The FIT-pull Method: an experimental tool to monitor the track measurements and the B proper time.

Public Note

Issue: 1
Revision: 0

Reference: LHCb-2007-032
Created: February 3, 2007
Last modified: June 19, 2007

Prepared by: G. Balbi, V. M. Vagnoni and S. Vecchi

The FIT-pull Method: an experimental tool to monitor the track measurements and the B proper time.

Abstract

In this note we describe a statistical tool, the *FIT-pull method*, that can test the reliability of the measurements of the tracks and the vertices on real and Monte-Carlo data without knowledge of the truth information. The basic mathematical formalism is derived from the Lagrange Multipliers method and briefly described. Several tests are performed to prove its validity in different situations. In particular, by using Monte-Carlo simulation, we demonstrate that the method can be applied to check if the measured tracks or vertices have biases or incorrect covariance matrices. For correct input measurements we obtain pull distributions with a normal Gaussian statistical form. In this case the B proper time value and its error, which is a function of the track and vertex measurements, are correctly calculated. However, in the case of incorrect measurements, for example due to a systematic error or to a scale factor of the covariance matrix, the pull distributions studied deviate from normal Gaussians and the B proper time measurement can be affected. In principle the method can, if necessary, be used to recover information from corrupted measurements. Its potential in this capacity is demonstrated for the particular case of the decay channel $B_d^0 \to \pi^+\pi^-$ with the reconstructed Monte-Carlo data produced in 2004.

Document Status Sheet

1. Document Title: The FIT-pull Method: an experimental tool to monitor the track measurements and the B proper time.			
2. Document Reference Number: LHCb-2007-032			
3. Issue	**4. Revision**	**5. Date**	**6. Reason for change**
Draft	1	February 16, 2007	First version. For internal reading only
Draft	2	March 20, 2007	Submitted to the LHCb editorial board.
Draft	3	May 15, 2007	Returned proofs from the LHCb editorial board.
Final	1	June 19, 2007	Final

Contents

List of Figures

1 Introduction

One of the most important features of the LHCb experiment is its ability to measure the B proper time very precisely ($\sigma_\tau \approx 40fs$) which is required to measure the fast $B_s/\bar{B}_s$ oscillations and to study precisely time dependent CP asymmetries. Hence, the correct measurement of the proper time of the B mesons and the evaluation of its uncertainty are key points in LHCb physics analyses.
The lifetime of a particular B meson, of mass M_B, can be calculated knowing its distance of flight (from the production vertex, PV to the decay vertex, SV) and its momentum ($\vec{\mathbf{p}}$):

$$\tau_B = \frac{M_B \cdot (\mathbf{P}\vec{\mathbf{V}} - \mathbf{S}\vec{\mathbf{V}}) \cdot \vec{\mathbf{p}}}{|\vec{\mathbf{p}}|^2 c} \ . \tag{1}$$

The PV is indirectly measured with a common vertex fit of all the tracks with segments in the Vertex Locator (VELO), while the SV and the B momentum are determined by a common vertex fit of the tracks from the stable B decay products.
In general we can say that the proper time is a function of the measured quantities, m_i, so $\tau_B = \tau_B(m_1, m_2, \cdots m_N)$. The error on the proper time can be obtained merely by error propagation, under the assumption that the m_i are Gaussian distributed:

$$\sigma_\tau^2 = \vec{\mathbf{J}_\tau} \cdot cov(\vec{m}) \cdot \vec{\mathbf{J}_\tau}^T \ , \tag{2}$$

where $\vec{\mathbf{J}} = (\frac{\partial \tau_B}{\partial m_1}, \ \frac{\partial \tau_B}{\partial m_2}, \ ..., \ \frac{\partial \tau_B}{\partial m_n})$ is the Jacobian and $cov(\vec{m})$ is the covariance matrix of the measurements.
Within the validity limits of these equations, we can calculate for each event the B proper time and its error. The proper time error provides an estimate of the resolution for each event.
Using the Monte-Carlo (indicated by MC in the following) simulated data it is straightforward to check the accuracy of the proper time estimate. This is performed by comparing the reconstructed value (τ_B) with the true B decay lifetime ($\hat{\tau}_B$). Hence we can calculate the statistical quantity, referred as MCPull in the following:

$${}^{MC}Pull(\tau_B) = \frac{\tau_B - \hat{\tau}_B}{\sigma_\tau} \ , \tag{3}$$

which is distributed as a normal Gaussian if the measurements and the errors are correct.
Clearly, on real data we cannot apply this statistical test since the true B decay lifetime is unknown. For this reason it is very important to develop some experimental tools which, at least indirectly, test or study the reliability of the measurements and the resolutions.
The LHCb collaboration has studied different strategies to retrieve the proper time resolution from

real data, for example by studying the proper time distribution of the $J/\psi \to \mu^+\mu^-$ produced directly in pp collisions [1] or analyzing the decays of the control channels $B \to J/\psi K^*$ and $B^\pm \to J/\psi K^\pm$[2]. The aim of these studies is to find a parametrization of the resolution as a function of different kinematical observables.
In this note we discuss the possibility of using a kinematical/vertex fitter as a tool to test the input measurements and, at least indirectly, the reliability of the proper time measurements. Such a tool, which can be applied to the experimental data, would be a useful building block for LHCb analyses.

This note starts with a general description of the kinematical and geometrical fits and the definition of the FITPull quantities (section 2). Then it discuss some specific features of the fits used in the LHCb applications (section 3). Section 4 and 5 report detailed studies of the FIT-pull method in case that the initial measurements are generated with a well-known distributions (*fake measurements*), and in the case when reconstructed MC informations are used, respectively. In section 5 we consider the possibility to use the FIT-pull method to recover corrupted input measurements by means of an iterative procedure that uses the FITPull distribution mean and sigma parameters. Finally, in section 6 we show how the reconstructed B proper time measurement, in the $B_d \to \pi^+\pi^-$ channel, can be affected by systematic errors or scale factors in the error estimate if the particles or vertex measurements are corrupted.

2 Constrained kinematical and geometrical fit

Constrained fits are widely used in high energy physics experiments to get the best estimates of some relevant information from a set of measurements, or for testing the compatibility of the data to a given hypothesis. They can be used for track reconstruction, vertex reconstruction or physics analyses, where one needs to select events of a given decay. Constrained fits can be solved by looking for the set of unknown parameters which minimize the χ^2 according to the measurements and to the given constraints. The mathematical/statistical formalism used is based on the least squares (maximum likelihood) and the Lagrange multiplier method, which will be briefly summarized in the next section.
Furthermore, in addition to this functionality, in some cases the constrained fits can also be useful to test the accuracy of the measurements. If the error distributions are Gaussian and the constraint equations are "quasi-linear" within the errors, the distribution of the normalized residuals (referred to in this note as the FITPull) are normal Gaussians (mean=0, sigma=1). Any deviation from the expected shape can be related to a wrong input measurement: for example a biased value or covariance matrix entries that differ from their true values by a scale factor. In this context, the constrained fits can also be useful for data and resolution calibration.

2.1 A statistics reminder: the Least squares and the Lagrange multipliers method

Let us consider a set of N independent measurements, m_i, of a given observable. Due to the finite experimental accuracy, the measured values deviate from the "true" ones, y, by a random amount defined by the error σ_i. If the error distribution $E_i = m_i - y$ is Gaussian with sigma σ_i, the best estimate of the true value $\hat{y}$ can be found by maximizing the likelihood as a function of y:

$$L(y) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma_i} exp(-\frac{(m_i - y)^2}{2\sigma_i^2}) \ , \tag{4}$$

or, equivalently, by minimizing the weighted sum of the distance squared (least squares):

$$S(y) = \sum_{i=1}^{N} \frac{(m_i - y)^2}{\sigma_i^2} \ . \tag{5}$$

In a more general approach we can consider several measurements of different observables that are related by functional relationships. In this case each measurement m_i corresponds to a true value y_i

that satisfies a set of constraint equations f_c. A dependence on additional parameters a_j, for which no direct measurement exists, can also be present:

$$f_c(y_1, \cdots, y_N, a_1, \cdots, a_p) = 0 \; , \qquad c = 1, \cdots, k \; . \tag{6}$$

The chisquare in equation (5) in this case transforms to the more generic expression:

$$S(\vec{\mathbf{y}}) = (\vec{\mathbf{m}} - \vec{\mathbf{y}})^T \mathbf{W} (\vec{\mathbf{m}} - \vec{\mathbf{y}}) = \Delta\vec{\mathbf{y}}^T \mathbf{W} \Delta\vec{\mathbf{y}} \; , \qquad \vec{\mathbf{y}}^{\mathbf{T}} = (y_1, \cdots, y_N) \; , \tag{7}$$

where $\mathbf{W}$ is the inverse covariance matrix associated to the measurements $\vec{\mathbf{m}}^{\mathbf{T}} = (m_1, \cdots m_N)$ ($\mathbf{W} = \mathbf{c}ov^{-1}(\vec{\mathbf{m}})$).

A simple way to include the equation constraints (6) in the least square search is through the Lagrange multiplier method: by introducing a new unknown scalar variable, the Lagrange multiplier λ_c, for each constraint, the method looks for the minimum of a linear combination of $S(\vec{\mathbf{y}})$ and $\vec{f}(\vec{\mathbf{y}}, \vec{\mathbf{a}})$ involving the multipliers as coefficients.

$$min\big(S(\vec{\mathbf{y}}) - 2\vec{\lambda}^T \vec{\mathbf{f}}(\vec{\mathbf{y}}, \vec{\mathbf{a}})\big) \; , \qquad \vec{\lambda}^{\mathbf{T}} = (\lambda_1, \cdots, \lambda_k) \; , \tag{8}$$

$$\begin{cases} \frac{\partial\left(S(\vec{\mathbf{y}}) - 2\vec{\lambda}^T \vec{\mathbf{f}}(\vec{\mathbf{y}}, \vec{\mathbf{a}})\right)}{\partial y_i} = 0 \; , & i = 1, \cdots, N \; , \\ \frac{\partial\left(S(\vec{\mathbf{y}}) - 2\vec{\lambda}^T \vec{\mathbf{f}}(\vec{\mathbf{y}}, \vec{\mathbf{a}})\right)}{\partial a_j} = 0 \; , & j = 1, \cdots, p \; , \\ \frac{\partial\left(S(\vec{\mathbf{y}}) - 2\vec{\lambda}^T \vec{\mathbf{f}}(\vec{\mathbf{y}}, \vec{\mathbf{a}})\right)}{\partial \lambda_c} = 0 \; , & c = 1, \cdots, k \; . \end{cases} \tag{9}$$

