Kinematic and vertex fitting package for the CMD-3 experiment

S.S. Gribanov and A.S. Popov

Budker Institute of Nuclear Physics, SB RAS,
Prospekt Akademika Lavrent’eva, 11, Novosibirsk, 630090 Russia
Physics Department, Novosibirsk State University,
Pirogova, 2, Novosibirsk, 630090 Russia

E-mail: S.S.Gribanov@inp.nsk.su

ABSTRACT: In this article, we discuss a new software package of kinematic and vertex fitting for the CMD-3 experiment at the VEPP-2000 electron-positron collider. The authors describe in detail the fitting algorithm, parametrization of four-momenta and trajectories of various particles and present the results of testing the fitting package using events of Monte Carlo simulation of various $e^+e^-$ annihilation processes. The authors also provide several examples of the fitting package validation using Gaussian simulation. Although the package discussed in this article is intended for the CMD-3 experiment, it can also be used in other experiments. The authors consider the described package as their first step towards a more universal and rigorous kinematic and vertex fitting package that can be used in future $e^+e^-$ experiments, such as the Super Charm-Tau factory.

KEYWORDS: Data processing methods; Particle identification methods; Software architectures (event data models, frameworks and databases)

ArXiv ePrint: 2208.11569

*Corresponding author.
8.2.1 Description of the $e^+e^- \rightarrow K_SK^+\pi^-$ hypotheses 36
8.2.2 Fitting the $e^+e^- \rightarrow K_SK^+\pi^-$ events 37
8.3 Hypothesis $e^+e^- \rightarrow \gamma \gamma$ 38
8.3.1 Description of the $e^+e^- \rightarrow \gamma \gamma \gamma$ hypothesis 38
8.3.2 Fitting the $e^+e^- \rightarrow \pi^0\gamma$ events 38
8.4 Hypothesis $e^+e^- \rightarrow \pi^+\pi^-\gamma\gamma$ 41
8.4.1 Description of the $e^+e^- \rightarrow \pi^+\pi^-\gamma\gamma$ hypothesis 41
8.4.2 Fitting the $e^+e^- \rightarrow \eta\pi^+\pi^-$, $\eta \rightarrow \gamma \gamma$ events 41
8.5 Hypothesis $e^+e^- \rightarrow \eta\pi^+\pi^-$, $\eta \rightarrow \pi^+\pi^-\pi^0_{\text{lost}}$ 42
8.5.1 Description of the $e^+e^- \rightarrow \eta\pi^+\pi^-$, $\eta \rightarrow \pi^+\pi^-\pi^0_{\text{lost}}$ hypothesis 42
8.5.2 Fitting procedure details 43
8.5.3 Fitting the $e^+e^- \rightarrow \eta\pi^+\pi^-$, $\eta \rightarrow \pi^+\pi^-\pi^0_{\text{lost}}$ events 43

9 Examples of Gaussian simulation 44
9.1 Hypothesis $e^+e^- \rightarrow \pi^+\pi^-\gamma\gamma$ 45
9.2 Hypotheses $e^+e^- \rightarrow XK^+\pi^-$, $X \rightarrow \pi^+\pi^-$ 48

10 Fitting package 52

11 Summary 53

1 Introduction

Kinematic and vertex fitting is widely used data analysis technique in particle physics experiments [1–10]. This technique can be used in order to separate events corresponding to different kinematic hypotheses and to reconstruct interaction and decay vertices. In addition, the use of this technique often leads to significant improvements in resolution of some quantities, such as various invariant masses. This paper describes kinematic and vertex fitting package developed for the CMD-3 experiment [11–16] located at the VEPP-2000 electron-positron collider [17, 18] with a maximum center-of-mass energy of 2 GeV. The discussed package is currently actively used in the study of a number of processes with the CMD-3 detector. This package can also be used in similar experiments. To do this, it is necessary to make a number of experiment-dependent changes related to the format of input data and probably to some particle parametrizations. The authors believe that it is quite easy to do, since the discussed package is divided into detector-dependent and detector-independent parts.

In general, any kinematic and vertex fitting algorithm consists in constrained minimization of the chi-square function:

$$
\chi^2(x) = \Delta x^\top \hat{C}^{-1} \Delta x,
$$

where $x \in \mathbb{R}^n$ is the vector of the optimization variables corresponding to measurable parameters, i.e. such parameters that can be measured in an experiment. The natural number $n$ is the number of such
parameters, the vector $\tilde{x}$ is the vector of these parameters, measured in a single event. The matrix $\hat{\tilde{C}}$ is the corresponding covariance matrix, obtained in the same event. The parameters measured by a detector are usually referred to one particle or another. Moreover, usually the parameters corresponding to different particles are measured independently. With these assumptions in mind, the chi-square from equation (1.1) can be written as

$$
\chi^2(x) = \sum_{i=1}^{N_p} \Delta x^{(i)} \top \hat{\tilde{C}}^{(i)} \Delta x^{(i)},
$$

(1.2)

where $i$ is the particle index, $N_p$ is the number of particles in a hypothesis, $x^{(i)}$ is the vector of the optimization variables corresponding to measurable parameters of the $i$-th particle. The vector $\tilde{x}^{(i)}$ is the vector of these parameters, measured in a single event. $\hat{\tilde{C}}^{(i)}$ is the corresponding covariance matrix. The notation $x^{(i)} \oplus x^{(j)}$ means the concatenation of the vectors $x^{(i)}$ and $x^{(j)}$, while the notation $\hat{\tilde{C}}^{(i)} \oplus \hat{\tilde{C}}^{(j)}$ means the direct sum of the matrices $\hat{\tilde{C}}^{(i)}$ and $\hat{\tilde{C}}^{(j)}$. The parameters measured by a detector can belong not only to particles, but also, for example, to the $e^+e^-$ interaction vertex. The contribution of these parameters can also be included in the chi-square given by equation (1.2).

In the case of kinematic and vertex fitting, equality constraints are commonly used. These constraints have the following form:

$$
f_k(y) = 0, \ k = 1, 2, \ldots, m,
$$

(1.3)

where $k$ is the constraint index, $f_k$ is the $k$-th constraint function, $m$ is the number of constraints in a hypothesis, $a \in \mathbb{R}^l$ is the vector of non-measurable optimization parameters, $l$ is the number of these parameters. In this article we mean that non-measurable parameters are those parameters that cannot be measured experimentally, but can be obtained as result of fitting. There is a number of such parameters: momentum components of a lost particle, time parameter of a charged particle trajectory, vertex coordinates, and some other parameters. Some examples of non-measurable parameters are given in sections 5 and 6.

Taking into account the chi-square function (1.1) or (1.2) and the equality constraints (1.3), one can formulate the following problem for finding the conditional extremum of the chi-square function:

$$
\begin{align*}
\left\{ \chi^2(x) \rightarrow \text{extremum}, \\
f_k(y) = 0, \ k = 1, 2, \ldots, m.
\right.
\end{align*}
$$

(1.4)

It is important to note that the initial values of the measurable parameters $x$ should be set equal to their measured values $\tilde{x}$. In this case, the minimization algorithm starts from the lowest point of the chi-square paraboloid, which increases the probability that the fit will converge to a global minimum. The criterion that the found extremum is a local minimum and not a local maximum is discussed in section 2.
2 Minimization algorithm

In this work the well-known method of Lagrange multipliers [19–21] is used in order to reduce the constrained least-squares optimization problem to the unconstrained one. The method of Lagrange multipliers is based on the concept of the Lagrange function. In the context of the problem (1.4) this function has the following form:

\[ L(y, \lambda) = \chi^2(x) + \sum_{k=1}^{m} \lambda_k f_k(y), \quad \lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{bmatrix}, \] (2.1)

where variables \( \lambda_1, \lambda_2, \ldots, \lambda_m \) are the Lagrange multipliers. Taking into the account the Lagrange function, the necessary condition for the existence of an extremum in the problem (1.4) can be formulated using the following theorem [19].

**Theorem 2.1 (necessary condition).** If \( y' \in \Omega \equiv \{y \in \mathbb{R}^{n+l} | f_k(y) = 0, \, k = 1, 2, \ldots, m\} \) is a conditional extremum point of the problem (1.4), then there is a nonzero vector \( \lambda' \in \mathbb{R}^m \) such that

\[ \nabla_y L(y, \lambda') \big|_{y=y'} = 0. \] (2.2)

In order to determine whether the found extremum is a local minimum or a local maximum, it is necessary to formulate sufficient conditions for an extremum [19].

**Theorem 2.2 (sufficient conditions).** Let \( y' \in \Omega \) and there is a vector \( \lambda' \), such that \( \nabla_y L(y, \lambda') \big|_{y=y'} = 0 \). If for any vector \( dy \in \mathbb{R}^{n+l}, \, dy \neq 0 \) that satisfies the conditions \( \nabla_y f_k(y) \big|_{y=y'} \cdot dy = 0, \, k = 1, 2, \ldots, m \), the following inequality holds

\[ \sum_{i=1}^{n+l} \sum_{j=1}^{n+l} \frac{\partial^2 L(y, \lambda')}{\partial y_i \partial y_j} \big|_{y=y'} dy_idy_j > 0, \] (2.3)

then the strict local minimum of the function \( \chi^2 \) on the set \( \Omega \) is reached at the point \( y = y' \). If the following inequality holds

\[ \sum_{i=1}^{n+l} \sum_{j=1}^{n+l} \frac{\partial^2 L(y, \lambda')}{\partial y_i \partial y_j} \big|_{y=y'} dy_idy_j < 0, \] (2.4)

then the strict local maximum is reached at the point \( y = y' \).

Thus, in order to find the conditional minimum of the \( \chi^2 \) function (1.2), one should follow the algorithm below. First of all, we need to compose the Lagrange function according equation (2.1). Then we need to use the necessary condition for the existence of an extremum (theorem 2.1), i.e. solve the following system of equations:

\[
\begin{align*}
\left\{ \begin{array}{l}
\nabla_y L(y, \lambda) = 0, \\
f_k(y) = 0, \, k = 1, 2, \ldots, m.
\end{array} \right. \tag{2.5}
\]
Let us note that the system of equations (2.5) can be rewritten in a form that is more convenient for further use:

\[
\nabla_q \mathcal{L}(y, \lambda) = 0, \\
q = y \oplus \lambda,
\]

where \( q \) is the vector of all parameters involved in the constrained chi-square minimization, including the Lagrange multipliers. Suppose that as a result of solving the system of equations (2.6), a pair of vectors is obtained: \( y = y' \) and \( \lambda = \lambda' \). This pair of vectors corresponds to a conditional local extremum of the chi-square function. Finally, it is necessary to check whether this pair of vectors corresponds to a local minimum or to a local maximum point. To do this, we can use inequalities (2.3) and (2.4).

To solve the system of equations (2.6), the package discussed in this paper uses Newton’s method [22–24]. This method is based on the following iterative approach:

\[
q_s = q_{s-1} - \hat{H}^{-1}(q_{s-1}) \nabla_q \mathcal{L}(y_{s-1}, \lambda_{s-1}), \quad s = 1, 2, \ldots,
\]

where \( q_s = x_s \oplus a_s \oplus \lambda_s \) is the vector of optimization parameters at the \( s \)-th step of the iterative procedure. \( q_0 = x_0 \oplus a_0 \oplus \lambda_0 \) is the vector of their initial values, \( \hat{H}(q_{s-1}) \) is the Hessian of the Lagrange function with respect to the variables \( q \), and \( \nabla_q \mathcal{L}(y_{s-1}, \lambda_{s-1}) \) is its gradient calculated at the point \( q_{s-1} \). The vector \( y_{s-1} \) is the concatenation of the vectors of measurable \( x_{s-1} \) and non-measurable \( a_{s-1} \) parameters at the \( (s-1) \)-th step of the optimization algorithm: \( y_{s-1} = x_{s-1} \oplus a_{s-1} \).

The initial values \( x_0 \) of the measurable parameters \( x \) are set equal to their measured values \( \tilde{x} \). The initial values \( \lambda_0 \) of the Lagrange multipliers are set to zero. The initial values \( a_0 \) of the non-measurable parameters can optionally be set to custom values.

The Hessian \( \hat{H}(q_{s-1}) \) of the Lagrange function has the following form:

\[
\hat{H}(q_{s-1}) = \left[ \frac{2\hat{M} + \hat{Q}_{s-1}}{J_{s-1}} \right] \partial^T \mathcal{J}_{s-1} \mathcal{L}(y_{s-1}, \lambda_{s-1}),
\]

where \( \mathcal{J}_{s-1} \) is the matrix containing the Hessians of the constraint functions with respect to the variables \( y \). Let us denote the Hessian of the \( k \)-th constraint function calculated at the point \( y_{s-1} \) as \( \hat{Q}_{s-1}^{(k)} \). In these notations, the matrix \( \hat{Q}_{s-1} \) has the following form:

\[
\hat{Q}_{s-1} = \sum_{k=1}^{m} \lambda_{k,s-1} \hat{Q}_{s-1}^{(k)},
\]

where \( \lambda_{k,s-1} \) is the value of the \( k \)-th Lagrange multiplier at the \( (s-1) \)-th step of the iterative procedure given by equation (2.7). The matrix \( J_{s-1} \) in equation (2.8) is the Jacobian of the constraint functions with respect to the variables \( y \). This Jacobian is also calculated at the point \( y_{s-1} \).

The constraints (1.3) do not explicitly depend on the parameters of particles and vertices. The constraint functions usually explicitly depend only on parametrizations of the particle four-momenta and trajectories (see section 6), as well as the vertex parametrization (see section 5). In this case, it
is convenient to use the chain rule to obtain partial derivatives of the constraints:

\[ \frac{\partial f_k(g_1(y), \ldots, g_{N_g}(y))}{\partial y_i} = \frac{\partial f_k}{\partial g_j} \frac{\partial g_j}{\partial y_i}, \]

(2.10)

where \( g_j \) is a function of some parametrization, \( N_g \) is the number of such functions, the repeated index \( j \) is summed up. Using a chain rule in the fitting package makes it possible to implement constraint, particle, and vertex classes independently. Thus, for example, adding a new particle class to the package will not change the implementation of the constraints, and vice versa.

In the case of the package discussed in this paper, the procedure given by equation (2.7) continues until one of the following conditions is met:

1. \( |\chi^2(x_s) - \chi^2(x_{s-1})| < \epsilon \) and \( \delta(y_s) < \epsilon \), where \( \delta(y_s) = \sum_{k=1}^{m} |f_k(y_s)| \) is the residual and \( \epsilon \) is some tolerance;
2. \( s > N_{\text{iter}} \), where \( s \) is the step number of the iterative procedure given by equation (2.7) and \( N_{\text{iter}} \) is the maximum number of iterations.

If the first condition (1) is met, then the fit has converged to a local extremum. If the second condition (2) is met, then the fit has not converged. The default tolerance \( \epsilon \) and maximum number of iterations \( N_{\text{iter}} \) used in the fitting package are equal to \( 10^{-4} \) and 20, respectively. These values can be easily adjusted if desired. Although the default value for the maximum number of iterations is equal to 20, the optimization algorithm usually converges in fewer iterations. For example, in the case of fitting the events of the \( e^+e^- \rightarrow K_SK^+\pi^- \), \( K_S \rightarrow \pi^+\pi^- \) process under the corresponding signal hypothesis (see sections 3 and 8.2.2), the optimization algorithm most often converges in three iterations. At the same time, in about 93% of all events, the algorithm converges in 2–6 iterations. In the case of fitting the events of the \( e^+e^- \rightarrow \eta\pi^+\pi^- \), \( \eta \rightarrow \gamma\gamma \) process (see section 8.4.2), the algorithm most often converges in two iterations. In about 98% of the \( e^+e^- \rightarrow \eta\pi^+\pi^- \), \( \eta \rightarrow \gamma\gamma \) events, the algorithm converges in 2–4 iterations.

In the implementation of the optimization algorithm, special attention is paid to periodic parameters. If, during optimization, such a parameter goes beyond the period boundaries, then the corresponding contribution to the chi-square becomes incorrect because the measured value of this parameter is within the period. Therefore, after each iteration, if the parameter goes beyond the period boundaries, then it is returned back using a shift by an integer number of periods. Since it is assumed that it does not matter for the constraint functions whether the parameter is inside the period or not, the parameter values before and after the shift are equivalent.

Another important feature of the fitting package is the ability to use upper and lower limits for parameters. If, during optimization, a parameter is out of range, it returns to its initial value, after which the optimization algorithm continues.

The fitting package provides an interface for fixing and releasing parameters. Fixing a parameter means that this parameter does not change during the execution of the optimization algorithm. This can be achieved in the following way. In the Hessian (2.8), all off-diagonal elements of rows and columns corresponding to a fixed parameter are set to zero. At the same time, the diagonal element of the Hessian corresponding to this parameter must be set equal to one. The component of the Lagrange function gradient corresponding to a fixed parameter must be set to zero (see equation (2.7)). It is this
approach to fixing parameters that is used in the considered fitting package. However, this approach
is not unique. For instance, fixed parameters can be excluded programmatically, i.e. one can consider
the Hessian with respect to free parameters only. Both approaches are equivalent, but the latter
approach makes it possible to reduce the size of the Hessian matrix. With the further development
of the fittings package, it is planned to switch to using this mechanism for fixing parameters.

The fitting package also provides an interface for enabling and disabling constraints. This
feature is coming very useful when it is necessary to perform a fit in several hypotheses (see section 3)
that differ from each other only in the set of constraints.

3 Hypotheses

Further, in this work the term hypothesis is often used. By hypothesis we mean a set of all particles,
vertices and constraints involved in the constrained chi-square minimization. The hypothesis also
depends on the state of particles and vertices, since some parameters of these entities can be fixed or
limited. The fitting package discussed in this paper has special classes responsible for describing
these hypotheses. Further, for simplicity, we name the hypotheses according to the particles they
contain, for example, as \( e^+ e^- \rightarrow \gamma \gamma \gamma \). One has to distinguish a hypothesis from a process, since the
same notation is used for both of them. The text of the paper always notes whether the reader is
dealing with the notation of a hypothesis or a process.

The paper also often uses the terms signal and background hypothesis. A hypothesis is called
a signal hypothesis for some process if it describes this process. Otherwise, the hypothesis is a
background hypothesis. There may be several signal hypotheses for some processes. Such hypotheses
can differ from each other, for example, by the set of constraints, as well as by the presence or absence
of lost particles. To show this, let us consider the \( e^+ e^- \rightarrow \eta \pi^+ \pi^- \) process. Since the final
state of this process contains two charged pions and two photons, then hypothesis \( e^+ e^- \rightarrow \pi^+ \pi^- \gamma \gamma \)
is a signal hypothesis for this process. Another signal hypothesis is \( e^+ e^- \rightarrow \eta \pi^+ \pi^- \), \( \eta \rightarrow \gamma \gamma \). This
hypothesis takes into account the intermediate particle \( \eta \). This particle can be taken into account
using the constraint on the two-photon invariant mass.

