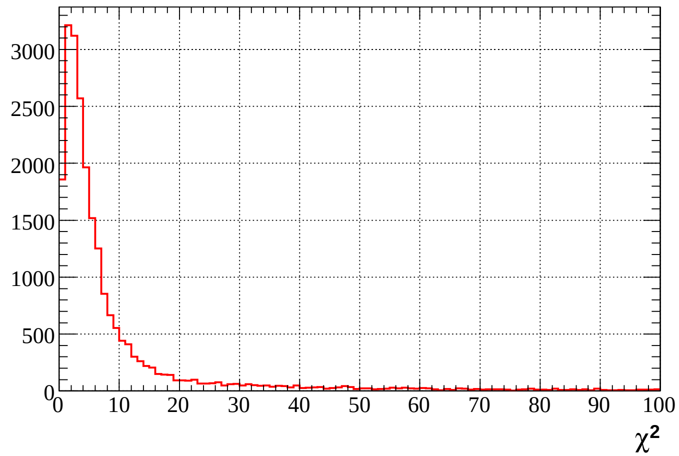
> If *T* is large enough constraints are satisfied automatically during the fitting procedure

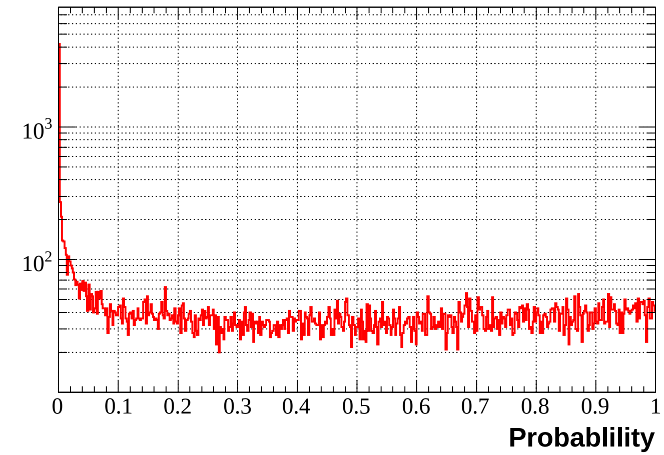
> In order to equalize penalty contributions from different constraints errors of constraints are formally calculated using error propagation

> Non-linear constraints of inequality type can be included

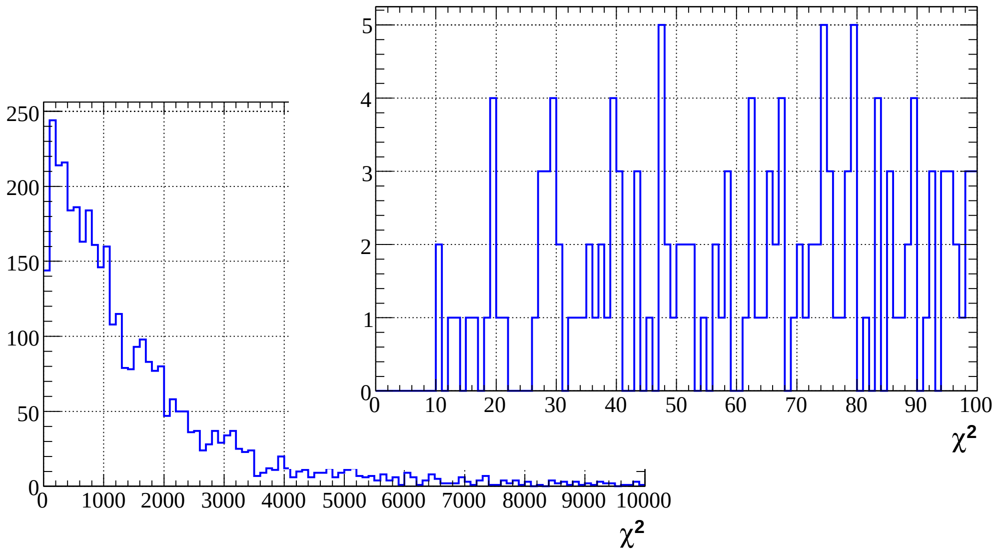
Results for elastic DVCS/BH and associated background

