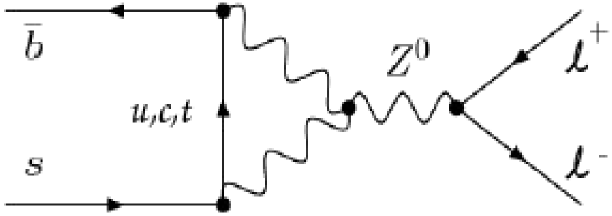
Measurement of the $B_s \rightarrow \mu\mu$ BF

Maren Ugland

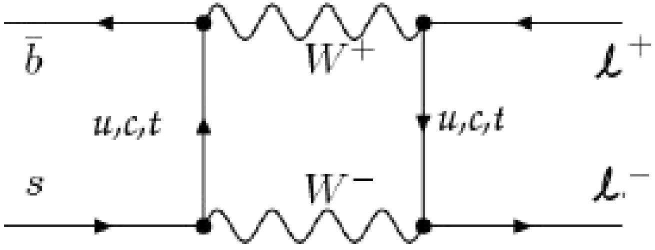
Bergen group meeting - 25.11.2009

Introduction

- Neutral B-meson decays to two leptons are very sensitive to new physics. Since $B_s^0 \rightarrow l^+ l^-$ involves a flavor changing neutral current, the lowest order contributions within the standard model occur at one-loop level.
- The photonic penguin does not contribute because C-parity violation is not possible with an $\ell = 0$ transition.
- The standard model Higgs penguin has a very small contribution, and can safely be neglected.
- The decay is Cabbibo suppressed because of the $b \rightarrow s$ transition, and helicity suppressed since angular momentum conservation forces the two leptons to have the same chirality.
- New physics may contribute to the branching fractions of neutral B-meson both by introducing new particles to the loops, and by allowing the transitions to occur at tree level.



Feynman penguin diagram



Feynman box diagram

Calculation of decay rates

- The decay of a heavy quark is a weak decay, but the calculation of decay rates is complicated by the fact that quarks do not exist as free objects, but are bound together in hadrons.
- Use operator product expansion (OPE) to separate the non-perturbative long-distance effects from the perturbatively calculable short-distance effects of the transition.
- In OPE the effective Hamiltonian describing the b→s transition is:

$$\mathcal{H}_{\text{eff}} = -\frac{G_F}{\sqrt{2}} V_{tb} V_{ts}^* \sum_i C_i(\mu) \mathcal{O}_i(\mu),$$

where the short distance effects are described by scale dependent Wilson coefficients, $C_i(\mu)$, and the long-distance effects are described by local operators O_i.

- From the effective Hamiltonian, we find the decay amplitude:

$$\mathcal{M} = \frac{G_F}{\sqrt{2}} V_{\text{CKM}} \sum_i C_i(\mu) \langle f | \mathcal{O}_i | B \rangle$$

The Wilson coefficients are universal in the sense that they do not depend on either the initial or final state.

Calculation of decay rates

- Semi-leptonic operators:

$$\mathcal{O}_{9V} = \frac{\alpha_{\rm em}}{2\pi}\left(\overline{Q}_\alpha \gamma_\mu P_L b_\alpha\right)\left(\ell^+\gamma^\mu \ell^-\right)$$
$$\mathcal{O}_{10A} = \frac{\alpha_{\rm em}}{2\pi}\left(\overline{Q}_\alpha \gamma_\mu P_L b_\alpha\right)\left(\ell^+\gamma^\mu \gamma_5 \ell^-\right)$$

- Values of corresponding Wilson coefficients (NNLO):
  $C_{9V} = 4.214$ and $C_{10A} = -4.312$

- When including physics beyond the standard model, the operators may be modified, or new operators with corresponding Wilson coefficients may be introduced.

- The largest contributions arise from models which introduce new scalar and pseudo-scaler particles, which may mediate the decay. (For instance Models with an extended Higgs sector.)