Some of the hypotheses discussed in the article contain a particle denoted by the letter \( X \). Throughout this paper the designation \( X \) is used for a neutral intermediate particle of unknown mass.
An example of such a hypothesis is \( e^+ e^- \rightarrow X K^+ \pi^- , X \rightarrow \pi^+ \pi^- \), which is a signal hypothesis for
the \( e^+ e^- \rightarrow K_S K^+ \pi^- \), \( K_S \rightarrow \pi^+ \pi^- \) process.

4 Chi-square distribution

As a result of fitting the events of a certain process, one can obtain the distribution of the chi-square
given by equation (1.2). The properties of this distribution need to be known as the chi-square is
often used in selection criteria, and analysis of the chi-square distribution is in some cases useful for
validating kinematic and vertex fitting packages. This raises the question of whether this distribution
is consistent with the chi-squared probability density function. This probability density function
(PDF) has the following form:

\[
f_{\chi^2}(t; \nu) = \begin{cases} 
\frac{1}{\nu^{\nu/2} \Gamma(\nu/2)} t^{\nu/2-1} e^{-t/2}, & t > 0; \\
0, & \text{otherwise},
\end{cases}
\] (4.1)
where the variable \( \nu \) has the meaning of the number degrees of freedom (NDF), and the variable \( t \) has the meaning of the chi-square value.

This section is dedicated to answering the question posed in the previous paragraph. Sections 4.1 and 4.2 discuss how linear and nonlinear constraints affect the properties of the chi-square distribution. In section 4.3, we briefly discuss how this distribution is affected by the non-Gaussian response of the detector. In section 4.4, we describe an approach to the verification of the fitting package based on a Gaussian simulation technique.

4.1 Chi-square in the case of linear constraints

Let us first consider the case where all constraints are linear. Suppose that these constraints have the following form:

\[
\hat{J} y + v = 0,
\]

where \( \hat{J} = \frac{\partial(f_1, \ldots, f_m)}{\partial(y_1, \ldots, y_{m+1})} \) is the Jacobian of constraint functions, \( v \) is some vector, \( \dim v = m \). The Lagrange function in this case has the form:

\[
L(y, \lambda) = (x - \bar{x})^\top \hat{C}^{-1} (x - \bar{x}) + \lambda^\top (\hat{J} y + v),
\]

\( y = x \oplus a \).

After substituting the Lagrange function (4.3) into equation (2.6), this equation will be reduced to the following system of linear equations:

\[
\begin{bmatrix}
2\hat{M} & \hat{J} \\
\hat{J} & 0_{m \times m}
\end{bmatrix} \Delta q = -\Delta g,
\]

where \( \hat{M} = \hat{C}^{-1} \oplus 0_{l \times l} \), \( \Delta q = \Delta y \oplus \Delta \lambda \), \( \Delta y = y - \bar{y} \), \( \Delta \lambda = \lambda - 0 = \lambda \). \( 0_{m \times m} \) and \( 0_{l \times l} \) are zero \( m \) by \( m \) and \( l \) by \( l \) zero matrices, respectively. The vector \( \bar{y} = \bar{x} \oplus \bar{a} \) has the meaning of the vector of initial values of the parameters \( y \). As discussed above, the vector \( \bar{x} \) at the same time has the meaning of the measured values of the parameters \( x \). The vector \( \Delta g \) can be written in the following form:

\[
\Delta g = \nabla_q L(\bar{y}, 0) = \begin{bmatrix}
0 \\
\vdots \\
0
\end{bmatrix} \begin{bmatrix}
0 \\
\vdots \\
0
\end{bmatrix} = \begin{bmatrix}
\Delta f_1 \\
\Delta f_2 \\
\vdots \\
\Delta f_m
\end{bmatrix},
\]

where \( \Delta f_k = f_k(\bar{y}) - f_k(y') = f_k(\bar{y}), f_k(y') = 0 \) and \( y' \) is a local extremum point.

Note that the matrix of the system of equations (4.4) is non-singular only if \( \text{rank}(\hat{J}) = m \). If the last equality is not satisfied, then linearly dependent rows and columns appear in the matrix of the system of linear equations (4.4), i.e. this matrix becomes degenerate. Note that the condition

\[\text{Thus, within the framework of the iterative algorithm given by equation (2.7), the optimization problem (1.4) is solved in one iteration in the case of linear constraints.}\]
rank($\hat{J}$) = $m$ must also hold in the nonlinear case, otherwise the Hessian (2.8) is degenerate. Thus, it makes sense to consider only functionally independent constraints (rank($\hat{J}$) = $m$). In addition, it is easy to show that the inequality $n + l \geq m$ must also hold, otherwise the constraints are functionally dependent: rank($\hat{J}$) $\leq \min(n + l, m)$ = $n + l < m$.

Let us represent the Jacobian as the concatenation of two matrices: $\hat{J} = \begin{bmatrix} \hat{J}_x & \hat{J}_a \end{bmatrix}$. The matrix $\hat{J}_x = \frac{\partial \{f_1, \ldots, f_m\}}{\partial \{x_1, \ldots, x_n\}}$ is the Jacobian of the constraint functions with respect to the parameters $x$, and the matrix $\hat{J}_a = \frac{\partial \{f_1, \ldots, f_m\}}{\partial \{a_1, \ldots, a_l\}}$ is the Jacobian of these functions with respect to the parameters $a$. Taking into account this form of the Jacobian, the system of equations (4.4) can be rewritten as follows:

$$
\begin{bmatrix}
2\hat{C}^{-1} & 0_{n \times l} \\
0_{l \times n} & 0_{l \times l} \\
\hat{J}_x & \hat{J}_a \\
\end{bmatrix}
\Delta q = -\Delta g.
$$

Equation (4.6) can be easily solved by inverting the corresponding block matrix. To do this, one can use the well-known Forbenius formula [25] for block matrix inversion:

$$
\hat{A}^{-1} = \begin{bmatrix} \hat{B} & \hat{C} \\ \hat{D} & \hat{F} \end{bmatrix}^{-1} = \begin{bmatrix}
\hat{B}^{-1} + \hat{B}^{-1} \hat{C} \hat{G}^{-1} \hat{D} \hat{B}^{-1} & -\hat{B}^{-1} \hat{C} \hat{G}^{-1} \\
-\hat{G}^{-1} \hat{D} \hat{B}^{-1} & \hat{G}^{-1} \\
\end{bmatrix},
$$

where $\hat{B}$ is a $n_1 \times n_1$ non-singular matrix, $\hat{F}$ is a $n_2 \times n_2$ matrix, and $\hat{G} = \hat{F} - \hat{D} \hat{B}^{-1} \hat{C}$. Taking into account the last formula, one can obtain the following result:

$$
\begin{bmatrix}
2\hat{C}^{-1} & 0_{n \times l} \\
0_{l \times n} & 0_{l \times l} \\
\hat{J}_x & \hat{J}_a \\
\end{bmatrix}
\begin{bmatrix}
2\hat{C}^{-1} & 0_{n \times l} \\
0_{l \times n} & 0_{l \times l} \\
\hat{J}_x & \hat{J}_a \\
\end{bmatrix}^{-1} = \begin{bmatrix}
\hat{K} \\
\hat{S} \\
\end{bmatrix},
$$

where $\hat{S}$ is the following matrix:

$$
\hat{S} = \begin{bmatrix}
0_{l \times l} & \hat{J}_a \\
\hat{J}_a & 0_{m \times m} \\
\end{bmatrix}.
$$

$\hat{K}$ is the concatenation of matrices $0_{l \times n}$ and $\hat{J}_x$:

$$
\hat{K} = \begin{bmatrix}
0_{l \times n} \\
\hat{J}_x \\
\end{bmatrix}.
$$

The matrix $\hat{G}$ in the case of equation (4.8) is as follows:

$$
\hat{G} = \hat{S} - \frac{1}{2} \hat{K} \hat{C} \hat{K}^\top.
$$
Thus, the solution for the vector $\Delta q$ has the form:

$$
\Delta q = \begin{bmatrix}
\frac{1}{2} \hat{C} - \frac{1}{2} \hat{C} \hat{K} \hat{C}^{-1} \hat{K} \hat{C}
\end{bmatrix}
\begin{bmatrix}
\begin{bmatrix}
0 \\
\vdots \\
0 \\
m
\end{bmatrix}
\end{bmatrix}
\begin{bmatrix}
\begin{bmatrix}
0 \\
\vdots \\
0 \\
m
\end{bmatrix}
\end{bmatrix}
$$

$$
= \begin{bmatrix}
\frac{1}{2} \hat{C} \hat{K} \hat{C}^{-1} \\
\hat{G}^{-1}
\end{bmatrix}
\begin{bmatrix}
\begin{bmatrix}
0 \\
\vdots \\
0 \\
m
\end{bmatrix}
\end{bmatrix}
= \hat{R} \hat{G}^{-1} \hat{P} \Delta f ,
$$

(4.12)

where matrix $\hat{R}$ is such that

$$
\hat{R} = \begin{bmatrix}
\frac{1}{2} \hat{C} \\
\hat{I}_{l+m}
\end{bmatrix},
$$

(4.13)

matrix $\hat{P}$ has the following form:

$$
\hat{P} = \begin{bmatrix}
\hat{0}
\hat{I}_m
\end{bmatrix},
$$

(4.14)

matrix $\hat{I}_k$ is the $k \times k$ identity matrix, $\Delta f = \begin{bmatrix}
\Delta f_1 \\
\Delta f_2 \\
\vdots \\
\Delta f_m
\end{bmatrix}$. The chi-square at the local minimum point $x'$ has the following form:

$$
\chi^2(x') = (x' - \hat{x})^T \hat{C}^{-1} (x' - \hat{x}) = \Delta x^T \hat{C}^{-1} \Delta x = \Delta q^T \hat{Y} \Delta q
$$

$$
= \Delta f^T \hat{P} \hat{G}^{-1} \hat{R} \hat{G}^{-1} \hat{P} \Delta f
$$

$$
= \Delta f^T \left( \frac{1}{4} \hat{\delta} \hat{C} \hat{K} \hat{C}^{-1} \hat{K} \hat{C} \hat{G}^{-1} \hat{P} \hat{G}^{-1} \hat{K} \hat{C} \hat{C}^{-1} \hat{K} \hat{C} \hat{G}^{-1} \hat{P} \right) \Delta f = \Delta f^T \hat{M} \Delta f ,
$$

(4.15)

where $\hat{Y} = \hat{C}^{-1} \oplus \hat{0}_{(l+m) \times (l+m)}$ and $\hat{M} = \frac{1}{4} \hat{P} \hat{G}^{-1} \hat{K} \hat{C} \hat{C}^{-1} \hat{K} \hat{C} \hat{G}^{-1} \hat{P}$.

### 4.1.1 Case $l = 0$

Let us consider first the case where the number of non-measurable parameters is zero, $l = 0$. In this case $\hat{G} = -\frac{1}{2} \hat{J}_x \hat{C} \hat{J}_x^T$. This matrix must be non-singular, otherwise the system of equations (4.6) is unsolvable.

Consider the random vector $\Delta f = \hat{J}_x \hat{x} + v$. Since this vector is linearly related to the random vector $\hat{x} \sim N(\mu_x, \hat{C})$, it is easy to show that it is also normally distributed. The expectation $E[.]$ of this vector has the following form: $E[\Delta f] = \hat{J}_x E[\hat{x}] + v = \hat{J}_x \mu_x + v \equiv 0$. In the previous formula, we used the identity $E[\hat{x}] = \mu_x$. The covariance matrix $\text{Cov}[.]$ for this vector can also be easily
obtained:

\[ \text{Cov} \{ \Delta f \}_{ij} = E[\{\Delta f - \mathbf{0}\}_i (\Delta f - \mathbf{0})_j] = E[(\hat{J}_x \hat{x} + \nu) (\hat{J}_x \hat{x} + \nu)] = (\hat{J}_x)_{ia} (\hat{J}_x)_{ib} E[(\hat{x})_a (\hat{x})_b] + \nu_i (\hat{J}_x \mu_x)_j =
\]

\[ (\hat{J}_x)_{ia} (\hat{J}_x)_{jb} E[(\hat{x} - \mu_x)_a (\hat{x} - \mu_x)_b] = \left( \hat{J}_x \hat{C} \hat{J}_x^\top \right)_{ij}. \]  

(4.16)

In the last equation it is assumed that summation is carried out over the repeated indices \(a\) and \(b\). Indices \(i\) and \(j\) are free matrix indices. Thus we got that \(\Delta f \sim \mathcal{N}(\mathbf{0}, \hat{J}_x \hat{C} \hat{J}_x^\top)\). Since the matrix \(\hat{C} = -\frac{1}{2} \hat{J}_x \hat{C} \hat{J}_x^\top\) is non-singular, the matrix \(\hat{J}_x \hat{C} \hat{J}_x^\top\) is also non-singular. Therefore, the random variable \(\Delta f \top \left( \hat{J}_x \hat{C} \hat{J}_x^\top \right)^{-1} \Delta f\) is distributed according to the chi-square distribution with rank \(\left( \hat{J}_x \hat{C} \hat{J}_x^\top \right) = m\) degrees of freedom.

On the other hand, according to equation (4.15), the following equality holds in the case of \(l = 0\):

\[ \chi^2(x') = \Delta x \top \hat{C}^{-1} \Delta x = \Delta f \top \hat{M} \Delta f = \Delta f \top \left( \hat{J}_x \hat{C} \hat{J}_x^\top \right)^{-1} \Delta f, \]  

(4.17)
i.e. the random variable \(\chi^2(x') = \Delta x \top \hat{C}^{-1} \Delta x\) has the same distribution as the random variable \(\Delta f \top \left( \hat{J}_x \hat{C} \hat{J}_x^\top \right)^{-1} \Delta f\). Thus, in the case of linear constraints and with \(l = 0\), the random variable \(\chi^2(x')\) is distributed according to the probability density function (4.1) with \(m\) degrees of freedom.

### 4.1.2 Case \(l \neq 0\)

Now consider the case when the number of non-measurable parameters is non-zero. The system of linear equations (4.2) describing the constraints can be rewritten as follows:

\[ \hat{J}_x \hat{x} + \hat{J}_a a + \nu = \mathbf{0}. \]  

(4.18)

Since \(\text{rank}(\hat{J}) = m\), the latter system contains \(m\) linearly independent equations. Let us multiply this system of equations on the left by the matrix \(\hat{J}_a \hat{a}\). Since \(\text{rank}(\hat{J}_a \hat{a}) = \text{rank}(\hat{J}_a) = l\), the \(l \times l\) matrix \(\hat{J}_a \hat{a}\) is a full rank matrix, i.e. this matrix is non-singular. Thus, non-measurable variables \(a\) can be expressed from equation (4.18) as follows:

\[ a = - (\hat{J}_a \hat{a})^{-1} \hat{J}_a \hat{J}_x + \nu. \]  

(4.19)

Further, non-measurable parameters can be excluded from equation (4.18) by substituting equation (4.19) back into equation (4.18). As a result of this substitution, one can obtain the following system of equations:

\[ \left( \mathbf{I} - \hat{J}_a (\hat{J}_a \hat{a})^{-1} \hat{J}_a \right) (\hat{J}_x \hat{x} + \nu) = \mathbf{0}. \]  

(4.20)

where \(\mathbf{I}\) is the identity matrix. The system of linear equations (4.20) contains \(m\) equations, but some of the equations in this system are linearly dependent. Since this system was obtained from a system of \(m\) linearly independent equations (4.18) by eliminating \(l\) non-measurable parameters, it contains only \(m - l\) linearly independent equations. Thus the case \(l \neq 0\) can be reduced to the case \(l = 0\). Therefore, by analogy with section 4.1.2, one can conclude that in the case of linear constraints and with \(l \neq 0\), the random variable \(\chi^2(x')\) is distributed according to the probability density function (4.1) with \(m - l\) degrees of freedom.
4.2 Chi-square in the case of nonlinear constraints

In the case of nonlinear constraints, the random vectors $x'$ and $\Delta f$ are expressed in terms of the random vector $\tilde{x}$ nonlinearly. Therefore, in the case of nonlinear constraints, the distribution of the random variable $\chi^2(x') = (x' - \tilde{x})^T \hat{C}^{-1} (x' - \tilde{x})$ is not consistent with probability density function (4.1).

The kinematic and vertex fitting implies extensive use of nonlinear constraints (see sections 6 and 7). Thus, one should expect that the distribution of the random variable $\chi^2(x')$ will not be consistent with the probability density function (4.1) even in the case of the Gaussian response of the detector. However, in the case of many hypotheses, nonlinear effects do not lead to significant distortion of the chi-square distribution. That is, despite the nonlinearity, the distribution of the random variable $\chi^2(x')$ in these cases is well described by the probability density function (4.1). At the same time, in the case of some hypotheses, the nonlinear effects can be so large that the distribution of the random variable $\chi^2(x')$ in these cases is not described by the probability density function (4.1). A detailed discussion of how constraint nonlinearity affects the $\chi^2(x')$ distribution is given in section 9 with examples of Gaussian simulation.

4.3 Chi-square in the case of Monte Carlo simulation and experimental data

It often happens that the uncertainties of some measurable parameters are of a non-Gaussian nature. The reason is the non-Gaussian response of the detector. This behavior of parameter uncertainties can be observed not only in experimental data, but also in Monte Carlo simulation events. In section 4.1, it is shown that the chi-square distribution is described by the probability density function (4.1) only if the parameter uncertainties are Gaussian. Thus, we expect that the chi-square distribution obtained for experimental events or Monte Carlo simulation events will not be consistent with the chi-squared probability density function. Another effect that distorts the chi-square distribution in the case of the CMD-3 experiment is the imperfect calibration of the covariance matrix $\hat{C}$. These effects are well illustrated by Monte Carlo simulation examples in section 8.

Since the Monte Carlo simulation does not perfectly describe the experimental data, the distributions of the random variable $\chi^2(x')$ in the experiment and simulation may be different. Suppose that in the analysis of some physical process, a selection criterion for $\chi^2(x')$ is used. Since the distributions of the random value $\chi^2(x')$ in experiment and simulation are different, the detection efficiencies in simulation and experiment are also different. The detection efficiency in the experiment is usually unknown, so the corresponding efficiency from the simulation is usually used instead. Since these efficiencies are different, this difference must be taken into account using efficiency corrections and finding the corresponding systematic uncertainties. Examples of detection efficiency corrections related to a chi-square selection criterion can be found, for example, in articles [26, 27].

