CP Violation Physics

Battling With Backgrounds

Lecture # 18

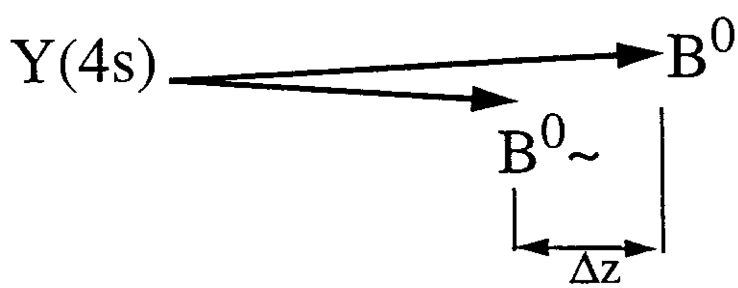
December 12, 1997

Theresa Champion

Setting the Scene

BaBar TDR (first page) :

- reconstruct decays of B^0 mesons into a wide range of exclusive final states with high efficiency and low background
- tag flavor of other B^0 meson with high efficiency and purity
- measure relative decay time of the two B^0 mesons



- there are many decay modes to consider
- the branching ratios tend to be small ($\sim 10^{-5}$)

Introduction

- types of background :
  main features
  relative importance

- information used to fight background

- methods used to optimise cuts

- effect on statistical reach

Illustration using an example from BaBar :

$$B^0 \to \rho\pi \to 3\pi \text{ final state}$$

Know Your Enemy.....

Backgrounds from :

- **qq~ continuum**

$e^+e^- \longrightarrow \Upsilon(4s) \longrightarrow B^0B^0\sim$

$\quad\hookrightarrow qq\sim \qquad\qquad \hookrightarrow B^+B^-$

...fight with event shape variables

- **combinatorics**

wrong combinations of particles which satisfy signal properties

...fight with kinematics/PID

- **physics channels**

eg. $B^0 \to K\pi\pi$ background to $B^0 \to \pi\pi\pi$

...fight with PID

- **machine/beam/external sources**

synchrotron radiation γ

scattered photons γ

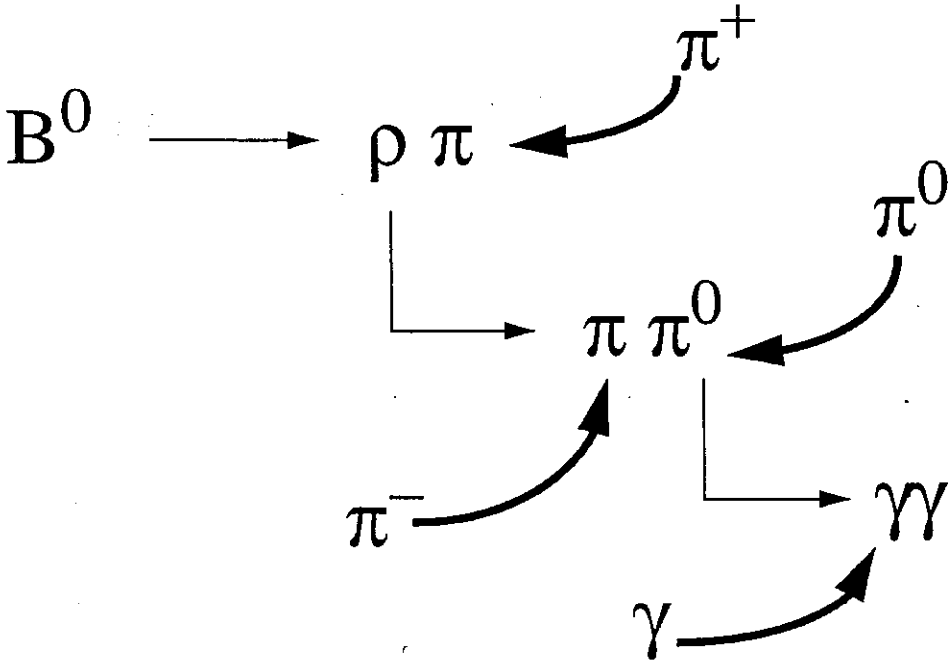
beam-gas interactions p, n, π^0, π^{+-}

radiative Bhabhas $e^+e^-\gamma$

cosmic rays $\mu^+\mu^-$

...fight with design/trigger/shielding

Combinatorial Background effects



Possible wrong combinations :

- the correct $\pi^+\pi^-$, but the wrong (real) π^0
- the correct $\pi^+\pi^-$, but one wrong γ in the π^0
- the wrong π^+ or π^-
- (for inclusive decays) the wrong assignment of the charged π to the ρ :

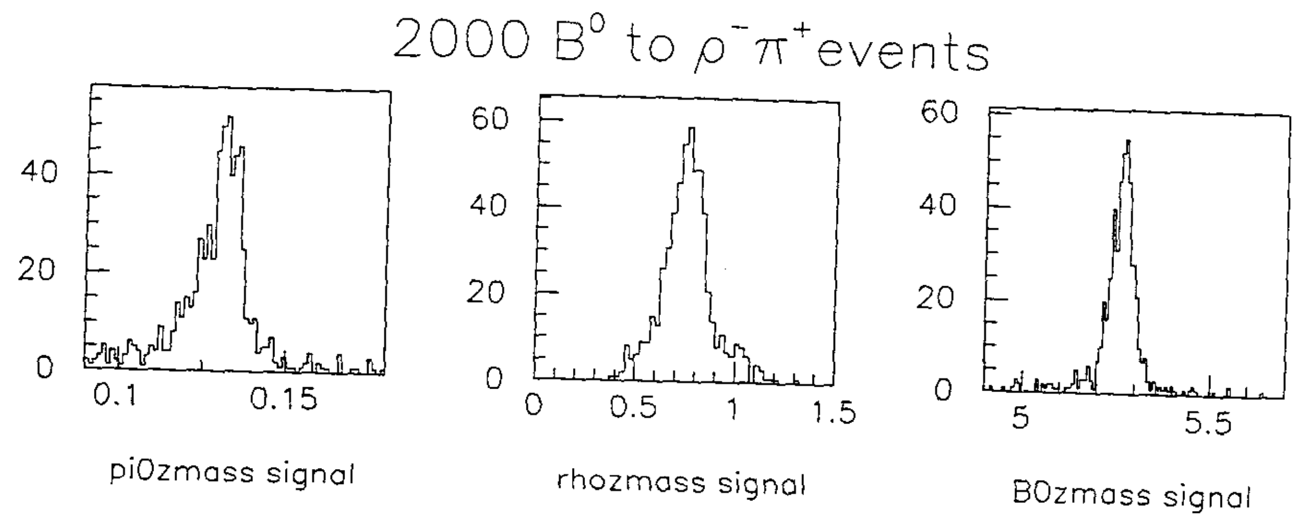
$B^0 \to \rho^+\pi^- \to (\pi^+\pi^0)\pi^-$
$B^0 \to \rho^-\pi^+ \to (\pi^-\pi^0)\pi^+$

