A search for baryon and lepton number violation in B decays using the BaBar dataset

SMU HEP Seminar

Southern Methodist University

M. Bellis

Stanford University

March 7^{th}, 2011

Outline

Outline

Motivation

- Our universe is matter...not anti-matter.

Motivation

- Our universe is matter...not anti-matter.
- t = early universe:
  - matter = anti-matter

Motivation

- Our universe is matter...not anti-matter.
- t = early universe:
  - matter = anti-matter
- t = now:
  - matter $\neq$ anti-matter
- How do we know?
  - Cosmic ray's are mostly matter.
  - γ-ray spectrum.
- Universe is compartmentalized?
  - *Very* difficult theoretically.

Sakharov conditions

- Andrei Sakharov (1967)
- "Violation of CP Invariance, c Asymmetry, and Baryon Asymmetry of the Universe."
- *Three conditions* required for matter (baryon) asymmetry. [1]

Sakharov conditions

- Andrei Sakharov (1967)
- "Violation of CP Invariance, c Asymmetry, and Baryon Asymmetry of the Universe."
- *Three conditions* required for matter (baryon) asymmetry. [1]
  1. Early universe is out of thermal equilibrium.
     - Rate for $\psi_{B=0} \rightarrow \psi_{B\neq 0}$ is different than for $\psi_{B\neq 0} \rightarrow \psi_{B=0}$.
     - The universe is cooling!

Sakharov conditions

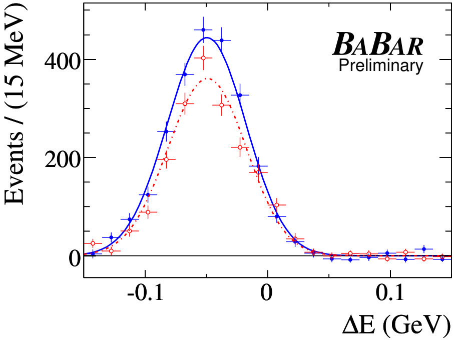
- Andrei Sakharov (1967)
- "Violation of CP Invariance, c Asymmetry, and Baryon Asymmetry of the Universe."
- *Three conditions* required for matter (baryon) asymmetry. [1]
  1. Early universe is out of thermal equilibrium.
     - Rate for $\psi_{B=0} \to \psi_{B\neq0}$ is different than for $\psi_{B\neq0} \to \psi_{B=0}$.
     - The universe is cooling!
  2. *C* and *CP*-violation
     - Decay rates are different for *matter* and *anti-matter*.

Sakharov conditions

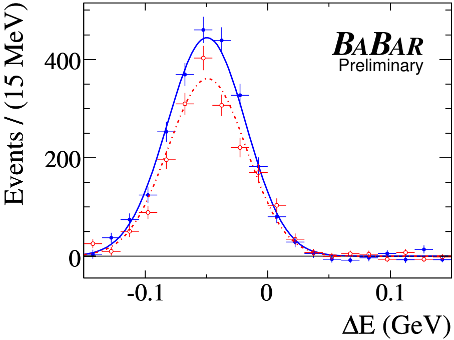
- Andrei Sakharov (1967)
- "Violation of CP Invariance, c Asymmetry, and Baryon Asymmetry of the Universe."
- *Three conditions* required for matter (baryon) asymmetry. [1]
  1. Early universe is out of thermal equilibrium.
     - Rate for $\psi_{B=0} \rightarrow \psi_{B\neq0}$ is different than for $\psi_{B\neq0} \rightarrow \psi_{B=0}$.
     - The universe is cooling!
  2. *C* and *CP*-violation
     - Decay rates are different for *matter* and *anti-matter*.
  3. Baryon number violation
     - Implies sum of baryons + anti-baryons is a *non-conserved quantity*.
- Let's look at these last two...

Direct CP-violation

- Direct CP-violation
  - $B^0 \to K^+\pi^-$
  - $\bar{B}^0 \to K^-\pi^+$



- Direct CP-violation
  - $B^0 \to K^+\pi^-$
  - $\bar{B}^0 \to K^-\pi^+$
- *Decay rates are different!*
- $A_{CP} \approx -0.1$
- Sakharov condition # 2!



Baryon number violation

- Baryon number violation actually *does* exist in the Standard Model.
  - Sphaleron, a non-perturbative process.
  - Occurs at *very* high temperatures. $T = 100\text{GeV}$. (10^{15}K)
  - Only found immediately after the big bang.
- Sakharov condition # 3!

Motivation

- Does this predict our asymmetric universe?
  - B for *baryon*
  - $B + \bar{B}$ annihilations in the early universe produced photons.
  - Asymmetry parameter (η).

$$\eta = \frac{N_B - N_{\bar{B}}}{N_\gamma} \approx 10^{-9}$$

Motivation

- Does this predict our asymmetric universe?
  - B for *baryon*
  - $B + \bar{B}$ annihilations in the early universe produced photons.
  - Asymmetry parameter (η).

$$\eta = \frac{N_B - N_{\bar{B}}}{N_\gamma} \approx 10^{-9}$$

- Combination of observed *CP*-violation and and theoretical BNV in Standard Model is insufficient by 10 orders of magnitude!!!

Motivation

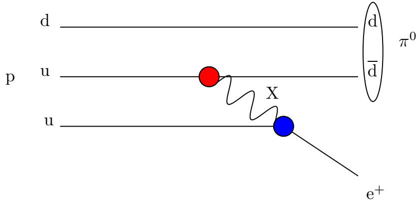
- Does this predict our asymmetric universe?
  - *B* for *baryon*
  - $B + \bar{B}$ annihilations in the early universe produced photons.
  - Asymmetry parameter (η).

$$\eta = \frac{N_B - N_{\bar{B}}}{N_\gamma} \approx 10^{-9}$$

- Combination of observed *CP*-violation and and theoretical BNV in Standard Model is insufficient by 10 orders of magnitude!!!
- Additional CP violation? *Much work out there...no smoking gun.*
- Additional BNV?
- 1974, "Unity of All Elementary Particle Forces.", Georgi and Glashow
  - Proton decay mediated by heavy bosons (X & Y) which couple to *quarks* and *leptons*.
- Many Grand Unified Theories $\Rightarrow$ BNV
- How would proton decay work?

Proton decay

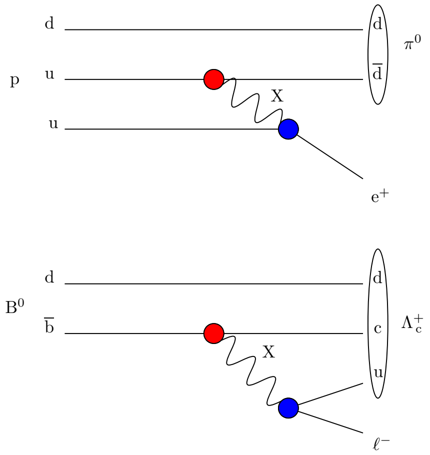
- X is $q = \frac{1}{3}$
  - $X \rightarrow q + q$
  - $X \rightarrow q + \ell$
- **B-L is conserved quantity.**



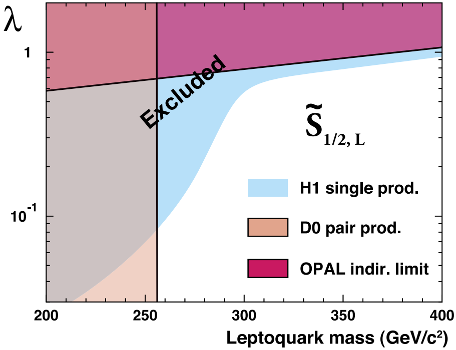
- X is $q = \frac{1}{3}$
  - $X \rightarrow q + q$
  - $X \rightarrow q + \ell$
- **B-L is conserved quantity.**
- Hypothesize a flavour/generation dependance to this interaction...
  - $B^0 \rightarrow \Lambda_c^+ \ell^-$
  - $B^+ \rightarrow \Lambda^0 \ell^+$
- $\ell = \mu$ or e

$$\begin{pmatrix} X_{u\bar{d}} & X_{c\bar{d}} & X_{t\bar{d}} \\ X_{u\bar{s}} & X_{c\bar{s}} & X_{t\bar{s}} \\ X_{u\bar{b}} & \mathbf{X_{c\bar{b}}} & X_{t\bar{b}} \end{pmatrix}$$

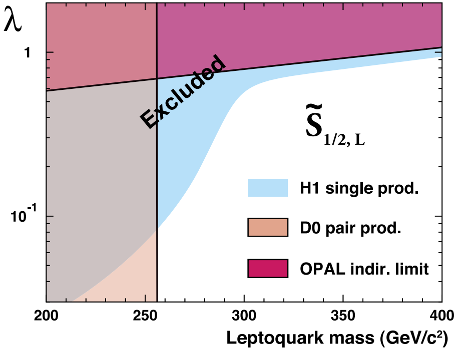
$$\begin{pmatrix} \mathbf{X_{\bar{u}e^-}} & X_{\bar{c}e^-} & X_{\bar{t}e^-} \\ \mathbf{X_{\bar{u}\mu^-}} & X_{\bar{c}\mu^-} & X_{\bar{t}\mu^-} \\ X_{\bar{u}\tau^-} & X_{\bar{c}\tau^-} & X_{\bar{t}\tau^-} \end{pmatrix}$$



- Experimental work
  - Proton lifetime $> 10^{32}$ years!
  - Tevatron and HERA searches for *"true"* leptoquarks [2].
    - M - Mass of the mediating leptoquark (X-boson)
    - λ - Yukawa coupling



- Experimental work
  - Proton lifetime $> 10^{32}$ years!
  - Tevatron and HERA searches for *"true"* leptoquarks [2].
    - M - Mass of the mediating leptoquark (X-boson)
    - λ - Yukawa coupling
- Theoretical work
  - *"Baryon number violation involving higher generations."*, Hou, et.al.[3]
    - Uses proton decay to constrain upper limits.
    - $\mathcal{B}(B^0) \to \Lambda_c^+ \ell^- < 4 \times 10^{-29}$
    - "Despite our findings, we believe it is still worth to look for BNV processes in τ, charm, B, and maybe in the future in top decays."
- **This analysis is the first search for $B \to \Lambda_{(c)} \ell$ decays.**



BNV in B decays

$$\ell = \mu \text{ or } e$$

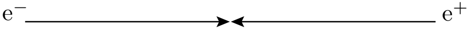
$$B^0 \rightarrow \Lambda_c^+ \ell^-$$

$$\rightarrow$$

$$\ell = \mu \text{ or } e$$

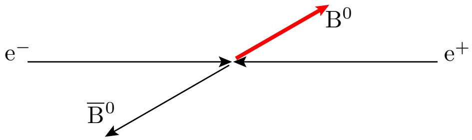
$$B^0 \rightarrow \Lambda_c^+ \ell^-$$

$$\rightarrow$$



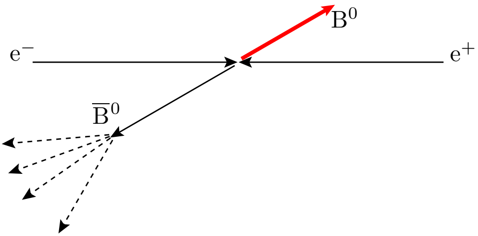
$$\ell = \mu \text{ or } e$$

$$B^0 \rightarrow \Lambda_c^+ \ell^-$$
$$\rightarrow$$



$$\ell = \mu \text{ or } e$$

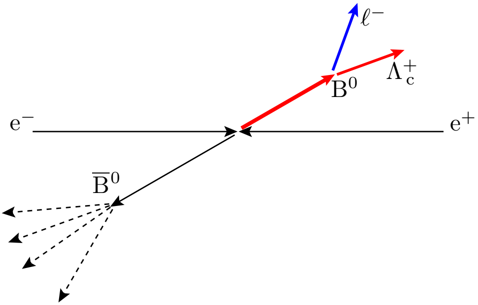
$$B^0 \rightarrow \Lambda_c^+ \ell^-$$
$$\rightarrow$$



$\ell = \mu$ or e

$B^0 \rightarrow \Lambda_c^+ \ell^-$

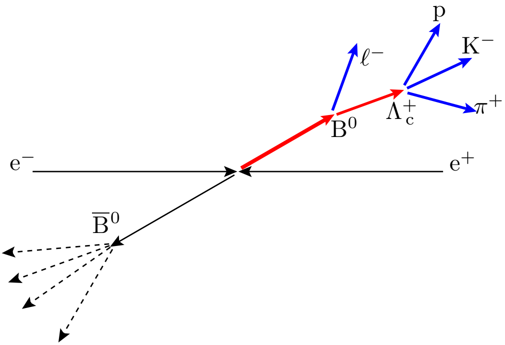
$\rightarrow$



BNV IN B DECAYS

$$\ell = \mu \text{ or } e$$

$$B^0 \rightarrow \Lambda_c^+ \ell^-$$
$$\Lambda_c^+ \rightarrow pK^- \pi^+$$



$\ell = \mu$ or e

$B^0 \rightarrow \Lambda_c^+ \ell^-$

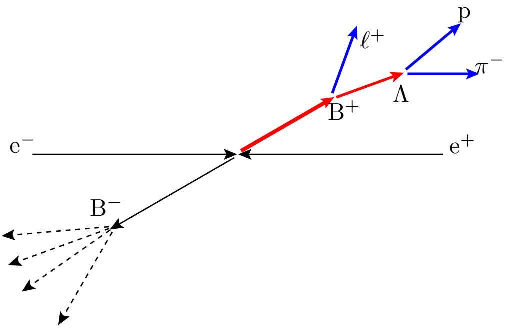
$\Lambda_c^+ \rightarrow p K^- \pi^+$

$B^- \rightarrow \Lambda^0 \ell^-$

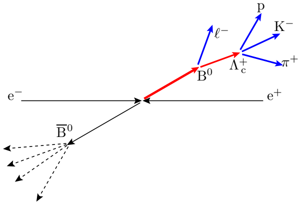
$\Lambda^0 \rightarrow p \pi^-$

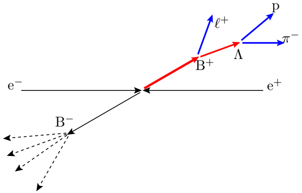
$B^- \rightarrow \bar{\Lambda}^0 \ell^-$

$\bar{\Lambda}^0 \rightarrow \bar{p} \pi^-$



- Experimental requirements:
  - Cleanly identify **4** charged particles.
  - Demonstrate they came from a B meson.



