Korea Institute of
Science and Technology Information
www.kisti.re.kr

The Study for $B^0 \to \ell^{\pm}\tau^{\mp}$ Decays at Belle Experiment

Kyungho Kim

khhkim4@kisti.re.kr

May 30th, 2022

Contents

Introductions

Movitations & Details

Motivation and Details

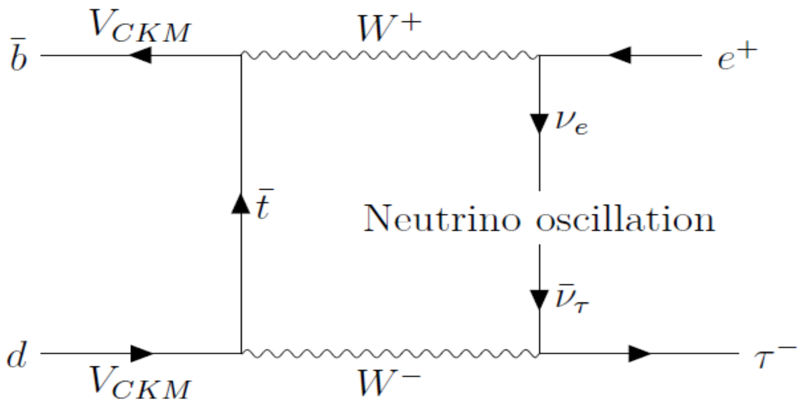
Lepton flavor violation in $B^0 \to \ell^\pm \tau^\mp$

- Forbidden in the Standard Model
- Predicted to occur in 'beyond the Standard Model' theories
  - Neutrino oscillation
    - $\Gamma(B^0 \to e^+\tau^-) = |V_{td}| \times |V_{tb}| \times \Gamma(W \to e\nu_e) \times \Gamma(W \to \tau\nu_\tau) \times \Gamma(\nu \text{ oscillation})$
      $= 1.06 \times 10^{-3} \times \Gamma(\nu \text{ oscillation})$
    - B$(B_s \to \ell^\pm \tau^\mp)$ significant constraints of neutrino mixing : sensitivity of 10^{-9} [1]

	B^0	$\to$	e^-	τ^+
e lepton number	0	$\neq$	1	0
τ lepton number	0	$\neq$	0	-1
	B^0	$\to$	μ^-	τ^+
μ lepton number	0	$\neq$	1	0
τ lepton number	0	$\neq$	0	-1

Research for $B^0 \to \ell^\pm \tau^\mp$

- BaBar collaboration (2008) [2]
  - Using hadronic tagging method with 378×10^6 $B\bar{B}$ pairs
  - $\Gamma(B^0 \to e^\pm \tau^\mp) < 2.8 \times 10^{-5}$
  - $\Gamma(B^0 \to \mu^\pm \tau^\mp) < 2.2 \times 10^{-5}$
- Belle collaboration (by Hulya Atmacan, 2021) [3]
  - Using hadronic tagging method(FR) with 771×10^6 $B\bar{B}$ pairs
  - $\Gamma(B^0 \to e^\pm \tau^\mp) < 1.6 \times 10^{-5}$
  - $\Gamma(B^0 \to \mu^\pm \tau^\mp) < 1.5 \times 10^{-5}$
- My research : using semileptonic tagging method(FEI) with 771×10^6 $B\bar{B}$ pairs



Feynman diagram of $B^0 \to e^+\tau^-$ with neutrino oscillation

Motivation and Details

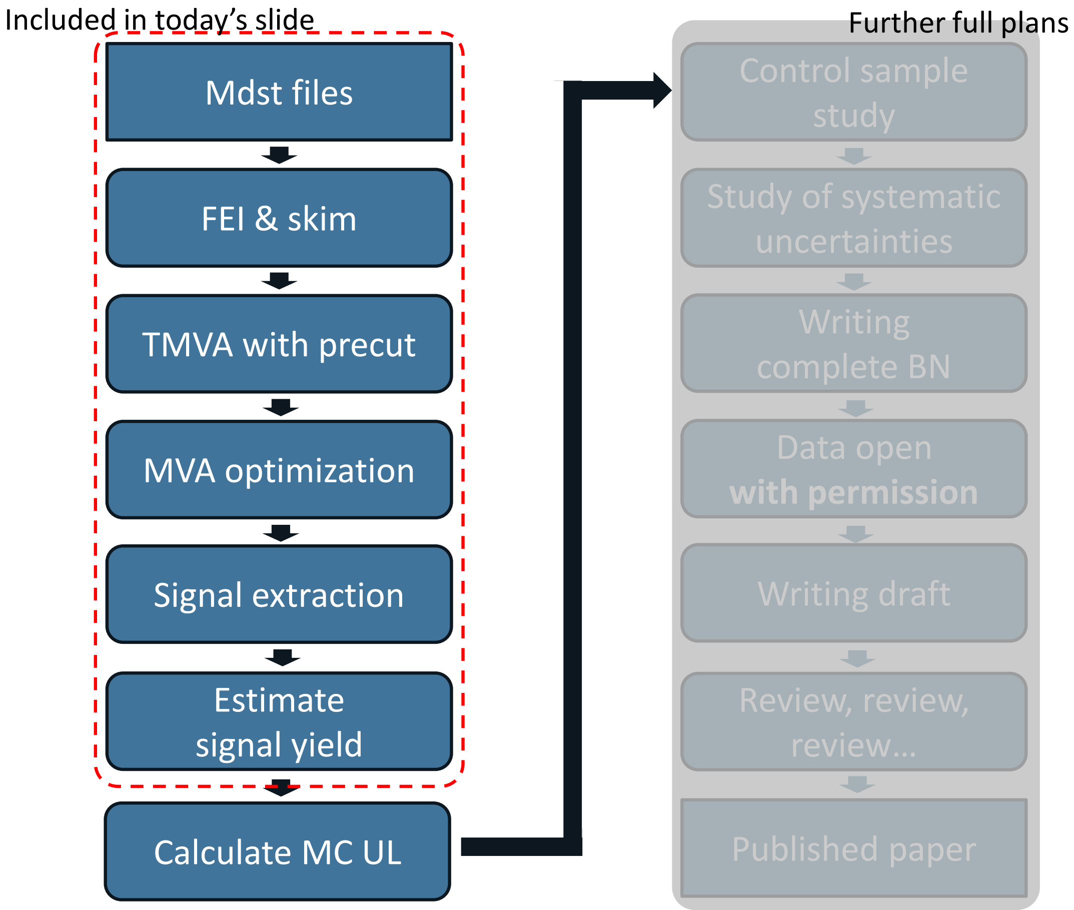
Selection of decay mode

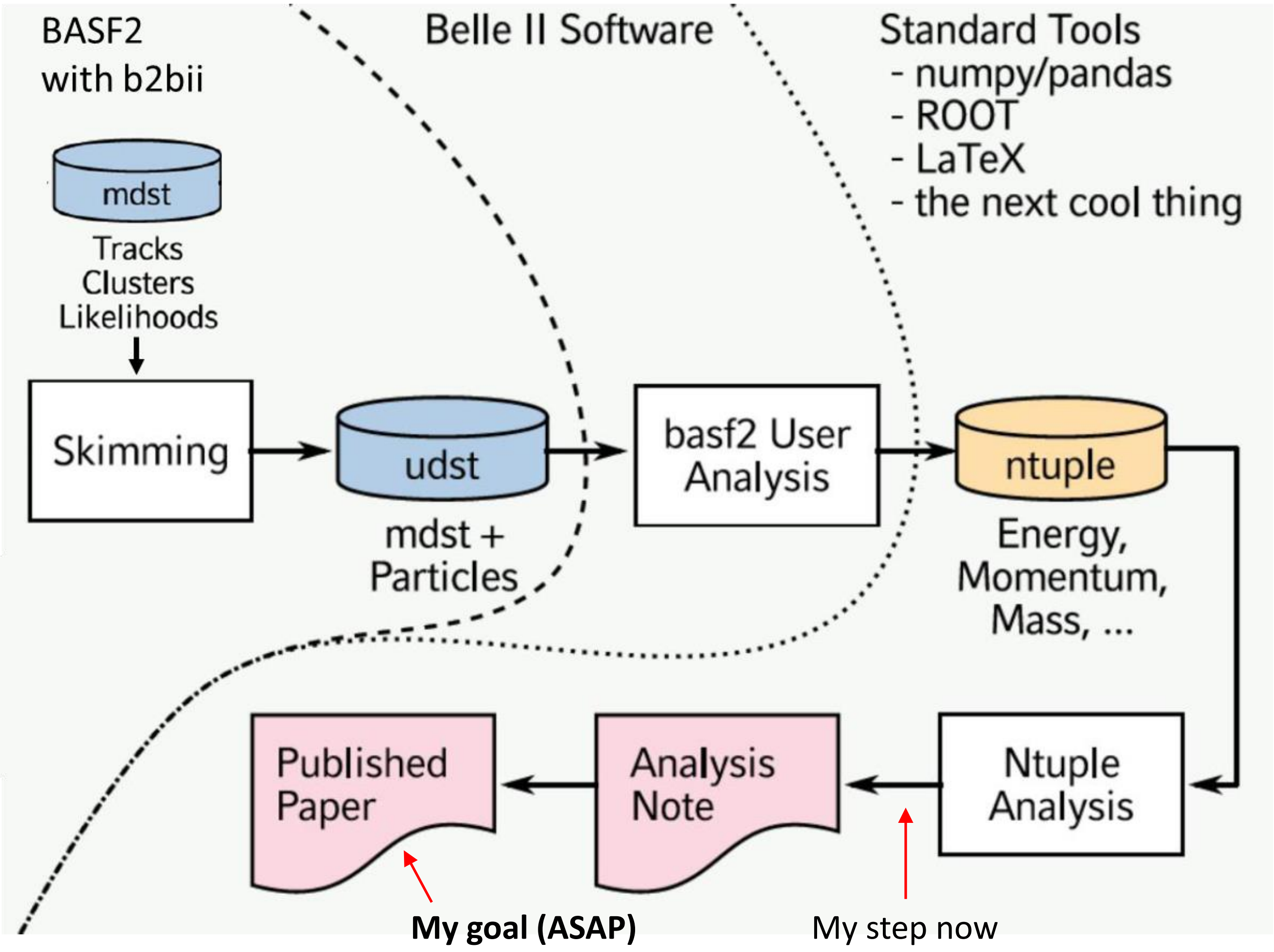
- Decay modes that τ to 1 lepton and 2 neutrinos : taking 35.2% of τ sub-decay
  - Expected to provide high purity and distinct kinematic signature with 2 leptons in final state
  - Considered mode
    - $B^0 \to e^+\tau^-(\tau^- \to e^-\overline{\nu_e}\nu_\tau)$ (e-e mode), $B^0 \to e^+\tau^-(\tau^- \to \mu^-\overline{\nu_\mu}\nu_\tau)$ (e-μ mode)
    - $B^0 \to \mu^+\tau^-(\tau^- \to e^-\overline{\nu_e}\nu_\tau)$ (μ-e mode), $B^0 \to \mu^+\tau^-(\tau^- \to \mu^-\overline{\nu_\mu}\nu_\tau)$ (μ-μ mode)

Amounts of samples for MC analysis

- Signal MC : 20M events (10 sets×2M/set) for each mode (Belle)
- Background samples (Belle)
  - Generic MC: $e^+e^- \to B\bar{B}$ (10 stream)
  - Generic MC: $e^+e^- \to q\bar{q}$ (6 stream)
  - rareB : $b \to s, d$ & leptonic decays (50 stream)
  - ulv : $b \to u\ell\nu$ decays (20 stream)

Process of the Analysis

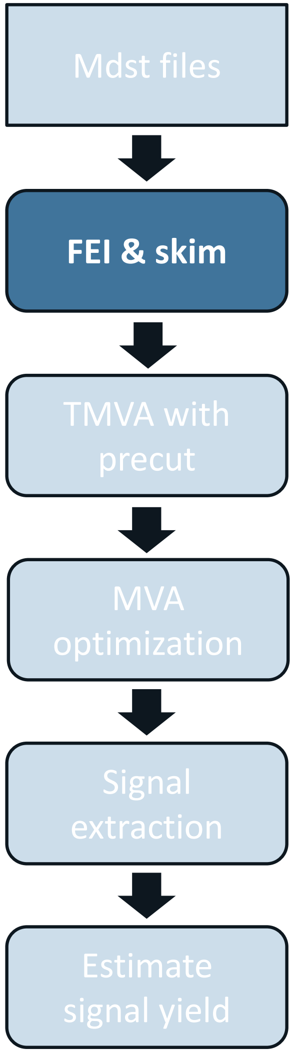




Full Event Interpretation

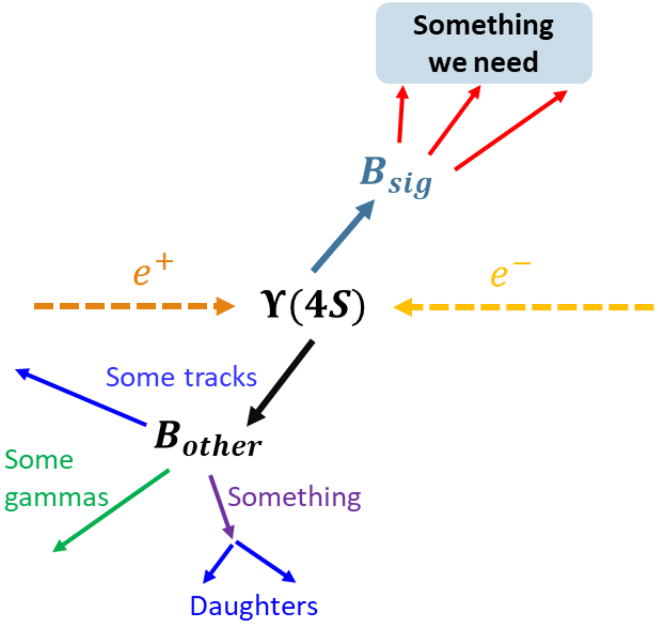
& Skim

How to Reconstruct Signal?

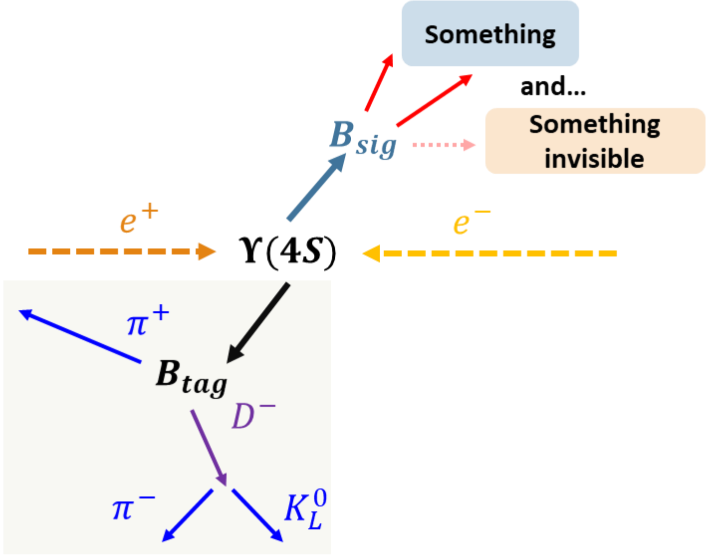


B meson analysis : reconstructing B inclusively or exclusively

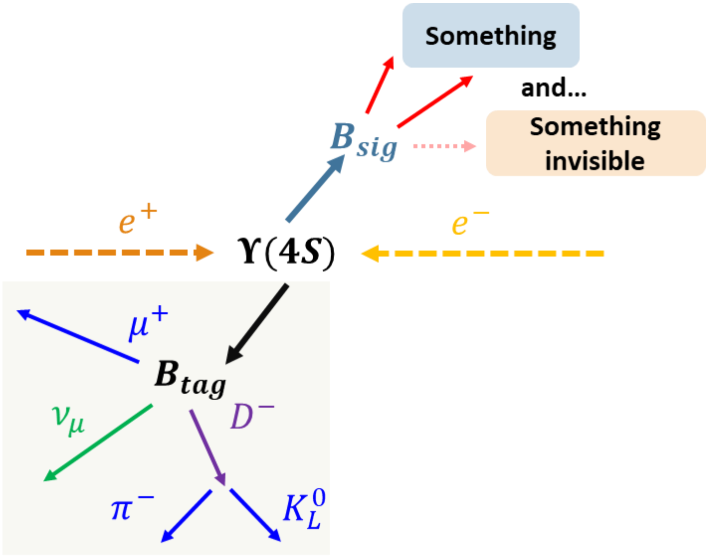
- Exclusive attempt : reconstructing one B at tag-side (B_{tag}) and next reconstruct signal-side B
- Inclusive attempt : only reconstructing B at signal-side



Schematic view of inclusive B study



Schematic view of B tagging via $B \to D\pi$ channel



Schematic view of B tagging via $B \to D\mu\nu_\mu$ channel

Inclusive vs. Exclusive B Studies

- Mdst files
- **FEI & skim**
- TMVA with precut
- MVA optimization
- Signal extraction
- Estimate signal yield

Tagging techniques and properties

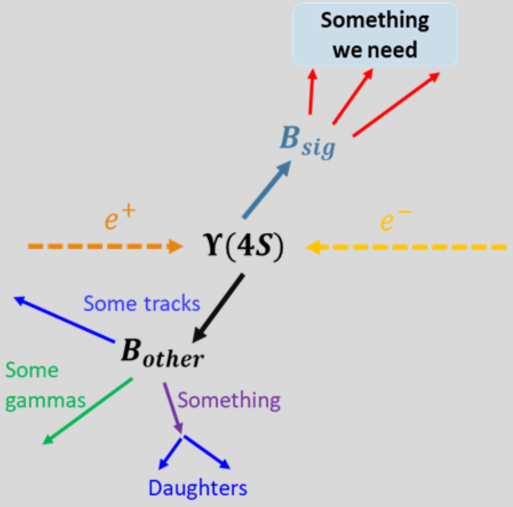
Purity →

← Efficiency

Inclusive

$B \rightarrow anything$

$\epsilon \approx O(100\%)$



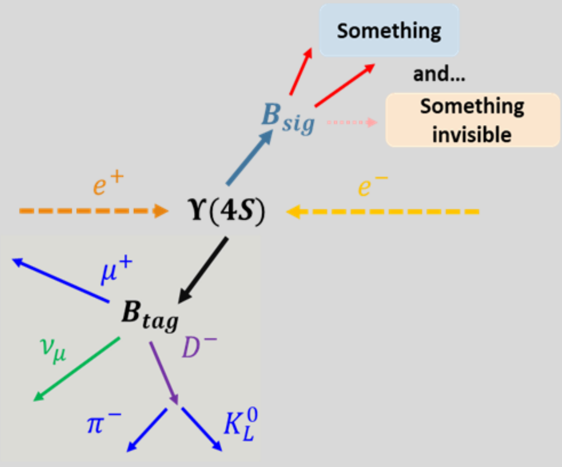
Large statistics

Large background (and candidates!)

