FERMILAB-TM-2425-E

DØNote 5229

Precise tuning of the b fragmentation for the DØ Monte Carlo

Y. Peters, K. Hamacher & D. Wicke

(Dated: August 29, 2006)

PYTHIA as used in DØ simulations has been tuned to e^+e^- data on b fragmentation. The tuning procedure is discussed and an optimal parameter sets for the LEP experiments and for SLD are given. We propose to use the LEP result for the reweighting of the DØ Monte Carlo and the SLD result as a valid alternative to determine systematic uncertainties due to the b fragmentation tuning.

Contents

I. INTRODUCTION

In a recent note [1] a first tuning of the b fragmentation within the DØ simulation (Pythia) has been described. In this tuning the parameter r_q of the so called Bowler fragmentation function:

$$f(z) \propto \frac{1}{z^{1+r_q b m_Q^2}} z^a \left(\frac{1-z}{z}\right)^a \exp\left(-\frac{-b(m_Q^2+p_\perp^2)}{z}\right) \tag{1}$$

was adjusted so to describe the average scaled momentum of current e^+e^- b fragmentation data. z is the fraction of $E+p_{\mathrm{long}}$ taken by the heavy hadron. The tuning was in fact restricted to the parameter r_q as this is the only parameter accessible in Pythia allowing a b tuning without, at the same time, modifying light quark fragmentation (touched by the parameters a and b). More involved tunings require, in principle, exhaustive cross-checks of the overall data description by Pythia.

The extended tuning described here was initially intended to provide an as good as possible tuning of the b sector for later use in the Monte Carlo production by adjusting also (beside r_q) the fragmentation parameters a and b of Pythia. It is well known that equally valid tunings can be reached for different choices of a and b. This is due to an almost perfect anti-correlation of a and b. In order to deviate least from the current light quark tuning we have therefore chosen to keep the difference $a-b$ fixed. Thus only the parameters a (taking $b = a + \mathrm{const} = a + 0.28$) and r_q were freely adjusted. It has been verified that this choice leaves enough freedom to allow good description of b fragmentation.

Meanwhile the reweighting of the (light cone) energy fraction z of the initial B hadron has been made possible in the DØ analysis. Therefore the above restriction to a and r_q could have been released when reweighting the Monte Carlo. Possible residual side effects of the reweighting of the light quark fragmentation or the b description are minimal as B hadron creation happens first in the fragmentation process in the model. A possible backslash due to conservation rules is thereby damped and tiny.

II. DESCRIPTION OF THE TUNING AND ITS RESULTS

The model has been tuned by adjusting the parameter to the measured $x_B = 2 \cdot E_B/E_{CM}$ distributions of Sld [2] and the Lep experiments [3–5]. A practical complication arising is that in [2–4] the bin-width of the B hadron energy fraction x_B has been chosen smaller (~ 0.05) than the reachable experimental resolution on $\Delta x_B \gtrsim 0.1$. This limited resolution is already due to physical effects, the identification of the b jet energy with the b quark energy and the imperfect reconstruction of the B hadrons due to losses of neutral particles etc. Due to this choice of the binning and the unfolding process performed in the analyses different bins show extremely strong correlations. Moreover the covariance matrices given with the results do not allow an inversion. Therefore these results cannot be used in a proper model-to-data χ^2 comparison. The result [5] in principle allows inversion, however this result is not finally published and only available through the accompanying theses. We have therefore chosen to use all available data and later infer the uncertainty of the tuning parameters from the comparison of the individual results. Bin correlations were neglected in the χ^2 calculation. To still allow inter-comparability the $\min(\chi^2/N_{df})$ obtained for each experimental result was rescaled to one.