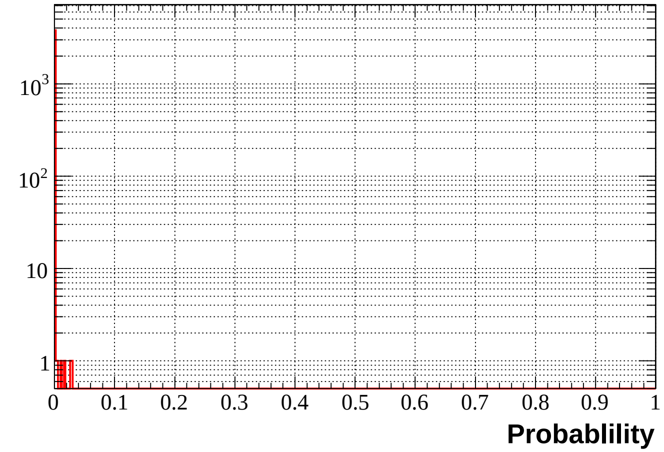
> For elastic DVCS/BH (Monte Carlo)



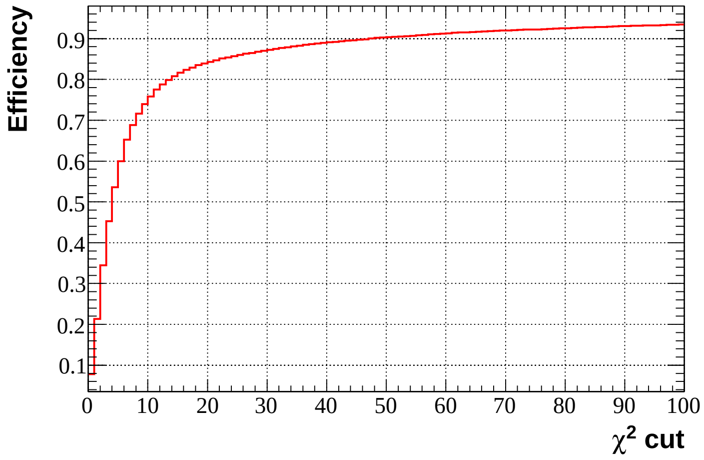


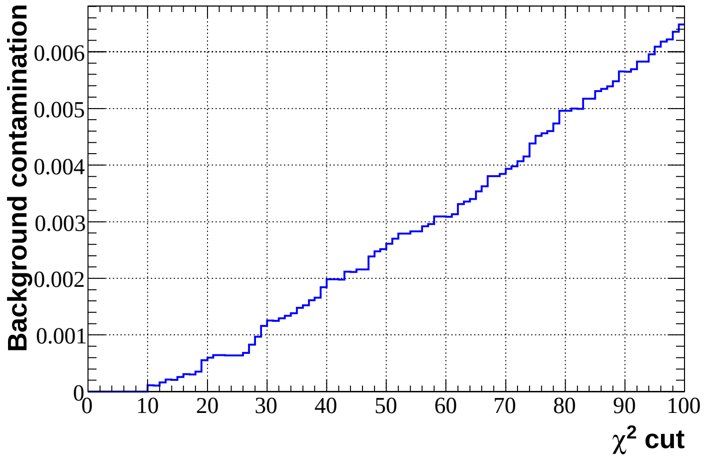
> For associated background





Efficiency of elastic event selection





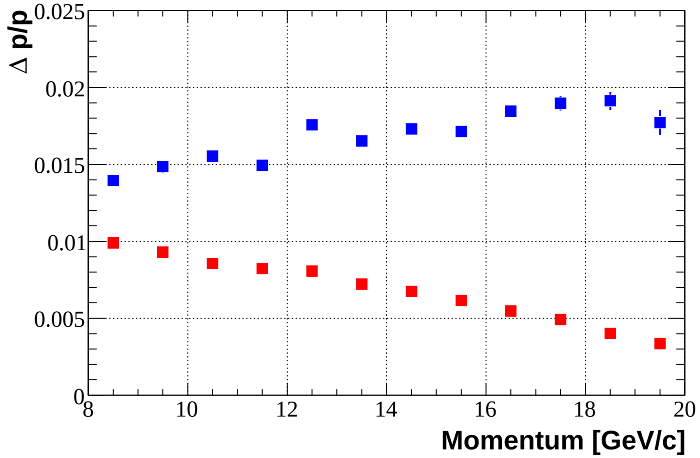
> Efficiency of event selection of elastic DVCS/BH events is high at a reasonable cut on chi-square

