An average b-quark fragmentation function from $e^+e^-$ experiments at the Z pole and recommendations for its use in other experimental environments

Eli Ben-Haïm
LPNHE-Paris
**Introduction**

*b-Fragmentation* = process by which b-quarks organize themselves into hadrons

(strong interaction, $\Delta t \sim 10^{-24}$ sec)
Introduction

| \( \sqrt{s} \) | \( m_b \) | \( \Lambda_{QCD} \) |

**Energy scale**

- **Hard process**
  - \( e^+ \gamma \)
  - \( Z^0 \)
- **Heavy quark (gluon radiation)**
- **Primary heavy hadrons**
  - \( \gamma, \pi, K \)
- **Weak decay**

**Observable particles**

**Other production mechanisms in hadron colliders** (e.g. \( q\bar{q} \rightarrow b\bar{b} \))
Introduction

Perturbative (calculable in QCD)

The boundary depends on the perturbative QCD computation (order and technique of resummation, Monte Carlo generator...)

Non-perturbative (non-calculable, usually described by a hadronisation model)

Large collinear logs

$$\alpha_s^n \log^{n-k} \left( \frac{S}{m^2} \right)$$

Large soft logs

$$\alpha_s^n \left( \frac{\log^{2n-1-k} \Delta}{\Delta} \right)$$

Perturbative (calculable in QCD) Non-perturbative (non-calculable, usually described by a hadronisation model)
**Introduction**

**Fragmentation function:** probability density function of a variable relating quark - hadron kinematics \( (x_B^{\text{weak}}, x_p^{\text{weak}}, z) \)

- **Directly measurable by experiments**
  - \( x_p^{\text{weak}} = \frac{p_B^{\text{weak}}}{p_{B,\text{max}}} = \frac{\sqrt{x_B^{\text{weak}}^2 - x_{\min}^2}}{\sqrt{1 - x_{\min}^2}} \)
  - \( x_B^{\text{weak}} = \frac{E_b^{\text{weak}}}{E_{\text{beam}}} \)

- **Defined within event generators**
  - \( z = \frac{(E + p_{\parallel})_B}{(E + p)_b} \)

Fold together two components: perturbative \( \otimes \) non-perturbative
Intermediate conclusions

- Ambiguous boundary between the perturbative and non-perturbative regimes
  ⇒ Perturbative and non-perturbative descriptions must be compatible
  ⇒ Only the measured fragmentation functions are non-ambiguous!

- Physics in high energy scales (production, perturbative regime) depend on the experimental environment whereas physics in low energy scale (non-perturbative regime, hadronisation) is expected to be independent of the experimental environment

- $e^+e^-$ colliders allowed to measure accurately the $b$-fragmentation function
Overview

- **New** DELPHI measurement of the fragmentation function
- Average distribution obtained at the Z pole (LEP+SLD)
- Hadronic-model independent extraction of the non-perturbative component
- Fits to hadronisation models: PYTHIA parameters to use in the TeVatron and the LHC
- Conclusions

arXiv:1102.4748 [hep-ex]
(DOI 10.1040/epjc/s 10052-011-1557-x)
DELPHI measurement of the fragmentation func.

Combination of results of the $x_B^{\text{weak}}$ distribution from two independent analyses, using different approaches.

**Regularised unfolding**
- Unfolds from the observed distribution in data the underlying fragmentation function. The two differ from each other due to effects of resolution, acceptance and backgrounds.

$$g(x_B^{\text{weak}}, x_B^{\text{rec}}) = \int R(x_B^{\text{weak}}, x_B^{\text{rec}}) f(x_B^{\text{weak}}) dx_B^{\text{weak}} + b(x_B^{\text{weak}}, x_B^{\text{rec}})$$
- Uses Neural Networks to reconstruct the b-hadron energy.

**Weighted fitting**
- b-hadron energy: assignment of tracks to the primary and secondary vertices.
  $$E_B = E_{\text{jet}} - E_{\text{primary vertex}}$$
- Apply weights to the z distribution in simulation.
- Fit weights to obtain best agreement between $p_B$ in data and simulation.
The results are independent of any initial assumption regarding the actual shape of the underlying fragmentation distribution in simulation.