- The standard model branching fraction:

$$\mathcal{B}(B_s^0 \to \ell^+\ell^-) = \frac{G_F^2\alpha^2}{64\pi^3 \sin^4\theta_W}|V_{tb}^* V_{ts}|^2 \tau_{B_s^0} M_{B_s^0}^3 f_{B_s^0}^2 \sqrt{1-\frac{4m_\ell^2}{M_{B_s^0}^2}}\left(-\frac{2m_\ell}{M_{B_s^0}} C_{10A}\right)^2$$

- The CKM dependence and most of the hadronic uncertainty may be eliminated by normalizing to the well measured meson mass difference, ΔM_s

Measurement of $\mathcal{B}(B_s \to \mu^+\mu^-)$

- In ATLAS, only modes with two muons are considered accessible, and of these only $B_s \to \mu\mu$ is expected to have a branching fraction large enough to be important during the first years of data taking

- Standard model prediction (normalized to ΔM_s):

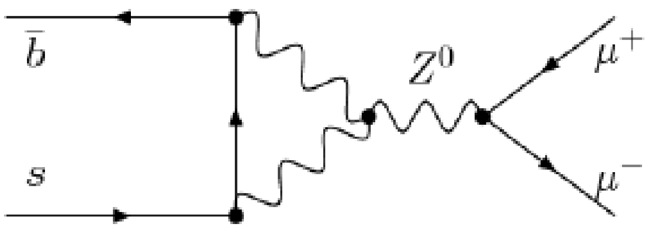
  $\mathcal{B}(B_s \to \mu^+\mu^-) = (3.35 \pm 0.32) \cdot 10^{-9}$ (hep-ph/0604057, 2006)

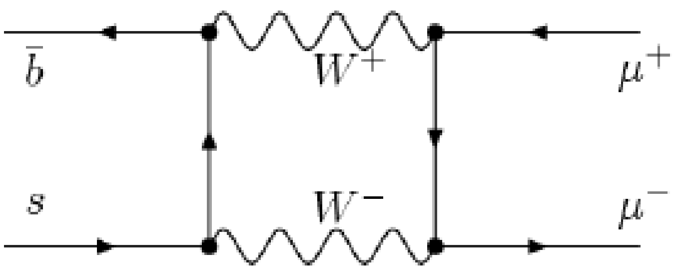
- Current best limit is set by CDF using 3.7fb^{-1} of data:

  $\mathcal{B}(B_s \to \mu^+\mu^-) = 4.3 \cdot 10^{-8}$ @ 95% CL (CDF Public Note 9892, 2009)

- D0 limit, using 1.3fb^{-1} of data:

  $\mathcal{B}(B_s \to \mu^+\mu^-) = 1.2 \cdot 10^{-7}$ @ 95% CL (hep-ex/0707.3997v1, 2007)





Measurement of $\mathcal{B}(B_s \to \mu^+\mu^-)$ in ATLAS

- Since the number of B_s mesons produced is not well measured, the preferred method for determining the branching fraction $B_s \to \mu\mu$ uses $B_s \to J/\psi(\mu\mu)\ \varphi(KK)$ and $B^+ \to J/\psi(\mu\mu)\ K^+$ as normalization modes

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = \frac{N_S}{\alpha_S \epsilon_S} \cdot \frac{\alpha_N \epsilon_N}{N_N} \cdot \frac{f_u}{f_s} \cdot \mathcal{B}(B^+ \to J/\psi K^+) \mathcal{B}(J/\psi \to \mu^+\mu^-)$$

where N_S (N_N) is the observed signal yield, α_S (α_N) is the dimuon trigger efficiency, and ϵ_S (ϵ_N) is the signal reconstruction efficiency for $B_s \to \mu^+\mu^-$ ($B^+ \to J/\psi K^+$).

- The ratio f_u/f_s represents the difference in b-quark fragmentation probability for producing a u- or an s-quark.

$f_u/f_s = 3.78 \pm 0.52$ (hep-ex/0704.3575, 2007)

Measurement of normalization modes

- $B^+ \to J/\psi\ K^+$ has a large branching fraction, which has been determined with high precision, and is reconstructed with reasonable efficiency in ATLAS.