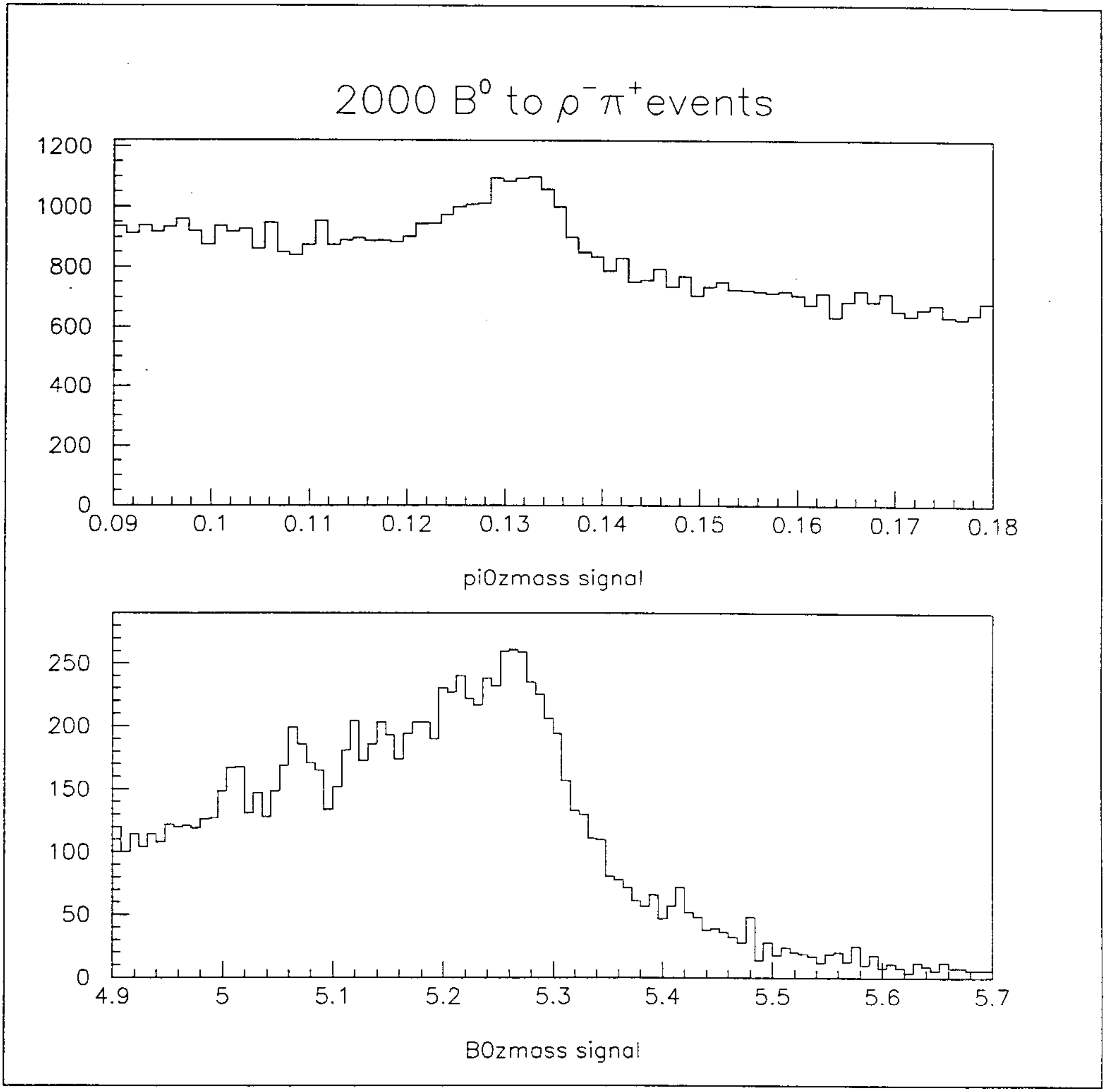
B^0 Candidate Reconstruction for the $\rho^+\pi^-$ Channel

$$B^0 \longrightarrow \rho^+\pi^-$$

$$\quad\hookrightarrow \pi^+\pi^0$$

$$\qquad\hookrightarrow \gamma\gamma$$



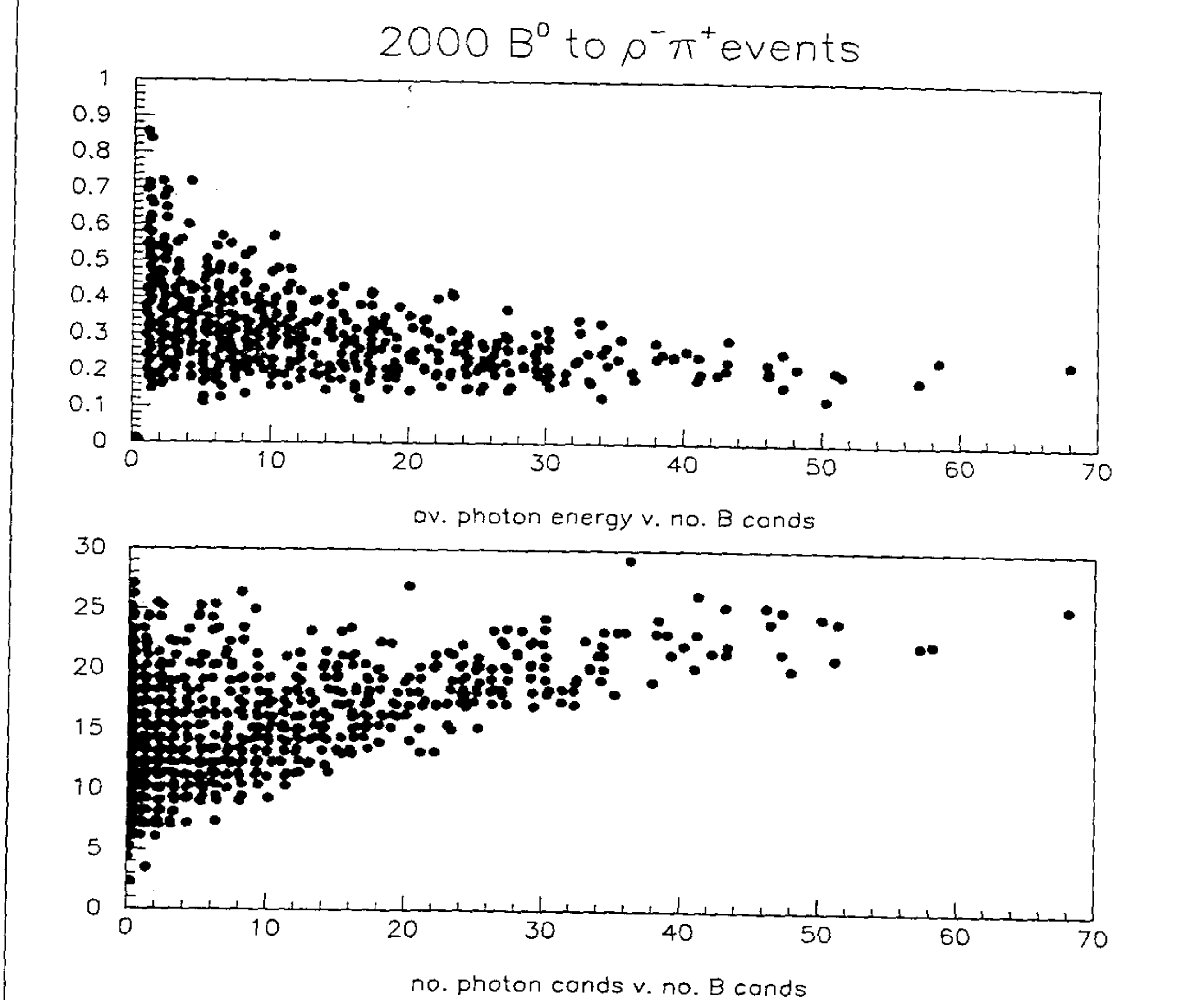


2000 B^0 to $\rho^-\pi^+$events
1200
1000
800
600
400
200
0
0.09
0.1
0.11
0.12
0.13
0.14
0.15
0.16
0.17
0.18
pi0zmass signal
250
200
150
100
50
0
4.9
5
5.1
5.2
5.3
5.4
5.5
5.6
5.7
B0zmass signal