4.4 Gaussian simulation

In sections 4.2 and 4.3, it is discussed that the nonlinearity of the constraints, the non-Gaussian response of the detector and the imperfect calibration of the covariance matrix $\hat{C}$ can lead to distortion

---

2Further, by Monte Carlo simulation we understand the simulation of events of any $e^+e^-$ annihilation process, taking into account the response of a detector.
of the chi-square distribution. However, possible errors in the implementation of the fitting package can also distort this distribution. Therefore, some testing procedure is needed to avoid such errors.

In order to verify the fitting procedure, the authors introduce the technique of Gaussian simulation. This verification technique eliminates the influence of the detector’s non-Gaussian response and the imperfect covariance matrix calibration on the chi-square distribution. The idea of Gaussian simulation is to redraw measurable parameters according to the multivariate normal distribution. A detailed description of this procedure is given in the next paragraph.

Let us consider one event of some process from an experiment or Monte Carlo simulation and assume that we have made a fit of this event under the corresponding signal hypothesis. The parameter values obtained after this fitting will be such that all kinematic and vertex constraints are satisfied. Therefore, it is possible to use the measurable parameters found as a result of this fitting as the corresponding mean values for random generation of similar events according to the multivariate normal distribution. The covariance matrix known from the initial event can be used as the covariance matrix corresponding to this distribution. After generating a sufficient number of events, the fitting procedure can be applied again to these events. As a result, one can obtain a chi-square distribution corresponding to uncertainties that are of a Gaussian nature. Examples of Gaussian simulation are discussed in section 9.

If the chi-square distribution obtained as a result of the Gaussian simulation turns out to be distorted, then this is due either to the nonlinearity of the constraints or to errors in the fitting package. The fact that in some cases the distortion of the distribution occurs due to the nonlinearity of the constraints can also be verified. A related discussion is given in section 9.2.

5 Vertices

In this paper, vertices are considered as separate entities. The reason for considering vertices as separate entities is that each vertex is always shared by some set of particles. Each vertex has three coordinates. The kinematic and vertex fitting package includes vertex classes corresponding to coordinates in Cartesian and cylindrical coordinate systems. The user can add vertices with custom parametrization corresponding to a different coordinate system. To do this, one needs to implement a new vertex class inherited from the base vertex class. The base vertex class contains virtual abstract methods that return the Cartesian coordinates of a vertex depending on its three parameters. This class also contains virtual abstract methods for obtaining gradients and Hessians of the Cartesian coordinates, depending on the vertex parametrization. In order to create a custom vertex class, the user has to implement the above abstract methods. In the case of the vertex corresponding to the Cartesian coordinate system, the implementation of the methods listed above is trivial. For example, functions that return the Cartesian coordinates of a vertex have the following form:

\[
x_v(a^{(v)}) = x^{(v)}, \\
y_v(a^{(v)}) = y^{(v)}, \\
z_v(a^{(v)}) = z^{(v)},
\]

(5.1)

\[
a^{(v)} = \begin{bmatrix} x^{(v)} \\ y^{(v)} \\ z^{(v)} \end{bmatrix},
\]
where $x_v$, $y_v$ and $z_v$ are functions that return the Cartesian coordinates of a vertex depending on the vector $a^{(v)}$ of its parameters $x^{(v)}$, $y^{(v)}$ and $z^{(v)}$, which are themselves the vertex Cartesian coordinates. In the case of a vertex parametrized according to the cylindrical coordinate system, the functions returning Cartesian coordinates are as follows:

$$
\begin{align*}
x_v(a^{(v)}) &= \rho^{(v)} \cos \phi^{(v)}, \\
y_v(a^{(v)}) &= \rho^{(v)} \sin \phi^{(v)}, \\
z_v(a^{(v)}) &= z^{(v)},
\end{align*}
$$

where parameters $\rho^{(v)}$, $\phi^{(v)}$ and $z^{(v)}$ are the vertex coordinates in the cylindrical coordinate system. Here and below, we do not present gradients and Hessians for vertex coordinates, particle four-momenta and trajectories, since the procedure for calculating them is well known. The reader can find them in the fitting package. Links to the relevant repositories are provided in section 10.

In the case of the CMD-3 experiment, vertex parametrization (5.1) is most often used. Such vertex parametrizations as (5.2) can be used in specific cases. For example, a vertex with parametrization (5.2) can be used if it is required to set limits for the $\rho^{(v)}$ parameter. Such a need may arise, for example, if one wants to reconstruct a vertex of a photon conversion on the beam pipe.

All coordinates of the $e^+e^-$ interaction vertex are considered as free measurable parameters. These parameters make an additional contribution to the chi-square given by equation (1.2). The part of the covariance matrix $\hat{C}$ corresponding to the $e^+e^-$ interaction vertex is diagonal. Its diagonal elements are the squares of the corresponding sizes of the interaction region. In the case of the CMD-3 experiment a typical longitudinal (along the beam axis) size of the interaction region is about $2.5–3.0$ cm. A typical transverse size of the interaction region is approximately $60–100$ $\mu$m.

All coordinates of decay vertices are treated as free non-measurable parameters and do not contribute to the chi-square.

6.1 Initial pseudo-particle

An initial pseudo particle is used in order to set the total four-momentum of all particles present in a certain hypothesis. In the case of the VEPP-2000 collider, the energies of the initial electrons and
positrons are the same with high accuracy, and the directions of their motion are opposite. For this reason, the four-momentum parametrization for the initial pseudo-particle in the case of the CMD-3 experiment can be written in the following form:

\[ P^{(i)}_{\text{pseudo}} = (E_{\text{c.m.}}, 0), \]  

where \( E_{\text{c.m.}} = 2E_{\text{beam}} \) is the center-of-mass energy, \( E_{\text{beam}} \) is the beam energy, \( i \) is the particle index. In the case of the CMD-3, the energy is set constant at each event. In principle, the user can release this parameter (make it free and measurable). In the case of the VEPP-2000, the typical beam energy spread is less than 1 MeV. For this reason, using a non-constant center-of-mass energy \( E_{\text{c.m.}} \) in the case of the CMD-3 experiment will not have a significant effect on the fit result.

### 6.2 Charged particles and photons

The parametrizations of a charged particle and photon were placed in a separate subsection, since the authors consider these parametrizations to be detector-dependent. For example, in the case of a photon, the detector-dependent part of the parametrization is contained at least in the description of the photon conversion point inside a calorimeter. In the case of the CMD-3 experiment, a cylindrical coordinate system is used to set the photon conversion point. The parametrization of a charged particle in the case of the CMD-3 experiment is described in section 6.2.1. This parametrization may also have some differences in the case of a different experiment. For example, in the case of the charged particle parametrization described in section 6.2.1, a constant magnetic field at any point of the drift chamber is used, i.e. the change in the magnetic field near the detector end-caps is not taken into account. This is due to the fact that the charged particle tracks themselves are reconstructed in the case of the CMD-3 experiment under the assumption of a uniform magnetic field. The inhomogeneity of the magnetic field in the CMD-3 experiment is usually taken into account by introducing a corresponding systematic uncertainty. Since the package of kinematic and vertex fitting discussed in this paper uses the parameters of already reconstructed tracks, it makes no sense to take into account the inhomogeneity of the magnetic field in this package. Section 6.2.1 discusses only the parametrization of a final charged particle. The parametrization of an intermediate charged particle is discussed in section 6.3.2.

#### 6.2.1 Charged particle

In the case of the experiment with the CMD-3 detector, the parametrizations of four-momenta and trajectories of final charged particles depend on five measurable parameters. These parameters are listed below:

- \( p_{\perp}^{(i)} \) is the radial component of momentum, i.e. the momentum component perpendicular to the magnetic field;
- \( \rho_c^{(i)} \) is the distance between the beam axis and track of a charged particle;
- \( z_c^{(i)} \) is \( z \)-coordinate of the track point closest to the beam axis;

---

3. This effect is within the detector resolution.
4. Furthermore, it is assumed that this field is directed along the \( Z \)-axis.
5. In this article, by track we mean the track of some charged particle, reconstructed using hits in the drift chamber.
• $\phi_c^{(i)}$ is the axial angle corresponding to the track point closest to the beam axis;

• $\theta_c^{(i)}$ is the polar angle of a charged particle momentum.

The parametrizations of a charged particle four-momentum and trajectory depend also on one non-measurable parameter $c t_{\text{out}}^{(i)}$. This parameter has a meaning of the distance along the charged particle trajectory from the charged particle origin vertex to the track point closest to the beam axis.

The parametrization of a charged particle four-momentum $\mathbf{P}_{c,\text{out}}^{(i)} = \left( E_c^{(i)}, \mathbf{p}_{c,\text{out}}^{(i)} \right)$ has the following form:

\[
E_c^{(i)} = \sqrt{p_\perp^{(i)} \left( 1 + \cot^2 \theta_c^{(i)} \right) + m_c^{(i)}},
\]

\[
p_{c,\text{out}}^{(i)} = \begin{bmatrix}
p_\perp^{(i)} \cos \left( \frac{\omega_c^{(i)}}{c} c t_{\text{out}}^{(i)} - \phi_c^{(i)} \right) \\
-p_\perp^{(i)} \sin \left( \frac{\omega_c^{(i)}}{c} c t_{\text{out}}^{(i)} - \phi_c^{(i)} \right) \\
p_\perp^{(i)} \cot \theta_c^{(i)}
\end{bmatrix},
\]

where the constant $q_c^{(i)}$ is the particle charge, measured in elementary charges; the constant $m_c^{(i)}$ is the mass of a charged particle, measured in GeV/$c^2$; the constant $B$ is the magnetic field, measured in T. Fraction $\omega_c^{(i)}/c$ is the fraction of cyclotron frequency to the speed of light; the constant $\kappa_c \approx 2.9979 \times 10^{-3}$ is the proportionality factor.\footnote{The value of this factor is determined by the speed of light and units of measurement of the quantities involved in equation (6.2). Here and below, we assume that energy is measured in GeV, magnetic field in T, time in seconds, and distance in cm.}

The parametrization of a charged particle trajectory $\mathbf{r}_{c,\text{out}}^{(i)} = \left[ x_{c,\text{out}}^{(i)}, y_{c,\text{out}}^{(i)}, z_{c,\text{out}}^{(i)} \right]^T$ is given by the following equation:

\[
x_{c,\text{out}}^{(i)} = x_{\text{beam}} + \left( R_c^{(i)} - \rho_c^{(i)} \right) \sin \phi_c^{(i)} + R_c^{(i)} \sin \left( \frac{\omega_c^{(i)}}{c} c t_{\text{out}}^{(i)} - \phi_c^{(i)} \right),
\]

\[
y_{c,\text{out}}^{(i)} = y_{\text{beam}} - \left( R_c^{(i)} - \rho_c^{(i)} \right) \cos \phi_c^{(i)} + R_c^{(i)} \cos \left( \frac{\omega_c^{(i)}}{c} c t_{\text{out}}^{(i)} - \phi_c^{(i)} \right),
\]

\[
z_{c,\text{out}}^{(i)} = z_c^{(i)} + \frac{p_\perp^{(i)} \cot \theta_c^{(i)} c t_{\text{out}}^{(i)}}{E_c^{(i)}},
\]

\[
R_c^{(i)} = \frac{p_\perp^{(i)}}{q_c^{(i)} B \kappa_c},
\]

where $x_{\text{beam}}$ and $y_{\text{beam}}$ are the $x$- and $y$- coordinates of the beam axis, respectively. In the case of the CMD-3 detector these coordinates considered as constants, since they are known with a high accuracy.
6.2.2 Photon

The parametrization of the photon four-momentum depends on seven parameters. Four of the seven parameters belong to the photon. The remaining three parameters are the coordinates $r_{\text{origin}}$ of its origin vertex. One of the four photon parameters is its energy $E^{(i)}_\gamma$. The remaining three photon parameters are the coordinates $\rho^{(i)}_\gamma$, $\phi^{(i)}_\gamma$ and $z^{(i)}_\gamma$ of the conversion point in a cylindrical coordinate system. The photon four-momentum $P^{(i)}_\gamma = (E^{(i)}_\gamma, p^{(i)}_\gamma)$ is given by the following equation⁷:

\[
p^{(i)}_\gamma = E^{(i)}_\gamma \frac{r^{(i)}_{\text{conv}} - r_{\text{origin}}}{\|r^{(i)}_{\text{conv}} - r_{\text{origin}}\|},
\]

\[
r^{(i)}_{\text{conv}} = \begin{bmatrix}
\rho^{(i)}_\gamma \\
\phi^{(i)}_\gamma \\
z^{(i)}_\gamma
\end{bmatrix}.
\]

Since equation (6.4) already requires a photon to fly from the origin vertex $r_{\text{origin}}$ to the conversion point $r^{(i)}_{\text{conv}}$, there is no need for a separate parametrization of a photon trajectory.

6.3 Intermediate particles

Along all intermediate particles in the case of the CMD-3 experiment, $K_S$ mesons require special attention due to their long lifetime. Parametrization of neutral intermediate particles, such as $K_S$ mesons, is given in section 6.3.1. Note that in the case of such particles as $\pi^0$ or $\eta$ mesons, the parametrization described in section 6.3.1 should not be used. In the terms of kinematic and vertex fitting, these particles decay at the origin vertex due to their short lifetime. To take these particles into account, it suffices to require a mass constraint on their decay products.

Due to the small radius (30 cm) of the drift chamber in the CMD-3 experiment, it practically makes no sense to consider hypotheses with intermediate charged particles. The probability of a charged $\pi$-meson or $K$-meson decay inside the drift chamber is low. There is always an odd number of charged particles among the decay products of a charged particle in order to conserve charge. The tracks of an initial charged particle and the charged products of its decay are well reconstructed if the decay vertex is near the center of the drift chamber radius, since in this case the probability that all tracks will have a sufficient number of hits is the highest. Thus, the number of charged particle decay events in which kinematic and vertex fitting can be successfully used is additionally limited by the track reconstruction efficiency. However, intermediate charged particles have been implemented in the discussed kinematic and vertex fitting package. In addition, it was verified that the use of these intermediate particles makes it possible, for example, to find events with $\pi^+ \rightarrow \mu^+ \nu_\mu$ decay⁸.

The parametrization of intermediate charged particles is discussed in detail in section 6.3.2.

6.3.1 Intermediate neutral particle

The parametrization of the intermediate neutral particle depends on seven parameters. Four parameters belong to this particle. The three remaining parameters are the coordinates $r_{\text{origin}}$ of

---

⁷In this article, by norm we always mean the $\ell^2$-norm: $\|b\| = \sqrt{\sum_{i=1}^{\dim b} b_i^2}$.

⁸Since the detection of intermediate charged particles in the case of CMD-3 is difficult due to the small size of the drift chamber, the authors do not further consider examples of hypotheses with such particles. The authors hope to study in more detail the possibility of using intermediate charged particles 6.3.2 in other experiments in the future.
its origin vertex. The first three parameters of an intermediate neutral particle are the Cartesian components of its momentum

\[ p_{\text{int. n.}}^{(i)} = \begin{bmatrix} p_x^{(i)} \\ p_y^{(i)} \\ p_z^{(i)} \end{bmatrix}, \]  

(6.5)

the last parameter \( \xi^{(i)} \) is proportional to the time interval between the particle origin and decay. The four-momentum parametrization for an intermediate neutral particle has the following form:

\[ p_{\text{int. n.}}^{(i)} = \sqrt{m_{\text{int. n.}}^{(i)} + p_{\text{int. n.}}^{(i)}^2}, \]  

(6.6)

where the constant \( m_{\text{int. n.}}^{(i)} \) is the particle mass. Trajectory parametrization of an intermediate neutral particle can be written as follows:

\[ r_{\text{int. n.}}^{(i)} = r_{\text{origin}} + \xi^{(i)} p_{\text{int. n.}}^{(i)} . \]  

(6.7)

All parameters \( p_x^{(i)} , p_y^{(i)} , p_z^{(i)} \) and \( \xi^{(i)} \) of an intermediate neutral particle are considered to be non-measurable.

### 6.3.2 Intermediate charged particle

The parametrizations of four-momenta and the trajectories of intermediate charged particles are very similar to the corresponding parametrizations in the case of final charged particles 6.2.1. In the case of an intermediate charged particle, the parametrizations of the output four-momentum (6.2) and the output trajectory (6.3) with respect to the origin vertex are exactly the same as for a final charged particle. The parametrizations of the input four-momentum and the input trajectory with respect to the decay vertex differ from the parametrizations (6.2) and (6.3) by replacing the parameter \( c t_{\text{out}}^{(i)} \) with the parameter \( c t_{\text{in}}^{(i)} \). As a result of this replacement, the parametrization of the input four-momentum \( p_{\text{c, in}}^{(i)} = (E_c^{(i)} , p_c^{(i)}) \) of an intermediate charged particle has the following form:

\[ E_c^{(i)} = \sqrt{p_\perp^{(i)}^2 (1 + \cot^2 \theta_c^{(i)}) + m_c^{(i)}}, \]

\[ p_{\text{c, in}}^{(i)} = \begin{bmatrix} p_\perp^{(i)} \cos \left( \frac{\omega_c^{(i)}}{c} c t_{\text{in}}^{(i)} - \phi_c^{(i)} \right) \\ -p_\perp^{(i)} \sin \left( \frac{\omega_c^{(i)}}{c} c t_{\text{in}}^{(i)} - \phi_c^{(i)} \right) \\ p_\perp^{(i)} \cot \theta_c^{(i)} \end{bmatrix}, \]  

(6.8)

\[ \frac{\omega_c^{(i)}}{c} = \frac{q_c^{(i)} B \kappa_c}{E_c^{(i)}}, \]
The parametrization of input trajectory can be written as follows:

\[
\begin{align*}
    x_{\text{c, in}}^{(i)} &= x_{\text{beam}} + \left( R_{\text{c}}^{(i)} - \rho_{\text{c}}^{(i)} \right) \sin \phi_{\text{c}}^{(i)} + R_{\text{c}}^{(i)} \sin \left( \frac{\omega_{\text{c}}^{(i)}}{c} ct_{\text{in}}^{(i)} - \phi_{\text{c}}^{(i)} \right), \\
    y_{\text{c, in}}^{(i)} &= y_{\text{beam}} - \left( R_{\text{c}}^{(i)} - \rho_{\text{c}}^{(i)} \right) \cos \phi_{\text{c}}^{(i)} + R_{\text{c}}^{(i)} \cos \left( \frac{\omega_{\text{c}}^{(i)}}{c} ct_{\text{in}}^{(i)} - \phi_{\text{c}}^{(i)} \right), \\
    z_{\text{c, in}}^{(i)} &= z_{\text{c}}^{(i)} + \frac{p_{\perp}^{(i)}}{q_{\text{c}}^{(i)}} \cot \theta_{\text{c}}^{(i)} \frac{ct_{\text{in}}^{(i)}}{E_{\text{c}}^{(i)}}, \\
    R_{\text{c}}^{(i)} &= \frac{p_{\perp}^{(i)}}{q_{\text{c}}^{(i)} B_{\text{c}}},
\end{align*}
\]

(6.9)

The parameter \(ct_{\text{in}}^{(i)}\) in equations (6.8) and (6.9) has the meaning of the distance along the charged particle trajectory from the track point closest to the beam axis to the decay vertex of this particle. Like the parameter \(ct_{\text{out}}^{(i)}\), the parameter \(ct_{\text{in}}^{(i)}\) is a non-measurable parameter.

