DIRAC NOTE 2008-07

How to extract the lifetime of pionium and $\left|a_0^0 - a_0^2\right|$ from the measurements of the pionium ionization probability

D. Drijard
CERN

M. Zhabitsky
JINR, Dubna, Russia

September 18, 2008

Abstract

The goal of this note is to describe the algorithms used in DIRAC to estimate the lifetime τ and the difference in scattering lengths $\left|a_0^0 - a_0^2\right|$. This note could then be used as a reference, saving lengthy explanations in the physics publications.

The results of the analysis consist of a set of N measurements of the breakup probabilities p_i, namely m_i $(i = 1, N)$ (see Tab. 1). The parent probability density function (pdf) of each measurement m_i is assumed to be a Normal law of mean p_i and standard deviation σ_i:

$$f_i(m_i|p_i, \sigma_i) = \frac{1}{\sigma_i\sqrt{2\pi}} \exp\left(-\frac{(m_i - p_i)^2}{2\sigma_i^2}\right). \tag{1}$$

Measurements corresponding to common experimental conditions (i. e. common expected breakup probabilities) have been regrouped using the standard average of the measurements. This is not a necessity, all measurements may be treated on the same footing without affecting the results. Only the statistical uncertainties σ_i are used in this calculation.

$$m_i = \frac{\sum_j m_{ij}/\sigma_{ij}^2}{1/\sigma_i^2}, \qquad \frac{1}{\sigma_i^2} = \sum_j \frac{1}{\sigma_{ij}^2}. \tag{2}$$

Table 1: Break-up probabilities and corresponding estimations $\hat{\tau}_i$ for different data periods

period	m_i	$\sigma_{\mathrm{stat}\,m}$	$\hat{\tau}_i$, fs	$\overline{\sigma}_{\mathrm{stat}\,\tau}$, fs	$\underline{\sigma}_{\mathrm{stat}\,\tau}$, fs
Ni 94 μm, 24 GeV/c [a]	0.4587	0.0224	3.1459	0.4804	0.4162
Ni2001b	0.4335	0.0390	2.5497	0.6961	0.5581
Ni2002_24	0.4124	0.0315	2.2333	0.4866	0.4110
Ni 98 μm, 24 GeV/c	0.42073	0.02451	2.3539	0.3886	0.3399
Ni2002_20	0.4643	0.0318	3.1658	0.6880	0.5656
Ni2003_20	0.4158	0.0412	2.3419	0.6845	0.5477
Ni 98 μm, 20 GeV/c	0.44619	0.02517	2.8309	0.4759	0.4108

* Experimental values correspond to "v4" analysis [1]

** Hereafter we keep excessive number of digits to ease a cross-check between independent calculations

[a] $B(\tau)$ is scaled by the factor 1/1.014 (target impurities correction)

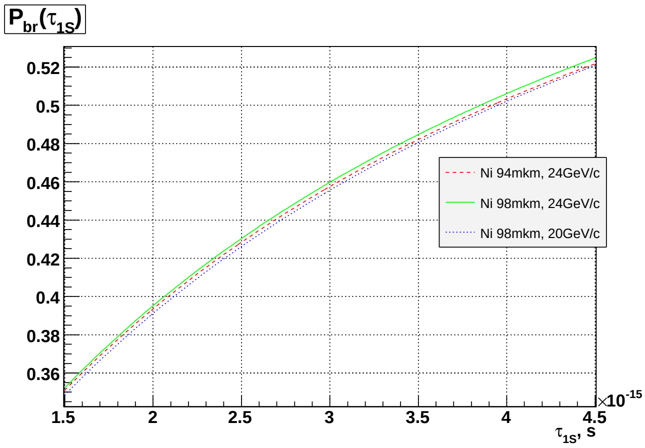

Figure 1: Functions $B_i(\tau)$ corresponding to the dependence of the break-up probability p averaged over the momentum spectra of produced pioniums for different targets and beam momenta

The relation between p_i and the lifetime τ, $p_i = B_i(\tau)$, depends on the target (material and thickness) and on the beam particle (nature and momentum) (see Fig. 1). This relation is by nature a one-to-one function (this is important for the following calculation). Its derivative is positive, $dB/d\tau > 0$, because evidently increasing the lifetime would decrease the decay probability hence increase the breakup probability (which is the competing evolution). Its calculation is described in [2].