If the constraint equations have a linear dependence on the parameters they can be rewritten in the matrix form:

$$\vec{\mathbf{f}}(\vec{\mathbf{y}}, \vec{\mathbf{a}}) = \mathbf{B}\vec{\mathbf{y}} + \mathbf{A}\vec{\mathbf{a}} = \vec{0} \; , \tag{10}$$

where $\mathbf{B}$ and $\mathbf{A}$ are $k \times N$ and $k \times p$ matrices respectively. In this case the solution of the constrained least square can be found in one step by solving the linear system:

$$\begin{cases} \mathbf{W}\Delta\vec{\mathbf{y}} + \mathbf{B}^T\vec{\lambda} &= \vec{0} \; , \\ \mathbf{A}^T\vec{\lambda} &= \vec{0} \; , \\ \mathbf{B}\vec{\mathbf{y}} + \mathbf{A}\vec{\mathbf{a}} &= \vec{0} \; . \end{cases} \tag{11}$$

In the case of non linear constraints they can be linearised by using a Taylor expansion close to a "good enough" solution $(\vec{\mathbf{y}}^0, \vec{\mathbf{a}}^0)$ and the problem can be solved iteratively. In this case the matrices $\mathbf{B}$ and $\mathbf{A}$ can be approximated using first derivative of the constraint equations with respect to $\vec{y}$ and $\vec{a}$:

$$\vec{\mathbf{f}}(\vec{\mathbf{y}}, \vec{\mathbf{a}}) \approx \vec{\mathbf{f}}(\vec{\mathbf{y}}^0, \vec{\mathbf{a}}^0) + \begin{pmatrix} \frac{\partial f_1^0}{\partial y_1} & \cdots & \frac{\partial f_1^0}{\partial y_N} \\ \cdots & \cdots & \cdots \\ \frac{\partial f_k^0}{\partial y_1} & \cdots & \frac{\partial f_k^0}{\partial y_N} \end{pmatrix} \begin{pmatrix} \Delta y_1 \\ \cdots \\ \Delta y_N \end{pmatrix} + \begin{pmatrix} \frac{\partial f_1^0}{\partial a_1} & \cdots & \frac{\partial f_1^0}{\partial a_p} \\ \cdots & \cdots & \cdots \\ \frac{\partial f_k^0}{\partial a_1} & \cdots & \frac{\partial f_k^0}{\partial a_p} \end{pmatrix} \begin{pmatrix} \Delta a_1 \\ \cdots \\ \Delta a_p \end{pmatrix} + \cdots \; . \tag{12}$$

If we are close enough to the solution then at each iteration cycle, α, we can linearise the equations around the values found at the previous step:

$$\begin{aligned} &\vec{\mathbf{f}}(\vec{\mathbf{y}}^\alpha, \vec{\mathbf{a}}^\alpha) + \mathbf{B}^\alpha(\Delta\vec{\mathbf{y}}^\alpha - \Delta\vec{\mathbf{y}}^{\alpha-1}) + \mathbf{A}^\alpha(\Delta\vec{\mathbf{a}}^\alpha - \Delta\vec{\mathbf{a}}^{\alpha-1}) \approx 0 \; , \\ &\mathbf{B}^\alpha\Delta\vec{\mathbf{y}}^\alpha + \mathbf{A}^\alpha\Delta\vec{\mathbf{a}}^\alpha = \vec{\mathbf{c}} \; , \\ &\vec{\mathbf{c}} = \mathbf{B}^\alpha\Delta\vec{\mathbf{y}}^{\alpha-1} + \mathbf{A}^\alpha\Delta\vec{\mathbf{a}}^{\alpha-1} - \vec{\mathbf{f}}(\vec{\mathbf{y}}^\alpha, \vec{\mathbf{a}}^\alpha) \; , \end{aligned} \tag{13}$$

where $\vec{\mathbf{c}}$ is a "residual" value which stops the iteration when the desired accuracy is reached.

In this case the best estimates of the parameters are found by solving the linear system:

$$\begin{cases} \mathbf{W}\Delta\vec{\mathbf{y}} + \mathbf{B}^T\vec{\lambda} &= 0 \; , \\ \mathbf{A}^T\vec{\lambda} &= 0 \; , \\ \mathbf{B}\Delta\vec{\mathbf{y}} + \mathbf{A}\Delta\vec{\mathbf{a}} &= \vec{\mathbf{c}} \; . \end{cases} \tag{14}$$

The solution of this system is obtained by iterating the calculation:

$$\begin{pmatrix} \Delta\vec{\mathbf{y}} \\ \Delta\vec{\mathbf{a}} \\ \vec{\lambda} \end{pmatrix} = \begin{pmatrix} \mathbf{W} & \mathbf{0} & \mathbf{B}^T \\ \mathbf{0} & \mathbf{0} & \mathbf{A}^T \\ \mathbf{B} & \mathbf{A} & \mathbf{0} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \vec{\mathbf{c}} \end{pmatrix} = \begin{pmatrix} \mathbf{C_{11}} & \mathbf{C_{21}}^T & \mathbf{C_{31}}^T \\ \mathbf{C_{21}} & \mathbf{C_{22}} & \mathbf{C_{32}}^T \\ \mathbf{C_{31}} & \mathbf{C_{32}} & \mathbf{C_{33}} \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \vec{\mathbf{c}} \end{pmatrix} \; . \tag{15}$$

When the desired accuracy is reached ($|c_i| \leq \epsilon$) we get the best estimate of the parameters $\hat{\vec{\mathbf{y}}} = \Delta\vec{\mathbf{y}} + \vec{\mathbf{y}}^0 = \Delta\vec{\mathbf{y}} + \vec{\mathbf{m}}$ and $\hat{\vec{\mathbf{a}}} = \Delta\vec{\mathbf{a}} + \vec{\mathbf{a}}^0$.

2.2 Definition of the Pull quantities

Once the least square solution is found, the covariance matrix of the parameters can also be calculated. This is obtained by propagating the errors according to the Jacobian[3].
The covariance matrix has the form :

$$\mathbf{V}\begin{pmatrix}\hat{\vec{\mathbf{y}}}\\ \hat{\vec{\mathbf{a}}}\\ \hat{\vec{\lambda}}\end{pmatrix}=\begin{pmatrix}\mathbf{C_{11}} & \mathbf{C_{21}}^T & \mathbf{0}\\ \mathbf{C_{21}} & \mathbf{C_{22}} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & -\mathbf{C_{33}}\end{pmatrix}. \tag{16}$$

In particular we are interested in the covariance matrix of the parameters $\vec{\mathbf{y}}$:

$$\mathbf{V}(\hat{\vec{\mathbf{y}}})=\mathbf{W}^{-1}-\mathbf{W}^{-1}\,\mathbf{B}^T\,\mathbf{W_B}\,\mathbf{B}\,\mathbf{W}^{-1}+\mathbf{W}^{-1}\mathbf{B}^T\,\mathbf{W_B}\,\mathbf{A}\,{\mathbf{W_A}}^{-1}\,\mathbf{A}^T\,\mathbf{W_B}\,\mathbf{B}\,\mathbf{W}^{-1}\,, \tag{17}$$

and in the covariance matrix of $\boldsymbol{\Delta}\vec{\mathbf{y}}$ which turns out to be:

$$\mathbf{V}(\boldsymbol{\Delta}\vec{\mathbf{y}})=\mathbf{W^{-1}}-\mathbf{V}(\hat{\vec{\mathbf{y}}})\,. \tag{18}$$

At this point we can define the normalized "stretch values" or "FITPulls" as:

$${}^{FIT}Pulls(y_i)=\frac{\Delta\vec{y}_i}{\sqrt{\mathbf{cov}_{ii}-\mathbf{V}(\hat{\vec{\mathbf{y}}})_{ii}}}\,. \tag{19}$$

If the measured data are Gaussian distributed and the linearization of the equation constraint is a good approximation within the range spread by the measurements, the FITPulls turn out to be distributed as normal Gaussians ($\mu=0$, $\sigma=1$). Likewise, it is reasonable to expect that if one of the conditions above is not satisfied a deviation from normality in their shape should appear. This feature represents the key point of the calibration method we are proposing in this note.

3 Some useful cases in LHCb analyses

We have shown that constrained least squares provide not only the best estimates of some parameters and of the chisquare value but they also allow the calculation of the normalized FITPull quantities which, under precise conditions, have a well known distribution.
We want to use the FITPull distributions to check if the measurements ($\vec{\mathbf{m}}$) and their corresponding errors ($\mathbf{cov}$) are correctly determined. Of course the method must first be validated on the specific cases which LHCb will work on. In particular we have to understand whether, for these cases, the assumptions concerning the linearization of the constraint equations and the Gaussian distribution of the measurements are valid.

The LHCb collaboration has developed several kinematical/geometrical fitters which implement, in different ways, the constraint equations. In this note we will consider the *Global Fitter Tool* which was developed by the authors (V.Vagnoni, A.Carbone and G. Balbi). This tool has the advantage that it is able to calculate the FITPulls and the proper time error by correctly using the correlations given by the constraints and the full covariance matrices of the input measurements. The Global Fitter is a general purpose fit tool which aims to fit, in one call, a complete multi vertex decay tree. This approach leads to similar results to those obtained from other tools which fit the decay tree step by step in cascade. Depending on the complexity of the decay tree the number and types of constraint equations may change.

3.1 Global Fitter constraint equations

In the LHCb convention a particle is defined by 5 track parameters (the x, y coordinates at a given reference plane $z=\tilde{z}$, the slopes t_x, t_y in the (x, z) and the (y, z) planes and the momenta p) [1]. From

[1] This convention was valid for the MC data produced in 2005 and 2004, that we are considering in this note. Starting from the 2006 data production a particle is represented by 7 parameters: x, y, z p_x, p_y, p_z and E.

these definitions it is easy to obtain the components p_x, p_y, p_z and the energy E.

$$p_x = p\,\frac{t_x}{\sqrt{1+t_x^2+t_y^2}}\ , \quad p_y = p\,\frac{t_y}{\sqrt{1+t_x^2+t_y^2}}\ , \quad p_z = p\,\frac{1}{\sqrt{1+t_x^2+t_y^2}}\ , \quad E = \sqrt{m^2+p^2}\ . \tag{20}$$

Any B decay can be described by a nested tree of decays, each one defined by a vertex and two or more decaying particles. The particles and vertices may or may not be measured and the particles can themselves decay to other particles. For each 'decay-unit' the equations are:

$$\begin{aligned} &x^i - x_V - t_x^i(\tilde{z} - z_V) = 0\ , \\ &y^i - y_V - t_y^i(\tilde{z} - z_V) = 0\ , \\ &\vec{p}_X - \textstyle\sum \vec{p}^{\,i} = 0\ , \\ &M_X^2 - (\textstyle\sum_i E^i)^2 + |\sum_i \vec{p}^{\,i}|^2 = 0\ , \end{aligned} \tag{21}$$

where the first two equations state that particle i originates from (or decays at) the vertex V, the third one applies momentum conservation and the last one constrains the vertex mass to the original particle mass M_X. Momentum and mass conservation are applied only if they are appropriate for the decay-unit being considered. For example, for inclusive decay studies or resonance decays with a large width these are not applied.
We can immediately note that the constraint equations couple quasi-independent groups of parameters:

- 1) vertex constraints couple the track parameters x^i, t_x^i (y^i, t_y^i) to the vertex ones x_V, z_V (y_V, z_V),
- 2) the mass and momentum conservation constraints couple the slopes t_x^i, t_y^i to the momentum p^i of the particles involved in the decay.