Average value of the DELPHI distribution:

$$\left\langle x_B^{\text{weak}} \right\rangle = 0.699 \pm 0.011$$
Averaged distribution obtained at the Z pole

Each of the 4 measurements of the fragmentation distribution used a different choice of binning and has a different number of effective degrees of freedom.
To obtain a combined distribution, a global fit was done using the parameterization:

\[ p_0 \left( p_1 x^{p_2} (1 - x)^{p_3} + (1 - p_1) x^{p_4} (1 - x)^{p_5} \right), \]

and cutting away non-significant degrees of freedom of the individual error matrices. The fit result:

<table>
<thead>
<tr>
<th>( x_B^{\text{weak}} )</th>
<th>( x_P^{\text{weak}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_1 )</td>
<td>( 12.97^{+0.77}_{-0.71} )</td>
</tr>
<tr>
<td>( p_2 )</td>
<td>( 2.67^{+0.15}_{-0.14} )</td>
</tr>
<tr>
<td>( p_3 )</td>
<td>( 2.29^{+0.19}_{-0.17} )</td>
</tr>
<tr>
<td>( p_4 )</td>
<td>( 1.45^{+0.28}_{-0.22} )</td>
</tr>
<tr>
<td>( p_5 )</td>
<td>( 0.663^{+0.035}_{-0.036} )</td>
</tr>
</tbody>
</table>

Uncertainties for \( x_B^{\text{weak}} \) (\( x_P^{\text{weak}} \)) are rescaled by 1.24 (1.37) to account for the dispersion of measurements, mainly between ALEPH and SLD.

Beside providing a good fit to data, this distribution has an analytical Mellin transformation (see following...)

Average value of the combined distribution:

\[ \langle x_B^{\text{weak}} \rangle = 0.7092 \pm 0.0025 \]
Model-independent extraction of the non-perturbative QCD component


Folding:
\[ f_{\text{predicted}}(x) = \frac{1}{x} \int f_{\text{pert.}}(z) \times f_{\text{non-pert.}} \left( \frac{x}{z} \right) \frac{dz}{z} \]

Measured perturbative (QCD, Monte Carlo)

\[ \tilde{D}(N) = \int_{0}^{1} dx \ x^{N-1} \ D(x) \]

Moments

Inverse Mellin Transformation

\[ D_{\text{non-pert.}}(x) = \frac{1}{2\pi i} \oint dN \ \frac{\tilde{D}_{\text{measured}}(N)}{\tilde{D}_{\text{pert.}}(N)} \ x^{-N} \]

Non-perturbative

x distributions

directly extracted, hadronisation-model independent

Non-perturbative

NO MODEL

Measured non-perturbative

Non-perturbative

Measured perturbative
Perturbative and non-perturbative components

Extracted non-perturbative component with NLL perturbative QCD


Folding the NLL perturbative QCD component with hadronisation models

NLL break-down ⇔ non-physical behaviour (x>1)

Folding a “non-physical” perturbative component with a “physical” hadronisation model (usually designed for use within event generators) does not reproduce the data!
Strong dependence of the extracted non-perturbative component on the perturbative framework!

- Higher order of QCD ⇒ peak displaced to higher $x$
- Dependence on the generator!

Perturbative and non-perturbative components

Perturbative QCD = PYTHIA 6.156 and JETSET 7.3
The DELPHI measurement was compared with expectations from different non-perturbative hadronisation models within PYTHIA 6.156. Only the Lund and Lund-Bowler models give reasonable descriptions of the data (Lund favoured). Similar results are observed with the world average distribution.

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameters</th>
<th>$\chi^2$/NDF</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peterson</td>
<td>$e_b = (4.06^{+0.46}_{-0.41}) \times 10^{-3}$</td>
<td>55.8/6</td>
<td>—</td>
</tr>
<tr>
<td>Lund</td>
<td>$a = 1.84^{+0.23}<em>{-0.21}$, $b = 0.642^{+0.073}</em>{-0.063}$ GeV$^{-2}$</td>
<td>9.8/5</td>
<td>92.2%</td>
</tr>
<tr>
<td>Lund-Bowler</td>
<td>$a = 1.04^{+0.14}<em>{-0.12}$, $b = 3.08^{+0.45}</em>{-0.39}$ GeV$^{-2}$</td>
<td>20.7/5</td>
<td>85.6%</td>
</tr>
</tbody>
</table>
To better understand this: recall the shape of the directly extracted non-perturbative component and compare with fitted hadronisation models.

