Theory perspectives on b-quark fragmentation in top-quark decays

Gennaro Corcella

INFN - Laboratori Nazionali di Frascati

1. **Introduction**
2. **QCD calculations and Monte Carlo codes for b-quark fragmentation**
3. **Tuning hadronization to e^+e^- data and predictions for LHC**
4. **Prospects with NLO+shower and C++ programs**
5. **Conclusions**

Based on work by G.C., F.Mescia, V. Drollinger, A.D. Mitov, M. Cacciari, K. Melnikov, S. Biswas, M.Schulze, LEP, SLD, ATLAS and CMS top/heavy-quark working groups

Reliable description of multiple radiation in top production and decay and of b-quark fragmentation is fundamental in the measurements of the top properties

b-fragmentation contributes to the systematic error on the measurements

LHC inclusive analyses (dilepton, lepton+jets and all-hadrons) propagate the uncertainty on b-fragmentation to the systematic error due to b-jet energy scale and b-tagging efficiency

$J/\psi+$ **lepton final states** $(10^3$**/year of high luminosity)**

$t \to bW$ **;** $b \to B \to J/\psi\ X$ **;** $J/\psi \to \mu^+\mu^-$ **;** $W \to \ell\nu_\ell$

A. Kharchilava, PLB 476 (2000) 73, R. Chierici and A. Dierlamm, CMS Note 2006/058

$m_{3\ell}^{\rm max} = 0.56\ m_t - 25.3$ GeV Systematics (theo + exp): $\Delta m_t({\rm syst}) \simeq 1.47$ GeV

b-fragmentation (PYTHIA+Peterson model): $\Delta m_t({\rm frag}) \simeq 0.51$ GeV

Several calculations and tools are available for bottom fragmentation in top decays, but not unique strategy for the systematic error: comparing two tuned codes/computations (PYTHIA vs. HERWIG, MCs vs. NLO), one program varying fragmentation parameters, etc.

Issues on whether sticking to e^+e^- fits or rather using hadron data

Some relevant calculations for top decays

A.Czarnecki, PLB 252 (1990) 467: Total NLO top decay width

A.Czarnecki and K.Melnikov, PRD59 (1999) 014036: Total top decay NNLO width

G.C. and A.Mitov, NPB623 (2001) 247 b-quark energy spectrum, collinear resummation of $\ln(m_t/m_b)$ and some soft-enhanced logarithms in the NLL+NLO approximation. Hadron corrections from e^+e^- data

M. Cacciari, G.C. and A.Mitov,JHEP 0212 (2002) 015:
As above, but with complete soft NLL resummation

S.Biswas, K.Melnikov and M.Schulze, JHEP 1008 (2010) 048:
NLO distributions with collinear resummation; hadronization by the above fits

A.Denner, S.Dittmaier, S.Kallweit and S.Pozzorini, JHEP 1210 (2012) 110:
NLO for off-shell top production and decays, interface with showers and hadronization in progress

J. Gao, C.S. Li and H.X. Zhu (SCET), PRL110 (2013) 042001;
M. Brucherseifer, F. Caola and K. Melnikov, JHEP 04 (2013) 059:
NNLO distributions for top decays for massless b, not yet b-hadronization

Standard parton shower generators (PYTHIA, HERWIG): LO+LL plus some NLLs at large x $(\Lambda_{\overline{\rm MS}} \to \Lambda_{\rm MC} = \Lambda_{\overline{\rm MS}} \exp(4K\beta_0))$

B-hadron production in top decays

$$t(q) \to b(p_b)W(p_W)\,(g_1(p_{g_1}),\ldots g_n(p_{g_n})) \quad , \quad b \to B \quad , \quad x_b = \frac{1}{1-m_W^2/m_t^2}\,\frac{2p_b\cdot q}{m_t^2}$$

$$\frac{d\sigma_{\rm had}}{dx_B}(B) = \frac{d\sigma_{\rm part}}{dx_b}(b) \otimes D_{np}(b \to B)$$

Narrow width approximation ($\Gamma_t \ll m_t$):

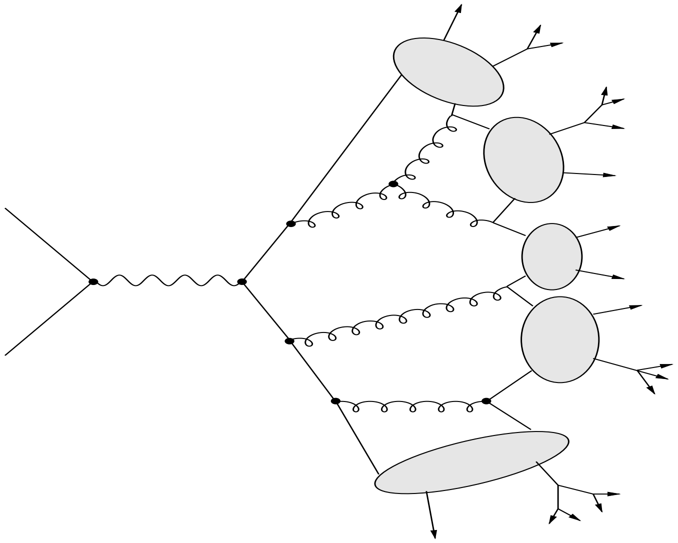
$$\frac{1}{\sigma}\frac{d\sigma_{\rm had}}{dx_B} \simeq \frac{1}{\Gamma}\frac{d\Gamma_{\rm had}}{dx_B} \quad ; \quad \frac{d\Gamma_{\rm had}}{dx_B} = \int_{x_B}^{1}\frac{dz}{z}\frac{d\Gamma_{\rm part}}{dz}(t \to b)D_{np}^B\left(\frac{x_B}{z}\right)$$

b-quark spectrum $d\Gamma/dx_b$ can be obtained by QCD calculations, e.g. carried out in Mellin moment space, or Monte Carlo generators

$D_{np}^B(x_B)$ is a non-perturbative fragmentation function with free parameters, e.g. Kartvelishvili and Peterson models:

$$D_{\rm K}(x,\alpha) = (1+\alpha)(2+\alpha)x(1-x)^\alpha \quad ; \quad D_{\rm P}(x,\epsilon) = \frac{N_P}{x\left[1-1/x-\epsilon/(1-x)\right]}$$

Hadronization models



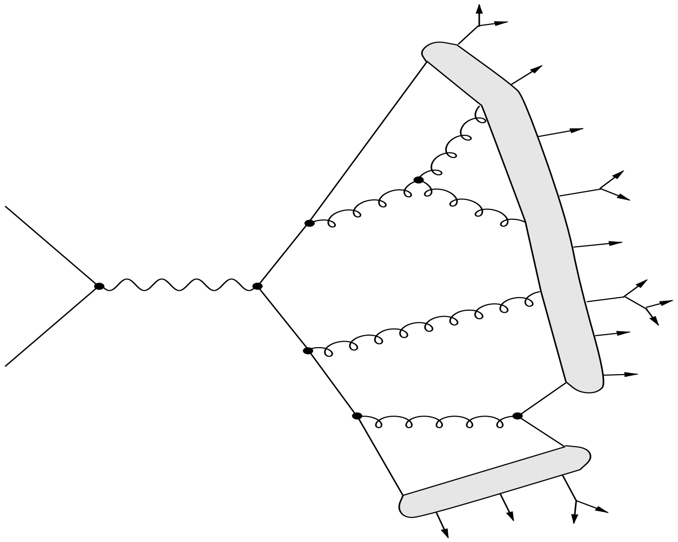
Cluster model (HERWIG)

Perturbative evolution ends at $Q^2 = Q_0^2$

Angular ordering $\Rightarrow$ colour preconfinement

Forced gluon splitting $(g \to q\bar{q})$

Colour-singlet clusters decay into the observed hadrons



String model (PYTHIA)

q and $\bar{q}$ move in opposite direction

The colour field collapses into a string, with uniform energy density

$q\bar{q}$ pairs are produced

The string breaks into the observed hadrons

Monte Carlo tunings involve hadronization and 'perturbative' parameters: shower cutoffs, $\Lambda_{\rm MC}$, m_g, etc.

