Lagrange multiplier method used in BESIII kinematic fitting*

YAN Liang(严亮)[1,2;1)] HE Kang-Lin(何康林)[1;2)] LI Wei-Guo(李卫国)[1;3)] BIAN Jian-Ming(边渐鸣)[1,2]
FU Cheng-Dong(傅成栋)[1] HUANG Bin(黄彬)[1,2] LIU Ying(刘颖)[3] LÜ Qi-Wen(吕绮雯)[4]
NING Fei-Peng(宁飞鹏)[4] SUN Sheng-Sen(孙胜森)[1] XU Min(徐敏)[5]
ZHANG Jian-Yong(张建勇)[1] ZHU Yong-Sheng(朱永生)[1]

[1] Institute of High Energy Physics, CAS, Beijing 100049, China
[2] Graduate University of Chinese Academy of Sciences, Beijing 100049, China
[3] Guangxi University, Nanning 530004, China
[4] Shanxi University, Taiyuan 030006, China
[5] Department of Modern Physics, University of Science and Technology of China, Hefei 230026, China

Abstract In this paper, the kinematic fitting with the Lagrange multiplier method has been studied for BESIII experiment. First we introduce the Lagrange multiplier method and implement kinematic constraints. Then we present the performance of the kinematic fitting algorithm. With the kinematic fitting, we can improve the resolution of track parameters and reduce the background.

Key words BESIII, kinematic fitting, Lagrange multiplier

PACS 07.05.kf, 02.60.Ed

1 Introduction

BESIII (Beijing Electron Spectrometer III[1]) is a powerfull facility for studying charmonium physics, D-physics, spectroscopy of light hadrons and τ-physics. In physics analysis, the kinematic fitting is an important tool. It mainly uses the particle's information in MDC (Main Drift Chamber) and EMC (Electromagnetic Calorimeter).

In kinematic fitting, we use some of physics laws to constraint the decay process, then we can get better particle parameters. In BESIII offline software [2], before using kinematic fitting, we can use vertex fitting first to update the decay vertex and track parameters. For example, considering the decay chains, $\psi(3770)\rightarrow D^0\overline{D^0}$, where $\overline{D^0}$ decays to the CP eigenstate $K^0_S\pi^0$ and D^0 decays to the hadronic mode $K^-\pi^+$. There are several constraints that can be applied: (1) the $\pi^+\pi^-$ pair from K^0_S decay must come from a common space point($2\times2-3=1$ constraint); (2) the momentum vector of $\pi^+\pi^-$ pair must be aligned with the position vector of the decay vertex relative to the interaction point($3-1$=2 constraints). The above constraints have been dealt with by decay vertex reconstruction [3]. In the kinematic fitting, we mainly use particle energy and momentum information, such as (1) the mass of the $\gamma\gamma$ pair has to be equal to the π^0 mass(1 constraint); (2) energy and momentum are conserved in the $D\overline{D}$ production (4 constraints); and (3) the mass of $K^0_S\pi^0$ has to be equal to the mass of $K^-\pi^+$ (1 constraint). After using the general algorithm which will be introduced in next section with such 9 constraints, their parameters are forced to satisfy the constraints, thereby improving the mass and momentum resolution of the D^0 and the $\overline{D^0}$. These resolution improvements will translate to a larger signal to background ratio and frequently elevate marginal signals to statistical significant re-

Received 24 March 2009, Revised 25 June 2009
* Supported by National Natural Science Foundation of China (10491300, 10491303, 10735080), Research and Development Project of Important Scientific Equipment of CAS (H7292330S7)

1) E-mail: yanl@ihep.ac.cn
2) E-mail: hekl@ihep.ac.cn
3) E-mail: liwg@ihep.ac.cn

sults. The importance of kinematic fitting to data analysis is demonstrated by its use in virtually all modern high energy physics experiments.