In addition we note that the constraint equations are not linear. A Taylor linear expansion is performed close to a first estimate of the solution which is calculated from the measured values. These features will show up and be discussed later for specific examples in sections 4 and 5.
The number of degrees of freedom of the fit, N_{dof}, depends on the complexity of the decay tree, which is defined by the number of constraints equations k and by the number of unmeasured parameters p and is given by $N_{dof} = k - p$.
In the following section, to validate the FIT-pull method, we will consider in detail the channel $B_d^0 \to \pi^+\pi^-$. This channel is rather simple to reconstruct since it features two high p_T pions detached from the primary interaction vertex. This decay is represented in figure 1. In this decay topology we apply 12 constraints: the $\pi^+\pi^-$ common SV (4 constraints); the B originating in the PV and decaying at the SV (4 constraints); mass constraint (1); and momentum conservation (3 constraints). We have 8 unmeasured parameters: the SV (3 parameters); and the 5 B track parameters. Hence, the fit has 4 degrees of freedom.

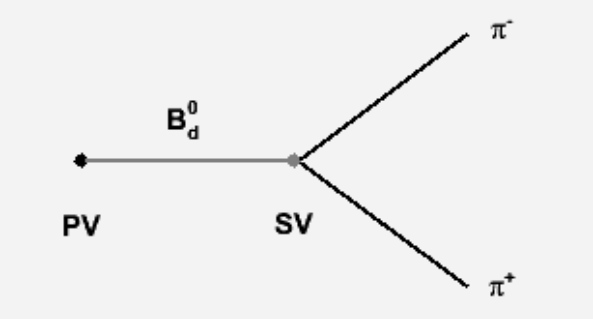

Figure 1 Schematic representation of the $B_d^0 \to \pi^+\pi^-$ decay channel: the measured quantities (PV and track parameters) are indicated in black, while in grey are the unmeasured ones (the secondary vertex SV and the B_d^0 parameters).

4 Validation of the FIT-pull method with input Gaussian distributions

In order to prove that the FIT-pull can be a valid tool to calibrate the tracks and the vertices in the LHCb analyses of real data we first have to show that in controlled situations they are distributed as *normal Gaussians*. For this reason we choose to initially work with perfectly Gaussian distributed measurements and pure signal events, avoiding the problems of background contamination or any non Gaussian dependence.

4.1 *Fake measurement* generation

For each test we consider a sample of ≈ 40000 MC events generated by the LHCb collaboration in 2004 (DC04 data production). For each event with a single pp collision we generate *fake measurements* by smearing the true generated particle and vertex information for the specific signal channel considered ($B \to \pi^+\pi^-$) according to a Gaussian resolution model. Correlations between different measurements can also be introduced as described further below. The Global fitter vertex tool is then applied with the appropriate configuration to test the specific decay tree. For each measurement we compute the FITPull, which is plotted on a histogram to study its statistical distribution. Since we only consider signal events and perfect Gaussian distributions, we apply a loose χ^2 cut ($\chi^2 < 1000$) to select events. This technique of generating fake measurement facilitates the test of the FIT-pull method in a range of situations, for example in the presence of systematic errors.

To obtain **Gaussian distributed "fake measurements"**, each measurement m_i is produced by smearing the MC true value t_i using the following equation:

$$\begin{array}{ll} m_i & = t_i + \sigma_i \cdot G^{\sigma=1}_{\mu=0} \ , \\ \mathbf{cov}_{ij} & = \rho_{ij} \cdot \sigma_i \cdot \sigma_j \ , \end{array} \tag{22}$$

where $G^{\sigma=1}_{\mu=0}$ is the normal Gaussian random generator and **cov** is the covariance matrix. Its elements are set to realistic values or parametric functions, which were obtained by MC studies on reconstructed particles and vertices:

$$\begin{array}{lll} \sigma_x = \sigma_y = \frac{\sigma_{IP}}{\sqrt{2}} = \frac{C_0 + C_1/p_t}{\sqrt{2}} \ , & C_0 = 0.014\, mm \ , & C_1 = 0.035 mm\ GeV \ , \\ \sigma_{t_x} = \sigma_{t_y} = C_2 \ , & C_2 = 0.4\, mrad \ , & \\ \sigma_p = C_3 |p| \ , & C_3 = 0.004 \ , & \\ \sigma_{V_x} = \sigma_{V_y} = C_4 \ , & C_4 = 0.010\, mm \ , & \\ \sigma_{V_z} = C_5 \ , & C_5 = 0.040\, mm \ . & \end{array} \tag{23}$$

If needed, the generation of correlated measurements is done with a simple linear transformation of independent measurements.

Adding a scale factor to the covariance matrix or a bias to a measurement.
If we wish to simulate the effect of introducing a bias, $BIAS_i$, to a measurement, m_i, or a scaling factor (SF_i) to the covariance matrix, equations (22) transform to:

$$\begin{array}{ll} m_i & = t_i + \sigma_i \cdot (G^{\sigma=1}_{\mu=0} + BIAS_i) \ , \\ \mathbf{cov}_{ij} & = \rho_{ij} \cdot \sigma_i \cdot \sigma_j / (SF_i \cdot SF_j) \ . \end{array} \tag{24}$$

In some of the tests performed, we set a charge dependent $BIAS_i$. By choosing $SF_i > 1$ we simulate the effect of an underestimated covariance matrix element.

Non Gaussian distributions.
In our studies we want to identify the validity limits of the method if the input measurements are not perfect Gaussians. In particular we consider the case of a generation of fake measurements according a double Gaussian distribution:

$$\begin{array}{ll} m_i & = t_i + \sigma_i \cdot (w \cdot G^{\sigma=1}_{\mu=0} + (1-w) \cdot G^{\sigma=3}_{\mu=0} + BIAS_i) \ , \\ \mathbf{cov}_{ij} & = \rho_{ij} \cdot \sigma_i \cdot \sigma_j / (SF_i \cdot SF_j) \ , \end{array} \tag{25}$$

which simulates the effect of the tails in the measurement distribution (see figure 2).

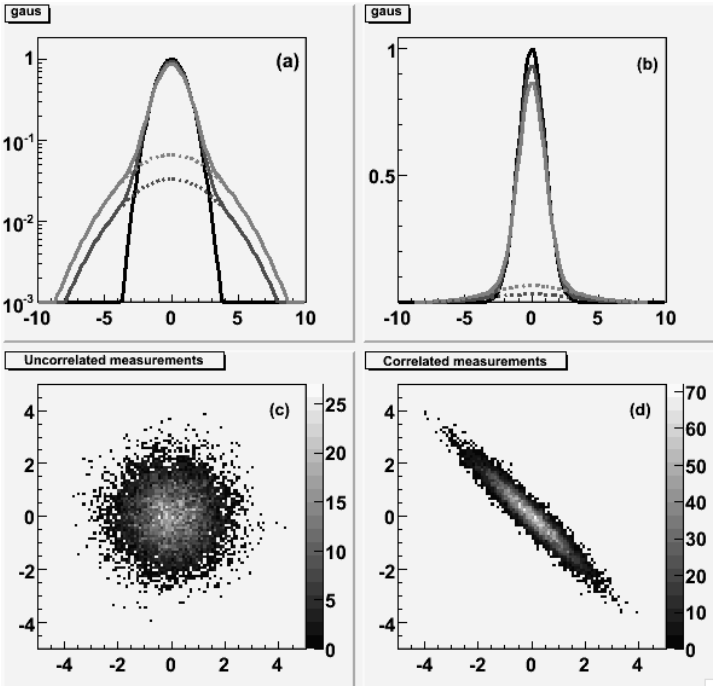

Figure 2 Fake measurement generation in case of single Gaussian (a) or double Gaussian distributions (b) ($w = 0.9$ in dark grey, $w = 0.8$ in light grey), for uncorrelated (c) or correlated measurements (d).

4.2 Results

In this section we test the FIT-pull method in a variety of input conditions in order to study the validity of the tool and characterize its performances. We start with some very simple tests and then we add further complexity later in this section.

4.2.1 Correct input data

By generating *fake measurements* according to equations (22) with $SF = 1$ and $BIAS = 0$ we test the fit performances in the ideal case. In this case, if the method is valid, we expect that the FITPulls follow a normal Gaussian distribution. In figure 3 for each input measurement (pion track parameters x, y, t_x, t_y, p and PV coordinates V_x, V_y and V_z) the mean (left) and sigma (right) values that fit the corresponding FITPull distributions are represented. The results agree with zero mean and unit sigma within $1 - 2\%$. Hence the method we are proposing is valid within this limit. In this situation the B proper time and error are correctly determined, as can be checked by a comparison with the MC generator true information. Figure 3, also reports the mean and the sigma values of the MCPull on proper time (see equation (3)), which are in perfect agreement with the expected values of 0 and 1.
This first test allows us to conclude that if the input measurements and the errors are correct, within $1 - 2\%$ the FITPulls are canonically distributed (normal Gaussian) and the proper time value and error are correctly calculated. This result is also achieved in the case where correlations between the measurements are included; for brevity this result is not shown.

4.2.2 Biased input data

If a vertex or a particle measurement is biased we would be able to identify it in the real data only if the FITPull associated to the corrupted measurement deviates from the canonical distribution. Following equation (24) we first test the effect of a single biased measurement on the FITPull distributions. In figure 4 are shown the results of tests where a bias on the particle's x (top), t_x (middle) and p (bottom) are introduced. In the first two cases a charge dependent bias is set.

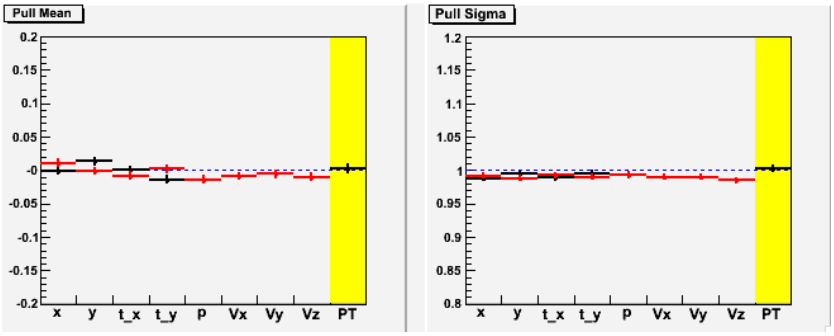

Figure 3 Graphical representation of the FITPull parameters: mean values (left) and sigma (right) of the FITPulls associated to each measurement associated to track (x, y, t_x, t_y, p) and PV (V_x, V_y, V_z), obtained by a Gaussian fit to the distributions. In the case of the track measurements the red and black points correspond to the positive and negative pions. On the yellow background are the MCPull values of the B proper time calculated with the fitted values. The input fake measurements are independently generated according to $BIAS = 0$ and $SF = 1$.