Multiple Candidates per Event

A large amount of combinatorial background is concentrated in relatively few events

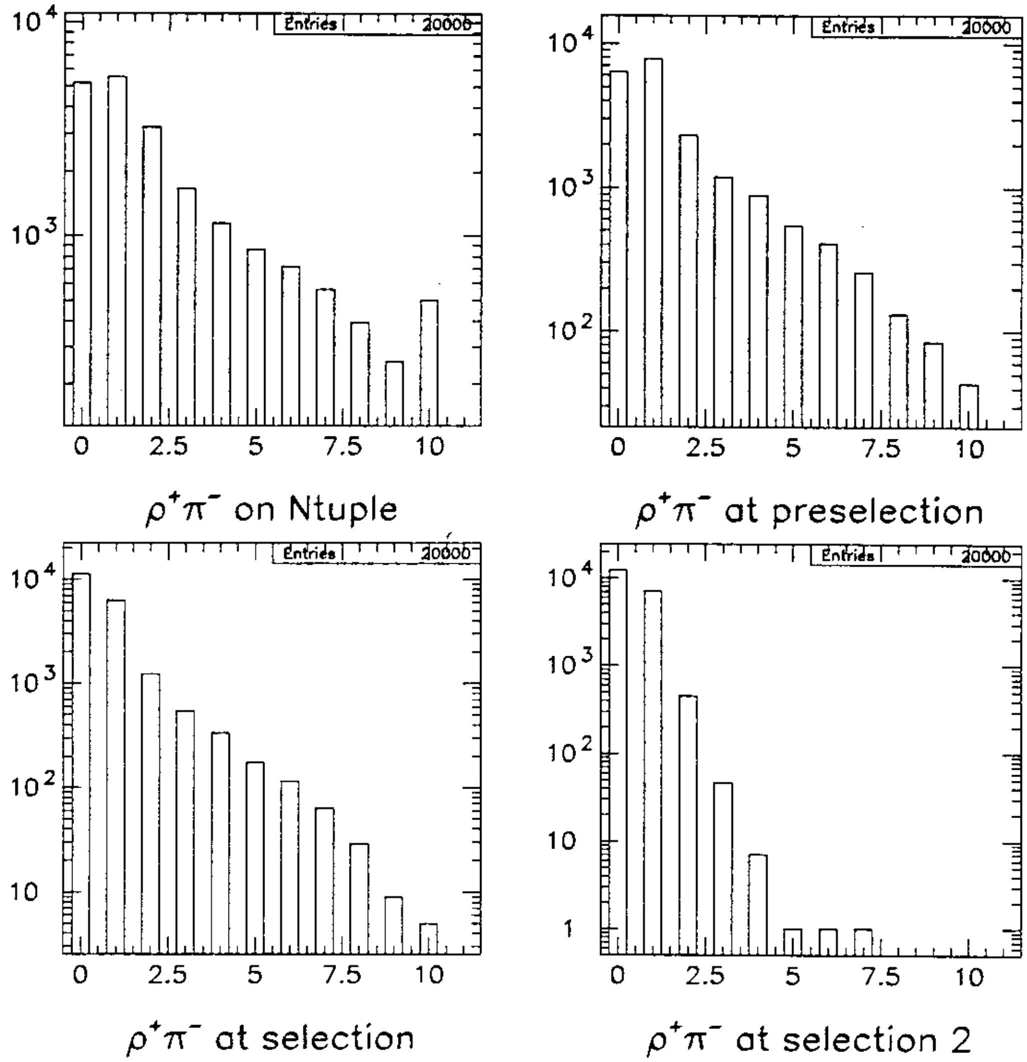
Need a "candidate quality" method for selecting the best candidate in each event
(a study has yielded suppression factors of 77% and 68% on u/d/s and c continuum respectively).



Reducing Combinatorial Background

G.Vasseur

Number of $\rho\pi$ candidates



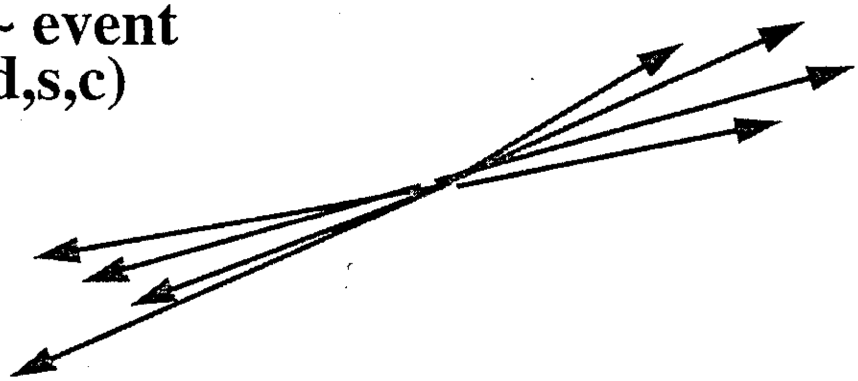
Ntuple : $5.0 < M_B < 5.6$ GeV/c^2 , $P^*_B < 1.0$ GeV/c
Preselection : $5.17 < M_B < 5.37$ GeV/c^2, $0.12 < P^*_B < 0.52$ GeV/c
Selection : cut on the Neural Network output (sg/bg(qq~)) to get a rejection of 10^5 on continuum (with no discrimination against combinatorial).
Selection 2 : cut on the two Neural Network outputs (sg/bg(qq~) and sg/bg(comb)) to get a rejection of 10^5 on continuum.