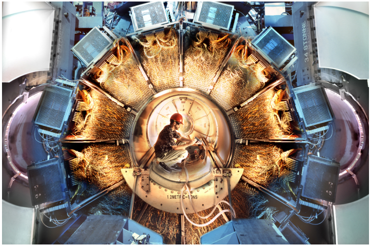


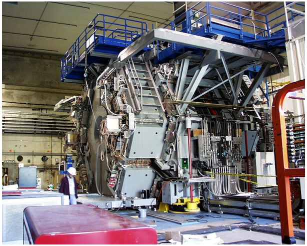
Outline

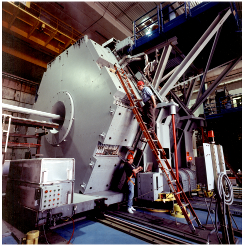
SLAC and BaBar

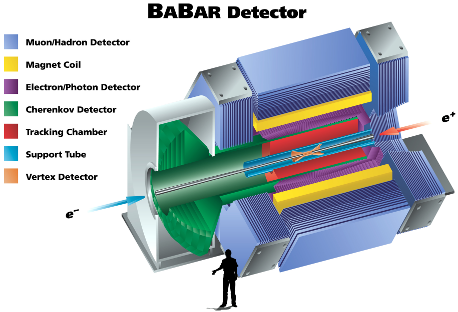
SLAC and BaBar

SLAC and BaBar

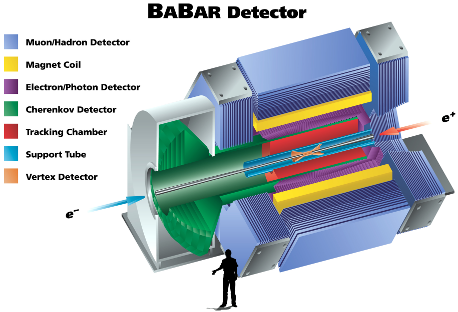




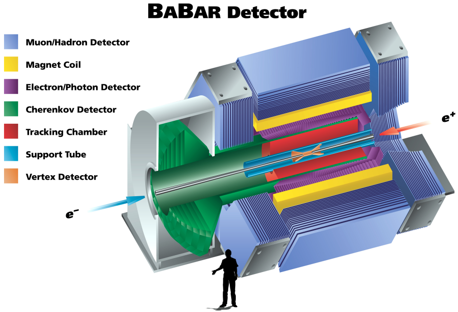




- BaBar at SLAC
- 1999-2008

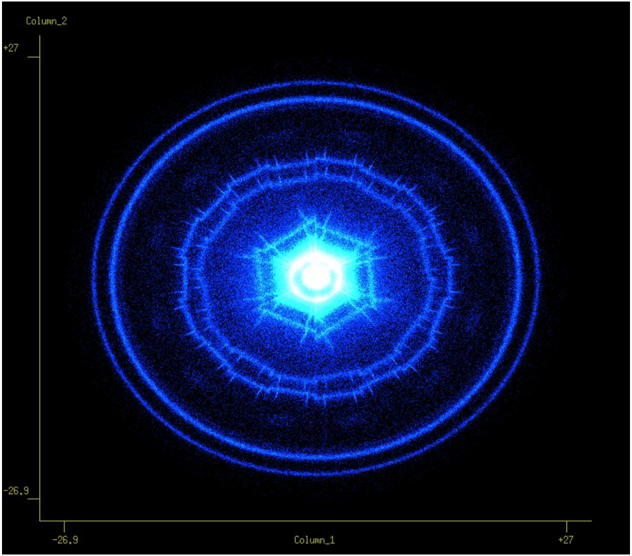


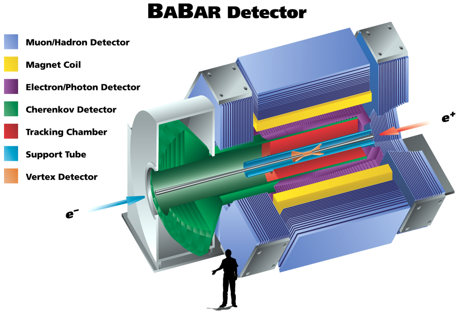
- BaBar at SLAC
- 1999-2008
- PEP-II asymmetric e^+e^- collider
  - Ran on $\Upsilon(4S)$
  - Instantaneous $\mathcal{L} = 10^{-34}$ cm^{-2} s^{-1}
- 1 billion B mesons



- BaBar at SLAC
- 1999-2008
- PEP-II asymmetric e^+e^- collider
  - Ran on $\Upsilon(4S)$
  - Instantaneous $\mathcal{L} = 10^{-34}$ cm^{-2} s^{-1}
- 1 billion B mesons
- *Backgrounds (or other physics!)*
  - $e^+e^- \to u\bar{u}/d\bar{d}/s\bar{s}/c\bar{c}$
  - $e^+e^- \to e^+e^-/\mu^+\mu^-/\tau^+\tau^-$
- 400+ papers and counting.
- Excellent momentum and spatial resolution.

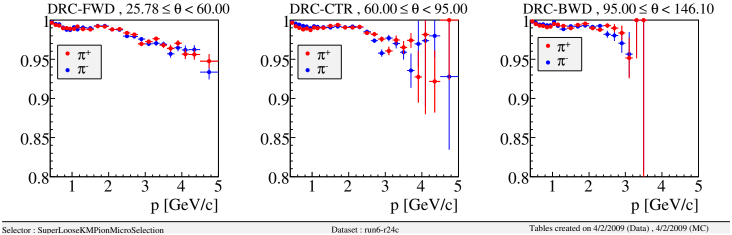
Pair production





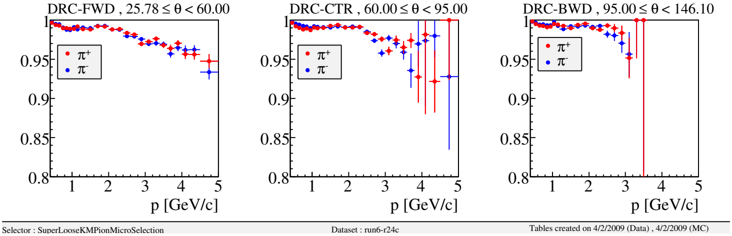
- Sophisticated particle ID.
- Inner silicon vertex detector.
  - Energy, spatial position (momentum)
- Drift chambers.
  - Energy, momentum
- Cerenkov detector (DIRC)
  - Velocity (π/K ID)
- Muon detection in outer region
  - Timing, spatial position
- All fed into a neural net algorithm.
- Gives analysts 6 levels.
  - For each particle type (e, μ, p, π, K)
  - Vary purity/efficiency.
  - Function of kinematics.

π efficiency

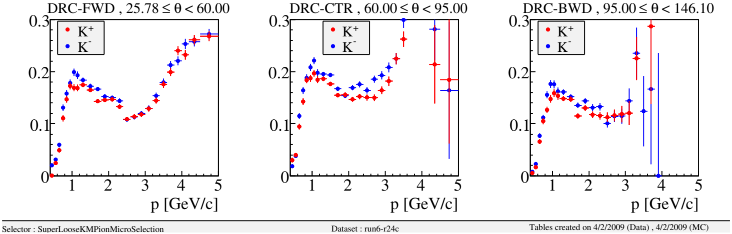


Particle ID (PID)

π efficiency

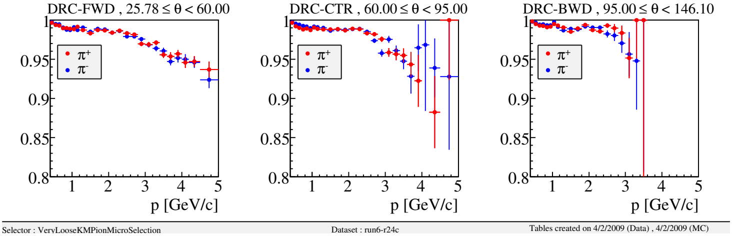


K contamination

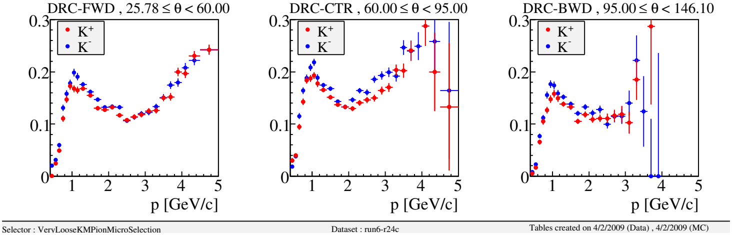


Particle ID (PID)

π efficiency

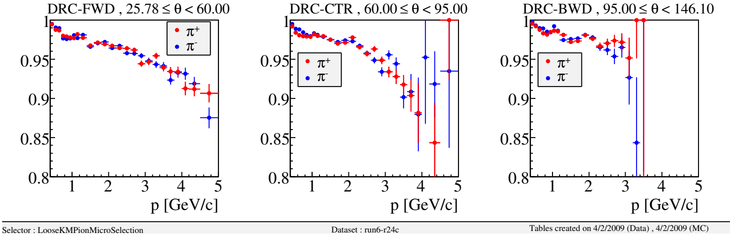


K contamination

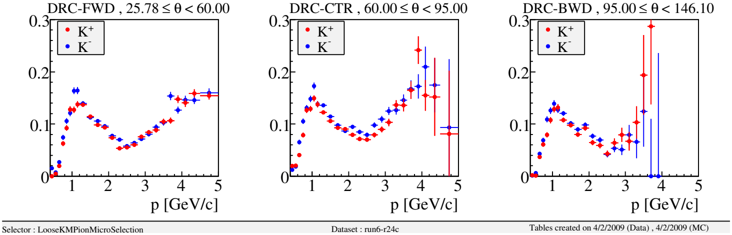


Particle ID (PID)

π efficiency

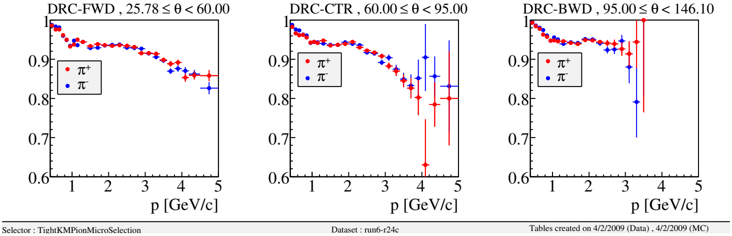


K contamination

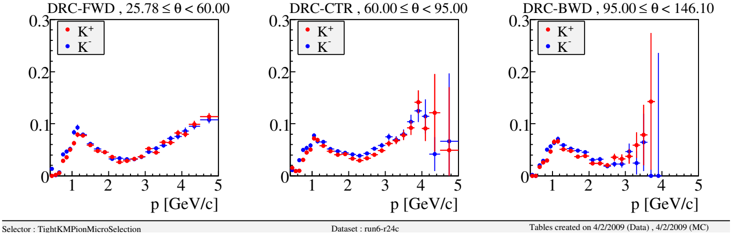


Particle ID (PID)

π efficiency

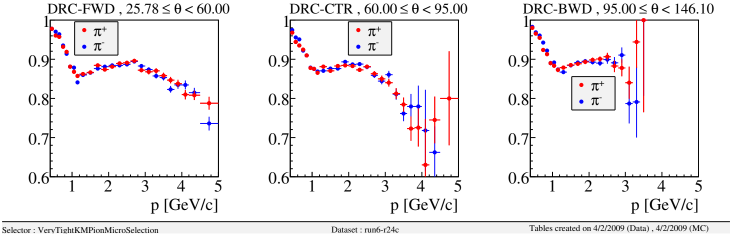


K contamination

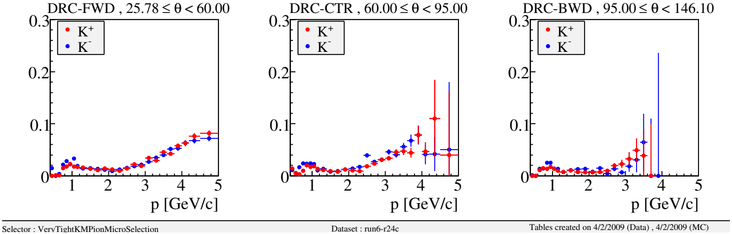


Particle ID (PID)

π efficiency

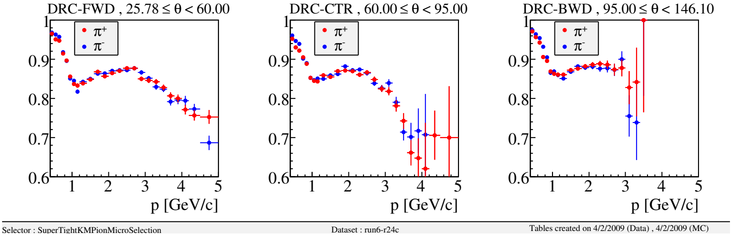


K contamination

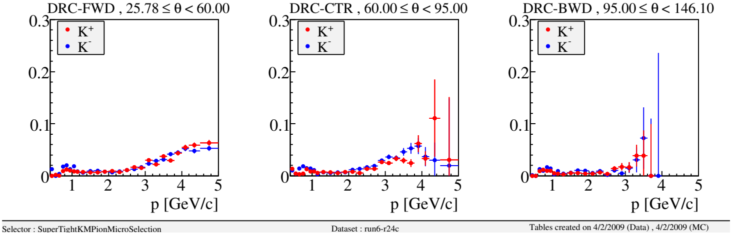


Particle ID (PID)

π efficiency



K contamination



Outline

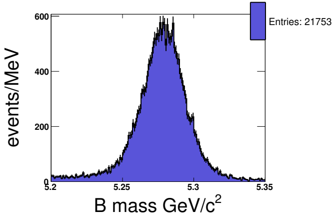
Beam constraints

- Unbinned extended likelihood fit.
- Fit using *appropriate kinematic variables*.

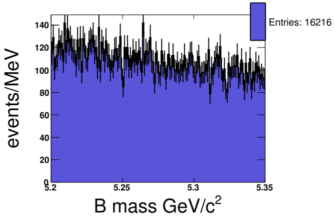
Beam constraints

- Unbinned extended likelihood fit.
- Fit using *appropriate kinematic variables*.
- *B* mass
- Beam energy resolution is better than B energy (combined track $\vec{p}$) resolution.