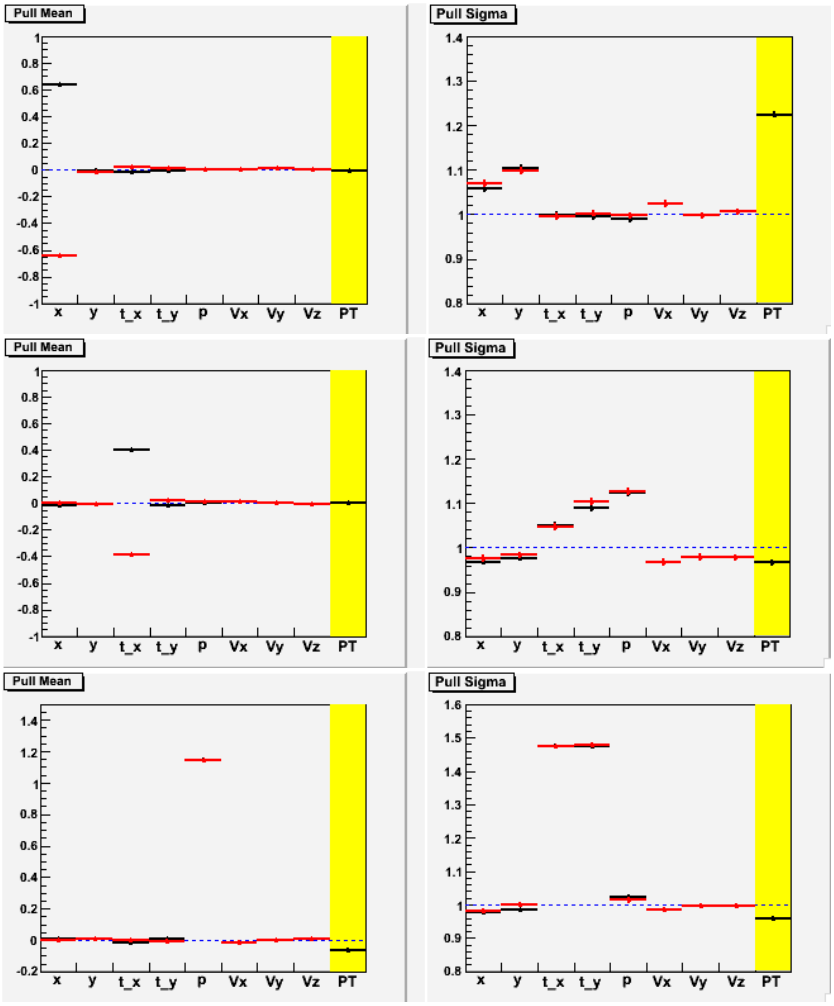

Figure 4 Graphical representation of the FITPull parameters (same graphical convention as in figure 3). The input fake measurement are independently generated with $SF = 1$ and $BIAS = \pm 1$ on x (top), $BIAS = \pm 1$ on t_x (middle) and $BIAS = 1$ on p (bottom). The sign of the input bias is given by the charge of the particle.

As one can see, in the case of a single biased measurement without correlations, the FITPulls easily identify the corrupted variable (mean value not equal to 0). The sign of the bias is correctly found from the mean, but the value is not equal to the bias: indeed, it depends on which measurement is corrupted, being 0.6, 0.4 and 1.2 in case of a x, t_x or p bias respectively. The bias also affects the FITPull sigmas, some of these now deviate significantly from unity, even though the input $SF = 1$. We have to keep in mind this effect in order to correctly interpret the FITPull outputs, this problem is discussed in more detail in section 5.
Note that a bias on y (t_y) gives results equal to the x (t_x) case, since the constraint equations are exactly symmetric.
From these tests it is also evident that the B proper time is not biased but, especially in the case of the input x bias, the calculated error has a scale factor, since the corresponding MCPull has a sigma $\neq 1$. If correlations $x - t_x$ and $y - t_y$ in the track parameters are considered[2] the FITPull output changes (see figure 5). Although the bias is only on x (correlated measurements don't necessarily mean correlated biases), the FITPull distributions have a mean$\neq 0$ on both x and t_x with reversed signs. The same thing happens in the case of a biased t_x. The reason for this behavior is the almost 100% correlation between the two measurements, which makes the fit mean values correlated even if the input ones are not. In fact we have to remember that in this way we are testing the constrained least square method in an unconventional way, given that one of the assumptions, that the input measurements are Gaussian distributed around the true value, is not fulfilled. Even in this case the proper time error show a sigma $ne1$ and, in case of a biased p, also a systematic deviation.
The effects of introducing a bias on PV measurements are plotted in figure 6. A bias on V_x also modifies the FITPulls x and t_x, while in the case of a bias on V_z only the FITPull of V_z is affected. This is due to the coupling introduced by the vertex constraint equations which, as we noted above, relate x and t_x to V_x. As a result the B proper time is biased if V_z is biased.
From the tests performed so far we can conclude that a bias on a measurement will make FITPulls appear not canonical. In some cases the corrupted FITPull identifies the corrupted measurement, but due to the couplings between variables this statement is not valid in general. In some cases, as expected, a bias on a measurement also affects the B proper time.

4.2.3 Scale Factor in the covariance matrix

In order to simulate an incorrect resolution we introduce a scale factor putting $SF = 2$ to alter the covariance matrix elements (see equation (24)) of x, t_x, p, V_x and V_z separately. The FITPull distributions are sensitive to the scale factor: in this case their sigma deviates from 1 as shown in figure 7, especially in the cases of SF on x, p and V_z. It should be noted that the mean values are still compatible with 0. In general any observed deviation from the canonical FITPull distribution is related to a wrong measurement error. However, a one-to-one correspondence between the measurement that has been corrupted and the affected FITPull distribution is present only in some cases.
Figure 7 also shows the effect on the B proper time resolution, which in the cases of SF on x, p and V_z, turns out to be affected significantly.

4.2.4 Double Gaussian error distribution

Usually the distribution of real measurements are only approximately Gaussian, since tails commonly show up. To be able to use the FIT-pull method on real data, we have to prove that tails do not modify the output distributions too much, or if they do, we need to establish the validity limits of our proposal.
In this case the fake measurements were generated by smearing MC truth informations according to double Gaussian distributions (see equation (25)), where a fraction w of events have correct measurements, while the remaining ones have under-estimated errors (by a factor 3). In this test all track parameters are modified simultaneously. In the previous tests we have seen that scale factors in the covariance matrix affect the FITPull sigmas, so we expect that the FITPull distributions will deviate by an amount which depends on the tail contribution. Indeed in the cases of $w \neq 1$ the FITPulls show a double Gaussian shape. An important result is that their main components are still canonical if the tail

[2]On MC data, by studying the MCPull correlation between the different track measurements at the *point on track* reference position, that is the extrapolation point closer to the beam line, we get $\rho_{x,t_x} = \rho_{y,t_y} = -0.95$.

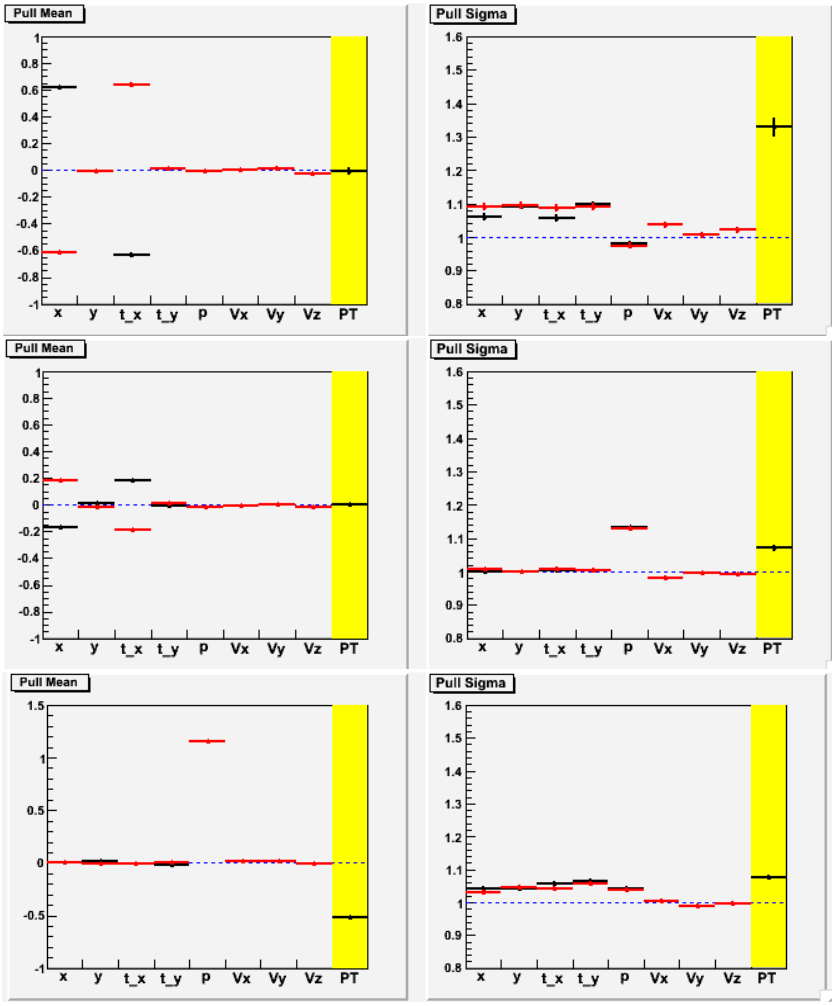

Figure 5 Graphical representation of the FITPull parameters (same graphical convention as in figure 3). The input fake measurements are generated with $SF = 1$ and $BIAS = \pm 1$ on x (top), $BIAS = \pm 1$ on t_x (middle) and $BIAS = 1$ on p (bottom) in case of correlated $x - t_x$ and $y - t_y$ track parameters. The sign is given by the particle's charge.