Models with low-x tail are inadequate…
A global fit of the Lund and Lund-Bowler models parameters using measurements from ALEPH, DELPHI, OPAL and SLD, within PYTHIA 6.156. (the $\chi^2$ minimised in this study was the sum of $\chi^2$ corresponding to the 4 results)

The fit clearly favours the Lund model over the Lund-Bowler one.

Lund parameters:

\begin{align*}
  a &= 1.48^{+0.11}_{-0.10} \\
  b &= 0.509^{+0.024}_{-0.023} \text{ GeV}^{-2}
\end{align*}

In PYTHIA it is impossible to separate the Lund parameters for different flavors

$\Rightarrow$ Reweight the $z$ distribution…
A comment

comparison between the directly extracted non-perturbative component and the results of fits within a Monte Carlo simulation

- Extraction and fits done within the same version of PYTHIA (6.156)
- The results of the two methods have been found to be similar in all tested cases.
- The direct extraction method can (at least...) help finding the way to the right parameters to use.
Comparison between $p_T$ spectra of the $B^+$ in generated 7 TeV pp collisions with two b-hadronisation models:

- Our average (Lund)
- Peterson with $\varepsilon = 0.003$

How does this affect LHC measurements?

$\frac{dN(\text{ave. Lund})}{dN(\text{Peterson})} \approx 15\%$ effect in $25 < p_T < 30$ GeV!

Larger effect expected in higher $p_T$
Summary and conclusions

We presented:

- *New* DELPHI measurement of the fragmentation function
- Average distribution obtained at the Z pole
- Study of the non-perturbative QCD component and hadronisation models


A few points to underline:

- Pay attention to the joint use of perturbative and non-perturbative components
- Some hadronisation models fail to describe e⁺e⁻ data (and are not expected to work in other environments)

---

Result for the world average Lund parameters to use in PYTHIA 6.156:

\[
a = 1.48^{+0.11}_{-0.10} \quad ; \quad b = 0.509^{+0.024}_{-0.023} \text{ GeV}^{-2}
\]

This result is expected to be valid in experimental environments other than LEP. It would be fruitful to check how it fits data in the LHC and the TeVatron.

- If you use another version of PYTHIA: extract the adequate parameters (all the information is given in the paper!) or try to contact us…
Backup
Variables

\[ z = \frac{(E + p_{||})_B}{(E + p)_b}, \quad (0 < z < 1) \]

\[ x_B^{\text{weak}} = \frac{E_B^{\text{weak}}}{E_{\text{beam}}} \quad \text{and} \quad \left( x_\text{min} < x_B^{\text{weak}} < 1 \; ; \; x_\text{min} = \frac{2m_B}{\sqrt{s}} \right) \]

\[ x_p^{\text{weak}} = \frac{p_B^{\text{weak}}}{p_{B,\text{max.}}} = \sqrt{x_B^{\text{weak}} - x_\text{min}^2} \quad \frac{1 - x_\text{min}^2}{\sqrt{1 - x_\text{min}^2}} \quad , \quad (0 < x_p^{\text{weak}} < 1) \]
## Delphi analysis – systematics (I)

### Table 2: Systematic uncertainty on the mean value of the unfolded $x_B^{\text{weak}}$ distribution. The total is the sum in quadrature of all contributions. The sign indicates the correlation between the change in an uncertainty source and the shift in the final result. Uncertainties assigned by turning a weight on/off have no sign.