2 General algorithm

In kinematic fitting,we use the Lagrange multiplier method [4] to deal with various types of constraints. Based on formulas in the applied fitting theory Ⅳ [5], we have developed the kinematic fitting algorithm. The final equations which we used in the kinematic fitting algorithm can be written as:

$$\begin{aligned} \alpha &= \alpha_0 - V_{\alpha_0} D^{\mathrm{T}} \lambda, \\ \lambda &= V_{\mathrm{D}} (D \delta\alpha_0 + d), \\ V_\alpha &= V_{\alpha_0} - V_{\alpha_0} D^{\mathrm{T}} V_{\mathrm{D}} D V_{\alpha_0}, \end{aligned} \tag{1}$$

where α represents the particle parameters after kinematic fitting, and α_0 means the initial particle parameters before kinematic fitting. $\delta\alpha_0$ is the difference between α and α_{A}, where α_{A} is the expanding point to the constraint equations. V_{α_0} is the error matrix of initial particle parameters and V_α is the updated error matrix. λ is called Lagrange multipliers. $V_{\mathrm{D}} = (D V_{\alpha_0} D^{\mathrm{T}})^{-1}$ is the $m \times m$ constraint covariance matrix. D and d are used to linearize the constraints equations. Different constraints have different expressions.

Finally, we get the χ^2 expression:

$$\chi^2 = \lambda^{\mathrm{T}} V_{\mathrm{D}}^{-1} \lambda = \lambda^{\mathrm{T}} (D \delta\alpha_0 + d). \tag{2}$$

Note that the χ^2 can be written as a sum of m distinct terms, one for each constraint. It can be shown that the new covariance matrix V_α has diagonal elements that are smaller than the initial covariance matrix V_{α_0}. In general, the nonlinearities of the constraint equations require that the kinematic fitting procedure be applied iteratively until satisfactory convergence is achieved. Track parameters and their errors, covariance matrices, fit information and other quantities can be obtained after fitting.

3 Track parameter representation

For kinematic fitting, it is important to choose a track representation that uses physically meaningful quantities. We adopt the 4-parameter W format, defined as $\alpha = (p_x, p_y, p_z, E)$, the 4-momentum, in the BESIII kinematic fitting software package. It is straight-forward to transfer the parameters of neutral tracks and their covariance to the W representation. It has been noted that the W format also has enough information to represent the general decays of particles.

3.1 Charged track representation

For charged track, 5-helix parameters are adopted in MDC track fitting program. The relation of helix and 4 momentum is quite simple:

$$\boldsymbol{\alpha}_0 = \begin{pmatrix} p_{0x} \\ p_{0y} \\ p_{0z} \\ E_0 \end{pmatrix} = \begin{pmatrix} -\sin\phi_0 \cdot Q/\kappa \\ \cos\phi_0 \cdot Q/\kappa \\ \lambda \cdot Q/\kappa \\ \sqrt{(1+\lambda^2)/\kappa^2 + m^2} \end{pmatrix}, \tag{3}$$

where ϕ_0 is the track direction polar angle, κ is the radius of helix, λ is the track direction closest to the z-axis point, Q is the particle charge, m is the mass of particle. The corresponding Jacobian matrix J_{W} can be easily calculated. The 4-momentum covariance matrix can be gotten from:

$$V_{\mathrm{W}} = J_{\mathrm{W}} \cdot V_{\mathrm{helix}} \cdot J_{\mathrm{W}}^{\mathrm{T}}. \tag{4}$$

3.2 Neutral track representation

Neutrals are reconstructed in the calorimeter with the position $(\phi, \cos\theta)$, an energy E, and a corresponding 3×3 covariance matrix. Supposing the decay vertex is fixed at the point of origin, the 4-momentum of neutral tracks, e.g., photons, can be easily determined by

$$p = \begin{pmatrix} E\sin\theta\cos\phi & E\sin\theta\sin\phi & E\cos\theta & E \end{pmatrix}.$$

