Investigating $B \rightarrow \tau\nu_\tau$ *at* **BaBar** *with New Statistical Techniques*

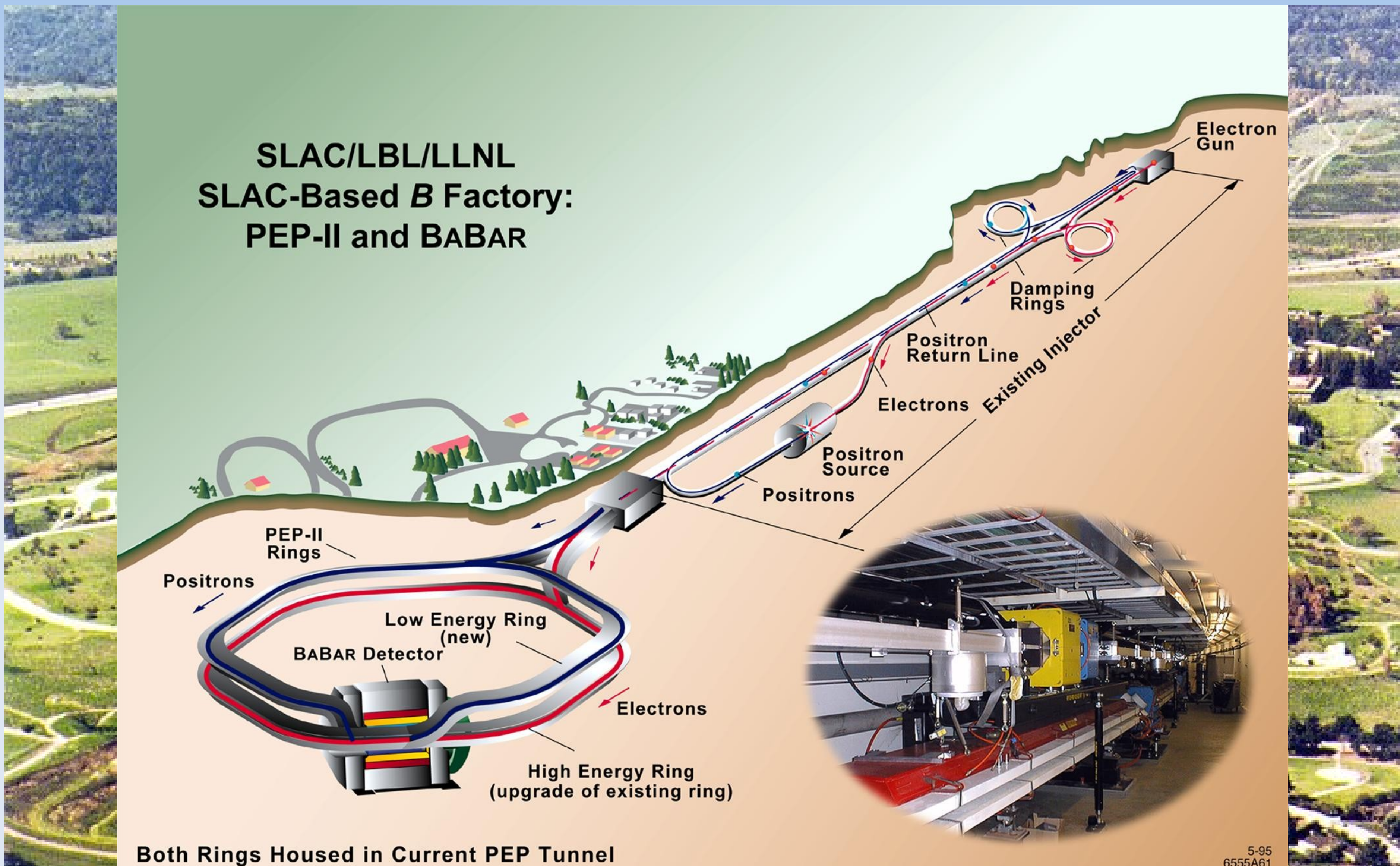
Matthew Barrett
Dept of Electronic and Computer Engineering
Brunel HEP group
Brunel University

Brunel
UNIVERSITY
WEST LONDON

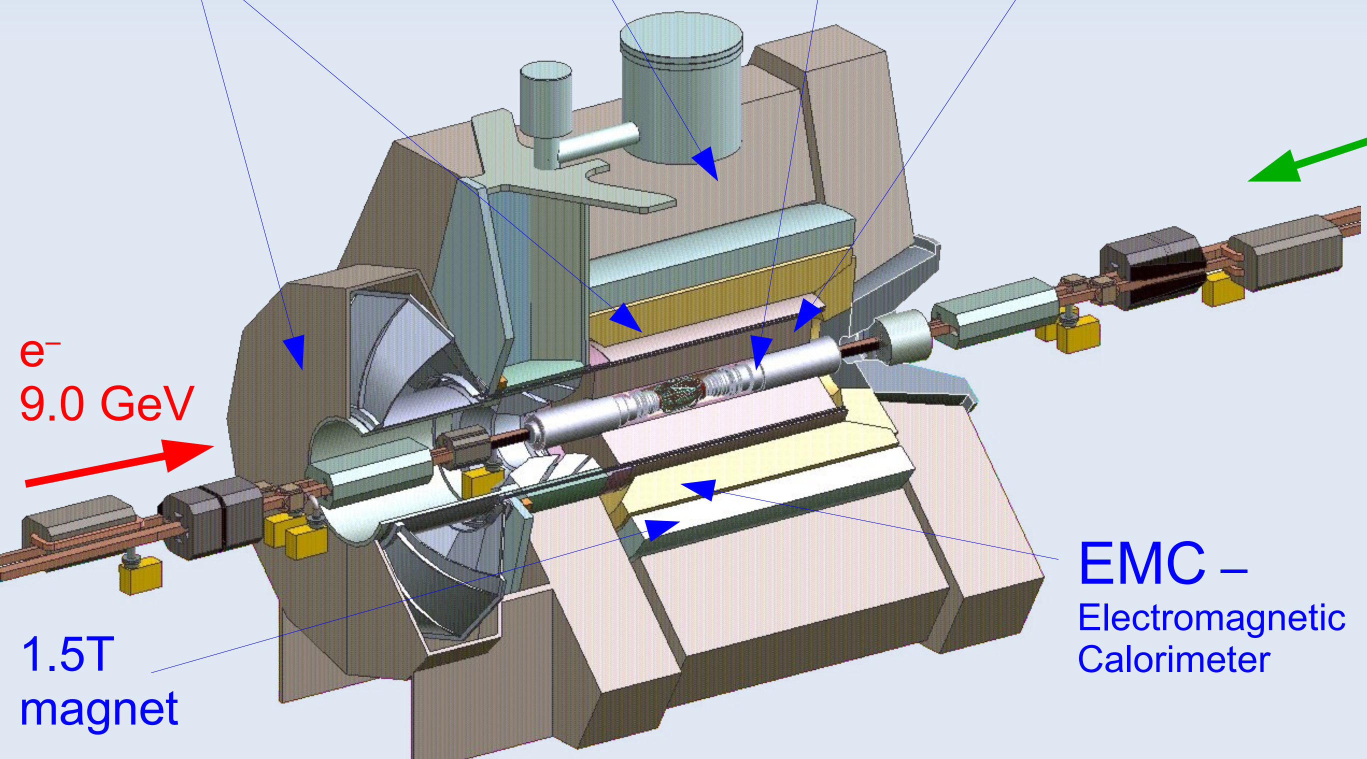
Outline of Talk

- The BaBar Experiment.
- $B \to \tau\nu$ – Why is it interesting?
- How to study $B \to \tau\nu$.
- Current Measurements from BaBar and Belle, and Summer 2008 updates.
- Improving the measurements with new statistical techniques.
- The future for BaBar and beyond...

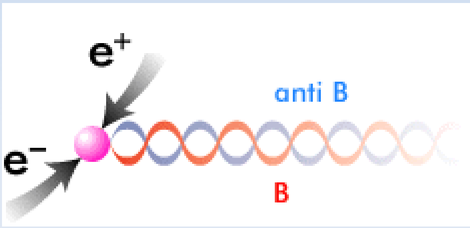
The BaBar Experiment



The BABAR Experiment



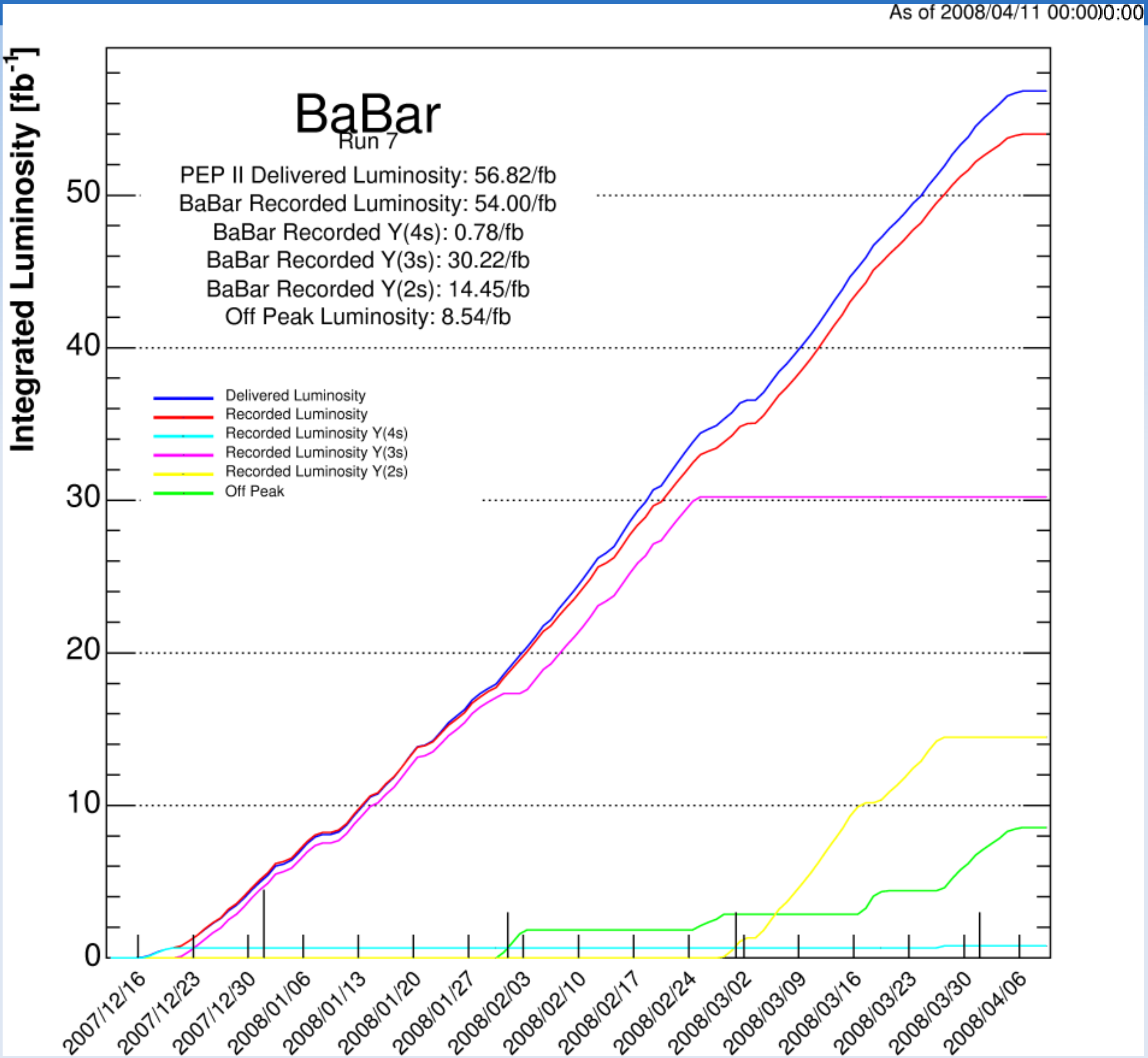
- Centre of Mass Energy = 10.58GeV.
- Mass of Υ(4S).



- Just above threshold for BB production.
- B mesons almost at rest.
- $\beta\gamma = 0.56$

The BaBar Experiment

- BaBar started data taking: 1999.
- Finished at 12:43 on April 7 2008.
- After running on Y(3S) and Y(2S).
- Off Peak: (mostly) 40 MeV below Y(4S).
  - No B mesons produced.
- Mass of Y(3S) = 10.355GeV/c^2.



Why Study $B \to \tau\nu$?

- Physics motivated by one equation:

$$\mathcal{B}(B^- \to \ell^- \bar{\nu}) = \frac{G_F^2 m_B}{8\pi} m_l^2 \left(1 - \frac{m_l^2}{m_B^2}\right)^2 f_B^2 |V_{ub}|^2 \tau_B$$

- Parameters of Note:
  - f_B – *B* meson decay constant.
  - Can only access via purely leptonic *B* decays.
  - Current value from Lattice QCD:
    $f_B = (189 \pm 27)$ MeV.