Semilpetonic

$B \rightarrow D^{(*)} \ell \nu_\ell$

$\epsilon \approx O(1\%)$



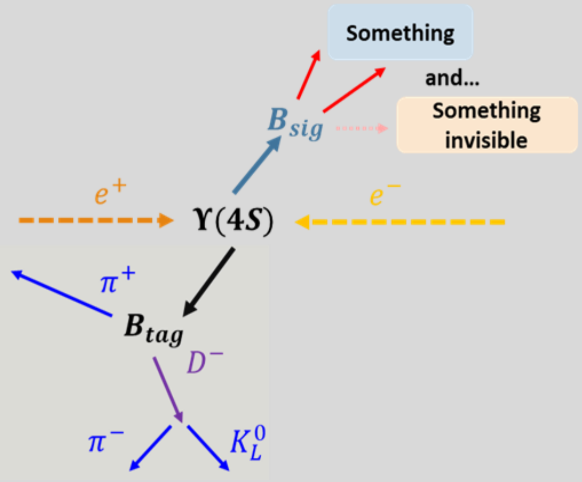
Mid-range reconstruction efficiency

Less information about B_{tag} due to neutrinos

Hadronic

$B \rightarrow$ hadrons

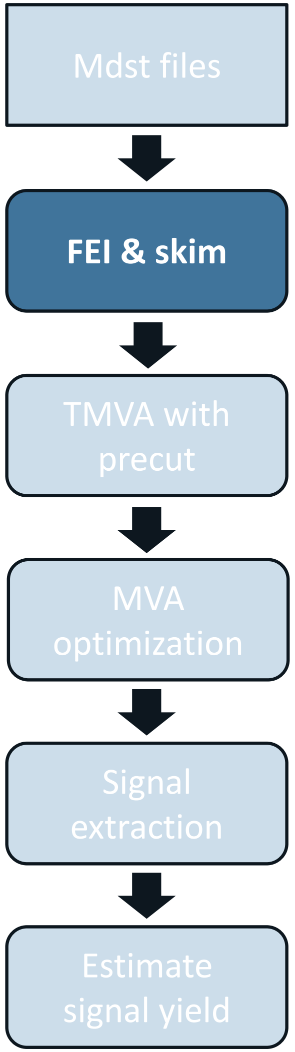
$\epsilon \approx O(0.1\%)$



Cleaner sample and signal kinematics

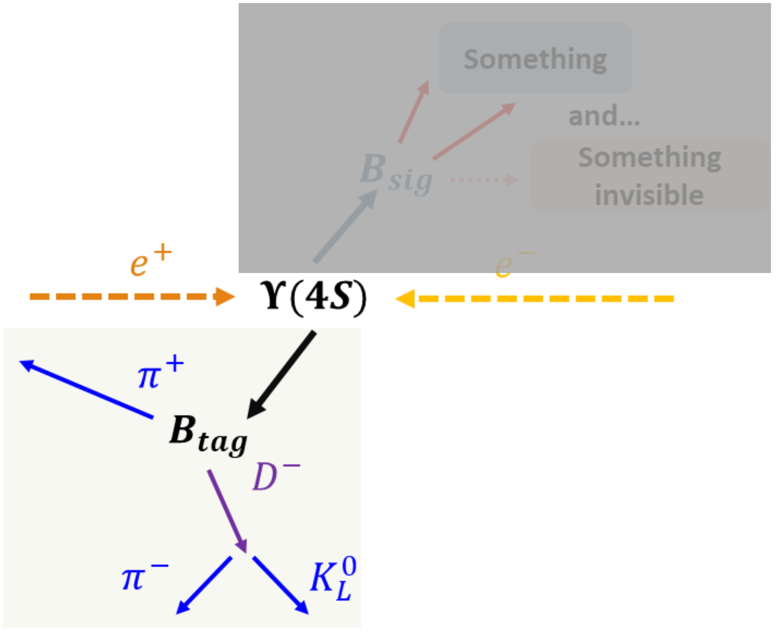
Lower tag-side efficiency

Reconstructing B with Machine Learning

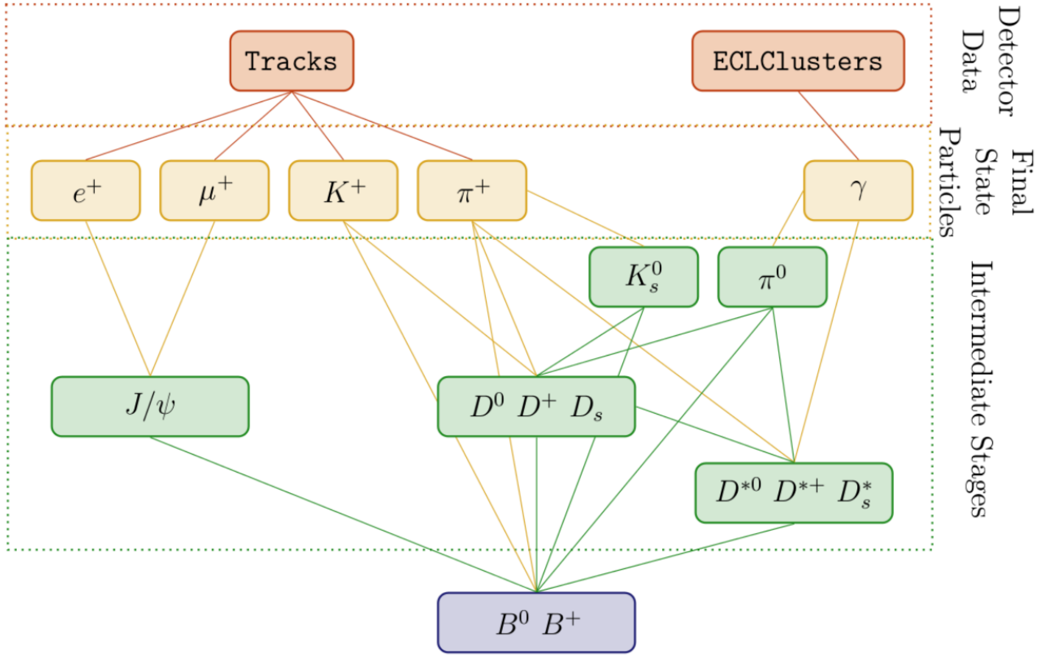


Full Event Interpretation [4]

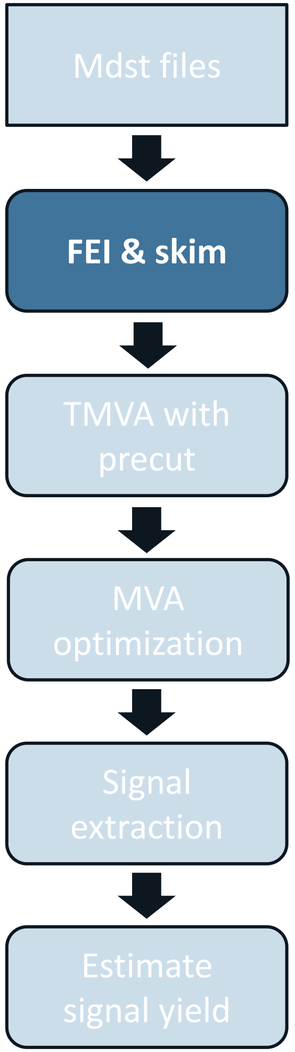
- B tagging with MVA (MultiVariate Analysis)
- Applied at each particle reconstruction & identification
- Trained on Fast Boosted Decision Tree (FastBDT)



Schematic view of B tagging via $B \to D\pi$ channel

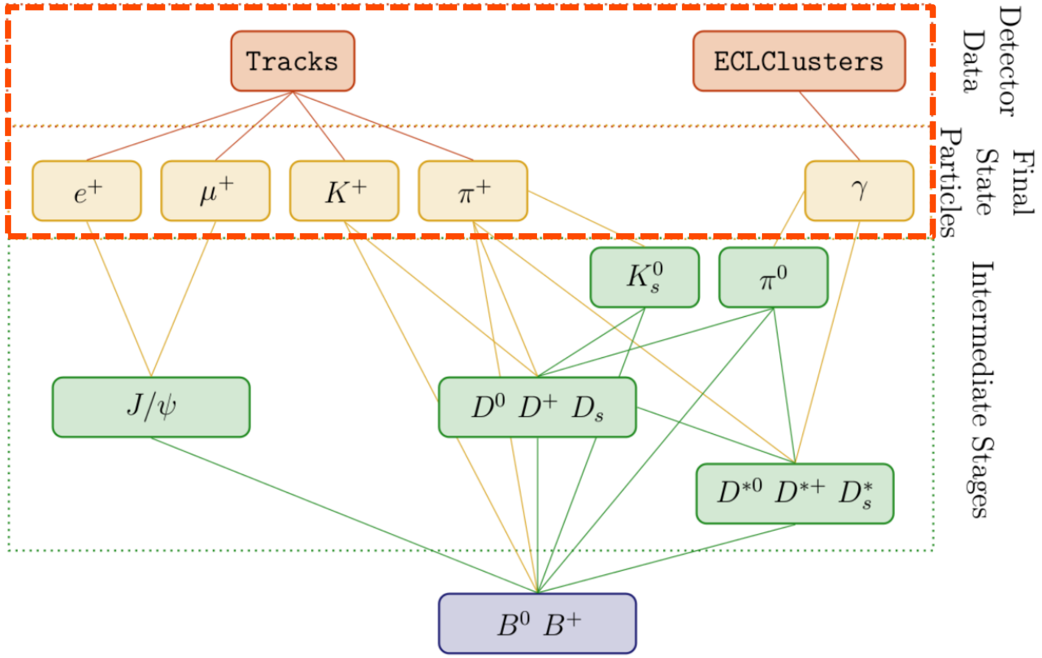


Training a Classifier for Final State Particle

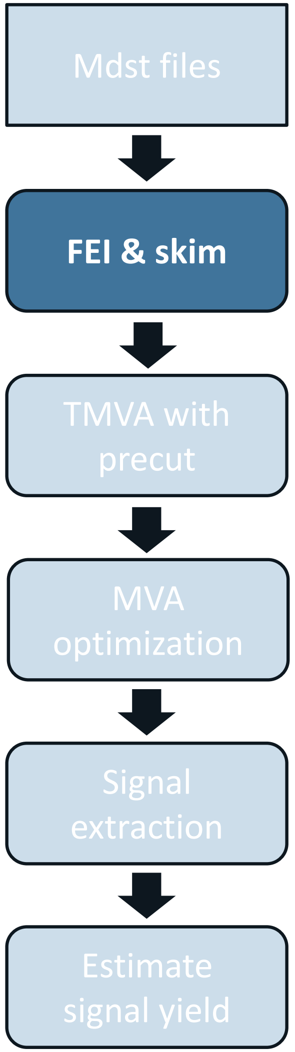


Final state particles : 4 charged particles (e, μ, π, K) and gamma

- Reconstructed with subdetector information
- Training variables
  - Charged particle (e, μ, π, K): track information from detectors + kinematic variables
  - Photon : deposited energy and shape information from calorimeters + kinematic variables
- Pre- and post-cut condition
  - To optimize computing resources
  - Before
    - 10 highest PID (e, μ)
    - 20 highest PID (π, K)
    - 40 highest energy (photon)
  - After: output > 0.01

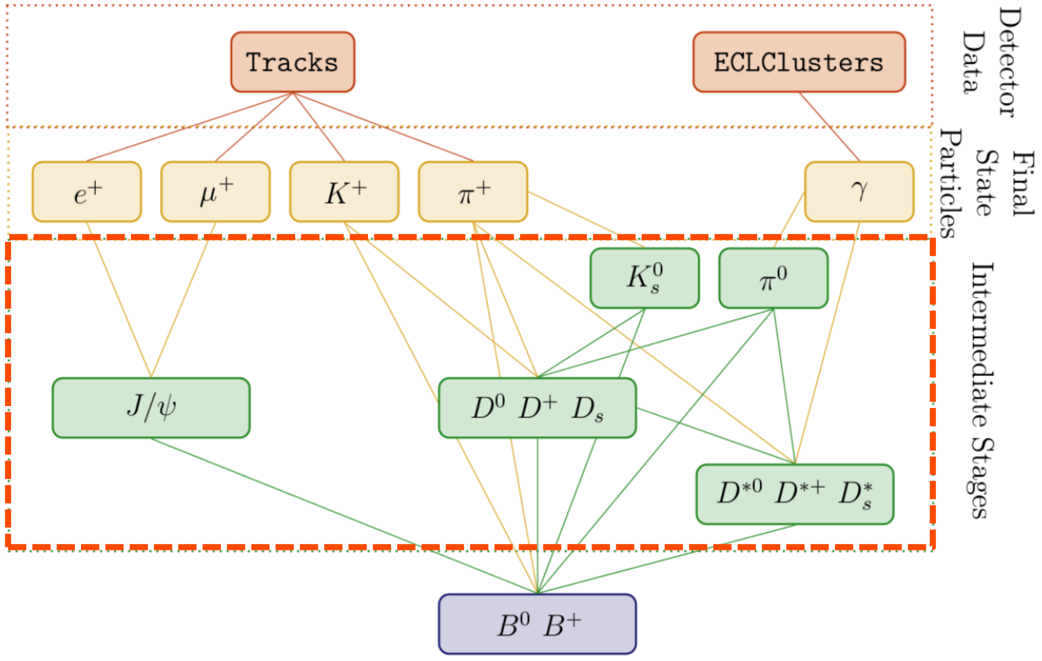


Training a Classifier for Intermediate Particles

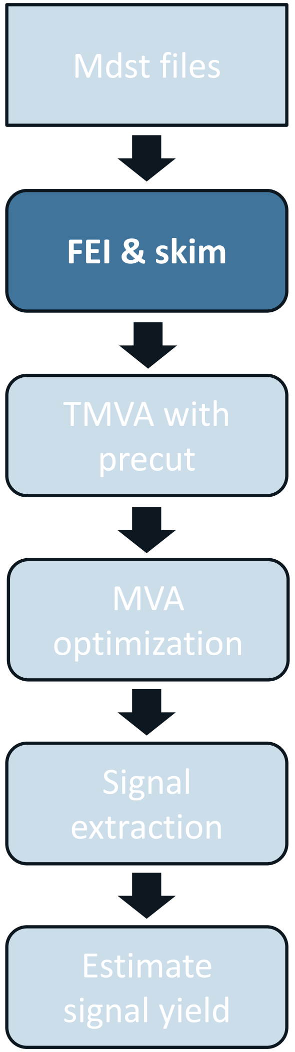


Training for particles as a part of B_{tag}

- Particle type : $\pi^0, K_S^0, J/\psi, D^0, D^+, D_S^+, D^{0*}, D^{+*}, D_S^{+*}$
- Training variable : kinematic variables + vertex variables + MVA probabilities of the daughters
- Pre- and post-cut condition
  - Pre-cut (20 per particles)
    - $P_{daughter}$ product (semileptonic or K-long included D)
    - Released energy (hadronic D*)
    - Mass differences (other particles)
  - Post-cut : probability of particle classifier
    - output > 0.01 (π^0, K_S^0) or > 0.001 (others)

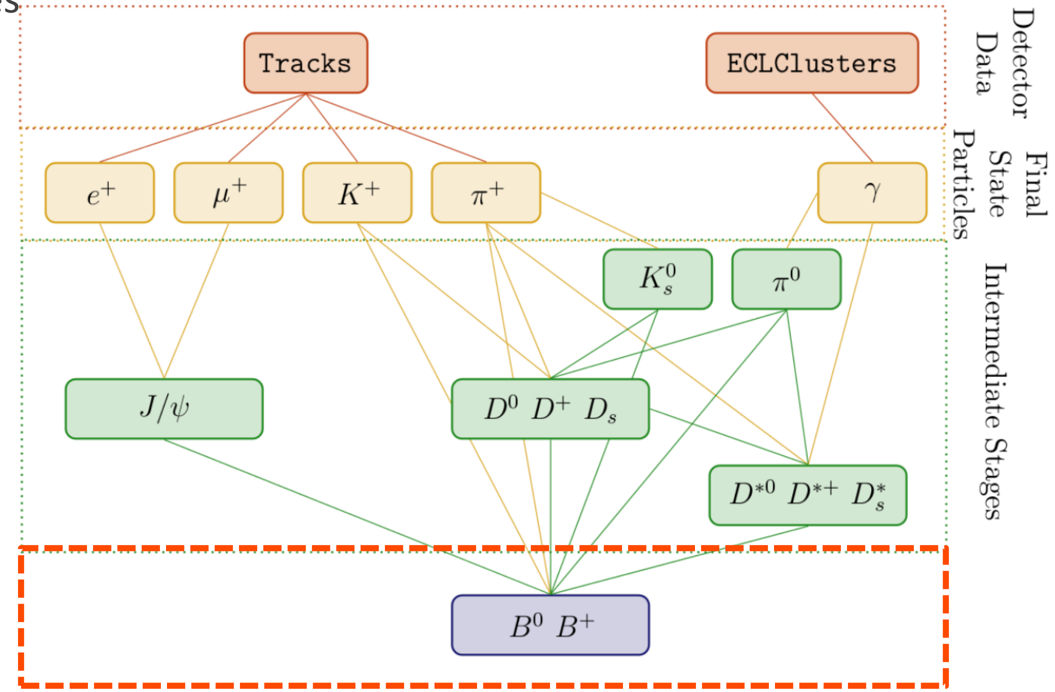


Training a Classifier for B Mesons



Final training for B mesons

- Training variables : kinematic variables + vertex variables + MVA probabilities of the daughters
- Not using any correlated variables with M_{bc} (hadronic) or or $\cos\theta_{B,D^{(*)}\ell}$ (semileptonic)
- Pre- and post-cut condition
  - Pre-cut : $P_{daughter}$ product (20 highest)
  - Post-cut : B candidates with 20 highest probabilities

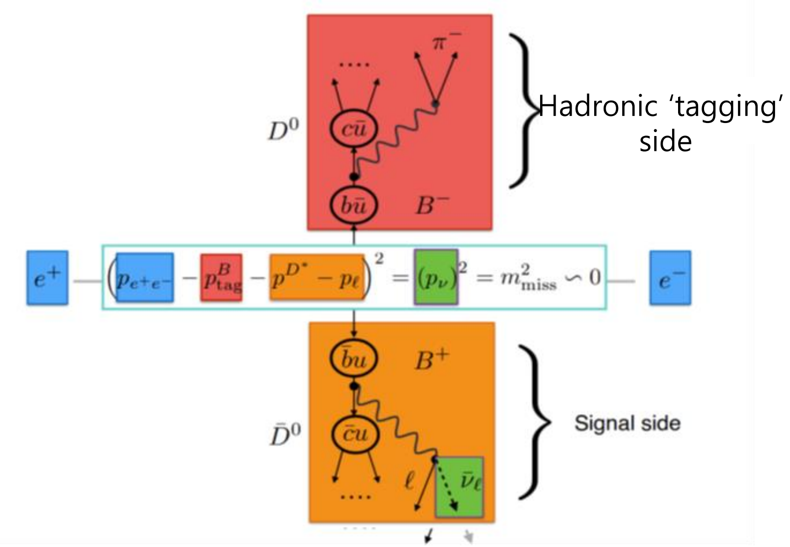


Advantages with FEI

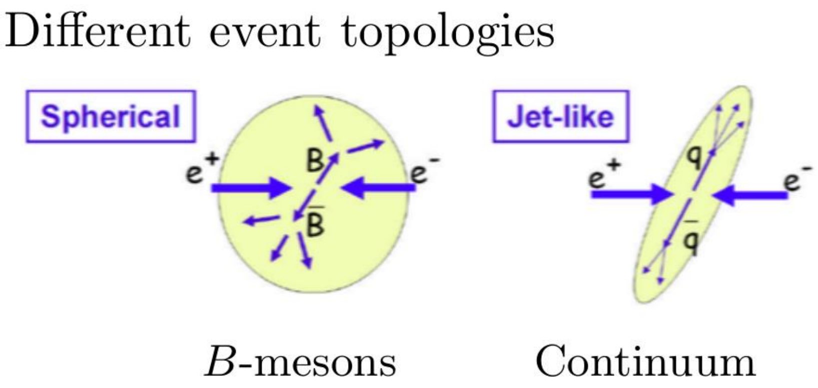
Mdst files → **FEI & skim** → TMVA with precut → MVA optimization → Signal extraction → Estimate signal yield

Advantages of this methods with machine learning

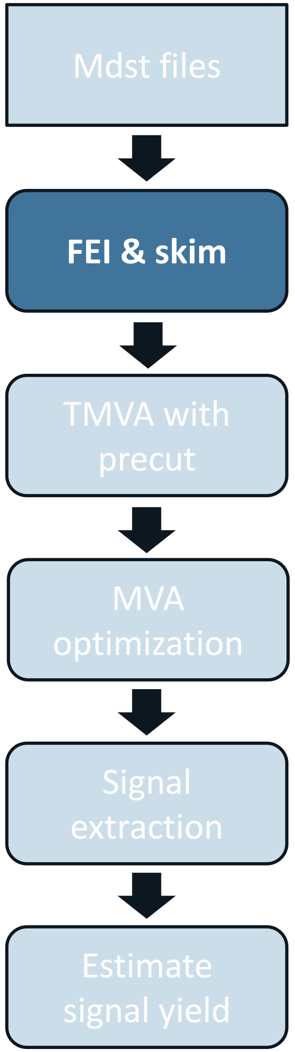
- Access additional information (signal kinematics, charge)
  - Optimized to study decay modes including invisible particles
- Requiring no additional particles in event (Completeness)
- Good continuum background suppression due to different event topologies