Signal process (Monte Carlo)



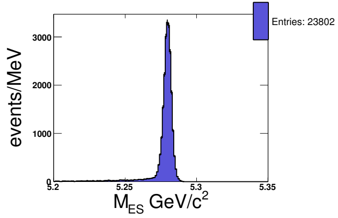
All background processes (Monte Carlo)



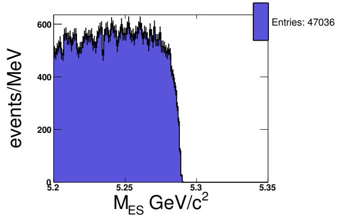
Beam constraints

- Unbinned extended likelihood fit.
- Fit using *appropriate kinematic variables*.
- B mass
- Beam energy resolution is better than B energy (combined track $\vec{p}$) resolution.
- $m_{ES} = \sqrt{\frac{1}{4}s - (p_B^*)^2}$

Signal process (Monte Carlo)



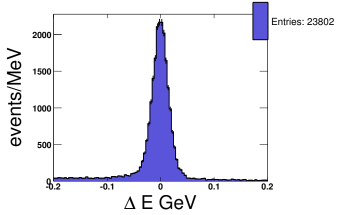
All background processes (Monte Carlo)



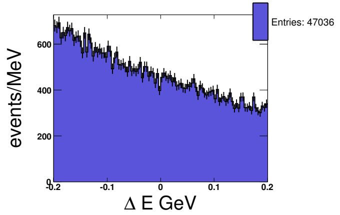
Beam constraints

- Unbinned extended likelihood fit.
- Fit using *appropriate kinematic variables*.
- B mass
- Beam energy resolution is better than B energy (combined track $\vec{p}$) resolution.
- $m_{ES} = \sqrt{\frac{1}{4}s - (p_B^*)^2}$
- $\Delta E = E_B^* - \frac{1}{2}\sqrt{s}$

Signal process (Monte Carlo)



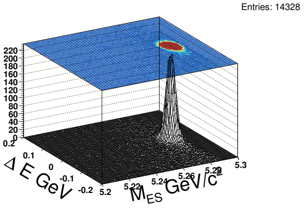
All background processes (Monte Carlo)



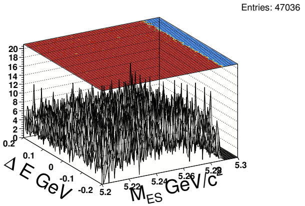
Beam constraints

- Unbinned extended likelihood fit.
- Fit using *appropriate kinematic variables*.
- B mass
- Beam energy resolution is better than B energy (combined track $\vec{p}$) resolution.
- $m_{ES} = \sqrt{\frac{1}{4}s - (p_B^*)^2}$
- $\Delta E = E_B^* - \frac{1}{2}\sqrt{s}$
- **Discriminating power in 2D plane.**
- Signal region is blinded in data analysis!

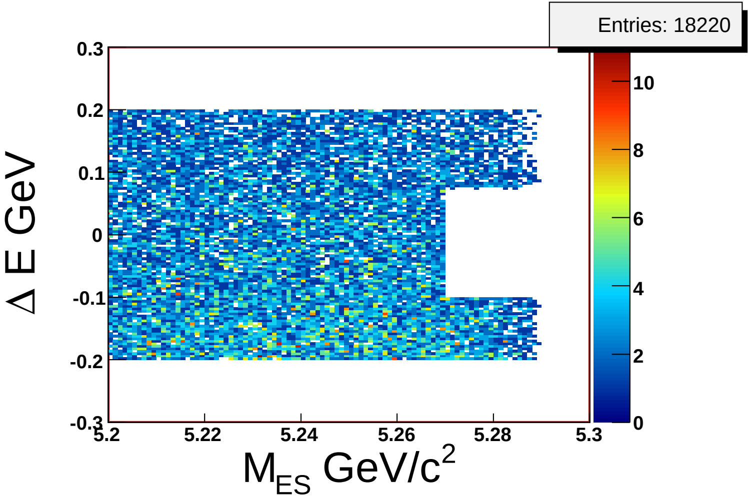
Signal process (Monte Carlo)



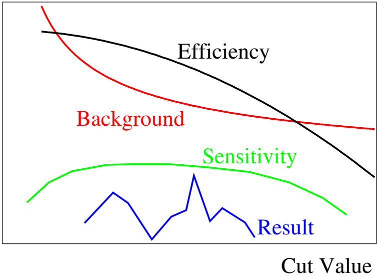
All background processes (Monte Carlo)



Our blind data

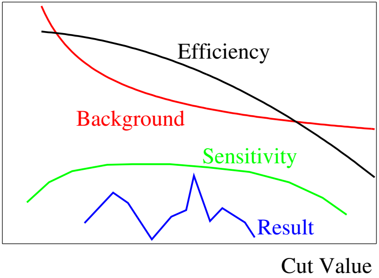


- Blind searches
  - Taken from Roodman, "*Blind analysis in particle physics*" [4]



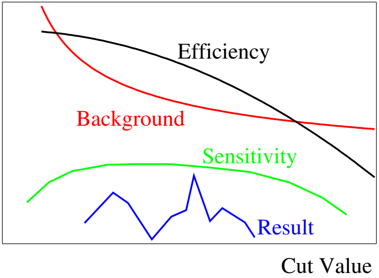
Blind searches

- Blind searches
  - Taken from Roodman, "*Blind analysis in particle physics*" [4]
- *Experimenter's bias.*



Blind searches

- Blind searches
  - Taken from Roodman, "*Blind analysis in particle physics*" [4]
- *Experimenter's bias.*
- Medical field: double blind trials.
- Electron e/m, Dunnington (1933)



Blind searches

- Blind searches
  - Taken from Roodman, "*Blind analysis in particle physics*" [4]
- *Experimenter's bias.*
- Medical field: double blind trials.
- Electron e/m, Dunnington (1933)
- *History of measurements?*

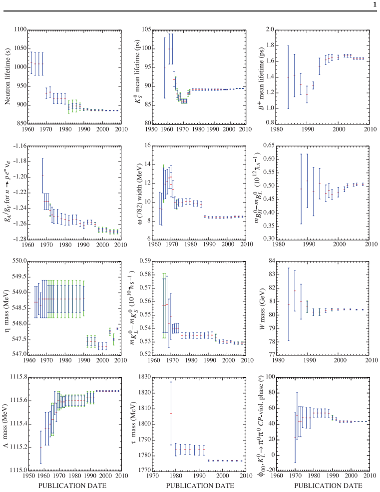


Figure 2: A historical perspective of values of a few particle properties tabulated in this *Review* as a function of date of publication of the *Review*. A full error bar indicates the quoted error; a thick-lined portion indicates the same but without the "scale factor."

Blind searches

- Blind searches
  - Taken from Roodman, "*Blind analysis in particle physics*" [4]
- *Experimenter's bias.*
- Medical field: double blind trials.
- Electron e/m, Dunnington (1933)
- *History of measurements?*
- How do you guard against this?
- Don't look!

1

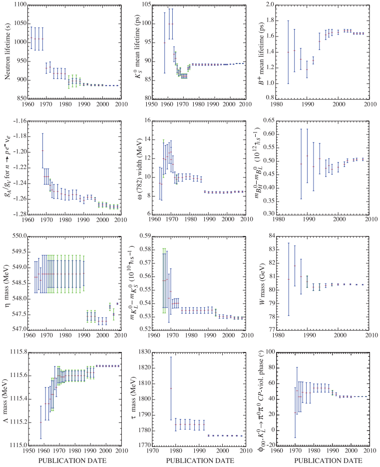
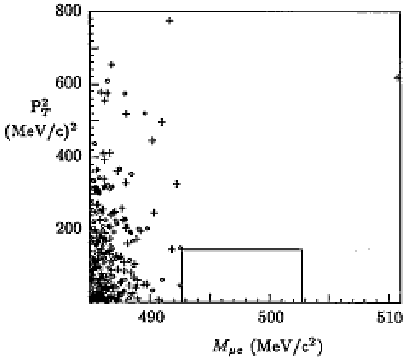


Figure 2: A historical perspective of values of a few particle properties tabulated in this *Review* as a function of date of publication of the *Review*. A full error bar indicates the quoted error; a thick-lined portion indicates the same but without the "scale factor."

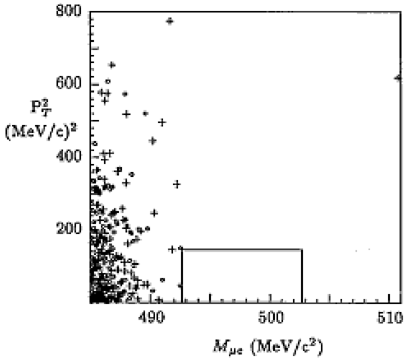
Blind searches

- Hidden signal box.
  - $K_L^0 \to \mu^{\pm} e^{\mp}$
  - Ariska, **PRL 70**, 1993



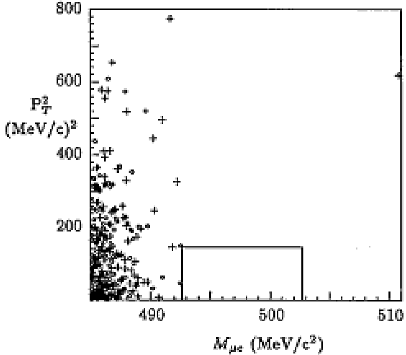
Blind searches

- Hidden signal box.
    - $K_L^0 \to \mu^{\pm} e^{\mp}$
    - Ariska, **PRL 70**, 1993
- Kinematics dictates region of interest.



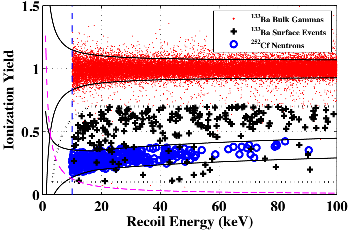
Blind searches

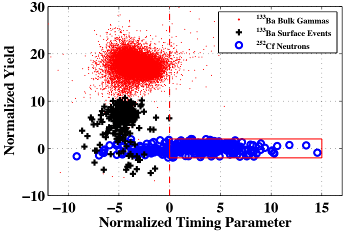
- Hidden signal box.
  - $K_L^0 \rightarrow \mu^\pm e^\mp$
  - Ariska, **PRL 70**, 1993
- Kinematics dictates region of interest.
- Other approaches: hidden answer, random noise.

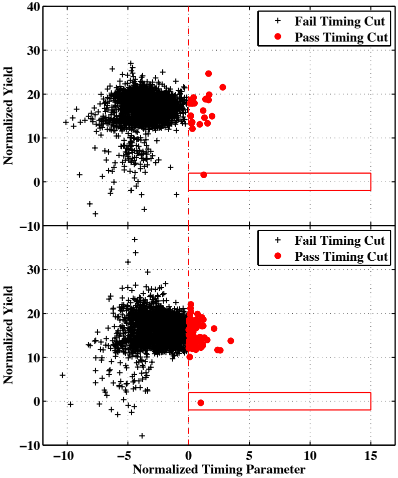


Blind search

CDMS (2009), dark matter search







Analysis overview

- Full dataset (435 fb^{-1})
  - Constrain vertex of B candidate.
  - Mass/vertex constrain Λ_c/Λ^0 candidate.
- Optimize PID selectors.
- Multivariate discriminator.
  - TMVA.
  - Input variables.
  - Choice of background sample.
  - Check for correlations with ΔE and m_{ES}.
- Check for fit bias.
- *Fit the unblinded data.*

- Punzi figure of merit.
  - Strike a balance between *setting upper limit for null results* and *observation of a small signal*.

$$\text{f.o.m.} = \frac{\epsilon_S}{\sqrt{B} + a/2}$$

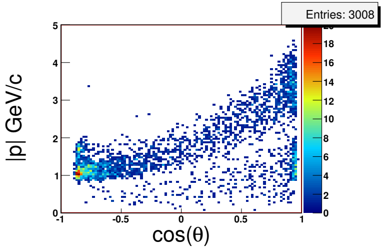
  - a is the significance (sigma) at which you want to make a final claim.
    - For this analysis, $a = 5$.
  - ϵ_S is the efficiency of the signal.
  - Don't need to know S (cross section), but we do need an idea of B (background).

Definition of MC samples

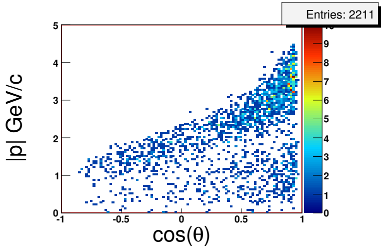
- During optimization, make extensive use of MC samples.
- GEANT4 simulation of the detector.
- Simulated signal events (assume no polarization of $(\Lambda_{(c)})$
- Background samples.
- $q\bar{q}$
    - $u\bar{u}/d\bar{d}/s\bar{s}/c\bar{c}$
- $B\bar{B}$
    - B^+B^-
    - $B^0\bar{B}^0$
- All generics
    - $q\bar{q} + B\bar{B}$ (weighted by relative cross sections)

Bhabba leakage

- Leakage from Bhabba events.
- Eliminated by requiring # charged tracks > 4.



(a) Sideband data.

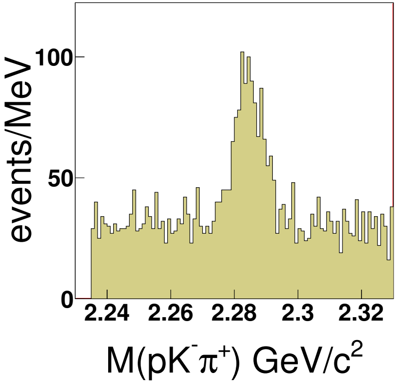


(b) Generic MC

Figure: $|p|$ vs. $\cos(\theta)$ for the π^- coming from the Λ^0 candidate.

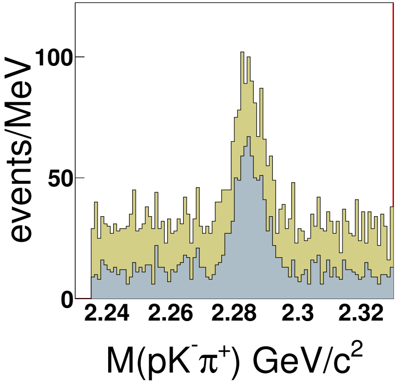
PID

- Optimize PID for kinematics.
- e.g. $\Lambda_c^+ \to pK^-\pi^+$
- Loosest PID



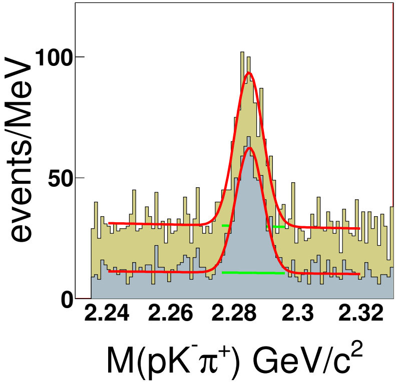
PID

- Optimize PID for kinematics.
- e.g. $\Lambda_c^+ \rightarrow pK^-\pi^+$
- Loosest PID
- Some set of PID criteria.