In the following, we shall represent sets by bold characters:

$$\boldsymbol{m} = [m_1, m_2, \ldots], \quad \boldsymbol{\sigma} = [\sigma_1, \sigma_2, \ldots], \quad \boldsymbol{p} = [p_1, p_2, \ldots]. \tag{3}$$

The likelihood of this set of measurements is the product of the individual functions f_i, an overall function of their common parameter τ. After replacing p_i by the breakup probability functions $B_i(\tau)$, and changing the name f to g, it reads:

$$g_i(m_i, \sigma_i, \tau) = f_i(m_i|p_i, \sigma_i), \tag{4}$$

$$L(\mathbf{m}, \boldsymbol{\sigma}, \tau) = \prod_i g_i(m_i, \sigma_i, \tau). \tag{5}$$

The classical estimate is $\hat{\tau}$, that value of τ that maximizes the likelihood. Equivalently, with evidently an identical result, one may minimize the log-likelihood (i. e. least-squares formulation):

$$\ln L(\mathbf{m}, \boldsymbol{\sigma}, \tau) = \sum_i \ln g_i(m_i, \sigma_i, \tau). \tag{6}$$

Please, notice that we started with the pdf f_i which have a probabilistic meaning: $f_i(m_i|p_i, \sigma_i)\, dm_i$ is the probability that m_i lies in a range dm_i about p_i. There is nothing of that sort in the likelihood $L(\mathbf{m}, \boldsymbol{\sigma}, \tau)$, although the same functions are used. There is no probabilistic statement about p_i around m_i.

Justification and properties of the Maximum Likelihood method may be found in [3]. The invariance property, important for our problem, is recalled here:

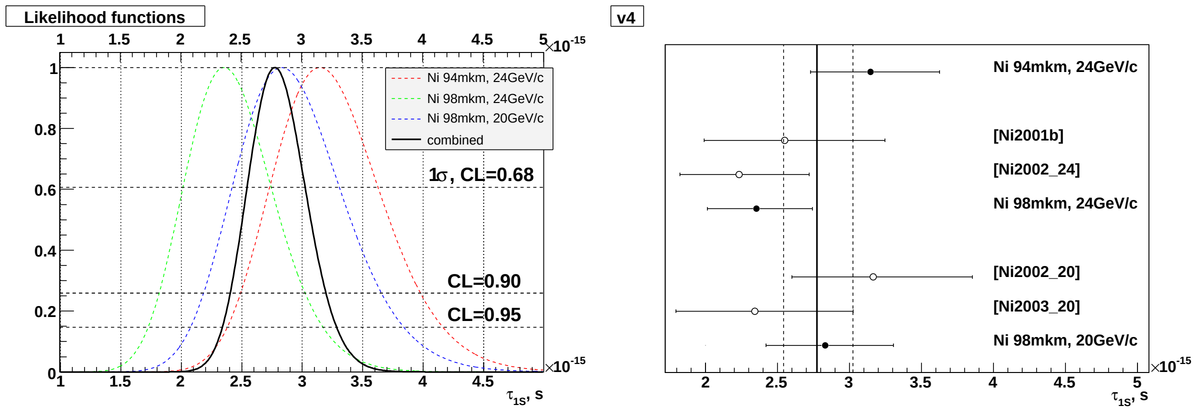

Figure 2: Likelihoods: individuals $g_i(m_i, \sigma_i, \tau)$ and the combined $L(\mathbf{m}, \boldsymbol{\sigma}, \tau)$ (left). Most probable value $\hat{\tau}$ with its statistical uncertainties for each period (right).

> The maximization of $L(x)$, with respect to x, and of $L(x(y))$, with respect to y, will give correlated results x_0 and y_0, $x_0 = x(y_0)$, a relation independent of the likelihood function. Indeed, if x_0 is such that it is solution of $dL/dx|_{x=x_0} = 0$ then y_0 such that $x_0 = x(y_0)$ is solution of $dL/dy|_{y=y_0} = dL/dx|_{x=x_0} \cdot dx/dy = 0$ because there $dL/dx|_{x=x_0} = 0$.

This is called *invariance* of the maximum likelihood estimation (see Eq. 8.11 in [3]). All this assumes $x(y)$ to be a one-to-one function, which is our case (more general cases may however be treated).

Fig. 2 shows a representation of the overall likelihood and the individual ones, restricted each to one measurement (i. e. functions g_i) versus the lifetime τ. For commodity of the graphical representations, all functions are rescaled so that their maxima are set to a common value, 1. This does not affect the τ-positions of the maxima because the scaling factors are independent of τ. The individual likelihoods are now simply the exponent terms of the functions g_i.