The decay particles are not always produced from the origin. The track parameters for neutrals have to be recalculated as:

$$\begin{pmatrix} \phi' \\ \lambda' \\ p \end{pmatrix} = \begin{pmatrix} \tan^{-1}\dfrac{y}{x} \\ \dfrac{z}{\sqrt{x^2+y^2}} \\ E \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_{\mathrm{clus}} - x_{\mathrm{beam}} \\ y_{\mathrm{clus}} - y_{\mathrm{beam}} \\ z_{\mathrm{clus}} - z_{\mathrm{beam}} \end{pmatrix}, \tag{5}$$

where $\vec{x}_{\mathrm{clus}} = (x_{\mathrm{clus}}, y_{\mathrm{clus}}, z_{\mathrm{clus}})$ is the cluster hit position in EMC that can be got by the EMC reconstruction, and $\vec{x}_{\mathrm{beam}} = (x_{\mathrm{beam}}, y_{\mathrm{beam}}, z_{\mathrm{beam}})$ is the interaction position which can be determined from vertex reconstruction. Since $\vec{x}_{\mathrm{clus}}$ and $\vec{x}_{\mathrm{beam}}$ are calculated independently, there is no correlation between $\vec{x}_{\mathrm{clus}}$ and $\vec{x}_{\mathrm{beam}}$. According to the error propagation formula, we have:

$$V_{\phi',\lambda'} = J_1 \cdot V_{\mathrm{clus}} \cdot J_1^{\mathrm{T}} + J_2 \cdot V_{\mathrm{beam}} \cdot J_2^{\mathrm{T}}, \tag{6}$$

where

$$J_1 = \frac{\partial(\phi', \lambda')}{\partial(\vec{x}_{\mathrm{clus}})}, \quad J_2 = \frac{\partial(\phi', \lambda')}{\partial(\vec{x}_{\mathrm{beam}})}$$

and $J_1=-J_2$, $V_{\rm clus}$ is the EMC cluster error matrix and $V_{\rm beam}$ is the beam vertex error matrix.

It is assumed that there is no correlation between energy measurement and position measurement, The final shower representation will be:

$$\begin{pmatrix}\phi'\\ \lambda'\\ E\end{pmatrix}\begin{pmatrix}V_{\phi',\lambda'} & 0\\ 0 & (\delta E)^2\end{pmatrix}.$$

In the kinematic fitting, we use 4-momentum format:

$$\begin{pmatrix}p_x\\ p_y\\ p_z\\ E\end{pmatrix}=\begin{pmatrix}\dfrac{p}{\sqrt{1+\lambda^2}}\cdot\cos\phi\\ \dfrac{p}{\sqrt{1+\lambda^2}}\cdot\sin\phi\\ p\cdot\lambda/\sqrt{1+\lambda^2}\\ \sqrt{p^2+m^2}\end{pmatrix}. \qquad (7)$$

In the same way, the covariance matrix of 4-momentum is deduced by $J\cdot V_{\phi,\lambda,p}\cdot J^{\rm T}$,where J is:

$$J=\begin{pmatrix}-p_y & -\dfrac{p_z\cdot p_x\cdot p_\perp}{p^2} & \dfrac{p_x}{p}\\ p_x & \dfrac{p_z\cdot p_y\cdot p_\perp}{p^2} & \dfrac{p_y}{p}\\ 0 & \dfrac{p_\perp^3}{p^2} & \dfrac{p_z}{p}\\ 0 & 0 & \dfrac{p}{E}\end{pmatrix}.$$

4 Applied kinematic constraints at BESⅢ

In this section, we will introduce some constraints which are used in our implementation. For various constraints, we write the linearized parameters D and d.

Suppose that we start with $\alpha_0 = \begin{pmatrix}\alpha_{01} & \alpha_{02} & \ldots & \alpha_{0n}\end{pmatrix}$ and

$$V_{0\alpha}=\begin{pmatrix}V_{01} & & & \\ & V_{02} & & \\ & & \ddots & \\ & & & V_{0n}\end{pmatrix},$$

where $\alpha=(p_x,p_y,p_z,E)$ is a 4-momentum vector. V_α is its covariance matrix.