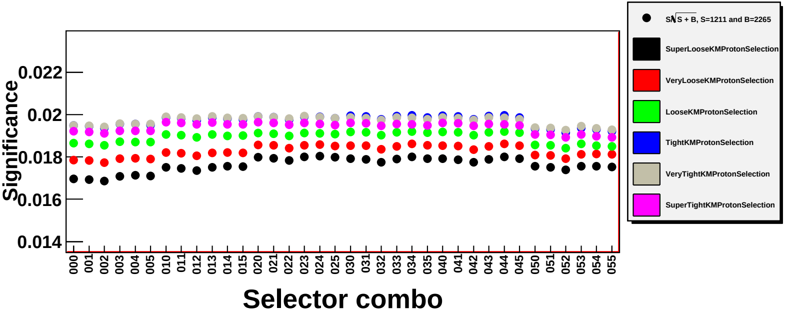


PID

- Optimize PID for kinematics.
- e.g. $\Lambda_c^+ \to pK^-\pi^+$
- Loosest PID
- Some set of PID criteria.
- Use Λ_c^+ *efficiency* and *background rejection* to optimize selection.



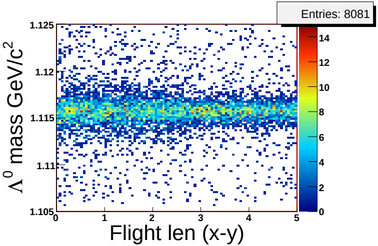
- PID selectors optimized for this signal.



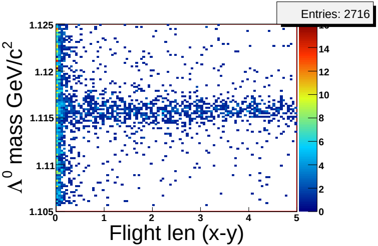
(a) Signal significance for the 216 PID selector combinations.

Figure: Optimization criteria for the PID selectors for the $B \to \Lambda_c^+ \mu$ mode.

- Λ: $c\tau = 7.89$ cm
- Pristine Λ^0 candidates after transverse flight length > 0.2cm.



(a) $\Lambda^0\mu$ Signal MC



(b) $\Lambda^0\mu$ Generic MC

FIGURE: Invariant mass of the Λ^0 candidate vs. the transverse flight length

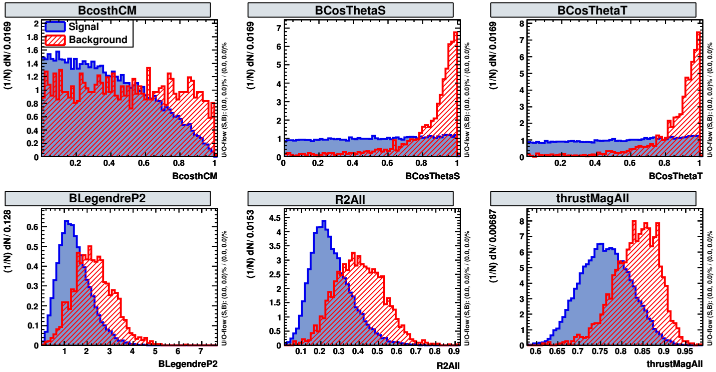
Training variables

- Explored *multivariate discriminators*.
- TMVA implementaion.
  - Toolkit for Multivariate Data Analysis with ROOT.
  - http://tmva.sourceforge.net/
- Pruned discriminating variables to most sensitive that *did not have high correlations with* $\Delta E/m_{ES}$.
- Six variables.
  - B cos(θ) CM
  - B cos(θ) sphericity wrt ROE sphericity
  - B cos(θ) thrust wrt ROE thrust
  - Legendre P2 (historical name)
    - Moments
    - Use ROE tracks, and B-thrust axis
  - Thrust all
  - R2 all
    - Ratio of Fox-Wolfram moments (0 and 2)

Training variables (data and MC)

Comparison between signal MC and $q\bar{q}$ MC.

$\Lambda_c \mu^-$



- Checked correlations with ΔE and m_{ES}.
  - **Bootstrap method** used to estimate significance of correlation coefficient.
  - Numerical procedure to estimate some *estimator*.
  - When you only have *one* data sample.
  - Originally applied to calculating the error of a correlation coefficent!
- Wound up using only 4 of the variables for Λ^0 modes.
- Orthogonally, checked discrimination power using $q\bar{q}$ MC or $q\bar{q} + B\bar{B}$ MC as background training sample.

- **Bootstrap method**
- Efron (1982)
- Numerical procedure to estimate some *estimator*.
- When you only have *one* data sample.
- Originally applied to calculating the error of a correlation coefficent!
- How does it work?

Bootstrap method

- Given some dataset, $\vec{x}$, of size n.
- Need error of some characteristic of that dataset: $\hat{\rho}$

- Given some dataset, $\vec{x}$, of size n.
- Need error of some characteristic of that dataset: $\hat{\rho}$
- Sample from original dataset to create multiple (1000) datasets of n entries.
  - $\vec{x} = (0, 1, 2, 3, 4)$

Bootstrap method

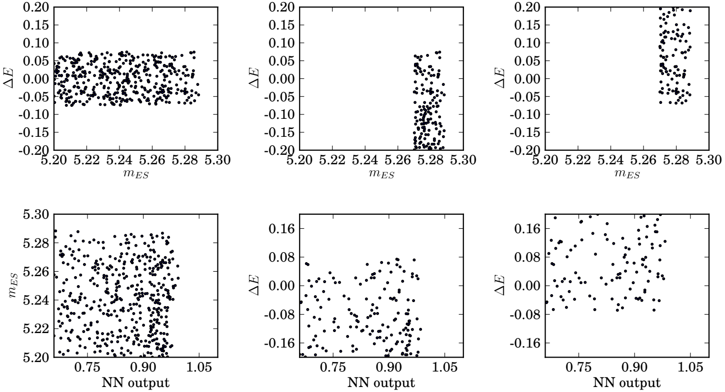
- Given some dataset, $\vec{x}$, of size n.
- Need error of some characteristic of that dataset: $\hat{\rho}$
- Sample from original dataset to create multiple (1000) datasets of n entries.
  - $\vec{x} = (0, 1, 2, 3, 4)$
    - $\vec{x}_0 = (4, 4, 0, 2, 3)$
    - $\vec{x}_1 = (1, 4, 3, 4, 0)$
    - $\vec{x}_2 = (3, 4, 2, 1, 4)$
    - $\vdots$

Bootstrap method

- Given some dataset, $\vec{x}$, of size n.
- Need error of some characteristic of that dataset: $\hat{\rho}$
- Sample from original dataset to create multiple (1000) datasets of n entries.
  - $\vec{x} = (0, 1, 2, 3, 4)$
    - $\vec{x}_0 = (4, 4, 0, 2, 3)$
    - $\vec{x}_1 = (1, 4, 3, 4, 0)$
    - $\vec{x}_2 = (3, 4, 2, 1, 4)$
    - $\vdots$
- For each dataset, calculate it's own $\hat{\rho}^*$.
- Use this distribution to quote a confidence interval (95% in upcoming examples)

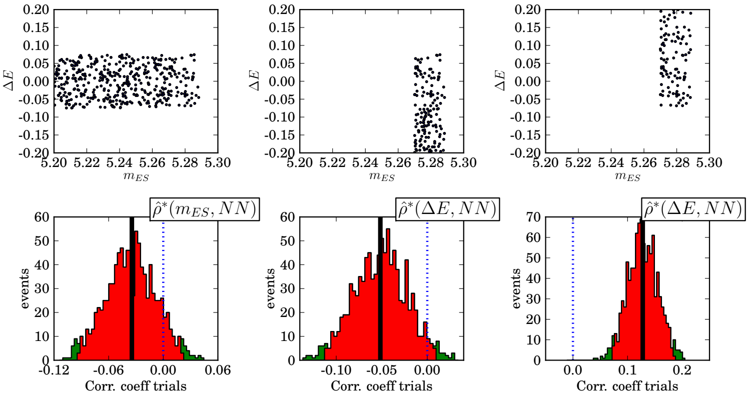
Correlation coefficients

Look in regions of $\Delta E / m_{ES}$ plane.



Correlation coefficients

Look in regions of $\Delta E/m_{ES}$ plane.



- **Red** region of histogram shows 95% confidence interval.
- **Black** solid line shows value of correlation coefficient for original dataset.
- **Blue** dashed line is at 0.

Summary of correlation coefficients

Decay Mode	Training sample	# vars.	Region 1	Region 2	Region 3
$\Lambda_c^+\mu^-$	$q\bar{q}$	6	(-0.09, 0.02)	(-0.11, 0.01)	**(0.07, 0.18)**
$\Lambda_c^+e^-$	$q\bar{q}$	6	(-0.11, 0.03)	**(-0.19,-0.07)**	(-0.02, 0.11)
$\Lambda^0\mu^-$	$q\bar{q}$	4	**(-0.18,-0.02)**	**(-0.28,-0.12)**	**(0.07, 0.23)**
Λ^0e^-	$q\bar{q}$	4	(-0.09, 0.08)	(-0.00, 0.15)	(-0.03, 0.16)
$\bar{\Lambda}^0\mu^-$	$q\bar{q}$	4	**(0.02, 0.15)**	(-0.07, 0.07)	**(-0.24,-0.11)**
$\bar{\Lambda}^0e^-$	$q\bar{q}$	4	(-0.12, 0.11)	**(-0.25,-0.02)**	**(-0.26,-0.04)**

TABLE: Confidence intervals of correlation coefficients for different modes/regions. Red intervals are inconsistent with 0

.

Doesn't appear to be able to create a peak!

Multivariate discriminator

- Summary of MVA studies.
  - Used **MLP neural net** implementation in TMVA
  - Used $q\bar{q}$ MC as background training sample.
  - Optimized cut on neural net output for $\Lambda^0 \ell$ modes.
  - Loose cut (90% signal efficiency) for $\Lambda_c^+ \ell$ modes.
    - Will include in fit as **third variable in fit** for $\Lambda_c^+ \ell$ modes.

Summary of remaining events

TABLE: Remaining events after all cuts for each decay mode for both fitting region (still blinded) and estimated signal region.

Decay Mode	Fitting region	Signal region
$\Lambda_c^+\mu^-$	900	18-25
$\Lambda_c^+e^-$	700	14-20
$\Lambda^0\mu^-$	350	7-10
Λ^0e^-	220	5-8
$\bar{\Lambda}^0\mu^-$	220	5-8
$\bar{\Lambda}^0e^-$	80	1-3

FITS

Unbinned extended maximum likelihood method.

$$\mathcal{L} = \frac{e^{-\nu}\nu^{n}}{n!} \times \prod_{i}^{n} \mathcal{P}(\vec{x}, \vec{k})$$

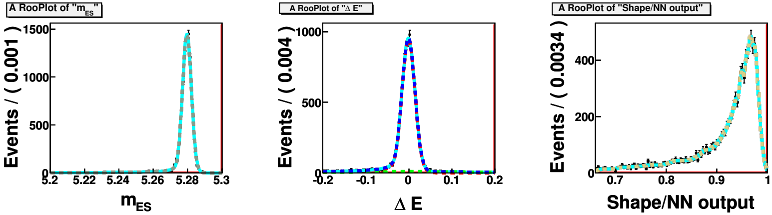
- PDF $\mathcal{P}(\vec{x}, \vec{k})$ provided for signal and background.
- Conversion factor ($\mathcal{F}$) turns number of signal events into branching fraction ($\mathcal{B}$).
  - $\mathcal{B}(B \to \Lambda\ell) = n_{\mathrm{sig}}/\mathcal{F}$
- We include a Gaussian constraint on conversion factor than incorporates systematic errors.
  - Takes into account asymmetric errors.
  - Incorporates systematics into the upper limit calculation.

$$LH = \frac{(\mathcal{F} - \mathcal{F}_{\mathrm{fit}})^2}{2\sigma_{\mathcal{F}}^2} - \sum_{i}^{n} \ln \mathcal{L}$$

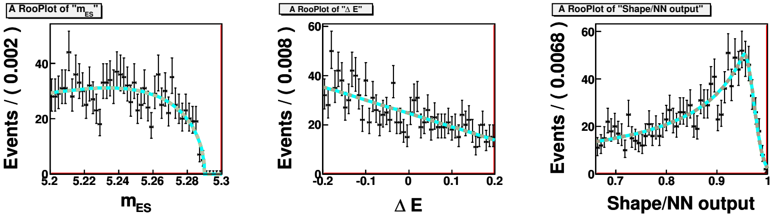
PDF descriptions

- m_{ES}
  - Signal PDF: **Crystal Ball function**.
  - Background PDF: **Argus function**.
- ΔE
  - Signal PDF: **Double Crystal Ball function**. Constrained to have the same mean.
  - Background PDF: **Linear function**.
- NN output (only in Λ_c fits)
  - Signal PDF: **RooKeysPdf**. Adaptive kernel estimation.
  - Background PDF: **Crystal Ball function**.

PDF DESCRIPTIONS



(a) Signal fit.



(b) Background fit.

FIGURE: $\Lambda_c \mu$ decay mode. Fits to the signal and generic MC.

Branching fraction to yield

$$\mathcal{B}(\mathcal{B} \rightarrow \text{baryon} + \text{lepton}) \quad = \quad \frac{N_{\text{sig}}}{\epsilon_{\text{sig}} \times \mathcal{B}_{\text{baryon}} \times N_{B\bar{B}} \times 2.0 \times \mathcal{B}_{\text{neut./chgB}}}$$

- N_{sig}: The number of signal events.
- ϵ_{sig}: Signal reconstruction efficiency.
- $\mathcal{B}_{\text{baryon}}$: Baryon branching fraction.
- $N_{B\bar{B}}$: The number of $B\bar{B}$ pairs.
- $\mathcal{B}_{\text{neut./chgB}}$: The branching fraction of the $\Upsilon(4S)$ to either a charged or neutral $B\bar{B}$ pair.

Systematics

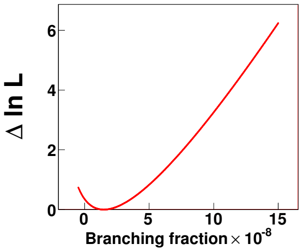
Table: Contributions to the systematic uncertainty on the branching fraction.

