Kinematic Fitting Technique—KinFit

V. Kurbatov[1†], V. Tokareva[2], D. Tsirkov[1]

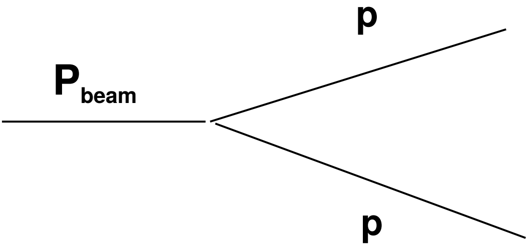
[1]Laboratory of Nuclear Problems, JINR, Dubna, Russia
[2]Institut für Kernphysik, Karlsruher Institut für Technologie, Karlsruhe, Germany

[†]kurbatov@jinr.ru

August 16, 2019

Introduction to KinFit

What is it needed for?



At least two hypotheses:

- $pp \to pp$
- $pp \to pp\pi^0$

Conservation laws should be used to select the correct hypothesis!

Problem formulation

Find kinematical parameters X_i that turn χ^2 to the minimum

$$\chi^2 = \sum_{k=1,j=1}^{n_p,n_p} (X_i - X_i^m) Z_{i,j} (X_j - X_j^m)$$

and satisfy the conservation law equations (constraints)

$$f_\lambda(\boldsymbol{X}) = 0; \quad \lambda = 1 \dots n_c.$$

$\boldsymbol{X}$ vector of kinematical parameters X_i;

n_p their number;

$Z_{i,j}$ inverse error matrix;

X_j^m measured values of parameters;

n_c number of conservation law equations.

History

$$\chi^2 = \sum_{k=1,j=1}^{n_p,n_p} (X_i - X_i^m) Z_{i,j} (X_j - X_j^m) \tag{1}$$

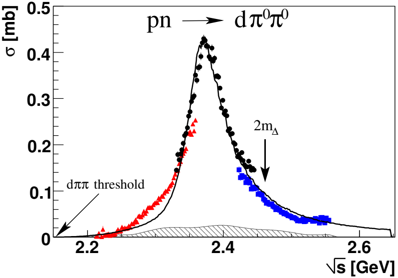
$$f_\lambda(\boldsymbol{X}) = 0; \quad \lambda = 1 \ldots n_c \tag{2}$$

[1] J.P. Berge, F.T. Solmitz and H.D. Taft, Rev. Sci. Instr. **32** 538 (1961);

[2] R. Bock, CERN 60-30 (1960).

The authors have shown that if X_j^m are distributed according to Gaussian and the hypothesis is true, then (1) has a χ^2 distribution. Its number of degrees of freedom (ndf) is equal to n_c after substituting X_j with the values that turn (1) to the minimum and satisfy (2).

Example of an application: WASA discovery



Cross section of $pn \to d\pi^0\pi^0$ as a function of pn mass

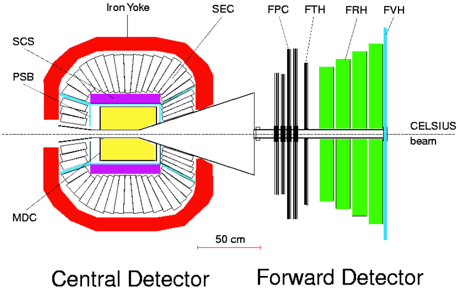
An example of the employment of a similar technique is the WASA observation of the resonance-like cross section behavior for the $pd \to d\pi^0\pi^0 p_{\text{sp}}$ reaction.

The authors identified the following chain of the processes:

$$p + d \to d + \pi^0 + \pi^0 + p_{\text{sp}} \to d + 2\gamma + 2\gamma + p_{\text{sp}}.$$

There were 12 equations, and $\boldsymbol{p}_{\text{sp}}$, $\boldsymbol{p}_{\pi_1^0}$, $\boldsymbol{p}_{\pi_2^0}$ were found with KinFit.

WASA setup while in CELSIUS



WASA setup while in CELSIUS.