Cluster and string models also used in conjunction with other Monte Carlo programs (MadGraph, ALPGEN, POWHEG, SHERPA, aMC@NLO)

Feasibility of PYTHIA to use Peterson, Kartvelishvili, Bowler/Lund models

Standard analyses use tuning to $e^+e^- \to b\bar{b}$ data and best-fit parameters to predict b-fragmentation in hadron collisions, such as top events

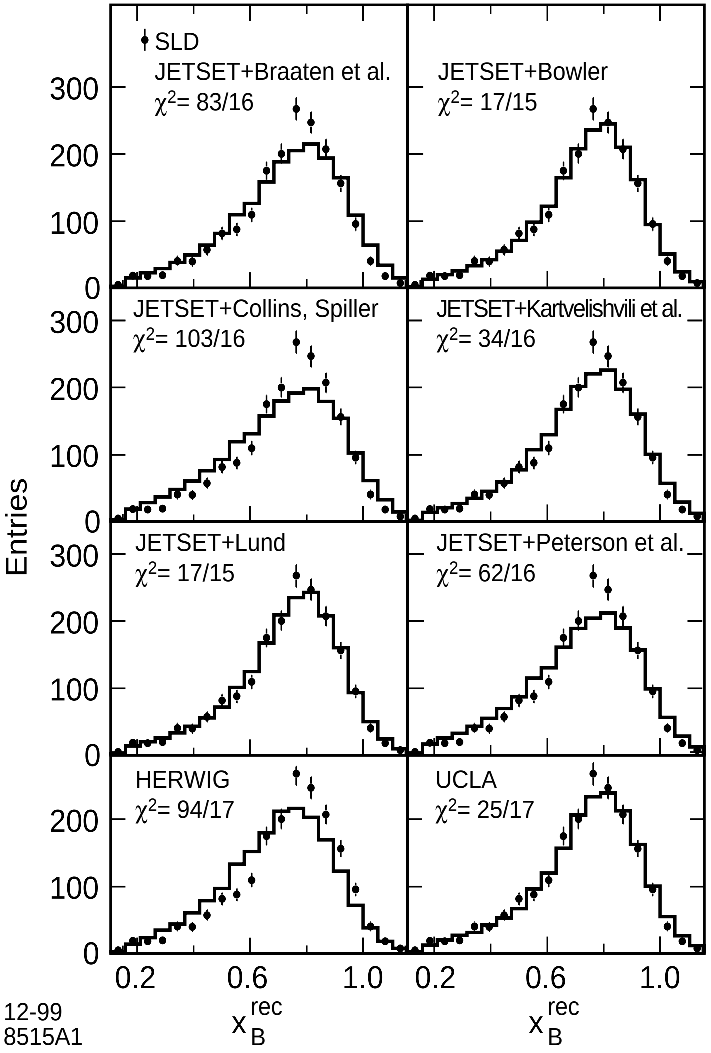
e^+e^- data are very precise and clean, but one lacks information on initial-state radiation and possible colour connection between initial and final state

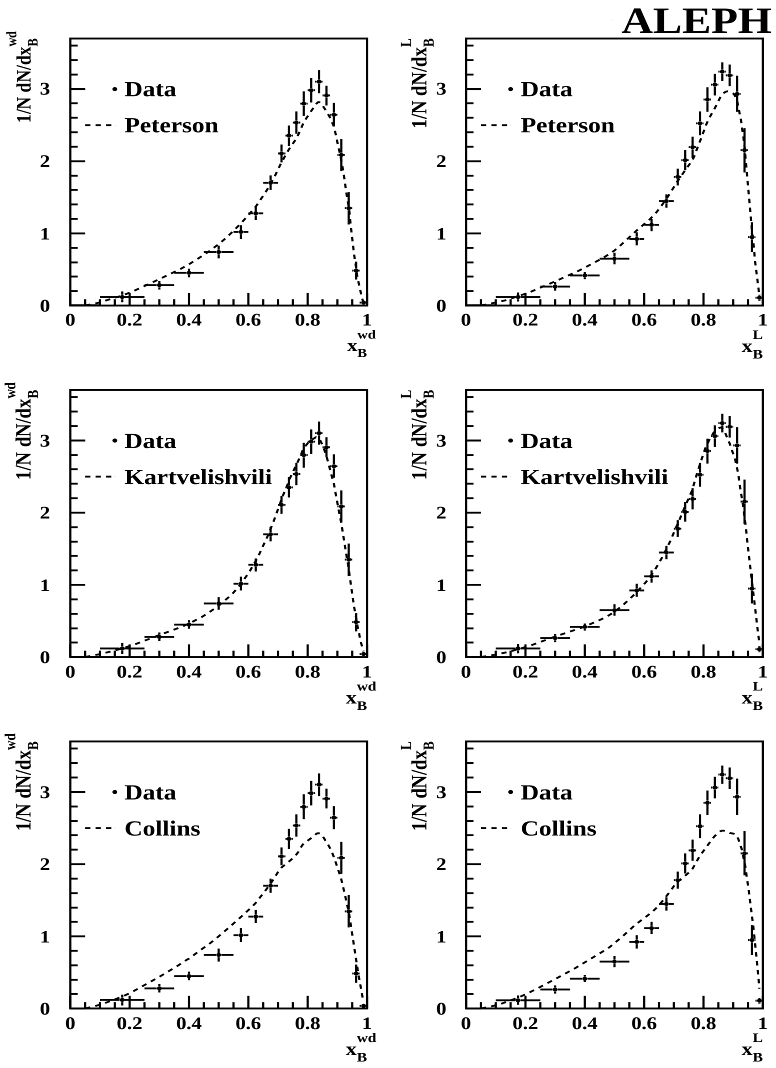
Suitable Lorentz-invariant observables relying on top decays typically show little dependence on production mechanism and initial-state effects

Tuning Monte Carlo models directly to LHC data, e.g. $t\bar{t}$, $b\bar{b}$ or $\gamma/Z + b$ not yet carried out, but can be useful to validate Monte Carlo programs and test hadronization models and factorization

Possibility to minimize impact of production phase and ISR by constructing clever observables, e.g. $p_T(B)/p_T(Z)$ in $Z + b$ events