Contribution	Decay mode					
	$\Lambda_c^+\mu$	$\Lambda_c^+ e$	$\Lambda^0\mu$	$\Lambda^0 e$	$\bar{\Lambda}^0\mu$	$\bar{\Lambda}^0 e$
B counting (%)	0.28	0.28	0.28	0.28	0.28	0.28
Charged/neutral B's (%)	1.24	1.24	1.24	1.24	1.24	1.24
Efficiency (MC stat.) (%)	0.33	0.33	0.30	0.30	0.30	0.30
$\Lambda_{(c)}$ Branching fraction (%)	26.00	26.00	0.78	0.78	0.78	0.78
Tracking eff. (%)	0.50	0.50	0.38	0.38	0.38	0.38
PID eff. (%)	2.70	2.10	2.50	1.70	2.50	1.70
Total (%)	26.21	26.16	3.05	2.54	3.02	2.49

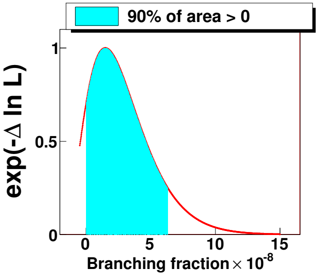
Upper limit calculation

- Perform likelihood scan as function of $\mathcal{B}$.
- Integrate under the curve *above* $\mathcal{B} = 0$.
- Interpret $\mathcal{B}$ at 90% of the area above 0 as upper limit.
- Define $\mathcal{B}_{\rm best}$ be the best solution for the branching fraction.

$$\Delta\mathcal{L} = \ln\mathcal{L}(\mathcal{B}_{\rm best}) - \ln\mathcal{L}(\mathcal{B})$$
$$y = e^{\Delta\mathcal{L}}$$



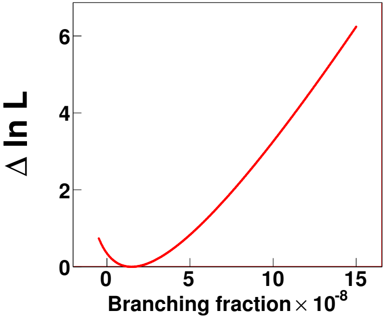
(a)



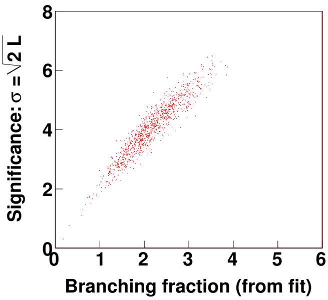
(b)

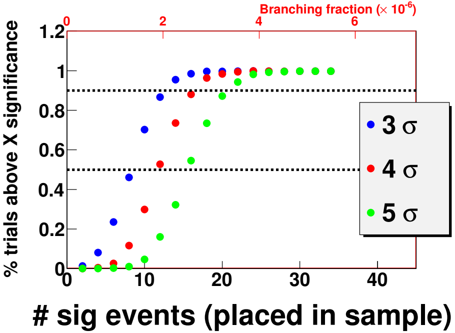
- Use ratio of likelihoods.
- Best $\mathcal{B}$ and $\mathcal{B} = 0$.

$$\sigma = \sqrt{2 \cdot (\ln \mathcal{L}(\mathcal{B}_{\mathrm{best}}) - \ln \mathcal{L}(\mathcal{B}_0))}$$



- Ran 100k's of toy studies to determine *bias* and *sensitivity*.
- Can run 1000 toy studies at a given branching fraction.
- Count what % show a $3\sigma, 4\sigma$ or 5σ observation.





Toy studies

- Summary of toy studies.
  - Possible bias?
    - Negligible, less than errors on yield.
  - Sensitivity to a 5σ discovery

$$\mathcal{B}(B \to \Lambda_c \ell) \approx 400 \times 10^{-8}$$
$$\mathcal{B}(B \to \Lambda^0 \ell) \approx 25 \times 10^{-8}$$

Simulated data

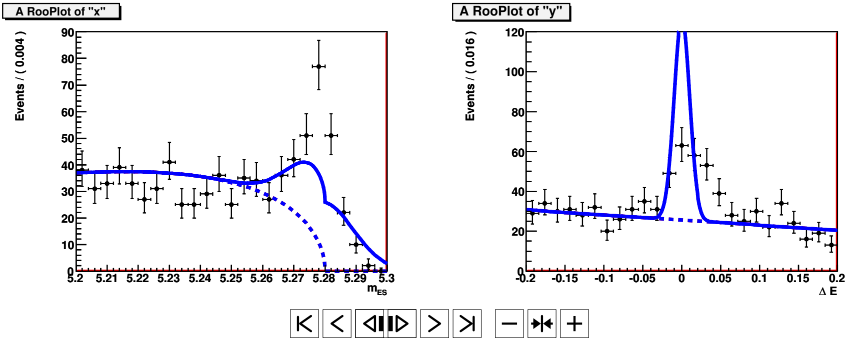
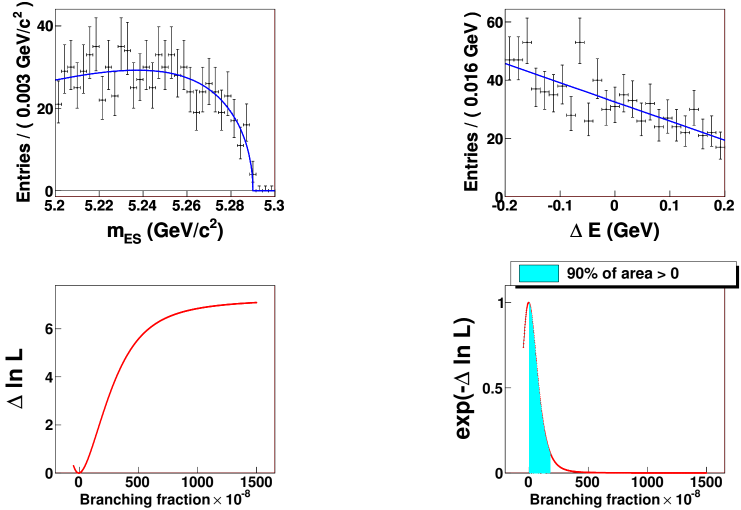
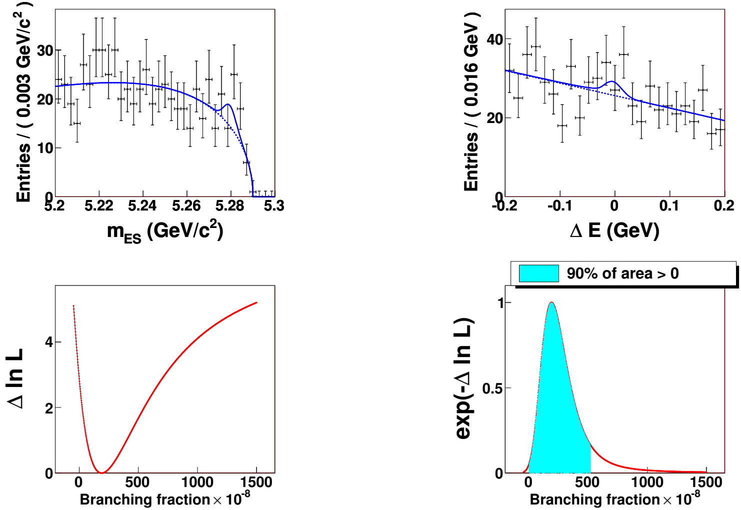


FIGURE: Fit to simulated data

$B \to \Lambda_c^+ \mu^-$

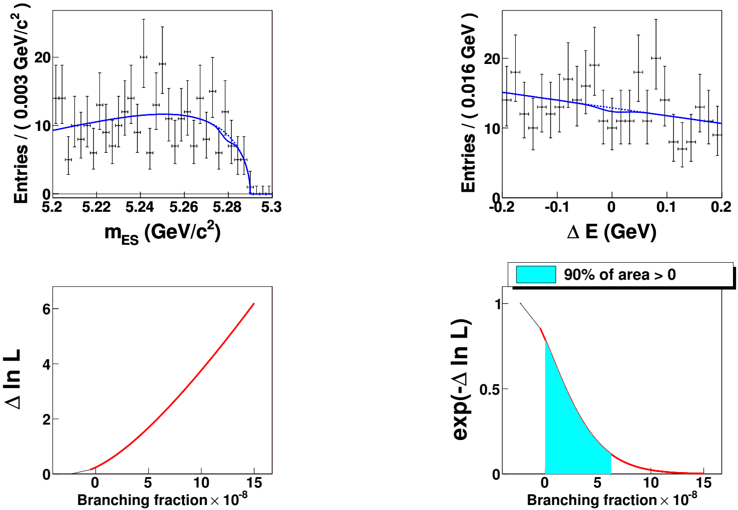


$B \to \Lambda_c^+ e^-$

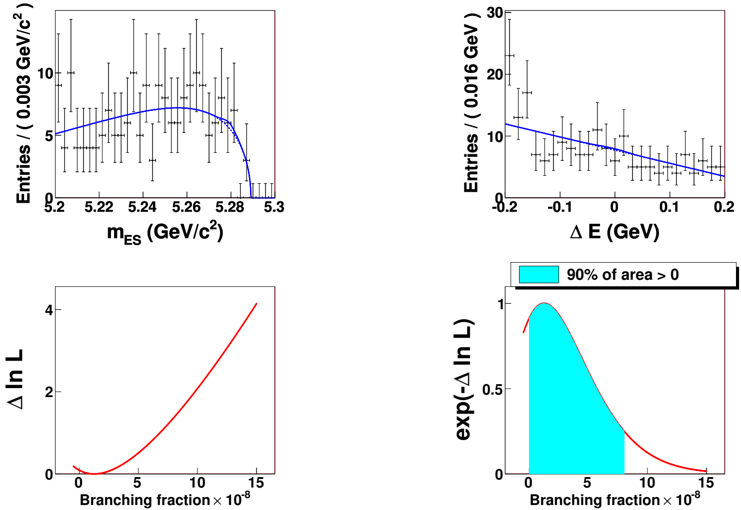


Results

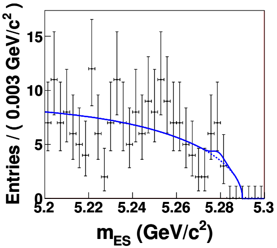
$B \to \Lambda^0 \mu^-$

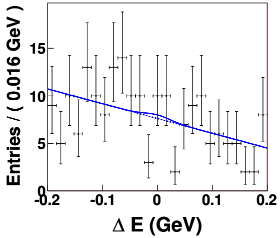


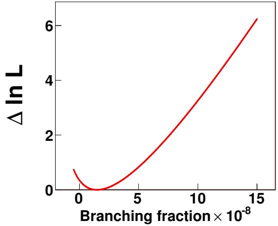
$B \to \Lambda^0 e^-$

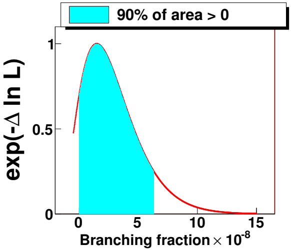


$B \to \bar{\Lambda}^0 \mu^-$

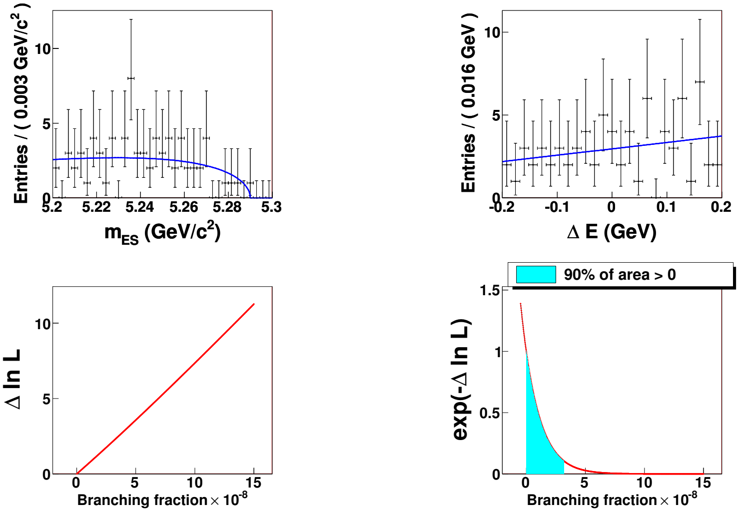








$B \rightarrow \bar{\Lambda}^0 e^-$



RESULTS

TABLE: Upper limits on branching fractions at 90% confidence level for the six decay modes.

Decay mode	Upper limit
$B^0 \to \Lambda_c^+ \mu$	170×10^{-8}
$B^0 \to \Lambda_c^+ e$	500×10^{-8}
$B^- \to \Lambda^0 \mu$	6.0×10^{-8}
$B^- \to \Lambda^0 e$	8.2×10^{-8}
$B^- \to \bar{\Lambda}^0 \mu$	6.3×10^{-8}
$B^- \to \bar{\Lambda}^0 e$	3.1×10^{-8}

Most signficant branching fraction:

$$\mathcal{B}(B^0 \to \Lambda_c^+ e^-) = (190^{+130}_{-94}) \times 10^{-8} \text{ at } 2.4\sigma$$

Summary

- Interesting physics analysis.
- Submitted to PRD-RC.

Summary

- Interesting physics analysis.
- Submitted to PRD-RC.
- Similar analysis still left to be done.
- Much more physics left in BaBar dataset!

Summary

- Interesting physics analysis.
- Submitted to PRD-RC.
- Similar analysis still left to be done.
- Much more physics left in BaBar dataset!
- **Thanks for your time!**

References

- A. D. Sakharov, Pisma Zh. Eksp. Teor. Fiz. **5**, 32 (1967) [JETP Lett. **5**, 24 (1967 SOPUA,34,392-393.1991 UFNAA,161,61-64.1991)].
- K. Nakamura [Particle Data Group], J. Phys. G **37**, 075021 (2010).
- W. S. Hou, M. Nagashima and A. Soddu, Phys. Rev. D **72**, 095001 (2005).
- A. Roodman, [physics/0312102]. HEP :: Search :: Help Powered by Invenio v1.0.0-rc0+ Problems/Questions to feedback@inspirebeta.net

Backup slides

Training variables

- Pruned discriminating variables to most sensitive that *did not have high correlations with* $\Delta E/m_{ES}$.
- Six variables.
    - B cos(θ) CM
    - B cos(θ) sphericity wrt ROE sphericity
    - B cos(θ) thrust wrt ROE thrust
    - Legendre P2
        - Legendre moments
        - Use ROE tracks, and B-thrust axis
    - Thrust all
    - R2 all
        - Ratio of Fox-Wolfram moments (0 and 2)
        - Whole event variables.
        - Both charged and neutral tracks.