Why Study $B\to\tau\nu$?

- Physics motivated by one equation:

$$\mathcal{B}(B^- \to \ell^-\bar{\nu}) = \frac{G_F^2 m_B}{8\pi} m_l^2 \left(1 - \frac{m_l^2}{m_B^2}\right)^2 f_B^2 |V_{ub}|^2 \tau_B$$

- Parameters of Note:
  - Mass of daughter lepton m_l.
  - Leads to helicity suppression:

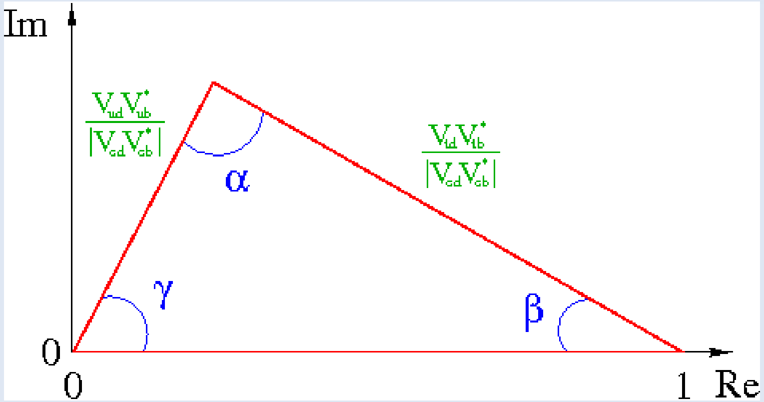
τ	: μ	: e
1	: 5×10^{-3}	: 10^{-7}

Why Study B→τν?

- Physics motivated by one equation:

$$\mathcal{B}(B^- \to \ell^- \bar{\nu}) = \frac{G_F^2 m_B}{8\pi} m_l^2 \left(1 - \frac{m_l^2}{m_B^2}\right)^2 f_B^2 |V_{ub}|^2 \tau_B$$

- Parameters of Note:
  - V_{ub} – CKM matrix element.
  - Current PDG value: $|V_{ub}| = (4.31 \pm 0.30) \times 10^{-3}$.
  - B meson oscillation frequency: $\Delta m_d \propto f_B^2 |V_{td}|^2$.
  - $\mathcal{B}(B \to \tau\nu) / \Delta m_d \propto |V_{ub}|^2 / |V_{td}|^2$

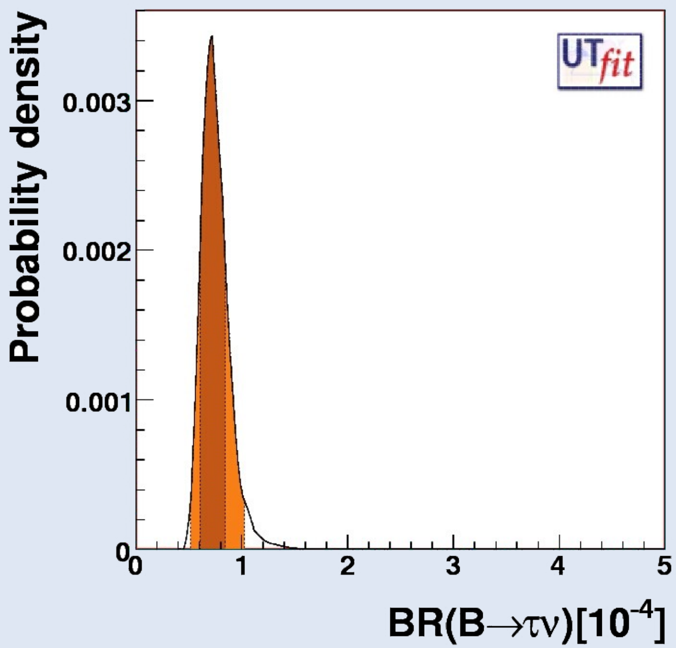


Why Study $B \to \tau\nu$?

- Physics motivated by one equation:

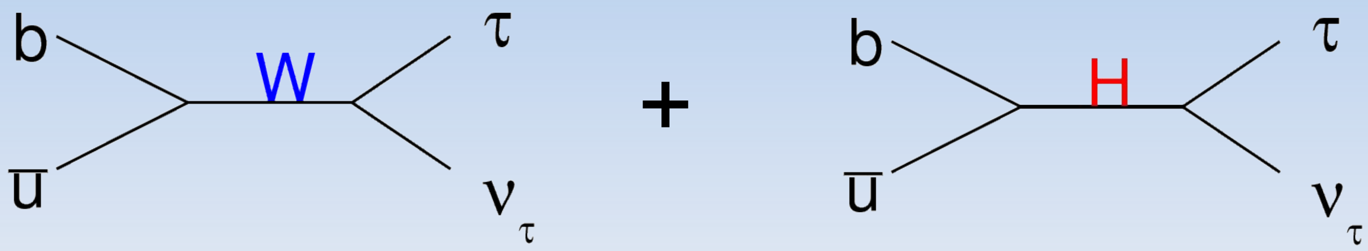
$$\mathcal{B}(B^- \to \ell^- \bar{\nu}) = \frac{G_F^2 m_B}{8\pi} m_l^2 \left(1 - \frac{m_l^2}{m_B^2}\right)^2 f_B^2 |V_{ub}|^2 \tau_B$$

- Standard Model Prediction:
  - $\mathcal{B}(B \to \tau\nu) = (1.6 \pm 0.4) \times 10^{-4}$.
- UTfit prediction:
  - $\mathcal{B}(B \to \tau\nu) = (0.85 \pm 0.14) \times 10^{-4}$. (Vub, no lattice)



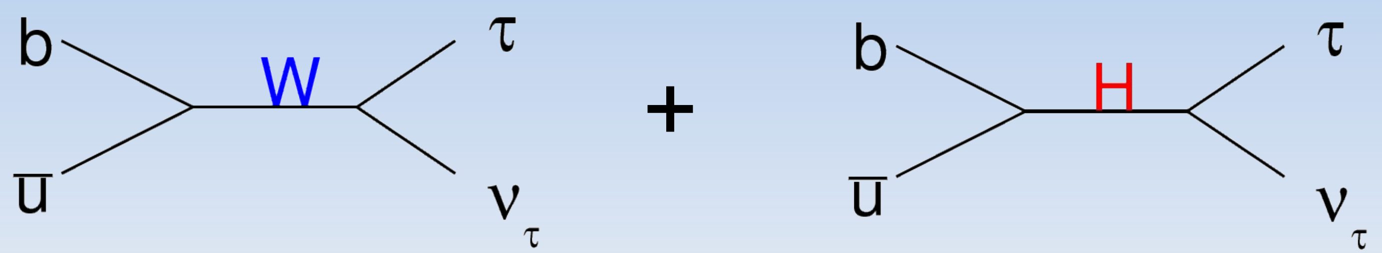
And Beyond the Standard Model?

- Additional Feynman diagram from Higgs boson:



And Beyond the Standard Model?

- Additional Feynman diagram from Higgs boson:



- Two Higgs Doublet Model (2HDM) and Minimal Supersymmetry (MSSM) lead to modified Branching fraction:

$$\mathcal{B}^{2HDM} = \mathcal{B}^{SM}\left(1 - \frac{m_B^2 \tan^2\beta}{m^2_{H^\pm}}\right)^2$$

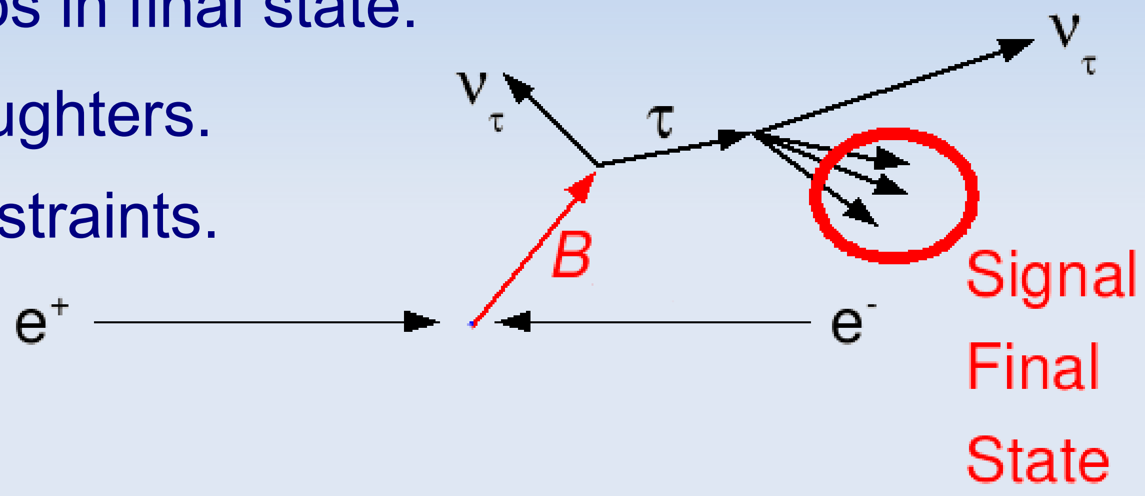
W.S.Hou PRD 48 2342 (1993)

$$\mathcal{B}^{MSSM} = \mathcal{B}^{SM}\left(1 - \left(\frac{m_B^2}{m^2_{H^\pm}}\right)\frac{\tan^2\beta}{1+\epsilon\tan\beta}\right)^2$$

- tan β – ratio of vacuum expectation values.

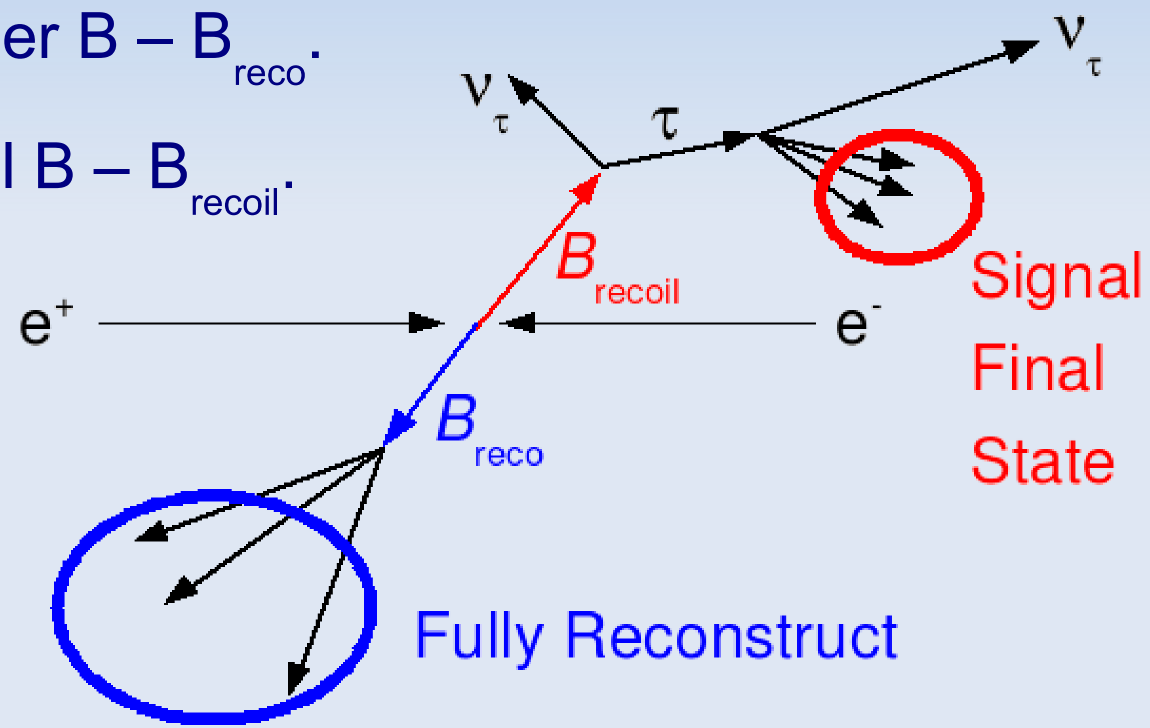
How to look for $B \rightarrow \tau\nu$

- Experimentally challenging:
  - Two or Three neutrinos in final state.
  - Only reconstruct τ daughters.
  - Lack of kinematic constraints.