4.1 Invariant mass constraint

The constraints are

$$d=E^2-p_x^2-p_y^2-p_z^2-m_{\rm c}^2=0,$$

where $m_{\rm c}$ that we constraint the invariant mass of several tracks to be this value. The linearization yields

$$D=\begin{pmatrix}-2p_x & -2p_y & -2p_z & 2E\end{pmatrix}.$$

In some sense, the mass constraint is most complicated, because it is quadratic without any linear terms. The mass constraint is one of the most common kinematic constraints, for example, $\pi^0/\eta\rightarrow\gamma\gamma$.

4.2 Total energy constraint

The constraints are

$$d=E-E_{\rm c}=0,$$

where $E_{\rm c}$ that we constraint a particle's energy to be this value. The linearization yields

$$D=\begin{pmatrix}0 & 0 & 0 & 1\end{pmatrix}.$$

The energy constraint is linear. In charm physics analysis, it's often used to reconstruct the beam-constraint mass of single charm tag. For example, in the decay of $\psi(3770)\rightarrow D\bar{D}$, each D meson carries a constant energy (beam energy) in the rest frame of $\psi(3770)$.

4.3 Total momentum constraint

The constraint is

$$d=\sqrt{p_x^2+p_y^2+p_z^2}-p_{\rm c}=0,$$

where $p_{\rm c}$ is that we constrain a particle's momentum to be this value. The linearization yields

$$D=\begin{pmatrix}p_x/p & p_y/p & p_z/p & 0\end{pmatrix}.$$

As the mass constraint, the total momentum is a nonlinear constraint. For example, in the study of $\psi(4030)\rightarrow D\bar{D}^*$, the momentum of the direct daughter D is a constant. This constraint can also be applied in the decay of $\psi(3770)\rightarrow D\bar{D}$.

4.4 Three momentum constraints

The constraints are

$$d=\begin{pmatrix}p_x-p_{cx}\\ p_y-p_{cy}\\ p_z-p_{cz}\end{pmatrix}=0.$$

The linearization yields

$$D=\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\end{pmatrix}.$$

The linear 3-momentum constraints are seldom used in physics analysis, for example, to distinguish the processes of $J/\psi\rightarrow p\bar{p}$, $J/\psi\rightarrow\pi^+\pi^-$ and $J/\psi\rightarrow$

K^+K^-. After the 3-momentum constraints fit, one can check the total momentum distribution.

4.5 Four momentum constraints

Four momentum constraints are the most common analysis tools for most of the physics topics at BESIII. The constraints are

$$d=\begin{pmatrix} p_x-p_{cx} \\ p_y-p_{cy} \\ p_z-p_{cz} \\ E-E_c \end{pmatrix}=0.$$

The linearization yields

$$D=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

In most of the analyse of J/ψ and $\psi(2S)$ physics, the decay daughters are required to satisfying the momentum-energy conversation. The 4-momentum constraints are helpful to improve the momentum, energy and mass resolution, and also useful to suppress background contaminations.

4.6 Equal mass constraint

In some analyse, for example, in $\psi(3770)\to D\bar{D}$ and $J/\psi\to\gamma\pi^0\pi^0$ the D and π^0 are reconstructed by their decay modes. If we want to check the signal/background linear shape without fixing their mass to the nominal value, we may require the additional constraint $m_D=m_{\bar{D}}$ and/or $(m_{\pi^0})_1=(m_{\pi^0})_2$. In another word, we want to add the constraint

$$d=(E_1^2-p_{1x}^2-p_{1y}^2-p_{1z}^2)-(E_2^2-p_{2x}^2-p_{2y}^2-p_{2z}^2)=0.$$

The linearization yields

$$D_1=\begin{pmatrix} -2p_{1x} & -2p_{1y} & -2p_{1z} & 2E_1 \end{pmatrix},$$
$$D_2=\begin{pmatrix} 2p_{2x} & 2p_{2y} & 2p_{2z} & -2E_2 \end{pmatrix},$$

where the D_1 and D_2 correspond to the derivatives to the two sets of track parameters.

5 Physics performance

The kinematic fitting with the Lagrange multiplier method uses the MDC information for charged tracks and EMC information for neutral particles through different kinematic constraints to update the track parameters including the energy, the momentum and their resolutions.