> Contamination of associated process is well below 1%

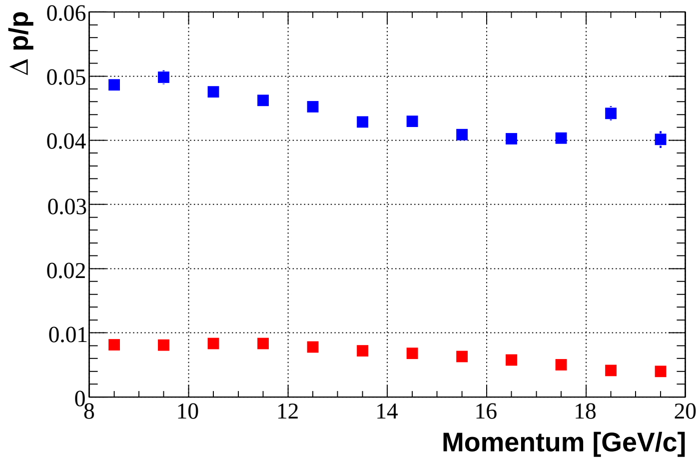
> In addition, improvement of resolutions

Improvement of kinematic parameter reconstruction

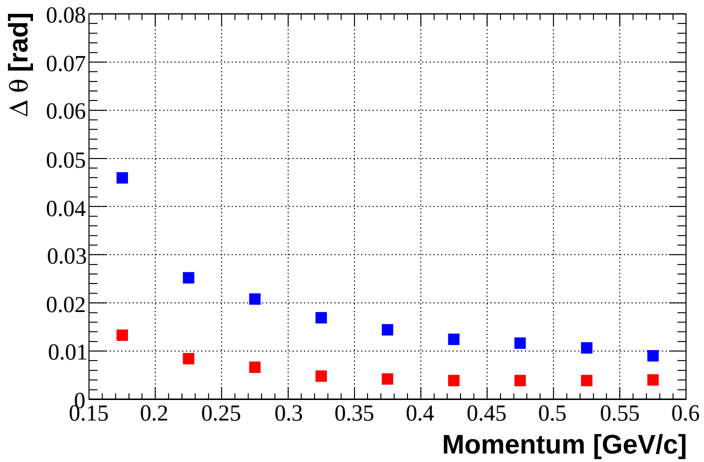
Momentum of electron



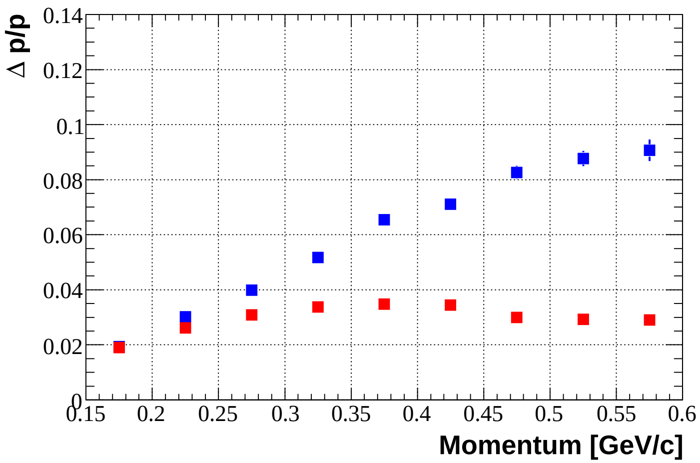
Energy of photon



Polar angle of proton

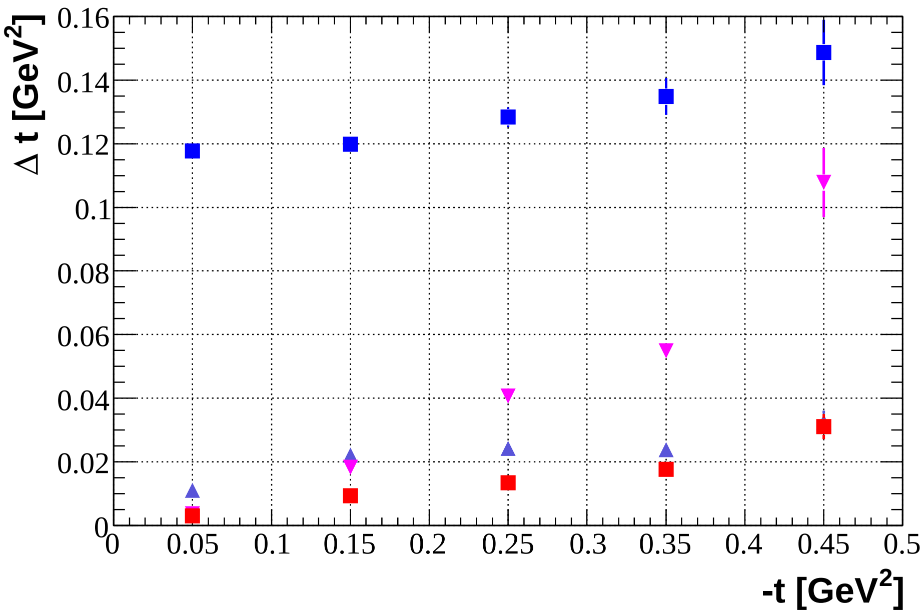


Momentum of proton



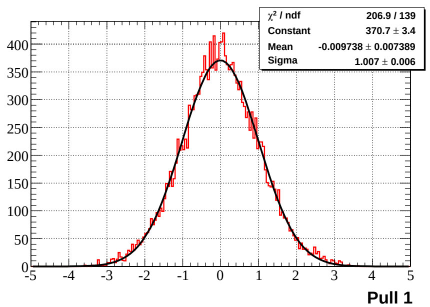
t resolution

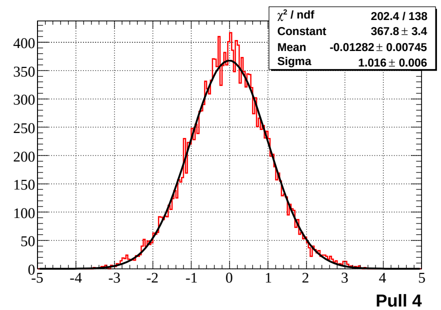
- Mandelstam $t=(p-p')^2$, important physical observable
- Blue squares – reconstruction using measured kinematic parameters of electron and photon
- Magenta triangles - reconstruction using measured kinematic parameters of recoil proton
- Light blue triangles - reconstruction using measured kinematic parameters of electron and photon (excluding photon energy) assuming proton mass
- Red squares – reconstruction using kinematic fitting