Practically, final 6.2.1 and intermediate 6.3.2 charged particles are implemented as a single class in the kinematic and vertex fitting package. In the case of a final charged particle the parameter \(ct_{\text{in}}^{(i)}\) is fixed, and the parameterizations of the input four-momentum and the input trajectory given by equations (6.8) and (6.9) are not used.

6.4 Lost particles

An important requirement for the kinematic and vertex fitting package is the ability to perform a fit with the hypothesis of lost particles.\(^9\) It should be noted that for a lost particle it makes sense to specify only a four-momentum, while specifying the trajectory of such a particle is meaningless. It should also be noted that the four-momentum parameterization is different for massive and massless particles. The four-momentum parameterization of a massive lost particle is discussed in section 6.4.1, while the parameterization in the case of a massless lost particle is given in section 6.4.2.

The parameters of the lost particles are non-measurable as they are not measured directly with a detector. In the case of the discussed package, this is achieved by specifying zero inverse covariance matrices corresponding to the parameters of these particles. In some cases the parameterizations from sections 6.4.1 and 6.4.2 can also be used for detected particles. It is assumed that the parametrization of the trajectories of such particles is of no interest, and the corresponding inverse covariance matrices are non-zero.

6.4.1 Massive particle

The parametrization of a massive lost particle has the following form:

\[
\mathcal{P}_{m \neq 0}^{(i)} = (\sqrt{m^{(i)} + \mathcal{P}^{(i)}}, \mathcal{P}^{(i)}) = \begin{bmatrix} p_{x}^{(i)} \\ p_{y}^{(i)} \\ p_{z}^{(i)} \end{bmatrix},
\]

(6.10)

where the parameters \(p_{x}^{(i)}, p_{y}^{(i)}\) and \(p_{z}^{(i)}\) are the momentum components of the particle, and the constant \(m^{(i)} \neq 0\) is its mass.

\(^9\)By lost particles, we mean final particles that are not detected.
6.4.2 Massless particle

The four-momentum parametrization of a lost massless particle cannot be chosen in the form (6.10) with \( m^{(i)} = 0 \). The disadvantages of the parametrization (6.10) in the case of a massless particle is that the denominators of the derivatives of this parametrization become very small if \( \mathbf{p}^{(i)} \) is close to zero. For this reason, in the case of a lost massless particle, a slightly different four-momentum parametrization is used. This parametrization has the following form:

\[
\mathbf{p}^{(i)}_{m=0} = (E^{(i)}, E^{(i)} \mathbf{n}^{(i)}), \quad \mathbf{n}^{(i)} = \begin{bmatrix}
\sin \theta^{(i)} \cos \phi^{(i)} \\
\sin \theta^{(i)} \sin \phi^{(i)} \\
\cos \theta^{(i)}
\end{bmatrix},
\]

where the parameter \( E^{(i)} \) is the particle energy, the parameters \( \theta^{(i)} \) and \( \phi^{(i)} \) are the polar and axial angles, respectively. These angles determine the direction of the particle momentum \( \mathbf{p}^{(i)} = E^{(i)} \mathbf{n}^{(i)} \).

7 Constraints

The discussed kinematic and vertex fitting package uses three different kinds of constraints. These are energy-momentum conservation constraints, vertex constraints and mass constraints.

The energy-momentum conservation constraints are imposed on a certain set of particles, and are divided into four different constraints: one constraint is needed for energy conservation, the other three are for momentum conservation. A detailed discussion of the energy-momentum conservation constraints is given in section 7.1.

The mass constraint is imposed on a certain set of particles in order to require that the invariant mass of these particles be equal to a certain value. The detailed description of the mass constraint is given in section 7.2.

The vertex constraints are imposed on some individual particles at the vertices of their origin and/or decay. These constraints require that the particle’s trajectory pass through the origin vertex and the decay vertex. A detailed discussion of the vertex constraints is given in section 7.3.

7.1 Energy-momentum conservation constraints

The energy-momentum conservation constraints have the following form:

\[
\sum_{a \in \mathcal{S}^{(j)}_+} \mathbf{p}^{(a)}_{\text{output}} - \sum_{b \in \mathcal{S}^{(j)}_-} \mathbf{p}^{(b)}_{\text{input}} = 0, \tag{7.1}
\]

where the sets \( \mathcal{S}^{(j)}_\pm \) are the sets of indices corresponding to the particles involved in the considered constraints, \( \mathbf{p}^{(a)}_{\text{output}} \) is the output four-momentum of the \( a \)-th particle, \( \mathbf{p}^{(b)}_{\text{input}} \) is the input four-momentum of the \( b \)-th particle. The set \( \mathcal{S}^{(j)}_- \) corresponds\(^\text{10}\) to particles before interaction or decay, while the set \( \mathcal{S}^{(j)}_+ \) corresponds to particles that are products of this interaction or decay. The index \( j \) in \( \mathcal{S}^{(j)}_\pm \) means the index of particle sets on which the constraints (7.1) are imposed. Thus, it is implied that, in the same hypothesis, the constraints given by equation (7.1) can be imposed on several different sets of particles. This circumstance is caused by the fact that the energy-momentum

\(^{10}\)Taking into account section 6.1, it can be concluded that the cardinality of the set \( \mathcal{S}^{(j)}_- \) is equal to one (as rule).
conservation laws often need to be written at each vertex. Thus, the sets $\mathcal{S}_k^{(j)}$ usually correspond to particles having a ($j$-th) common vertex. However, in some cases described below, this statement is not true. Note also that equation (7.1) is considered as four different constraints: one constraint is for energy conservation, and remaining three constraints are for conservation of three-dimensional momentum components. The discussed kinematic and vertex fitting package allows using all these constraints simultaneously, as well as only a part of them.

The energy-momentum constraints given by equation (7.1) are schematically shown in the figure 1 when one particle decays into two other particles. The vertex of the decay tree shown in this figure corresponds to one of the components ($\alpha = x, y, z, t$) of the four-momentum conservation constraints. The edges of this tree correspond to the particles on which the constraints are imposed. The solid edge corresponds to the initial particle, while the dashed edges correspond to its decay products. Equation (7.1) in the case of the example shown in figure 1 takes the following form:

$$-P_\text{input}^{(1)} + P_\text{output}^{(2)} + P_\text{output}^{(3)} = 0.$$  

Figure 2 shows an example of the energy-momentum conservation constraint tree in the case of the $e^+e^- \rightarrow K_S^0K^+\pi^-$, $K_s \rightarrow \pi^+\pi^-$ hypothesis. Label $e^+e^-$ denotes the initial pseudo-particle (see section 6.1) with the input four-momentum $P_\text{input}^{(e^+e^-)} = (2E_{\text{beam}}, 0)$. Constraint $[P_\alpha^{(1)}]$ ensures the conservation of the $\alpha$-th four-momentum component at the $e^+e^-$ interaction vertex, while constraint $[P_\alpha^{(2)}]$ does this at the $K_S$ meson decay vertex. According to the tree shown in figure 2, the intermediate particle $K_S$ is included in the four-momentum conservation constraints. In particular, this particle is included in the energy conservation constraint, which leads to the fact that the invariant mass of the decay products of this particle turns out to be fixed on the $K_S$ meson mass after the fit.

Sometimes it becomes necessary to use a hypothesis in which the invariant mass of decay products is not fixed after the fit. This result can be achieved if the momentum conservation constraints are applied to all particles, including intermediate ones, while the energy conservation constraint is applied only to the initial and final particles. Examples of energy-momentum conservation
Figure 2. The energy-momentum conservation constraint tree corresponding to the hypothesis $e^+e^- 	o K_S K^+ \pi^-$, $K_S \to \pi^+ \pi^-$. Label $e^+e^-$ in the tree denotes the initial pseudo-particle with the input four-momentum $P^{(e^+e^-)}_{\text{input}} = (2E_{\text{beam}}, \mathbf{0})$. The constraint $[P^{(1)}_{\alpha}]$ ensures the conservation of the $\alpha$-th four-momentum component at the $e^+e^-$ interaction vertex, $\alpha = x, y, z, t$. Constraint $[P^{(2)}_{\alpha}]$ ensures the conservation of the $\alpha$-th four-momentum component at the $K_S$ meson decay vertex. Since the considered hypothesis contains two $\pi^-$ mesons, they are denoted as $\pi^-_1$ and $\pi^-_2$.

Constraint trees for such a hypothesis are shown in figure 3. This hypothesis corresponds to the $e^+e^- \to X K^+ \pi^-$, $X \to \pi^+ \pi^-$ process, where particle $X$ is unknown. Two trees are shown in the figure 3. The tree shown in figure 3(a) describes how three momentum conservation constraints are applied, while the tree in figure 3(b) describes how the energy conservation constraint is applied. An example of using this hypothesis can be found in section 8.1.

7.2 Mass constraint

The mass constraint is used in order to require that the invariant mass of some set of particles be equal to a certain value. This constraint can be written as follows:

$$ \left( \sum_{i \in S_{+}^{(j)}} P_{\text{output}}^{(i)} \right)^2 - m_{\text{target}}^2 = 0, \quad (7.2) $$

where $S_{+}^{(j)}$ is some set of particle indices, $P_{\text{output}}^{(i)}$ is the output four-momentum of the $i$-th particle, and $m_{\text{target}}$ is the target value of the invariant mass.

On the one hand, the introduction of mass constraint (7.2) is redundant, since the energy-momentum conservation constraints 7.1 together with the use of intermediate particles 6.3 make it possible to achieve the same result. Indeed, suppose that we are dealing with a hypothesis in which the $\pi^0$ meson decays into two photons. Since the $\pi^0$ meson lifetime is short, the $\pi^0$ meson decay vertex is close to its origin vertex. The resolution of the (CMD-3) detector does not allow

---

$^a$Particle $X$ is unknown in the sense that the value of its mass does not affect the result of the fitting. This property of particle $X$ is determined not by its parametrization, but by the specific configuration of energy-momentum conservation constraints. As for the intermediate particle $K_S$, the parameterizations described in section 6.3.1 are used for particle $X$. Thus, the particle $K_S$ can be used as a particle $X$. The designation $X$ is introduced only to emphasize that the mass of such a particle does not affect anything.
Figure 3. The momentum (a) and energy (b) conservation constraint trees corresponding to the hypothesis $e^+e^- \rightarrow XK^+\pi^-$, $X \rightarrow \pi^+\pi^-$, where $X$ is unknown intermediate massive particle. Vertices $[p_k^{(1)}]$ and $[p_k^{(2)}]$ of the first tree represent the three-dimensional $(k = x, y, z)$ momentum conservation constraints at the $e^+e^-$ interaction vertex and at the $X \rightarrow \pi^+\pi^-$ decay vertex, respectively. This tree contains all particles, including intermediate ones. Vertex $[E^{(3)}]$ of the second tree corresponds to the energy conservation constraint. This tree contains all particles except intermediate ones.

distinguishing one vertex from another. Thus, in terms of the kinematic and vertex fitting, the $\pi^0$ meson decays at the origin vertex. In this case, the use of vertex constraints 7.3 for the intermediate $\pi^0$ meson does not make sense. The parameter $\xi^{(i)}$ of the intermediate $\pi^0$ meson will not be used in this case and can be fixed (see equation (6.7)). However, eight energy-momentum conservation constraints can be imposed on the intermediate $\pi^0$ meson in this case. Four constraints correspond to the laws of energy-momentum conservation at the origin vertex, and four others correspond to such laws at the decay vertex.

On the other hand, the use of four-momentum parametrizations from section (6.3) in the case of short-lived intermediate particles is not justified because it leads to a needless increase in the number of minimization parameters. In addition to the parameters of the intermediate particle, eight more Lagrange multipliers will be involved in the minimization procedure in this case. If, however, a short-lived intermediate particle is not introduced directly and the mass constraint (7.2) is used, then only one additional Lagrange multiplier corresponding to this constraint will be involved in the minimization procedure. Thus, the conclusion is that in the case of short-lived intermediate particles, it is better to use the mass constraint (7.2) if necessary.
7.3 Vertex constraints

Vertex constraints are used to require particle trajectories to pass through origin and decay vertices. The vertex constraints that require a particle to fly out of its origin vertex have the following form:

\[ r_{\text{output}}^{(i)} - r_{\text{origin}} = 0, \]  

(7.3)

where \( r_{\text{output}}^{(i)} \) is the parametrization of the output trajectory of the \( i \)-th particle, i.e. the trajectory corresponding to the escape of \( i \)-th particle from its origin vertex \( r_{\text{origin}} \). The vertex constraints, which require a particle to fly to its decay vertex, can be written as follows:

\[ r_{\text{input}}^{(i)} - r_{\text{decay}} = 0, \]  

(7.4)

where \( r_{\text{input}}^{(i)} \) is the parametrization of the input trajectory of the \( i \)-th particle, i.e. the trajectory along which the particle flies into its decay vertex \( r_{\text{decay}} \).

Among all particle kinds discussed in section 6, there are particle kinds for which trajectories are not defined. Such particle kinds are the initial pseudo particle from section 6.1, the lost particles from section 6.4, and the photon from section 6.2.2. However, it should be noted that the photon flies from its origin vertex to its conversion point according to the parametrization of its momentum (6.4). For the final charged particle 6.2.1, only the output trajectory (6.3) is given, while for the intermediate charged particle 6.3.2, both the output (6.3) and input (6.9) trajectories are given. In the case of the discussed package of kinematic and vertex fitting, only the input trajectory (6.7) is specified for the neutral intermediate particle 6.3.1.

8 Examples

This section provides examples of applying the kinematic and vertex fitting package in various hypotheses to the simulated\(^\text{12}\) events of various \( e^+e^- \) annihilation processes. In all examples considered in this section, with the exception of the example described in subsection 8.1.4, simulations are done taking into account initial state radiation (ISR). Since the kinematic hypotheses considered in the article do not contain ISR photons, they are not strict signal hypotheses if they are applied to events with ISR photons. As a result, some of the distributions given in this section are distorted. However, the distortion of chi-square distributions is not fully explained by neglecting ISR photons in kinematic hypotheses. Verification of this statement is given in subsection 8.1.4.

Among the examples below, hypotheses with intermediate long-lived neutral particles (\( K_S \) mesons), hypotheses containing charged particles only, hypotheses containing photons only, hypotheses containing both charged particles and photons, as well as hypotheses with lost particles are discussed in detail. The following subsections present mainly the results of applying the fitting algorithm under signal hypotheses. However, in some cases, examples of the fitting under background hypotheses are also given.

In some cases, the kinematic and vertex fitting procedure includes several fits per event. This is due to the particle combinatorics. For example, it may turn out that there are several ways to match

\(^{12}\text{By simulated events in this section, we mean the events of Monte Carlo simulation, which takes into account the response of the CMD-3 detector.}\)
tracks from the drift chamber with charged particles of some hypothesis. Further, the set of all fits in an event is called the fitting procedure. If at least one of the fits in an event converges to a local minimum, then the corresponding fitting procedure is said to have converged to a local minimum. If none of the fits converge, then the fitting procedure is said to have failed. If among all convergent fits there are only those that have converged to a local maximum, then the fitting procedure is said to have converged to a local maximum. The figures given in this section correspond only to those events in which the kinematic and vertex fitting procedure converged to a local minimum.

8.1 Hypotheses $e^+e^- \rightarrow XK \pi^\pm$, $X \rightarrow \pi^+\pi^-$

8.1.1 Description of the $e^+e^- \rightarrow XK\pi^\pm$, $X \rightarrow \pi^+\pi^-$ hypotheses

Section 8.1 presents the results of applying the package of kinematic and vertex fitting under two charged conjugate hypotheses $e^+e^- \rightarrow XK^\mp\pi^\mp$, $X \rightarrow \pi^+\pi^-$, where $X$ is an unknown intermediate neutral particle decaying into two charged pions: $X \rightarrow \pi^+\pi^-$. These hypotheses were developed in order to select the events of the $e^+e^- \rightarrow KS K^\mp \pi^\mp$ process. The particle $X$ is called unknown because the energy conservation constraints are not applied at the $e^+e^-$ interaction vertex and the decay vertex. The energy conservation constraint is applied only to final and initial particles in these hypotheses. Therefore, the mass of the particle $X$ is not contained in the Lagrange function (2.1) and, as a consequence, does not participate in the constrained minimization of the chi-square function (1.2). For this reason, the invariant mass of the $X \rightarrow \pi^+\pi^-$ decay products, calculated using parameters obtained from the fitting is not fixed (e.g. on the $K_S$ mass). For more details, see sections 6.3.1, 7.1 and figure 3.

The hypotheses under discussion have the following structure.

- Each hypothesis has two vertices. The first vertex is the $e^+e^-$ interaction vertex, the second vertex is the decay vertex of the particle $X$. All coordinates of the $X \rightarrow \pi^+\pi^-$ decay vertex are free non-measurable parameters. All coordinates of the $e^+e^-$ interaction vertex are considered as free measurable parameters and contribute to the chi-square.

- The hypothesis $e^+e^- \rightarrow XK^\mp\pi^\mp$, $X \rightarrow \pi^+\pi^-$ requires the presence of four final particles and one intermediate particle $X$. The intermediate particle $X$ originates at the $e^+e^-$ interaction vertex and decays at its decay vertex into two final particles $\pi^+$ and $\pi^-$. Another final $\pi^-$ comes from the $e^+e^-$ interaction vertex. The final $K^+$ also originates at this vertex. The charge conjugate hypothesis $e^+e^- \rightarrow XK^-\pi^+$ differs from the $e^+e^- \rightarrow XK^+\pi^-$ hypothesis in that $K^-$ and $\pi^+$ propagate from the $e^+e^-$ interaction region instead of $K^+$ and $\pi^-$. Both hypotheses also contain the initial pseudo-particle that provides the $e^+e^-$ four-momentum (see section 6.1).