There are two important variables to check the kinematic fitting process: pull [6] and χ^2. The pull of the $i^{\rm th}$-track parameter is defined as:

$$(\text{pull})_i=\frac{\alpha_i-\alpha_{0i}}{\sqrt{(V_{\alpha_0})_{ii}-(V_\alpha)_{ii}}}. \tag{8}$$

The resulting χ^2 that is obtained with m constraints is distributed like a standard χ^2 with m degrees of freedom, if Gaussian errors apply. Of course, since the track errors are only approximately Gaussian, the actual distribution will have more events in the tail than predicted by Gaussian approximation. Still, knowledge of the distribution allows one to define the reasonable χ^2 cuts.

5.1 $J/\psi\to\rho^0\pi^0$

In the $J/\psi\to\rho^0\pi^0$ with $\rho^0\to\pi^+\pi^-$ and $\pi^0\to\gamma\gamma$, because of the narrow width of the J/ψ peak, we can constrain that all the tracks in the final state should satisfy the four momentum constraints. Another constraint, namely the invariant mass of photon pair equals the π^0 mass, can be added.

5.1.1 Performance of kinematic constraints

Figure 1 shows the distribution of χ^2 and the probability of χ^2 which is the cumulative distribution function of the χ^2 variable. But the existence of the photons made the problem more complicated, so in the analysis of decay which contains gamma we did some adjustment, and that could influence the final result.

To check the kinematic fit, we draw the pull of the helix parameters for charged tracks of π^+.

In Fig. 1, we notice that the mean of the χ^2 consistency is 0.3083, which is not around 0.5 as the theory expected. Because the probability of χ^2 should be the flat distribution in the range of (0,1). In Fig. 2, ϕ_0 is the track direction polar angle, κ is the radius of helix. In the pull distribution of κ, the mean value is apart from the 0 apparently. The main reason is that the initial values of the parameters and their errors are not very accurate. As mentioned in a Gaussian-sum filter for vertex reconstruction [7]: The model is linear and all random noise is Gaussian. In that case, the Lagrange multiplier estimation is unbiased and has minimum variance, residuals and standardized residuals (pulls) of estimated quantities have Gaussian distributions, and the minimum of the objective function obeys a χ^2-distribution. Despite the fact that some of the checked parameters distributions are not perfect, the physics values we care about are worth trusted. For non-linear models or non-Gaussian noise, it is still the optimal linear estimator. The π^0 mass distribution is good according to the mean value and its error compared with the EMC information read directly from the reconstruction.

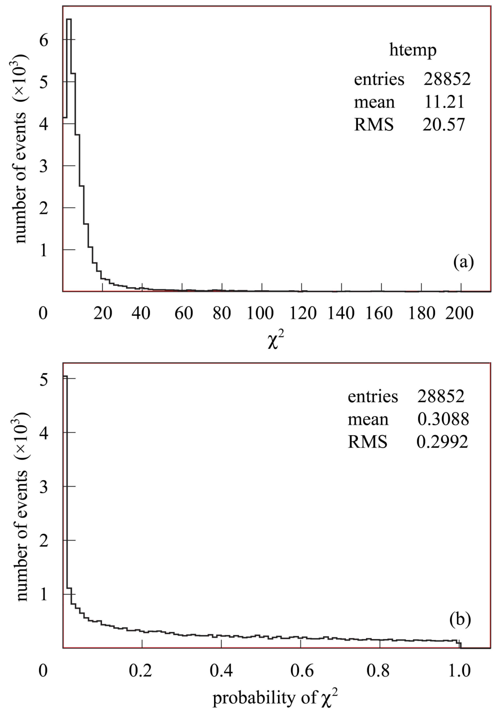

Fig. 1. χ^2 distribution of four momentum constraints in Lagrange multiplier method (a); Probability of χ^2 (b), for $J/\psi \rightarrow \rho^0\pi^0$ channel.