The tuning was then performed by producing Pythia events for a grid of parameter choices ($0.2 \leq a \leq 1.8$, step size 0.05 and $0.6 \leq r_q \leq 1$, step size 0.05). For each parameter setting 50k events were generated with DØ flavour Pythia in e^+e^- mode (the complete parameter setting is given in the appendix). The distribution of the B hadron energies were calculated in the binning proposed by the individual experiments. In correspondence with the inclusive analyses for DELPHI, OPAL and SLD all observed B hadrons were used, for ALEPH only weakly decaying B mesons (corresponding to the more exclusive analysis) were used. Then, for each parameter choice and data set the resulting χ^2 was calculated. For the above explained reason these χ^2 values were later renormalised such that for the optimal description of the model $\min(\chi^2/N_{df}) = 1$. Here in accordance with the expected optimal resolution $\Delta x_B \gtrsim 0.1$ the number of equivalent data points N_{data} was set to 10.

In the vicinity of the optimal parameter choice the so determined χ^2 is expected to show a parabolic dependence on the parameters. Due to the complicated dependence of the fragmentation function on a, b and r_q this expectation is only fulfilled very close to the optimum. In order to interpolate between the chosen parameter points an analytic description of χ^2 depending on a and r_q was fit to the following ansatz:

$$f_{\chi^2} = A + B \cdot r_q + C \cdot r_q^2 + D \cdot r_q^3 \qquad \text{where}$$

$$A = p_1 + p_2 \cdot y + p_3 \cdot y \qquad y = a - 0.9$$

$$B = p_4 + p_5 \cdot y \qquad C = p_6 + p_7 \cdot y \qquad D = p_8 + p_9 \cdot y$$