Overtraining test

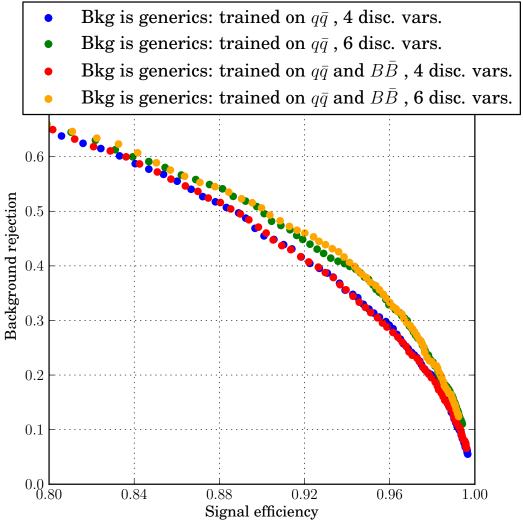
- For one mode ($\bar{\Lambda}^0 e^-$), some correlation with both $\Delta E / m_{ES}$ in *generic MC background*.
- But none in *signal MC*.
- Previous studies showed **MLP** classifier was not sensitive to *overtraining*.
- Questions remain:
  - For the Λ_c modes, which training sample gives us better sensitivity?
  - Is there a danger of creating an artificial peak?

Overtraining test

- For one mode ($\bar{\Lambda}^0 e^-$), some correlation with both $\Delta E/m_{ES}$ in *generic MC background*.
- But none in *signal MC*.
- Previous studies showed **MLP** classifier was not sensitive to *overtraining*.
- Questions remain:
  - For the Λ_c modes, which training sample gives us better sensitivity?
  - Is there a danger of creating an artificial peak?
- These are somewhat correlated.
- There may be some concern about two of the training variables: R2 (all), Thrust (all)
- Our approach:
  - For each mode run *4 TMVA training sessions.*
    - Train using $q\bar{q}$ and 4 variables.
    - Train using $q\bar{q}$ and 6 variables.
    - Train using $q\bar{q} + B\bar{B}$ and 4 variables.
    - Train using $q\bar{q} + B\bar{B}$ and 6 variables.
  - Generate *sig eff. vs. bkg. rej.* using signal MC and all the generic MC.
  - Generate *sig eff. vs. bkg. rej.* using signal MC and sideband data for comparison.

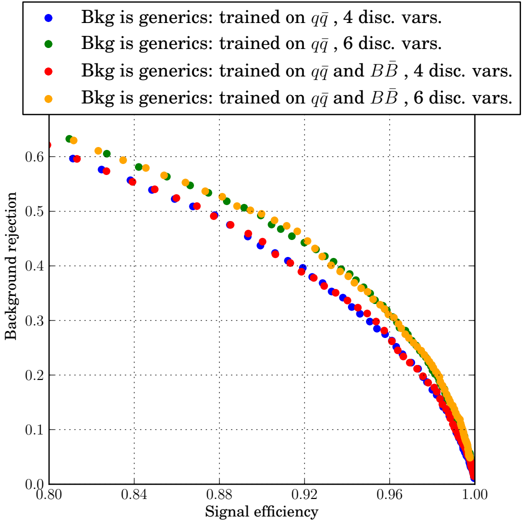
Training sample/variable choice

Curves generated with generic MC and sideband data: $\Lambda_c \mu^-$ (zoomed in)



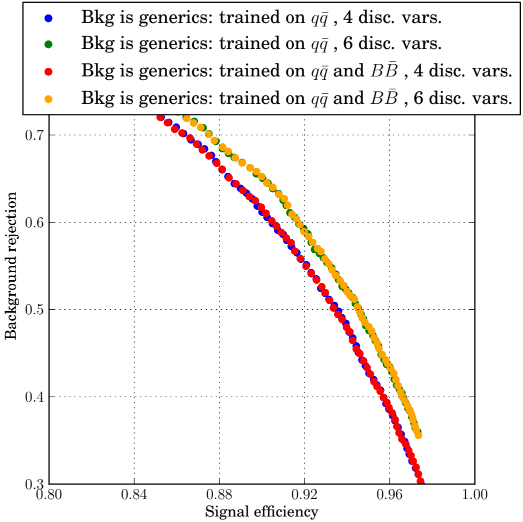
Training sample/variable choice

Curves generated with generic MC and sideband data: $\Lambda_c e^-$ (zoomed in)

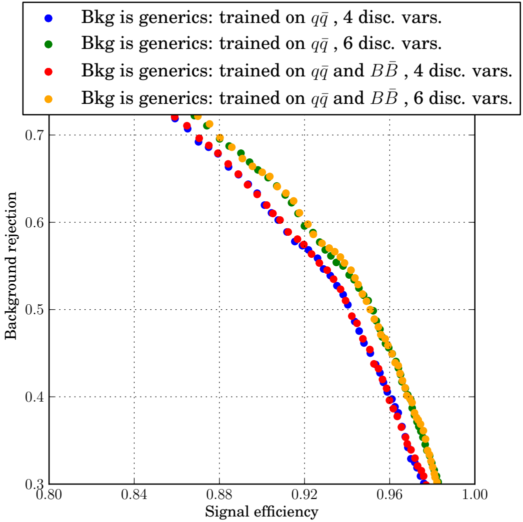


Training sample/variable choice

Curves generated with generic MC and sideband data: $\Lambda^0\mu^-$

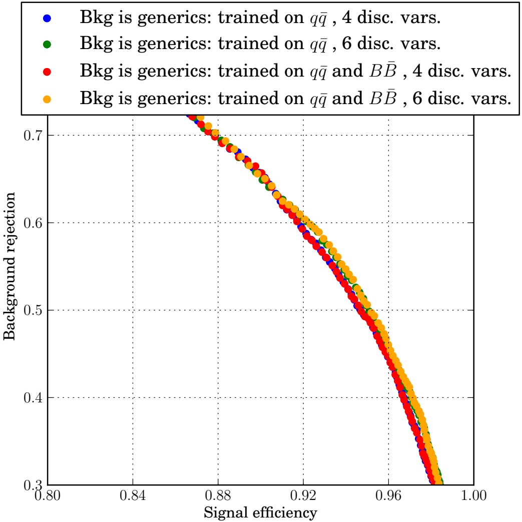


Curves generated with generic MC and sideband data: $\Lambda^0 e^-$



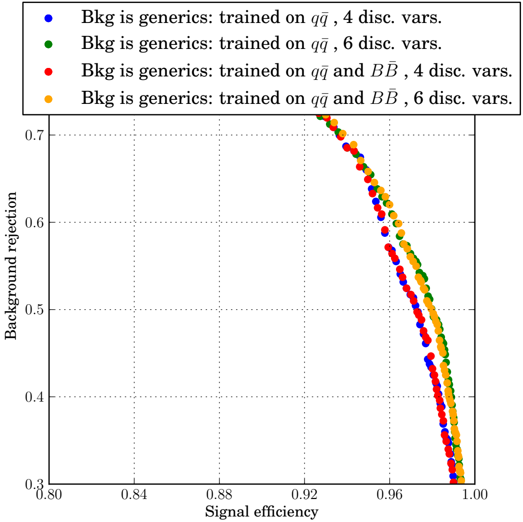
Training sample/variable choice

Curves generated with generic MC and sideband data: $\bar{\Lambda}^0 \mu^-$



Training sample/variable choice

Curves generated with generic MC and sideband data: $\bar{\Lambda}^0 e^-$

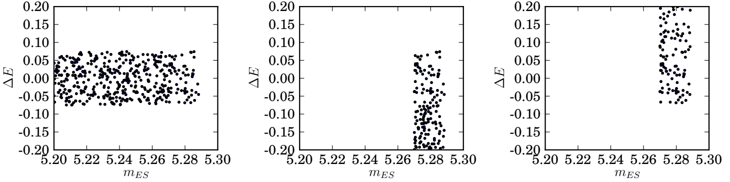


Training sample/variable choice

- At this point it seems like I can use any of the combinations.
- Little difference amongst them all.
- However...are any of these combinations producing an artifical peak?
- Can see if the MLP (neural net) output is correlated with $m_{ES}/\Delta E$.
- But how to do that?
- Look at *correlation coefficents* between neural net output and $\Delta E/m_{ES}$ plane.
- How to isolate *signal* region of $\Delta E/m_{ES}$ plane?

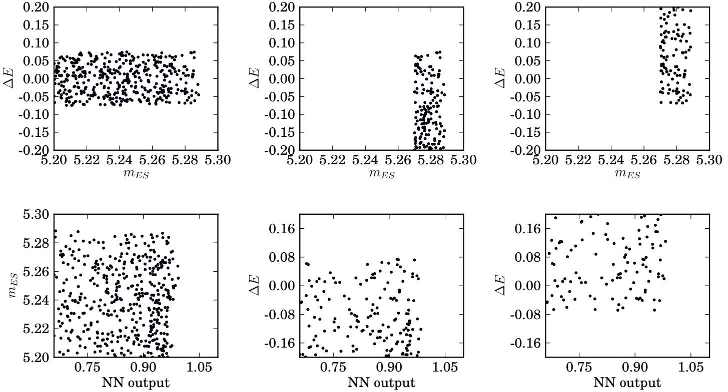
Correlation coefficients

Look in regions of $\Delta E/m_{ES}$ plane.



Correlation coefficients

Look in regions of $\Delta E/m_{ES}$ plane.



- Danger would be a positive, positive, negative correlation, respectively, for these three regions.
- But how to determine if correlation is **significant**?

NEXT STEPS

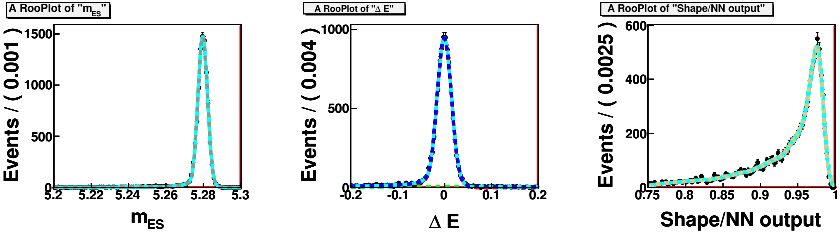
- Return to fitting procedure.
- Questions to answer with toy studies.
  - How the upper limit be calculated? review...
  - How will significance of "signal" be determined? review...
  - What is my sensitivity to a signal? NEW!
  - What is the chance of a false positive? NEW!
  - Are there any inherent bias' in the fitting procedure? NEW!

Next steps

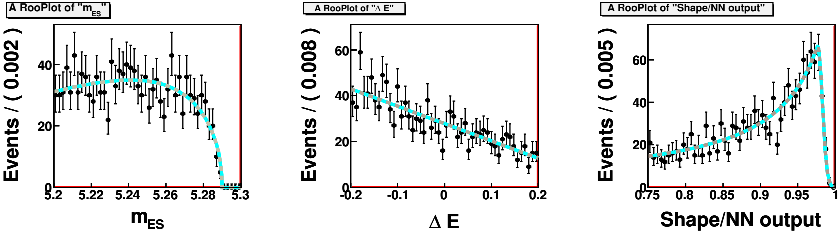
- Return to fitting procedure.
- Questions to answer with toy studies.
    - How the upper limit be calculated? review...
    - How will significance of "signal" be determined? review...
    - What is my sensitivity to a signal? NEW!
    - What is the chance of a false positive? NEW!
    - Are there any inherent bias' in the fitting procedure? NEW!
- Take a look at some trials...

$B \to \Lambda_C^+ \mu$

- Signal: Fit

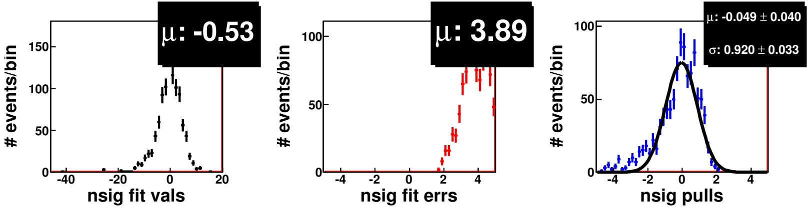


- Background: Fit



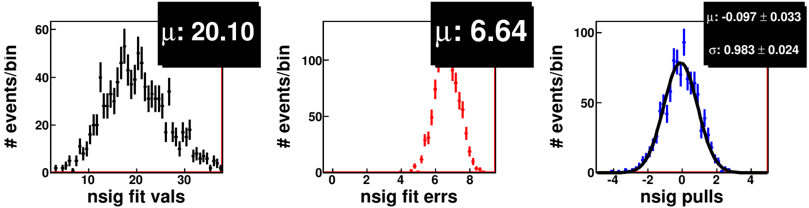
$B \to \Lambda_C^+ \mu$

- 1000 trials
- 1400 background (Poisson fluctuated)
- 0 signal

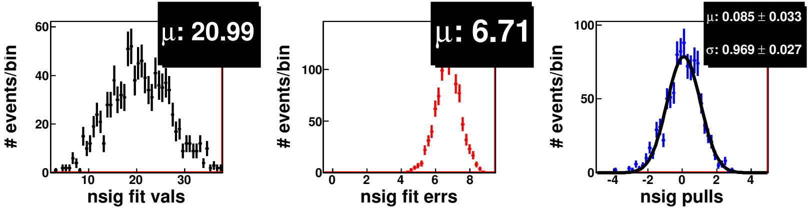


$B \to \Lambda_C^+ \mu$

- 1000 trials
- 1400 background (Poisson fluctuated)
- 20 signal (**Toy**, Poisson fluctuated and fixed number)



- 1000 trials
- 1400 background (Poisson fluctuated)
- 20 signal (**Full simulation**, Poisson fluctuated and fixed number)

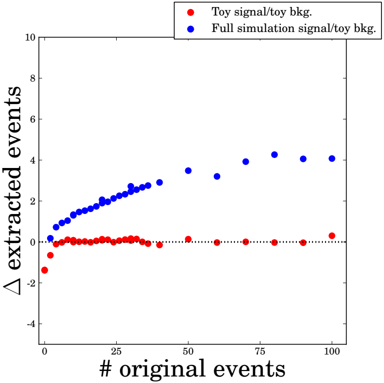


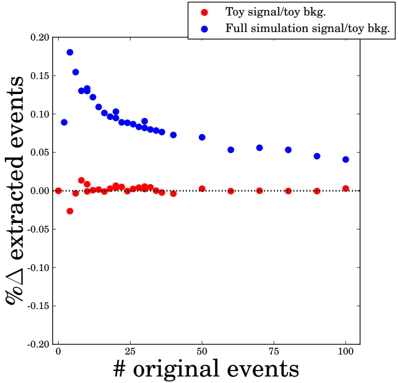
$B \to \Lambda_C^+ \mu$

- How to summarize this?
- As a function of **number of embedded signal events** plot:
  - Number of extracted (fit) events for toy signal.
  - Number of extracted (fit) events for fully simulated signal.
- Because of issues that have only recently come up, I will show this for the *full 3D fit* and for a *2D fit to the* $\Delta E/m_{ES}$ *plane only*.
- Both of these have a loose cut on the NN output ($\approx$ 90% signal efficiency)