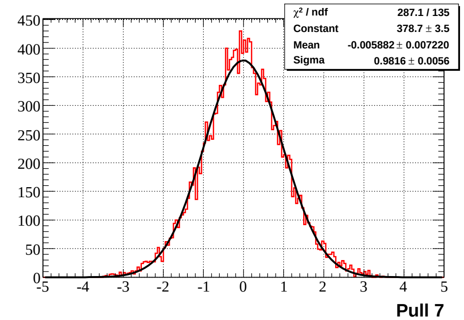


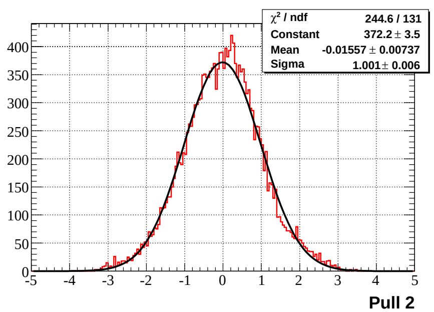
Pull distributions

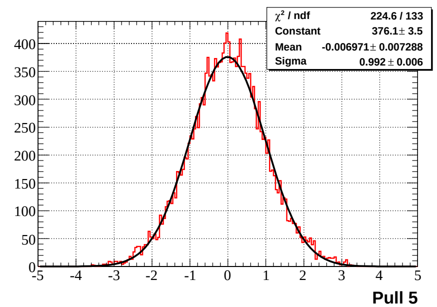
Electron | Photon | Proton

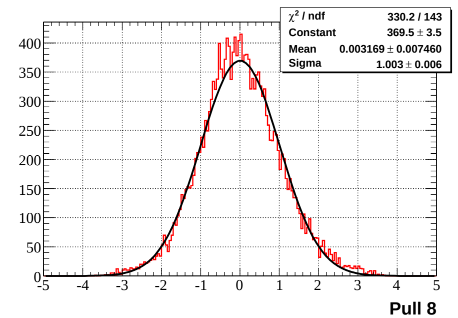


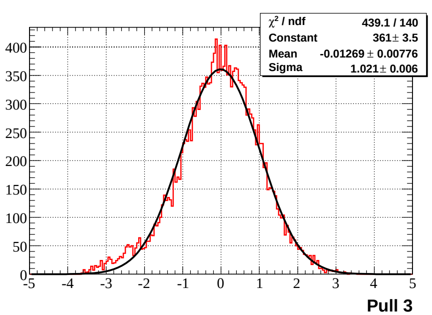


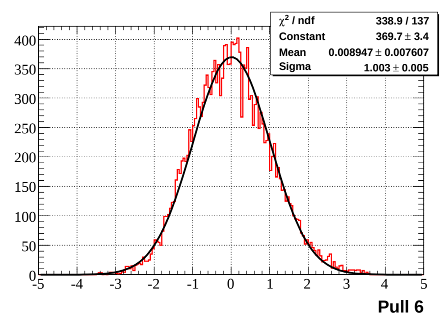


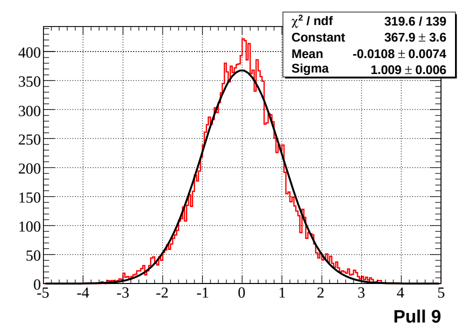












Selection of elastic DVCS/BH events (HERMES data)

> Kinematic fitting is developed and tested on Monte-Carlo

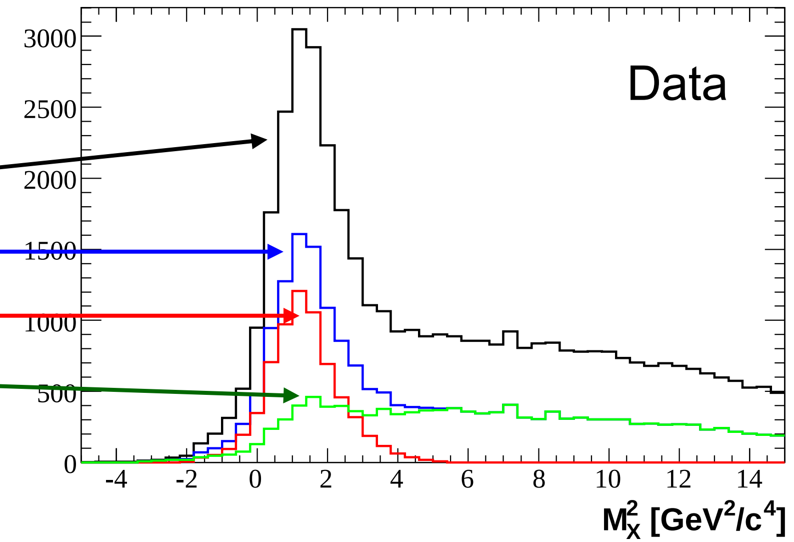
- 3 particles detected → 4 constraints from energy-momentum conservation
- Allows to suppress the associated processes and semi-inclusive background to negligible level

> Applied for data for physics analysis

- Systematic studies in progress
- First physics results expected soon

> Missing mass distribution

- No requirement for Recoil
- Positively charged Recoil track
- Kinematic fit probability > 1%
- Kinematic fit probability < 1%



Summary

> Kinematic fitting is a powerful tool of event selection and reconstruction

> Many applications
  - Improvement of resolution
  - Better understanding of measurement errors, search for possible sources of systematic uncertainties
  - Detector calibration

> Kinematic fitting technique is well developed and described in literature

> Different problems could appear in different experiments, solutions are not always straightforward

Literature (personal choice)

- W.E.Deming, Statistical Adjustment of Data
- S.Brandt, Data Analysis
- R.Frühwirth, M.Regler, R.K.Bock, H. Grote, D. Notz, Data Analysis Techniques for High-Energy Physics
- J.P.Berge, F.T.Solmitz, H.D.Taft, Rev. Scient. Instr. 32 N5 (1961) 538
- R.Boeck CERN 60-30
- W.G.Davis, Nucl. Instr. Meth A 386 (1997) 498
- S.N.Dymov et al, Nucl. Instr. Meth A 440 (2000) 431
- M.Williams, C.A.Meyer, CLAS-NOTE 2003-017
- P.Avery, http://www.phys.ufl.edu/~avery
- V.Bloebel, http://www.desy.de/~blobel/