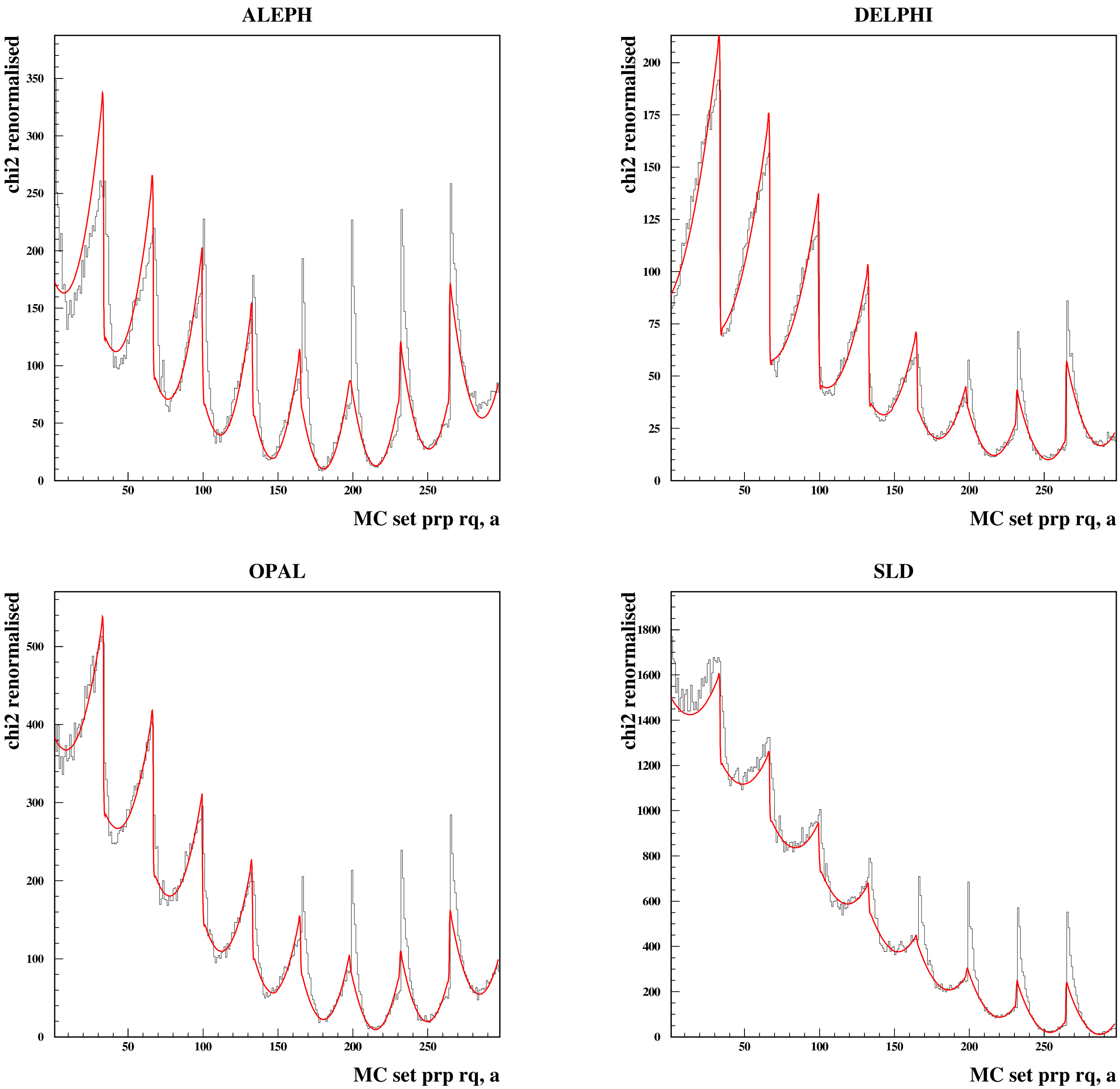

FIG. 1: Renormalised χ^2 of the Monte Carlo to data comparison for the LEP experiments and SLD. The sawtooth structure is due to the 1-dimensional representation of the 2-dimensional (a vs. r_q) problem. Each parabola refers to the variation of $a(0.2, \dots, 1.8)$ for different fixed $r_q(0.6, 0.65, \dots, 1.)$. The lines represent the fit described in the text.

The higher order "parameter" terms in the expression f_{χ^2} are sufficiently described by straight lines, the constant term due to the general behaviour of χ^2 close to the minimum by a parabola. The above ansatz reasonably describes χ^2 in the vicinity of the optimal parameters (see Fig. 1). In order not to influence the optimum by large χ^2 values and the then inadequate mathematical description χ^2 values bigger than 20 (40 for SLD) times the minimal value were excluded from the fit.

The optimal parameter choices for the different experiments can be read off from the fitted function shown in Fig. 2 and are given in Tab. 1 (central values are given by the diamonds, the errors by the 95% C.L. ellipses. The inner (red) error ellipses in Fig. 2 correspond to 65% and 95% C.L. definitions. Note, however, that due to the renormalisation

Experiment	r_q	a
ALEPH	0.86 ± 0.02	0.97 ± 0.13
DELPHI	0.94 ± 0.03	1.21 ± 0.24
OPAL	0.90 ± 0.02	1.03 ± 0.15
A+O+D	0.897 ± 0.013	1.03 ± 0.08
SLD	0.98 ± 0.01	1.30 ± 0.09

TABLE 1: Fit parameters deduced from the different experimental results.

of χ^2 (induced by the incompletely known experimental errors) the exact C.L. is unknown. The parameters deduced from the ALEPH and OPAL data, and to reduced extent also the DELPHI data agree well. Due to the χ^2 rescaling and as the DELPHI data in general is less well described by the model the uncertainties for DELPHI are slightly increased. As the data are consistent the χ^2 of the Lep results added. This sum is shown in Fig. 2. The SLD data suggests slightly different parameter values.

The data are compared to the Monte Carlo prediction in detail in Fig. 3. The agreement of the LEP data fit and the individual data sets is very good for the ALEPH and OPAL results and slightly less satisfactory for DELPHI (Note that some deviations for very small and high x are partly due to missing (zero) data points). Although the fit parameters disagree for the SLD fit the actual model predictions are also quite similar in this case. The fit parameters resulting from our tuning to the average x [1] results in a wider peak of the fragmentation function.

On first view the r_q parameters from this tuning and from [1] appear completely inconsistent. This is not so, however, as is demonstrated in Fig. 4 showing $\langle x_B \rangle$ as function of r_q and a.