Continuum Background

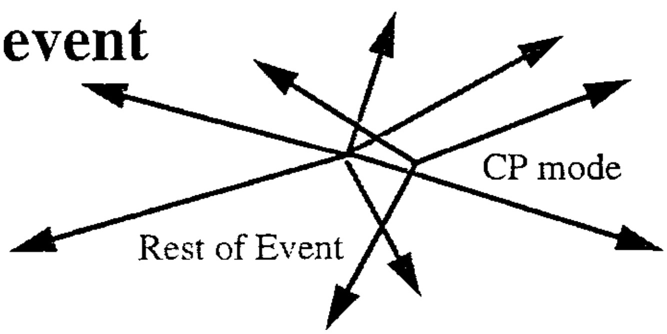
use event shape variables for discrimination

<u>Upsilon cms</u>

**qq~ event
(u,d,s,c)**

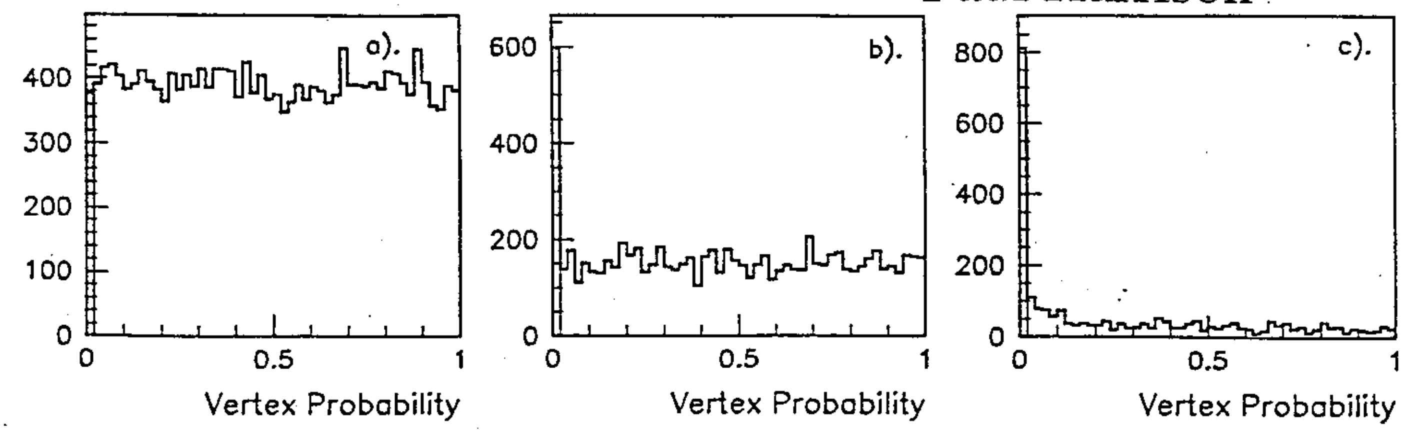


B^{0}B^0~ event

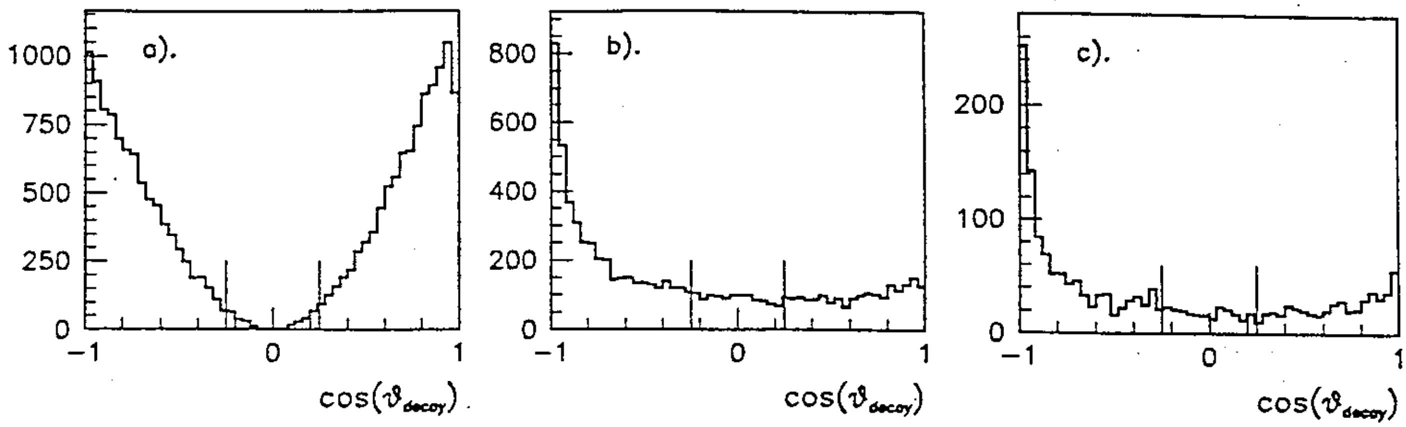


Distributions demonstrating the distinction between light-quark continuum & cc~ events

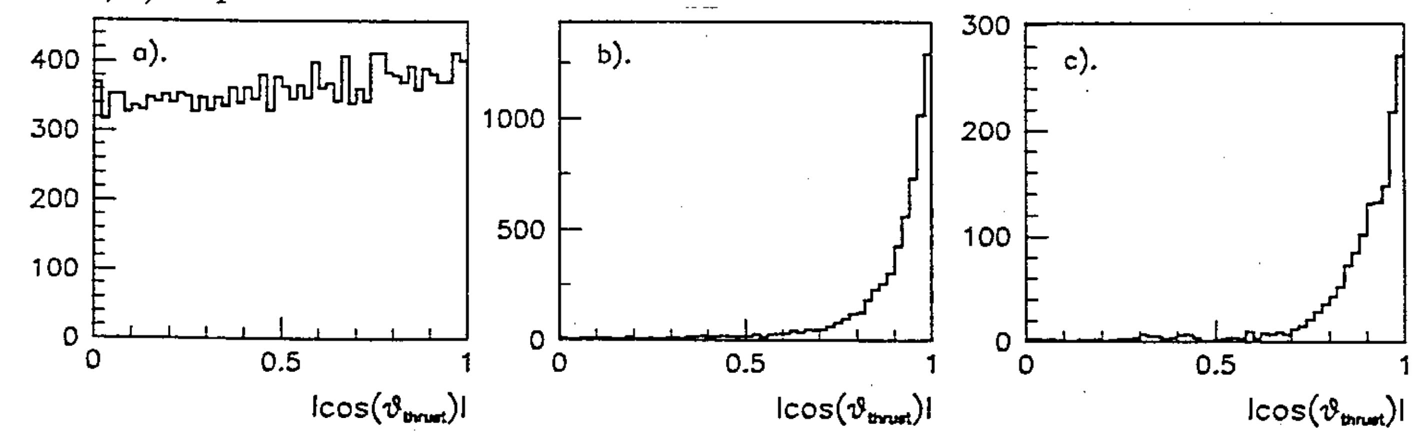
Paul Harrison



Vertex χ^2 probabilities for a) $B^0 \to \rho^{\pm}\pi^{\mp}$, b) light-quark continuum events, c) $c\bar{c}$ pair events.



Distributions of $\cos\theta_{decay}$ for a) $B^0 \to \rho^{\pm}\pi^{\mp}$, b) light-quark continuum events, c) $c\bar{c}$ pair events.



$\cos\theta_{thr}$, distributions in a) $B^0 \to \rho^{\pm}\pi^{\mp}$, b) light-quark continuum events c) $c\bar{c}$ pair events.

Topological Variables

- **Thrust**
- **Sphericity**
- **Aplanarity**
- **Fox-Wolfram Moments**
- **Cones (CLEO)**

considerations :

- which frame of reference?
  usually Upsilon cms

- which subset of the event?
  global
  CP mode
  rest of event (the tagging mode)
  only charged tracks

- **Thrust**

The thrust axis of an event is the vector **v**, such that the sum of the scalar products of normalised **v** with the particle momenta has a maximum value.
Thrust, T, is related to the maximum value by,

$$T = max\left[\frac{\sum_i \left|\frac{v}{|v|} \bullet p_i\right|}{\sum_i |p_i|}\right]$$

highly directional event -> T ~ 1
isotropic event => T = 0.5

In a typical background event for a two-body decay, the decay products of the B candidate will each lie in one of the two jets, and will therefore be approximately back-to-back. Thus the decay axis of the B candidate will be roughly colinear with the thrust axis for the rest of the event. Whereas, for a true signal event, the B decay axis will be uncorrelated with the thrust axis of the rest of the event, which in that case comes from the decay of the other B meson.