In the case of a Normal law, the confidence level corresponding to the range $(\mu - n\sigma, \mu + n\sigma)$ is:

$$\mathrm{CL} = \int_{\mu-n\sigma}^{\mu+n\sigma} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx. \tag{7}$$

The value of the likelihood $L(x, \mu, \sigma)$ at $x = \mu \pm n\sigma$ (away "n sigmas" from its maximum) is

$$L = L_{\max} \exp\left(-n^2/2\right). \tag{8}$$

For an arbitrary likelihood function $L_x(x)$, one may always make a change of variable on x, $y = g(x)$, such that $L_y(y)$ is a Normal law $N(y|\mu_y, \sigma_y)$. Thus the same calculation as above holds, still based on the invariance property. In particular:

$$\mu_y = g(\mu_x), \quad L_x(\mu_x - x_{\mathrm{low}}) = L_y(\mu_y - n\sigma_y) \quad \text{and} \quad L_x(\mu_x + x_{\mathrm{sup}}) = L_y(\mu_y + n\sigma_y). \tag{9}$$

The range $(\mu_y - n\sigma_y, \mu_y + n\sigma_y)$ transforms back in an x range, $(\mu_x - x_{\mathrm{low}}, \mu_x + x_{\mathrm{sup}})$, BUT the latter is NOT any more symmetrical about μ_x. The τ-ranges corresponding to different confidence levels are indicated in the figures and in Tab. 2.

The estimation of the lifetime reads

$$\hat{\tau} = \left(2.7738 \, {}^{+0.2497}_{-0.2302}\Big|_{\text{stat}}\right) \cdot 10^{-15} \text{ s.} \tag{10}$$

The systematical errors have not yet been considered. They are here common to all measurements, thus should be considered as a correlated effect while, until now, the measurements were considered independent from each other. The likelihood of the set of measurements cannot any more be the product of individual likelihoods. One now considers a multidimensional Normal distribution. Let $\sigma_{\text{stat}\,i}$ represent the statistical uncertainty on m_i and σ_{sys} the systematical error common to all measurements. Let U be a column matrix (and U^T its transposed, a line matrix) defined by $U_i = m_i - p_i$.

We use the standard notations:

$$E[x] \qquad \text{expectation value of } x, \tag{11}$$

$$V[x] = E[(x - E[x])^2] \qquad \text{variance of } x, \tag{12}$$

$$\text{cov}[x, y] = E[(x - E[x]) \cdot (y - E[y])] \qquad \text{covariance of } x \text{ and } y. \tag{13}$$

By construction $E[U_i] = 0$. Let G be the error matrix on U:

$$G_{ij} = \text{cov}(U_i, U_j) = E[U_i U_j], \tag{14}$$

and H be the inverse of G. Then one defines the scalar

$$M = U^T H U \tag{15}$$

and the likelihood

$$L = \exp(-M/2) \tag{16}$$

This expression is the same as defined earlier when G is diagonal, i. e. $G_{ij} = \delta_{ij}\, {\sigma_{\text{stat}}}_i^2$. We merge in quadrature statistical and systematical uncertainties, so that the variance V of U_i is

$$G_{ii} = V[U_i] = E\left[[U_i - E[U_i]]^2\right] = E\left[U_i^2\right] = {\sigma_{\text{stat}}}_i^2 + \sigma_{\text{sys}}^2. \tag{17}$$

Using the fact that the systematical error on $U_i - U_j$ is 0, the off-diagonal terms, $G_{ij} = E[U_i U_j]$, can be estimated from:

$$\begin{aligned} V[U_i - U_j] &= {\sigma_{\text{stat}}}_i^2 + {\sigma_{\text{stat}}}_j^2 = V[U_i] + V[U_j] - 2\,\text{cov}(U_i, U_j) \\ &= {\sigma_{\text{stat}}}_i^2 + \sigma_{\text{sys}}^2 + {\sigma_{\text{stat}}}_j^2 + \sigma_{\text{sys}}^2 - 2E[U_i U_j] = {\sigma_{\text{stat}}}_i^2 + {\sigma_{\text{stat}}}_j^2 + 2\sigma_{\text{sys}}^2 - 2E[U_i U_j], \end{aligned} \tag{18}$$

hence

$$G_{ij} = E[U_i U_j] = \sigma_{\text{sys}}^2 + \delta_{ij} {\sigma_{\text{stat}}}_i^2. \tag{19}$$

G is thus the sum of the diagonal matrix corresponding to the statistical uncertainty and a matrix whose all terms are equal to σ_{sys}^2.