- **MDC** The Mini-Drift Chamber
- **SCS** Superconducting Solenoid
- **PSB** Plastic Scintillator Barrel
- **SEC** CsI (Na) Electromagnetic Calorimeter
- **FPC** Proportional Counter straw chamber (tracker)
- **FTH** The Forward Trigger Hodoscope
- **FRH** The Forward Range Hodoscope
- **FVH** The Forward Veto Hodoscope

Realization and Restrictions of proposed method

WASA used the method of Lagrange multipliers:

$$\chi^2 = \sum_{k=1,j=1}^{n_p,n_p} (X_i - X_i^m) Z_{i,j} (X_j - X_j^m) + 2 \sum_{\lambda=1}^{n_c} \alpha_\lambda f_\lambda(\boldsymbol{X}) \qquad (3)$$

Here α_λ are arbitrary multipliers to be found during minimization; both X_i and α_λ are varied.

Some shortcomings of the method:

- Measurement errors are assumed to have Gaussian distribution;
- In (3) the kinematical parameters themselves are used. In the experiment we obtain a number of primary observables like hit coordinates. Often we have limited knowledge on their errors, and they may be far from Gaussian;
- Thus, for applying the technique (3) one should somehow find the matrix $Z_{i,j}$;
- In other words we have the problem of error propagation.

KinFit in JINR in 1960-s

- FUMILI by S.N. Sokolov and I.N. Silin;
- Penalty function method by V.I. Moroz [V.I. Moroz, JINR, P-1958 (1965)].

$$\chi^2 = \sum_{k=1,j=1}^{n_p,n_p} (X_i - X_i^m) Z_{i,j} (X_j - X_j^m) + T \sum_{\lambda=1}^{n_c} (f_\lambda / \Delta(f_\lambda))^2$$

T: large number;

$\Delta(f_\lambda)$: "error" of the constraint.

The idea is if $T \to \infty$ the parameter estimates approach the true ones.

Drawbacks of this method:

- Selection of the value T?
- The resulting value of χ^2 and parameters are distorted and one should control it.

Later in last half of 60-s JINR switched to the method of Lagrange multipliers used in CERN.

Generalization of the method

Goal: bypass the propagation error problem.
[3] A.J. Ketikian, . . . , V.S. Kurbatov *et al.*, NIM A **314** 572 (1992).

$$\chi^2 = \frac{1}{2} \sum_{i=1,j=1}^{n_f,n_f} (C_i(\boldsymbol{X}) - C_i^m) Q_{i,j} (C_j(\boldsymbol{X}) - C_j^m) \qquad (4)$$

and satisfying the constraints

$$f_\lambda(\boldsymbol{X}) = 0; \quad \lambda = 1 \ldots n_c.$$

$C_i(\boldsymbol{X})$: observables (functions of kinematical parameters $\boldsymbol{X}$);
C_i^m: measured values of observables;
$Q_{i,j}$: inverse error matrix.

If errors have Gaussian distribution and the hypothesis is true, then (4) has χ^2 distribution with ndf $= n_f - n_p + n_c$, n_p is the number of kinematical parameters, i.e. the dimensionality of $\boldsymbol{X}$.

Method of elimination of differentials

In the neighborhood of parameter values $\boldsymbol{X_0}$: the function

$$\chi^2 = \frac{1}{2} \sum_{i=1,j=1}^{n_f,n_f} (C_i(\boldsymbol{X}) - C_i^m) Q_{i,j} (C_j(\boldsymbol{X}) - C_j^m)$$

is approximated by a quadratic form

$$F = F_0 + \boldsymbol{G} \cdot \Delta \boldsymbol{X} + \frac{1}{2} \Delta \boldsymbol{X}^T \cdot Z \cdot \Delta \boldsymbol{X}, \tag{5}$$

and the constraints $\boldsymbol{f}(\boldsymbol{X}) = 0$ by

$$\boldsymbol{f}(\boldsymbol{X}) = \boldsymbol{f}(\boldsymbol{X}_0) + D \cdot \Delta \boldsymbol{X} = 0. \tag{6}$$

- $\boldsymbol{G}$: a vector of derivatives;
- Z: a matrix of second derivatives over $\boldsymbol{X}$;
- D: a matrix of constraint derivatives over $\boldsymbol{X}$ with n_c rows and n_p columns.

Method of elimination of differentials

$$\begin{aligned}\boldsymbol{f}(\boldsymbol{X}) &= \boldsymbol{f}(\boldsymbol{X}_0) + D \cdot \Delta\boldsymbol{X} \\ &= \boldsymbol{f}(\boldsymbol{X}_0) + D_1 \cdot \Delta\boldsymbol{X}_f + D_2 \cdot \Delta\boldsymbol{X}_c.\end{aligned}$$

D_1: sub-matrix of D with n_c rows and $n_p - n_c$ columns;

D_2: sub-matrix of D with n_c rows and n_c columns.