Calcultaion of m^2_{miss} using information of p^B_{tag}

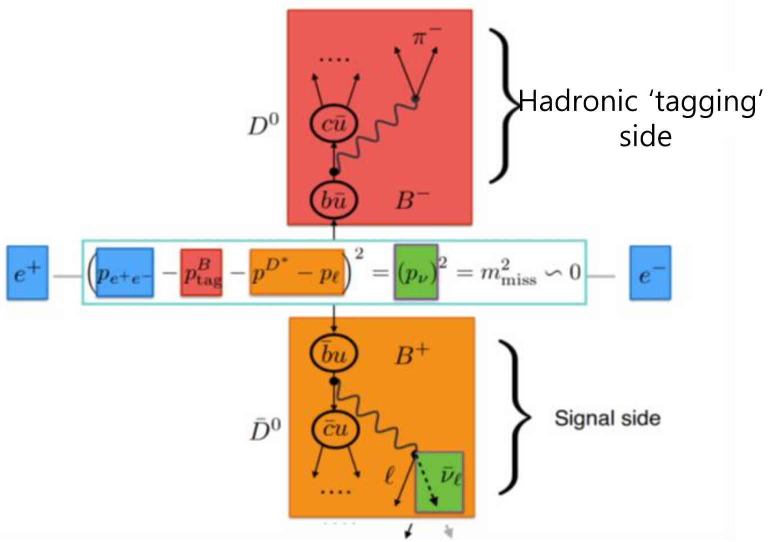


Advantages with FEI

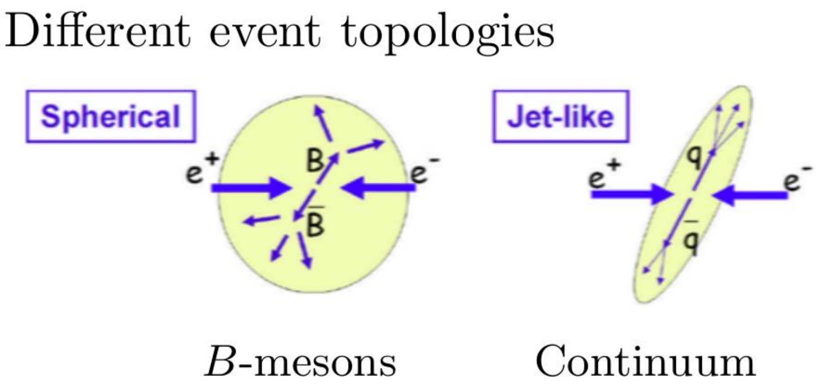


Advantages of this methods with machine learning

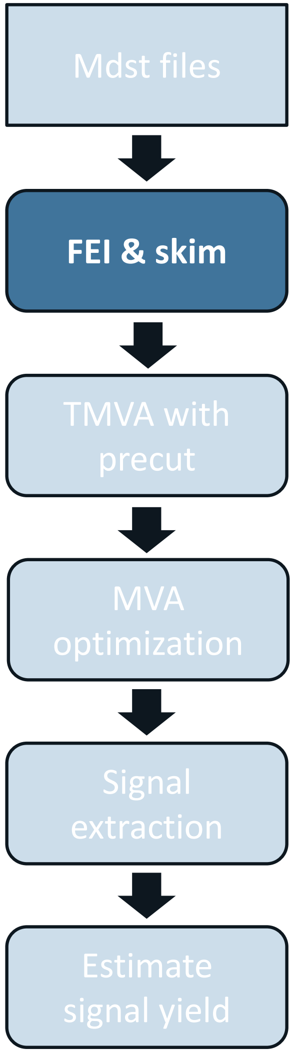
- Access additional information (signal kinematics, charge)
  - Optimized to study decay modes including invisible particles
- Requiring no additional particles in event (Completeness)
- Good continuum background suppression due to different event topologies



Calcultaion of m^2_{miss} using information of p^B_{tag}



FEI and Skim



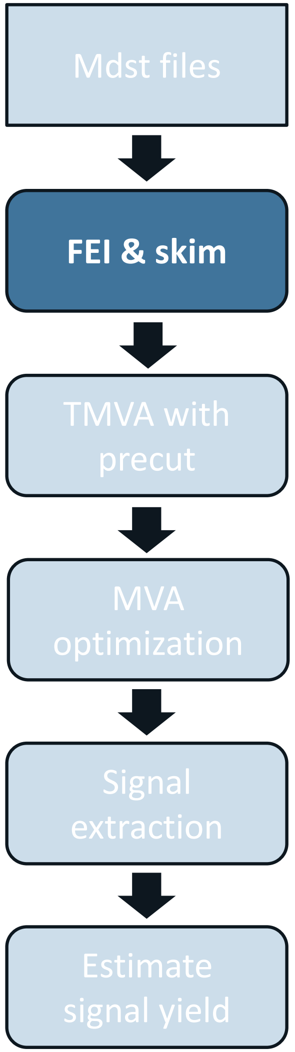
Skim conditions after FEI

- FEI probability: loose cut to check the distribution of other variables
- dr, |dz|
  - Distance on (xy-plane or z-axis) between particle creation point and interaction point(IP)
  - B^0 and τ will be decayed near IP
- PID cut: taking electron and muon with high probability
- Momentum cut
  - p_{PL} cut
  - p_{SP} cut: threshold cut for skim

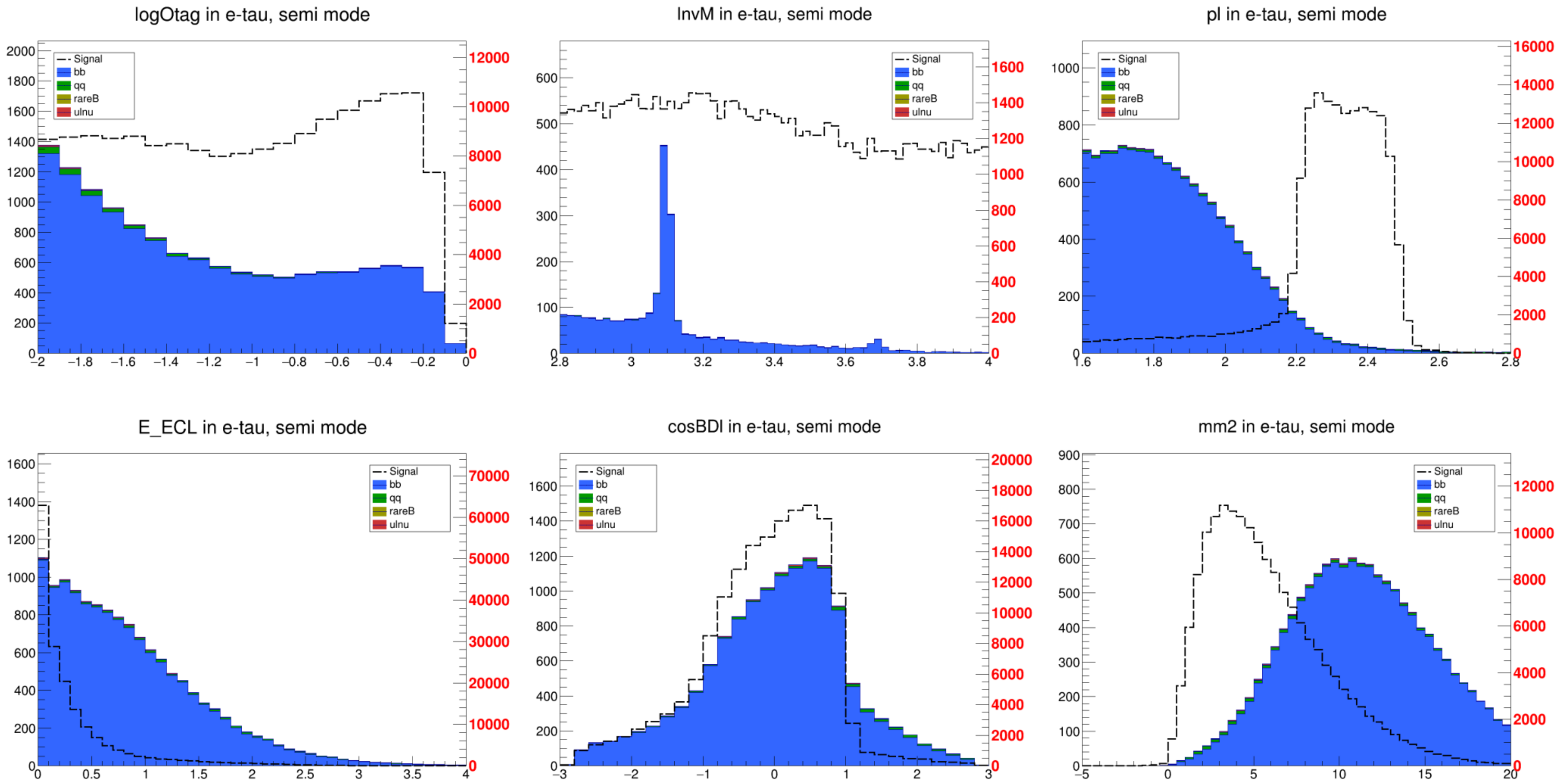
$\log_{10}(O_{tag})$ (O_{tag}: prob. of tagging)	dr (cm)	|dz| (cm)
> -2	< 2	< 4

Particle	PID for e and μ	p_{PL} (GeV/c)	p_{SP} (GeV/c)
e	eID > 0.9, eID > μID	> 1.45	> 0.2
μ	μID > 0.9, μID > eID		> 0.6

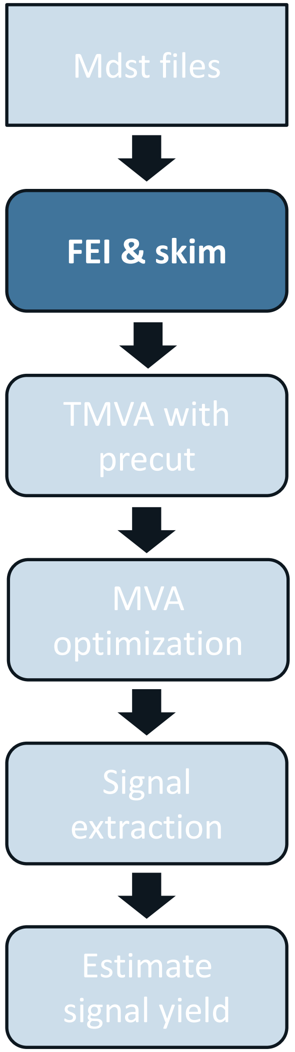
Distributions of Variables



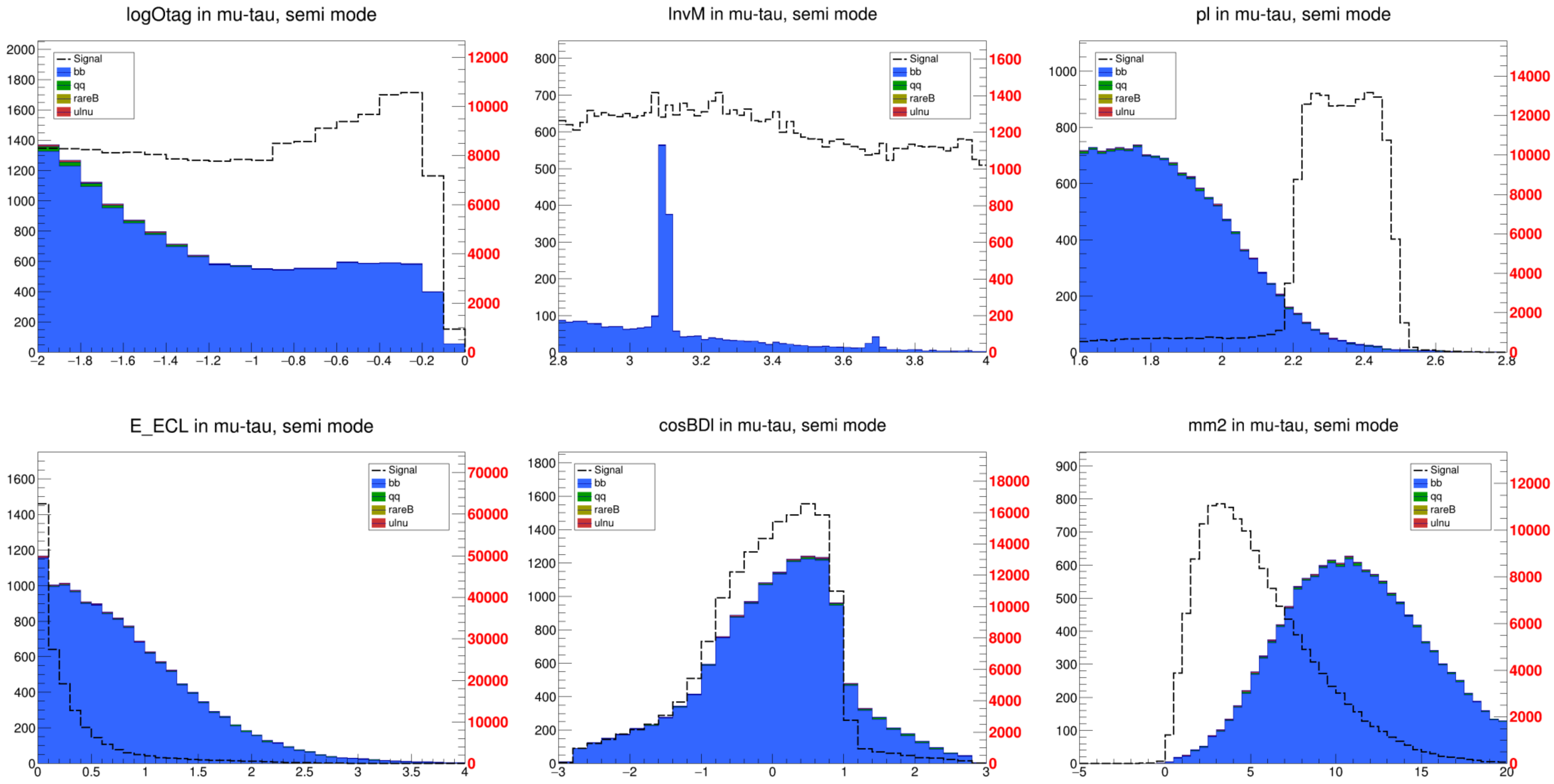
Distribution with e-τ mode



Distributions of Variables

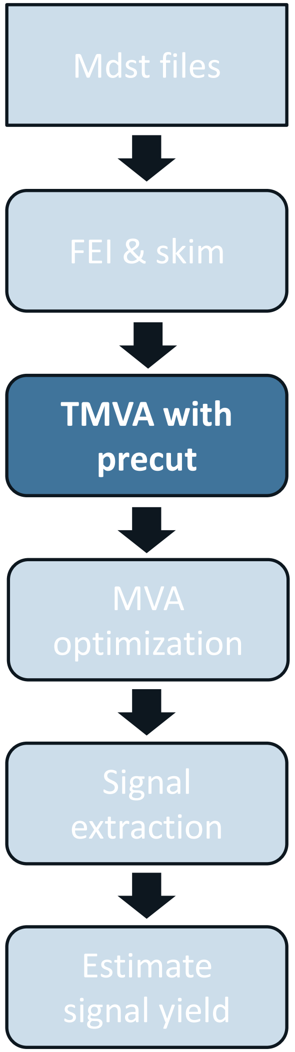


Distribution with μ-τ mode



TMVA & Optimization

TMVA and Conditions



TMVA [5]: Toolkit for MultiVariate data Analysis with ROOT

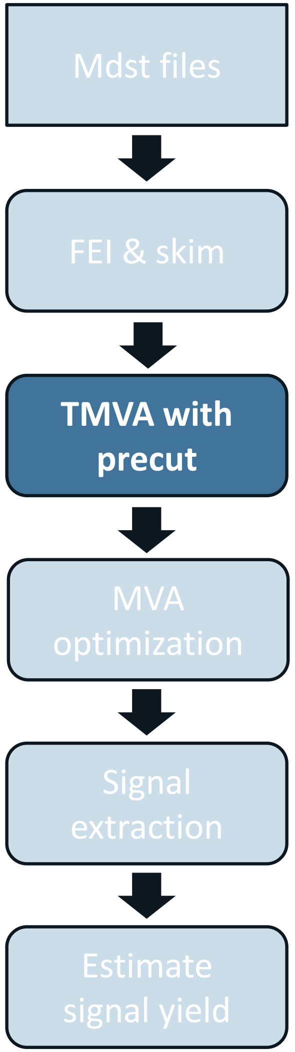
- ROOT-integrated project that provides a machine learning environment

Training condition

- Input variable : $\cos\theta_{B,D^{(*)}\ell}, E_{ECL}, m_{miss}^2$
- Precut : in table
- Training sample
  - 2 streams of background samples
  - 1 sets of signal MC samples
    - All signal MCs are 10 sets.
  - Why 2 streams & 2M?
    - Sufficient amounts of samples to well-train methods

Variables	Precut	Etc.
$M_{\ell+\tau D}$	Not in $3.0<M_{\ell+\tau D}<3.2$, $3.6<M_{\ell+\tau D}<3.8$	
$\log O_{tag}$	$\log O_{tag} \geq -2$	
p_ℓ^*	$1.6 \leq p_\ell^* \leq 2.8$	Signal extraction variable
E_{ECL}	$E_{ECL} \leq 3$	TMVA Input variables
$\cos\theta_{B,D^{(*)}\ell}$	$-1 \leq \cos\theta_{B,D^{(*)}\ell} \leq 1$	
m_{miss}^2	$-5 \leq m_{miss}^2 \leq 20$	