Bottom-quark fragmentation at the Z^0 pole





LEP tuning of PYTHIA+Peterson used in $J/\psi + \ell$ analysis

Best-fit parameters not the same, e.g. $\epsilon_b = 0.0033$ (ALEPH), 0.0055 (SLD); $\alpha_K = 11.9$ (OPAL), 13.7 (ALEPH), 10.0 (SLD)

G. C. and V. Drollinger, NPB (2005): weakly-decaying B-hadron data from OPAL (mesons and baryons), ALEPH (only mesons) and SLD (mesons and baryons)

HERWIG	PYTHIA
CLSMR(1) $=$ 0.4 (0.0)	
CLSMR(2) $=$ 0.3 (0.0)	PARJ(41) $=$ 0.85 (0.30)
DECWT $=$ 0.7 (1.0)	PARJ(42) $=$ 1.03 (0.58)
CLPOW $=$ 2.1 (2.0)	PARJ(46) $=$ 0.85 (1.00)
PSPLT(2) $=$ 0.33 (1.00)	
χ^2/dof $=$ 222.4/61 (739.4/61)	χ^2/dof $=$ 45.7/61 (467.9/61)

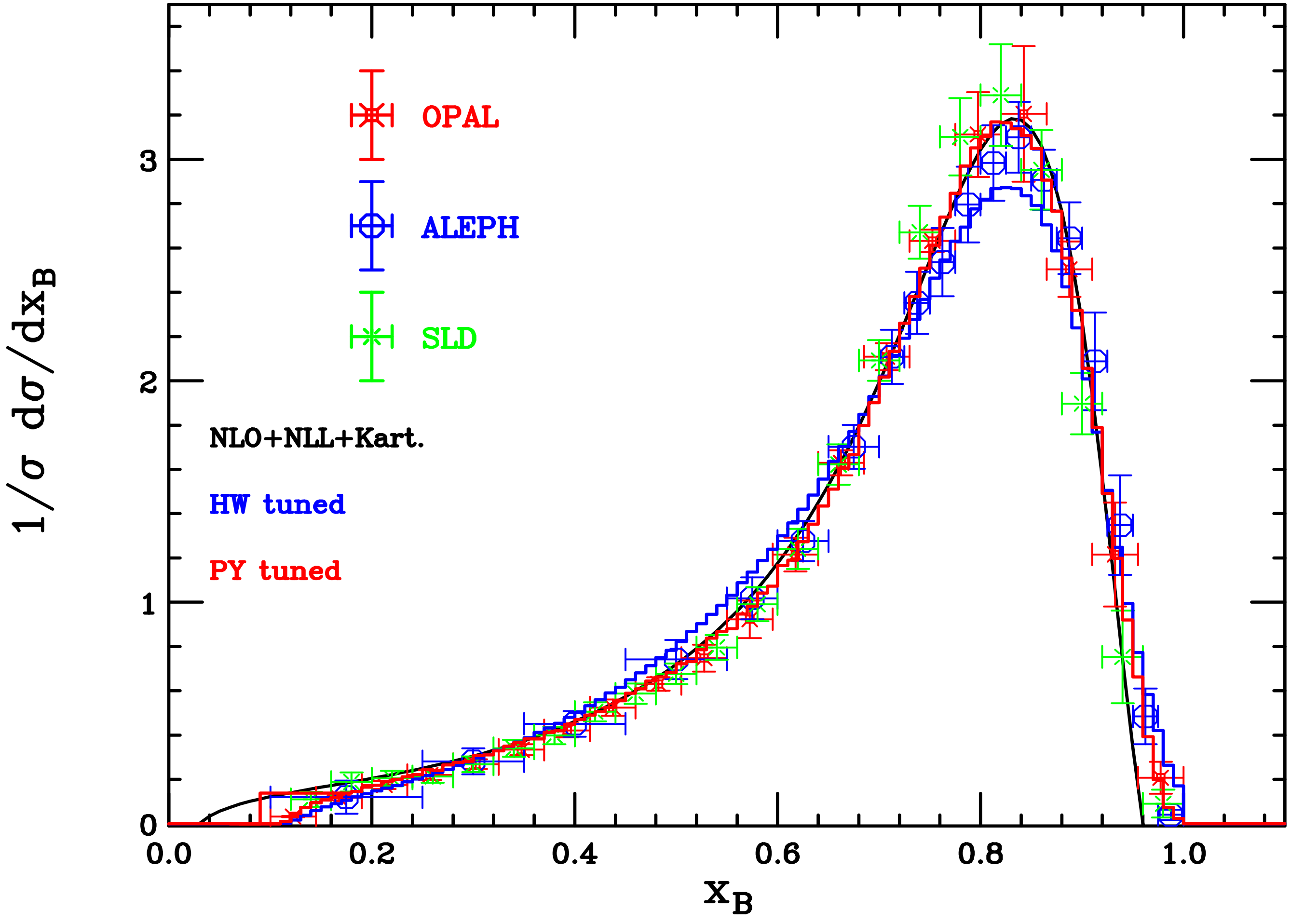
Lund/Bowler fragmentation function (PYTHIA):

$$f_B(z) \propto \frac{1}{z^{1+brm_b^2}}(1-z)^a \exp(-bm_T^2/z)$$

HERWIG tuned parameters describe hadron gaussian smearing **(CLSMR)**, baryon/meson **(CLPOW)** and decuplet/octet **(DECWT)** ratios, mass spectrum of b-like clusters **(PSPLT)**

Our PYTHIA tuning in ATLAS jet-energy measurement (EPJ C73 (2013) 2304) and as a cross-check for top analyses