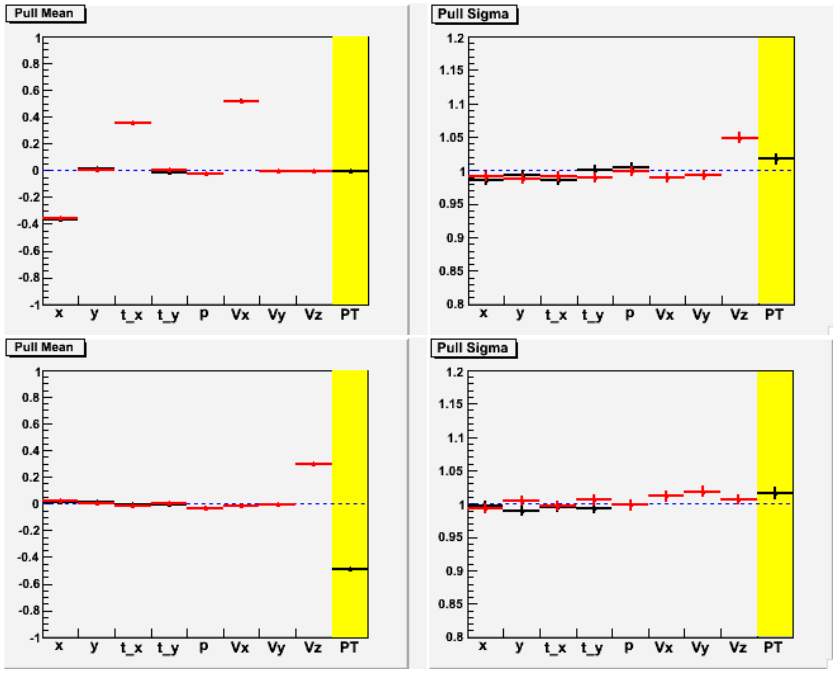

Figure 6 Graphical representation of the FITPull parameters (same graphical convention as in figure 3). The input fake measurement are generated with $SF = 1$ and $BIAS = 1$ on V_x (top) and V_z (bottom) in case of correlated $x - t_x$ and $y - t_y$ track parameters.

contribution does not exceed $10-15\%$, while they start to deviate significantly for larger contributions (see figure 8). So we can say that the method is still a good tool provided that the tail contributions are less than $\approx 10\%$. Figure 9 summarizes the dependence of the B proper time MCPull of its main and second Gaussian contribution as a function of w. Again, for $w \leq 10\%$ most of the events have the correct proper time and error.

We should emphasize that the second Gaussian contribution to the FITPull and to the B proper time MCPull distributions is due to the incorrect estimation of errors for the observed data. The FIT-pull method, being a "statistical" method, can only control a sample of several measurements and indicate whether most of the events are well measured. Nothing can be done to eliminate or correct data populating the second Gaussian. A cut on the fit χ^2 can only eliminate part of these events (the one populating the tails that contribute to the larger χ^2).

5 Validation of the FIT-pull method with the reconstructed tracks

In the previous sections we have shown that the FIT-pull method can be used to monitor the measured quantities and their estimated errors. However, real tracks will be a much less controlled enviroment to work with: phase space dependence, non Gaussian effects, correlations and background are some possible "complications" that can reduce the effectiveness of the FIT-pull method capability to test the measurements.
For this reason in this section we test the FIT-pull method in a more realistic scenario, by using the reconstructed tracks from the full LHCb MC simulation. Figure 10 shows the MCPull mean and sigmas as a function of the reconstructed momenta, for the pions produced in $B_d^0 \to \pi^+\pi^-$ decays[3].

In the DC04 simulation data a coding error occurred [4][4] and the reconstructed tracks show a systematic deviation from the true information, even if the track fitting was optimized. In particular x, t_x and

[3] In this case we request that the reconstructed particles are associated to the pions of the B decay (cheated selection) in order to evaluate MCPulls correctly

[4] The problem was due to a mismatch in the settings of the *deltaChord*, *deltaIntersection* and *deltaOnestep* parameters that control the simulation of a track inside the magnetic field and the physical volumes in the LHCb simulation code *Gauss*. The problem was solved in subsequent MC productions.

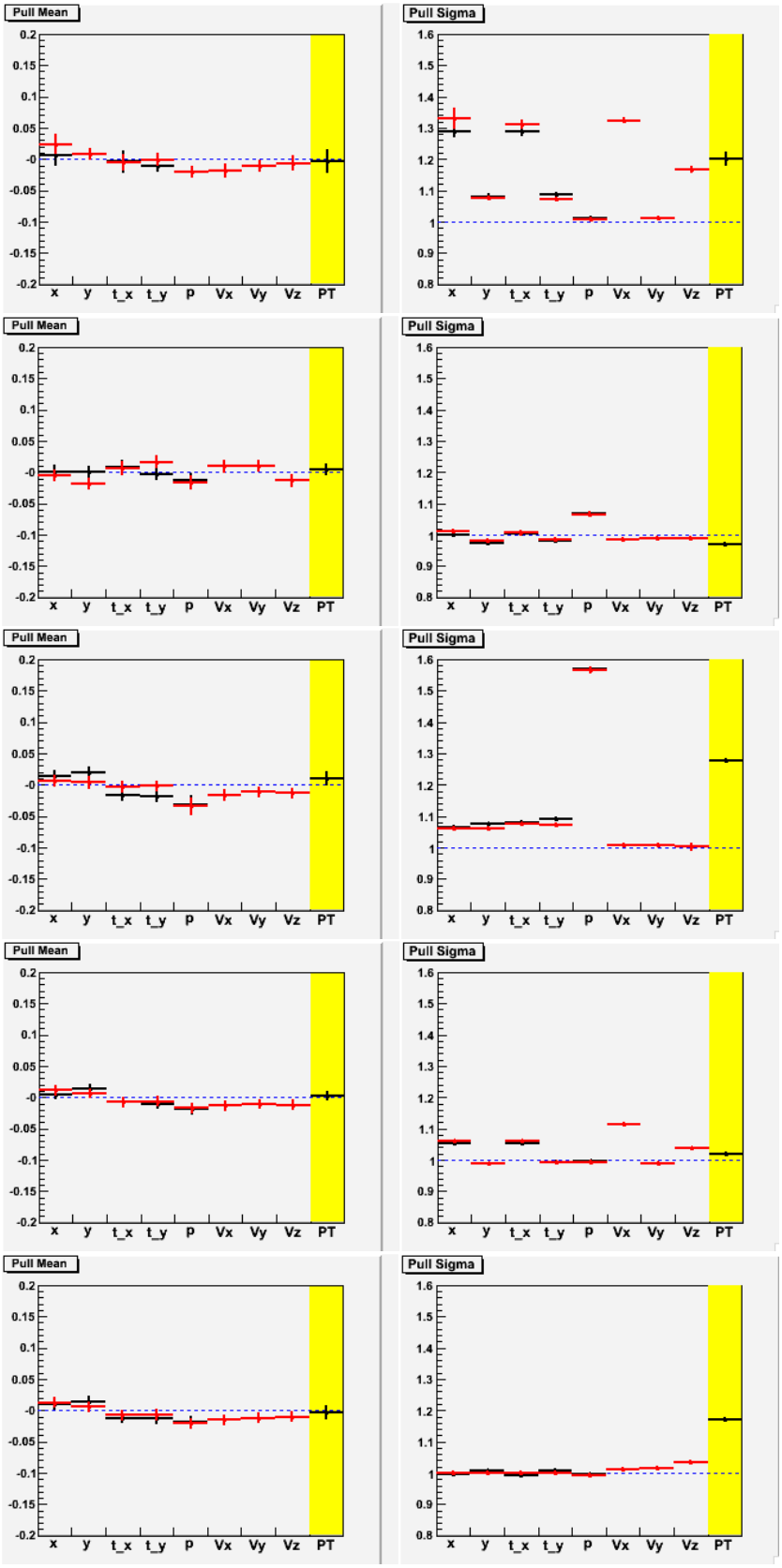

Figure 7 Graphical representation of the FITPull parameters (same graphical convention as in figure 3). From top to bottom rows: the input fake measurements are generated with $SF = 2$ and $BIAS = 0$ on x, t_x, p, V_x and V_z in case of correlated $x - t_x$ and $y - t_y$ track parameters.

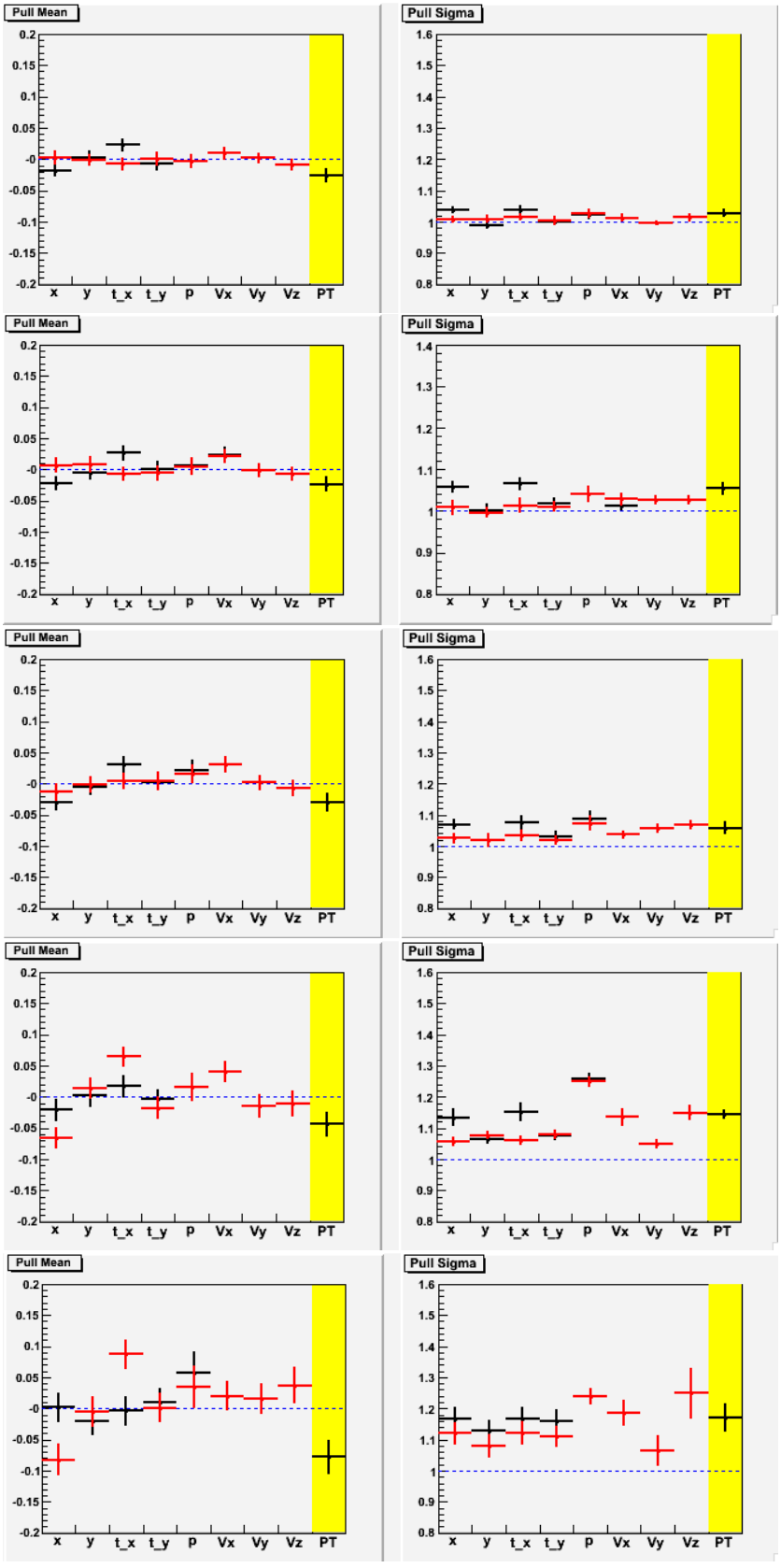

Figure 8 FITPull double Gaussian fit parameters corresponding to the main Gaussian contribution. The same notation of the previous figures is chosen. Input data were generated with increasing tail contribution in all track measurements: from top to bottom 5%, 10%, 15%, 30% and 40%.