TMVA Methods

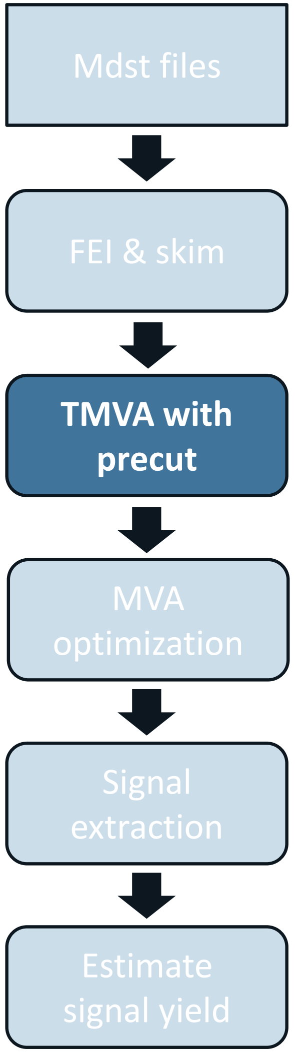


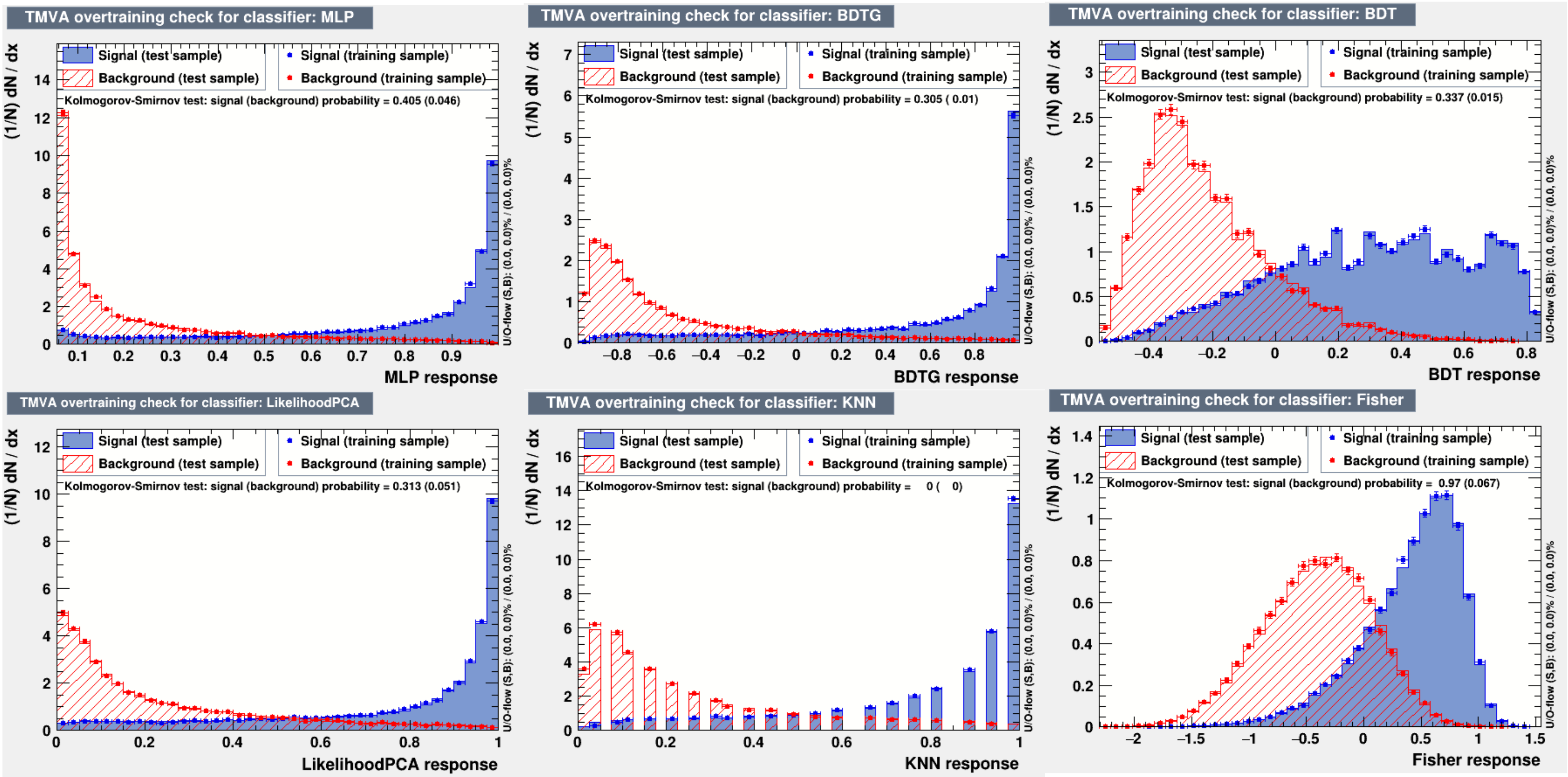
Training condition

- Method
  - BDT: Boosted Decision Tree (AdaBoost)
  - BDTG: Boosted Decision Tree (Gradient boost)
    - Decision tree: a set of sequential decision to classification
    - AdaBoost(Adaptive boosting): Weight to bad-classificated groups
    - Gradient boost: Gradual training with residual
  - LikelihoodPCA: Linear decomposition with likelihood
  - KNN: classification with the status of nearest neighbor
  - MLP: Multi-Layer Perceptron (neural network)
  - Fisher discriminants: Finding 1-dim. space to classify

Variables	Precut	Etc.
$M_{\ell+\tau D}$	Not in $3.0<M_{\ell+\tau D}<3.2$, $3.6<M_{\ell+\tau D}<3.8$	
$\log O_{tag}$	$\log O_{tag} \geq -2$	
p_ℓ^*	$1.6 \leq p_\ell^* \leq 2.8$	Signal extraction variable
E_{ECL}	$E_{ECL} \leq 3$	TMVA Input variables
$\cos\theta_{B,D^{(*)}\ell}$	$-1 \leq \cos\theta_{B,D^{(*)}\ell} \leq 1$	
m_{miss}^2	$-5 \leq m_{miss}^2 \leq 20$	

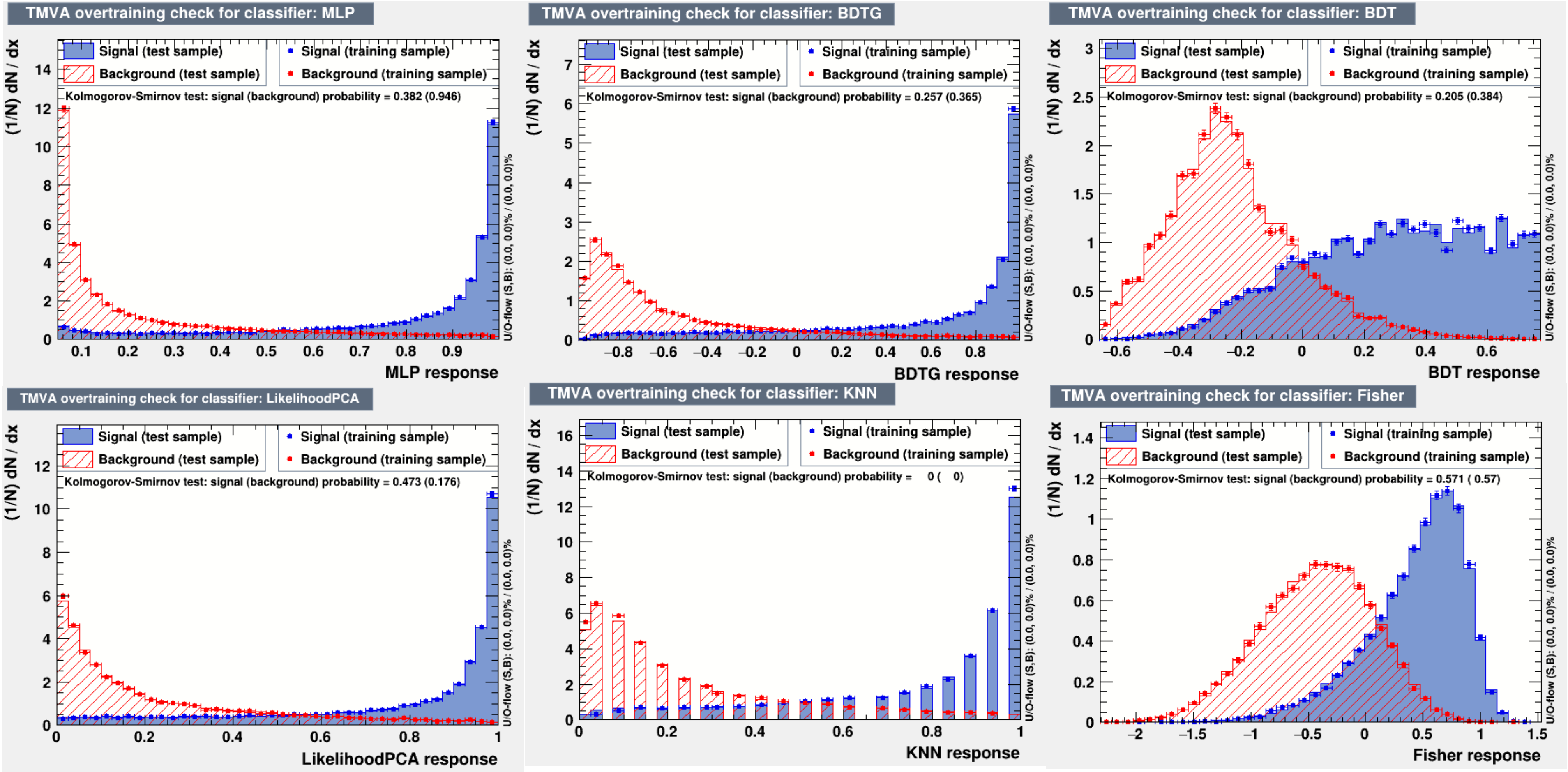
TMVA Distributions: e-τ





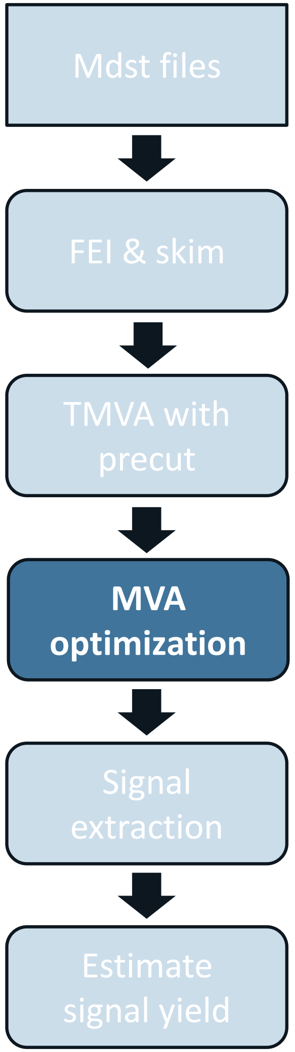
TMVA Distributions: μ-τ

- Mdst files
- FEI & skim
- **TMVA with precut**
- MVA optimization
- Signal extraction
- Estimate signal yield



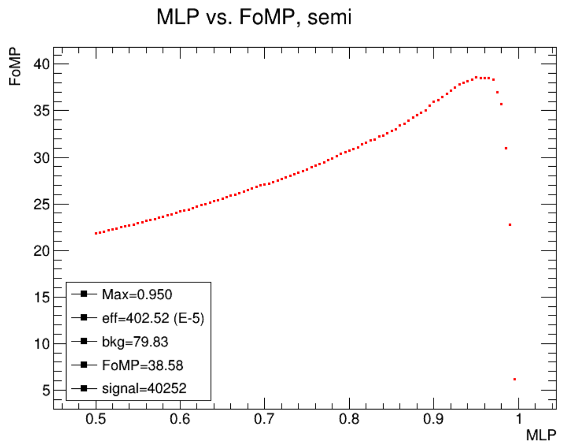
TMVA Distributions: e-τ

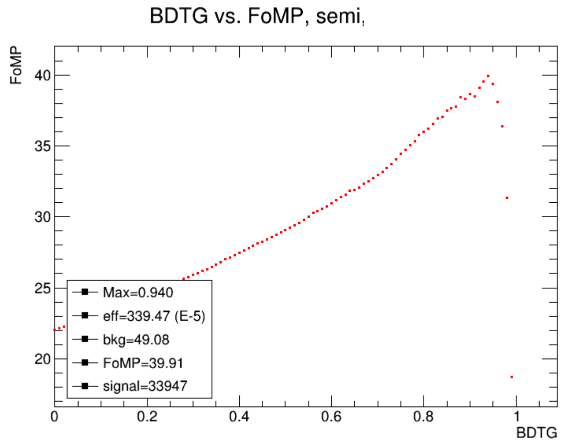
Method	MLP	BDTG	BDT	LPCA	KNN	Fisher
E_{sig} (E-5)	402.52	339.47	**335.16**	363.28	541.65	409.48
N_{bkg}	79.83	49.08	**46.17**	64.55	234.95	90.21
FoMP	38.58	39.91	**40.23**	38.1	32.19	37.23

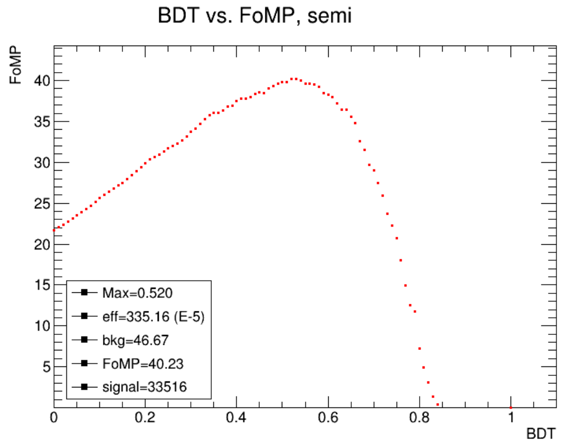


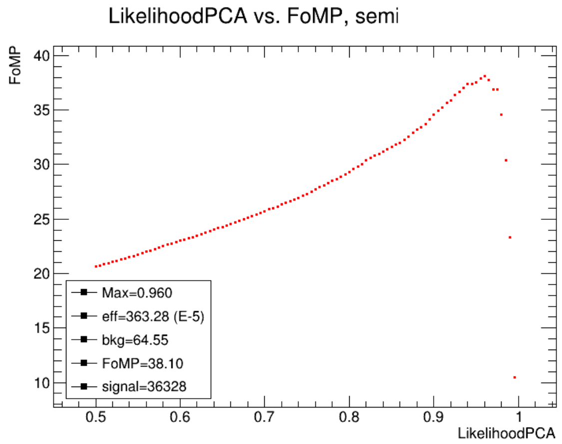
Optimization of output cut with e-τ

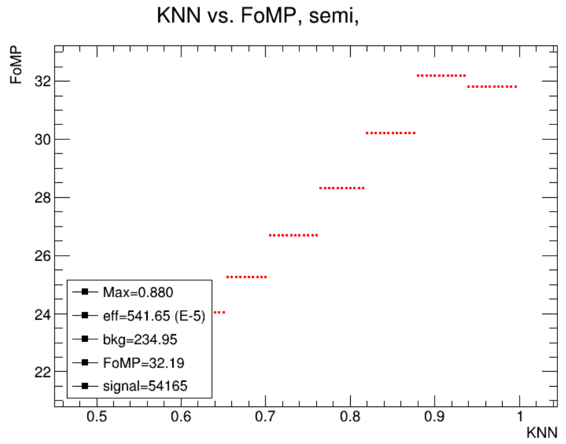
- Punzi figure of merit distribution for optimization of MVA output cut: $\frac{\epsilon_{sig}}{a/2+\sqrt{N_{bkg}}}$ $(a = 3$: the number of sigmas)

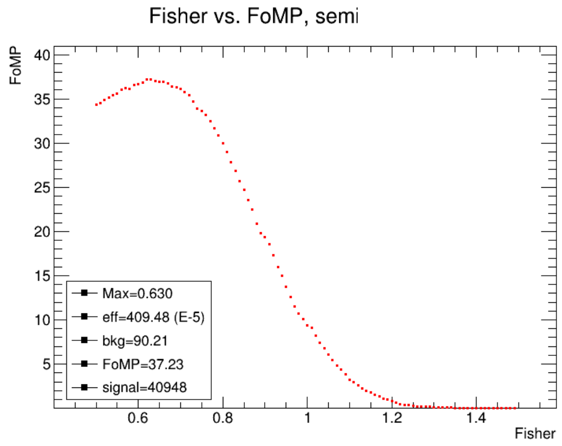






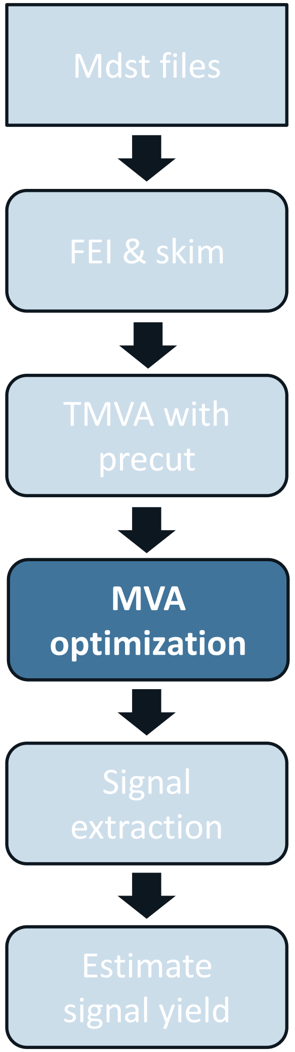






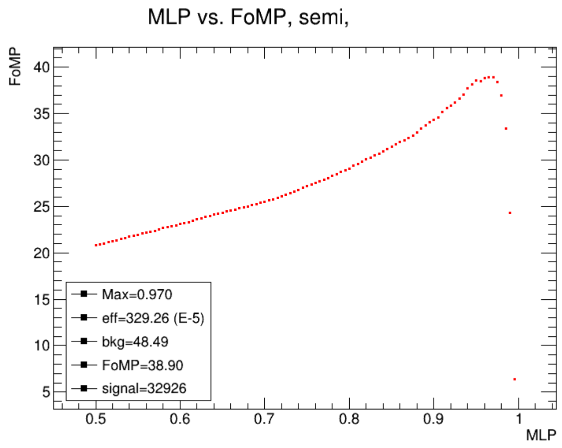
TMVA Distributions: μ-τ

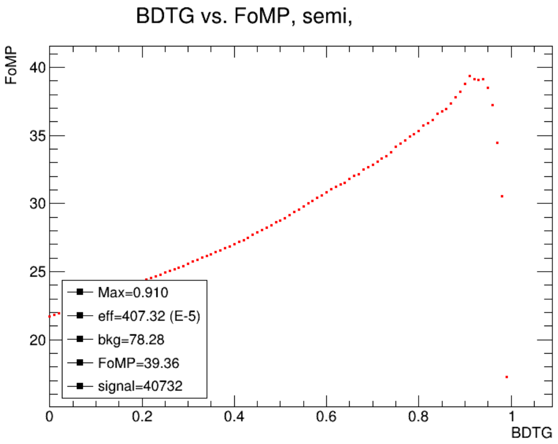
Method	MLP	BDTG	**BDT**	LPCA	KNN	Fisher
E_{sig} (E-5)	329.26	407.32	**333.96**	322.23	515.08	445.54
N_{bkg}	48.49	78.28	**48.75**	48	214.1	111.35
FoMP	38.9	39.36	**39.37**	38.23	31.93	36.97

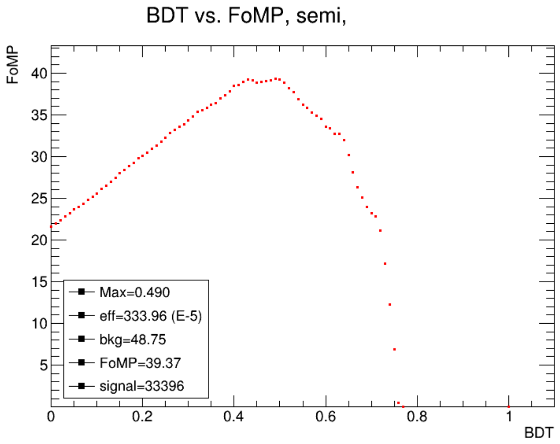


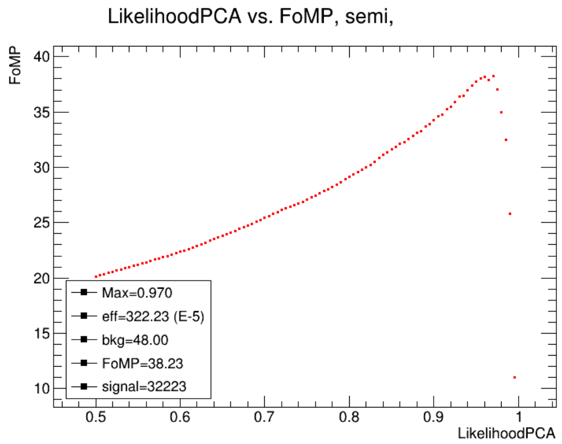
Optimization of output cut with μ-τ

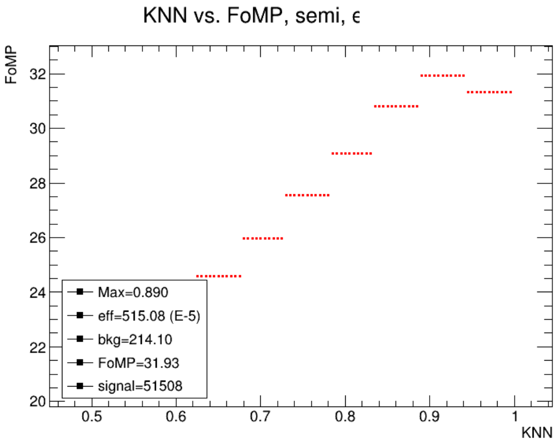
- Punzi figure of merit distribution for optimization of MVA output cut: $\frac{\epsilon_{sig}}{a/2+\sqrt{N_{bkg}}}$ $(a = 3$: the number of sigmas)