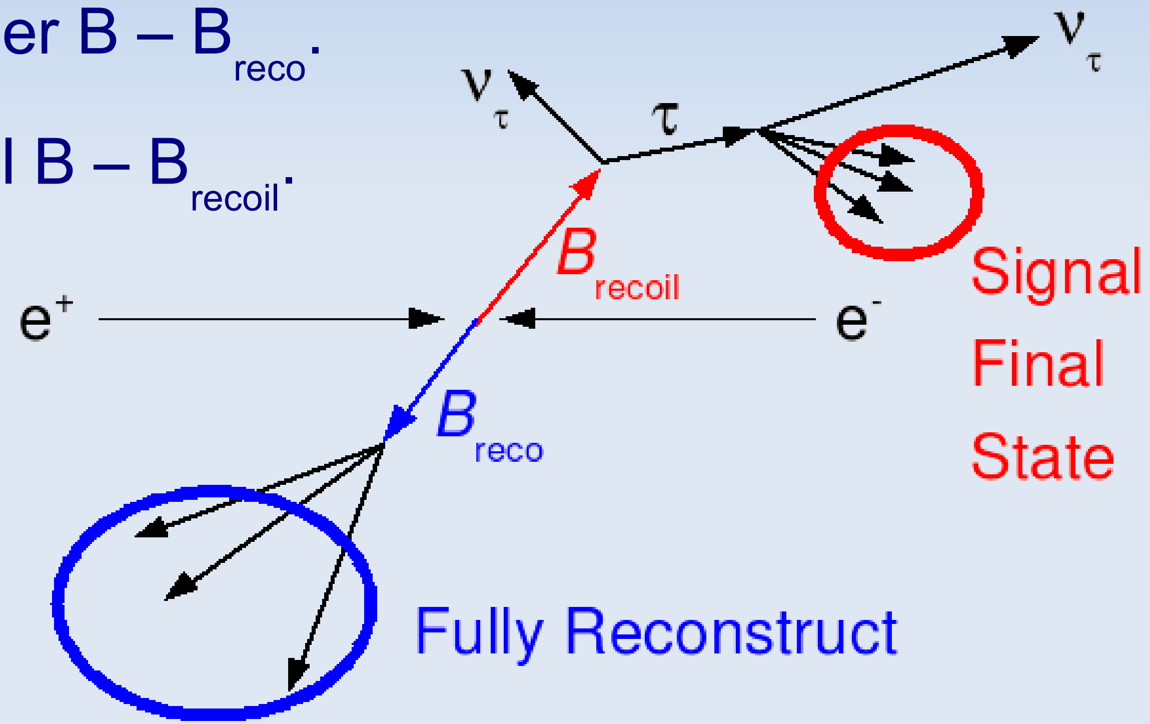
How to look for $B \rightarrow \tau\nu$

- Recoil Analysis technique:
- Fully Reconstruct the other B – B_{reco}.
- This constrains the signal B – B_{recoil}.
- Two different types:
  - Hadronic tag:
    $B \rightarrow DX$ (X = Hadrons – $\pi^{\pm}$, π^0, $K^{\pm}$, K_S)
  - SemiLeptonic tag*:
    $B \rightarrow Dl\nu X$ ($X = \gamma$, π^0, or nothing)
    *fully reconstruct except the neutrino.



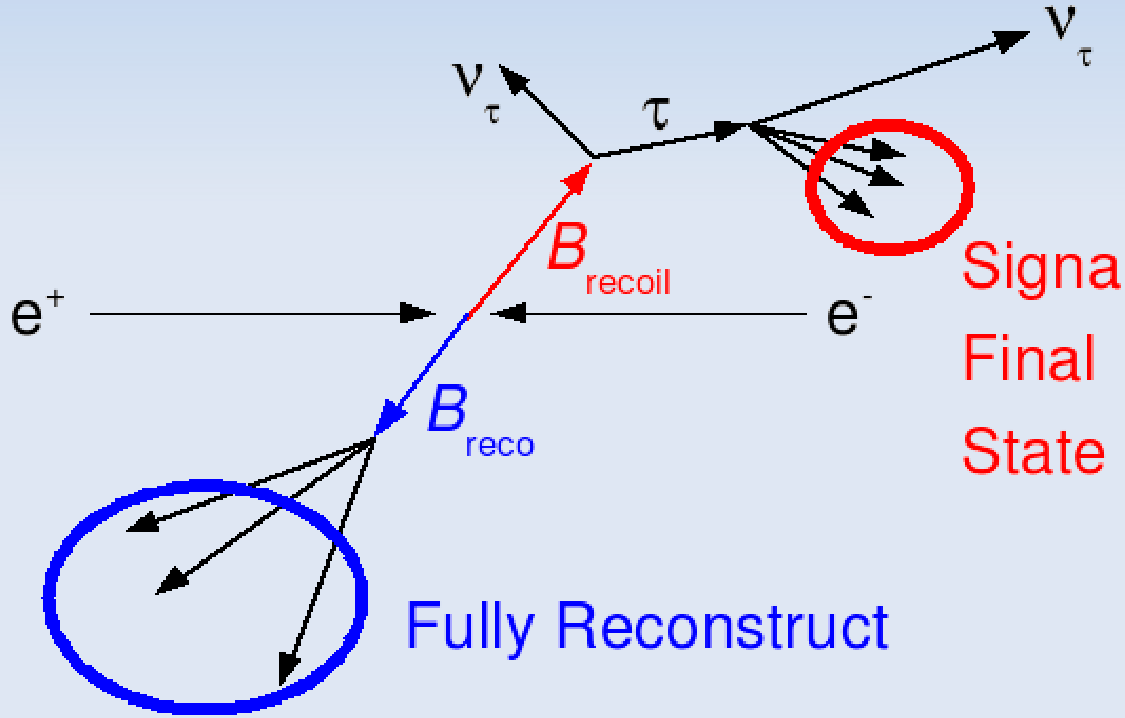
How to look for $B \to \tau\nu$

- Recoil Analysis technique:
- Fully Reconstruct the other B – B_{reco}.
- This constrains the signal B – B_{recoil}.
- Two different types:
  - Hadronic tag:
    $B \to DX$ (X = Hadrons – $\pi^{\pm}$, π^0, $K^{\pm}$, K_S)
  - SemiLeptonic tag*:
    $B \to Dl\nu X$ ($X = \gamma$, π^0, or nothing)
    *fully reconstruct except the neutrino.



How to look for $B \to \tau\nu$

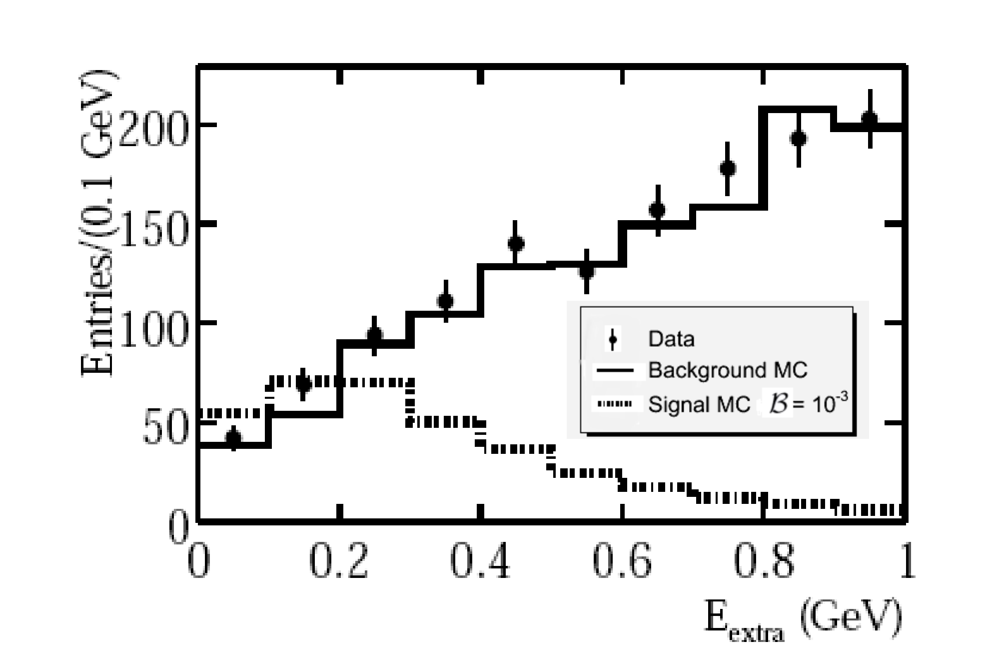
- τ is reconstructed in five modes:
- $\tau^- \to e^- \nu_e \nu_\tau$
- $\tau^- \to \mu^- \nu_\mu \nu_\tau$
- $\tau^- \to \pi^- \nu_\tau$
- $\tau^- \to \rho^- (\pi^- \pi^0) \nu_\tau$
- $(\tau^- \to a_1^- (\pi^+ \pi^- \pi^-) \nu_\tau)$



a_1 is only used in most recent analysis.

The E_{Extra} Variable

- Most discriminating variable available.
- Sum of Energy deposited in Calorimeter,
  that is not attributed to any reconstructed particle.
- Should be (close to) zero for true signal events.
- Background typically much higher.
- Moreover – used to define signal box.

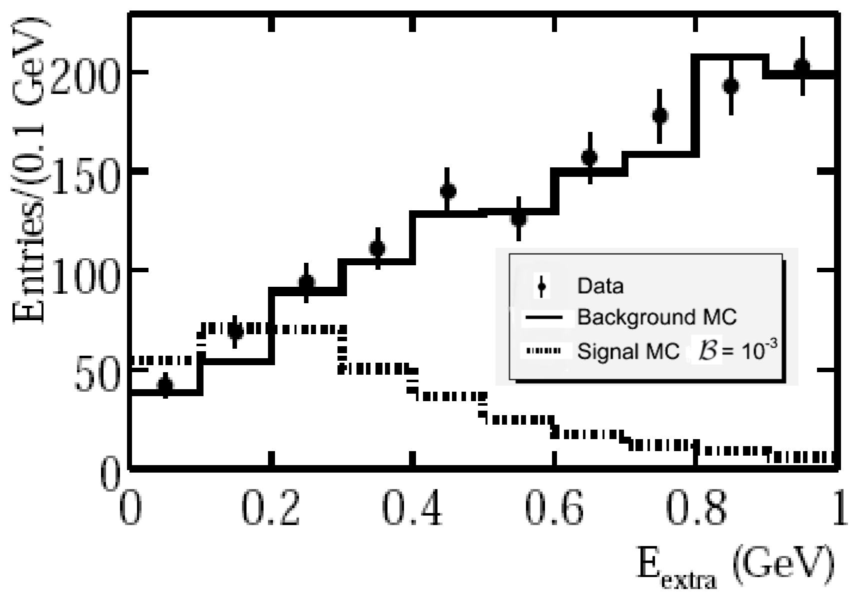


Current Results
and
Summer 2008 Updates

Semileptonic Tags

- Used 383 x 10^6 BB pairs.
- Carry out Likelihood fit to yield in four tau channels.
- $\mathcal{B}(B\to\tau\nu)$= (0.9 ± 0.6(stat) ± 0.1(syst)) x 10^{-4}.
- 90% CL UL: $\mathcal{B}(B\to\tau\nu)$< 1.7 x 10^{-4}.
- $f_B \cdot |V_{ub}| = (7.2^{+2.0}_{-2.8}(\text{stat.}) \pm 0.2(\text{syst.})) \times 10^{-4}$

τ decay mode	Expected background events	Observed events in on-resonance data
$\tau^+ \to e^+\nu\overline{\nu}$	44.3 ± 5.2	59
$\tau^+ \to \mu^+\nu\overline{\nu}$	39.8 ± 4.4	43
$\tau^+ \to \pi^+\overline{\nu}$	120.3 ± 10.2	125
$\tau^+ \to \pi^+\pi^0\overline{\nu}$	17.3 ± 3.3	18
All modes	221.7 ± 12.7	245

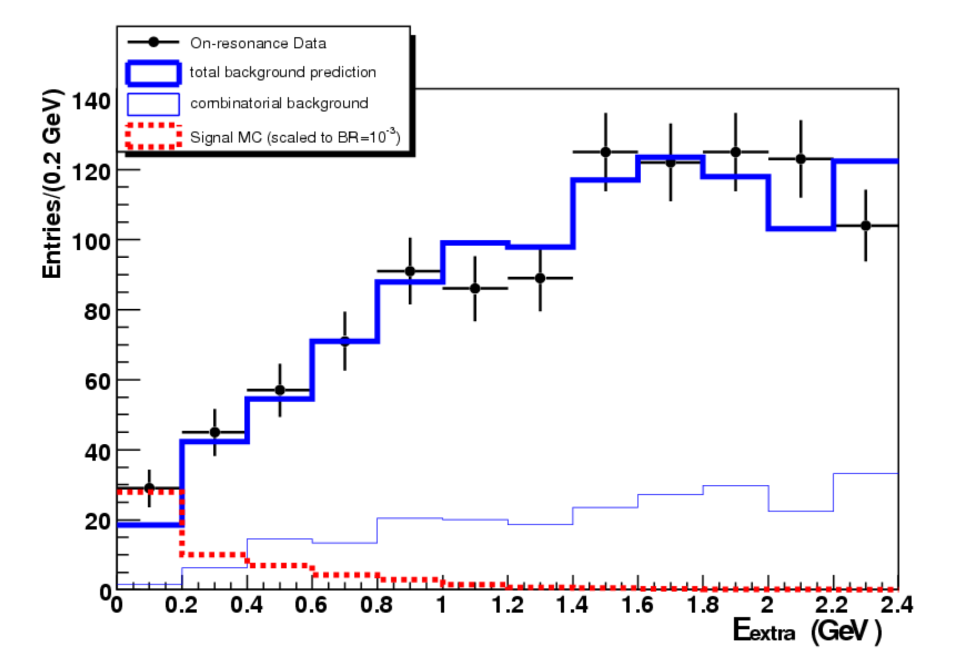


Phys.Rev.D76:052002, 2007
arXiv:0705.1820 [hep-ex]