The Monte Carlo reweighting factors obtained from the fragmentation function used at Monte Carlo generation stage and resulting from the optimal tuning (see Fig. 5) is moderate. Therefore the loss of statistical precision induced by the reweighting is tolerable.

III. CONCLUSION

The available e^+e^- data on the b fragmentation function has been used to tune the DØ Monte Carlo. The parameter sets deduced from the LEP experiment are consistent. Although the parameter set obtained from the SLD data differs the actual fragmentation functions predicted by Pythia agree reasonably. Because of the consistency of the LEP results we propose to use the corresponding parameters as central the setting for the reweighting of the DØ Monte Carlo. The alternate choice of the SLD parameters should serve to determine the systematic uncertainty due to the tuning of b fragmentation in the DØ analyses.

Appendix: Parameter settings used in the tuning

Pythia v6.323 in p17 was used with default parameter settings. For the Bowler FF the initially relevant parameters were $a = 0.3$, $b = 0.58$ and $r_q = 1.0$. The D0RunII version was p17.09.01 . Pythia was run with the packages d0_mcpp_gen, d0_mcpp_gen p17-br-09 mcpp_gen p17-br-15, isajet v7-72-00, mcpp p17-br-13, and mc_analyze. Within the described framework the collision was changed from $p\bar{p}$ to e^+e^- and Pythia was recompiled. The most relevant parameters used are listed in Tab. 2. All remaining parameters were left at the default values. Initial state photon radiation was switched off.

parameter	used value	description
MSTP(81)	1	turn on multiple interaction
MSTJ(11)	3,4 or 5	choice of fragmentation function
PMAS(4,1)	1.55	mass of c-quark
PMAS(5,1)	4.75	mass of b-quark
MSTJ(22)	2	cut-off on decay length
PARJ(46), PARJ(47)	r_q	Bowler FF parameter for c- and b-quark, respectively
PARJ(55)	$-\epsilon_b$	Peterson FF parameter
PARJ(41)	a	Lund FF parameter
PARJ(42)	b	Lund FF parameter

TABLE 2: parameter and switches for Pythia 6.2 in $e^+e^- \to Z \to b\bar{b}$; for details see [6].