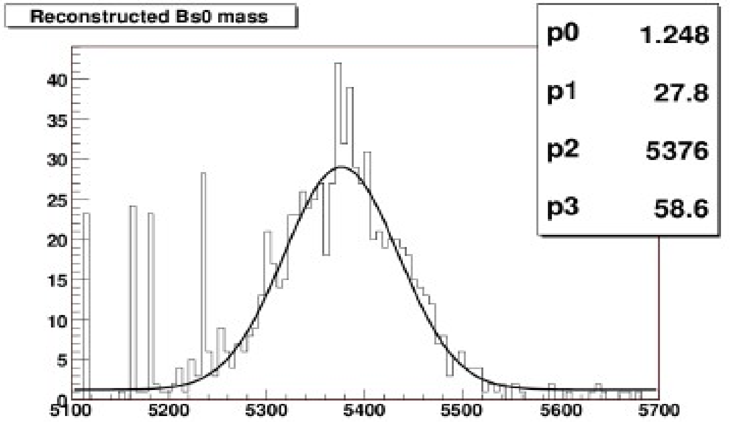
- $B_s \to J/\psi\ \varphi$ is expected to yield a lower precision, but has the advantage of being independent of the b-quark fragmentation probability.

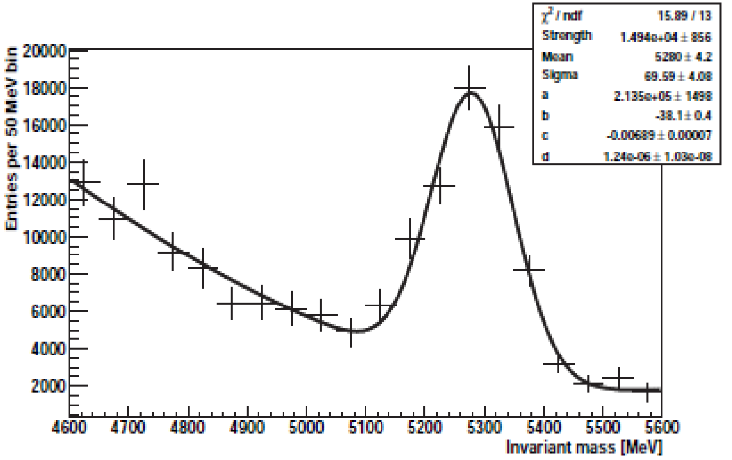
- The reconstructed invariant mass of $B_s \to J/\psi(\mu\mu)\ \varphi(KK)$:
  5376.2 ± 2.6 MeV

  PDG value: 5366,3 ± 0.6 MeV.
  Difference is understood.

- The reconstructed invariant mass of $B^+ \to J/\psi(\mu\mu)\ K^+$:
  5280 ± 4.2 MeV

  PDG value: 5279.15 ± 0.31 MeV.





Reconstructed B_s (B^+) mass, fitted with a Gaussian and a polinomial.

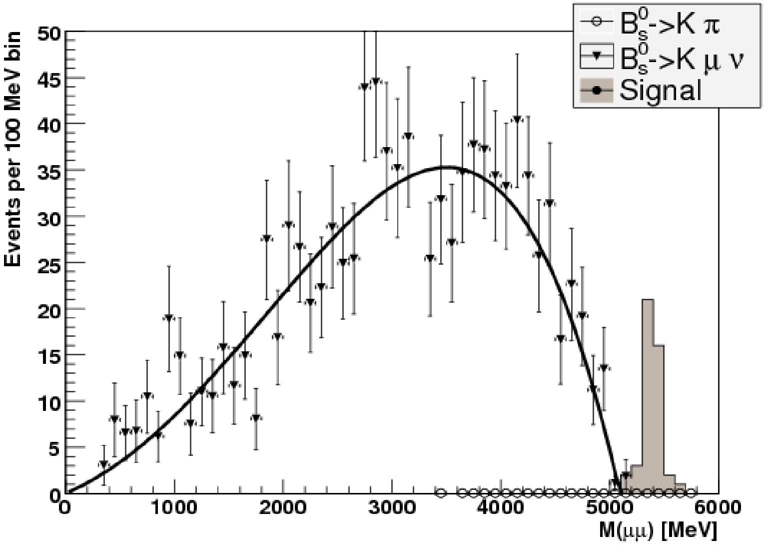
Measurement of $\mathcal{B}(B_s \rightarrow \mu^+\mu^-)$ in ATLAS

- Two analysis strategies are applied:
  - Cut-based analysis — - Simple procedure, easy to understand.
    - Provides a cross check for multivariate analysis
  - Boosted decision tree (BDT) analysis — - Requires excellent understanding of data
    - Increases efficiency by 35 % with respect to cut-based analysis

- Background from exclusive semi-leptonic or hadronic two-body decays is negligible.