- **Sphericity**

$$S^{\alpha\beta} = \frac{\sum_i p^{\alpha} p^{\beta}}{\sum_i |p_i|^2}$$

$\alpha, \beta = 1,2,3$ correspond to x, y, z components

Sphericity is defined as,

$$S = 3/2\ (\lambda_2 + \lambda_3)$$

where λ_2 and λ_3 are the two smaller eigenvalues of the diagonalized sphericity tensor.

The Sphericity Axis is defined by the eigenvector $\mathbf{v_1}$, corresponding to the largest eigenvalue λ_1

Sphericity is a measure of the summed $\mathbf{p_T}^2$ with respect to the event axis

- interval [0,1]
- highly directional events => low sphericity
- isotropic events sphericity = 1.

- **Aplanarity**

Aplanarity is a measure of the transverse momentum component out of the event plane. It is related to the largest eigenvalue of the sphericity tensor by,

$A = 3/2\ \lambda_3$

interval [0, 0.5]

Take the cosine of the angle between the sphericity axis of the rest of the event and :

- each B decay product
- the B direction
- the sphericity axis of the B decay product system

- **Fox-Wolfram Moments**

$$H_l = \sum_{i,j} \frac{|p_i||p_j|}{E_{vis}^2} P_l(\cos\theta_{ij})$$

$\mathbf{p}_{i,j}$ are the particle momenta
θ_{ij} is the opening angle between particles i and j
P_l are the Legendre polynomials
E_{vis} is the total visible energy of the event

Neglecting particle masses: $H_0 = 1$

If momentum is balanced: $H_1 = 0$

For 2-jet events: $H_l \sim 1$, l-even, $H_l \sim 0$, l-odd

Take the ratio of Fox-Wolfram 2nd to 0th moments for :

- the whole event
- the rest of the event

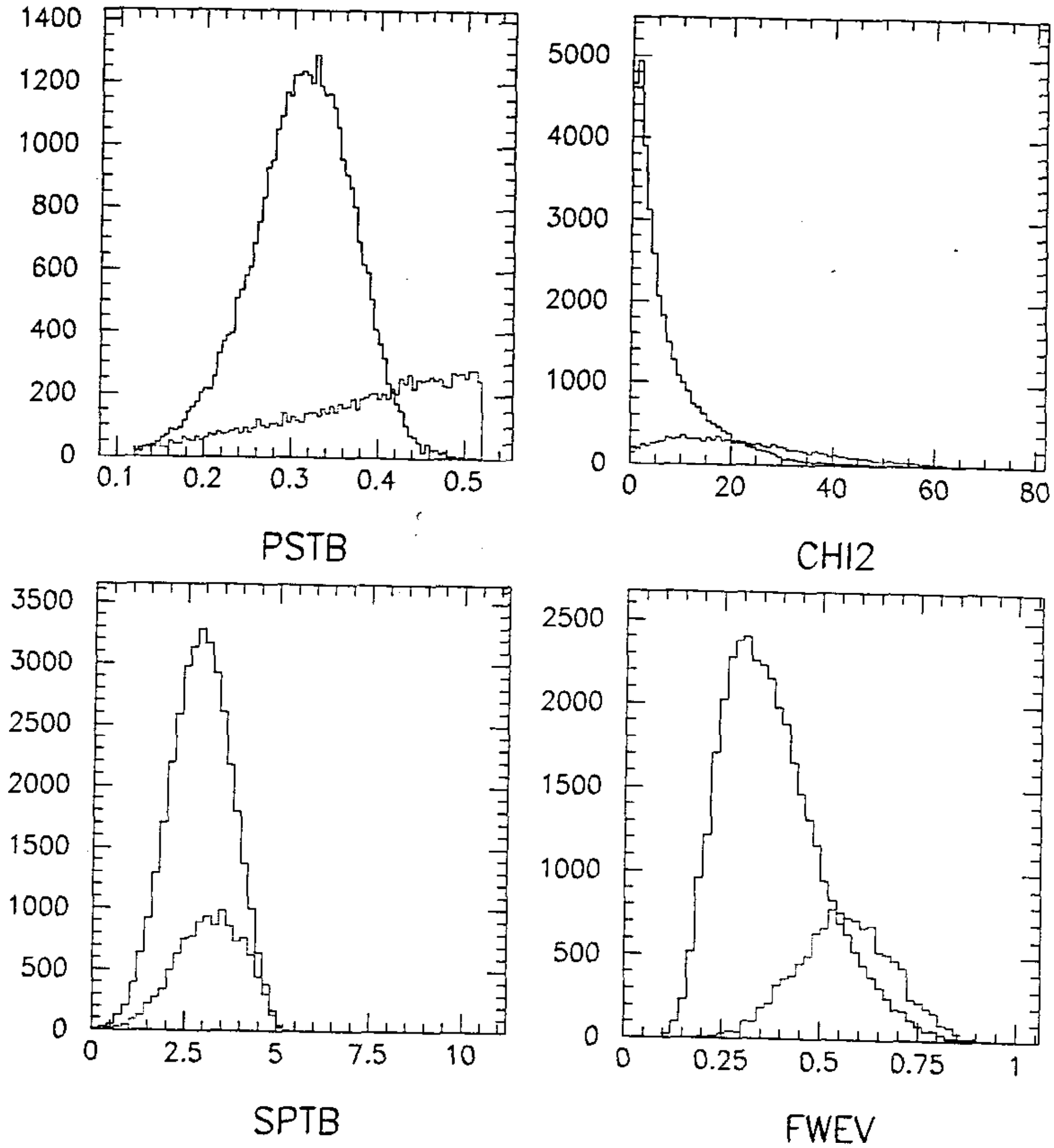
Other Discriminating Variables

- B mass
- B momentum in the $\Upsilon(4s)$ cms
- mass χ^2, formed from summing over residuals of masses of B and its decay products
- angle between the B decay product directions in $\Upsilon(4s)$ cms
- sum of transverse momenta wrt the B direction of the rest of the event
- difference between the B decay product momenta in the $\Upsilon(4s)$ cms
- sum of the magnitudes of the momenta of the B decay products in the $\Upsilon(4s)$ cms
- cosine of the angle of each B decay product momentum in the B cms wrt the B direction