Hadronic Tags

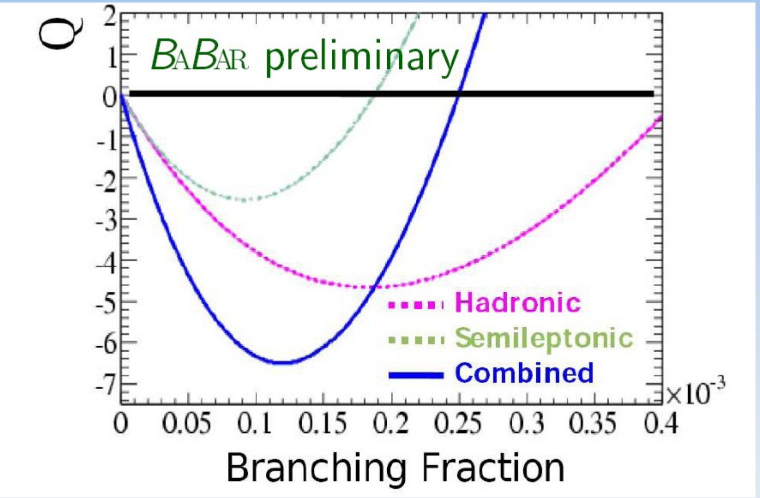
- Also uses 383×10^6 BB pairs.
- Measured Branching fraction:
- $\mathcal{B}(B^+ \to \tau^+ \nu) = 1.8^{+1.0}_{-0.9}(\text{stat.+bkg}) \pm 0.3(\text{syst.})) \times 10^{-4}$.
- 90% CL Upper Limit:
  $\mathcal{B}(B \to \tau\nu) < 3.4 \times 10^{-4}$.
- $\mathcal{B}$ also calculated from likelihood ratio fit to the individual tau channel yields.
- $f_B \cdot |V_{ub}| = (10.1^{+2.8}_{-2.5}(\text{stat.}) \pm 0.8(\text{syst.})) \times 10^{-4}$ GeV

τ decay mode	Expected background	Observed
$\tau^+ \to e^+\nu\overline{\nu}$	1.47 ± 1.37	4
$\tau^+ \to \mu^+\nu\overline{\nu}$	1.78 ± 0.97	5
$\tau^+ \to \pi^+\overline{\nu}$	6.79 ± 2.11	10
$\tau^+ \to \pi^+\pi^0\overline{\nu}$	4.23 ± 1.39	5
All modes	14.27 ± 3.03	24



Phys.Rev.D77:011107, 2008
arXiv:0708.2260 [hep-ex]

Combined Result



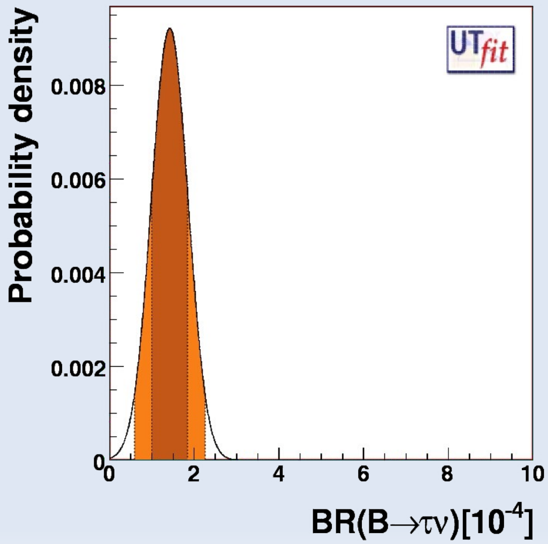
- Combine semileptonic and hadronic results.
- Statistically independent.
- Extend likelihood ratio technique used in both to determine combined result:

$$\mathcal{B}(B^+ \to \tau^+ \nu) = (1.20^{+0.40}_{-0.38}(\text{stat.})^{+0.29}_{-0.30}(\text{bkg syst.}) \pm 0.22(\text{syst.})) \times 10^{-4},$$

2.6σ significance.

Belle result: (Hadronic tags) $\mathcal{B} = (1.79^{+0.56}_{-0.49}\ {}^{+0.39}_{-0.46}) \times 10^{-4}$ (3.5σ)

PRL 97, 251802 (2006)



- UTfit combined value:
- $\mathcal{B}(B \to \tau\nu) = (1.41 \pm 0.43) \times 10^{-4}$.

Summer 2008 Updates

- BaBar update to semileptonic tagged analysis:
  - $\mathcal{B}(B^+ \rightarrow \tau^+\nu_\tau) = (1.8 \pm 0.8 \pm 0.1) \times 10^{-4}$
  - Significance: 2.4σ

Mode	Expected Background (N_{BG})	Observed Events (N_{obs})
$\tau^+ \rightarrow e^+\nu_e\overline{\nu}_\tau$	91 ± 13	148
$\tau^+ \rightarrow \mu^+\nu_\mu\overline{\nu}_\tau$	137 ± 13	148
$\tau^+ \rightarrow \pi^+\overline{\nu}_\tau$	233 ± 19	243
$\tau^+ \rightarrow \pi^+\pi^0\overline{\nu}_\tau$	59 ± 9	71

- Updated combined result:

$$\mathcal{B}(B^+ \rightarrow \tau^+\nu_\tau) = (1.8 \pm 0.6) \times 10^{-4}$$ (3.2σ)

arXiv:0809.4027v1 [hep-ex]

- New Belle semileptonic tagged analysis:

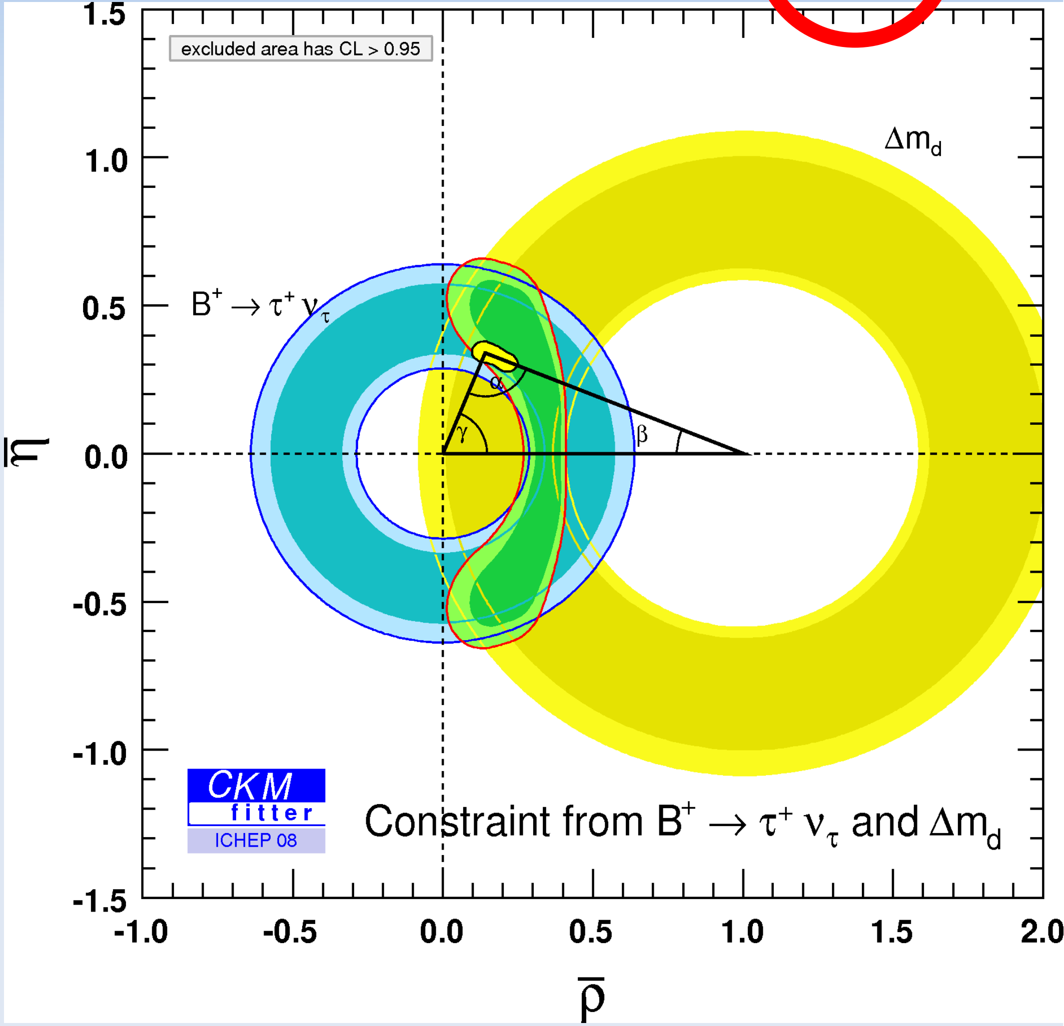
$$\mathcal{B}(B^- \rightarrow \tau^-\overline{\nu}_\tau) = (1.65^{+0.38}_{-0.37}(\mathrm{stat})^{+0.35}_{-0.37}(\mathrm{syst})) \times 10^{-4}$$ (3.8σ)

arXiv:0809.3834v1 [hep-ex]

Constraint on Unitarity Triangle

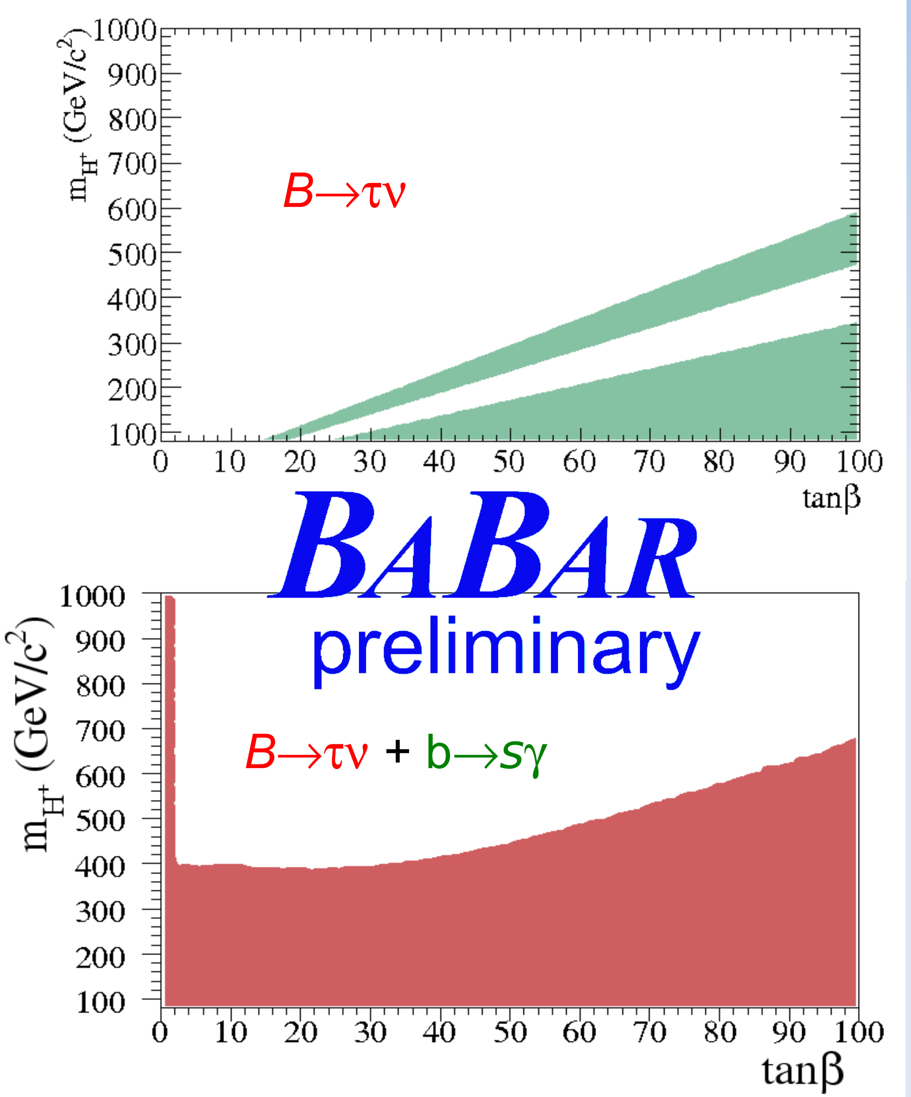
$$\mathcal{B}(B^- \to \ell^- \bar{\nu}) = \frac{G_F^2 m_B}{8\pi} m_l^2 \left(1 - \frac{m_l^2}{m_B^2}\right)^2 f_B^2 |V_{ub}|^2 \tau_B$$

- Combine $B \to \tau\nu$ with Δm_d measurements to constrain CKM ratio $|V_{ub}|/|V_{td}|$.
- f_B cancels – least well known value.
- Shown as a graphical constraint on Unitarity Triangle.
- Consistent with SM.