<table>
<thead>
<tr>
<th>uncertainty class</th>
<th>item</th>
<th>$\delta (x_B^{\text{weak}})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>technical</td>
<td>number of degrees of freedom</td>
<td>+0.0025</td>
</tr>
<tr>
<td>selection cuts and backg. dependence</td>
<td>$g \rightarrow b\bar{b}$</td>
<td>+0.0004</td>
</tr>
<tr>
<td>reconstructed energy</td>
<td>neutral energy smearing</td>
<td>0.0023</td>
</tr>
<tr>
<td></td>
<td>fragmentation track multiplicity</td>
<td>+0.0030</td>
</tr>
<tr>
<td>reconstructed energy</td>
<td>b-decay track multiplicity</td>
<td>−0.0004</td>
</tr>
<tr>
<td></td>
<td>neutral multiplicity</td>
<td>+0.0010</td>
</tr>
<tr>
<td></td>
<td>hemisphere scaled energy $E_{\text{hem}}/E_{\text{beam}}$</td>
<td>0.0003</td>
</tr>
<tr>
<td>b-physics modelling</td>
<td>b-hadron lifetimes</td>
<td>−0.0004</td>
</tr>
<tr>
<td></td>
<td>b-hadron production fractions</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>hemisphere quality</td>
<td>−0.0018</td>
</tr>
<tr>
<td></td>
<td>$B^{**}$ rate</td>
<td>−0.0018</td>
</tr>
<tr>
<td></td>
<td>$B^{**}$ $Q$-value dependence</td>
<td>0.0003</td>
</tr>
<tr>
<td></td>
<td>$K^0$ rate</td>
<td>+0.0005</td>
</tr>
<tr>
<td></td>
<td>$B^*$ rate</td>
<td>−0.0001</td>
</tr>
<tr>
<td></td>
<td>semileptonic decay rate</td>
<td>−0.0001</td>
</tr>
<tr>
<td></td>
<td>wrong sign charm rate</td>
<td>+0.0001</td>
</tr>
<tr>
<td></td>
<td>c- and b-quark efficiency</td>
<td>0.0001</td>
</tr>
<tr>
<td>calibration stability &amp; simulation statistics</td>
<td>calibration periods</td>
<td>0.0025</td>
</tr>
<tr>
<td></td>
<td>finite simulation statistics</td>
<td>0.0005</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>0.0060</td>
</tr>
</tbody>
</table>

### Regularized unfolding
### Table 6: Systematic uncertainty on the mean value of the $x_B^{\text{weak}}$ distribution in the weighted fitting analysis.

The total is the sum in quadrature of all contributions. The sign indicates the correlation between the change in an uncertainty source and the shift in the final result. Uncertainties assigned by turning a weight on/off have no sign.

<table>
<thead>
<tr>
<th>uncertainty class</th>
<th>item</th>
<th>$\delta \langle x_B^{\text{weak}} \rangle$</th>
</tr>
</thead>
<tbody>
<tr>
<td>technical</td>
<td>fitted function shape</td>
<td>0.0015</td>
</tr>
<tr>
<td></td>
<td>curvature parameter in $\chi^2$</td>
<td>0.0020</td>
</tr>
<tr>
<td></td>
<td>b-tagging selection cut</td>
<td>0.0010</td>
</tr>
<tr>
<td></td>
<td>non-b background level</td>
<td>0.0012</td>
</tr>
<tr>
<td>selection cuts and backg. dependence</td>
<td>jet clustering parameter value</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>ambiguous energy level</td>
<td>0.0034</td>
</tr>
<tr>
<td></td>
<td>secondary vertex multiplicity</td>
<td>0.0012</td>
</tr>
<tr>
<td></td>
<td>$g \rightarrow b\bar{b}$</td>
<td>-0.0001</td>
</tr>
<tr>
<td>reconstructed energy</td>
<td>corrections on tracks</td>
<td>0.0017</td>
</tr>
<tr>
<td></td>
<td>jet multiplicity</td>
<td>-0.0001</td>
</tr>
<tr>
<td>b-physics modelling</td>
<td>b-hadron lifetimes</td>
<td>-0.0005</td>
</tr>
<tr>
<td></td>
<td>$B^{**}$ rate</td>
<td>-0.0008</td>
</tr>
<tr>
<td></td>
<td>b-decay track multiplicity</td>
<td>-0.0008</td>
</tr>
<tr>
<td>calibration stability</td>
<td>calibration periods</td>
<td>0.0038</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>0.0064</td>
</tr>
</tbody>
</table>

**Weighted fitting**
Delphi combination: correlated systematics

- neutral energy smearing in the regularised unfolding analysis with ambiguous energy level in the weighted fitting analysis;
- $g \rightarrow b\bar{b}$ branching fraction;
- $B^{**}$ production rate;
- $b$-hadron lifetimes;
- $b$-decay track multiplicity;
- $b$-hadron production fractions;
- wrong sign charm rate.
Another comparison of the $p_T(B)$ spectrum

- Comparison between $p_T$ spectra of the $B^+$ in generated 7 TeV pp collisions with two b-hadronisation models:

- Our average (Lund) Peterson with $\varepsilon = 0.01$ (used in some LHC analyses)

- ~20% effect in $20 < p_T < 30$ GeV! Larger effect expected in higher $p_T$