One can express $\Delta\boldsymbol{X}_c$ as a function of $\Delta\boldsymbol{X}_f$:

$$\Delta\boldsymbol{X}_c = \boldsymbol{R} + S \cdot \Delta\boldsymbol{X}_f \tag{7}$$

and substitute it into (5):

$$\begin{aligned}F &= F_0 + \boldsymbol{G} \cdot \Delta\boldsymbol{X} + \frac{1}{2}\Delta\boldsymbol{X}^T \cdot Z \cdot \Delta\boldsymbol{X} \\ &= F_0' + \boldsymbol{G'} \cdot \Delta\boldsymbol{X}_f + \frac{1}{2}\Delta\boldsymbol{X}_f^T \cdot Z' \cdot \Delta\boldsymbol{X}_f.\end{aligned} \tag{8}$$

Thus, we get a quadratic form depending $\boldsymbol{X}_f$ with the dimensionality $n_p - n_c$, and have the dimensionality of the problem reduced.

Details

$$\Delta \boldsymbol{X}_c = \boldsymbol{R} + S \cdot \Delta \boldsymbol{X}_f; \tag{7}$$

$$F = F_0' + \boldsymbol{G'} \cdot \Delta \boldsymbol{X}_f + \frac{1}{2} \Delta \boldsymbol{X}_f^T \cdot Z' \cdot \Delta \boldsymbol{X}_f; \tag{8}$$

$$F_0^{'} = F_0 + \sum_{k=1}^{n_c} R_k \left[G_{n_f+k} + \frac{1}{2} \sum_{l=1}^{n_c} Z_{n_f+k,n_f+l} R_l \right];$$

$$G_i^{'} = G_i + \sum_{k=1}^{n_c} G_{n_f+k} S_{k,i} + \sum_{k=1}^{n_c} R_k \left[Z_{n_f+k,i} + \sum_{l=1}^{n_c} S_{l,i} Z_{n_f+l,n_f+k} \right];$$

$$Z_{i,j}^{'} = Z_{i,j} + \sum_{k=1}^{n_c} \left[S_{k,i} Z_{n_f+k,j} + S_{k,j} Z_{i,n_f+k} \right] + \sum_{k=1,l=1}^{n_c,n_c} S_{k,i} Z_{n_f+k,n_f+l} S_{l,j}.$$

Realization of the method

The FUMILI algorithm has been extended with the method of elimination of differentials.

FUMILI basics:

$$\chi^2 = \frac{1}{2} \sum_{i=1,j=1}^{n_f,n_f} (C_i(\boldsymbol{X}) - C_i^m) Q_{i,j} (C_j(\boldsymbol{X}) - C_j^m) \tag{4}$$

in the neighborhood of $\boldsymbol{X}_0$ expands to

$$F = F_0 + \boldsymbol{G} \cdot \Delta\boldsymbol{X} + \frac{1}{2}\Delta\boldsymbol{X}^T \cdot Z \cdot \Delta\boldsymbol{X}, \tag{5}$$

Requirement for the minimum: the first derivatives of (5) over parameters should equal zeros.

Thus, the formula for the parameters steps leading to the minimum is:

$$\Delta\boldsymbol{X} = -Z^{-1} \cdot \boldsymbol{G}. \tag{9}$$

FUMILI basics, continued

The matrix of the second derivatives (hessian) should be positively defined.
For (4) the matrix of second derivatives is:

$$Z_{i,j} = \sum_{k=1,l=1}^{n_f,n_f} \left[\frac{\partial C_k}{\partial X_i} \frac{\partial C_l}{\partial X_j} + \frac{\partial^2 C_k}{\partial X_i \partial X_j} \left(C_l(\boldsymbol{X}) - C_l^m \right) \right] Q_{k,l}. \tag{10}$$

To ensure the hessian is positively defined, FUMILI utilizes the following trick: the second term in (10) is discarded, and such a matrix is always positively defined.
FUMILI employment in the practice over many years has shown its simplicity and reliability over enormous variety of the problems.