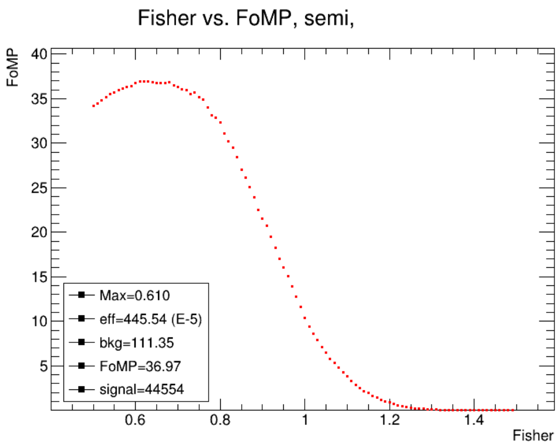








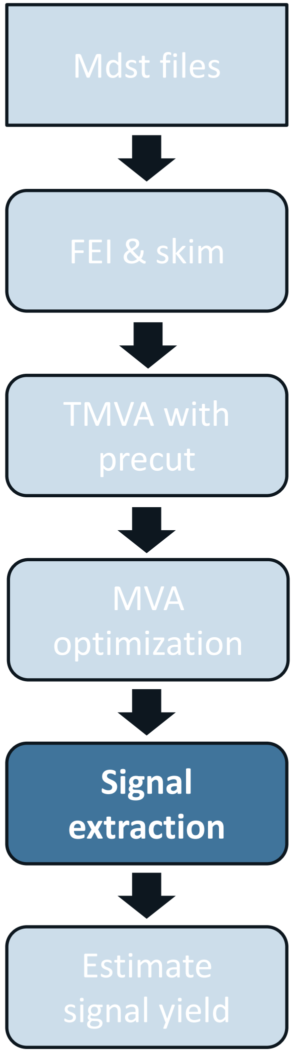




Signal Extraction

& Validation with ToyMC

p_ℓ^* Distribution



p_ℓ^* distributions after TMVA output cut

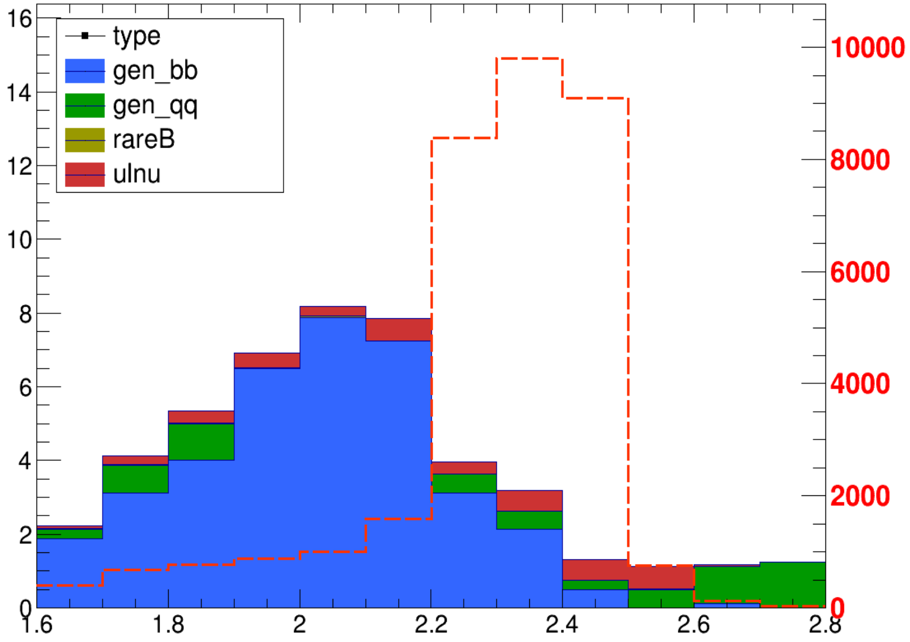
- Signal region : 2.2 < p_ℓ^* < 2.5

Signal distribution
: Gaussian-like shape with left-side tail

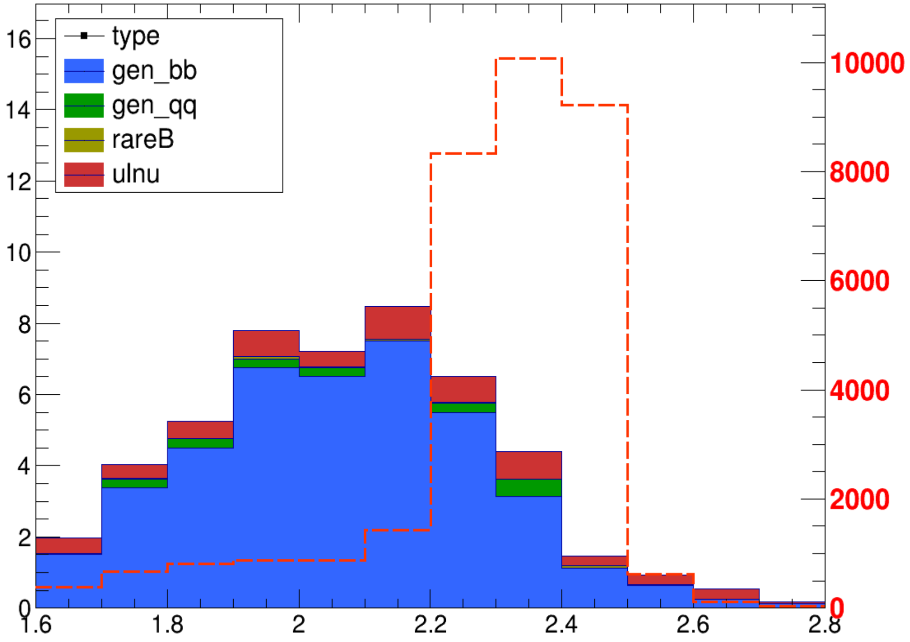
Background distribution

- e-τ mode: Gaussian with tail at right side
- μ-τ mode: Gaussian

p_ℓ^* distribution of e-τ mode

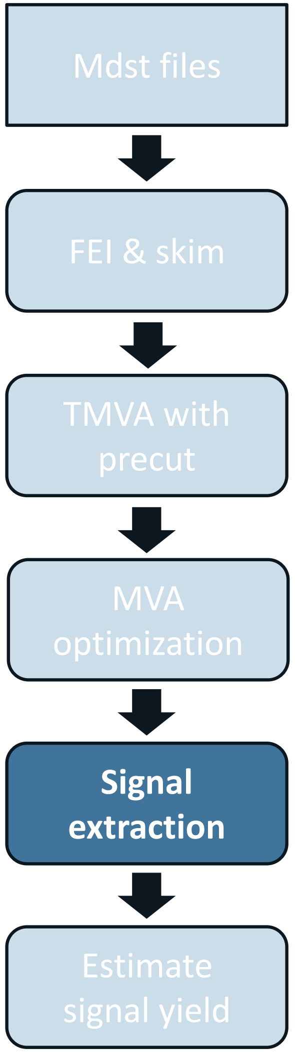


p_ℓ^* distribution of μ-τ mode

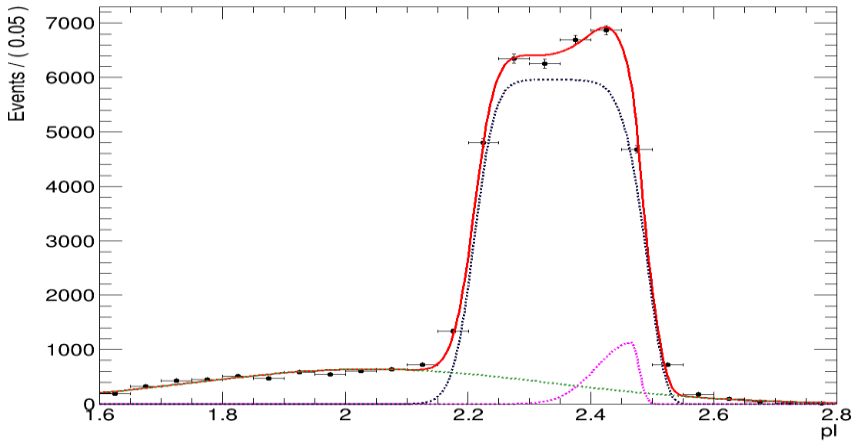


Fit of p_ℓ^* Distribution

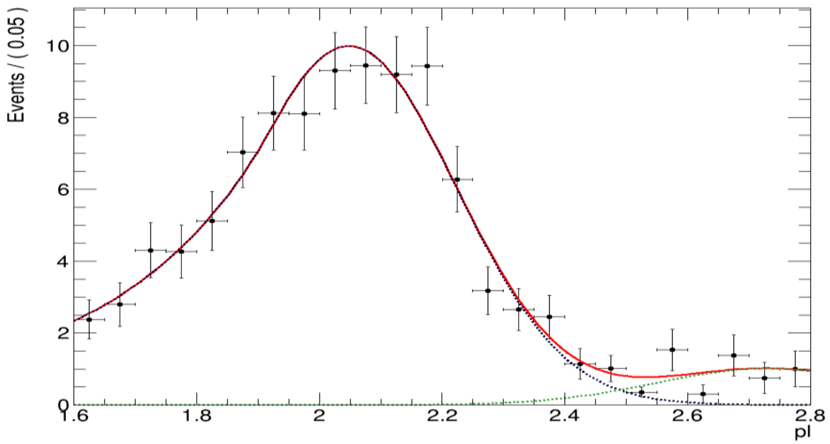
Types	e-τ	μ-τ
Signal	Convoluted function + Asymmetric gaussian + gaussian	
Background	Crystal Ball function + gaussian	Crystal Ball function



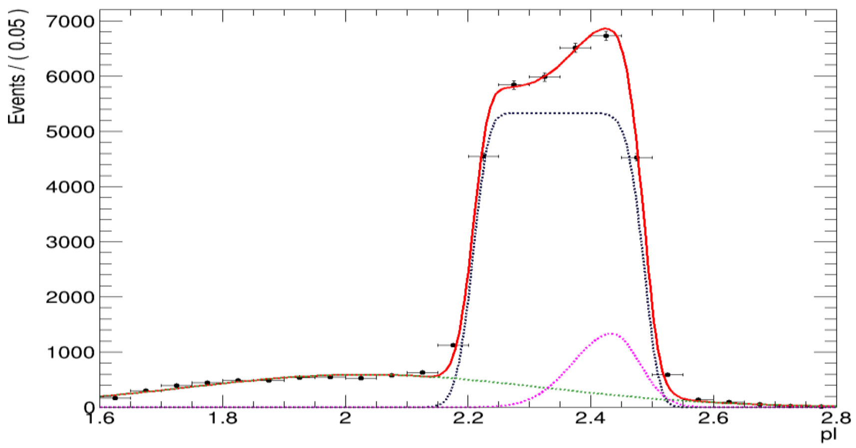
Signal fit of p_ℓ^* distribution (e-τ mode)



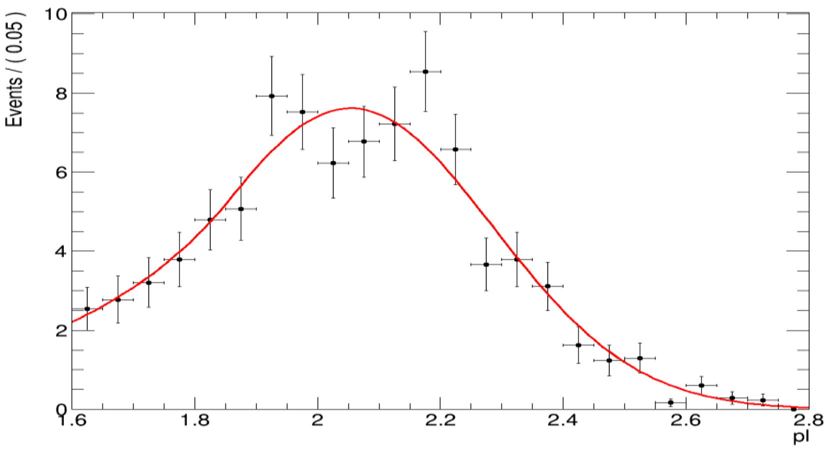
Background fit of p_ℓ^* distribution (e-τ mode)



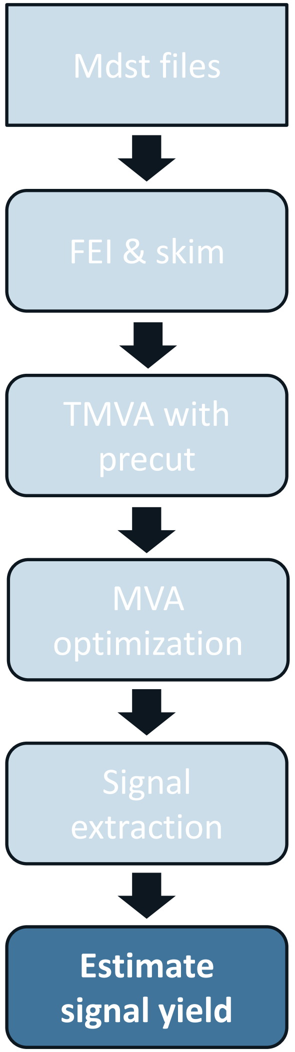
Signal fit of p_ℓ^* distribution (μ-τ mode)



Background fit of p_ℓ^* distribution (μ-τ mode)



ToyMC Study with Nsig



Fit validation with toyMC

- Fit 10,000 ensembles with 10,000 events from background PDF.
- N_{sig} is generated from signal PDF.

Problems with toyMC study

- Large variation → Requiring more research
- No results with CL 90% at N < 20…

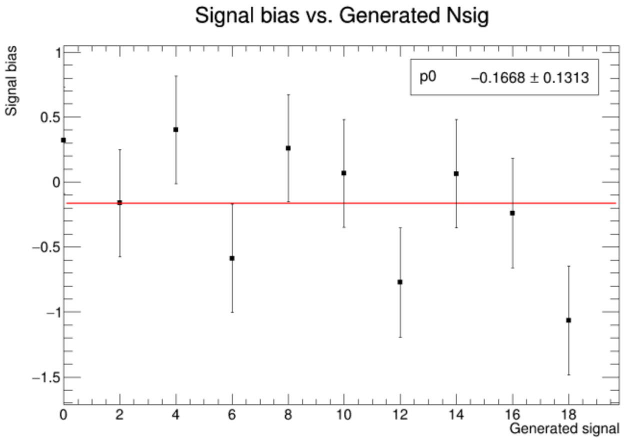
Plan for problems

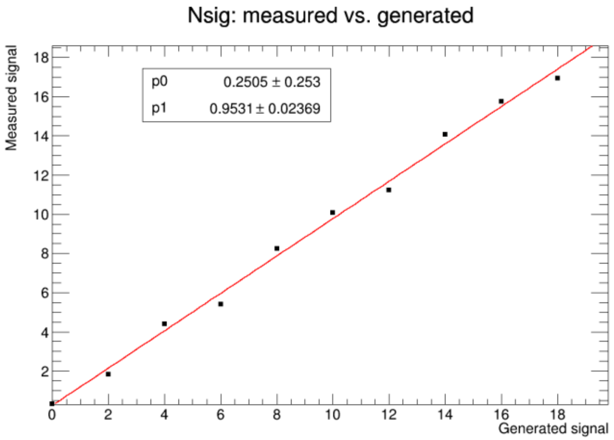
- Histogram PDF fit of signal MC
- Using fit validation - not toyMC but 1-stream set of backgrounds
  - 4 background sets are available.

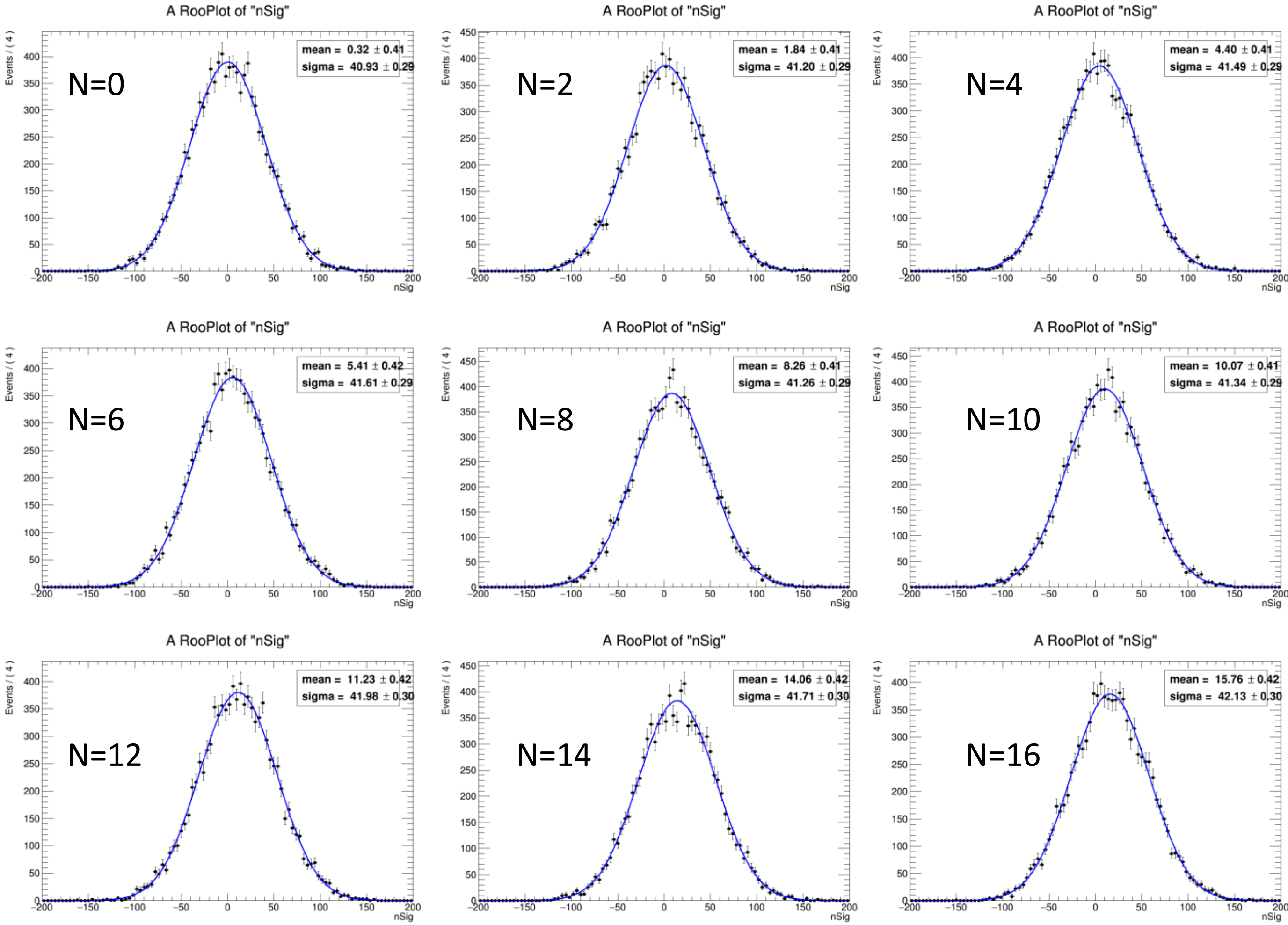
ToyMC Study with Nsig

- Mdst files
- FEI & skim
- TMVA with precut
- MVA optimization
- Signal extraction
- **Estimate signal yield**