Comparing tuned HERWIG and PYTHIA

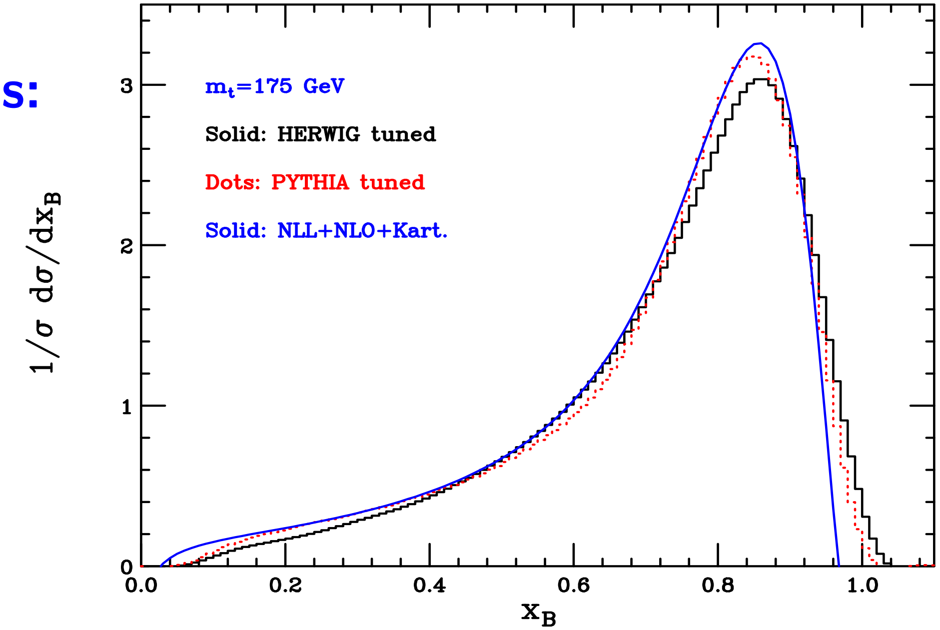


NLO+NLL calculation with Kartvelishvili model: M.Cacciari and S.Catani '01

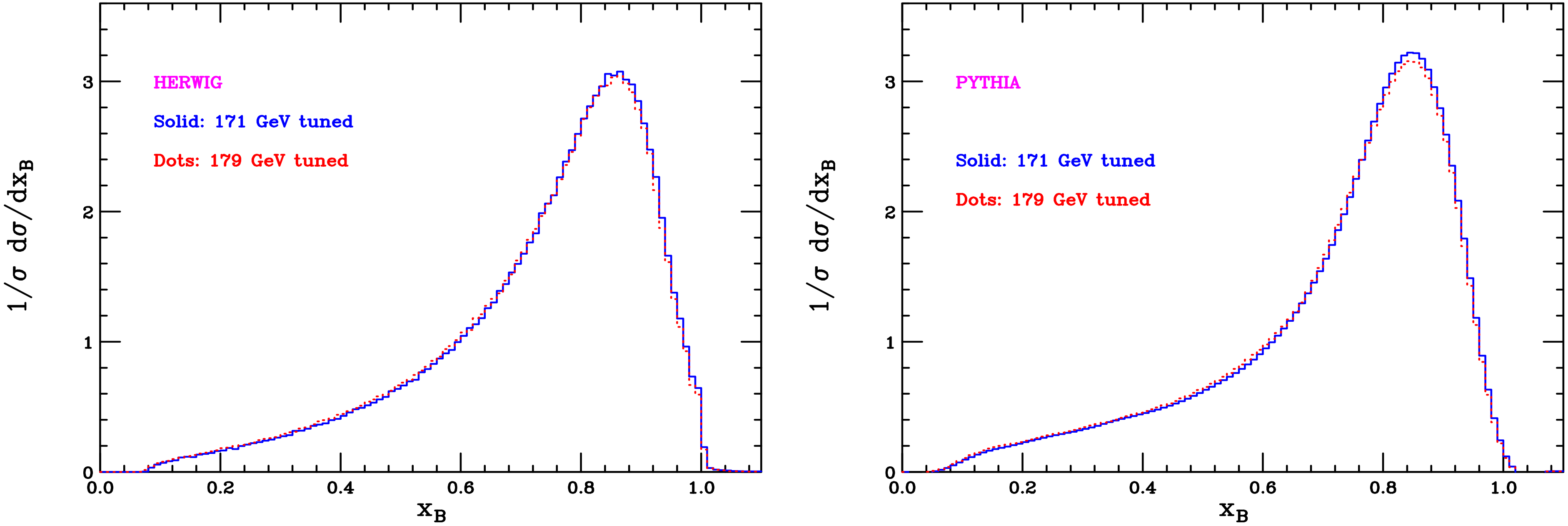
$$D_{\rm np}(x_B,\alpha) = (1+\alpha)(2+\alpha)x_B(1-x_B)^\alpha$$

Best fit ($0.18 \le x_B \le 0.94$): $\alpha = 17.178 \pm 0.303, \quad \chi^2/\mathrm{dof} = 46.2/53$

B-hadron spectrum in top decays:



Mild dependence on the top mass in both HERWIG and PYTHIA:



Discussion with CMS/ATLAS folks: x_B hard to measure experimentally

Results in moment space

$$\Gamma_N = \int_0^1 dz\ z^{N-1} \frac{1}{\Gamma}\frac{d\Gamma}{dz}(z)$$

e^+e^- annihilation $\sigma_N^B = \sigma_N^b D_N^{np}$

σ_N^B measured ; σ_N^b calculated ; D_N^{np} fitted

top decay: $\Gamma_N^B = \Gamma_N^b D_N^{np} = \Gamma_N^b \sigma_N^B / \sigma_N^b$

Fits to DELPHI data (DELPHI 2002-069 CONF 603)

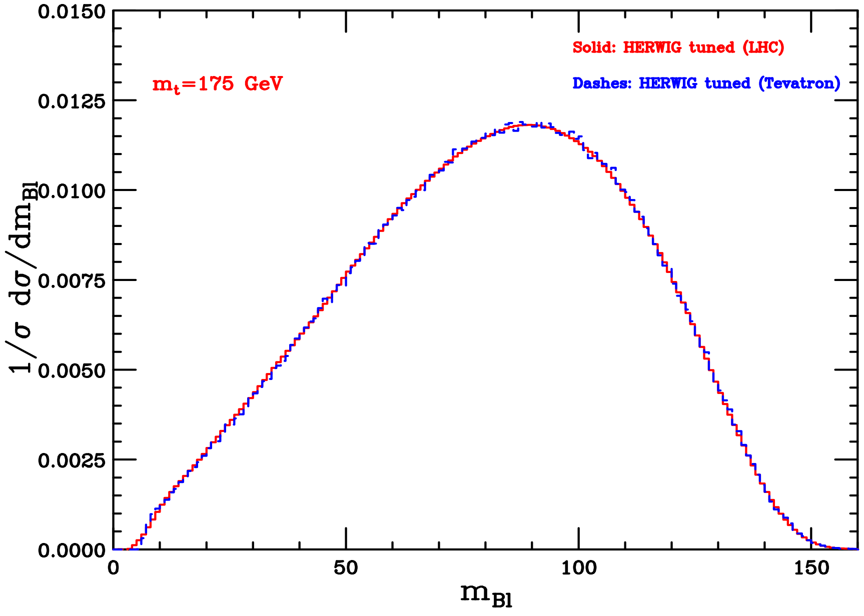
	$\langle x \rangle$	$\langle x^2 \rangle$	$\langle x^3 \rangle$	$\langle x^4 \rangle$
e^+e^- data σ_N^B	0.7153±0.0052	0.5401±0.0064	0.4236±0.0065	0.3406±0.0064
e^+e^- NLL σ_N^b	0.7801	0.6436	0.5479	0.4755
D_N^{np}	0.9169	0.8392	0.7731	0.7163
e^+e^- HW σ_N^B	0.7113	0.5354	0.4181	0.3353
e^+e^- PY σ_N^B	0.7162	0.5412	0.4237	0.3400
t-dec. NLL Γ_N^b	0.7883	0.6615	0.5735	0.5071
t-dec. NLL $\Gamma_N^B = \Gamma_N^b D_N^{\mathrm{np}}$	0.7228	0.5551	0.4434	0.3632
t-dec. HW Γ_N^B	0.7325	0.5703	0.4606	0.3814
t-dec. PY Γ_N^B	0.7225	0.5588	0.4486	0.3688

b-fragmentation is process independent, but top events are different from e^+e^- collisions, because of ISR, pdfs, colour connection between initial and final states