Implications for New Physics

- Exclusions in m_H – tan β plane.
  - m_H – Charged Higgs mass.
  - tan β – ratio of v.e.v. of 2HD.
- Plots shown for region above direct search limit from LEP.
- Can be combined with measurement of $b \to s\gamma$.
- $B \to \tau\nu$ more useful at higher values of tan β.



Plots from S. Sekula

Multivariate Analysis

Multivariate Analysis

- Use a combination of many variables to select events.
- Make use of correlations between variables.
- Use combination of weakly classifying variables that could not be cut on.
    - Examples of Multivariate Classifiers include:
      Fisher Discriminant, Neural Net
      Boosted Decision Tree, Random Forest
- Increase signal efficiency and/or background rejection.

Multivariate Analysis Packages

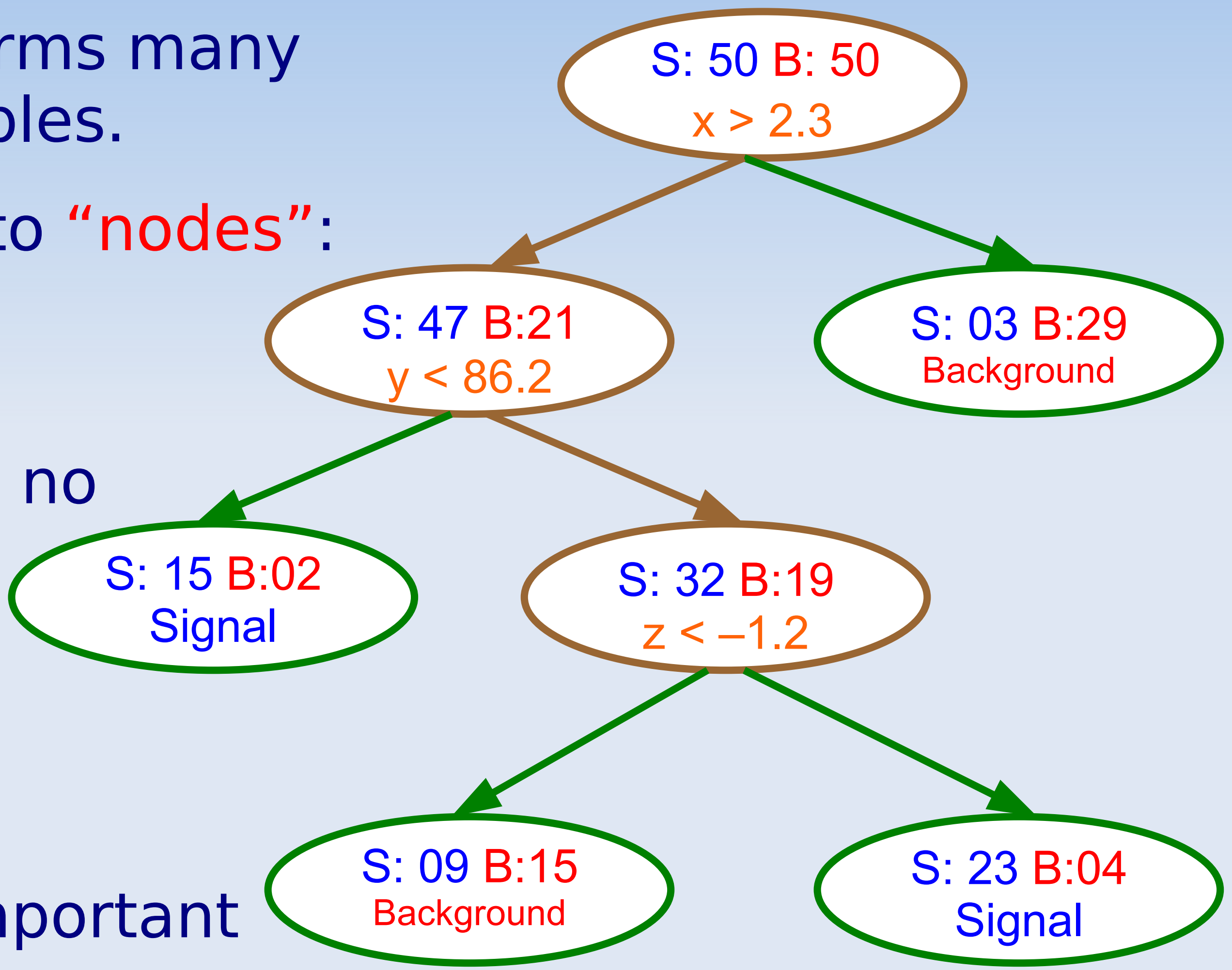
- Two packages commonly used in Particle Physics.
- TMVA Toolkit for MultiVariate Analysis:
  - http://tmva.sourceforge.net/
  - Developed mainly at CERN.
  - Incorporated in recent releases of ROOT (5.11+).
- StatPatternRecognition:
  - https://sourceforge.net/projects/statpatrec
  - Developed by Ilya Narsky (Caltech).
  - Fully compatible with ROOT.

General Strategy for MVA

- The chosen classifier must be trained.
- Three steps – divide available data (typically Monte-Carlo) into three datasets.
  - Training
  - Validation – check, and optimise training parameters.
  - Testing – realistic evaluation of performance.
- Example division of data: 50%:25%:25%.
- Separate samples reduces danger of over-training.
- Testing sample used for all performance plots shown.

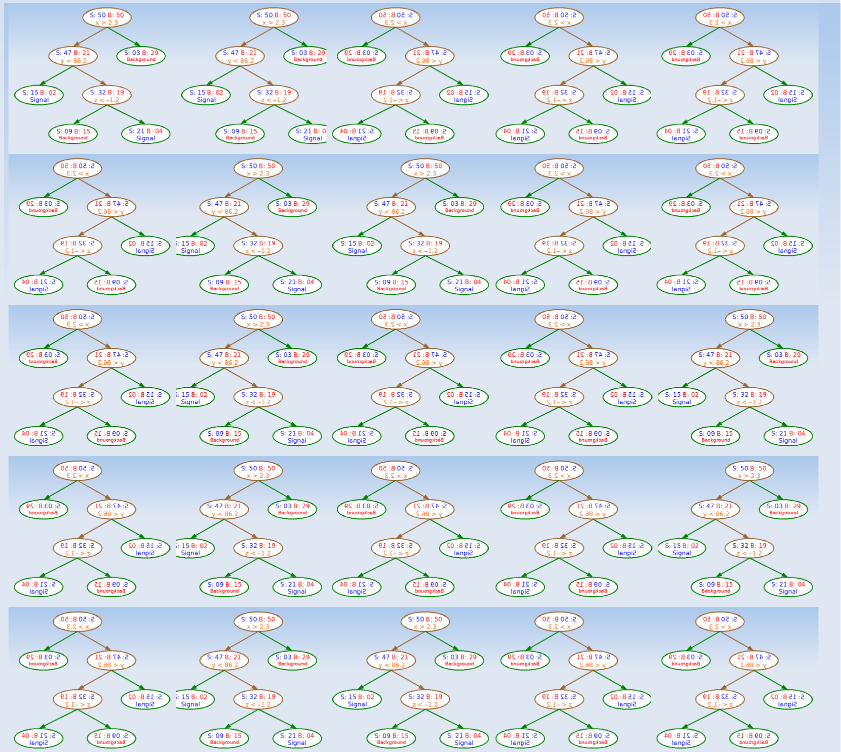
Decision Tree

- Decision tree performs many cuts on input variables.
- Separate events into "nodes": signal-like or background-like.
- Keep splitting, until no new nodes can be created.
- Terminal nodes called "leaves".
- Leaf size is most important training variable.



Boosted Decision Tree

- Boosting – over a specified number of cycles:
  - increase weight of misclassified events
  - decrease weight of correctly classified events.
- Increases predictive power.
- Boosted decision tree can no longer be easily visualised.
- Advantages:
  - Can cope with very correlated variables and useless inputs.
  - No “Curse of dimensionality”.

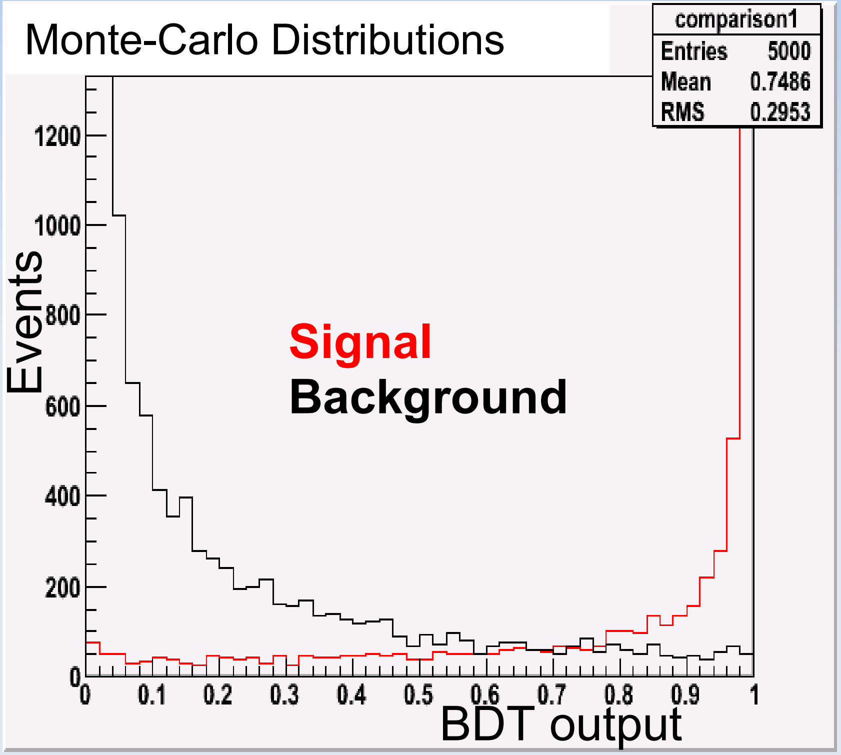


Bagging and Random Forests

- Bagging - Bootstrap AGGregatING.
- Bootstrapping – sampling with replacement.
- Train classifiers on bootstrap replicas of training data.
- Overall response is average of each classifier training.
- Bootstrapping the input dimensions (variables) as well is called a Random Forest.
  - "De-correlates" variables.
- Important training parameters are Leaf size, and number of input dimensions to sample.

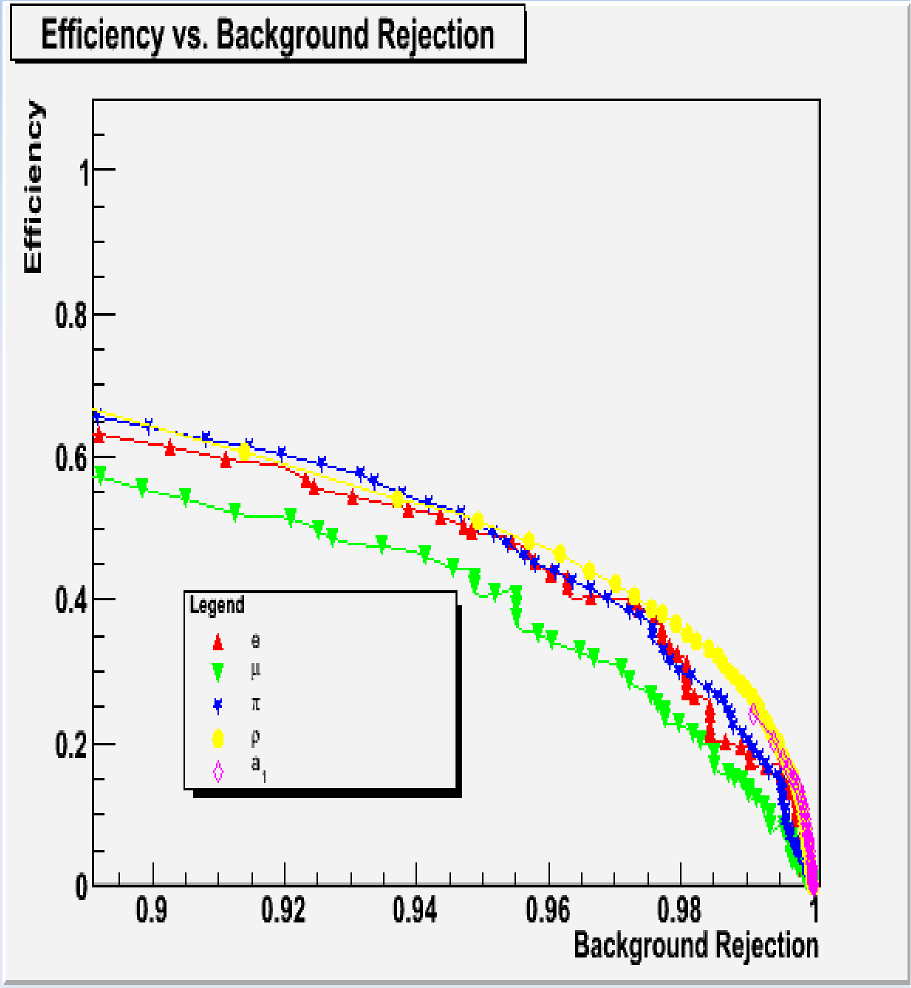
Using a Boosted Decision Tree for B→τν

- Use a BDT to classify events.
- Train for each τ mode.
- Use many weakly discriminating variables such as:
  - ρ, a_1 candidate mass,
  - Momentum of τ daughter,
  - $\cos\theta_{miss}$...
- Use 11-18 variables in training (τ mode dependent).
- E_{Extra} is not used, so it can be analysed separately.