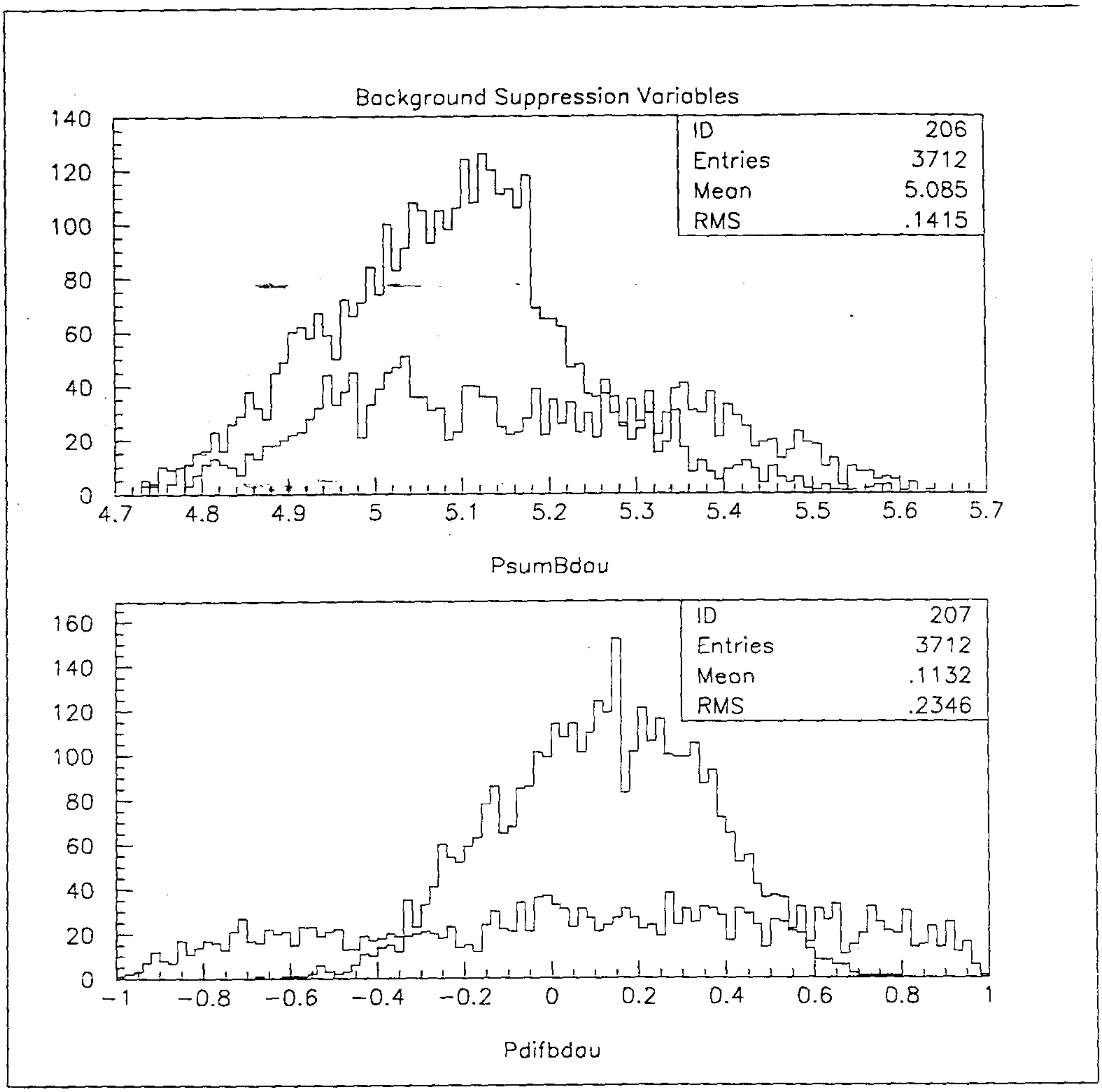
Discriminating Variable Distributions

Georges Vasseur

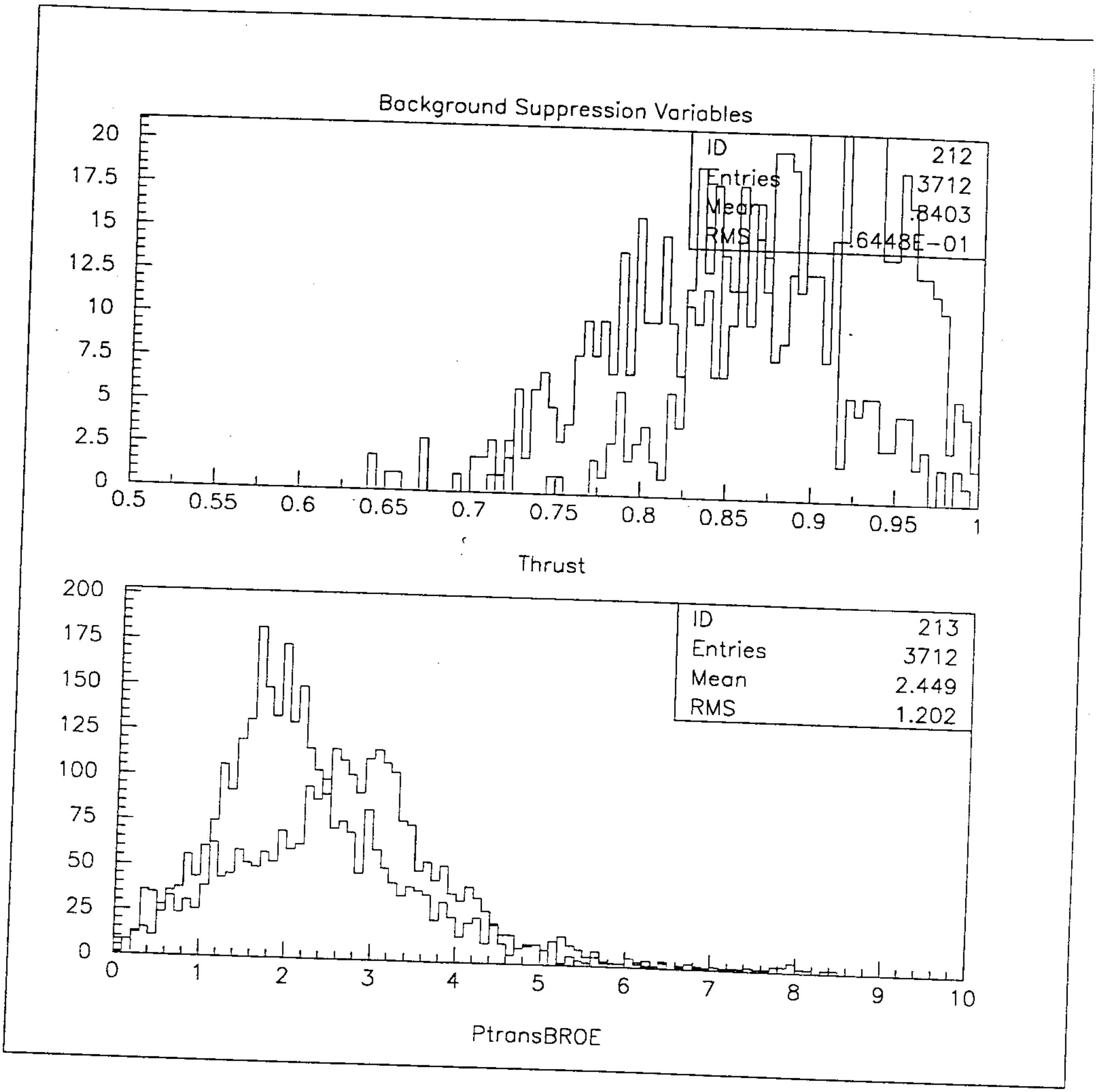
$B \rightarrow \rho\pi$



i) momentum of B0 in Upsilon cms
ii) mass χ^2 for CP decay mode particles
iii) Sum of transverse momentum of rest of event
iv) Fox Wolfram moment (second order/zeroth order)



Background Suppression Variables
ID 206
Entries 3712
Mean 5.085
RMS .1415
PsumBdou
ID 207
Entries 3712
Mean .1132
RMS .2346
Pdifbdou


Background Suppression Variables
ID 212
Entries 3712
Mean .8403
RMS .6448E-01
20
17.5
15
12.5
10
7.5
5
2.5
0
0.5
0.55
0.6
0.65
0.7
0.75
0.8
0.85
0.9
0.95
1
Thrust
ID 213
Entries 3712
Mean 2.449
RMS 1.202
200
175
150
125
100
75
50
25
0
0
1
2
3
4
5
6
7
8
9
10
PtransBROE

Practicalities of fighting qq~ background

Problem - need many qq~ events to estimate effects

- require a condensed sample (ie. ~1%) containing qq~ events of the most significant background for the channels considered

- each event contains a B-decay candidate for at least one of the CP-modes.