- At each vertex, three-dimensional momentum conservation constraints are imposed. In total, each of two hypotheses has 6 constraints on the conservation of the three-dimensional momentum components, i.e. three constraints per each vertex.

- Each hypothesis has only one energy conservation constraint. Only the final particles and the initial pseudo-particle are involved in this constraint.
- Each hypothesis requires three vertex constraints for each final particle (one constraint for each trajectory component). Three vertex constraints are also required for the intermediate particle $X$. See, sections 6.2.1, 6.3.1 and 7.3 for more details. In total, each of two hypotheses contains 15 vertex constraints.

In total, each of two discussed hypotheses contains 22 constraints. Each of the hypotheses contains 34 free parameters: 23 measurable and 11 non-measurable parameters.

The similar hypothesis with $K_S$ instead of $X$ is discussed in section 8.2.

### 8.1.2 Fitting procedure details

Further, the considered hypotheses are applied to the simulated events of the signal process $e^+e^- \rightarrow K_SK^\mp \pi^\mp$ and the background process $e^+e^- \rightarrow \pi^+\pi^-\pi^+\pi^-$. Only four-track events are used in order to demonstrate how the fitting works under the discussed hypotheses. Moreover, only such four tracks are considered, which correspond to the total electric charge equal to zero. At the same time, for demonstration purposes, the pre-separation of tracks into kaon and pion ones is not performed. The presence of four charged particles implies the presence of combinatorics associated with the fact that it is not known in advance which track from a certain event corresponds to one or another charged particle. Therefore, it is necessary to perform a separate fit for each mapping of particles to their tracks. All possible variants of such mappings are given in table 1. The table shows for each hypothesis there are four possible mappings of charged particles into tracks. Since it is not known in advance which of the signal processes in a particular event, in each event the fit is performed in both hypotheses.

In order to demonstrate how the fitting works, for each hypothesis, among the four mappings of particles to their tracks, the one in which the fit gives the smallest chi-square is selected. Fits in each event, the fit is performed eight times.

In order to demonstrate how the fitting works under the discussed hypotheses. Moreover, only such four tracks are considered, which correspond to the total electric charge equal to zero. At the same time, for demonstration purposes, the pre-separation of tracks into kaon and pion ones is not performed. The presence of four charged particles implies the presence of combinatorics associated with the fact that it is not known in advance which track from a certain event corresponds to one or another charged particle. Therefore, it is necessary to perform a separate fit for each mapping of particles to their tracks. All possible variants of such mappings are given in table 1. The table shows for each hypothesis there are four possible mappings of charged particles into tracks. Since it is not known in advance which of the signal processes $e^+e^- \rightarrow K_SK^\mp \pi^\mp$ or $e^+e^- \rightarrow K_SK^-\pi^+$ took place in a particular event, in each event the fit is performed in both hypotheses $e^+e^- \rightarrow XK^\pm \pi^\mp$. Thus, in each event, the fit is performed eight times.

In order to demonstrate how the fitting works, for each hypothesis, among the four mappings of particles to their tracks, the one in which the fit gives the smallest chi-square is selected. Fits in the charge conjugate hypotheses $e^+e^- \rightarrow XK^+\pi^-$ and $e^+e^- \rightarrow XK^-\pi^+$ give the same chi-square distribution when applied to signal events. Therefore, among these hypotheses, the hypothesis corresponding to the minimum chi-square is selected. It should be noted that only those fits that

| Hypothesis | $e^+e^- \rightarrow XK^+\pi^-$ | $e^+e^- \rightarrow XK^-\pi^+$ |
|------------|-------------------------------|-------------------------------|
| Origin vertex | $K^+$ | $\pi^-$ | $\pi^+$ | $\pi^-$ | $K^+$ | $\pi^+$ | $\pi^+$ | $\pi^-$ |
| Particle | $t_1^+$ | $t_1^-$ | $t_2^+$ | $t_2^-$ | $t_1^+$ | $t_1^-$ | $t_2^+$ | $t_2^-$ |
| Track combinations | $t_1^+$ | $t_2^+$ | $t_1^+$ | $t_2^+$ | $t_1^-$ | $t_2^-$ | $t_1^-$ | $t_2^-$ |

The signal events for the $e^+e^- \rightarrow XK^+\pi^-$ hypothesis are the events of the $e^+e^- \rightarrow K_SK^\mp \pi^\mp$ process, while the signal events for the $e^+e^- \rightarrow XK^-\pi^+$ hypothesis are the events of the $e^+e^- \rightarrow K_SK^-\pi^+$ process.

Thus, among the results of eight fits in each event, the result that corresponds to the minimum chi-square is accepted.
Figure 4. Chi-square distribution (a) and invariant mass distributions of the $X \rightarrow \pi^+\pi^-$ decay products (b). All distributions were obtained for the events of the $e^+e^- \rightarrow K_SK^+\pi^-$ simulation. The chi-square distribution is shown as black solid line and corresponds to the fitting procedure described in section 8.1.2. The red dash-dotted line shown in figure (a) is the result of fitting the chi-square distribution with the chi-squared probability density function given by equation (4.1). The fitting function depends on two parameters. One of these parameters is the normalization factor $N$, the other is the number of degrees of freedom $\nu$. Both parameters were free during fitting. The dashed line in figure (b) corresponds to the invariant mass distribution obtained using the parameters found with the fitting procedure described in section 8.1.2. The solid line corresponds to the invariant mass distribution obtained using the vertex fit only. This vertex fit is applied to the pair of pions from the $X \rightarrow \pi^+\pi^-$ decay. The tracks corresponding to these pions are considered to be known from the fitting procedure described in section 8.1.2.

have converged to a local minimum are accepted for consideration. If in the event none of the fits converged to a local minimum, then such an event is ignored.

8.1.3 Fitting the $e^+e^- \rightarrow K_SK^+\pi^-$ events

This subsection presents the results of the fitting the simulated $e^+e^- \rightarrow K_SK^+\pi^-$ events under the $e^+e^- \rightarrow XK^+\pi^-$, $X \rightarrow \pi^+\pi^-$ hypotheses. The $e^+e^- \rightarrow K_SK^+\pi^-$ events were simulated at the center-of-mass energy of 1792.9 MeV and the magnetic field of 1 T. The fitting procedure converged to a local minimum in 24770 events out of 25224. In 147 events, the fitting procedure converged to a local maximum. In the rest 307 events, this procedure did not converge.

Figure 4(a) shows the chi-square distribution (black solid line) for events where the fitting procedure described in section 8.1.2 has converged to a local minimum. If the influence of the constraint nonlinearity is small, then according to sections 4.1.2 and 8.1.1, we expect the number of degrees of freedom to be $m - l = 22 - 11 = 11$. Chi-square distribution that corresponds to the
chi-squared probability density function given by equation (4.1) must have a mean value equal to the number of degrees of freedom. However, the mean value of the chi-square distribution shown in figure 4(a) is significantly larger than the expected number of degrees of freedom. The reason is that the chi-square distribution shown in figure 4(a) is distorted due to the non-Gaussian response of the detector and the imperfect calibration of the covariance matrix $\hat{C}$. Partially, the distortion of this distribution is caused by the nonlinearity of the constraints, as well as neglecting ISR photons in the kinematic hypotheses.

Although the mean value of the chi-square distribution shown in figure 4(a) is significantly larger than expected one, the peak position of this distribution is quite close to the expected number of degrees of freedom. This is easy to see by comparing the chi-square distribution (black solid line) with the red dash-dotted line. The dash-dotted line is obtained by fitting the chi-square distribution with the chi-square probability density function given by equation (4.1). It can be seen from the figure that the fitting curve does not describe the chi-square distribution well. The fitting curve reaches its maximum at point $\chi^2 = \nu \approx 15.7$. The maximum of the chi-square distribution (black solid line) is located slightly to the left of this point.

Effects related with the non-Gaussian response of the detector, the imperfect calibration of the covariance matrix $\hat{C}$, and constraint nonlinearity are discussed in section 9. This discussion is based on the examples of the Gaussian simulation. The technique of the Gaussian simulation is introduced in section 4.4. This technique eliminates the influence of the detector response and the calibration of the covariance matrix on the shape of the chi-square distribution. The effects associated with the absence of ISR photons in kinematic hypotheses are also excluded in the Gaussian simulation. Therefore, in cases where the nonlinearity of the constraints can be neglected, the chi-square distribution obtained using the Gaussian simulation is well described by the chi-squared probability density function. In particular, section 9.2 shows that in the case of a Gaussian response, the number of degrees of freedom of the chi-square distribution corresponding to this example is indeed 11.

Figure 4(b) shows the invariant mass distributions of the two pions from the $X \rightarrow \pi^+\pi^-$ decay. The dashed line corresponds to the invariant mass distribution obtained using the particle parameters found with the kinematic and vertex fitting procedure described in section 8.1.2. The solid line corresponds to the invariant mass distribution obtained using the vertex fit only. This vertex fit is applied to the pair of the pions from the $X \rightarrow \pi^+\pi^-$ decay. The tracks corresponding to these pions are considered to be known from the kinematic and vertex fitting procedure. It can be seen from the figure that kinematic fitting significantly improves the resolution of the invariant mass. The mean values of the invariant mass for the both distributions are close to the $K_S$ meson invariant mass.

Figure 5(a) shows the distribution of the distance between the $e^+e^-$ interaction vertex and the $X \rightarrow \pi^+\pi^-$ decay vertex. The coordinates of both vertices are found using the kinematic and vertex fitting procedure. The shape of the considered distribution, as expected, is close to exponential. However, this distribution is not exactly exponential, since $K_S$-mesons have a non-trivial energy spectrum.

The kinematic fitting technique makes it possible to separate charged kaons from pions, since the masses of these particles differ significantly. However, this technique does not allow achieving the desired level of the $K-S\pi$ separation. For example, in the case of fitting the $e^+e^- \rightarrow K_SK^\pm\pi^\mp$ events under the $e^+e^- \rightarrow XKK^\pm\pi^\mp$ hypotheses, kaons and pions can be misassigned at the $e^+e^-$

\[\text{At least when using the considered hypotheses.}\]
interaction vertex. These misassignments can be seen in figures 5(b) and 5(c). The figures show the dependencies of $dE/dx$ on the particle momentum for those particles that the fitting procedure identified as kaons and pions. Figure 5(b) corresponds to particles that have been identified as kaons, while figure 5(c) corresponds to pions. In each of the figures, one can see typical dependencies corresponding to both particle types (true kaons and pions). Figure 5(d) shows the dependence of $dE/dx$ on the particle momentum for those particles that were identified as pions from the $X \rightarrow \pi^+\pi^-$ decay. In this figure, one can see only a typical dependence corresponding to true pions. Thus, we can conclude that in this case there is a significant misassignment of kaons and pions at the $e^+e^-$ interaction vertex, while pions from the $X \rightarrow \pi^+\pi^-$ decay are not misassigned as kaons.

Since it is known from the simulation which of processes $e^+e^- \rightarrow K_SK^+\pi^-$ or $e^+e^- \rightarrow K_SK^-\pi^+$ took place in each event, it is possible to indicate the proportion of events in which the fitting procedure incorrectly found the signal hypothesis. This proportion is approximately equal to 34.6%.

It should be noted that the misassignment of kaons and pions has a dramatic effect on the results of the fitting procedure described in section 8.1.2. This misassignment leads to a significant distortion of the fitting parameters. In the considered case, not only the parameters of particles flying from the $e^+e^-$ interaction vertex are distorted, but also the parameters of pions from the $X \rightarrow \pi^+\pi^-$ decay vertex. This happens because the parameters of different particles are related through the
constraints. Figures 6(a) and 6(b) show the distributions of the invariant masses of pions from the $X \rightarrow \pi^+\pi^-$ decay. Both figures were obtained using the $e^+e^- \rightarrow K_SK^\pm\pi^\mp$ simulation events. However, figure 6(a) was obtained for events in which the fitting procedure chose the wrong signal hypothesis, and figure 6(b) corresponds to those events in which the signal hypothesis was chosen correctly. It can be seen from these figures that the resolution of the invariant mass is significantly worse if the signal hypothesis was chosen incorrectly.

Thus, one can conclude that an additional $K$-$\pi$ separation procedure is required to avoid the misassignment of kaons and pions. This procedure may be based on the use of $dE/dx$, for example. The $K$-$\pi$ separation procedure must be used before the kinematic and vertex fitting. After performing such a procedure, it will be known which tracks correspond to kaons with the highest probability. Thus, in each event, one know the mapping of the kaon to its track. The sign of the kaon charge determines the signal hypothesis that will be used for the fitting in a certain event. As a result, the number of fits in a single event can be significantly reduced. Instead of mappings from table 1, it is enough to consider only permutations of tracks for the pions of the same charge, i.e. there will be only two fits per event. Among these two fits, as in the case of fitting procedure described in section 8.1.2, one can choose the fit with the smallest chi-square.

In this work, we do not use any additional $K$-$\pi$ separation procedure, but for demonstration purposes we can use the fact that in each event of the simulation we know which of the final states $K_SK^\pm\pi^\mp$ or $K_SK^-\pi^+$ occurred. Thus, it is possible to run fitting procedure for each event of the $e^+e^- \rightarrow K_SK^\pm\pi^\mp$ simulation under the right signal hypothesis. For a given signal hypothesis, it

---

**Figure 6.** Invariant mass distributions of the $X \rightarrow \pi^+\pi^-$ decay products. The distributions shown in the figure were obtained for the simulated events of the $e^+e^- \rightarrow K_SK^\pm\pi^\mp$ process. The meaning of the solid and dashed lines is the same as in figure 4(b). Figure (a) corresponds to events in which the fitting procedure converged to an incorrect signal hypothesis, while figure (b) corresponds to events in which the procedure converged to the correct signal hypothesis.
is necessary to consider only four mappings from table 1, i.e. do four fits in each event. Using the correct signal hypothesis in each event eliminates the misassignment of kaons and pions at the $e^+e^-$ interaction vertex. Taking into account that there is also no misassignment of kaons and pions from the $X \rightarrow \pi^+\pi^-$ decay (see figure 5(d)), it can be obtained that with this approach there is no misassignment of kaons and pions. The fitting procedure in this case converges to a local minimum in 24525 events out of 25224, i.e. in most events where the wrong hypothesis was previously chosen, there was a fit under the correct hypothesis with a slightly larger chi-square.

Some of the results of using the fitting procedure described in the previous paragraph are shown in figure 7. Figure 7(a) shows the chi-square distribution. Figure 7(b) shows the distributions of the invariant mass of pions from the $X \rightarrow \pi^+\pi^-$ decay. Unlike figure 4(b), figure 7(b) was obtained by fitting each simulation event using the correct signal hypothesis only. It can be seen from figure 7(b) that the distribution of the invariant mass obtained using the pion parameters after fitting has somewhat changed compared to the similar distribution shown in figure 4(b) (see dashed line). The left slope of the peak became steeper, and the peak itself is slightly higher, which indicates a slight improvement in mass resolution. However, the distribution in figure 7(b) has non-Gaussian tails, as in the case of a similar distribution in figure 4(b), which leads to the fact that the standard deviations for the histograms in both figures are approximately the same.

It can be concluded what is the accuracy of vertex reconstruction using the kinematic and vertex fitting procedure by comparing the vertex coordinates known from the simulation with similar coordinates found using the fitting procedure. Figure 8(a) shows the distribution of the difference between

---

**Figure 7.** Chi-square distribution (a) and invariant mass distributions of the $X \rightarrow \pi^+\pi^-$ decay products (b). This figure is similar to figure 4, but the fitting procedure is performed under the correct signal hypothesis only. In each event of the $e^+e^- \rightarrow K_S K^+\pi^-$ simulation, it is known which of the two signal hypotheses ($e^+e^- \rightarrow X K^+\pi^-$ or $e^+e^- \rightarrow X K^-\pi^+$) is the correct one.
Figure 8. Comparison of the $x$-coordinate of the $e^+e^-$ interaction vertex found using the fitting procedure with the same coordinate known from the simulation. This figure corresponds to the events of the $e^+e^- \rightarrow K_SK^-\pi^+$, $K_S \rightarrow \pi^+\pi^-$ simulation. The kinematic and vertex fitting procedure is applied to these events under the correct signal hypothesis ($e^+e^- \rightarrow XK^+\pi^-$ or $e^+e^- \rightarrow XK^-\pi^+$), known from the simulation. The difference between the $x$-coordinate of the $e^+e^-$ interaction vertex found using the fitting procedure and the same coordinate known from the simulation is shown in figure (a) with a black solid line. The pull distribution for this coordinate is shown in figure (b) with a black solid line. The dash-dotted lines in both figures (a) and (b) indicate the results of fitting the listed distributions using the probability density function of the normal distribution multiplied by the normalization factor $N$. This factor, mean value $\mu$ and standard deviation $\sigma$ are free fitting parameters.

the $x$-coordinate of the $e^+e^-$ interaction vertex found using the kinematic and vertex fitting procedure with the same coordinate known from the simulation. To rule out misassignment of kaons and pions in the $X \rightarrow \pi^+\pi^-$ decay vertex, the fitting procedure was carried out under the correct signal hypothesis.

Figure 8(b) shows the pull distribution for the $x$-coordinate of the $e^+e^-$ interaction vertex. Further, by pull distribution we mean the distribution of the quantity $\Delta \zeta / \sqrt{\text{Var}(\zeta^{\text{fit}})}$, where $\zeta = x, y$ or $z$ is the vertex coordinate, $\Delta \zeta$ is the difference between the $\zeta$-coordinate of the vertex found using the kinematic and vertex fitting procedure with the same coordinate known from the simulation, $\text{Var}(\zeta^{\text{fit}})$ is the variance of the $\zeta$-coordinate found using the fitting package. This variance is the corresponding diagonal element of the covariance matrix, defined as $2\hat{\mathcal{H}}^{-1}$, where $\hat{\mathcal{H}}$ is the Hessian of the Lagrange function calculated at the last iteration of Newton’s method (see equation (2.7)).