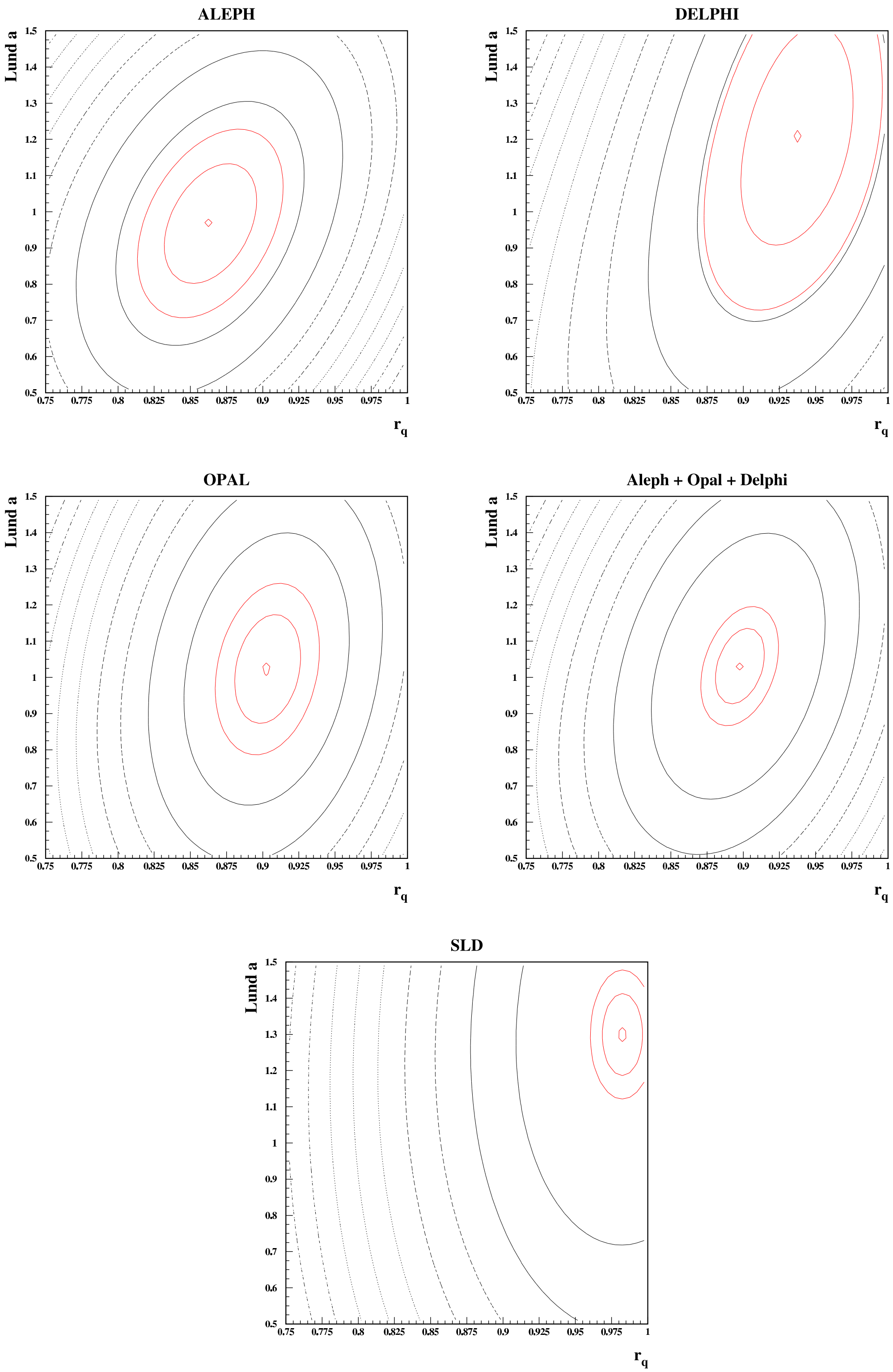

FIG. 2: Fitted χ^2 for the LEP experiments and SLD as function of r_q and a. ALEPH + OPAL + DELPHI represents the sum of χ^2 of the LEP experiments. The red diamond represents the optimum χ^2, the inner two (red) ellipses are the 65% and 95% C.L. ellipses. The black lines are arbitrary contours.