- apply to a set of channels with similar requirements for background studies.

Optimising the Cuts

Four combining methods are implemented in CORNELIUS :

Parametrized Approach (PA) [Ref. 3]
A statistical method, where correlations among the variables are accounted for, but are not fully exploited.

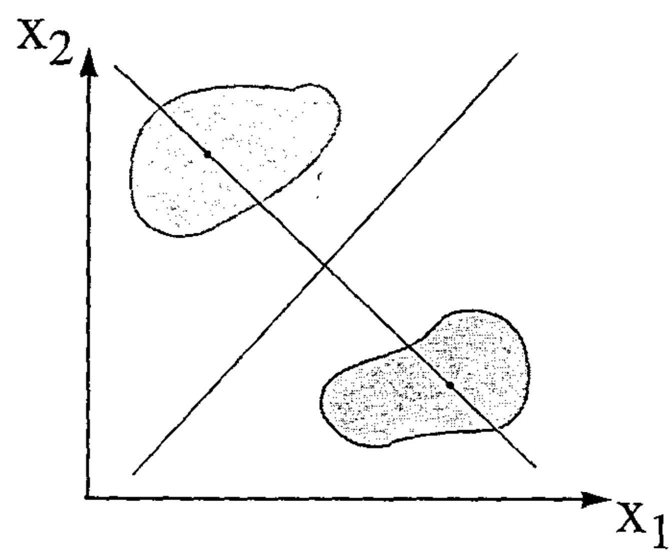
Fisher approach (FI)
Mahalanobis approach (MA) [Ref. 5]
Two linear discriminant methods, actually almost degenerate in their outputs (in this application)

Neural Network method (NN) [Ref. 4]
A non-linear method, which is capable of fully exploiting the correlations between variables.

Linear Discriminant Analysis

Form a single discriminating variable from a linear combination of selected variables

- want the best separation between classes (eg. sg and bg)



-define an axis such that the projection of classes onto it results in best separation

Fisher and Mahalanobis discriminants differ only in method of normalisation

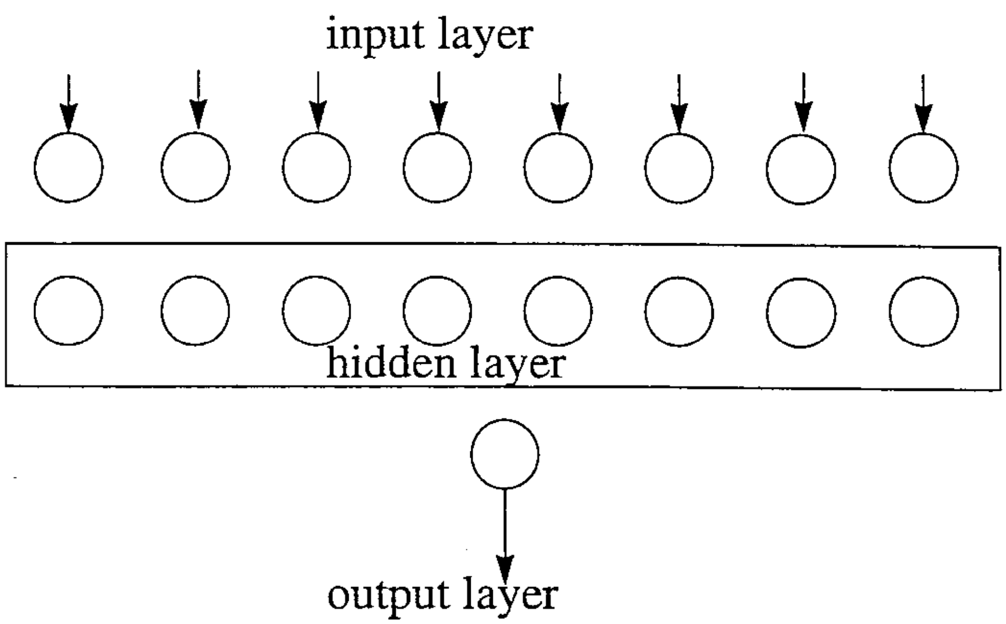
Artificial Neural Network

- generalisation to non-linear discrimination

- only as good as training distributions

- traditionally hard to understand systematics

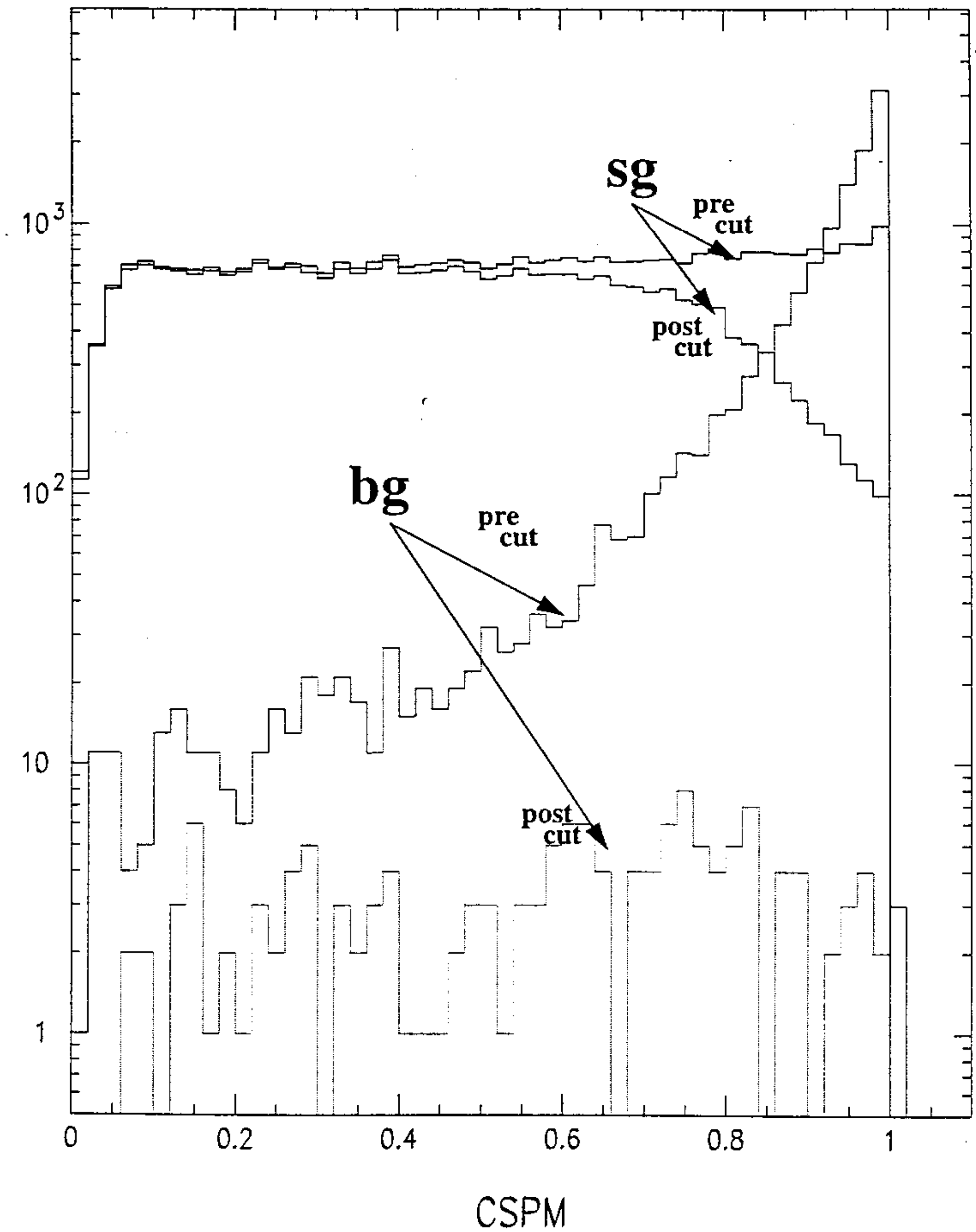
- simple ANN (one hidden layer) is sufficient for a classification task (ie. distinguishing sg and bg)