In the case of the pull distribution shown in figure 8(b), it can be seen that the corresponding standard deviation is about twice greater than one. As in the case of the chi-square distribution, this fact is mainly explained by the non-Gaussian response of the detector and the imperfect calibration.
Figure 9. Comparison of the $x$-coordinate of the $X \rightarrow \pi^+\pi^-$ decay vertex found using the fitting procedure with the same coordinate known from the simulation. This figure corresponds to the events of the $e^+e^- \rightarrow K_S K^+\pi^-$, $K_S \rightarrow \pi^+\pi^-$ simulation. The kinematic and vertex fitting procedure is applied to these events under the correct signal hypothesis ($e^+e^- \rightarrow KK^+\pi^-$ or $e^+e^- \rightarrow KK^-\pi^+$), known from the simulation. The difference between the $x$-coordinate of the $X \rightarrow \pi^+\pi^-$ decay vertex found using the fitting procedure and the same coordinate known from the simulation is shown in figure (a) with a black solid line. The pull distribution for this coordinate is shown in figure (b) with a black solid line. The dash-dotted lines in both figures (a) and (b) indicate the results of fitting the listed distributions using the probability density function of the normal distribution multiplied by the normalization factor $N$. This factor, mean value $\mu$ and standard deviation $\sigma$ are free fitting parameters.

of the matrix $\hat{C}$. The discussed standard deviation should be close to one if the detector response is Gaussian and the covariance matrix is well calibrated. The verification of the latter statement is given in section 9.2 using Gaussian simulation of the $e^+e^- \rightarrow K_S K^+\pi^-$ process. As a result of this verification, the standard deviation of the quantity $\Delta \zeta / \sqrt{\text{Var}(\zeta)}$ is close to one, which indicates that the fitting procedure correctly reconstructs the vertices.

Figure 9(a) shows the distribution of the difference between the $x$-coordinate of the $X \rightarrow \pi^+\pi^-$ decay vertex found using the kinematic and vertex fitting procedure with the same coordinate known from the simulation. Figure 9(b) shows the pull distribution corresponding to this coordinate. Figure 10 is similar to figure 8, but corresponds to the $z$-coordinate of the $e^+e^-$ interaction. Figure 11 corresponds to the $z$-coordinate of the $X \rightarrow \pi^+\pi^-$ decay vertex. For the same reason as in the case of figure 8(b), the standard deviations corresponding to the pull distributions shown in figures 9(b), 10(b) and 11(b) are greater than one.
8.1.4 Fitting the events of the $e^+e^- \rightarrow K_S K^\pm \pi^\mp$ simulation with ISR disabled

In the previous subsection, it is shown that the chi-square distribution obtained as a result of fitting the events of the $e^+e^- \rightarrow K_S K^\pm \pi^\mp$, $K_S \rightarrow \pi^+\pi^-$ simulation under the correct signal hypothesis ($e^+e^- \rightarrow X K^\pm \pi^\mp$ or $e^+e^- \rightarrow X K^- \pi^+$), which is known from the simulation. The difference between the $z$-coordinate of the $e^+e^-$ interaction vertex found using the fitting procedure and the same coordinate known from the simulation is shown in figure (a) with a solid line. The pull distribution for this coordinate is shown in figure (b) with a solid line. The dash-dotted lines in both figures (a) and (b) indicate the results of fitting the listed distributions using the probability density function of the normal distribution multiplied by the normalization factor $N$. This factor, mean value $\mu$ and standard deviation $\sigma$ are free fitting parameters.

In this regard, it is interesting to know what contribution to this distortion is caused by the absence of ISR photons in the kinematic hypotheses. In order to answer this question, we consider the events of the $e^+e^- \rightarrow K_S K^\pm \pi^\mp$, $K_S \rightarrow \pi^+\pi^-$ simulation, in which initial state radiation is turned off. The center-of-mass energy and the magnetic field in this simulation are the same as in the previous subsection. The events of this simulation are fitted under the correct signal hypothesis ($e^+e^- \rightarrow X K^\pm \pi^\mp$, $X \rightarrow \pi^+\pi^-$) known from the simulation. In the case of the considered example, the fitting procedure converged to a local minimum in 24091 events out of 24839. This procedure converged to a local maximum in 279 events. In the remaining 469 events, the procedure did not converge.
The results of the fitting are shown in figure 12. The distributions shown in this figure should be compared with similar distributions shown in figure 7 (in the case of figure 7 ISR is enabled in the events of the $e^+e^- \rightarrow KS K^+\pi^-$ simulation).

Figure 12(a) shows the chi-square distribution corresponding to the events of the $e^+e^- \rightarrow KS K^+\pi^-$ simulation with ISR disabled. Although neither the simulated events nor the kinematic hypotheses contain ISR photons, this distribution is not consistent with the chi-squared probability density function. As a result, it can be concluded that the distortion of the chi-square distribution is not fully explained by the effects associated with initial state radiation. The same conclusion can be drawn for other kinematic hypotheses discussed below.

Figure 12(b) shows the distributions of the invariant mass of the $X \rightarrow \pi^+\pi^-$ decay products. The solid line corresponds to the distribution obtained as a result of the vertex fitting, while the dashed line corresponds to the distribution obtained as a result of kinematic and vertex fitting under the $e^+e^- \rightarrow X K^+\pi^+$, $X \rightarrow \pi^+\pi^-$ hypotheses. Both distributions were obtained for the events of the

---

**Figure 11.** Comparison of the $z$-coordinate of the $X \rightarrow \pi^+\pi^-$ decay vertex found using the fitting procedure with the same coordinate known from the simulation. This figure corresponds to the events of the $e^+e^- \rightarrow KS K^+\pi^-$, $KS \rightarrow \pi^+\pi^-$ simulation. The kinematic and vertex fitting procedure is applied to these events under the correct signal hypothesis ($e^+e^- \rightarrow XK^+\pi^-$ or $e^+e^- \rightarrow XK^-\pi^+$), known from the simulation. The difference between the $z$-coordinate of the $X \rightarrow \pi^+\pi^-$ decay vertex found using the fitting procedure and the same coordinate known from the simulation is shown in figure (a) with a black solid line. The pull distribution for this coordinate is shown in figure (b) with a black solid line. The dash-dotted lines in both figures (a) and (b) indicate the results of fitting the listed distributions using the probability density function of the normal distribution multiplied by the normalization factor $N$. This factor, mean value $\mu$ and standard deviation $\sigma$ are free fitting parameters.
Figure 12. Chi-square distribution (a) and invariant mass distributions of the $X \rightarrow \pi^+\pi^-$ decay products (b). The distribution is obtained for the events of the $e^+e^- \rightarrow K_S K^+\pi^-$, $K_S \rightarrow \pi^+\pi^-$ simulation fitted under the $e^+e^- \rightarrow X K^+\pi^-$, $X \rightarrow \pi^+\pi^-$ hypotheses. This figure is similar to figure 7, but ISR is disabled in the events of the $e^+e^- \rightarrow K_S K^+\pi^-$ simulation.

$e^+e^- \rightarrow K_S K^+\pi^-$ simulation with ISR disabled. The distribution indicated by the dashed line in figure 12(b) can be compared with the similar distribution shown in figure 7(b). The distribution shown in figure 7(b) is distorted with respect to the distribution shown in figure 12(b) because the simulated events in the case of the distribution shown in figure 7(b) contain ISR photons, which are not taken into account in the $e^+e^- \rightarrow X K^+\pi^-$, $X \rightarrow \pi^+\pi^-$ hypotheses. To avoid such distortion of invariant mass distribution, one can use kinematic hypotheses containing ISR photons. However the use of such hypotheses usually worsens the resolution of invariant masses, since ISR photons in the case of CMD-3 are considered as lost particles.

8.1.5 Fitting the $e^+e^- \rightarrow \pi^+\pi^-\pi^+\pi^-$ events

This subsection provides an example of applying the fitting procedure described in section 8.1.2 under the $e^+e^- \rightarrow X K^+\pi^-$, $X \rightarrow \pi^+\pi^-$ hypotheses to the simulated $e^+e^- \rightarrow \pi^+\pi^-\pi^+\pi^-$ events. The center-of-mass energy and the magnetic field for these events are the same as in the previous section. The process $e^+e^- \rightarrow \pi^+\pi^-\pi^+\pi^-$ is one of the main background processes in the study of the $e^+e^- \rightarrow K_S K^+\pi^-$ process with the CMD-3 detector. It should be noted that the $e^+e^- \rightarrow \pi^+\pi^-\pi^+\pi^-$ process does not contain charged kaons in the final state and, therefore, is also a background process with respect to the $e^+e^- \rightarrow X K^+\pi^-$ hypotheses. The fitting procedure converged to a local minimum in 44739 events out of 45199. Some results of the fitting are shown in figure 13.

Figure 13(a) shows the chi-square distribution obtained by applying the fitting procedure to the events of the $e^+e^- \rightarrow \pi^+\pi^-\pi^+\pi^-$ simulation. Since the $e^+e^- \rightarrow \pi^+\pi^-\pi^+\pi^-$ is the background
Figure 13. Chi-square distribution (a) and distribution of the distance between the $e^+e^-$ interaction vertex and the $X \rightarrow \pi^+\pi^-$ decay vertex (b). Both distributions are obtained by applying the kinematic and vertex fitting procedure to the events of the $e^+e^- \rightarrow \pi^+\pi^-\pi^+\pi^-$. The corresponding fitting procedure is described in section 8.1.2.

process with respect to the considered hypotheses, the chi-square distribution turned out to be wide (compare with figure 4(a)).

Figure 13(b) shows the distribution of the distance $\Delta r_{\text{fit}}$ between the $e^+e^-$ interaction vertex and the $X \rightarrow \pi^+\pi^-$ decay vertex. The distribution was obtained for the $e^+e^- \rightarrow \pi^+\pi^-\pi^+\pi^-$ events using the fitting procedure described in section 8.1.2. The distance is calculated using the vertex coordinates found as a result of the fitting. Since in the $e^+e^- \rightarrow \pi^+\pi^-\pi^+\pi^-$ process all pions originate from the $e^+e^-$ interaction vertex, the large $\Delta r_{\text{fit}}$ distances are not expected. Indeed, it can be seen from the figure that most of the events belong to region $\Delta r_{\text{fit}} \lesssim 5$ mm.

Finally, we can conclude that the chi-square and $\Delta r_{\text{fit}}$ distributions can be successfully used to select the events of the $e^+e^- \rightarrow K_S K^\mp \pi^\mp$ process and suppress the corresponding background processes, such as the $e^+e^- \rightarrow \pi^+\pi^-\pi^+\pi^-$ process.

8.2 Hypotheses $e^+e^- \rightarrow K_S K^\mp \pi^\mp$, $K_S \rightarrow \pi^+\pi^-$

8.2.1 Description of the $e^+e^- \rightarrow K_S K^\mp \pi^\mp$ hypotheses

In this section we present the results of fitting under $e^+e^- \rightarrow K_S K^\mp \pi^\mp$, $K_S \rightarrow \pi^+\pi^-$ hypotheses. The difference between section 8.1 and this section is that in this section the intermediate neutral particle ($K_S$) is known. By the fact that the intermediate particle is known, we mean that this particle directly participates in the energy-momentum conservation constraints, i.e. its mass influences the fitting results. The energy-momentum conservation constraint tree corresponding to the $e^+e^- \rightarrow K_S K^+\pi^-$, $K_S \rightarrow \pi^+\pi^-$ hypothesis has the form shown in figure 2. For the charge-conjugate hypothesis, there is a similar tree, but with oppositely charged particles. Thus, energy conservation constraints are imposed in each of the two vertices and include all particles related with one or another vertex. This is the only difference between the hypotheses $e^+e^- \rightarrow K_S K^\mp \pi^\mp$ and $e^+e^- \rightarrow XK^\pm \pi^\mp$ (see subsection 8.1.1).
Figure 14. Chi-square distribution (a) and invariant mass distributions of the $X \to \pi^+\pi^-$ decay products (b). This figure is similar to figure 4, but the kinematic and vertex fitting is done under the $e^+e^- \to K_S K^\pm \pi^\mp$, $K_S \to \pi^+\pi^-$ hypotheses.

The fitting procedure with the $e^+e^- \to K_S K^\pm \pi^\mp$ hypotheses is completely similar to the fitting procedure with the $e^+e^- \to X K^\pm \pi^\mp$ hypotheses described in section 8.1.2. This procedure does not contain any additional $K$-$\pi$ pre-separation algorithm and is used for demonstration purposes only.

8.2.2 Fitting the $e^+e^- \to K_S K^\pm \pi^\mp$ events

This subsection discusses the results of applying the fitting procedure under the $e^+e^- \to K_S K^\pm \pi^\mp$ hypotheses to the events of the $e^+e^- \to K_S K^\pm \pi^\mp$ process. The fitting procedure converged to a local minimum in 24645 events out of 25224. In 217 events, this procedure converged to a local maximum. In the rest 362 events, the fitting procedure did not converge. Some of the fitting results are shown in figure 14.

Figure 14(a) shows the chi-square distribution for the events of the $e^+e^- \to K_S K^\pm \pi^\mp$ simulation. This distribution corresponds to the $e^+e^- \to K_S K^\pm \pi^\mp$ hypotheses. The expected number of degrees of freedom is $l - m = 23 - 11 = 12$. However, as in the case of figure 4(a), the distribution in figure 14(a) does not correspond to the expected number of degrees of freedom due to the reasons listed in section 8.1.3.

Figure 14(b) shows the two-pion invariant mass distribution for pions from the $K_S \to \pi^+\pi^-$ decay. As in the case of figure 4(b), the dashed line indicates the invariant mass distribution obtained using the pion parameters found with the kinematic and vertex fitting procedure. Since the intermediate particle ($K_S$) in the considered hypotheses has a certain mass and participates in the energy-momentum conservation constraints, this distribution turned out to be fixed on the value of the $K_S$-meson mass. The solid line denotes the invariant mass distribution obtained using the
\( K_S \rightarrow \pi^+\pi^- \) decay vertex fit. The pion tracks from the \( K_S \)-meson decay are assumed to be known from the kinematic and vertex fitting procedure.

8.3 Hypothesis \( e^+e^- \rightarrow \gamma\gamma\gamma \)

8.3.1 Description of the \( e^+e^- \rightarrow \gamma\gamma\gamma \) hypothesis

In section 8.3, we present the results of applying the kinematic and vertex fitting package under the \( e^+e^- \rightarrow \gamma\gamma\gamma \) hypothesis to the events of the \( e^+e^- \rightarrow \pi^0\gamma, \pi^0 \rightarrow \gamma\gamma \) simulation. The \( e^+e^- \rightarrow \gamma\gamma\gamma \) hypothesis has the following form.

- The hypothesis contains a single vertex. This vertex is the \( e^+e^- \) interaction vertex. All coordinates of this vertex are considered as free measurable parameters and contribute to the chi-square.
- The hypothesis contains three final particles. These particles are photons described in section 6.2.2. The hypothesis contains also the initial pseudo-particle (see section 6.1), which is used in order to provide the total four-momentum of the initial particles (\( e^+e^- \)).
- All of the above particles are involved in four energy-momentum conservation constraints.
- The vertex constraints described in section 7.3 are not imposed on the photons, because the photons are already emitted from the vertex of their origin due to the parametrization of their momentum (6.4).

In total, the \( e^+e^- \rightarrow \gamma\gamma\gamma \) hypothesis contains 4 constraints and 15 free parameters. All of these parameters are measurable.

8.3.2 Fitting the \( e^+e^- \rightarrow \pi^0\gamma \) events

In this subsection, we discuss the results of fitting the simulated \( e^+e^- \rightarrow \pi^0\gamma, \pi^0 \rightarrow \gamma\gamma \) events under the \( e^+e^- \rightarrow \gamma\gamma\gamma \) hypothesis. The events of this simulation correspond to the center-of-mass energy equal to 782.7 MeV. For the demonstration purposes, only events with three photon clusters in the calorimeter are considered. Thus, since all three photons are equivalent, there is a single mapping of photons to their clusters, i.e. only one fit is performed in each event. The fitting procedure converged to a local minimum in all events.

Figure 15(a) shows a chi-square distribution corresponding to the example discussed in this section. The mean value of the distribution shown in the figure is approximately 5.9. This value is greater than the expected number of degrees of freedom \((m - l = 4 - 0 = 4)\). The fact that the chi-square distribution is distorted is due to the reasons listed in section 8.1.3.

Figure 15(b) shows histograms with the two-photon invariant mass distributions. The histogram indicated by the solid line is obtained using the measured photon parameters, and the histogram indicated by the dashed line is obtained using the photon parameters found using the fitting procedure. In the case of the histogram marked with the solid line, it is assumed that photons propagate from the center of the detector. Two-photon invariant mass distributions are shown for the reason that we would like to see a pion peak on them. Since in the \( e^+e^- \rightarrow \gamma\gamma\gamma \) hypothesis, it is unknown which photons are the pion decay products, these histograms contain the invariant masses of all different
Figure 15. Chi-square distribution (a) and two-photon invariant mass distributions (b). All distributions are obtained as a result of fitting the events of the $e^+e^- \rightarrow \pi^0 \gamma, \pi^0 \rightarrow \gamma\gamma$ simulation under the $e^+e^- \rightarrow \gamma\gamma\gamma$ hypothesis. In figure (b), the dashed line corresponds to the invariant mass distribution obtained with kinematic and vertex fitting procedure. The distribution indicated by the solid curve in the same figure is obtained using initial photon parameters and assuming that photons propagate from the center of the detector.

photon pairs, i.e. each histogram has three entries per event. Thus, in these histograms, in addition to the pion peak, there is a contribution from the combinatorial background. It is seen from the figure that the fitting procedure significantly improves the two-photon invariant mass resolution.

The hypothesis $e^+e^- \rightarrow \gamma\gamma\gamma$ is quite simple. In this hypothesis, the energy-momentum conservation constraints are only imposed, and there are no explicit vertex constraints. Let us show how the energy-momentum conservation constraints hold. Figure 16 shows the distributions of the left-hand sides of the energy-momentum conservation constraints (7.1). The solid lines indicate the distributions obtained using the measured particle parameters. In the case of distributions indicated by solid lines, it is assumed that photons propagate from the center of the detector. The dashed lines denote the distributions obtained using the particle parameters found with the fitting procedure. The dashed distributions are fixed at zero, i.e. energy-momentum conservation constraints are satisfied with high accuracy. Similar distributions for the energy-momentum conservation constraints, as well as for vertex constraints, can also be shown in the case of more complex hypotheses. However, in this paper, we restrict ourselves to the distributions shown in figure 16.

---

There are no vertex constraints described in section 7.3 and corresponding to their own Lagrange multipliers. The vertex constraints hold implicitly due to the parametrization of the photon momentum (see equation (6.4)).
Figure 16. Distributions demonstrating the fulfillment of the energy-momentum conservation constraints in the case of the $e^+e^- \rightarrow \gamma\gamma\gamma$ hypothesis. The distributions are obtained for the events of the $e^+e^- \rightarrow \pi^0\gamma$, $\pi^0 \rightarrow \gamma\gamma$ simulation. Figures (a), (b), and (c) show the $x$-, $y$-, and $z$- components of the total momentum, respectively. Figure (d) shows the difference between total energy and center-of-mass energy. The solid lines show the distributions obtained using the initial parameters of the photons under the assumption that the photons propagate from the center of the detector. The dashed lines indicate the distributions obtained using the parameters found by fitting.
8.4 Hypothesis $e^+e^- \rightarrow \pi^+\pi^-\gamma\gamma$

8.4.1 Description of the $e^+e^- \rightarrow \pi^+\pi^-\gamma\gamma$ hypothesis

Section 8.4.1 discusses the results of applying kinematic and vertex fitting under the $e^+e^- \rightarrow \pi^+\pi^-\gamma\gamma$ hypothesis. This hypothesis has the following form.