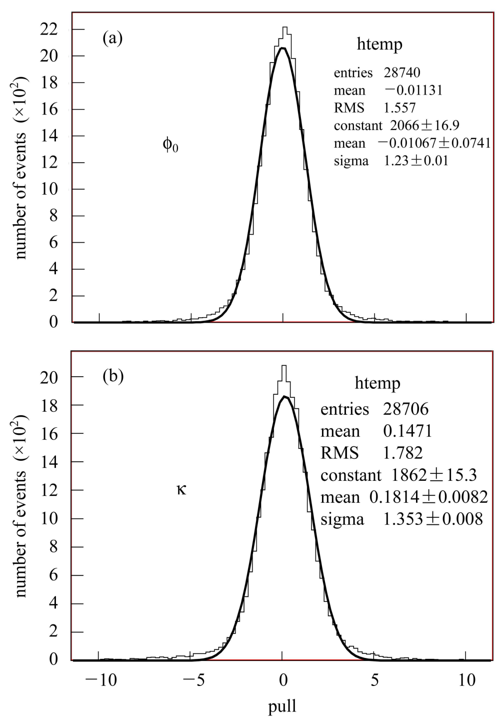

Fig. 2. The distribution of π^+ pull in helix format.

Besides the four momentum constraints, we add an invariant mass constraint in addition which requires that the two photon's invariant mass be equal to the π^0 mass. The degree of the freedom of the χ^2 distribution should be 5.

5.1.2 The efficiency of kinematic fitting

Another important quantity which should be studied is the efficiency of the kinematic fitting. First of all, we use the decay $J/\psi \rightarrow \rho^0\pi^0$ with $\rho^0 \rightarrow \pi^+\pi^-$ and $\pi^0 \rightarrow \gamma\gamma$ to check the efficiency for 4-momentum constraint and resonance constraint mentioned above. Before the kinematic fitting is applied, we selected the signals by PID, vertex fitting and other simple cuts such as total tracks and charge. In the 50 thousand Monte Carlo events, the number of events which pass the preliminary selection criteria is N_1 30887. Secondly, we defined the three types of efficiencies. The number of events passing 4-constraint fit (4c) is N_2 and 5c N_3. The efficiencies are defined as $\varepsilon_1 = \dfrac{N_2}{N_1}$, $\varepsilon_2 = \dfrac{N_3}{N_1}$, $\varepsilon_3 = \dfrac{N_3}{N_2}$. In the kinematic fitting, we can get the χ^2 value for each fitting, the distribution of χ^2 is dependent on the degree of the freedom of the total constraints in theory. From Fig. 3, the efficiency is dependent on χ^2 cut. It is easy to understand that the tight cut can reduce the efficiency. But we noticed that even if the χ^2 cut is loosely enough, the efficiency can not reach 100%. The reasons which affect the efficiency could be:

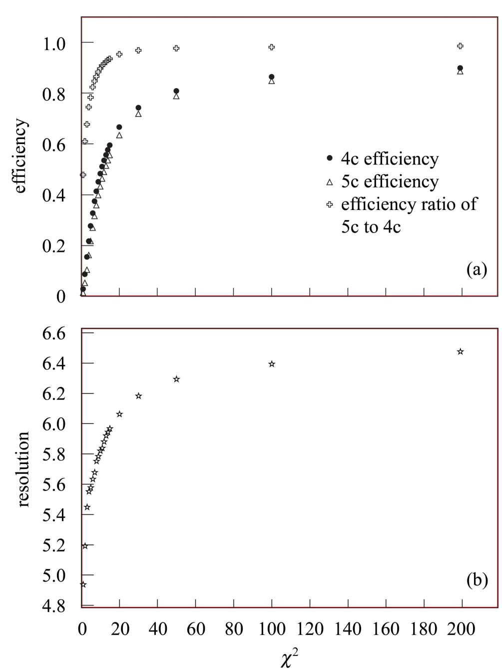

Fig. 3. Efficiencies vs χ^2 (a); resolution of π^0 mass vs χ^2 (b).