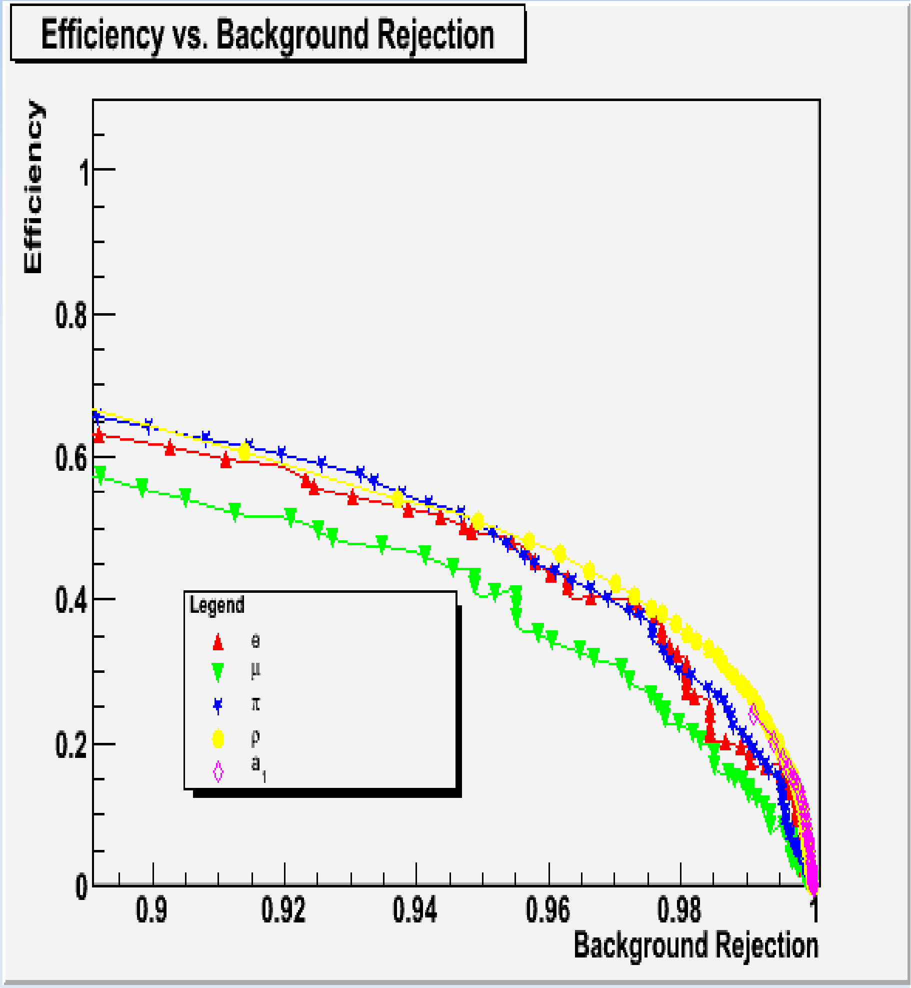
Using a Boosted Decision Tree for B→τν

- Raw signal/background distribution not most useful.
- Calculate Signal Efficiency and Background rejection for different cuts.
- Plot Signal Efficiency against Background rejection.
- Very high background rejection can be obtained:
  At cost of lower signal efficiency.



Using a Boosted Decision Tree for $B \to \tau\nu$

- Standard cuts perform very well in electron mode – very difficult to beat with MVA.
- The other τ decay modes show some promise of improvement using MVA.

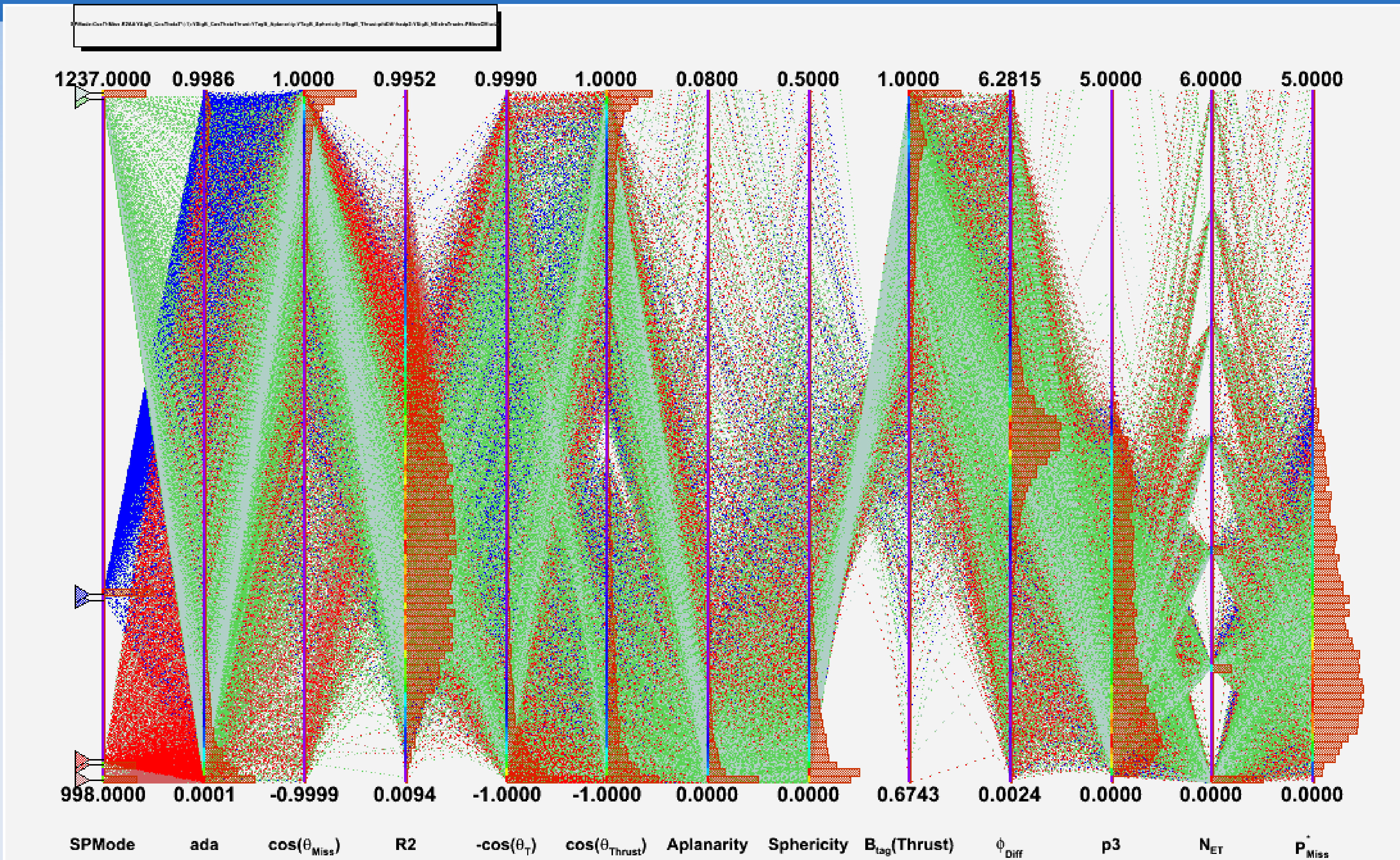


Visualisation of Parameters

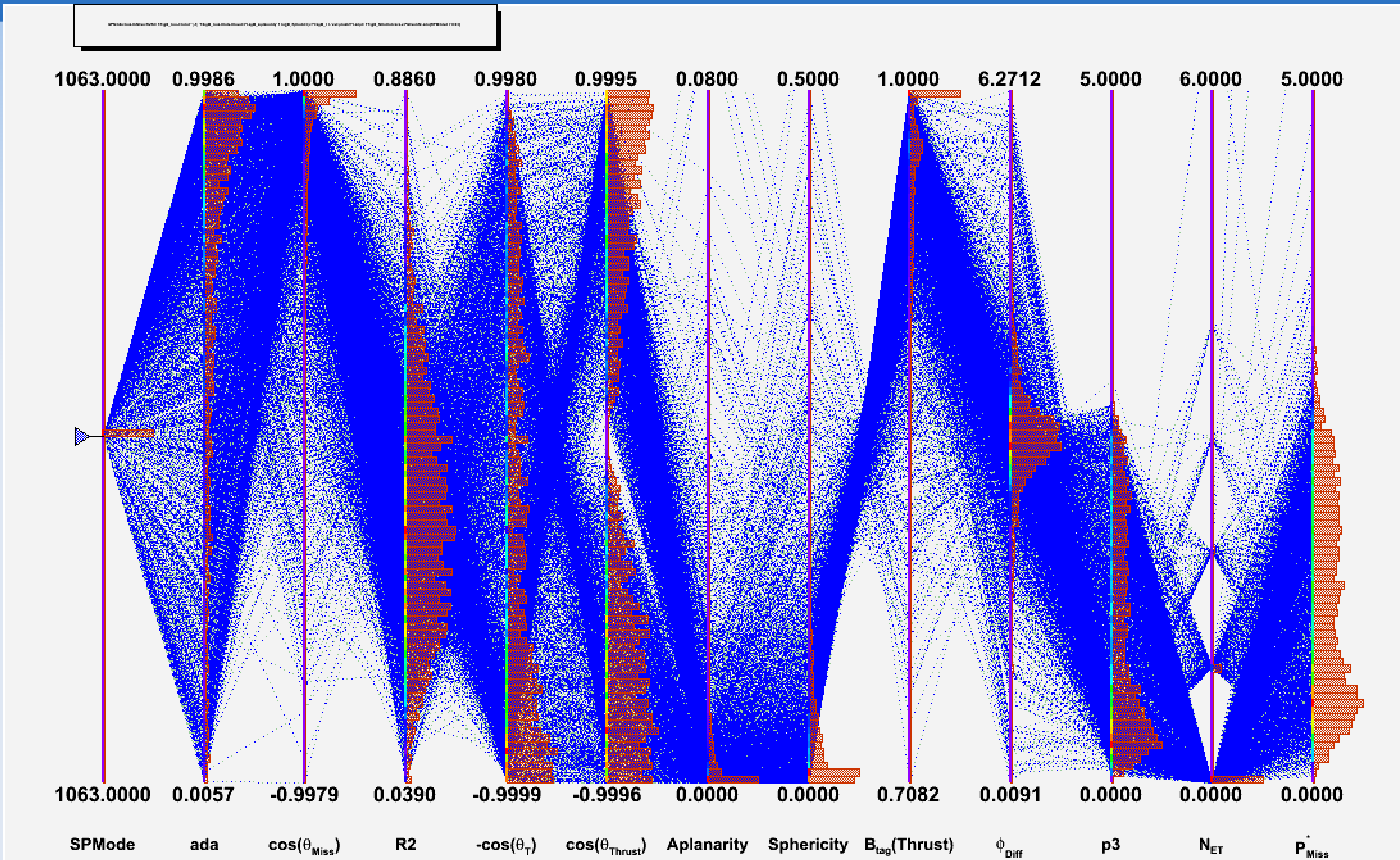
- Multi-dimensional problems are difficult to visualise.
- More dimensions → More Difficult to visualise.
- Parallel Coordinates are a visualisation method.
  - One (parallel) axis for each variable.
  - Each event is represented by a line.
- Background types represented by a different colours.
- Colour Scheme used in plots:

  Signal uds cc B^0B^0 B^+B^-
- Available in ROOT 5.17 (and above).
- Example is shown for variables for π mode.