Cosine of the angle between the B decay axis and the sphericity axis of the rest of the event

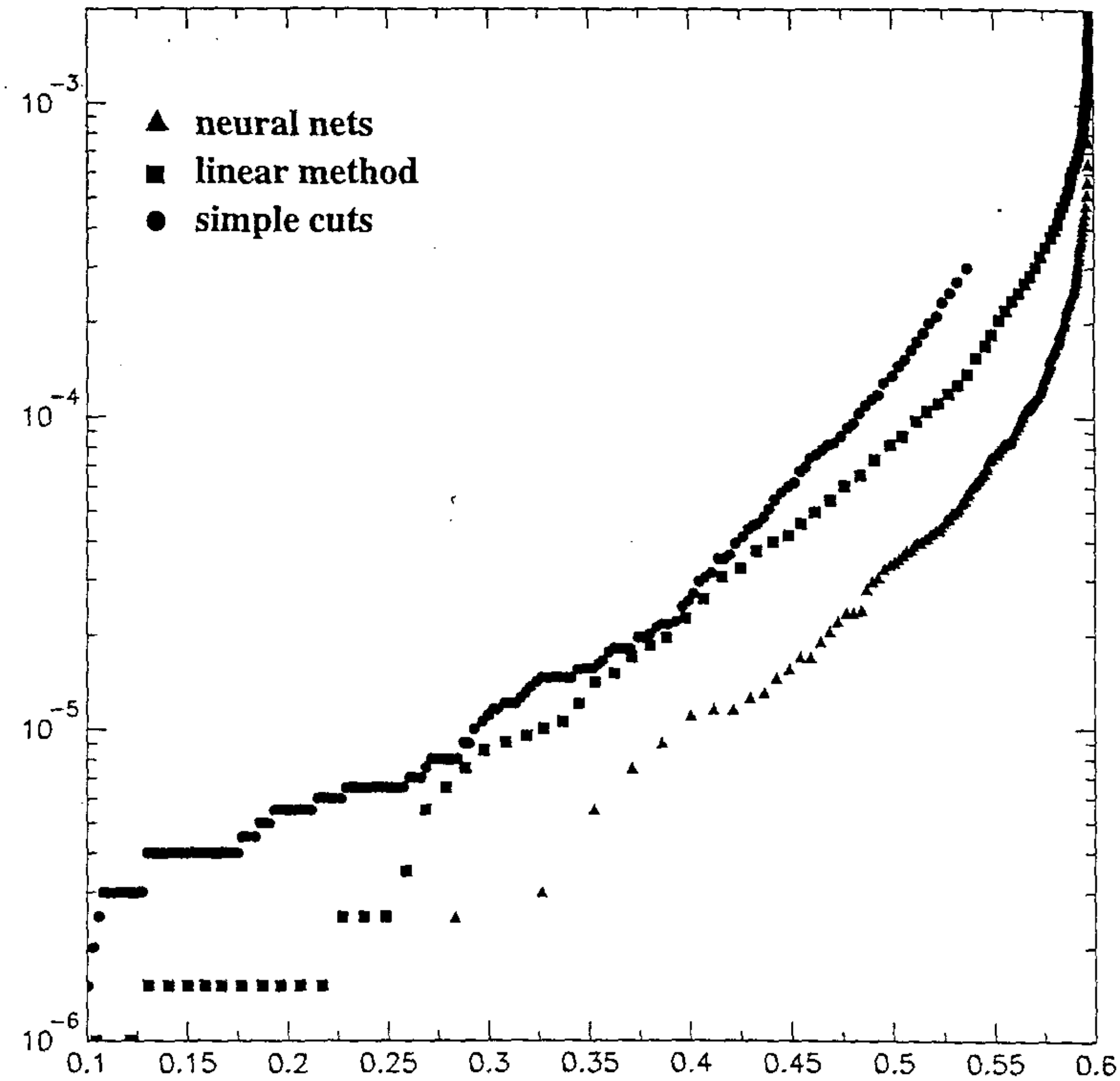
$B \rightarrow \rho\pi$

Georges Vasseur



Effect of a cut at 0.9 on the neural network
Efficiency of background v. signal

Georges Vasseur



Efficiency for signal for a 10^5 continuum background rejection

Method	simple cuts	linear	neural net
Efficiency	28.6	31.8	38.4

Summary

Effective background suppression is a key issue for BaBar in order to achieve our objectives

A range of methods can be employed to fight the different types of background :

- PID
- kinematic constraints
- topological variables

- multivariate analyses

Inclusion of background events in the final sample used for the CP-fit results in an increase in statistical uncertainty.

statistical reach : $sg^2/(sg + bg)$

References

[1] "Study of $B^0 \to J/\psi K^0_S$ and $B^0 \to \rho^{+-}\pi^{+-}$", P. Harrison, Babar Note #218

[2] "CORNELIUS user-guide" S. Versille, F. Le Diberder

[3] "Treatment of weighted events in a likelihood analysis of oscillations or CP-violation" D.E. Jaffe, F.,Le Diberder, M.H. Schune, Babar Note #132

[4] "Tagging studies for BaBar experiment using leptons with an Artificial Neural Network approach" A. Gaidot, Ch. Yeche

[5] "Linear Discriminant Analysis and Neural Network approach for Babar Tagging " A. Gaidot, Ch. Yeche

[6] "Pythia and Jetset, Physics and Manual" T. Sjostrand, Computer Physics Commun. 82 (1994) 74.