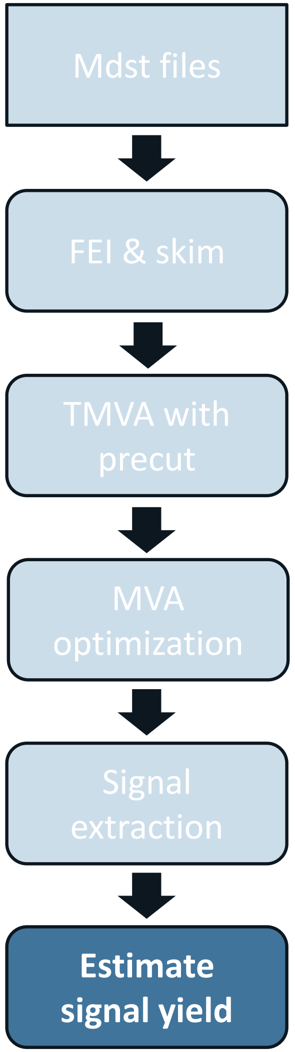
e-τ mode



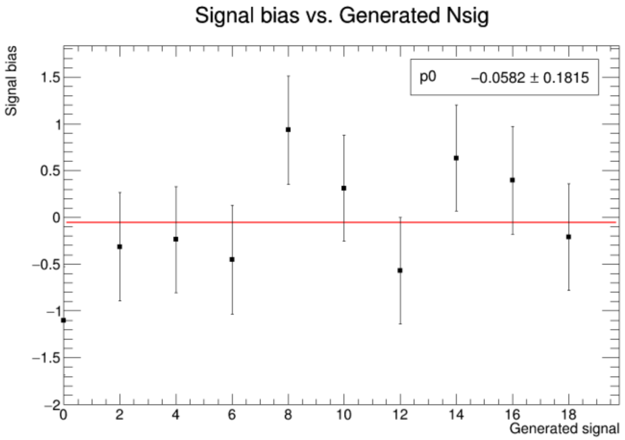


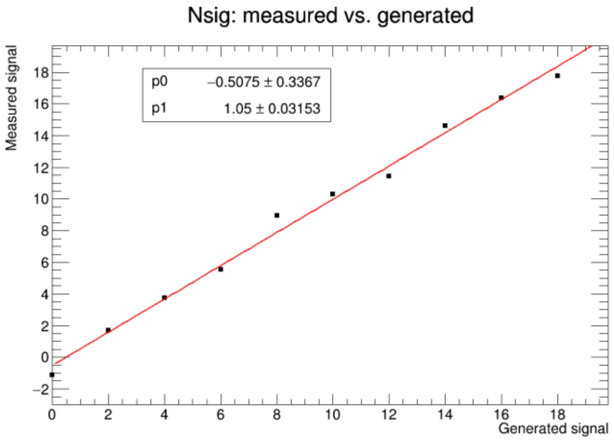


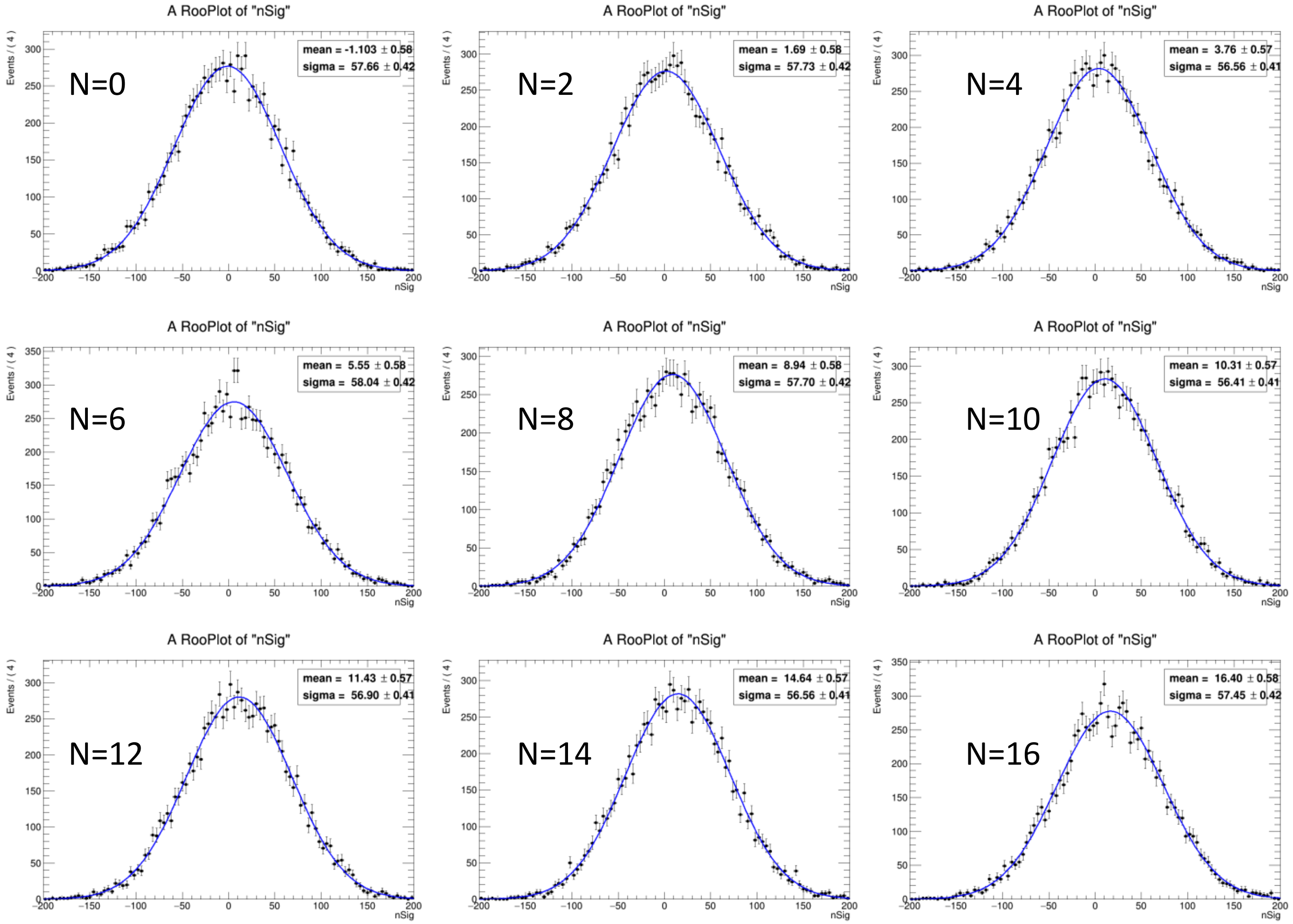
ToyMC Study with Nsig



μ-τ mode







Summary & Plan

Summary

- The study of forbidden decay mode $B^0 \to \ell^{\pm}\tau^{\mp}$ would be an answer of searching 'New Physics'.
- FEI and TMVA are used to optimize events and suppress backgrounds of Belle generic and special decays.
- The PDFs of signal and background are constructed by fit of distributions of p_ℓ^* from each l-tau decay modes
- ToyMC study is done with ensembles, and it remains additional study for variation of distributions.

Plan

- Histogram PDF fit and ToyMC analysis for validation of fit
  - Or, validation with 1-stream set of backgrounds
- Estimation of MC upper limit

References

[1] Xiao-Gang He, G. Valencia, Yili Wang, Lepton flavor violating τ and B decays and heavy neutrinos, *Phys. Rev. D* 70 (2004), 113011

- The benchmark for significant constraints from B decay is a sensitivity of 10^{-9} for $B(B_s \to \tau^{\pm}\ell^{\mp})$ and of 10^{-8} for $B(b \to s\tau^{\pm}\ell^{\mp})$.

[2] B. Aubert et al. (BaBar Collaboration), Searches for the decays $B^0 \to \ell^+\tau^-$ and $B^+ \to \ell^+\nu$ $(\ell = e, \mu)$ using hadronic tag reconstruction, *Phys.Rev.D* 77 (2008), 091104

[3] H. Atmacan et al. (Belle Collaboration), Search for $B^0 \to \tau^{\pm}\ell^{\mp}(\ell = e, \mu)$ with a hadronic tagging method at Belle, *Phys.Rev.D* 104 (2021) 9, L091105

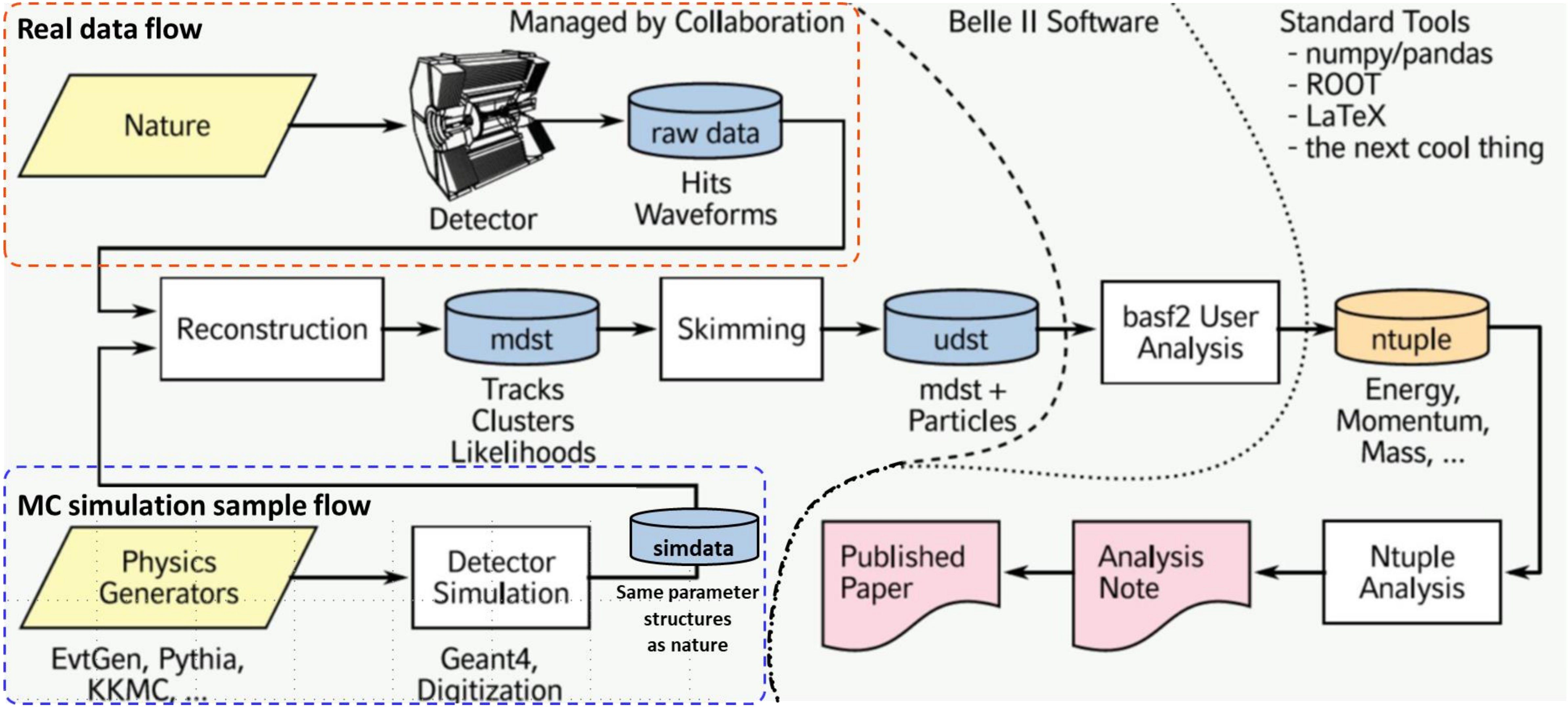
[4] T. Keck et al., The Full Event Interpretation - An Exclusive Tagging Algorithm for the Belle II Experiment, *Computing and Software for Big Science volume 3* (2019), 6

[5] A. Hoecker et al., TMVA - Toolkit for Multivariate Data Analysis, *arXiv:physics/0703039* (2007)

Thank you!

Backup

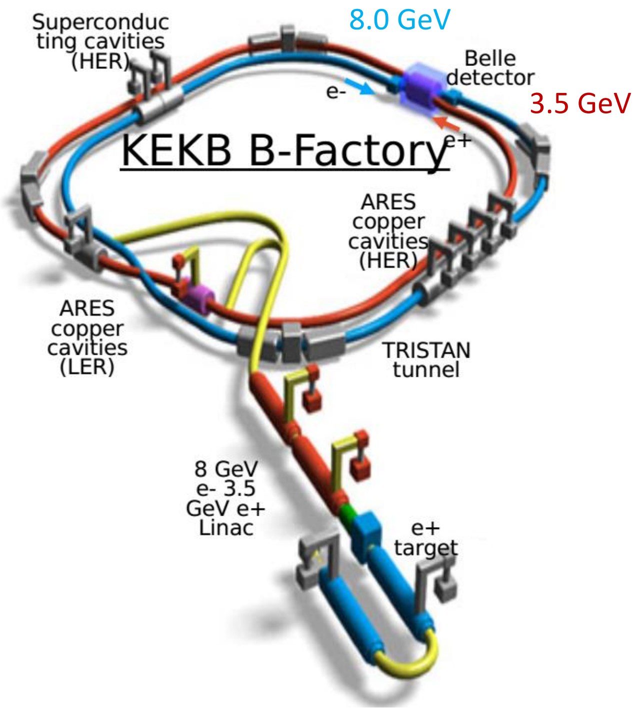
Belle (II) Data Flow Overview

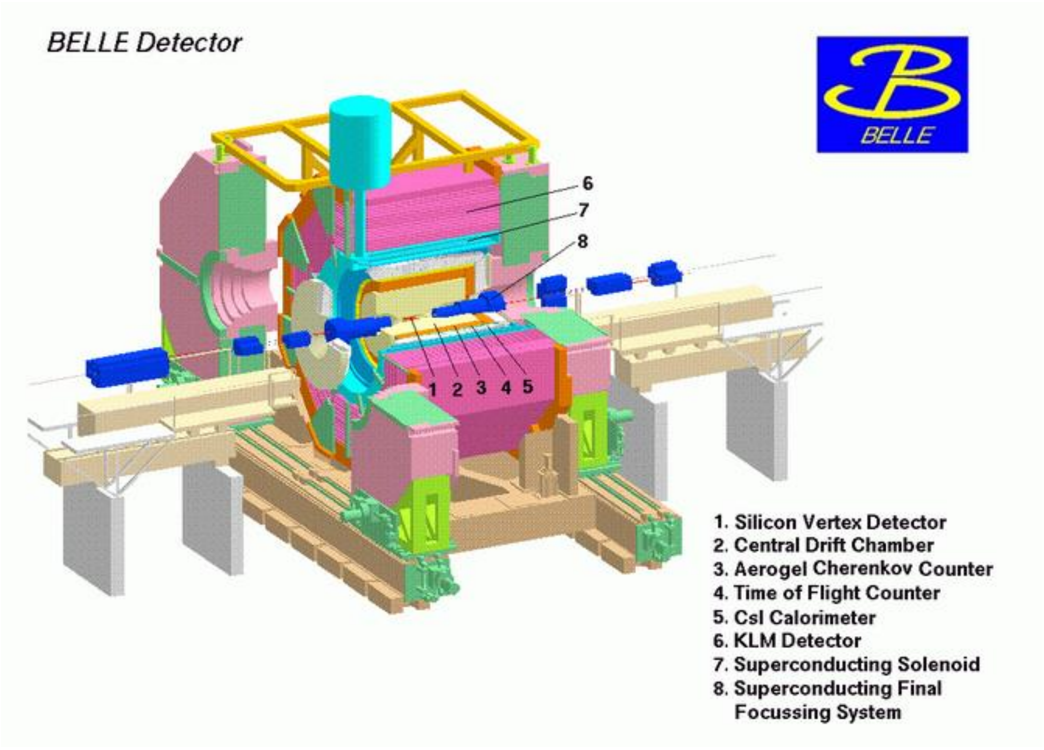


Belle Detector

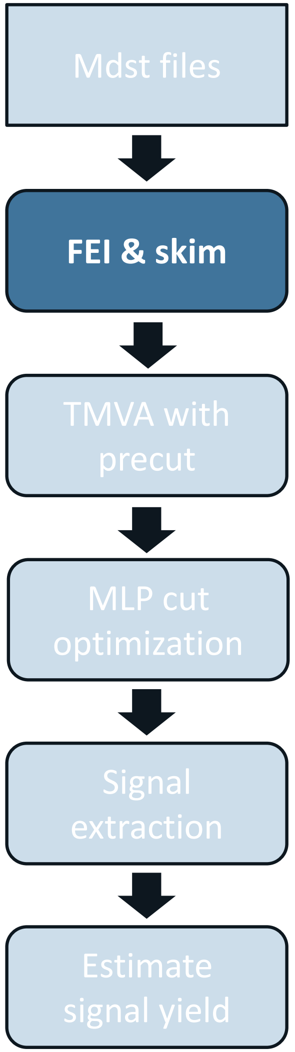
Data collected with Belle detector

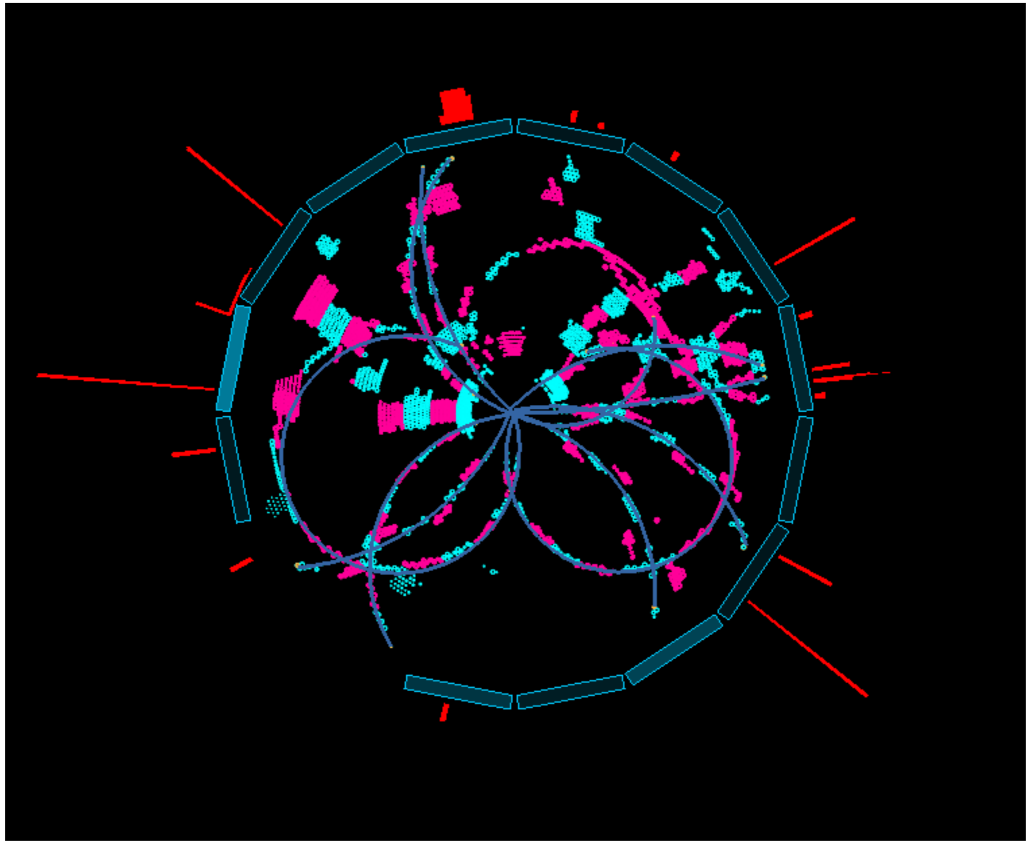
- At KEKB asymmetric e^+e^- collider : 3.5 GeV e^+ & 8 GeV e^-
  - CM energy = 10.58 GeV to make Υ(4S), that decays to BB pair with 96% rate
- Total 711 fb^{-1} of data (772M BB pairs) collected at Υ(4S)
- Covering 17~150 degree of beam-parallel and 360 degree of beam-proportional direction





Combinatorics





Example with one event with 10 tracks

- Assuming 5 positively and 5 negatively charged
- Reconstruction of $D^+ \to K^-\pi^+\pi^+$
  - $\binom{5}{2}\binom{5}{1} = 50$ possible combinations
- Reconstructing $B^0 \to D^+(\to K^-\pi^+\pi^+)\pi^-$
  - $\binom{4}{1} \times 50 = 200$ combinations
- Considered D meson decays
  - 15 for D^0, 11 for D^+