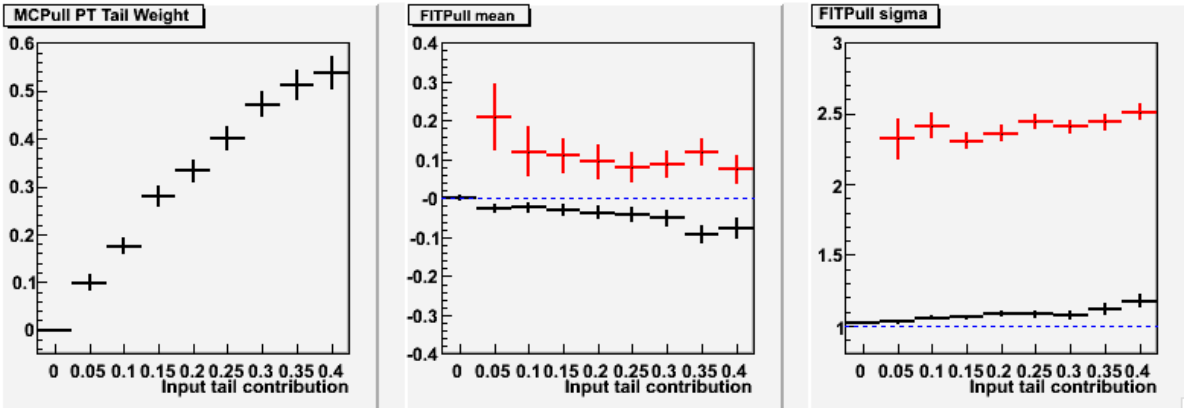

Figure 9 Parameters of a double Gaussian fit to Proper time MCPull in tests with double Gaussian distributed input measurements with different tail contributions. Left: second Gaussian contribution; center: mean values of the main (black) and second (red) Gaussian; right: sigma values of the main (black) and second (red) Gaussian. Their dependence is plotted as a function of the input tail contribution.

p show momentum and charge dependent biases. Moreover the reconstructed momentum is slightly underestimated ($SF \approx 1.1$).

However, we can take advantage of this coding error to see if the FIT-pull method monitor is able to observe this effect. From a sample of 100000 $B_d^0 \to \pi^+\pi^-$ MC data events we combine all the reconstructed $\pi^+\pi^-$ pairs and perform a fit with the hypothesis of $B_d^0 \to \pi^+\pi^-$ with the B_d^0 originating from the primary vertex[5]. Only combinatorial background is considered. Most of the bacground is suppressed by applying a cut of $\chi^2 < 10$, but still a fraction remains: the ratio of background to signal plus background events $B/(S+B)$ is ≈ 0.067. For simplicity, we consider here events with only one pp collision with the PV coordinates generated randomly around the true values, as in the previous section.

The FITPulls corresponding to each measurement were fitted with a double Gaussian shape in momentum slices and the parameters of the main Gaussian are represented in figure 11. As we expected, the bias on x, t_x and p shows up modifying the FITPull mean values. Both x and p biases are found with the correct sign, while in the case of t_x it appears with opposite sign. The reason for this could be connected to the $x - t_x$ correlation and the dominance of the x bias with respect to the t_x one. In fact simple tests with fake measurements have shown that a bias on x modifies x and t_x FITPull mean values by almost the same amount and with opposite sign (see figure 5). The same thing happens in the case of a bias on t_x. Quantitatively the absolute FITPull shift is smaller, so we can expect that the overall effect is dominated by the x bias and correlation. The deviation from 1 of the FITPull sigmas can also be explained as due to the input bias. In section 4.2.2 we observed that input biases modified also the FITPull sigmas.
Therefore we can conclude that also in this case the FIT-pull method is succesful in revealing some measurement errors. In this case also the B proper time would be incorrectly reconstructed: in figure 12, the left pad shows the MCPull distribution of the reconstructed proper time. It can be fitted with a double Gaussian distribution ($w \cdot G1 + (1-w) \cdot G2$), where the G2 component is mostly due to combinatorial background. The main component G1 shows an overall $\sigma \approx 1.27$ (parameter p3) which suggests that proper time error is underestimated, while the $BIAS$ is negligible $= 0.04\sigma_{PT} \approx 1.6fs$ (parameter p2). In the left plot it is evident that the sigma indeed depends on the pion momenta, like input measurement biases.

[5] In this case the standard selection is performed as it will be for the real data analysis

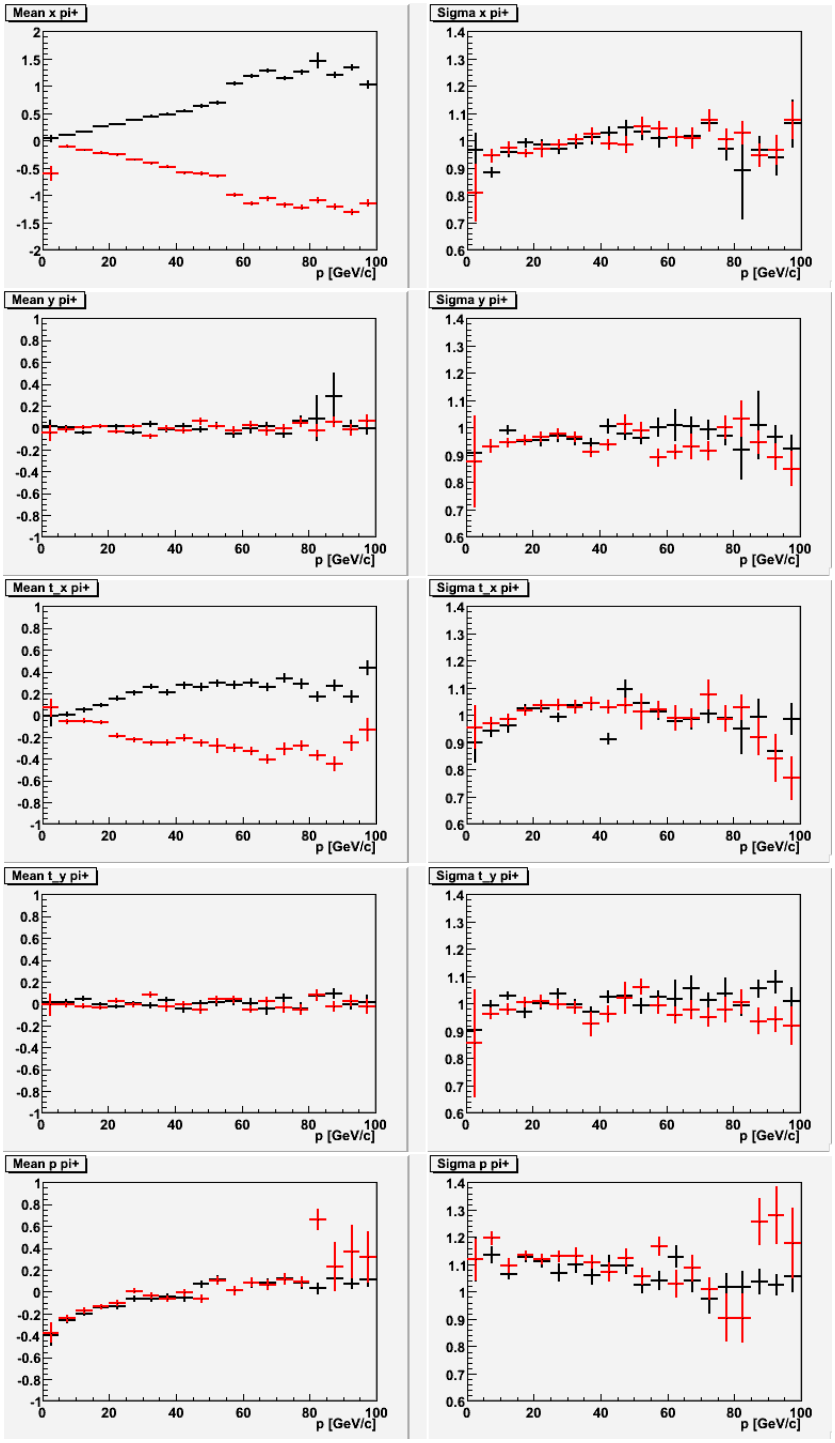

Figure 10 MCPull mean values (left) and sigma (right) associated to the reconstructed track $(x,\ y,\ t_x,\ t_y,\ p)$ measurements, obtained by a double Gaussian fit to the distributions: red and black data correspond to the main Gaussian contribution for π^+ and π^- respectively.