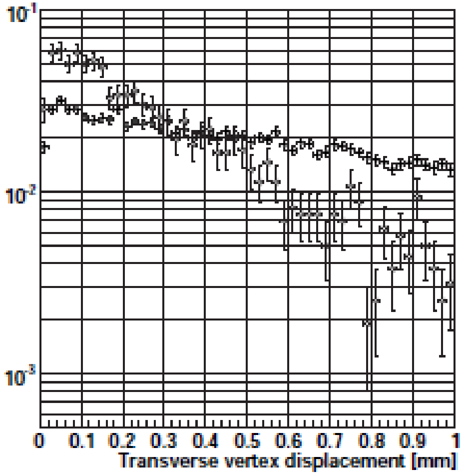
- Combinatorial background (e.g. muons from other B-hadrons) is dominant source of background.

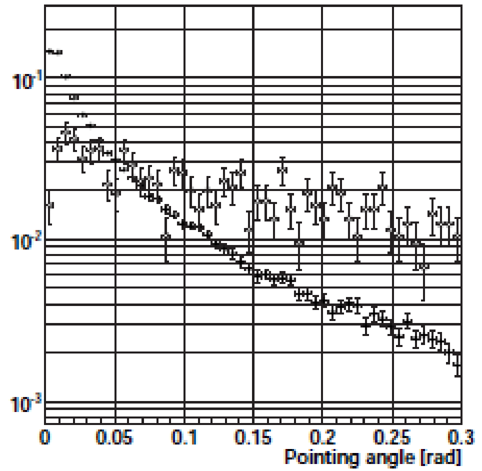
- Analysis is based on previous work by Trygve Buanes (MC study at 7TeV beam energy)

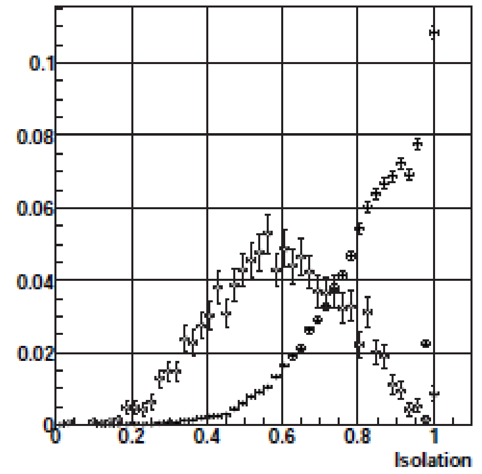


Signal selection

- For the cut-based analysis there are three important discrimination variables:
  - Transverse decay length
  - Pointing angle
  - Isolation



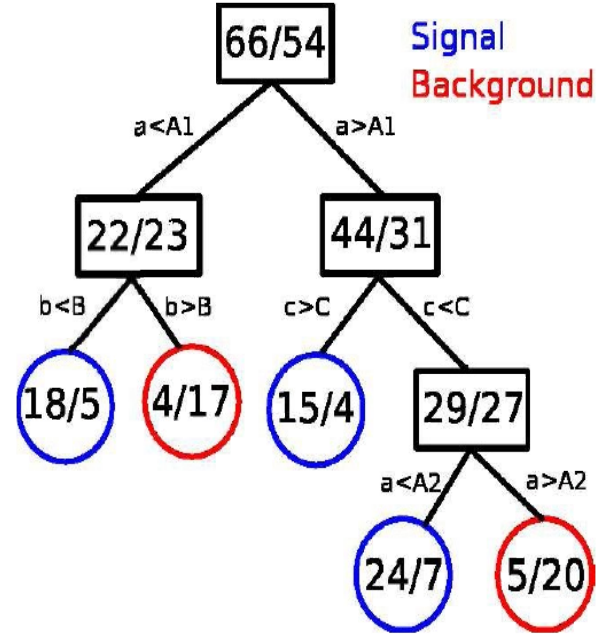




Signal is shown as solid points, while the open circles represent the combinatorial background.