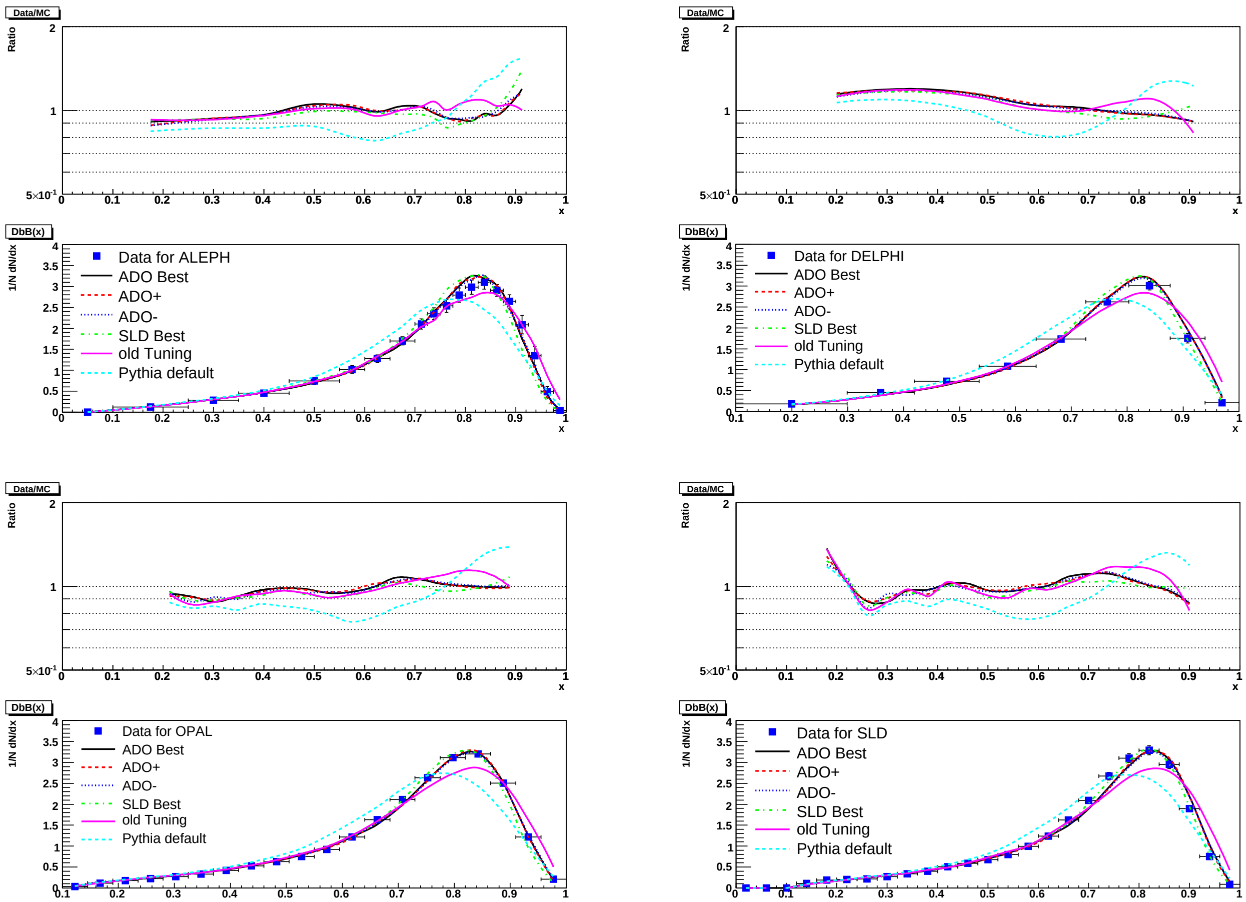

FIG. 3: Comparison (lower plots) of the data with different predictions of PYTHIA. Corresponding data to Monte Carlo ratio (upper plots). Black line: LEP tuning, red line: LEP plus errors, green line: LEP minus errors, blue line SLD. Purple: tuning described in [1].

[1] Y. Peters, K. Hamacher and D. Wicke, DØ note 4968.
[2] K. Abe *et al.* [SLD Collaboration], Phys. Rev. Lett. **84**, 4300 (2000) [arXiv:hep-ex/9912058],
[3] A. Heister *et al.* [ALEPH Collaboration], Phys. Lett. B **512**, 30 (2001) [arXiv:hep-ex/0106051],
[4] G. Abbiendi *et al.* [OPAL Collaboration], Eur. Phys. J. C **29**, 463 (2003) [arXiv:hep-ex/0210031],
[5] G. Barker *et al.* [DELPHI Collaboration], 2002-69 Conf 603. E. Ben-haim, "The b quark fragmentation function, from LEP to TeVatron," thesis, Univ. Paris 6, 2004, FERMILAB-THESIS-2004-50
[6] T. Sjöstrand, L. Lönnblad, S. Mrenna and P. Skands, "PYTHIA 6.3: Physics and manual," arXiv:hep-ph/0308153.