Example for π Variables

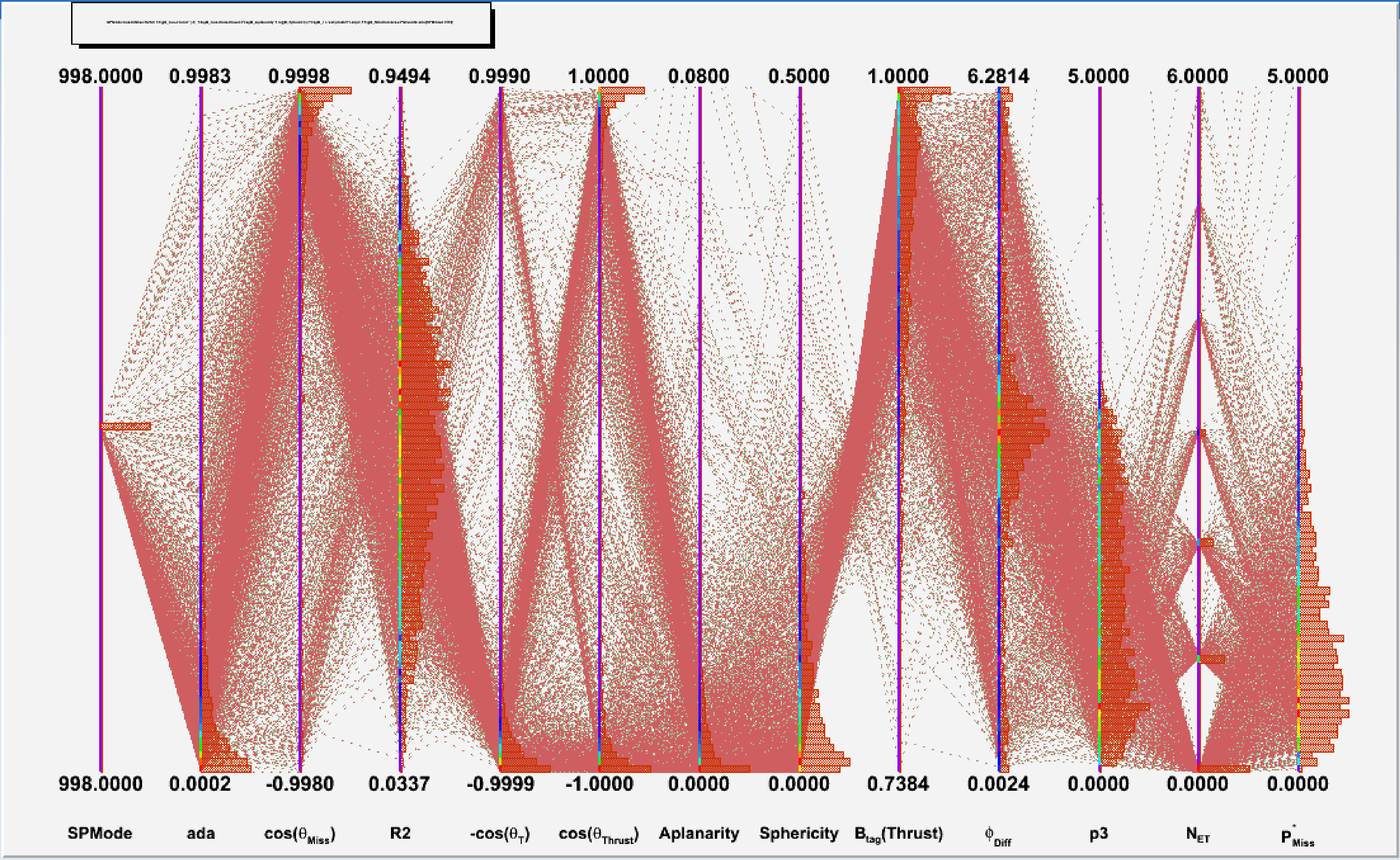


Signal Only

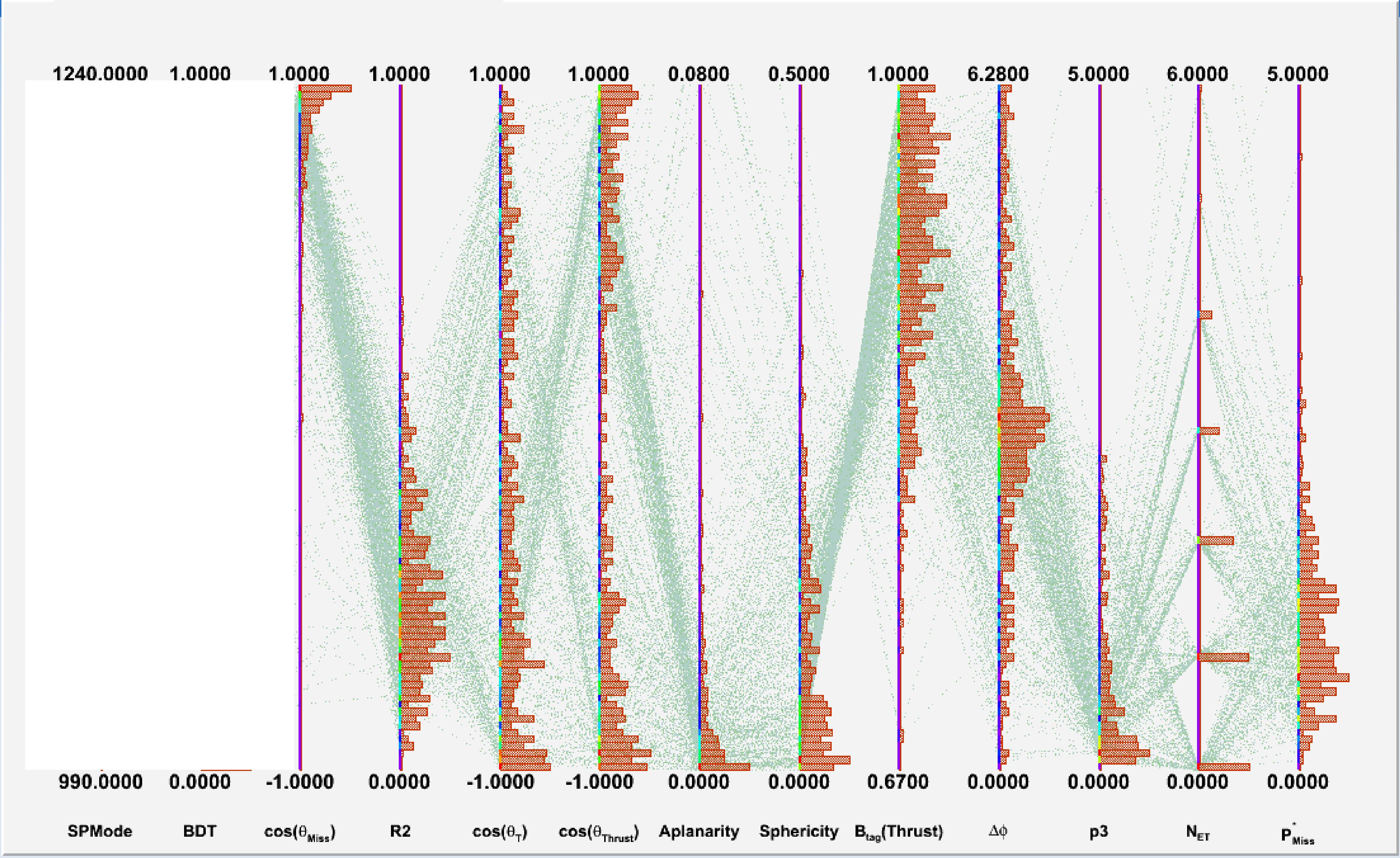


1063.0000 0.9986 1.0000 0.8860 0.9980 0.9995 0.0800 0.5000 1.0000 6.2712 5.0000 6.0000 5.0000
1063.0000 0.0057 -0.9979 0.0390 -0.9999 -0.9996 0.0000 0.0000 0.7082 0.0091 0.0000 0.0000 0.0000
SPMode
ada
$\cos(\theta_{Miss})$
R2
$-\cos(\theta_T)$
$\cos(\theta_{Thrust})$
Aplanarity
Sphericity
B_{tag}(Thrust)
ϕ_{Diff}
p3
N_{ET}
P^*_{Miss}

Light Continuum (uds) Only



All MC types



1240.0000
1.0000
1.0000
1.0000
1.0000
1.0000
0.0800
0.5000
1.0000
6.2800
5.0000
6.0000
5.0000
990.0000
0.0000
-1.0000
0.0000
-1.0000
-1.0000
0.0000
0.0000
0.6700
0.0000
0.0000
0.0000
0.0000
SPMode
BDT
cos(θMiss)
R2
cos(θT)
cos(θThrust)
Aplanarity
Sphericity
Btag(Thrust)
Δϕ
p3
NET
P*Miss

Prospects

- BaBar has collected its full dataset of Υ(4S) decays.
- The next sets of analyses carried out aim to be the definitive BaBar analyses.
- Work is ongoing to incorporate as many improvements as possible during this intense analysis period.
- $B \rightarrow \tau\nu$ will continue to be a subject of great interest at potential at the next generation of proposed B-factories: SuperB and SuperKEKB.

Summary

- The decay $B \to \tau\nu$ can be used to measure parameters unavailable to other B decays, and to constrain the Unitarity Triangle.
- Constraints on New Physics – Charged Higgs.
- Babar and Belle have seen evidence of this decay.

$$\mathcal{B}(B^+ \to \tau^+ \nu) = (1.20^{+0.40}_{-0.38}(\text{stat.})^{+0.29}_{-0.30}(\text{bkg syst.}) \pm 0.22(\text{syst.})) \times 10^{-4},$$

$$\mathcal{B} = (1.79^{+0.56}_{-0.49}\ {}^{+0.39}_{-0.46}) \times 10^{-4}$$

- Summer 2008 updates:

$$\mathcal{B}(B^+ \to \tau^+ \nu_\tau) = (1.8 \pm 0.6) \times 10^{-4}$$

$$\mathcal{B}(B^- \to \tau^- \overline{\nu}_\tau) = (1.65^{+0.38}_{-0.37}(\text{stat})^{+0.35}_{-0.37}(\text{syst})) \times 10^{-4}$$

- New methods could hopefully move this closer to a discovery.

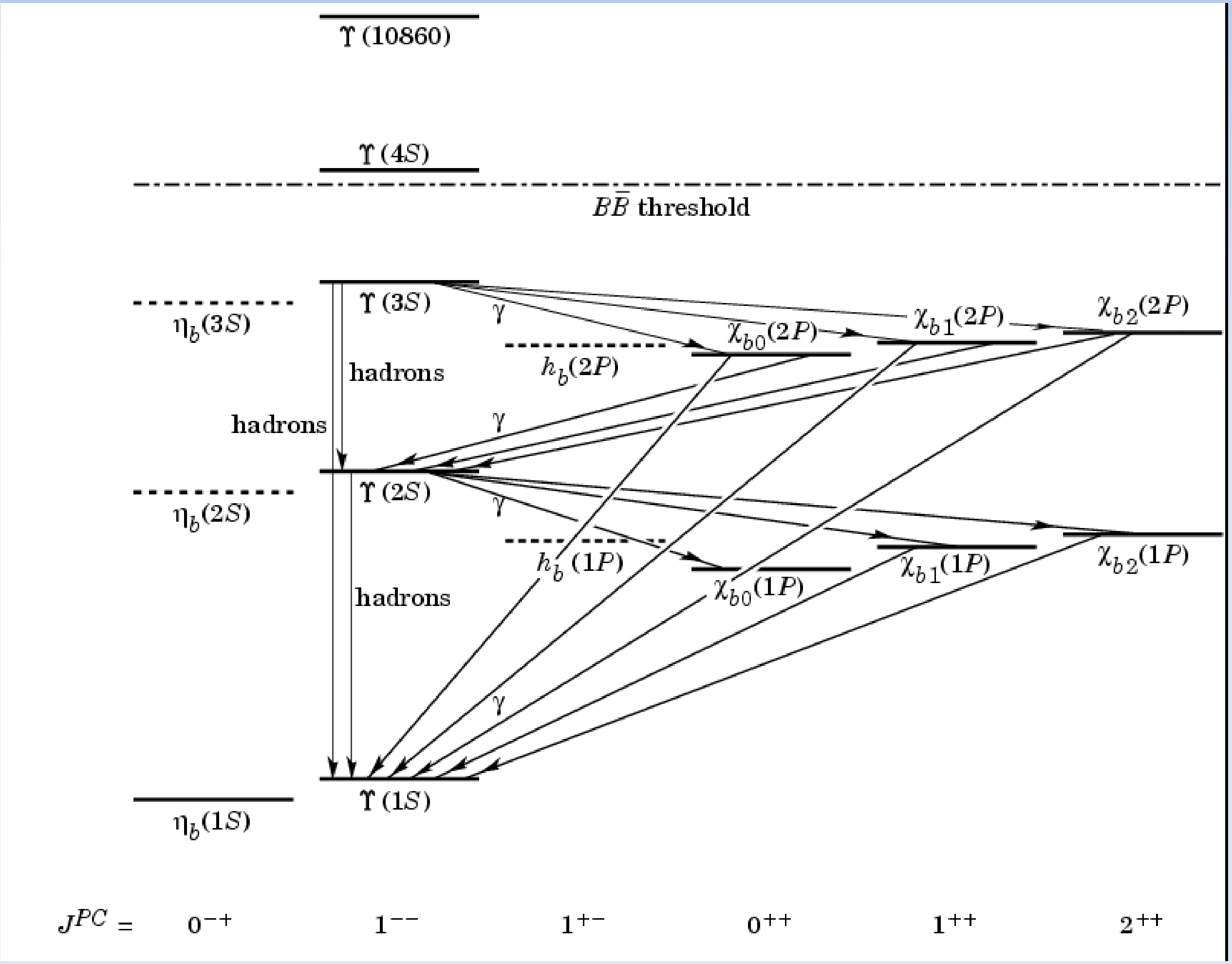
Back-Up Slides

Υ(nS) Physics

- Taken 30fb^{-1} at Υ(3S) resonance, ~90M Υ(3S) events.
- ~10× the previous largest sample.
- Taken ~15fb^{-1} at the Υ(2S) resonance, ~100M events.
- Standard Model:
  - Search for new states;
  - Bottomonium Spectroscopy.
- Beyond the Standard Model:
  - Low mass Higgs.
  - Lepton Flavour violation.
  - Low mass Dark Matter