- The BDT analysis makes use of a large number of additional selection variables.

Boosted Decision Trees



- Decision trees use a known set of signal and background events, finds the variable that gives the best S/B separation, and divides the sample into 2 subsets (nodes).
  The procedure is then repeated for each subset, until some stopping criterion is reached.

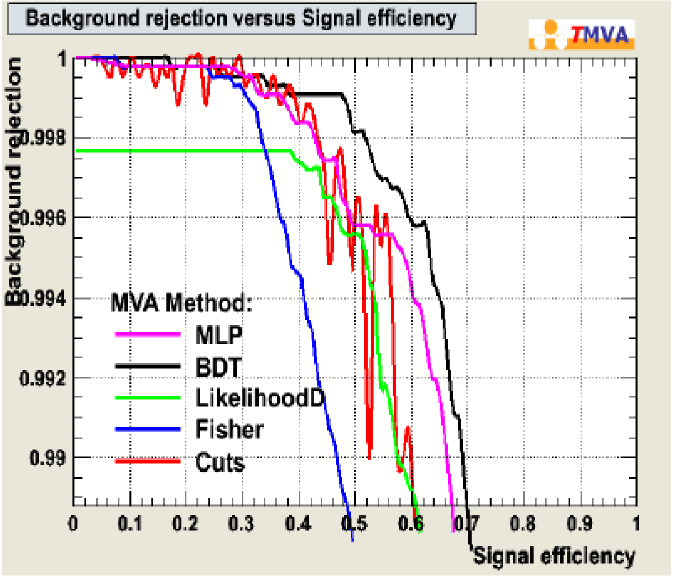
- There are 3 main classes of stopping criteria:
  - Purity of end nodes
  - Size of end nodes
  - Number of end nodes

- An important difference from a cut-based analysis, is that the decision tree allows a single variable to be used several times (with different thresholds)

- After training the tree on a known sample, it is ready to be used to separate signal from background events in a sample of unknown composition.

Boosted Decision Trees

- Boosting is introduced because the final tree is very sensitive to fluctuations in the training sample.
- Rather than just making one tree, a large number of trees are made. Events that are misclassified in one tree are reweighted, or boosted, before making the next tree so that the probability is enhanced for the new tree to classify these events correctly.
- After training, events are classified by majority vote, but giving more weight to the trees with a better S/B separation.
- Usually check the BDT performance on a test sample to make sure it is not over trained.
- Performance is as good, or better than, alternative multivariate analyses.



Next steps

- Redo complete simulation study for 5TeV beam energy.
- Include more variables in the BDT.
- Determine the combinatorial background for 5TeV beam energy (sample of 1M events available).
- 'Real data' is expected soon. (3.5TeV beam energy?)
  When data comes, determine combinatorial background from sidebands.

Conclusion

- First results by end of next year?

  Assuming 200 pb-1 of data, we expect to be able to measure an upper limit
  $\mathcal{B}(B_s \rightarrow \mu^+\mu^-) < 8 \cdot 10^{-8}$ @ 90% CL

  CDF+D0 expect to set a limit $\sim 2 \cdot 10^{-8}$ (2010)

- Improvements on current best limit, probably a few more years?
  Need 1-2fb^{-1} to be competitive with CDF+D0

- To measure the standard model prediction, $\mathcal{B}(B_s \rightarrow \mu^+\mu^-) = (3.35 \pm 0.32) \cdot 10^{-9}$,
  $\sim 10fb^{-1}$ of data is needed.

- LHCb claims evidence (3σ) with $\sim 4fb^{-1}$ of data.

(Keep in mind that the 2 experiments do not collect data at the same rate!)

→ Competitive with LHCb with early data.

THE END!