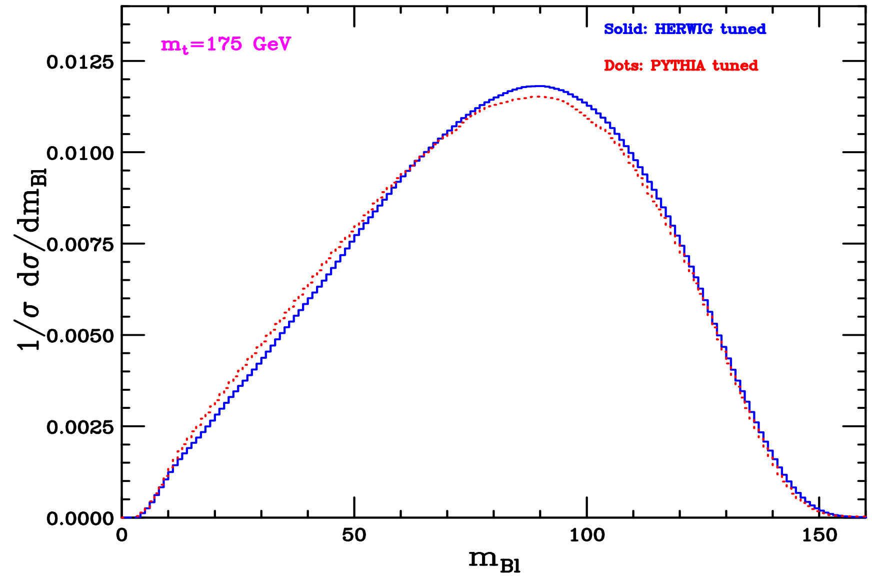
Lorentz-invariant quantities: mild dependence on ISR and colour flow

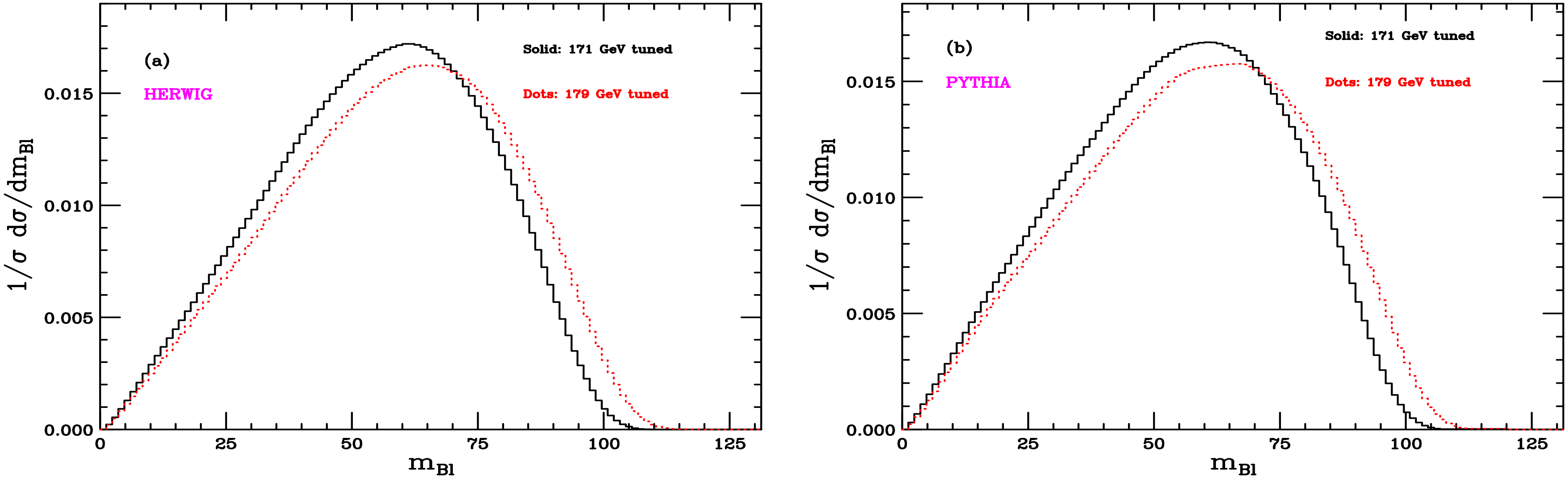
Example: $m_{B\ell}$ invariant-mass spectrum in dilepton channel is the same for LHC, Tevatron and e^+e^- collisions (G.C., M.L. Mangano and M.H. Seymour, JHEP 0007 (2000) 004)

m_t (GeV)	$\langle m_{B\ell}\rangle_{\rm LHC}$ (GeV)	$\sigma_{\rm LHC}$ (GeV)	$\langle m_{B\ell}\rangle_{\rm TeV}$ (GeV)	$\sigma_{\rm TeV}$ (GeV)
171	91.13	26.57	91.18	26.51
173	92.42	26.90	92.31	26.90
175	93.54	27.29	93.41	27.29
177	94.61	27.66	94.65	27.73
179	95.72	28.04	95.64	28.00



$m_{B\ell}$ **according to tuned HERWIG and PYTHIA** (G.C. and F.Mescia, EPJ '10)





Other observables, e.g. B-hadron p_T, though useful to gauge the acceptance, depend on ISR and top-production phase

Mellin moments - full $m_{B\ell}$ spectrum - tuned HERWIG and PYTHIA

HERWIG:

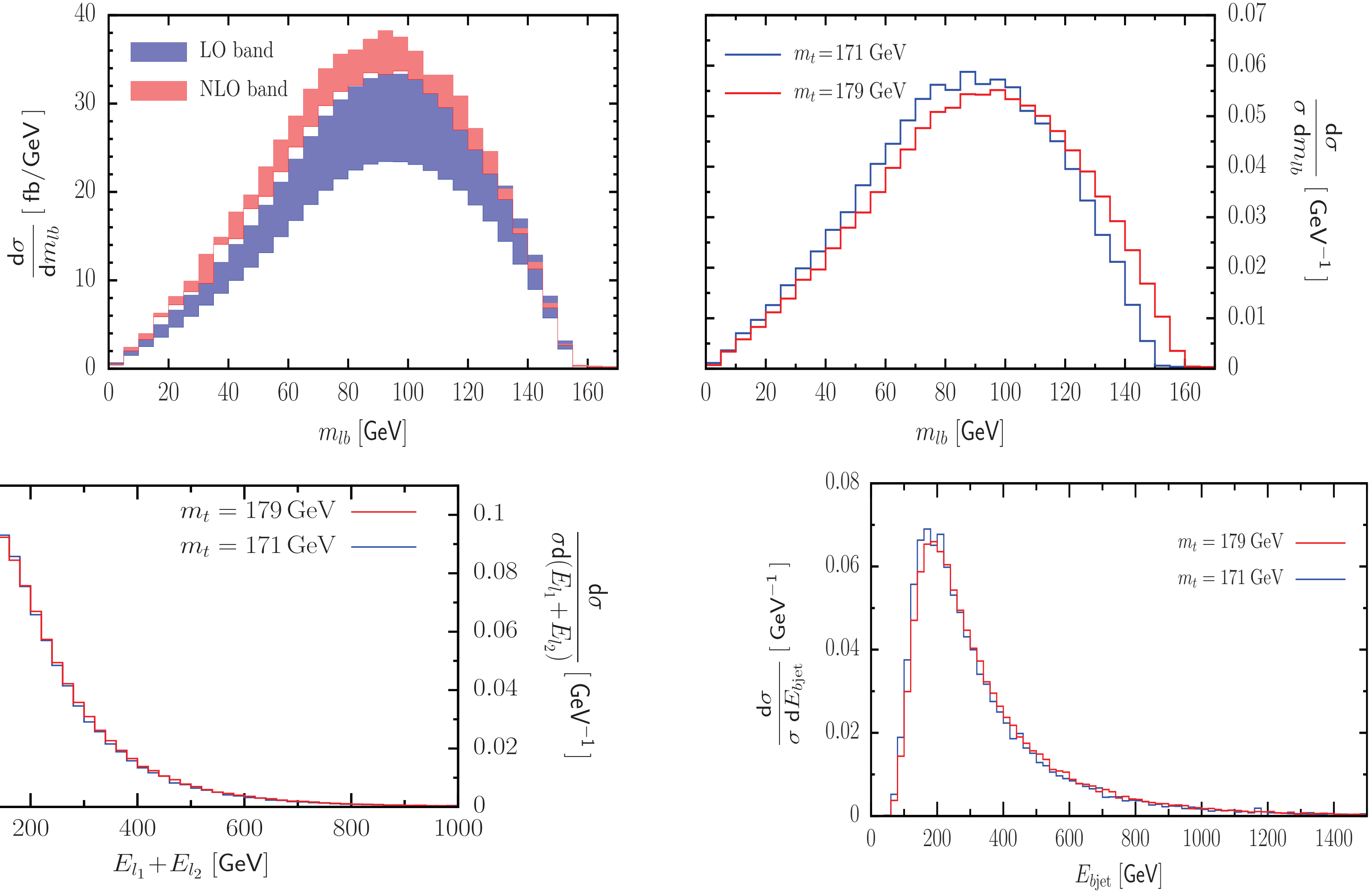
m_t (GeV)	$\langle m_{B\ell} \rangle$ (GeV)	$\langle m_{B\ell}^2 \rangle$ (GeV^2)	$\langle m_{B\ell}^3 \rangle$ (GeV^3)	$\langle m_{B\ell}^4 \rangle$ (GeV^4)
171	78.39	7.01×10^3	6.82×10^5	7.02×10^8
173	79.52	7.22×10^3	7.12×10^5	7.43×10^8
175	80.82	7.45×10^3	7.46×10^5	7.91×10^8
177	82.02	7.67×10^3	7.79×10^5	8.37×10^8
179	83.21	7.89×10^3	8.13×10^5	8.86×10^8

PYTHIA:

m_t (GeV)	$\langle m_{B\ell} \rangle$ (GeV)	$\langle m_{B\ell}^2 \rangle$ (GeV^2)	$\langle m_{B\ell}^3 \rangle$ (GeV^3)	$\langle m_{B\ell}^4 \rangle$ (GeV^4)
171	77.17	6.85×10^3	6.62×10^5	6.81×10^8
173	78.37	7.06×10^3	6.94×10^5	7.23×10^8
175	79.55	7.27×10^3	7.25×10^5	7.67×10^8
177	80.70	7.48×10^3	7.56×10^5	8.12×10^8
179	81.93	7.71×10^3	7.91×10^5	8.61×10^8

Results on NLO invariant mass, lepton and b-energy spectra

(Biswas, Melnikov and Schulze, ’10)



Hadronization correction by using fits in G.C. and A.D.Mitov ’01 (no soft res.)

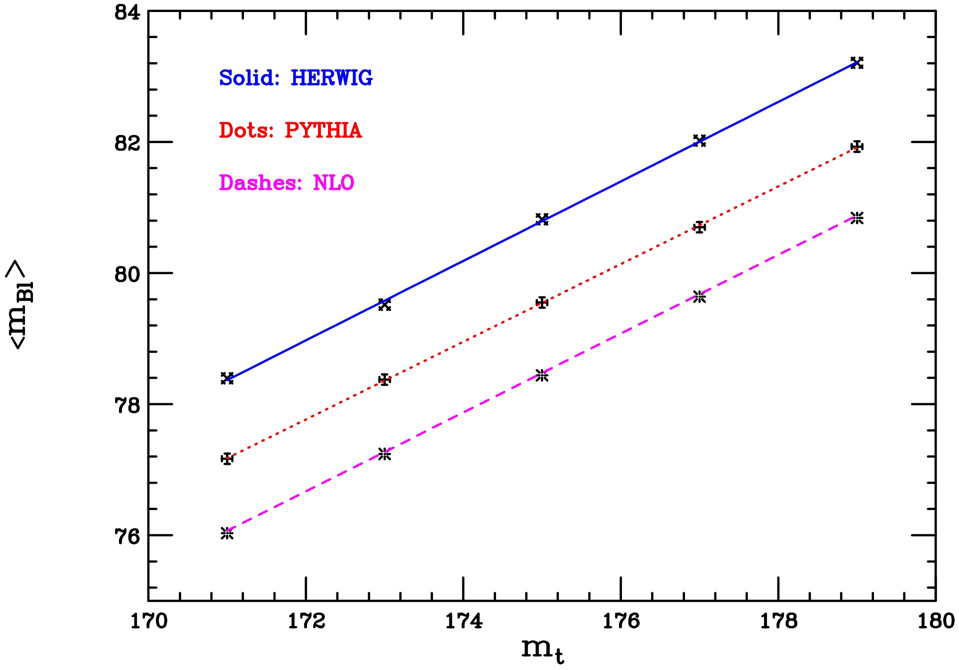
$m_{B\ell}$ independent of production only if no cuts, e.g. on W decay products, are applied

Linear fits to extract m_t from $m_{B\ell}$

HERWIG: $\langle m_{B\ell}\rangle_{\rm H} \simeq -25.31\ {\rm GeV} + 0.61\ m_t$; $\delta = 0.043$ GeV

PYTHIA: $\langle m_{B\ell}\rangle_{\rm P} \simeq -24.11\ {\rm GeV} + 0.59\ m_t$; $\delta = 0.022$ GeV