The estimation of the uncertainty (confidence level for n=1) is done in the following way. The calculations are done using successively statistical and total uncertainty, the latter being the combination in quadrature of statistics and systematics. This procedure is repeated using negative and positive components of the uncertainties on measurements

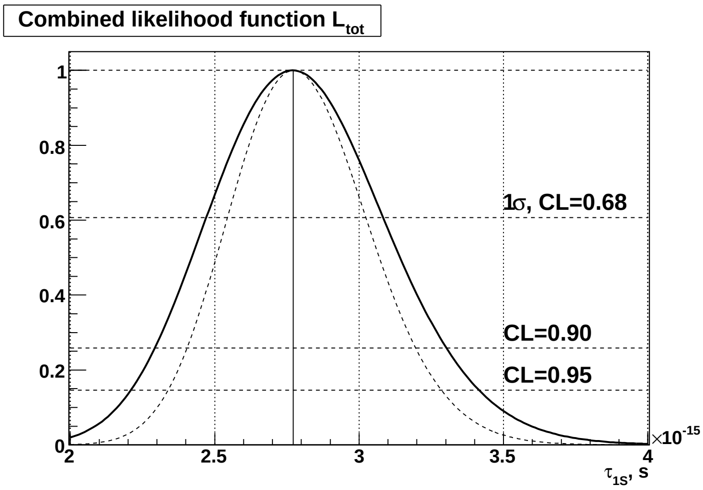

Figure 3: Likelihood function $L_{\rm tot}(\mathbf{m}, \boldsymbol{\sigma}, \tau)$, which corresponds to the combination of all measurements including their statistical and systematic uncertainties. Statistical-only $L_{\rm stat}(\mathbf{m}, \boldsymbol{\sigma}, \tau)$ is shown by a dashed line

Table 2: Confidence intervals of the estimation $\hat{\tau}$

CL	$\underline{\tau}_{\rm stat}$, fs	$\overline{\tau}_{\rm stat}$, fs	$\underline{\tau}_{\rm tot}$, fs	$\overline{\tau}_{\rm tot}$, fs
erf $(1/\sqrt{2})$	2.5437	3.0236	2.4707	3.0834
0.90	2.4045	3.1962	2.2926	3.3021
0.95	2.3391	3.2842	2.2097	3.4147

and keeping the corresponding lower and higher limits of the range of the confidence levels respectively. There are thus 4 calculations:

$$\underline{\sigma}_{{\rm stat}\,m} \Rightarrow \underline{\sigma}_{{\rm stat}\,\tau}, \qquad \overline{\sigma}_{{\rm stat}\,m} \Rightarrow \overline{\sigma}_{{\rm stat}\,\tau}, \tag{20}$$

$$\underline{\sigma}_{{\rm tot}\,m} \Rightarrow \underline{\sigma}_{{\rm tot}\,\tau}, \qquad \overline{\sigma}_{{\rm tot}\,m} \Rightarrow \overline{\sigma}_{{\rm tot}\,\tau}. \tag{21}$$

In our calculations $\underline{\sigma}_{{\rm stat}\,m} = \overline{\sigma}_{{\rm stat}\,m}$, while $\underline{\sigma}_{{\rm stat}\,\tau}$ and $\overline{\sigma}_{{\rm stat}\,\tau}$ are different due to the non-linearity of $B_i(\tau)$. The total errors are naturally $\underline{\sigma}_{{\rm tot}\,\tau}$ and $\overline{\sigma}_{{\rm tot}\,\tau}$. The systematical errors are

$$\underline{\sigma}_{{\rm sys}_\tau} = \sqrt{\underline{\sigma}_{{\rm tot}\,\tau}^{\,2} - \underline{\sigma}_{{\rm stat}\,\tau}^{\,2}}, \tag{22}$$

$$\overline{\sigma}_{{\rm sys}_\tau} = \sqrt{\overline{\sigma}_{{\rm tot}\,\tau}^{\,2} - \overline{\sigma}_{{\rm stat}\,\tau}^{\,2}}. \tag{23}$$