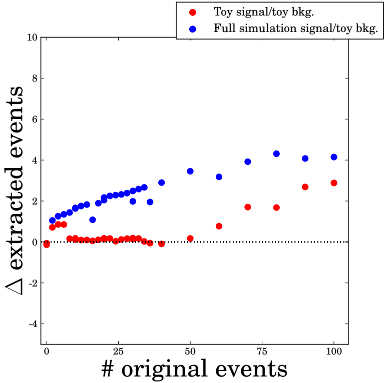
Sensitivity

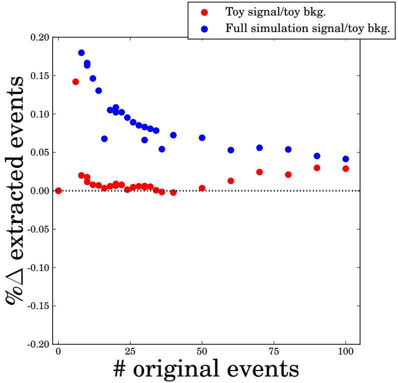
$\Lambda_c^+ \mu^-$ 3D fit





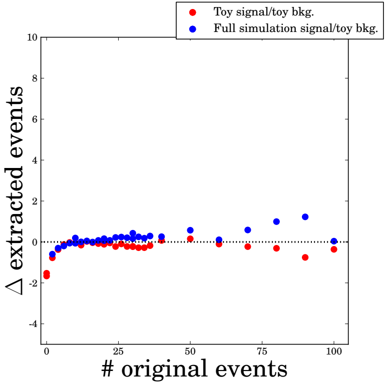
$\Lambda_c^+ \mu^-$ 2D fit

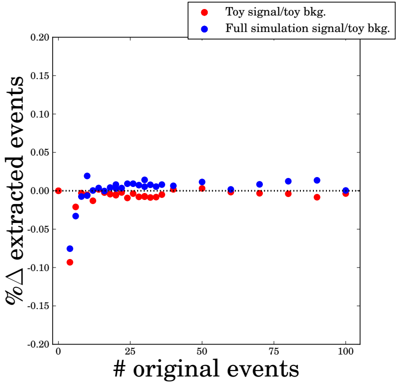




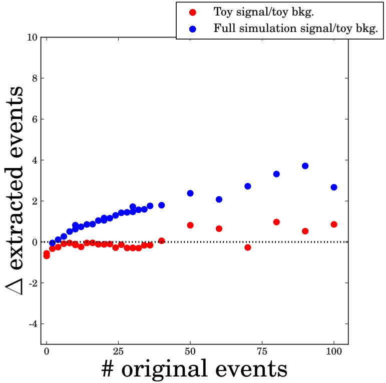
Sensitivity

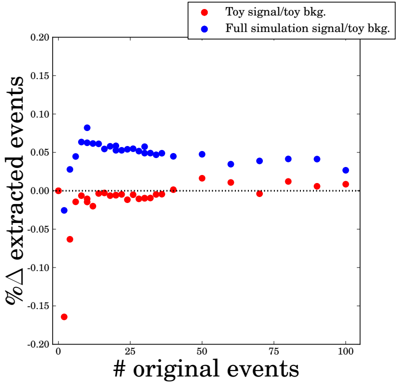
$\Lambda_c^+ e^-$ 3D fit





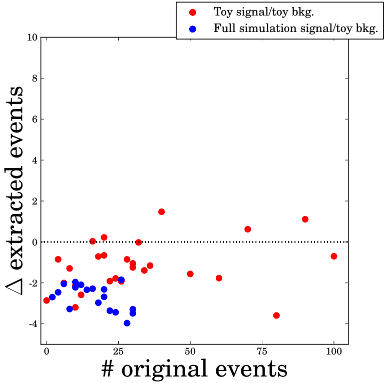
$\Lambda_c^+ e^-$ 2D fit

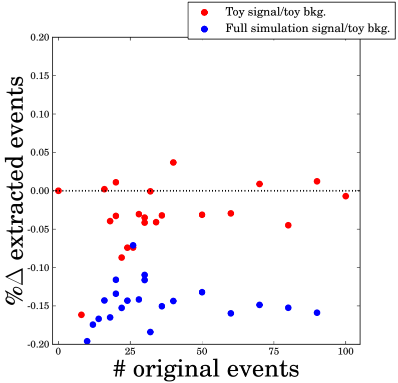




Sensitivity

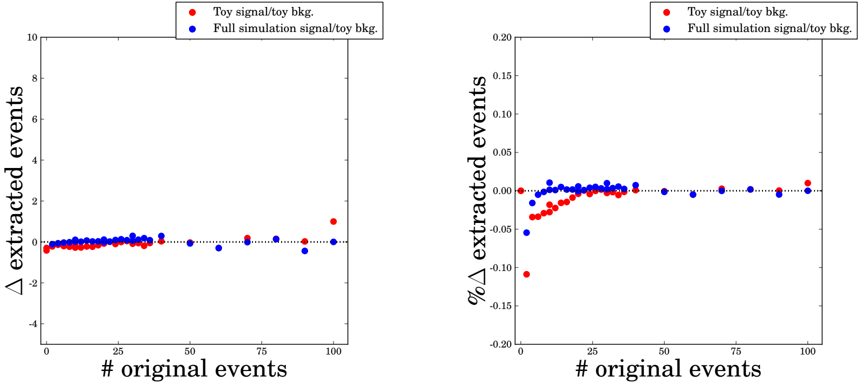
$\Lambda^0 \mu^-$ 3D fit



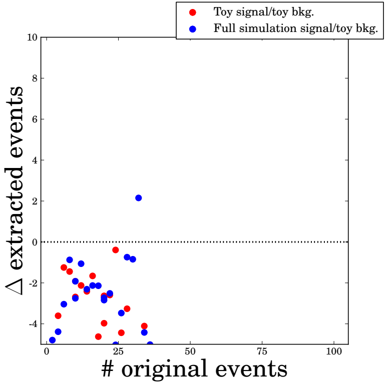


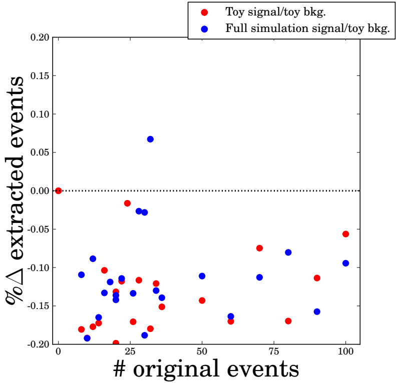
Sensitivity

$\Lambda^0\mu^-$ 2D fit



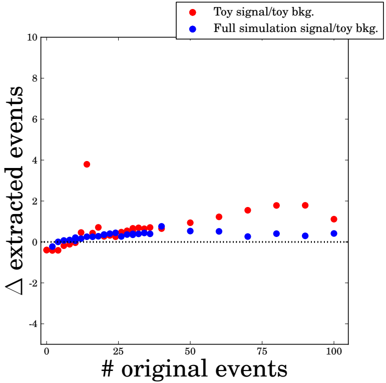
$\Lambda^0 e^-$ 3D fit

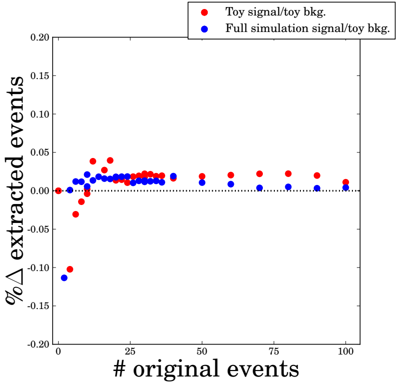




SENSITIVITY

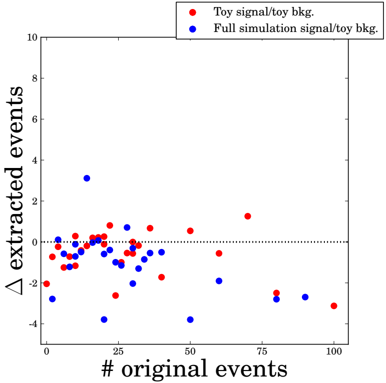
$\Lambda^0 e^-$ 2D fit

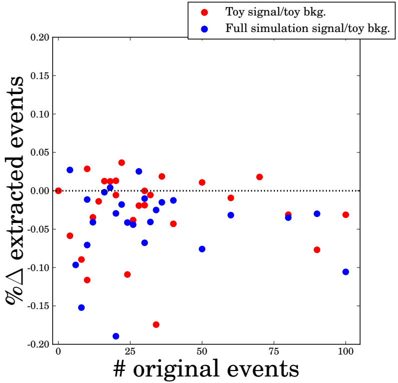




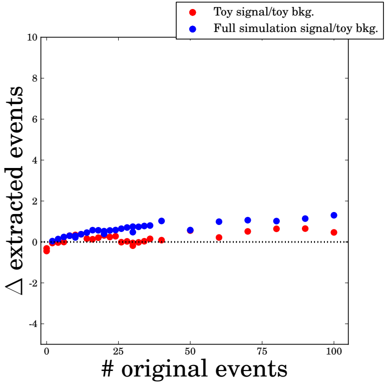
Sensitivity

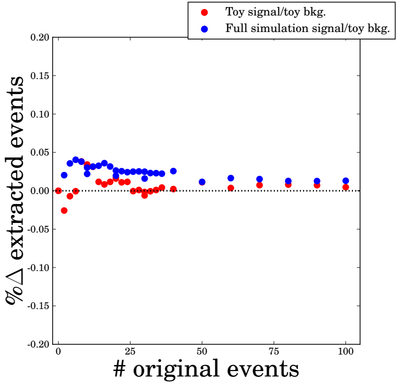
$\bar{\Lambda}^0\mu^-$ 3D fit



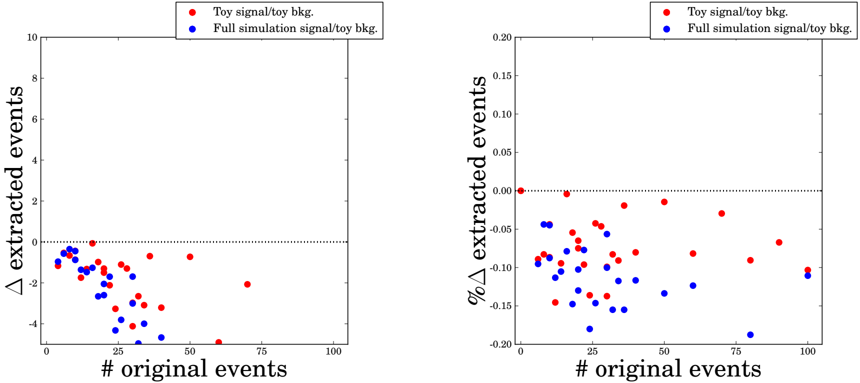


$\bar{\Lambda}^0\mu^-$ 2D fit

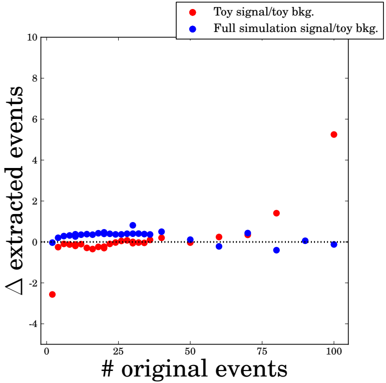


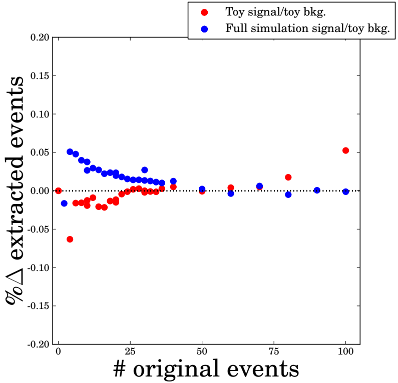


$\bar{\Lambda}^0 e^-$ 3D fit



$\bar{\Lambda}^0 e^-$ 2D fit



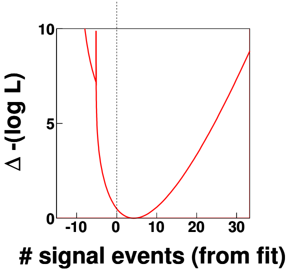


$B \to \Lambda_C^+ \mu$

- Λ_c modes.
  - $\Lambda_c^+ \mu^-$: (+5,+10)% on extracted signal events (both 2D and 3D).
  - $\Lambda_c^+ e^-$: (+3,6)% on extracted signal events (2D only).
- Λ^0 modes are unstable in 3D fits. About 25% fits don't converge.
  - $\Lambda^0 \mu^-$: (-3,0)% on extracted signal events (2D).
  - $\Lambda^0 e^-$: (-10,+2)% on extracted signal events (2D).
  - $\bar{\Lambda}^0 \mu^-$: (+1,+4)% on extracted signal events (2D).
  - $\bar{\Lambda}^0 e^-$: (0,+5)% on extracted signal events (2D).
- How does this map onto sensitivity?
- Review UL and significance calculation...

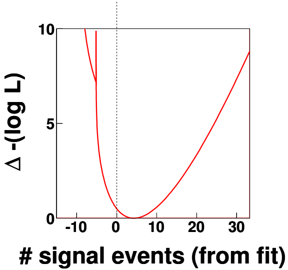
Upper limit

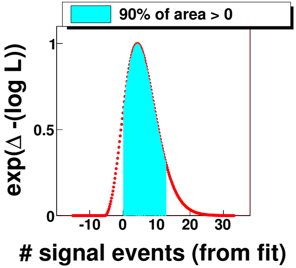
- Likelihood scan.
  - n = Number of signal events.
  - n_0 = Number of signal events, best solution.
  - $y = -\ln\mathcal{L}(n) - -\ln\mathcal{L}(n_0)$



Upper limit

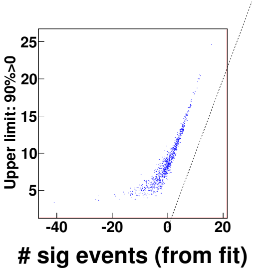
- Likelihood scan.
  - n = Number of signal events.
  - n_0 = Number of signal events, best solution.
  - $y = -\ln \mathcal{L}(n) - -\ln \mathcal{L}(n_0)$
- Upper limit
  - Area under likelihood curve.
  - $y = e^{-\ln \mathcal{L}(n) - -\ln \mathcal{L}(n_0)}$
  - Integrate under curve to find total area *above* 0.
  - Integrate under curve up above 0 up to to 90% of this area.
  - This is the UL at 90% confidence.
  - Plan to publish curve for others to draw their own conclusions.