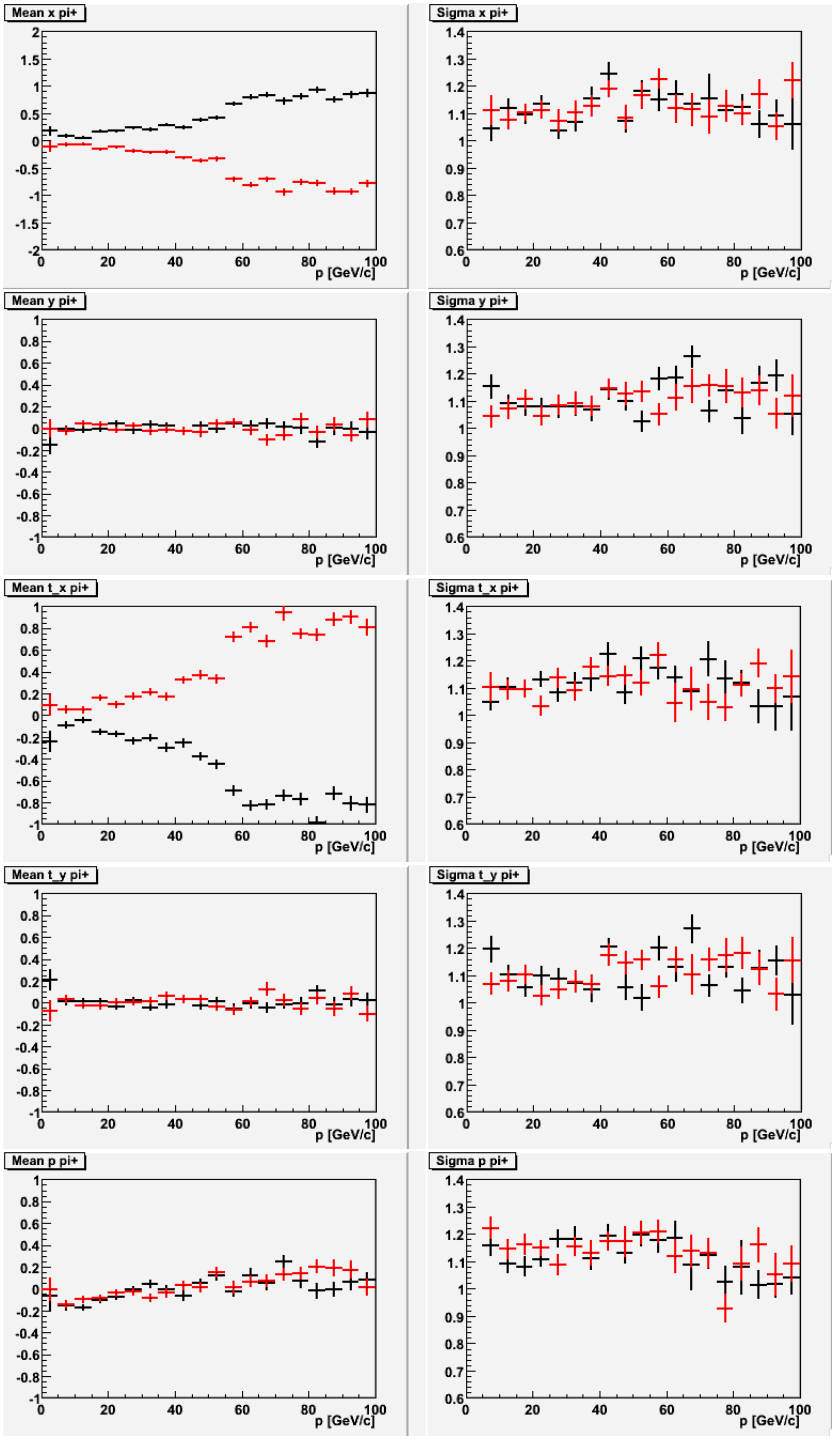

Figure 11 FITPull mean values (left) and sigma (right) associated to the reconstructed track $(x,\, y,\, t_x,\, t_y,\, p)$ measurements, obtained by a double Gaussian fit to the distributions: red and black data correspond to the main Gaussian contribution for π^+ and π^- respectively.

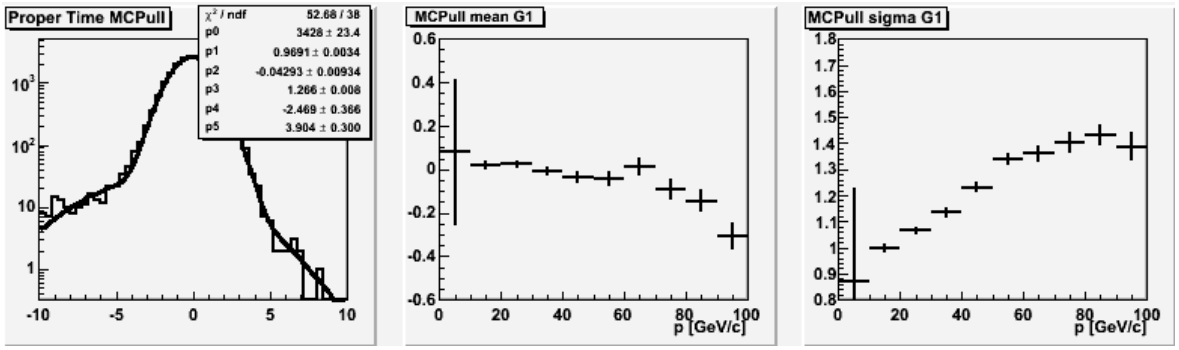

Figure 12 B proper time MCPull distribution (left) of the $B \to \pi^+\pi^-$ events selected with a $\chi^2 < 10$. Center and right pads show the mean and sigma MCPull parameters of the main Gaussian as a function of the pion momenta.

6 Potential for recovery of the measurements

All the tests done so far demonstrate that the FIT-pull method can be used on real data to investiagte the distribution of measurement values and errors down to fraction of resolution scale. Therefore it is a rather sensitive monitor of the measurement reliability. With some limits, the FITPulls can also indicate which kind of problem (bias or scale factor) the measurements may have. If the problem concerns the track parameters then ideally further optimization of track fitting may correct it. However, if this does not happen, we can investigate the possibility to use the FIT-pull method to recover the wrong measurements.
Given the correlation between the FITPull output we observed, and the fact that quantitatively the FITPulls mean and sigma do not represent the input bias or scale factor, we choose to apply iteratively correction cycles in which, at each step k, an input measurement (m_i^k, cov_{ij}^k) is corrected by the following equations:

$$\begin{aligned} m_i^k &= m_i^{k-1} + \mathbf{B_i^{k-1}}\sqrt{\mathbf{cov_{ii}^{k-1}}}\,, \\ \mathbf{cov_{ij}^{k}} &= \mathbf{S_i^{k-1}S_j^{k-1}cov_{ij}^{k-1}}\,, \end{aligned} \tag{26}$$

where $\mathbf{B_i^{k-1}}$ and $\mathbf{S_i^{k-1}}$ represent respectively the FITPull mean and sigma obtained at iteration cycle $k-1$ for the measurement i. The iteration cycle stops when the FITPull mean and sigma values reasonably agree with 0 and 1. Since at each iteration we correct the updated measurements, the final n$-th$ correction to the input bias is approximately the sum of the single $\sum_{k=0}^{n}\mathbf{B_i^k}$, while the correction factor to the input scale factor will be given by the $\prod_{k=0}^{n}\mathbf{S_i^k}$.
Driven by the experience obtained with the tests on fake measurements, we decide to apply first corrections to fix the biases, then to recover the scale factors. Moreover, given the self correlations between track measurements, we choose to correct x first, then t_x and p. With these criteria, we consider the case of reconstructed tracks of the previous section. The FITPull mean values shown in figure 11 are linearly fitted as a function of the particle momentum and the best parameters fed to the correction cycles. During the bias correction we also observe an improvement of the sigmas. After a few ($n \approx 10$) correction cycles the FITPull parameters are reasonably ($\leq 10\%$) compatible with mean$= 0$ and sigma$= 1$. In this situation, we can judge if the correction worked by looking at the MCPull distributions: figure 13 summarizes all the track measurements corrected by the procedure described. We can see that, except t_x bias, which is still biased, all the other measurements recovered almost completely: the mean and sigma values are correct at $\sim 10\%$ level. In this situation the B proper time value and error are better calculated, being the MCPull $\sigma = 1.04$ and its momentum dependence almost cancelled, as is shown in figure 14.

7 B proper time resolution and calibration

So far we have shown that the FIT-pull technique can be used to test if the measurements and errors of the measured event parameters in a given decay channel are distributed as expected. If they are

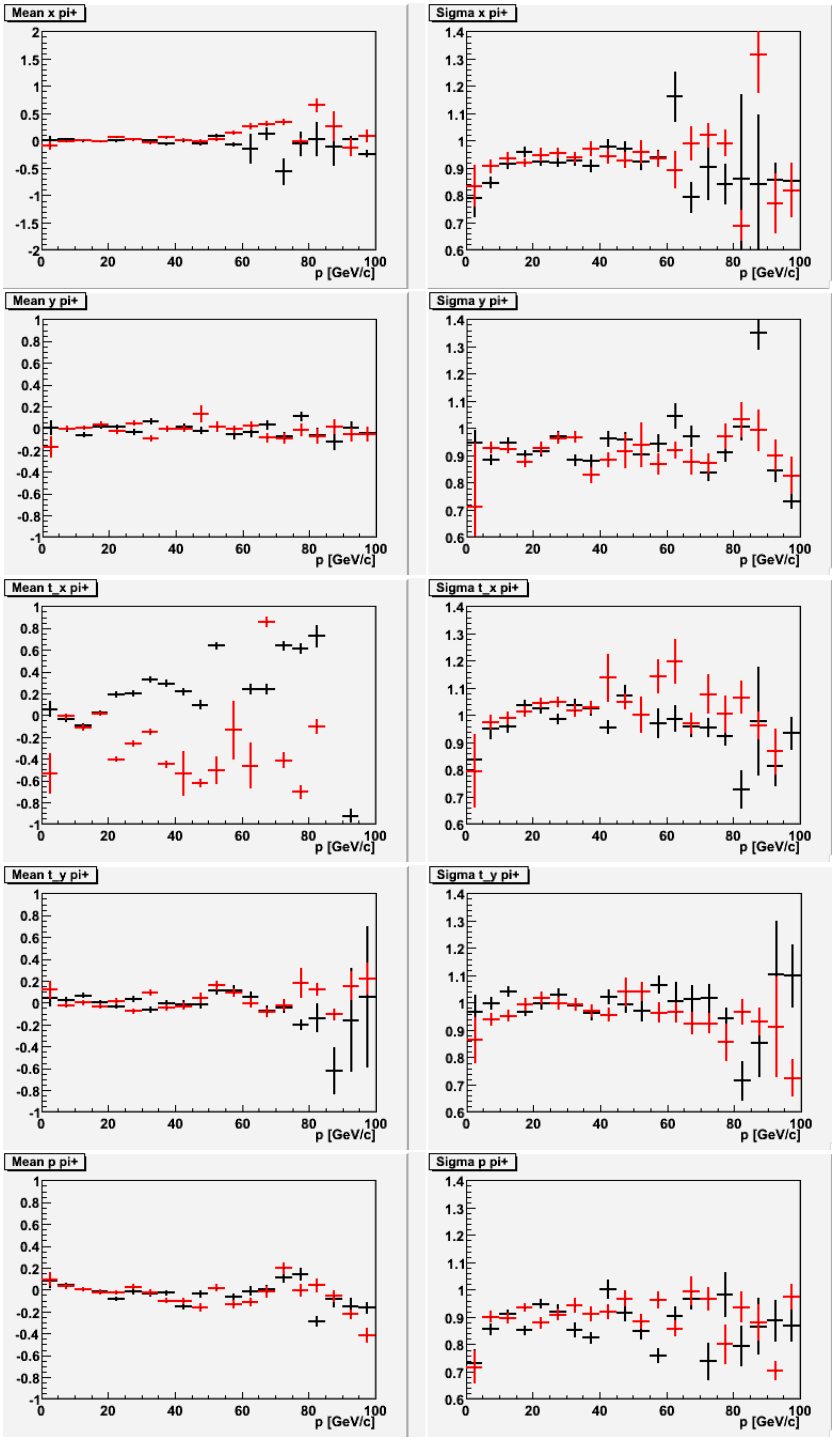

Figure 13 MCPull mean values (left) and sigma (right) associated to the reconstructed tracks $(x,\, y,\, t_x,\, t_y,\, p)$ measurements after the correction cycles based on the FITPull distributions. Values are obtained by a double Gaussian fit to the distributions: red and black data correspond to the main Gaussian contribution for π^+ and π^- respectively.