For numerical calculations we use $\underline{\sigma}_{{\rm sys}_m} = 0.012$ and $\overline{\sigma}_{{\rm sys}_m} = 0.010$ according to [4], which gives

$$\underline{\sigma}_{{\rm sys}_m} = 0.012 \quad \Rightarrow \quad \hat{\tau} = 2.7715\ {}_{-0.3008}\big|_{\rm tot}\ {\rm fs} \tag{24}$$

$$\overline{\sigma}_{{\rm sys}_m} = 0.010 \quad \Rightarrow \quad \hat{\tau} = 2.7722\ {}^{+0.3111}\big|_{\rm tot}\ {\rm fs} \tag{25}$$

These calculations give values of the solution τ which are slightly different between the cases with and without systematical errors. The differences are of the order of 1% of the uncertainties, thus negligible. These differences would be strictly 0 if the derivatives with respect to τ of the breakup probability functions $B_i(\tau)$ were equal to each other at the position of the solution $\hat{\tau}$. These derivatives are indeed very close to each other. In the general case the inverse of the square of the standard deviation on τ can be estimated from the formula

$$\frac{1}{\sigma^2_{\mathrm{tot}\,\tau}} = \frac{\sum\limits_{i=1}^{n} \frac{(B'_i)^2}{\sigma_i^2} + \sigma^2_{\mathrm{sys}} \sum\limits_{i=1}^{n} \sum\limits_{j=i+1}^{n} \frac{\left(B'_i - B'_j\right)^2}{\sigma_i^2 \sigma_j^2}}{1 + \sigma^2_{\mathrm{sys}} \sum\limits_{i=1}^{n} \frac{1}{\sigma_i^2}}. \tag{26}$$

Finally the estimation of the pionium lifetime in the ground state reads

$$\hat{\tau} = \left(2.77\ {}^{+0.25}_{-0.23}\Big|_{\mathrm{stat}}\ {}^{+0.19}_{-0.19}\Big|_{\mathrm{sys}}\right) \cdot 10^{-15}\ \mathrm{s} = \left(2.77\ {}^{+0.31}_{-0.30}\Big|_{\mathrm{tot}}\right) \cdot 10^{-15}\ \mathrm{s}. \tag{27}$$

The relation between lifetime and scattering lengths is:

$$\tau = \frac{1}{\Gamma_{\mathrm{tot}}} = \frac{\mathrm{Br}_{2\pi^0}}{\Gamma_{2\pi^0}} = \frac{\mathrm{const}}{|a_0 - a_2|^2}, \tag{28}$$

where (see [5])

$$\mathrm{const} = \frac{9}{2\alpha^3 m^2_{\pi^+}} \frac{1}{\sqrt{m^2_{\pi^+} - m^2_{\pi^0} - \frac{1}{4} m^2_{\pi^+} \alpha^2}} \frac{1}{1 + \delta_\Gamma}. \tag{29}$$

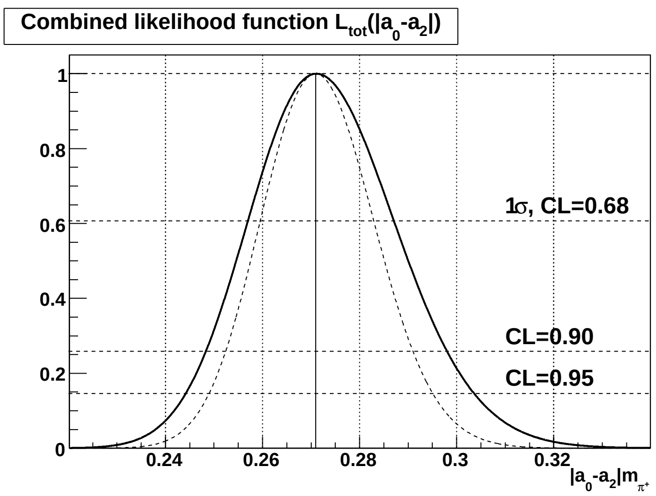

Figure 4: Likelihood function $L_{\mathrm{tot}}\left(\mathbf{m}, \boldsymbol{\sigma}, |a_0^0 - a_0^2|\right)$, which corresponds to the combination of all measurements including their statistical and systematic uncertainties. Statistical-only $L_{\mathrm{stat}}\left(\mathbf{m}, \boldsymbol{\sigma}, |a_0^0 - a_0^2|\right)$ is shown by a dashed line