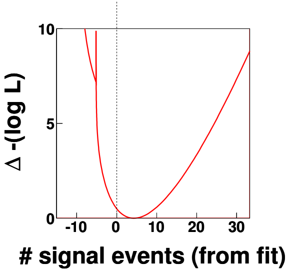


Upper limit

- 1000 studies
  - Background from PDF (1400 events, Poisson fluctuated)
  - No signal events.

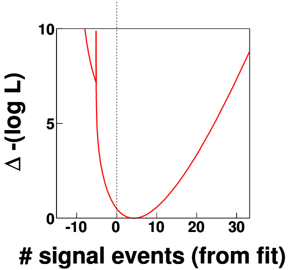


- Likelihood scan.
  - n = Number of signal events.
  - n_0 = Number of signal events, best solution.
  - $y = -\ln\mathcal{L}(n) - -\ln\mathcal{L}(n_0)$



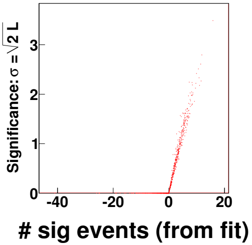
Significance of signal

- Likelihood scan.
  - n = Number of signal events.
  - n_0 = Number of signal events, best solution.
  - $y = -\ln \mathcal{L}(n) - -\ln \mathcal{L}(n_0)$
- Significance
  - *If there is no signal in nature, what are the odds that the background fluctuates to give a false peak?*
  - How many σ's is the extracted signal yield from **0**.
  - Assume $\mathcal{L}$ is Gaussian in region of best solution.
  - $\mathcal{L}(n) = e^{-n^2/2\sigma^2}$
  - $\sigma = \sqrt{2(\ln \mathcal{L}(0) - \ln \mathcal{L}(n_0))}$

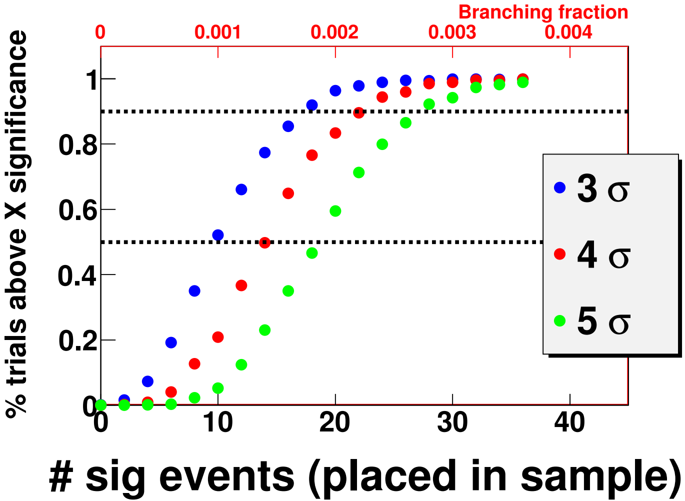


Significance of signal

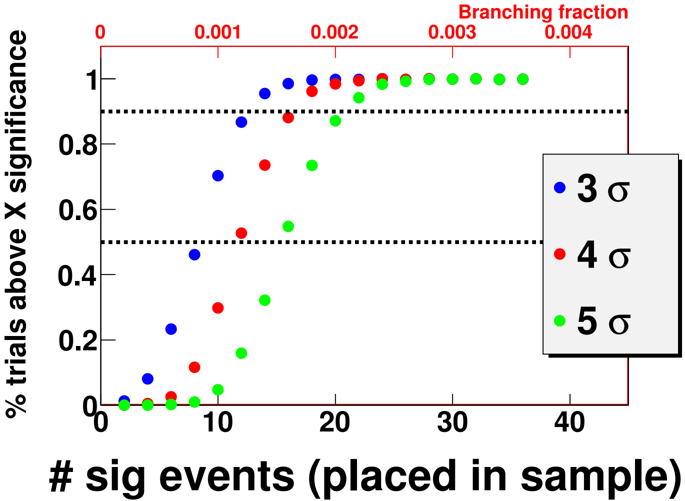
- 1000 studies
  - Background from PDF (800 events, Poisson fluctuated)
  - No signal events.
- Negative yields have 0 significance.
- Number of trials greater than some significance?
  - 3σ: 1
  - 4σ: 0
  - 5σ: 0
- APOLOGIES: In upcoming plots, ignore **RED** axis at top. Currently a plotting artifact that will eventually be calculated correctly.



$\Lambda_c^+ \mu^-$ Toy signal (3D),

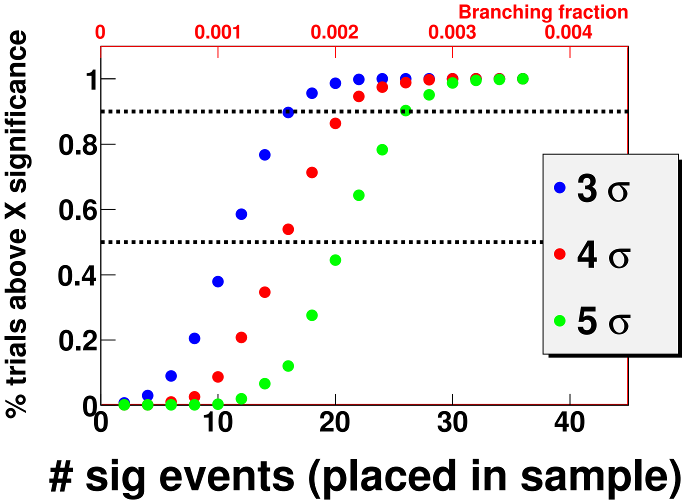


$\Lambda_c^+ \mu^-$ Toy signal (3D), Full simulation (3D),



Sensitivity

$\Lambda_c^+ \mu^-$ Toy signal (3D), Full simulation (3D), Full simulation (2D)

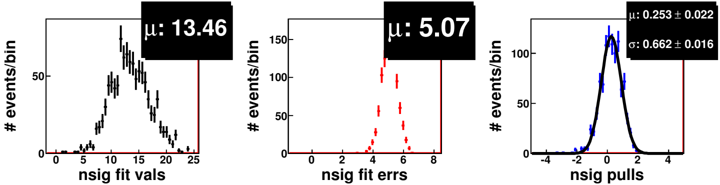


Sensitivity

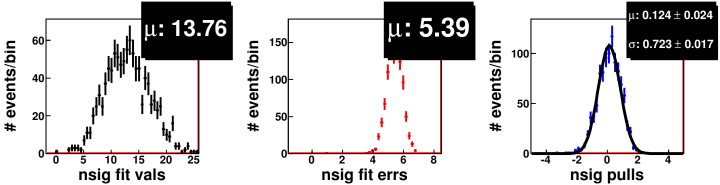
- The 3D and 2D fit, on average, extract same number of events, but the 2D fit has *less sensitivity*.
- Remind ourselves that the significance depends not just on the number of signal events extracted by the fit, but also the contours of the likelihood function itself.

Sensitivity

$\Lambda_c^+\mu^-$ Full simulation for signal, 12 embedded events (not Poisson fluctuated), 3D fit

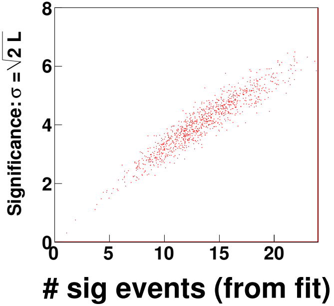


$\Lambda_c^+\mu^-$ Full simulation for signal, 12 embedded events (not Poisson fluctuated), 2D fit

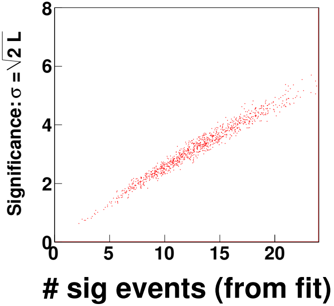


Sensitivity

$\Lambda_c^+\mu^-$ Full simulation for signal, 12 embedded events (not Poisson fluctuated), 3D fit



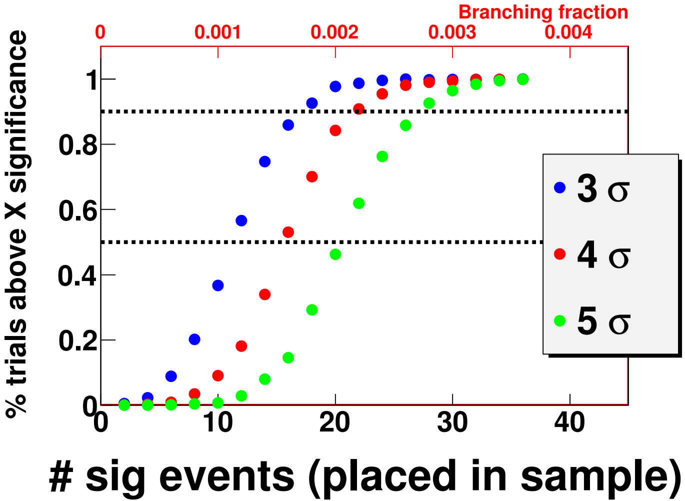
$\Lambda_c^+\mu^-$ Full simulation for signal, 12 embedded events (not Poisson fluctuated), 2D fit



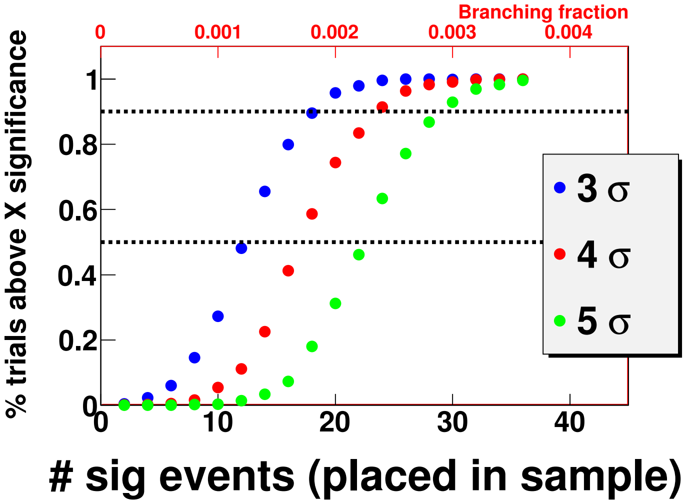
Sensitivity

- Understand the difference.
- Implication?
- Define sensitivity as number of events (branching fraction) that 90% of the time, gives $> 5\sigma$ observation.
- Fitting with 2D loses *only some* sensitivity (5-15%).
- Other modes?

$\Lambda_c^+ e^-$ Full simulation (3D),



$\Lambda_c^+ e^-$ Full simulation (3D), Full simulation (2D)

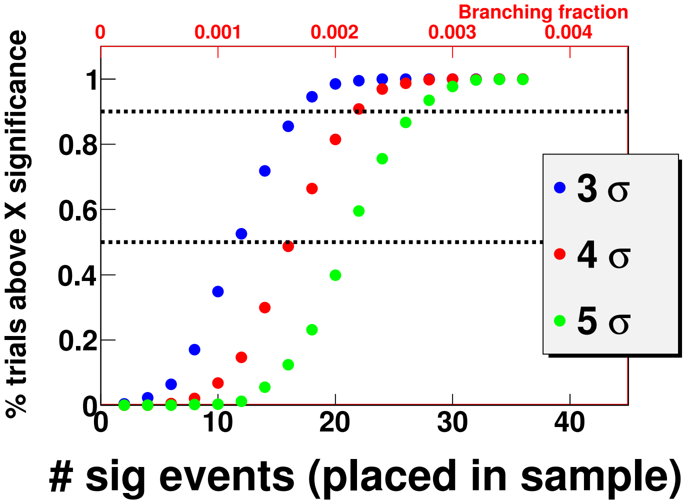


Sensitivity

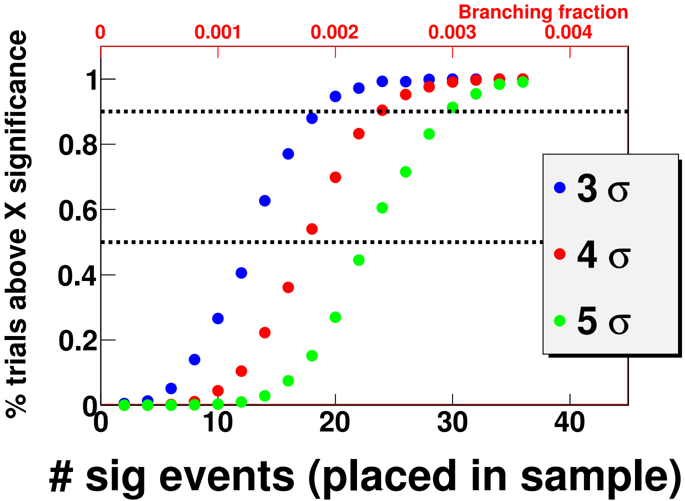
- For Λ^0 modes, show only 2D fits (full simulation)

Sensitivity

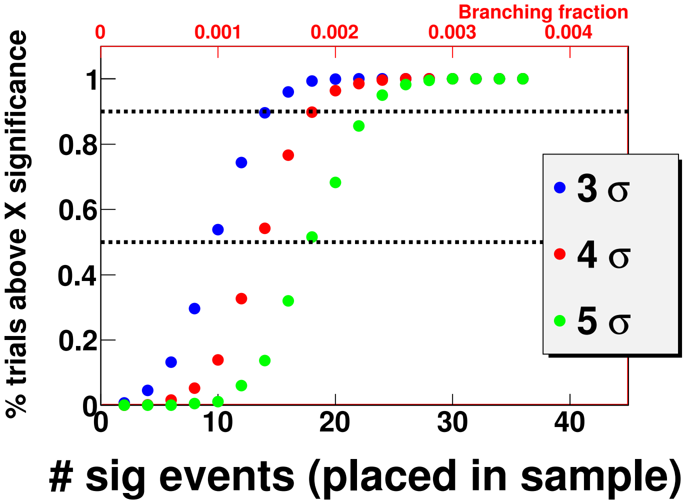
$\Lambda^0\mu^-$ Full simulation (2D)



$\Lambda^0 e^-$ Full simulation (2D)



$\bar{\Lambda}^0\mu^-$ Full simulation (2D)



$\bar{\Lambda}^0 e^-$ Full simulation (2D)