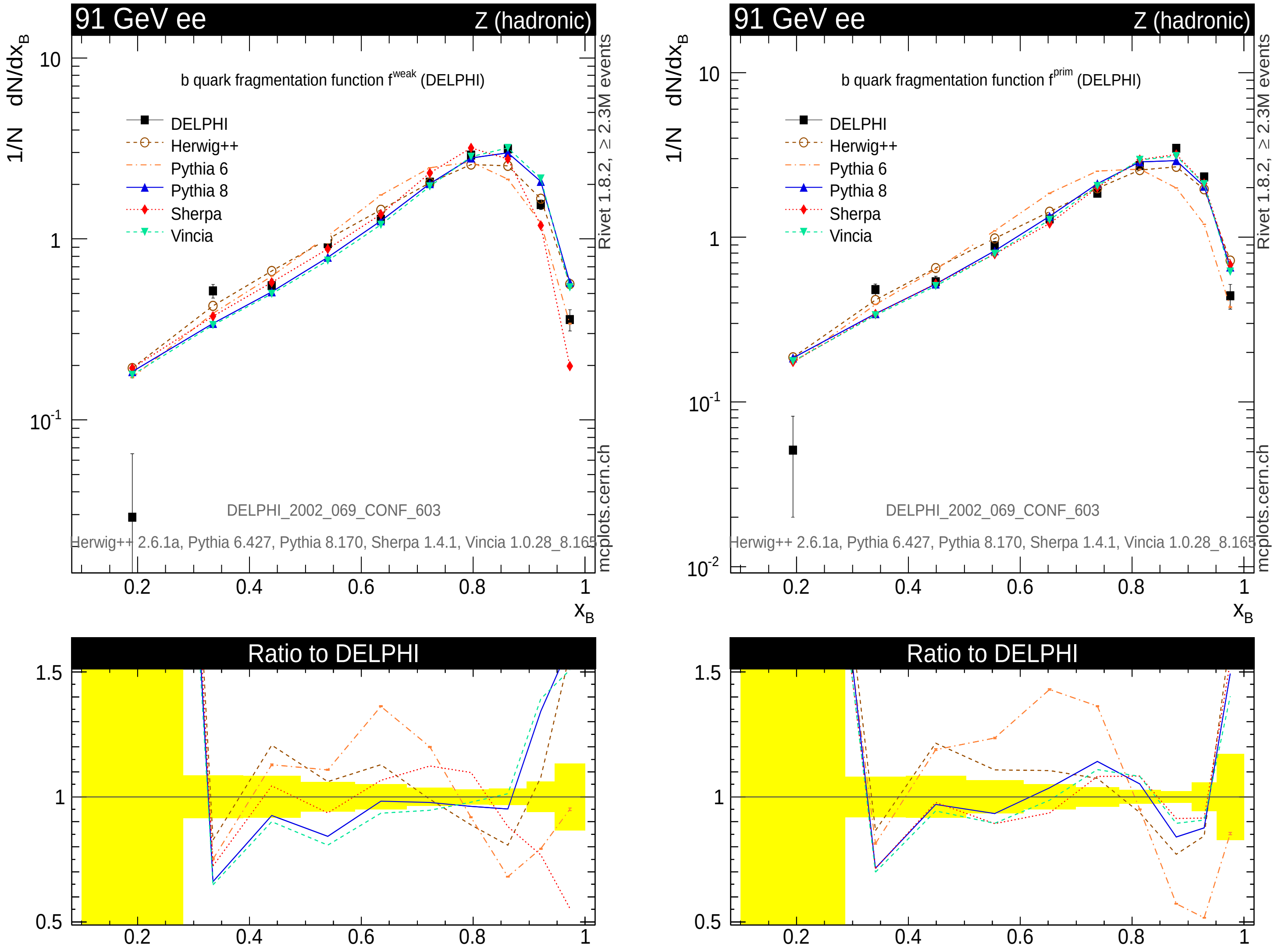
NLO: $\langle m_{B\ell}\rangle_{\rm NLO} \simeq -26.7\ {\rm GeV} + 0.60\ m_t$; $\delta = 0.004$ GeV



$\Delta\langle m_{B\ell}\rangle_{\rm H,P} \simeq 1.2$ GeV ; $\Delta\langle m_{B\ell}\rangle_{\rm H,NLO} \simeq 2.2$ GeV ; $\Delta\langle m_{B\ell}\rangle_{\rm P,NLO} \simeq 1.1$ GeV

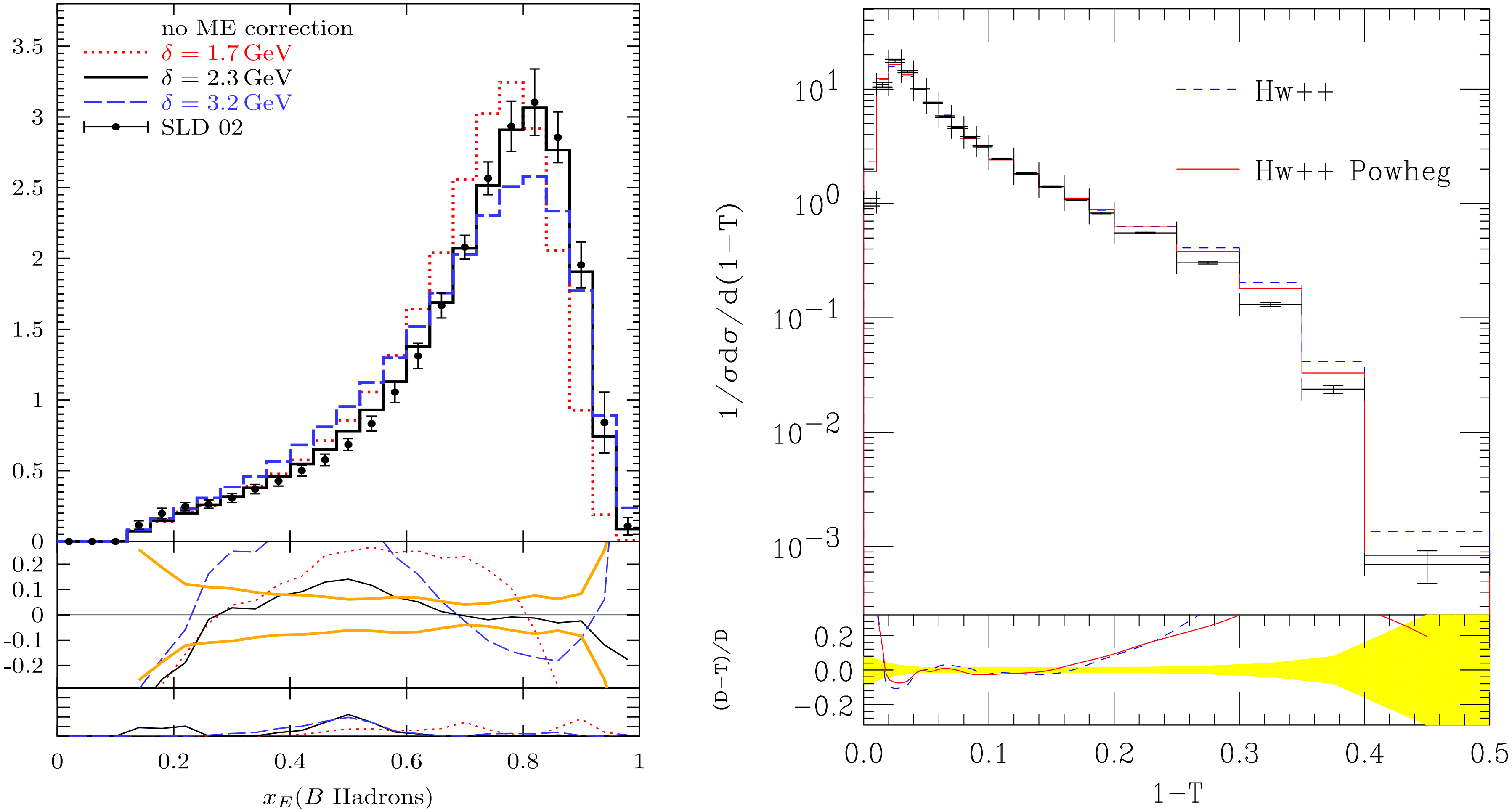
NLO+showers for top decays or C++ codes may shed light on this discrepancy