Table 3: Confidence intervals of the estimation $|\widehat{a_0 - a_2}|$

CL	$\underline{\lvert a_0 - a_2 \rvert}_{\mathrm{stat}}$	$\overline{\lvert a_0 - a_2 \rvert}_{\mathrm{stat}}$	$\underline{\lvert a_0 - a_2 \rvert}_{\mathrm{tot}}$	$\overline{\lvert a_0 - a_2 \rvert}_{\mathrm{tot}}$
erf $(1/\sqrt{2})$	0.2595	0.2829	0.2570	0.2871
0.90	0.2524	0.2910	0.2483	0.2980
0.95	0.2490	0.2950	0.2442	0.3036

The invariance of the maximum likelihood method recalled above allows to obtain directly the best value of $|a_0 - a_2|$ when that of τ has been found. The results are shown in Tab. 3 and in Fig. 4. The theoretical error induced by the uncertainty of the factor $\delta_\Gamma = (5.8 \pm 1.2) \cdot 10^{-2}$ is negligible:

$$\widehat{|a_0 - a_2|} m_{\pi^+} = \left(0.2709 \; {}^{+0.0119}_{-0.0114}\Big|_{\rm stat} \; {}^{+0.0015}_{-0.0015}\Big|_{\rm theor}\right) . \tag{30}$$

Final result reads

$$\widehat{|a_0 - a_2|} m_{\pi^+} = 0.271 \; {}^{+0.012}_{-0.011}\Big|_{\rm stat} \; {}^{+0.010}_{-0.008}\Big|_{\rm sys} = 0.271 \; {}^{+0.016}_{-0.014}\Big|_{\rm tot} . \tag{31}$$

The results of the algorithm described above were tested by applying it to a sample of simulated experiments. Starting from an assumed $\tau_{\rm true}$, one calculates the breakup probabilities p_i ($i = 1, 3$ as in the experiment) resulting from the target-beam conditions of the experiment. For each simulated experiment, one generates pseudo-measurements $m_i = p_i + \sigma_{{\rm stat}\,i} + \sigma_{\rm sys}$, where $\sigma_{{\rm stat}\,i}$ and $\sigma_{\rm sys}$ are Normal random numbers of mean 0 and rms $\sigma_{{\rm stat}\,i}$ and $\sigma_{\rm sys}$, measurement uncertainties and systematic error respectively.
The parameters used are:

- Sample size: 10000 simulated experiments. 206 cases were rejected because the LifeTime fitted was outside the range $(1, 5)$ fs.
- $\tau_{\rm true} = 2.8$ fs
- $\sigma_{{\rm stat}\,i} = (0.0224, 0.0252, 0.0245)$
- $\sigma_{\rm sys} = 0.015$.

The fitted lifetime $\hat{\tau}$ has an average difference with the generated value $\tau_{\rm true}$ of 0.009 fs while its uncertainty estimated by the fit is 0.31 fs and the equivalent rms observed from simulation is 0.29 fs. The average difference between lifetimes fitted using only statistical uncertainties or the complete error matrix is smaller than 0.00001 fs.

Lifetime		Uncertainty	
Generated	Fit-average	Fit-estimation	Fit-rms
2.8	2.791	0.305	0.288

The confidence levels define how frequently a Lifetime-range would include the true value. For each simulated experiment and for each choice of confidence level the corresponding Lifetime-range is defined and the number of cases which include the true value are counted. The table 4 shows the good agreement between expectations and observations.

In conclusion, the statistical study confirms the correctness of the method used.

Table 4: Confidence levels of the simulated experiments

Confidence levels in %	
Calculated	Simulated
68.3	67.9
86.6	86.8
90.0	90.7
95.0	96.2
95.5	96.7
98.8	99.2
99.0	99.3
99.7	99.8
99.9	99.9

References

[1] V. Yazkov, private communication.

[2] M. Zhabitsky, *Direct calculation of the probability of pionium ionization in the target*, Phys. At. Nucl. **71**, No. 6, 1040 (2008); arXiv:0710.4416 [hep-ph].

[3] F. E. James, **Statistical Methods in Experimental Physics**, World Scientific, New Jersey, London, 2006.

[4] V. Yazkov. *Investigation of systematic errors for analysis with DC and ScFi.* DIRAC NOTE 08-04.

[5] J. Gasser, V. E. Lyubovitskij, A. Rusetsky and A. Gall, Phys. Rev. D **64**, 016008 (2001).