Model example of constrained fit

$$\mathrm{PDF}(x, y) = (1 + \alpha_1 \cdot x + \alpha_2 \cdot y)/(1 + 0.5 \cdot \alpha_1 + 0.5 \cdot \alpha_2)$$

Area: $0 < x < 1$ and $0 < y < 1$;
True values: $\alpha_1 = 0.5$ and $\alpha_2 = 0.8$;
Events: 10^5;
Constraint: $\alpha_1 + \alpha_2 = 1.3$.

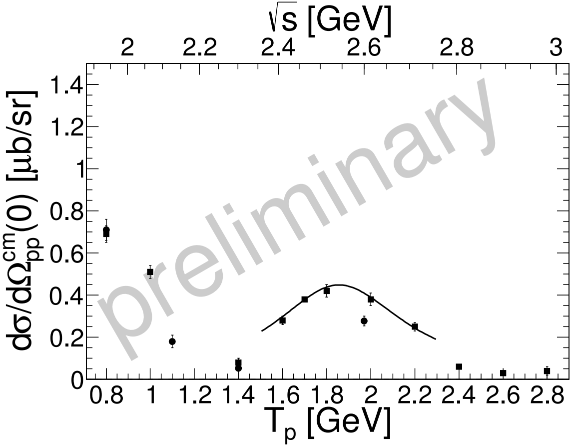
The values of the estimates for the constrained and unconstrained cases. Errors cited are those calculated by the program.

parameter	constrained option	unconstrained option
α_1	0.501 ± 0.013	0.515 ± 0.023
α_2	0.799 ± 0.013	0.815 ± 0.026

- In both cases the estimates are within one calculated error of true values;
- Calculated errors in constrained option are two times less than in unconstrained;
- The values of estimates in constrained option are much nearer to the true one.

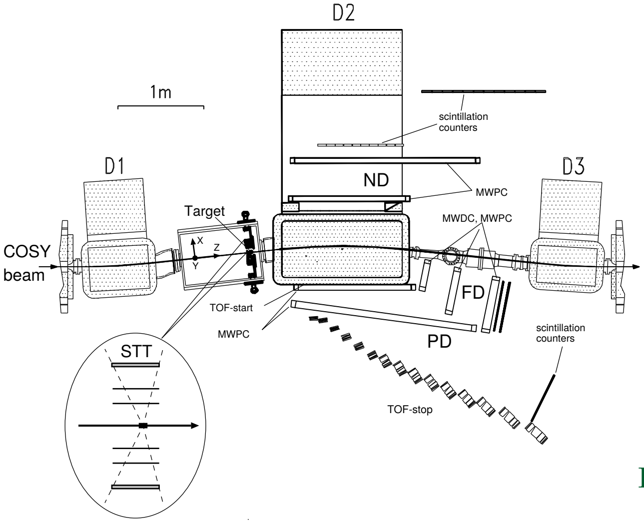
KinFit with elimination of differentials at ANKE

We extensively used KinFit while processing experimental data of the reaction $pp \to pp_S\pi^0$, employing FUMILI extended with the method of elimination of differentials [EPJ Web of Conferences 204, 08008 (2019)].



Forward differential cross section $d\sigma/d\Omega^{\rm cm}_{pp}(0)$ of the $pp \to pp_S\pi^0$ reaction; pp_S is a proton pair with $E_{pp} < 3$ MeV, so that the protons are mainly in $^1S_0$ state. The cross section exhibits a peak at the energy $\sqrt{s} \approx 2.65$ GeV.

ANKE setup



FD: forward detector;

PD: positive detector;

ND: negative detector;

STT: silicon tracking telescope;

D2: main spectrometric magnet;

D1, D3: other ANKE magnets.

Conclusion

- KinFit is an essential technique for a modern particle physics experiment;
- An approach for constrained KinFit using the method of elimination of differentials has been developed several years ago;
- The approach is self-evident and could be applied directly for any iterative method of gradient minimization using χ^2-like functionals;
- The software realization of the method has been developed, extending the FUMILI minimization package;
- The method has been tested using both the model and real experimental cases;
- Three approaches for constrained KinFit have been discussed in the talk: Lagrange multipliers, penalty function and elimination of differentials;
- Future SPD software might contain all the three KinFit methods with an option for a user to switch between them.

Thank you for your attention!

Any questions?