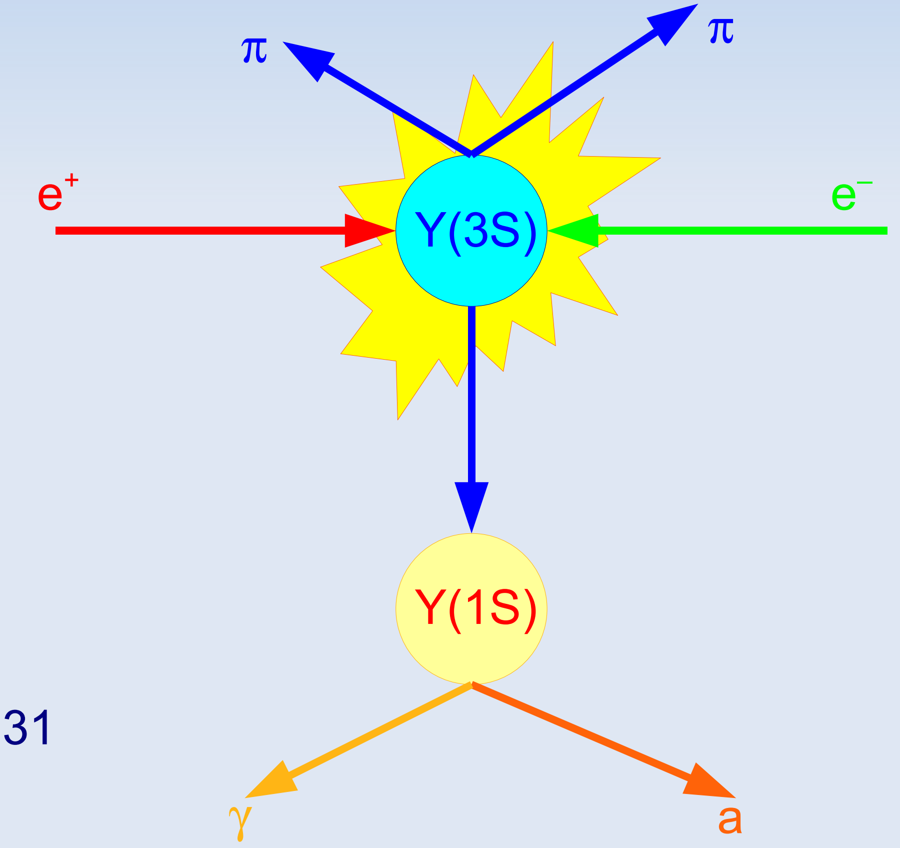
Υ(nS) Physics - Bottomonium

- Solid lines: Discovered.
- Dashed lines: Predicted.
- Most predicted states accessible.
- Known states have few measured branching fractions.



Y(nS) *Physics - Light Higgs*

- Recent work in NMSSM interested in low mass CP-odd Higgs (a).
- Avoids direct LEP constraints.
- Would decay to $\tau\tau$, light hadrons or charmed hadrons depending on mass.



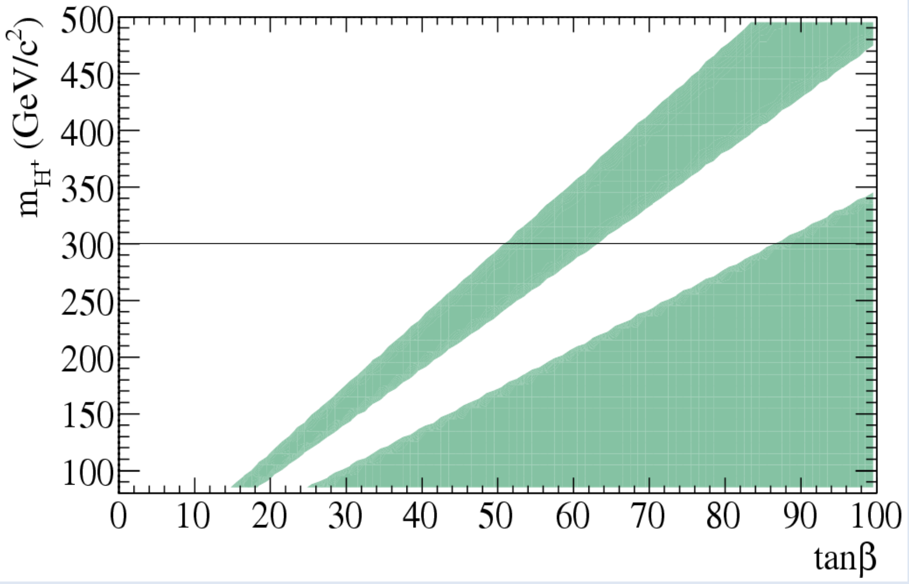
- Hiller, hep-ph/0404220
- Dermisek, Gunion, McElrath, hep-ph/0612031

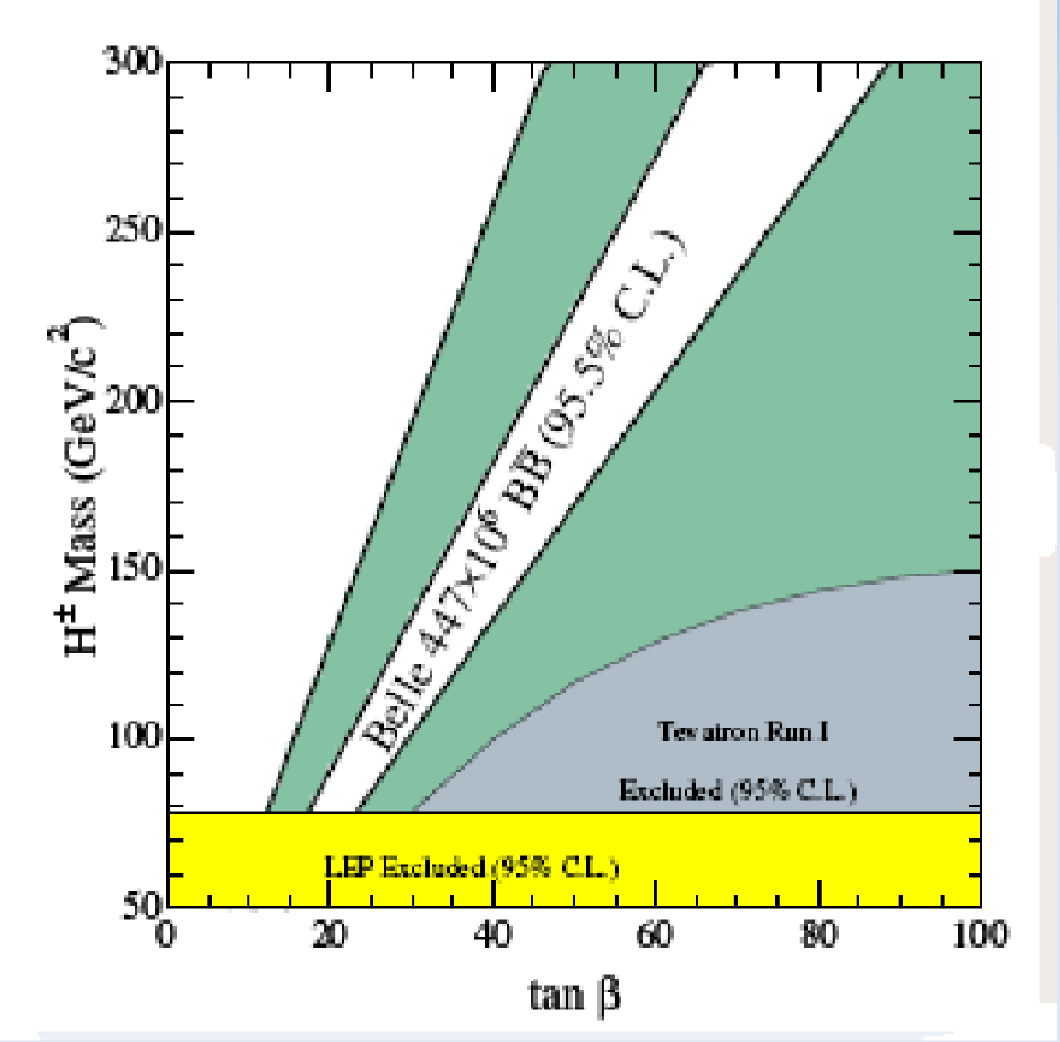
Υ(nS) Physics - Leptons

- Measure leptonic decays of Υ(nS).
- Different rates for e.g. $\mathcal{B}(\Upsilon(nS)) \rightarrow \tau^{+}\tau^{-}$ and $\mathcal{B}(\Upsilon(nS)) \rightarrow \mu^{+}\mu^{-}$ would be departure from Lepton Universality.
- Could be caused by low mass Higgs.
- Also search for lepton flavour violation, e.g. $\mathcal{B}(\Upsilon(3S)) \rightarrow \tau^{+}\mu^{-}$.

Belle $B \to \tau\nu$

- Comparison of BaBar and Belle exclusions from $B \to \tau\nu$.





Unitarity Triangle

- Weak eigenstates ≠ Flavour eigenstates (Strong, EM).
- Two generations of quarks described by Cabibbo matrix:

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}$$

- CKM matrix describes quark mixing with 3 generations.
- Apply Unitary condition $V^{\dagger}V = I$.

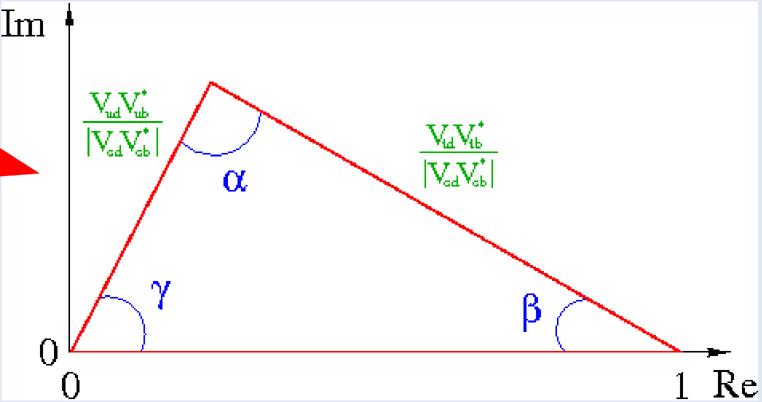
$$V_{\mathrm{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}$$

- 9 equations, e.g.

$$V_{ud}V_{ud}^{*} + V_{us}V_{us}^{*} + V_{ub}V_{ub}^{*} = 1.$$

$$V_{ud}V_{ub}^{*} + V_{cd}V_{cb}^{*} + V_{td}V_{tb}^{*} = 0$$

- Gives Unitarity Triangle.
- Measure angles α, β, and γ and lengths of sides.



MVA method comparison

- Summary Slide by Ilya Narsky.

	Neural Net	RBF	SVM	Trees (CART)	Boosted and bagged trees	MARS	k-NN	VAB
Predictive power	good	good	good	poor	good	fair	good	good
Ability to deal with irrelevant inputs	poor	poor	poor	good	good	good	poor	fair
Interpretability	poor	poor	poor	good	poor	fair	poor	poor
Curse of dimensionality	good	poor	fair	good	good	good	poor	good
Computational scalability with adding new dimensions	poor	poor	poor	good	good	fair	poor	horrible
Training stability	poor	poor	poor	fair	good	good	○	good
Response time	good	fair	fair	good	fair	fair	poor	poor

● good ● fair ● poor ● horrible

- Part of talk available on SPR homepage: http://www.hep.caltech.edu/~narsky/spr.html