C++ generators for b-quark fragmentation in e^+e^- annihilation (mcplots.cern.ch)



Discrepancies call for tuning hadronization models

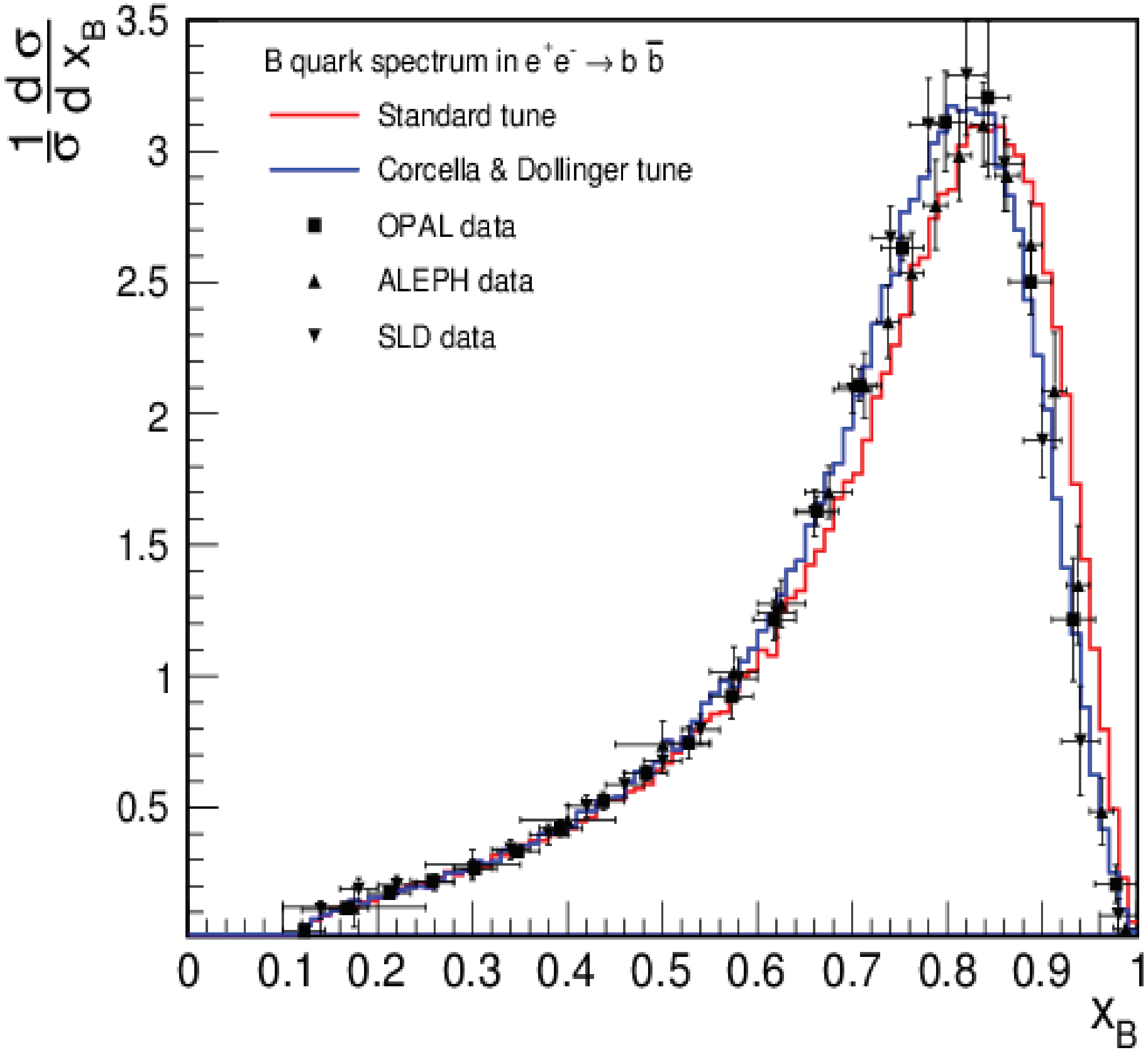
HERWIG++ : improved fragmentation model and a few processes according to POWHEG method



Left: comparison of HERWIG++ with SLD data on B-hadron energy fraction (δ: shower cutoff)

Right: comparison of HERWIG++, with $e^+e^- \to q\bar{q}$ in the POWHEG scheme, vs. DELPHI data on thrust distribution (1-T)

G.C., F. Mescia, D. Torres Moreno and K. Tywoniuk: HERWIG++ and PYTHIA 8



PYTHIA 8 - Default: $\chi^2/\mathrm{dof} = 189/56$; tuned: $\chi^2/\mathrm{dof} = 28/56$

In progress: using HERWIG++ with and without POWHEG e^+e^- annihilation

Perspectives in (N)NLO+shower codes

POWHEG: top production at NLO, not yet e^+e^- annihilation

NLO corrections to top decays in POWHEG are available, but turned off in default version: work in progress

aMC@NLO: NLO e^+e^- and hadronic top production for stable top quarks

No full NLO corrections to top decays, but finite-width effects available, for single-top production including both resonant and non-resonant (non-top) diagrams (talk by R. Frederix)

Off-shellness impact competitive with NLO corrections to top decays

Code for $pp \to b\bar{b}\ell^+\nu_\ell\ell^-\bar{\nu}_\ell$ including resonant and non-resonant (non-top) diagrams soon available

Perspectives to include shower and hadronization in the NLO calculation with off-shell tops by A.Denner et al, '12

Interest by the authors of NNLO exclusive top decays (J.Gao et al. '12, F.Caola et al.'12) to include b-fragmentation at NNLO, but not straightforward (e.g. A.Mitov and K.Melnikov NNLO perturbative fragmentation functions in the $\overline{\mathrm{MS}}$ factorization scheme)

Conclusions and outlook

Bottom fragmentation in top decays as a source of uncertainty in the measurement of the top properties in inclusive (b-tagging and b-energy scale) and exclusive analysis ($J/\psi + \ell$)

LO+shower codes and NLO calculations for b-fragmentation, tuning hadronization models to e^+e^- data

Predictions for top decays exhibit some discrepancies, mostly driven by unsatisfactory tunings

Preliminary results with object-oriented codes exhibit better description of b-fragmentation in e^+e^- collisions after the tuning

Perspectives:

Comparing tuned PYTHIA and HERWIG++ can be a valuable way to estimate b-hadronization systematics

Extending the analysis to NLO+showers tools (POWHEG and aMC@NLO with off-shell effects) and ultimately NNLO calculations

Tuning fragmentation parameters directly to LHC data ($t\bar{t}$, $b\bar{b}$, $Z/\gamma + b$) and comparison with e^+e^- fits to test factorization and quality of hadronization models