1) the χ^2 cut, but loose cut also affects the resolution of results;

2) the error matrix of track parameters also can influence the fitting, for charged tracks, we need the accurate MDC kalman tracking information [8];

3) because of the asymmetry of photon energy deposit shape, it can not be treated as Gaussian. The dynamic error should be introduced.

Considering the resolution and efficiency together, the 40 for χ^2 cut is recommended.

5.2 The parameters improvement using kinematic fitting

After using the kinematic fitting, we not only reduce the background, but also could revise the track parameters to get better resolutions. We compare the π^0 mass before kinematic fitting with the mass after kinematic fitting in the decay of $J/\psi \to \rho^0\pi^0$, as shown in Fig. 4.

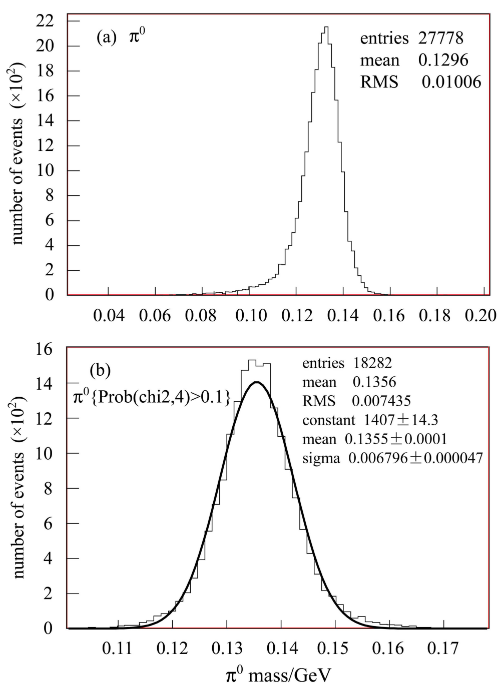

Fig. 4. The distribution of π^0 mass before kinematic fitting (a); the distrubition of π^0 mass after kinematic fitting (b).

Figure 4 shows the mean value of the invariant mass of two gammas from π^0 is far from the real value and the resolution is about 10 MeV. After kinematic fitting, the mean value is close to the PDG value and the resolution changes to 6.7 MeV. Through kinematic fitting, we get more accurate parameter value and better resolution.

6 Summary

The Lagrange multiplier method has been implemented for the kinematic fitting of the BESIII offline software. It has been validated with Monte Carlo simulation, with charged particles and neutral particles in various constraints. Through these tests, the particle parameters' value and resolution have been improved after kinematic fitting. From Fig. 4, we can see that the four momentum constraints can clearly improve the resolution of π^0 mass and the mean value of mass distribution. In the foundation of parameters with Gaussian error, the pull of parameters in kinematic fitting is $(0,1)$ normal Gaussian distribution. By adjusting the χ^2 cut, we can change the purity and efficiency. So we need to choose a proper cut depending on different situations. And we also should adjust the error matrix element to make the pulls closer to the normal distribution. For the constraints with error, the Lagrange multiplier method can not handle such problems, especially when the error distribution is not Gaussian. The next step is to use the Kalman filter method [9] to deal with it, and also can deal with the missing parameter situation.

References

1 BESIII Design Report. Interior Document in Institute of High Energy Physics, 2004
2 LI W D, LIU H M et al. The Offline Software for the BESIII Experiment. Proceeding of CHEP06. Mumbai, 2006
3 XU M, HE K L et al. Chinese Physic C, 2009, **33**(06): 428–435
4 Avery P. Applied Fitting Theory Ⅰ: General Least Squares Theory, CBX-91-72, 1991
5 Avery P. Applied Fitting Theory Ⅳ: Formulas for Kinematic Fitting, CBX-98-37, 1992
6 Frodeson A G et al. Probability and Statistics in Particle Physics, Universitetsforlaget, Bergen-Oslo-Tromso, 1979
7 Speer T, Frühwirth R. Computer Physics Communications, 2006, **174**(12): 935–947
8 WANG J K et al. Chinese Physic C, 2009, **33**(03): 210–216
9 Kalman R E. Transaction of the ASME – Journal of Basic Engineering, 1960, **82**(Series D): 35–45