Particle	pre-cut			post-cut					
e^+	10	highest	e-ID	5	highest	σ	and	0.01	$< \sigma$
μ^-	10	highest	μ-ID	5	highest	σ	and	0.01	$< \sigma$
π^+	20	highest	π-ID	10	highest	σ	and	0.01	$< \sigma$
K^+	20	highest	K-ID	10	highest	σ	and	0.01	$< \sigma$
γ	40	highest	E	20	highest	σ	and	0.01	$< \sigma$
π^0	20	lowest	$\lvert M - M_{\pi^0}\rvert$	10	highest	σ	and	0.01	$< \sigma$
K_S^0	20	lowest	$\lvert M - M_{K_S^0}\rvert$	10	highest	σ	and	0.01	$< \sigma$
K_L^0	20	lowest	$\lvert M - M_{K_L^0}\rvert$	10	highest	σ	and	0.01	$< \sigma$
D^0 (had)	20	lowest	$\lvert M - M_{D^0}\rvert$	10	highest	σ	and	0.001	$< \sigma$
D^0 (sem)	20	highest	$\prod_i \sigma_i$	10	highest	σ	and	0.001	$< \sigma$
D^0 (klong)	20	highest	$\prod_i \sigma_i$	10	highest	σ	and	0.001	$< \sigma$
D^+ (had)	20	lowest	$\lvert M - M_{D^+}\rvert$	10	highest	σ	and	0.001	$< \sigma$
D^+ (sem)	20	highest	$\prod_i \sigma_i$	10	highest	σ	and	0.001	$< \sigma$
D^+ (klong)	20	highest	$\prod_i \sigma_i$	10	highest	σ	and	0.001	$< \sigma$
D^{+*} (had)	20	lowest	$\lvert Q - Q_{D^{+*}}\rvert$	10	highest	σ	and	0.001	$< \sigma$
D^{+*} (sem)	20	lowest	$\lvert Q - Q_{D^{+*}}\rvert$	10	highest	σ	and	0.001	$< \sigma$
D^{+*} (klong)	20	lowest	$\lvert Q - Q_{D^{+*}}\rvert$	10	highest	σ	and	0.001	$< \sigma$
D_s^+ (had)	20	lowest	$\lvert M - M_{D_s^+}\rvert$	10	highest	σ	and	0.001	$< \sigma$
D_s^+ (klong)	20	highest	$\prod_i \sigma_i$	10	highest	σ	and	0.001	$< \sigma$
D_s^{+*} (had)	20	lowest	$\lvert Q - Q_{D_s^{+*}}\rvert$	10	highest	σ	and	0.001	$< \sigma$
D_s^{+*} (klong)	20	lowest	$\lvert Q - Q_{D_s^{+*}}\rvert$	10	highest	σ	and	0.001	$< \sigma$
B^+ (had)	20	highest	$\prod_i \sigma_i$	20	highest	σ			
B^+ (sem)	20	highest	$\prod_i \sigma_i$	20	highest	σ			
B^+ (klong)	20	highest	$\prod_i \sigma_i$	20	highest	σ			
B^0 (had)	20	highest	$\prod_i \sigma_i$	20	highest	σ			
B^0 (sem)	20	highest	$\prod_i \sigma_i$	20	highest	σ			
B^0 (klong)	20	highest	$\prod_i \sigma_i$	20	highest	σ			

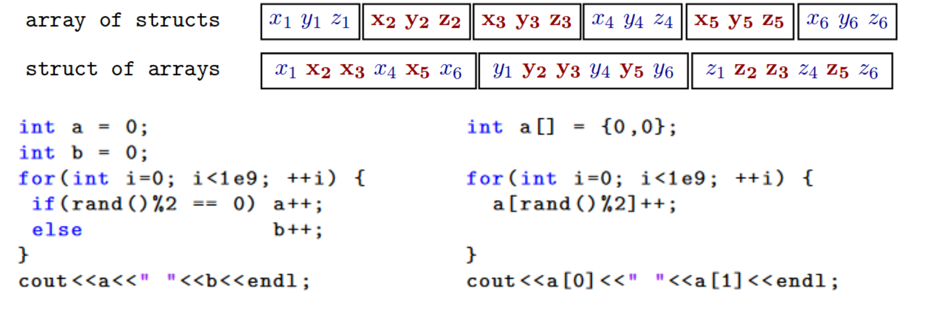
FastBDT : Classification Algorithm of FEI

Requirements for FEI classification algorithm

- Fast during fitting and application
- Robust enough to be trained in an automated environment
- Can be reliably used by non-experts

FastBDT : BDT with speed-optimized and cache-friendly implementations for multivariate classification

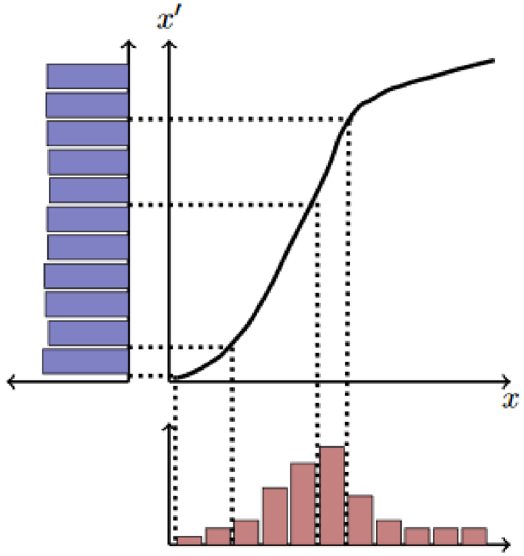
- Most time-cost part : calculation of cumulative probability histograms (CPH) to find the best-cut at each node of the tree
- Trial to reduce run time
  - Storing data as an array of structs
  - Computing cumulative probability histograms (CPH) of nodes in the same layer of the tree simultaneously
  - BDT cut decisions optimized based on equal frequency bins



(a) Straight-forward implementation – Execution time 10.1 sec

(b) If statement replaced by array lookup – Execution time 6.9 sec

Array of structs and example of optimized code modification



Concepts of equal-frequency binning

Benchmark of FastBDT and Others for FEI

Benchmarks of reconstruction using $D^0 \rightarrow K^- \pi^+ \pi^0$

- Fitting time measured about 28 features and 355,000 events
- Inference time measured about 28 features and 3,900,000 events

do nothing during the fitting phase → Trivial

Method	Fitting time in s	Inference time in s	AUC ROC	WeightFile size in KB
Trivial	0.2	4.9	0.066	2
Stochastic Gradient Boosted Decision Tree				
FastBDT	3.7	6.9	0.435	58
SKLearn-BDT	32.1	7.8	0.429	69
XGBoost	18.0	11.4	0.415	34
TMVA-BDT	19.8	16.5	0.297	101
Artificial Neural Network				
SKLearn-NN	27.6	7.2	0.401	32
Tensorflow	201.9	9.4	0.399	30
NeuroBayes	112.3	75.4	0.377	182
FANN	50.6	7.1	0.316 ± 0.061	21
TMVA-NN	510.6	16.8	0.156	53

FEI Tagging Performance with MC

Maximum tag-side efficiency of FEI and other reconstruction algorithm

- FR (Full Reconstruction) at Belle I, with NeuroBayes module
- SER (Semi-Exclusive-Reconstruction) at BaBar, with neural network

Tag	FR	SER	FEI Belle	FEI Belle II
Hadronic B^+	0.28%	0.4%	0.76%	0.66%
SL B^+	0.31%	0.3%	1.80%	1.45%
Hadronic B^0	0.18%	0.2%	0.46%	0.38%
SL B^0	0.34%	0.6%	2.04%	1.94%

FEI Performance Check

Distribution of the kinematic variable of B_{tag} at Belle II MC

- 180M BB pair signal and $1ab^{-1}$ scaled backgrounds from BB, e^+e^- to qq pair and $\tau^+\tau^-$

Distribution of

$$M_{bc} = \sqrt{E_{CM/2}^2 - p_{B_{tag}}^2}$$

(beam-strained mass)

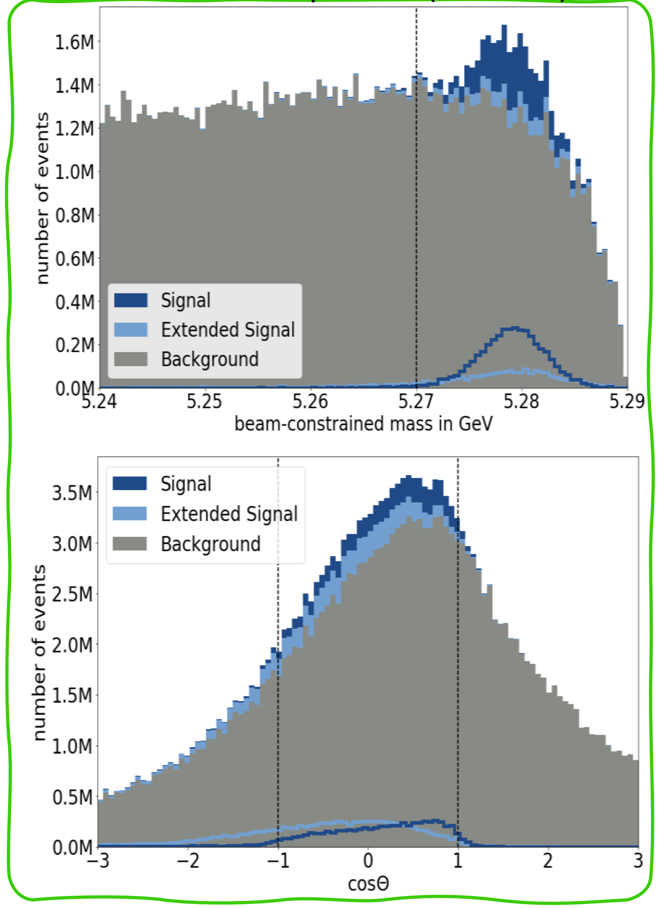
Distribution of

$$\cos\theta_{B,D^{(*)}\ell} = \frac{p'_B \cdot p_{D^{(*)}\ell}}{|p'_B||p_{D^{(*)}\ell}|},$$

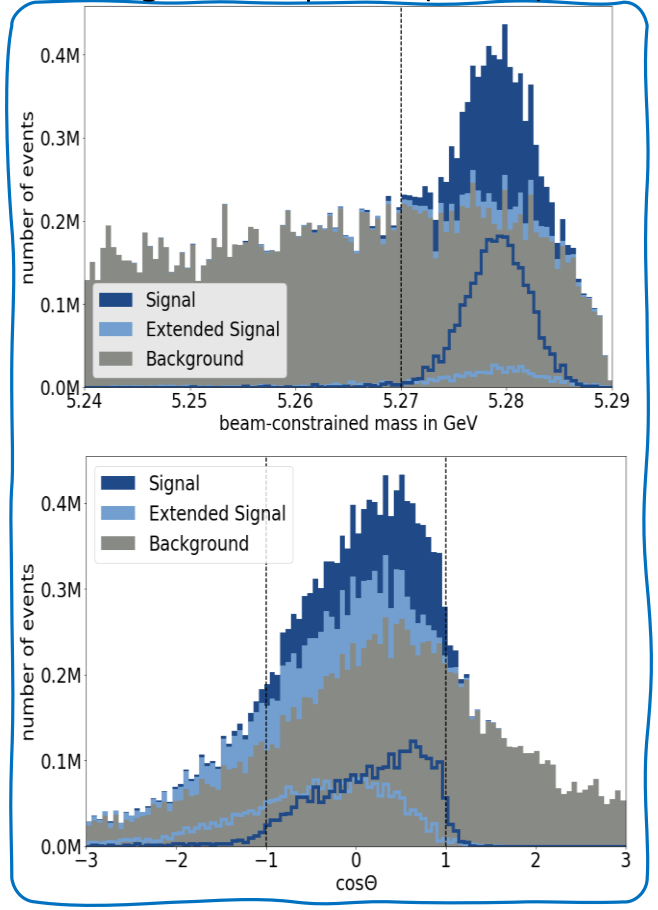
$$|p'_B| = \sqrt{E_{CM/2}^2 - m_{B_{true}}^2}$$

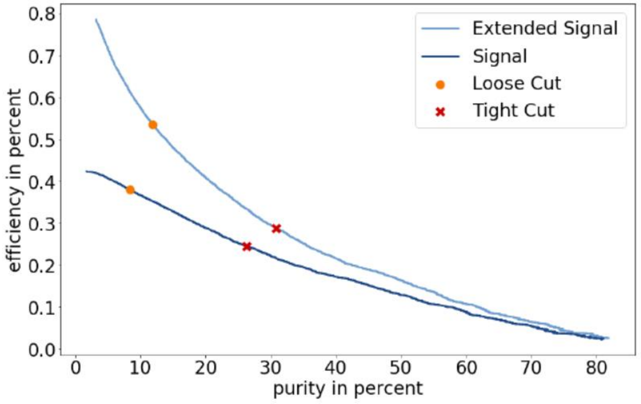
(cosine of angle difference between merged B daughters and nominal B)

Loose FEI output cut (σ > 0.01)

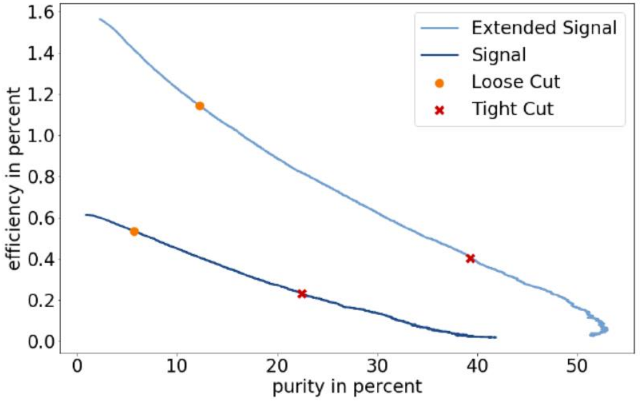


Tight FEI output cut (σ > 0.1)





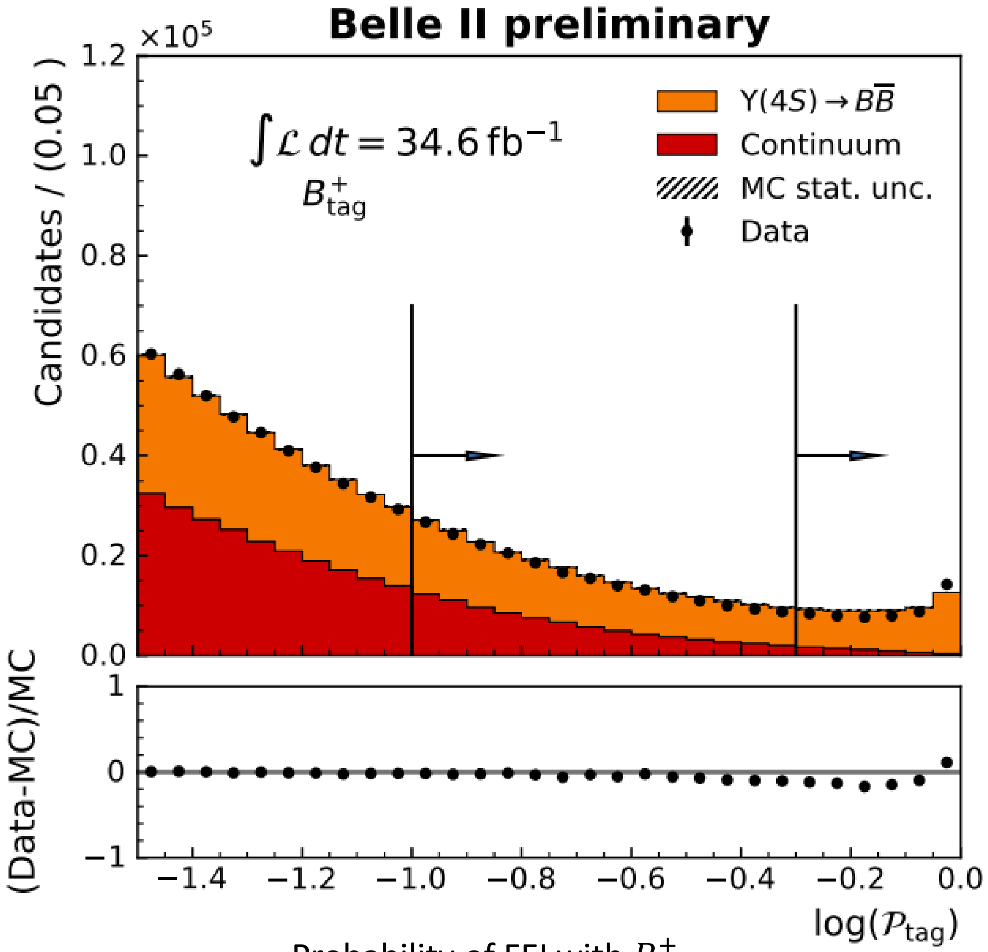
(a) Hadronic Tag



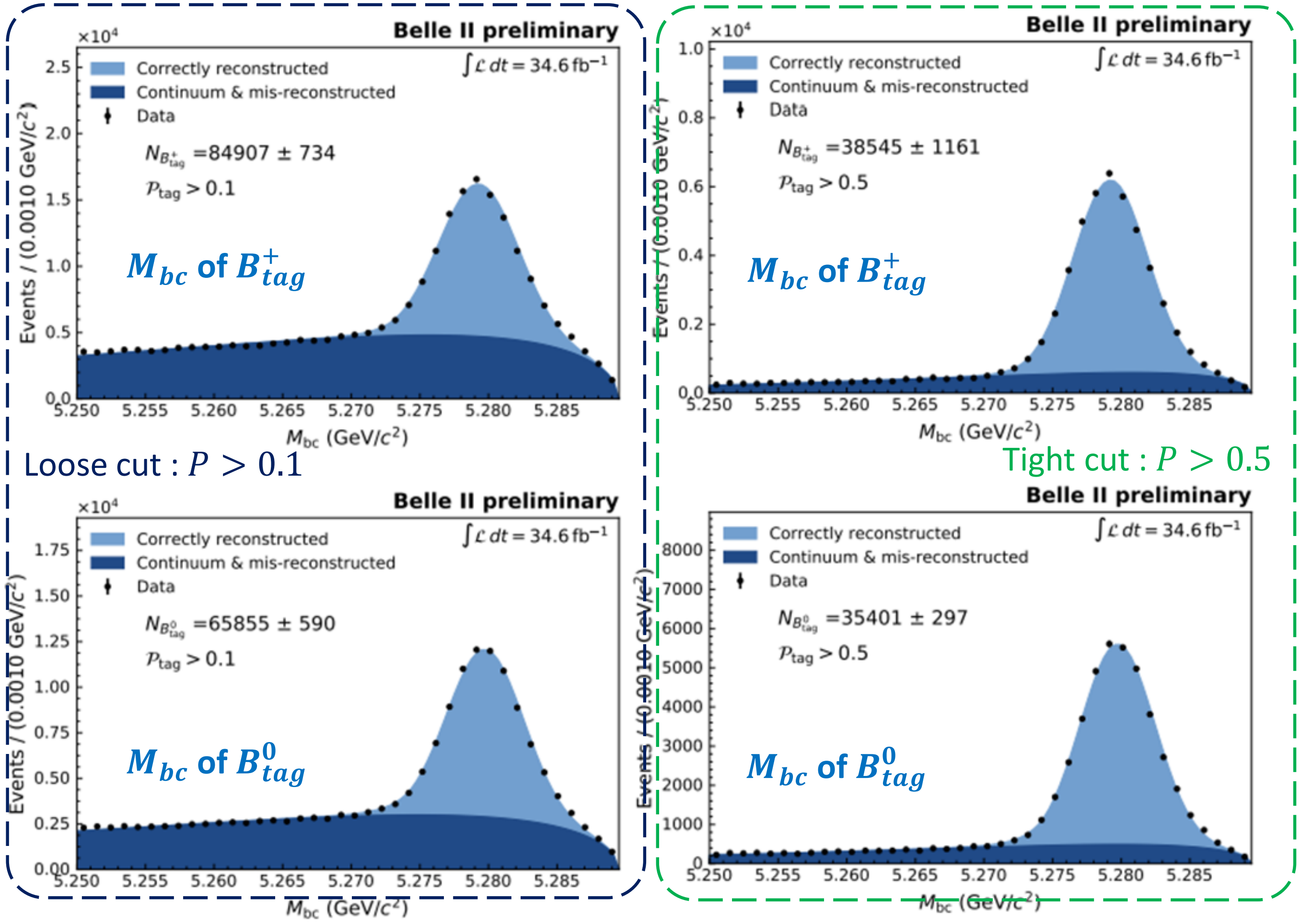
(b) Semileptonic Tag

Application : Hadronic FEI Performance ($B \to X\ell\nu$)