- The hypothesis has a single vertex, which is the $e^+e^-$ interaction vertex. All coordinates of this vertex are considered as free measurable parameters and contribute to the chi-square.

- The hypothesis requires the presence of four final particles and one initial pseudo-particle 6.1. The list of final particles consists of two photons (see section 6.2.2) and two oppositely charged pions (see section 6.2.1).

- There are four constraints on the energy-momentum conservation in this hypothesis. All particles are involved into these constraints.

- Three vertex constraints are imposed on each charged particle. No vertex constraints are imposed on the photons, because the photons already fly out of the vertex due to the parametrization of their momenta.

In total, the $e^+e^- \rightarrow \pi^+\pi^-\gamma\gamma$ hypothesis contains 10 constraints. This hypothesis contains 23 free parameters. Among these parameters, 21 parameters are measurable and 2 parameters are non-measurable.

8.4.2 Fitting the $e^+e^- \rightarrow \eta\pi^+\pi^-$, $\eta \rightarrow \gamma\gamma$ events

In this subsection we present the results of applying the fitting package to the events of the $e^+e^- \rightarrow \eta\pi^+\pi^-$, $\eta \rightarrow \gamma\gamma$ simulation. These events correspond to the center-of-mass energy of 1.84 GeV. For demonstration purposes, only those events are passed to the input of the fitting procedure, in which there are at least two photon clusters in the calorimeter. If there are only two photon clusters in an event, then a single fit is performed in this event. If the number of photon clusters is more than two, then separate fits are performed for each different mapping of photons to their clusters. Permutations of two photon clusters in each pair are not considered, because they do not affect the fitting result. In the case of the considered example, the fitting procedure converged in 26918 events out of 29199. In 2199 events the fitting procedure did not converge, while in 82 events this procedure converged to a local maximum. Some fitting results are shown in figure 17.

Figure 17(a) shows the chi-square distribution corresponding to the example discussed in this section. The mean value of the distribution is approximately equal to 23, while the expected number of degrees of freedom is $m - l = 10 - 2 = 8$. As in the previous examples, the mean value of this distribution does not match the expected number of degrees of freedom. The reasons for the distribution being distorted are the same as those listed in section 8.1.3.

The fitting procedure significantly improves the resolution of the two-photon invariant mass. The distributions of the two-photon invariant mass are shown in figure 17(b). The distribution indicated by the dashed line was obtained using the particle parameters found with the kinematic and vertex fitting procedure. The distribution shown by the solid line was obtained using the measured photon parameters. In the case of the latter distribution, the vertex is found from the vertex fit of the pair of tracks in the drift chamber.
Figure 17. Chi-square distribution (a) and two-photon invariant mass distributions (b). The distributions are obtained by fitting the events of the $e^+e^- \rightarrow \eta\pi^+\pi^-$, $\eta \rightarrow \gamma\gamma$ simulation under the $e^+e^- \rightarrow \pi^+\pi^-\gamma\gamma$ hypothesis. The dashed line in figure (b) corresponds to the invariant mass distribution obtained using the parameters found with the kinematic and vertex fitting procedure. The solid line corresponds to the invariant mass distribution obtained using the vertex fit only. This fit is applied to the two tracks of charged pions in the drift chamber. In the case of the latter distribution, the photons are assumed to propagate from the vertex found by vertex fitting.

8.5 Hypothesis $e^+e^- \rightarrow \eta\pi^+\pi^-$, $\eta \rightarrow \pi^+\pi^-\pi^0_{\text{lost}}$

8.5.1 Description of the $e^+e^- \rightarrow \eta\pi^+\pi^-$, $\eta \rightarrow \pi^+\pi^-\pi^0_{\text{lost}}$ hypothesis

In section 8.5, we discuss the $e^+e^- \rightarrow \eta\pi^+\pi^-$, $\eta \rightarrow \pi^+\pi^-\pi^0_{\text{lost}}$ hypothesis. In this hypothesis the $\eta$-meson is not explicitly contained in the form of an intermediate particle. This is due to its short life time. The presence of a short-lived intermediate particle in this case is emulated by the mass constraint (7.2). The considered hypothesis has the following structure.

- The hypothesis has a single vertex. This vertex is the $e^+e^-$ interaction vertex. All coordinates of this vertex are considered as free measurable parameters and contribute to the chi-square.

- The hypothesis contain five final particles. The four final particles are charged pions (see section 6.2.1), and the fifth particle is a lost $\pi^0$. This $\pi^0$ is represented as the lost massive particle described in section 6.4.1. In order to provide the four-momentum of initial particles, the hypothesis also contains the initial pseudo-particle (see section 6.1).

- All of the particles listed above are involved in four energy-momentum conservation constraints.

- Each charged pion has three vertex constraints. In total, the hypothesis contains 12 vertex constraints.
Table 2. Mappings between charged particles and their tracks in the case of the \( e^+e^- \rightarrow \eta\pi^+\pi^- \), \( \eta \rightarrow \pi^+\pi^-\pi^0_{\text{lost}} \) hypothesis. \( t_1^+ \) and \( t_1^- \) are tracks of positively charged particles, \( t_2^+ \) and \( t_2^- \) are tracks of negatively charged particles.

| Intermediate particle | \( \eta \) |
|-----------------------|------------|
| Final particle        | \( \pi^+ \)| \( \pi^- \)| \( \pi^+ \)| \( \pi^- \)| \( \pi^0_{\text{lost}} \)|
| Track combinations    | \( t_1^+ \)| \( t_1^- \)| \( t_2^+ \)| \( t_2^- \)| \( t_1^+ \)| \( t_2^- \)| \( t_2^+ \)| \( t_1^- \)|

- Three pions (\( \pi^+\pi^-\pi^0 \)) are involved in the mass constraint.

In total, the \( e^+e^- \rightarrow \eta\pi^+\pi^- \), \( \eta \rightarrow \pi^+\pi^-\pi^0_{\text{lost}} \) hypothesis contains 17 constraints. This hypothesis contains 30 free parameters. Among these parameters, 23 parameters are measurable and 7 parameters are non-measurable.

8.5.2 Fitting procedure details

For demonstration purposes, only events with four tracks are passed to the input of the fitting procedure. Moreover, these tracks must correspond to a neutral combination of charged particles. Due to the presence of the mass constraint, charged pions are not equivalent in the \( e^+e^- \rightarrow \eta\pi^+\pi^- \), \( \eta \rightarrow \pi^+\pi^-\pi^0_{\text{lost}} \) hypothesis. This is due to the fact that only two charged pions are involved in this constraint. Therefore, the fitting procedure under the considered hypothesis must take into account different mappings of charged pions into their tracks. These mappings are listed in table 2. Among all possible combinations from this table, the combination corresponding to the minimum chi-square is selected. Note that in order for the point of initial optimization parameters to lie closer to the conditional minimum point, the initial momentum of the lost \( \pi^0 \) is set equal to the initial missing momentum of four charged pions.

8.5.3 Fitting the \( e^+e^- \rightarrow \eta\pi^+\pi^- \), \( \eta \rightarrow \pi^+\pi^-\pi^0 \) events

The fitting procedure described above is applied to the events of the \( e^+e^- \rightarrow \eta\pi^+\pi^- \), \( \eta \rightarrow \pi^+\pi^-\pi^0 \) simulation. These events correspond to the center-of-mass energy of 1.88 GeV. For demonstration purposes, we completely ignore the information about the presence of photons, i.e. even if photons from the \( \pi^0 \rightarrow \gamma\gamma \) decay were detected, we still consider the neutral pion to be lost. The fitting procedure converged to a local minimum in 12751 events out of 12796. This procedure converged to a local maximum in 8 events and did not converge in 37 events. The corresponding chi-square distribution is shown in figure 18(a). The mean value of this distribution is approximately equal to 37, while the expected number of degrees of freedom is \( m - l = 17 - 7 = 10 \). That is, the distribution shown in the figure does not fit the chi-squared probability density function with 10 degrees of freedom. The fact that the chi-square distribution is distorted is due to the reasons listed in section 8.1.3.

Figure 18(b) shows the distributions of the three-pion invariant mass. These three pions are the \( \eta \rightarrow \pi^+\pi^-\pi^0 \) decay products. The distribution indicated by the dashed line was obtained using the
Figure 18. Chi-square distribution (a) and three-pion invariant mass distributions (b). The distributions are obtained by fitting the events of the $e^+e^- \rightarrow \eta \pi^+\pi^- \eta \rightarrow \pi^+\pi^-\pi_0$ simulation under the $e^+e^- \rightarrow \eta \pi^+\pi^-$, $\eta \rightarrow \pi^+\pi^-\pi_0^{\text{lost}}$ hypothesis. The dashed line in figure (b) corresponds to the invariant mass distribution obtained using the parameters found with the kinematic and vertex fitting procedure. The solid line corresponds to the invariant mass distribution obtained using the vertex fit only. This fit is applied to the four tracks of charged pions in the drift chamber. In the case of the latter distribution, the missing momentum of the four charged pions is chosen as the momentum of the neutral pion.

particle parameters found with the kinematic and vertex fitting procedure. This distribution turned out to be fixed on the $\eta$-meson mass, since the corresponding mass constraint is used in the hypothesis. The distribution indicated by the solid line was obtained using the parameters of the charged pions found from the corresponding vertex fit. In the case of the latter distribution, it is assumed that the momentum of the lost neutral pion was set equal to the missing momentum of the four charged pions.

9 Examples of Gaussian simulation

In this section, we discuss kinematic and vertex fitting of events obtained using Gaussian simulation described in section 4.4. Subsection 9.1 provides examples of Gaussian simulations of the $e^+e^- \rightarrow \eta \pi^+\pi^-$, $\eta \rightarrow \gamma\gamma$ process. The events of these simulations are fitted under the $e^+e^- \rightarrow \pi^+\pi^-\gamma\gamma$ hypothesis. Subsection 9.2 provides examples of Gaussian simulations of the $e^+e^- \rightarrow K_SK^+\pi^-$ process. The events of these simulations are fitted under the $e^+e^- \rightarrow KK^+\pi^-$, $K_S \rightarrow \pi^+\pi^-$ hypotheses. The examples discussed in this section are of interest for testing the fitting package, as well as for observing how constraint nonlinearity can distort a chi-square distribution.
Figure 19. Chi-square distribution in the case of the $e^+e^- \rightarrow \eta\pi^+\pi^-$, $\eta \rightarrow \gamma\gamma$ Gaussian simulation. The events of this simulation were fitted under the $e^+e^- \rightarrow \pi^+\pi^-\gamma\gamma$ hypothesis (see section 8.4.1). All the constraints corresponding to this hypothesis were enabled. The solid line is the $\chi^2$ histogram, and the dash-dotted line is the fitting function $bNf_{\chi^2}(t; \nu)$ proportional to the chi-square PDF with $\nu$ degrees of freedom. Constant $b$ is the bin width. Number of events $N$ and the number degrees of freedom $\nu$ are the free fitting parameters.

9.1 Hypothesis $e^+e^- \rightarrow \pi^+\pi^-\gamma\gamma$

This subsection presents the results of fitting the events of the Gaussian simulation of the $e^+e^- \rightarrow \eta\pi^+\pi^-$, $\eta \rightarrow \gamma\gamma$ process. As the initial event, a certain event of the $e^+e^- \rightarrow \eta\pi^+\pi^-$, $\eta \rightarrow \gamma\gamma$ Monte Carlo simulation is chosen. It should be noted that the same Monte Carlo simulation is used as in section 8.4.2. The kinematic and vertex fitting procedure under the $e^+e^- \rightarrow \pi^+\pi^-\gamma\gamma$ hypothesis (see section 8.4.1) was applied once to the initial event. This procedure made it possible to obtain measurable particle parameters satisfying the constraints. Then the initial event was redefined: the measured parameters of the particles were replaced by the corresponding parameters obtained with the fitting procedure (see section 4.4). Then these parameters were drawn multiple times according to the corresponding multivariate normal distribution. As a result of such draws, the events of the Gaussian simulation were obtained. These events were fitted under the $e^+e^- \rightarrow \pi^+\pi^-\gamma\gamma$ hypothesis. Figure 19 shows the chi-square distribution corresponding to the events of this Gaussian simulation. This distribution corresponds to the case when all constraints are enabled in the $e^+e^- \rightarrow \pi^+\pi^-\gamma\gamma$ hypothesis. A list of these constraints is given in section 8.4.1. The hypothesis contains $m = 10$ constraints and $l = 2$ free non-measurable parameters. If all constraints were linear, one would expect the chi-square distribution to be described by the chi-squared probability density function (4.1) with $m - l = 8$ degrees of freedom (see section 4.1.2). However, all the constraints of the considered hypothesis are nonlinear. Despite this, the chi-square distribution is well described by the chi-squared probability density function. This statement was verified by fitting the distribution to the probability density function (4.1) multiplied by the normalization factor $bN$, where $b$ is the bin width of the chi-square histogram and $N$ is the number of events in the Gaussian simulation. In the fitting function, the number of degrees of freedom is 126.2435.
Figure 20. Distribution (a) of the quadratic form $\frac{1}{2} \Delta y^\top \hat{Q} \Delta y$ and two-dimensional distribution (b): $\frac{1}{2} \Delta y^\top \hat{Q} \Delta y$ versus $\chi^2$. These distributions correspond to the events of the $e^+e^- \rightarrow \eta \pi^+\pi^-$, $\eta \rightarrow \gamma\gamma$ Gaussian simulation. These events were fitted under $e^+e^- \rightarrow \pi^+\pi^-\gamma\gamma$ hypothesis. All constraints of this hypothesis were enabled.

freedom and $N$ are considered as free parameters. The chi-square histogram in figure 19 is indicated by a solid line, while the fitting function is indicated by a dash-dotted line. The fitting parameters are shown in the figure inside the dash-dotted box. The fitting function indeed describes the histogram well. The number of degrees found with the fit is $\nu = 8.0015 \pm 0.0038$, i.e. it is equal to 8 with high accuracy.

To estimate the magnitude of the nonlinearity of the constraints, one can obtain the distribution of the quadratic form $\frac{1}{2} \Delta y^\top \hat{Q} \Delta y$, where $\hat{Q}$ is the $\hat{Q}_{S-1}$ from equations (2.8) and (2.9) calculated after the minimization is completed. As noted above, the matrix $\hat{Q}$ contains the constraint Hessians. In the case of linear constraints, this matrix is zero. The distribution of the quadratic form $\frac{1}{2} \Delta y^\top \hat{Q} \Delta y$ is shown in figure 20(a). This distribution was obtained under the conditions described in the previous paragraph. It can be seen from the figure that the mean value of the quadratic form $\frac{1}{2} \Delta y^\top \hat{Q} \Delta y$ is approximately equal to 1.7, which is significantly less than the mean value ($\nu = 8$) of the chi-square distribution. The standard deviation of the considered distribution is approximately equal to 0.49.

From equation (2.8) it is clear that in order to estimate the magnitude of the nonlinearity of the constraints, the quadratic form $\frac{1}{2} \Delta y^\top \hat{Q} \Delta y$ should be compared with the chi-square $\chi^2 = \Delta x^\top \hat{C}^{-1} \Delta x$. The corresponding two-dimensional histogram is shown in figure 20(b). It can be seen from the figure that the quadratic form $\frac{1}{2} \Delta y^\top \hat{Q} \Delta y$ is usually several times less than the chi-square. It is shown below that in those cases where the nonlinearity of the constraints leads to a significant distortion of the chi-square distribution, the $\frac{1}{2} \Delta y^\top \hat{Q} \Delta y$ distribution is significantly wider.

Chi-square distributions similar to those shown in figure 19 can be obtained when some of the constraints are disabled while fitting the events of the Gaussian simulation. So, for example, figure 21(a) shows the distribution in the case when the energy conservation constraint is disabled. Figure 21(b) shows the chi-square distribution in the case when the energy conservation and $z$-momentum conservation constraints are disabled. Figure 21(c) shows the chi-square distribution in the case when the energy, $x$-momentum, and $y$-momentum conservation constraints are disabled. Finally, figure 21(d) shows the distribution in the case when all four-momentum conservation constraints are disabled. If the constraints were linear, one would expect the listed distributions to follow the chi-square distribution (4.1) with $m - l$ degrees of freedom. Despite the fact that the constraints are nonlinear, the distributions in figures 21(a), 21(b), 21(c) and 21(d) are consistent
Figure 21. Chi-square distributions. These distributions correspond to the events of the $e^+e^- \rightarrow \pi^+\pi^-$, $\eta \rightarrow \gamma\gamma$ Gaussian simulation. The events of this simulation were fitted under the $e^+e^- \rightarrow \pi^+\pi^-\gamma\gamma$ hypothesis. During the fitting, some constraints were disabled. Figure (a) corresponds to the case when the energy conservation constraint is disabled. Figure (b) corresponds to the case when the energy and $z$-momentum conservation constraints are disabled. Figure (c) corresponds to the case when the energy, $x$- and $y$-momentum conservation constraints are disabled. Finally, figure (d) corresponds to the case when all four-momentum conservation constraints are disabled. The solid line indicates the $\chi^2$ histogram, while the dash-dotted line indicates the fitting function, similar to that used in figure 19. The parameters of the fitting functions, as well as the fit quality ratios $\chi^2$/NDF, are shown in the dash-dotted boxes.

with the chi-squared probability density function (4.1). The numbers of degrees of freedom in these cases are 7, 6, 5 and 4, respectively.

As mentioned above, a certain single event of the $e^+e^- \rightarrow \pi^+\pi^-\eta$, $\eta \rightarrow \gamma\gamma$ Monte Carlo simulation was chosen as the initial event. However, it should be noted that the fitting results given in this subsection do not depend on the choice of the initial event. However, in the cases where the nonlinearity of the constraints can manifest itself strongly, the results of the Gaussian simulation can significantly
Figure 22. Chi-square distributions corresponding to the events of the $e^+e^- \rightarrow K_S^0\pi^\pm$, $K_S \rightarrow \pi^+\pi^-$ Gaussian simulations. These events were fitted under the correct signal hypothesis: $e^+e^- \rightarrow X K^+\pi^-$, $X \rightarrow \pi^+\pi^-$ or $e^+e^- \rightarrow X K^-\pi^+$, $X \rightarrow \pi^+\pi^-$. The $\chi^2$ histogram is indicated by a solid line, while the dash-dotted line is the fitting function. This fitting function is proportional to the chi-squared probability density function (4.1). Figure (a) corresponds to the case of the Gaussian based on one of the Monte Carlo simulation events. Figures (b) corresponds to the case of the Gaussian simulation based on another Monte Carlo simulation event. Figure (c) corresponds to the case of the Gaussian simulation based on the third event of the Monte Carlo simulation. Finally, figure (d) corresponds to the case of the Gaussian simulation based on all Monte Carlo simulation events. The fitting function describes the histogram well only in figure (a).

depend on the choice of the initial event. Such a case is described in detail in the next subsection.