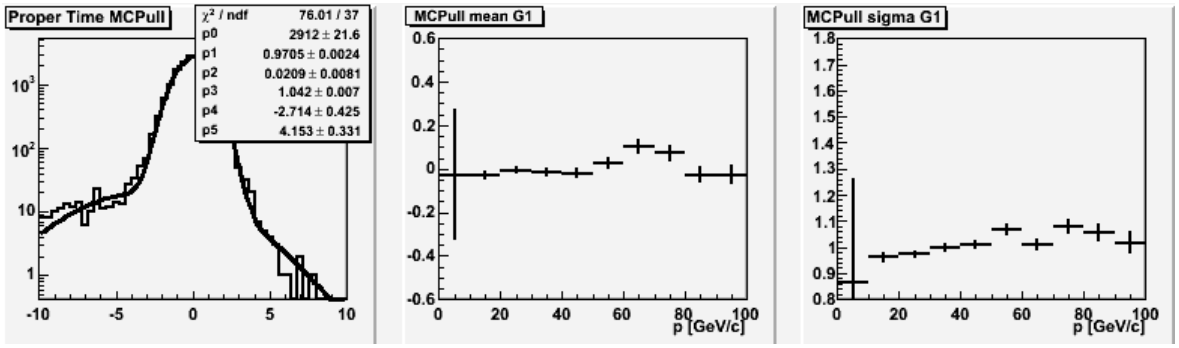

Figure 14 B proper time MCPull distribution and parameters after the correction.

canonically distributed they guarantee that the input measurements are correct and, indirectly, that the B proper time measurement is reliable. On the other hand, if the FITPulls are not canonical it means that some input measurements, or their error estimates, have a problem, which may also affect the proper time measurement.

In this section we want to provide a more quantitative study on the dependence of the proper time calculation on the input measurement bias or scale factor. For this reason, again utilising the fake measurement generation technique, we vary the SF and the $BIAS$ of the most relevant measurements and we plot the proper time MCPull parameters mean and sigma.

The plots in the left column in figure 15 summarize the proper time dependence on the input bias (ranging from -2.0 to 2.0) of single measurements V_z, x, t_x and p. As previously, in the case of x and t_x the bias observed in the means is charge dependent. Correspondingly in the right column of the figure are plotted the results of scale factor dependence (with values 0.2, 0.5, 1.0, 2.0 and 5.0).
As we already commented in section 4, the proper time has a different response to the applied bias or scale factor. Assuming that to have a good B proper time measurement the MCPull mean should be in the range $[-0.1, 0.1]$ and the sigma within $[0.9, 1.1]$, these plots can help us to to fix some limits to the input $BIAS$ or SF that can be tolerated. Furthermore, the MC study can translate these limits into requirements on the FITPull parameters that should be achieved on real data.

Conclusions

In this note we have described a technique to use kinematical/geometrical fits and their output FITPulls to test the input measurements and their errors on real data. The distributions of the FITPulls are normal Gaussians if the input measurements and errors are correctly defined, while they deviate from being normal Gaussians in the presence of systematic deviations ($BIAS$) or scale factors (SF) in the input measurements or errors. Unfortunately the determination of the affected measurement is not always unique due to the the fit constraints and the correlations between the measurements.
The validity of this method has been checked in several tests performed with MC data corresponding to the decay channel $B_d^0 \to \pi^+\pi^-$.
The studies made on data generated according to a well known distribution (fake measurements) allowed us to understand the features and the limits of the method in a simple way. In the case of Gaussian distributed measurements without correlation a biased input measurement can be easily identified since it modifies the mean value of the only corresponding FITPull. In this case the mean value is proportional to the input bias, but the coefficient depends on the type of measurement. For this reason a quantitative evaluation of the input bias is not straightforward. In case of correlations between the input measurements the one to one correspondence between the incorrect measurement and the non canonical FITPull remains valid only in case of p track measurement. A scale factor to the input measurements shows up as a $\sigma \neq 1$ of the FITPull distribution, but, as in the previous case, a precise and quantitative evaluation of the input SF is not possible. However, also in the case that the input measurements have a double Gaussian distribution, like in most of more realistic cases, the FITpull method is successful in monitoring the correctness of the main Gaussian component provided that the secondary Gaussian contribution is less than 10%.

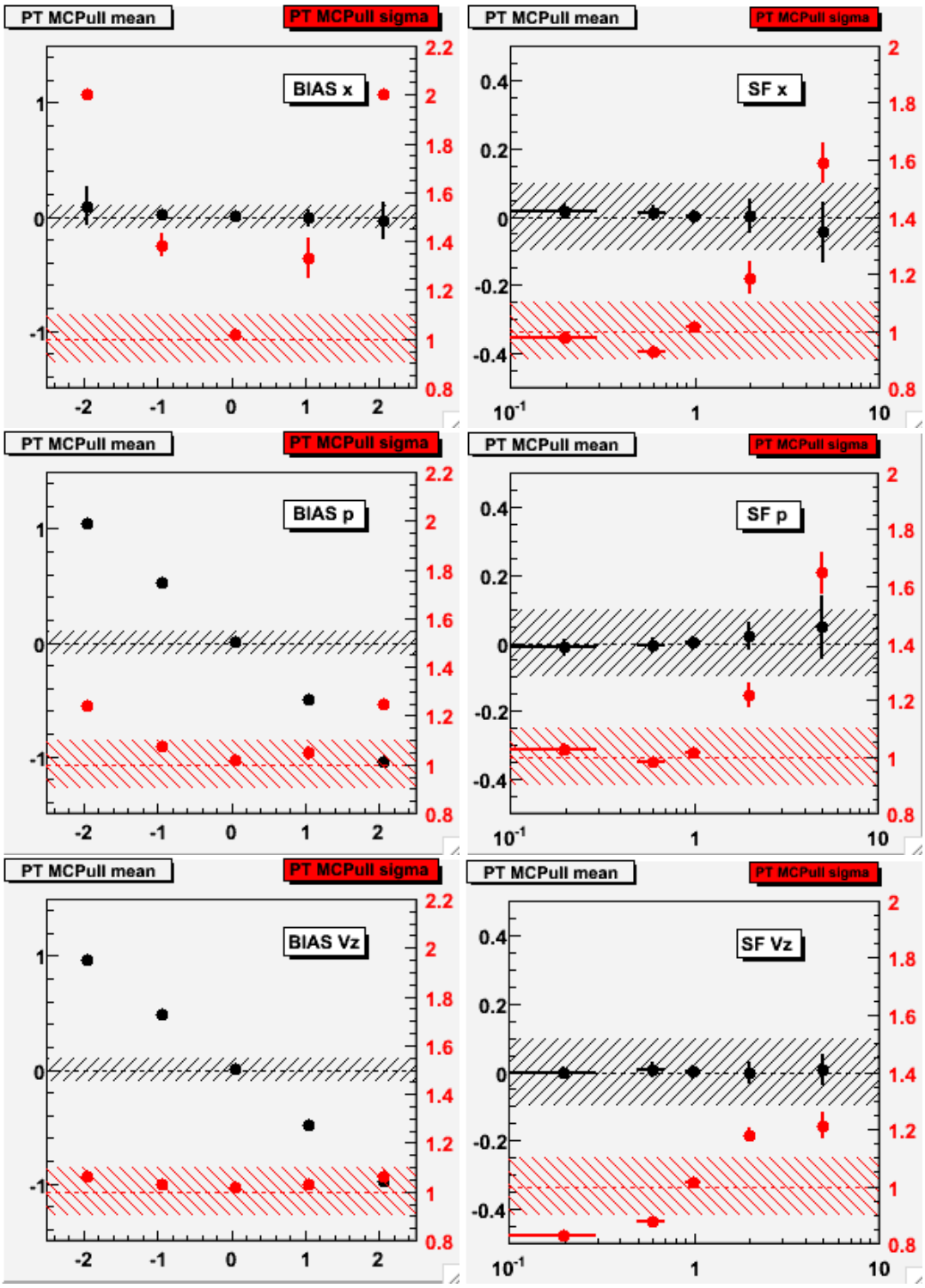

Figure 15 B proper time MCPull parameters mean (black) and sigma (red) as a function of input biases (left plots) and scale factors (right plots) on track measurement x, t_x, p and on vertex. The simulated events correspond to $B \to \pi^+\pi^-$ channel and the measurements are obtained by a Gaussian smearing of the MC truth informations (Fake measurements). The dashed areas correspond to the range of mean and sigma MCPull values that we assume as tolerance level for good proper time measurement.

Tests on reconstructed MC simulation data for the decay channel $B_d^0 \to \pi^+\pi^-$ have demonstrated that also in a more realistic case the FIT-pull monitor is useful to discover incorrect measurements. In this case we have also shown that a partial recovery of the corrupted measurements can be obtained by means of an iterative correction procedure. The input biases on x and p can be corrected up to a mean value of $\sim 10\%$ of the input errors, while the bias on the track slope t_x cannot be eliminated due to its strict correlation with the x track measurement. In this case we have shown that the B_d^0 proper time also improves, having a MCPull of $\sigma = 1.27$ before and $\sigma = 1.04$ after the correction.
Of course a precise determination of the FITPull parameters requires large statistics (especially if one wants to study their phase space dependence) and low background contamination. For this reason the use of the FIT-pull method to monitor or to recover measurements cannot be performend on the signal decay itself. In a note in preparation we will discuss the possibility of using the FIT-pull method in the control channel $J/\psi \to \mu^+\mu^-$ on real data to monitor the input measurements (tracks and vertices). This approach has the advantage that a high statistic and low background sample can be processed. In case the control sample analysis identifies track or vertex biases, or errors in the uncertainties, a review of the track fit or the vertex fit will be performed first. In the case that is not possible to recover the problem at its origin, we can potentially use the corrections found by the FIT-pull method in the control channel and export them to the physics channel.

Aknowledgments

The authors would like to thank Jeroen Van Hunen for the useful suggestions and support for this study and Chris Parkes for his patient and precise corrections.

8 References

[1] B.Carron PhD Thesis, EPFL Lausanne (2007)

[2] P. Vankov and G. Raven, Proper time resolution modeling, LHCb-2007-55, Public note.

[3] R. Fruhwirth, et al., Data analysis techniques for high-energy physics, Cambridge monographs on particle physics, nuclear physics and cosmology, 2000;
R.K. Bock, et al., Formulae and methods in experimental data evaluation, European Physical Society, 1984.

[4] M.Merk and J. Van Tilburg, private communication.