Used data & MC samples : 34.6 fb^{-1} data, samples of 100 fb^{-1} generic BB decay and 100 fb^{-1} generic qq decay



Probability of FEI with B^+_{tag}



Distribution of M_{bc}

Application : SL B⁰ FEI Performance ($B^0 \to \ell^{\pm}\tau^{\mp}$)

Amounts of sample

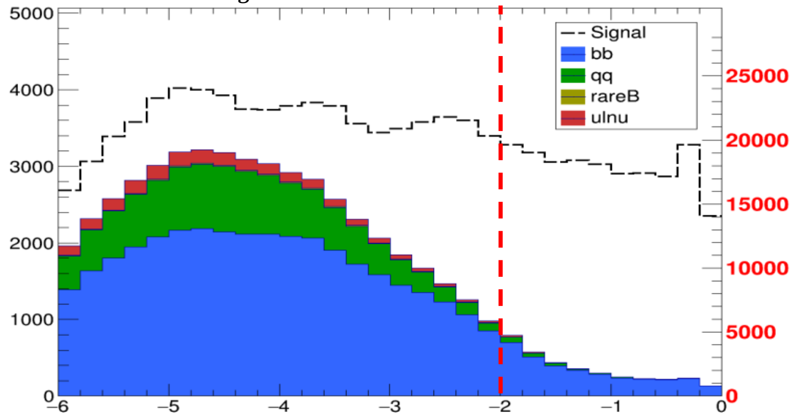
Type	Signal	bb	qq	Rare B	$u\ell\nu$
Amounts (million)	20	754	2350	5.66	5.47

- FEI output cut : logOtag > -2 (FEI output > 0.01)

Decay mode	$e-e$	$e-\mu$	$\mu-e$	$\mu-\mu$
# of signal	639195	409620	567436	368965
generic bb	167558.3	84725.3	151721.1	91767.2
generic qq	45065.2	9030.5	19905.5	14111.7
rare $B\bar{B}$ decay	865.6	263.2	546.4	411.5
$b \to u\ell\nu$	5136.8	3064.3	4139.8	2856.4

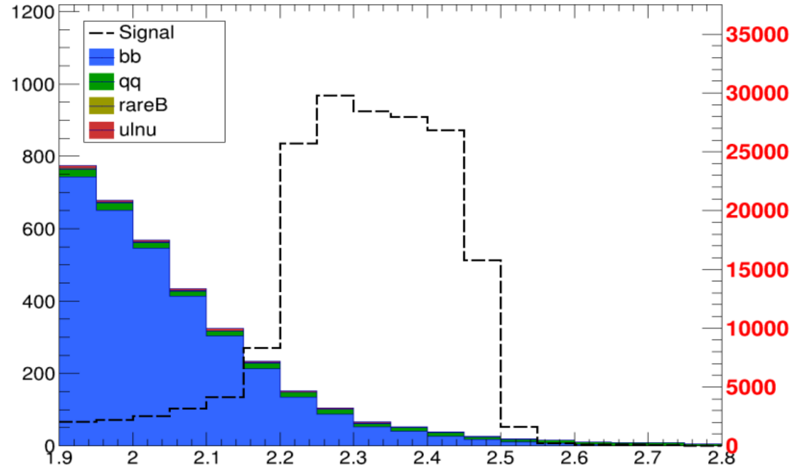
Amounts of the signal and background events after FEI and best candidate selection

$\log_{10} O_{tag}$ of $B^0 \to e^{\pm}\tau^{\mp}(\to e^{\mp}\nu\nu)$ mode



Distribution of FEI output and precut

p_ℓ^* of $B^0 \to e^{\pm}\tau^{\mp}(\to e^{\mp}\nu\nu)$ mode

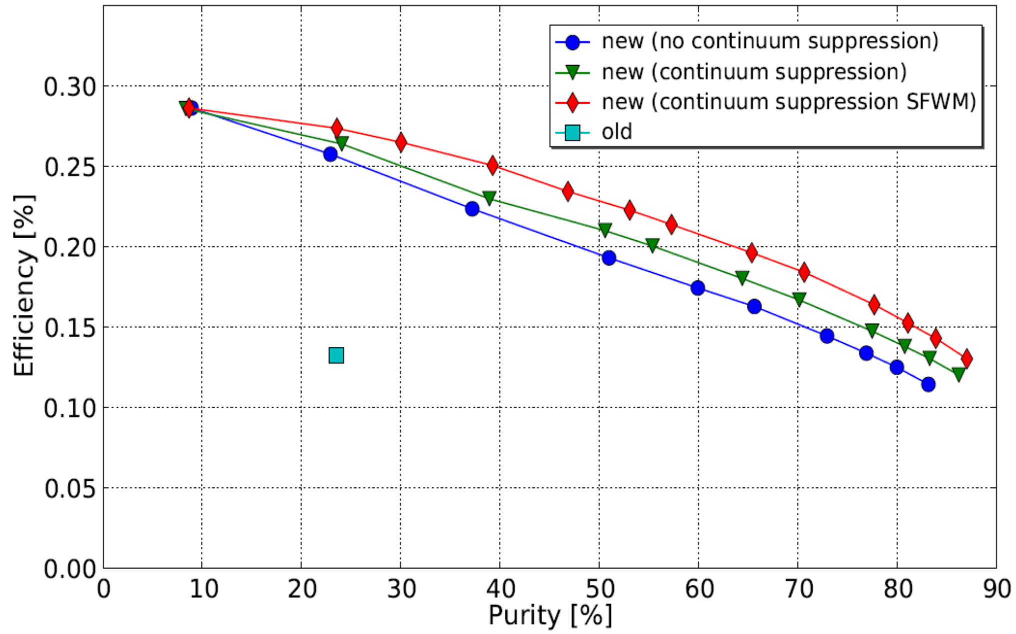


Distribution of p_ℓ^* after FEI output cut

Performance Check : Full Reconstruction

Reconstruction with Full Reconstruction

- Applied with 771.6M $B\bar{B}$ pair event samples
- 2.1M $B^{\pm}$ and 1.4M B^{0} remained as maximum efficiency case (0.28% for $B^{\pm}$ and 0.18% for B^{0})



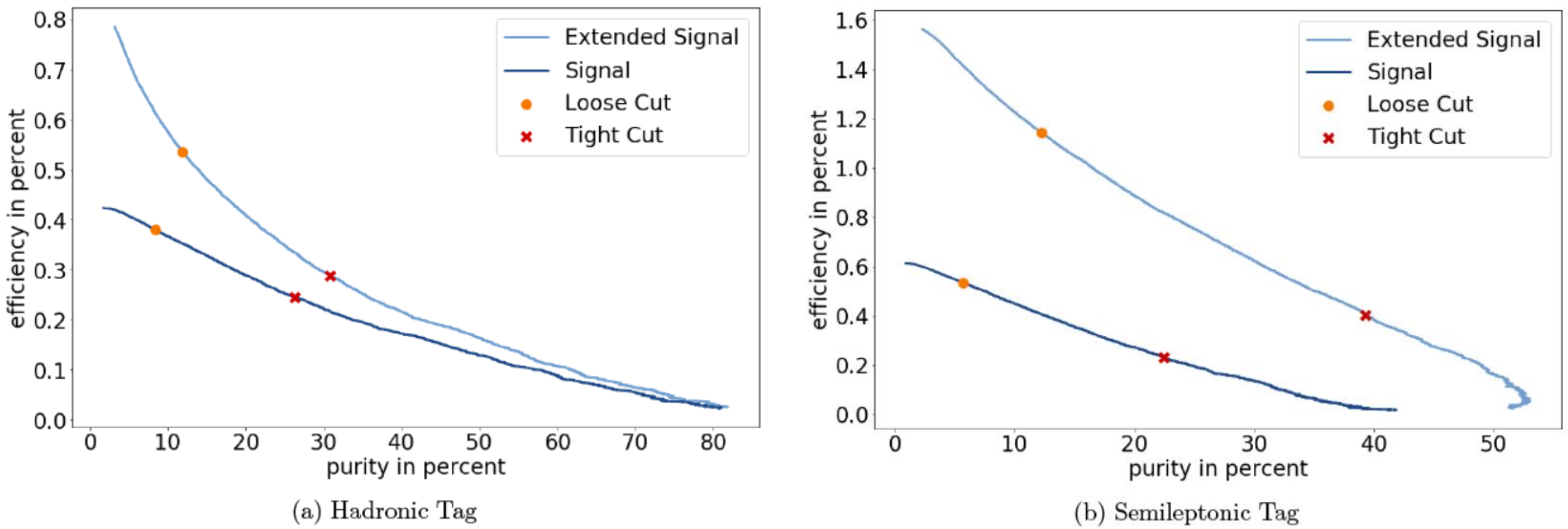
FR Purity-efficiency plot for B^{+} mesons

From https://arxiv.org/abs/1102.3876

Performance Check : Full Event Interpretation

Reconstruction with Full Event Interpretation

- 180M BB pair signal
- $1ab^{-1}$ (1.4 times to FR luminosity) scaled backgrounds from BB, e^+e^- to qq pair and $\tau^+\tau^-$



FEI Purity vs. efficiency distribution

From https://publikationen.bibliothek.kit.edu/1000078149

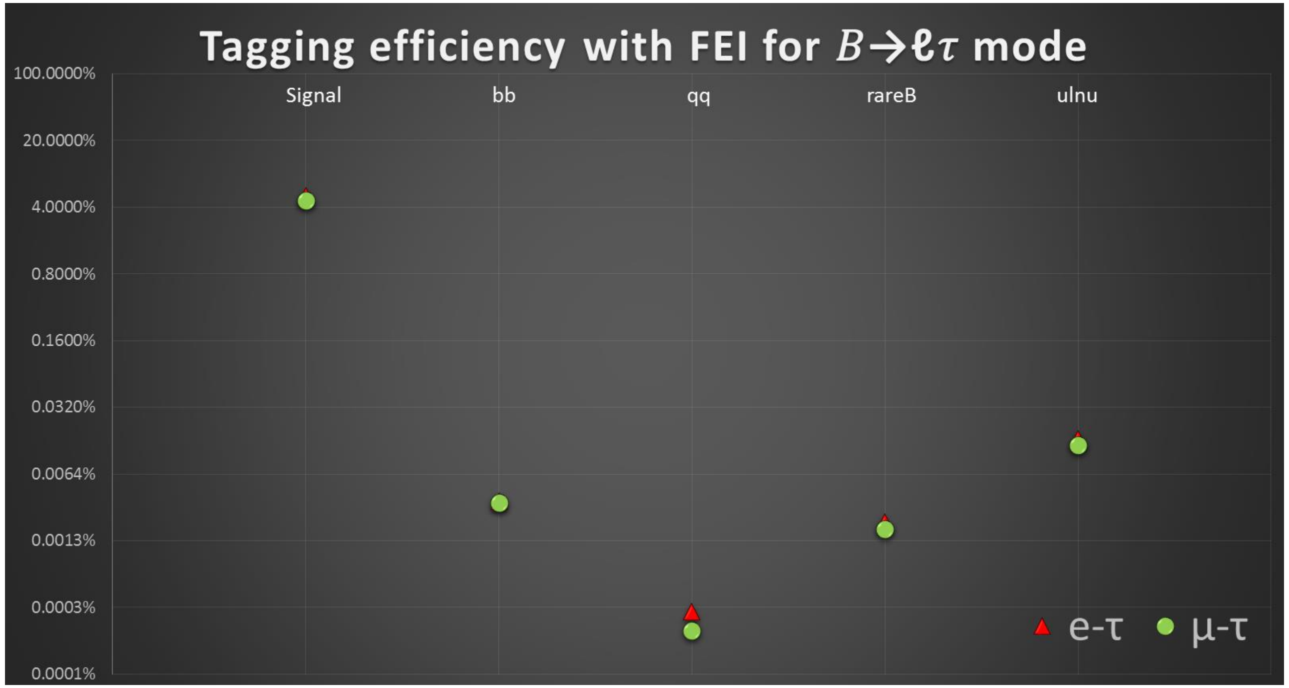
Application: $B^0 \to \ell^\pm \tau^\mp$ Signal Mode

The properties of the signal

- τ decay should include invisible particles (neutrino)
- Lack of signal mode for B reconstruction: needs more information for quality control

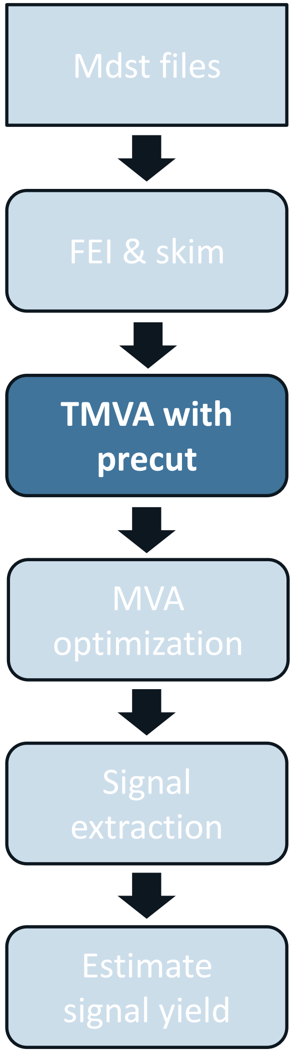
4% signal tagging efficiency & background suppression

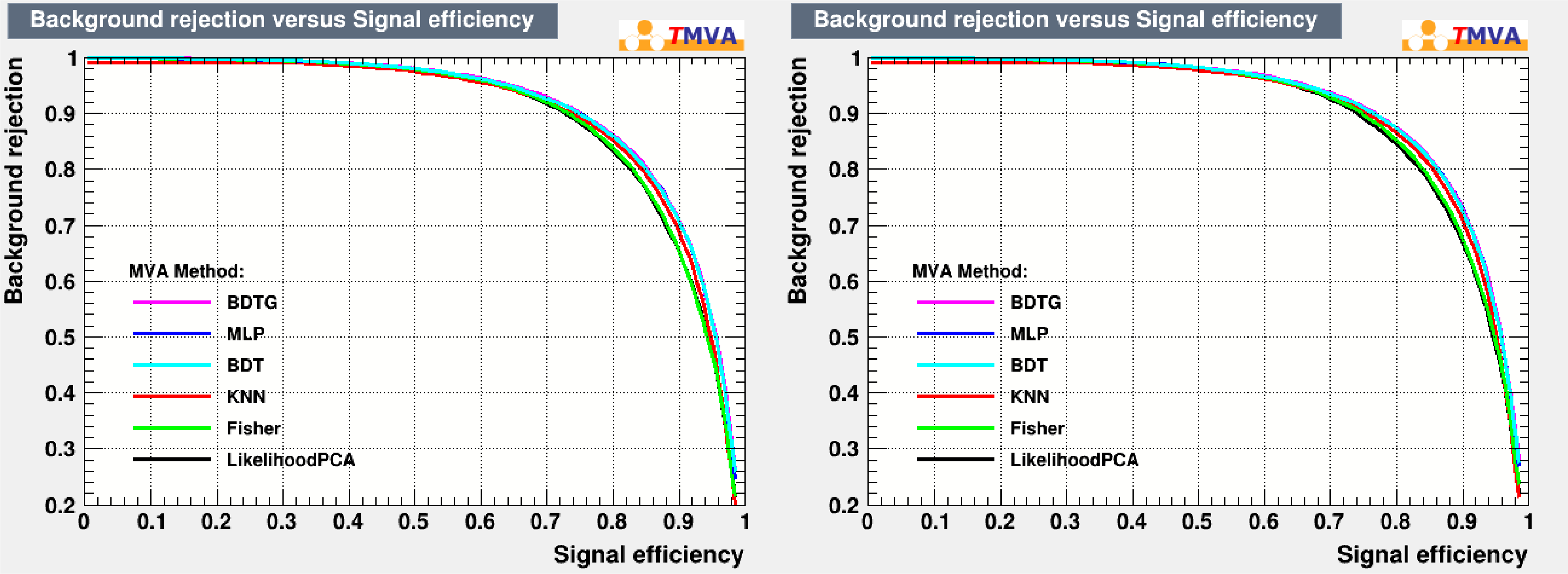
Studying this modes now for new physics study (lepton flavor violation research)



From https://www.kps.or.kr/conference/event/content/program/search_result_abstract_poster.php?id=4254&tid=503

TMVA Distributions: ROC Curve





BRIEF Estimation of MC Upper Limit

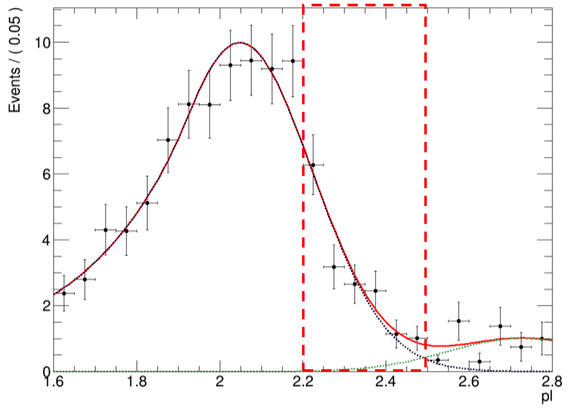
Estimating signal yields from background PDF

- The amounts of backgrounds in the signal region

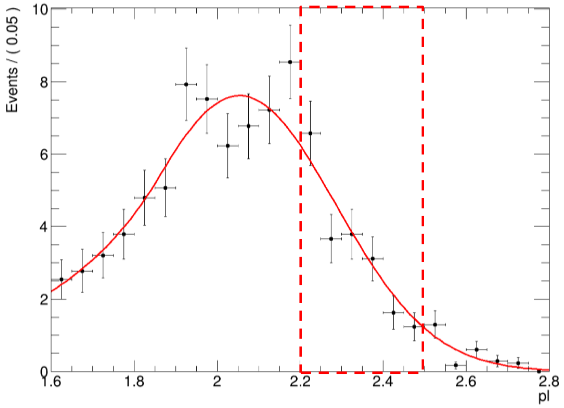
(2.2 < p_ℓ^* < 2.5)

	e-τ	μ-τ
MC event	16.7	20.00
PDF region	17.34	20.95

PDF for e-τ



PDF for μ-τ



Calculation of the branching faction

- $\Gamma\left(B^0 \to \ell^\pm\tau^\mp\right) = \frac{N_{obs}-N_{bkg}^{exp}}{\epsilon_{sig}^{\ell-\tau}\times N_{B\bar{B}}}$
  - $\epsilon_{sig}^{\ell-\tau} = \epsilon_{sig}^{\ell-e} \times \Gamma(\tau \to e\nu\nu) + \epsilon_{sig}^{\ell-\mu} \times \Gamma(\tau \to \mu\nu\nu)$
    $= \epsilon_{sig}^{\ell-e} \times 0.1782 + \epsilon_{sig}^{\ell-\mu} \times 0.1739$
  - N_{obs} and N_{bkg}^{exp} : summation of sub-mode ($\ell - e$, $\ell - \mu$ mode)
  - By calculating MC upper limit of branching fraction, the upper limit of $\left[N_{obs} - N_{bkg}^{exp}\right]$ is calculated by POLE

	e-τ	μ-τ
ϵ_{sig}	7.66E-4	7.27E-4
N_{obs}	17	20
N_{bkg}^{exp}	17.34	20.95
POLE range $\left[N_{obs} - N_{bkg}^{exp}\right]$	0 - 7.89	0 – 7.82
MC UL	1.36E-5	1.43E-5