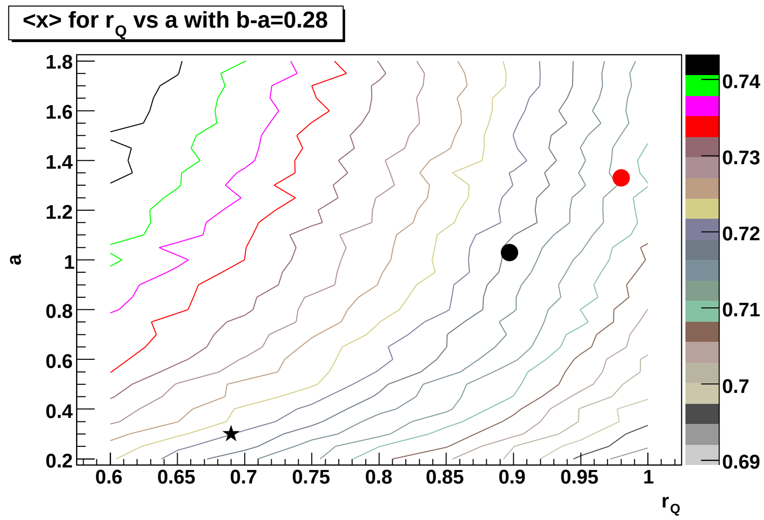

FIG. 4: $\langle x_B \rangle$ as function of r_q and a. The star denotes the result from [1], the black bullet is the result from LEP data and the read bullet from SLD data.

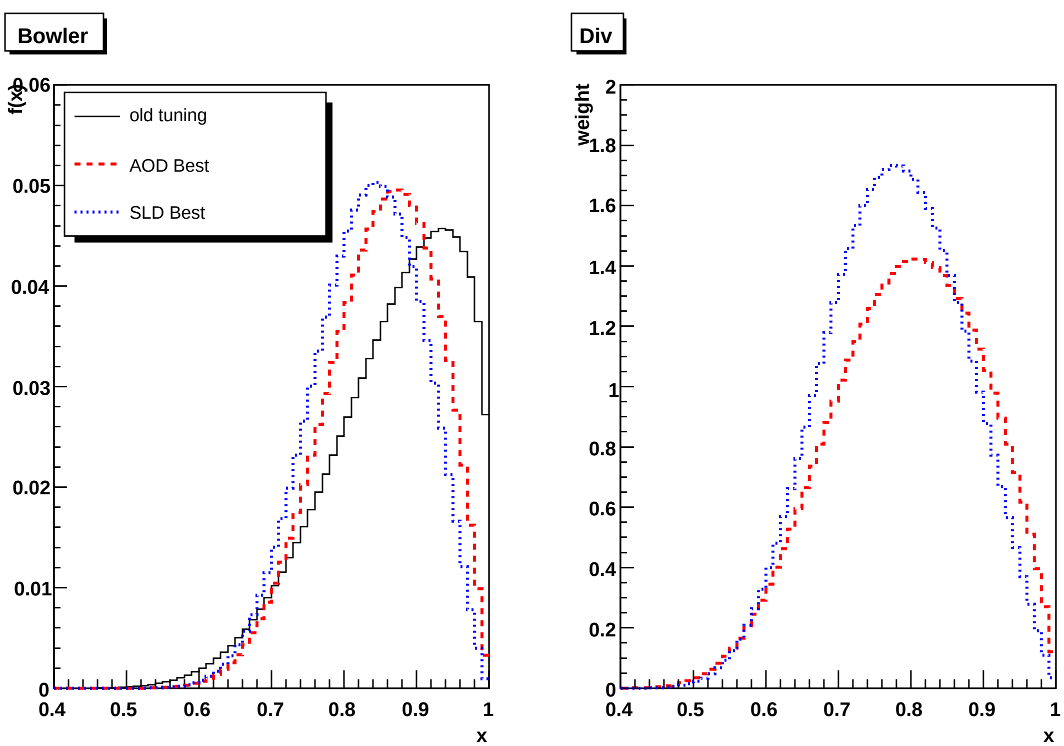

FIG. 5: Bowler fragmentation functions as determined by the different tunings (left) and weighting factors for reweighting the events generated as described in [1] to the LEP average and SLD tunings.