9.2 Hypotheses $e^+e^- \rightarrow XK^0\pi^\pm$, $X \rightarrow \pi^+\pi^-$

This subsection discusses the fitting results corresponding to a number of the $e^+e^- \rightarrow K_S^0K^\pm\pi^\mp$, $K_S \rightarrow \pi^+\pi^-$ Gaussian simulations. Each Gaussian simulation corresponds to a different initial
event. As in the previous subsection, the events of the corresponding Monte Carlo simulation (see section 8.1.3) were used as these initial events. Further, to obtain the events of each Gaussian simulation, a procedure similar to that described in the first paragraph of the previous subsection was performed. In the case of the examples presented below, one of the \( e^+e^- \rightarrow XK\pi^\pm \), \( X \rightarrow \pi^+\pi^- \) hypotheses is used, while fitting the simulated events. Moreover, when fitting each simulation event, only the correct signal hypothesis is used to avoid misassignments of kaons and pions.

Figure 22(a) shows the chi-square distribution corresponding to the events of the first Gaussian simulation. The parameters of the particles and their covariance matrices in the corresponding initial event are such that the effects of the constraints nonlinearity are not large enough to lead to a significant distortion of the chi-square distribution. It is seen from the figure that the chi-square histogram is well described by the probability density function (4.1). Moreover, the number of degrees of freedom is equal to 11, as expected.

In the case of the Gaussian simulation obtained using another initial event, the parameters of the particles in this initial event turned out to be such that they led to significant nonlinearity effects. The corresponding chi-square histogram is shown in figure 22(b). It can be seen from the figure that this histogram is not described by the probability density function (4.1).

Figure 22(c) shows the chi-square histogram for an even more dramatic case than in figure 22(b). This figure corresponds to the events of the third Gaussian simulation obtained using a third initial event different from the first two. In this case, the distribution is much wider than in figures 22(a) and 22(b). This distribution is not described by the probability density function (4.1).

Although the nonlinearity of the constraints tends to a wider chi-square distribution, even in the worst case shown in figure 22(c), the chi-square distribution peaks at \( \chi^2 \sim 10 \). Comparing this distribution with the chi-square distribution shown in figure 13(a) and obtained for background events, one can conclude that even in the worst case, signal and background events can be separated using a chi-square selection criterion. However, it is clear that the separation power is the smaller, the wider the signal chi-square distribution.

In connection with the above, the question arises, what is the ratio between the number of events in the Monte Carlo simulation, leading to the distributions in figure 22(a) and figures 22(b), 22(c). This question can be answered using a slightly different kind of Gaussian simulation than the one described above. Instead of doing a Gaussian simulation based on a single Monte Carlo simulation event, let us draw one Gaussian simulation event for each event of the Monte Carlo simulation. The chi-square distribution corresponding to such a Gaussian simulation is shown in figure 22(d). Although this distribution is not consistent with the probability density function (4.1), it can be concluded that the Monte Carlo simulation is dominated by events that correspond to distributions in figures 22(a) and 22(b).

Figure 23(a) shows the \( \frac{1}{2}\Delta y^\dagger \hat{Q} \Delta y \) distribution. This distribution corresponds to the Gaussian simulation in the case of figure 22(a). The mean value of this distribution is approximately equal to 2.8, which is significantly less than the mean value (\( \nu = 11 \)) of the corresponding chi-square distribution. The standard deviation of the considered \( \frac{1}{2}\Delta y^\dagger \hat{Q} \Delta y \) distribution is approximately equal to 2.3. Figure 23(b) shows the two-dimensional distribution of the \( \frac{1}{2}\Delta y^\dagger \hat{Q} \Delta y \) versus the \( \chi^2 \).

Figure 24 shows distributions similar to those shown in figure 23, but corresponds to the case of the second and third Gaussian simulations, i.e. figures 22(b) and 22(c), respectively. Figure 24(a) shows the \( \frac{1}{2}\Delta y^\dagger \hat{Q} \Delta y \) distribution for the events of the second Gaussian simulation. The standard deviation of this distribution is about two times larger than in the case of the distribution shown in figure 23(a). As shown
Figure 23. Distribution (a) of the quadratic form $\frac{1}{2} \Delta y^\top \hat{Q} \Delta y$ and two-dimensional distribution (b): $\frac{1}{2} \Delta y^\top \hat{Q} \Delta y$ versus $\chi^2$. These distributions correspond to the events of the $e^+e^- \rightarrow K_S K^\pm \pi^\mp$, $K_S \rightarrow \pi^\mp \pi^\pm$ Gaussian simulation. This Gaussian simulation corresponds to the case of the chi-square distribution shown in figure 22(a).

Figure 24. Distributions (a), (c) of the quadratic form $\frac{1}{2} \Delta y^\top \hat{Q} \Delta y$ and two-dimensional distributions (b), (d): $\frac{1}{2} \Delta y^\top \hat{Q} \Delta y$ versus $\chi^2$. These figures correspond to the events of the two different $e^+e^- \rightarrow K_S K^\pm \pi^\mp$, $K_S \rightarrow \pi^\mp \pi^\pm$ Gaussian simulations (second and third). Figures (a) and (b) correspond to the events of the second simulation, while figures (c) and (d) correspond to the events of the third simulation. The second Gaussian simulation corresponds to the chi-square distribution shown in figure 22(b), while the third Gaussian simulation corresponds to the distribution shown in figure 22(c).
in figure 22(b), such a deviation of the $\frac{1}{2} \Delta y \tilde{Q} \Delta y$ is sufficient to skew the corresponding chi-square distribution. Figure 24(b) shows the two-dimensional distribution of the $\frac{1}{2} \Delta y \tilde{Q} \Delta y$ versus the $\chi^2$ for the case of the second Gaussian simulation. Figures 24(c) and 24(d) show distributions similar to those shown in figures 24(a) and 22(b), respectively. Figures 24(c) and 24(d) correspond to the case of the third Gaussian simulation. Figure 24(c) shows that the distribution in this case is much wider than in the case of the first and second Gaussian simulations. In addition, this distribution is strongly asymmetric with respect to zero. Moreover, the long tail on the right slope of this distribution corresponds to large chi-square values. This can be seen from the two-dimensional distribution shown in figure 24(d).

We can conclude from the examples presented in section 9 that if the uncertainties of the parameters are of Gaussian nature and the nonlinearity of the constraints can be neglected, then the chi-square distributions are described by the corresponding probability density function with the expected number of degrees of freedom.

Another interesting question is what the pull distributions for vertex coordinates look like if the detector response is Gaussian and the covariance matrix $\tilde{C}$ is well calibrated. This question can be answered by applying the kinematic and vertex fitting procedure to Gaussian simulation events. Examples of such distributions are shown in figures 25 and 26. Figure 25(a) shows the pull distribution obtained for the $x$-coordinate of the interaction vertex. Figure 25(b) shows the pull distribution obtained for the $x$-coordinate of the $X \rightarrow \pi^+\pi^-$ decay vertex. Figures 26(a) and 26(b) show similar pull distributions corresponding to the $z$-coordinates of these vertices. The pull distributions shown in figures 25 and 26 were obtained for the same Gaussian simulation events as the distribution shown in figure 22(a). It is seen from figures 25 and 26 that the standard deviations corresponding to the distributions from these figures are close to one, as expected. This fact confirms that the kinematic and vertex fitting procedure correctly reconstructs the coordinates of the vertices.

**Figure 25.** Pull distributions for vertex coordinates. The distributions are obtained for the events of the same Gaussian simulation as the distribution shown in figure 22(a). The distribution shown in figure (a) corresponds to the $x$-coordinate of the $e^+e^-$ interaction vertex. The distribution shown in figure (b) corresponds to the $x$-coordinate of the $X \rightarrow \pi^+\pi^-$ decay vertex.
Figure 26. Pull distributions for vertex coordinates. The distributions are obtained for the events of the same Gaussian simulation as the distribution shown in figure 22(a). The distribution shown in figure (a) corresponds to the \( z \)-coordinate of the \( e^+e^- \) interaction vertex. The distribution shown in figure (b) corresponds to the \( z \)-coordinate of the \( X \to \pi^+\pi^- \) decay vertex.

Section 8.1.3 shows that the standard deviations in the case of pull distributions corresponding to the events of simulation (that takes into account the CMD-3 response) are significantly greater than one. This is mainly due to the non-Gaussian response of the detector and the imperfect calibration of the covariance matrix \( \hat{C} \). In section 8.1.3, this statement is demonstrated for the simulated events of the \( e^+e^- \to K_SK^+\pi^-, K_S \to \pi^+\pi^- \) process fitted to the \( e^+e^- \to XK^+\pi^-, X \to \pi^+\pi^- \) hypotheses. However, this statement is also true for events of other processes fitted to other hypotheses. Moreover, in the case of Gaussian simulations, the standard deviations corresponding to the pull distributions obtained for the events of these processes are always close to one.

10 Fitting package

The kinematic and vertex fitting package consists of two parts. The first part is the base fitting package KFBase. The KFBase package contains the optimizer class and an abstract base hypothesis class. This package also contains classes for all constraints, vertices and particles, except for charged particles and photons. The second part is the package of kinematic and vertex fitting for the CMD-3 experiment. This package is called KFCmd and depends on the base fitting package. The KFCmd package contains classes of charged particles and photons. It also contains a class responsible for reading input data, as well as classes with the implementation of some hypotheses. The packages KFBase and KFCmd, as well as their documentation, can be found in the repositories https://github.com/sergeigribanov/KFBase and https://github.com/sergeigribanov/KFCmd, respectively.

Both packages are written in the C++ (ISO/IEC 14882:2017) programming language and depend on a number of other packages such as ROOT [28] (version 6.26) and Eigen 3 [29] (version 3.4.0). The Eigen 3 header only library is used for matrix computations, while the ROOT framework is mainly used for file manipulations. The authors also provide a docker image that contains a number
of examples of using the fitting package. The examples can be run in Jupyter notebooks based on
the Python 3 kernel. The fitting package does not currently have a Python API. Instead of such an
API, the PyROOT [30] functionality is currently used. The repository with the docker image can be
found at the link https://github.com/sergeigribanov/kinfit-cmd3-docker.

11 Summary

In this paper, the authors review in detail the algorithm of operation and the capabilities of the
kinematic and vertex fitting package for the CMD-3 experiment. The paper provides a description
of the constraints available in the fitting package, as well as the parameterizations of particles and
vertices. The work provides a number of examples that show the operation of the fitting package using
most of the constraints and particle parametrizations. These examples consider hypotheses containing
charged particles and photons in the final state, hypotheses with intermediate and lost particles. It
should be noted that the fitting package can be easily extended with additional constraints, particle and
vertex parameterizations. Despite the fact that this package was developed for the CMD-3 experiment,
the authors believe that after some adaptation the package can be used in similar experiments.

In addition, the fitting package was tested using the Gaussian simulation technique. This kind
of testing is very useful, as it helps to eliminate some issues at the stage of package development.
Possible mistakes in the gradients and Hessians of constraints, as well as of the particle four-
momentum and trajectories, can be examples of such issues. Since some gradients and Hessians can
be very cumbersome, testing for this kind of issues requires special attention.

The authors consider the discussed kinematic and vertex fitting package as their first step towards
a more rigorous and universal fitting package. In the future, it is planned to pay special attention
to the optimization algorithm. The fitting package currently uses Newton’s method. This requires
inverting the Hessian matrix at each iteration of the optimization procedure. The more optimization
parameters the hypothesis contains, the more resource-intensive the Hessian inversion becomes. The
authors plan to consider the possibility of using quasi-Newtonian methods [22, 23], in which the
Hessian inversion does not occur at every iteration. In the case of hypotheses with a large number
of vertices, the authors also plan to study optimization algorithms based on the Kalman filtering,
similar to the one discussed in the papers [8, 9]. In addition, the authors plan to also explore the
possibility of using inequality constraints [22, 31, 32] in some kinematic and vertex fitting cases.

Acknowledgments

The authors are grateful to the members of the CMD-3 collaboration for discussion of the considered
topic and numerous suggestions. The authors also express special gratitude to the members of the
collaboration who are already actively using and testing the fitting package. The work described in
sections 2-7 is supported by Russian Science Foundation grant No. 19-72-20114. The work related
to the testing of the fitting package is supported by the Russian Foundation for Basic Research grant
No. 20-02-00496 A.
References

[1] J.M. Bauer, Kinematic fit for the radiative Bhabha calibration of BaBar’s electromagnetic calorimeter, Tech. Rep. SLAC-PUB-8650 (2000).

[2] K. Prokofiev and T. Speer, A kinematic and a decay chain reconstruction library, https://twiki.cern.ch/twiki/pub/CMSPublic/SWGuideKinematicVertexFit/CMS_IN_2004_20.pdf, 2004.

[3] K. Prokofiev and T. Speer, A kinematic and a decay chain reconstruction library, in proceedings of the 14th International Conference on Computing in High-Energy and Nuclear Physics, Interlaken, Switzerland, 27 September–1 October, 2004, pp. 411–414.

[4] L. Yan et al., Lagrange multiplier method used in BESIII kinematic fitting, Chin. Phys. C 34 (2010) 204.

[5] P. Jörg, The kinematically constrained fit, in Exploring the Size of the Proton by Means of Deeply Virtual Compton Scattering at CERN, Springer International Publishing (2018), pp. 61–76.

[6] J. Sundermann and T. Göpfert, KinFitter – A Kinematic Fit with Constraints, https://github.com/goepfert/KinFitter.

[7] R. Smith and J. Bishop, A Fast Universal Kinematic Fitting Code for Low-Energy Nuclear Physics: FUNKI_FIT, MDPI Phys. 1 (2019) 375 [arXiv:1911.10200].

[8] W.D. Hulsbergen, Decay chain fitting with a Kalman filter, Nucl. Instrum. Meth. A 552 (2005) 566 [physics/0503191].

[9] Belle-II Analysis Software Group collaboration, Global decay chain vertex fitting at Belle II, Nucl. Instrum. Meth. A 976 (2020) 164269 [arXiv:1901.11198].

[10] Y. Radkhorrami and J. List, Kinematic fitting for ParticleFlow Detectors at Future Higgs Factories, PoS PANIC2021 (2022) 398 [arXiv:2111.14775].

[11] F. Grancagnolo et al., Drift chamber for the CMD-3 detector, Nucl. Instrum. Meth. A 623 (2010) 114.

[12] B. Khazin, Physics and detectors for VEPP-2000, Nucl. Phys. B Proc. Suppl. 181–182 (2008) 376.

[13] R.R. Akhmetshin, D.N. Grigoriev, V.F. Kazanin, S.M. Tsaregorodtsev and Y.V. Yudin, Status of the endcap BGO calorimeter of the CMD-3 detector, Phys. Atom. Nucl. 72 (2009) 477.

[14] G.P. Razuvaev et al., Calorimetry at the CMD-3 Detector, in proceedings of CERN-BINP Workshop for Young Scientists in e^+e^- Colliders, Geneva, Switzerland, 22–25 August 2016, pp. 71–76.

[15] A.V. Anisenkov et al., Energy calibration of the barrel calorimeter of the CMD-3 detector, 2017 JINST 12 P04011.

[16] A.V. Anisenkov et al., Barrel calorimeter of the CMD-3 detector, Nucl. Instrum. Meth. A 732 (2013) 463.

[17] V.V. Danilov, P.M. Ivanov, I.A. Koop, I.N. Nesterenko, E.A. Perevedentsev, D.N. Shatilov et al., The concept of round colliding beams, Conf. Proc. C 960610 (1996) 1149.

[18] D. Berkaev et al., VEPP-2000 operation with round beams in the energy range from 1 GeV to 2 GeV, Nucl. Phys. B Proc. Suppl. 225-227 (2012) 303.

[19] A. Ioffe and V. Tihomirov, Theory of Extremal Problems, Studies in mathematics and its applications 6, North-Holland (1979).

[20] D.P. Bertsekas, Constrained optimization and Lagrange Multiplier methods, first edition, Academic Press (1982).

[21] E. Herman and G. Strang, Calculus, vol. Volume 3, OpenStax (2018).
[22] P.E. Gill, W. Murray and M.H. Wright, *Practical Optimization*, Emerald Group Publishing Limited (1982).

[23] A.F. Izmailov and M.V. Solodov, *Newton-Type Methods for Optimization and Variational Problems*, Springer Series in Operations Research and Financial Engineering, Springer International Publishing (2014), 10.1007/978-3-319-04247-3.

[24] N.K. Vishnoi, *Algorithms for Convex Optimization*, Cambridge University Press (2021), 10.1017/9781108699211.

[25] V.N. Bukov and S.V. Goryunov, *Inversion and canonization of block matrices*, *Math. Notes* **79** (2006) 614.

[26] S.S. Gribanov et al., *Measurement of the $e^+e^- \rightarrow \eta\pi^+\pi^-$ cross section with the CMD-3 detector at the VEPP-2000 collider*, *JHEP* **01** (2020) 112 [arXiv:1907.08002].

[27] SND collaboration, *Measurement of the $e^+e^- \rightarrow \eta\pi^+\pi^-$ cross section in the center-of-mass energy range 1.22-2.00 GeV with the SND detector at the VEPP-2000 collider*, *Phys. Rev. D* **91** (2015) 052013 [arXiv:1412.1971].

[28] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, *Nucl. Instrum. Meth. A* **389** (1997) 81.

[29] G. Guennebaud, B. Jacob et al., *Eigen v3*, https://eigen.tuxfamily.org, 2010.

[30] M. Galli, E. Tejedor and S. Wunsch, *A New PyROOT: Modern, Interoperable and More Pythonic*, *EPJ Web Conf.* **245** (2020) 06004.

[31] E.G. Birgin and J.M. Martínez, *Practical Augmented Lagrangian Methods for Constrained Optimization*, Society for Industrial and Applied Mathematics (2014), 10.1137/1.9781611973365.

[32] H.A. Eiselt and C.-L. Sandblom, *Nonlinear Optimization*, International Series in Operations Research and Management Science, Springer International Publishing (2019), 10.1007/978-3-030-19462-8.