R(D) and R(D^*) with an inclusive tagging method at Belle II

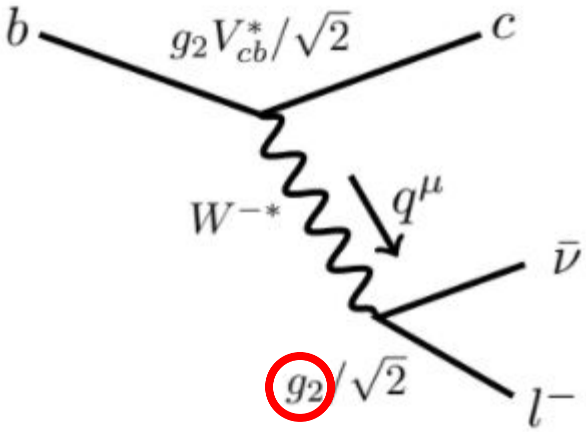
Boyang Zhang
zhangboy@hawaii.edu
2022 Belle II Summer Workshop

Contents

R(D) & R(D*)

Motivation:

1. Direct tests for Lepton Flavor Universality in the weak interaction;



$$\Gamma\left(B \to D\, e\, \upsilon_e\right) \propto \left(\frac{g_2}{M_W}\,\frac{g_e}{M_W}\right)^2 |V_{cb}|^2\, m_B^5\, F_{B\to D}^{e,\,\mu}\left(q^2\right) PS^e$$

$$R(D^*)_{light} = \frac{BF\left(B \to D^*\, e\, \upsilon_e\right)}{BF\left(B \to D^*\, \mu\, \upsilon_\mu\right)} = \frac{\Gamma\left(B \to D^*\, e\, \upsilon_e\right)}{\Gamma\left(B \to D^*\, \mu\, \upsilon_\mu\right)} = \left(\frac{g_e}{g_\mu}\right)^2 \frac{F_{B\to D^*}^{e,\,\mu}\left(q^2\right)}{F_{B\to D^*}^{e,\,\mu}\left(q^2\right)}\,\frac{PS^e}{PS^\mu}$$

$$\mathcal{R}(D^{(*)})_{\rm light} = 1.01 \pm 0.01 \pm 0.03$$

Waheed et al. Belle Collaboration 2019

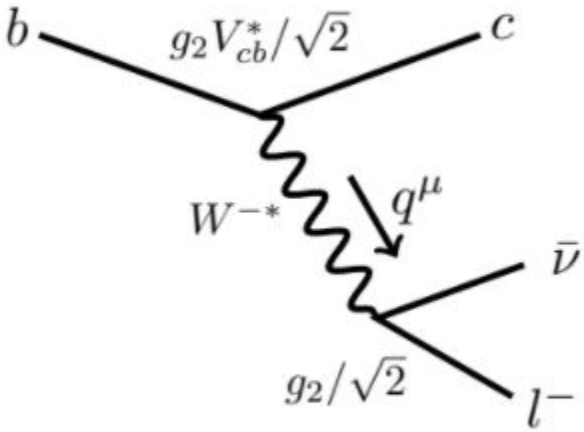
$$R(X_{e/\mu})^{p_\ell^*>1.3\mathrm{GeV}} = 1.033 \pm 0.010^{\rm stat} \pm 0.020^{\rm syst}$$

ICHEP, H. Junkerkalefeld, Belle II, 2022

R(D) & R(D*)

Motivation:

1. Direct tests for Lepton Flavor Universality in the weak interaction;
2. Sensitive probes for new physics (e.g. leptoquarks);
3. Can be measured with high precision (recon the signal and norm modes with the same procedure, most of systematic errors cancel, except lepton efficiency etc.);



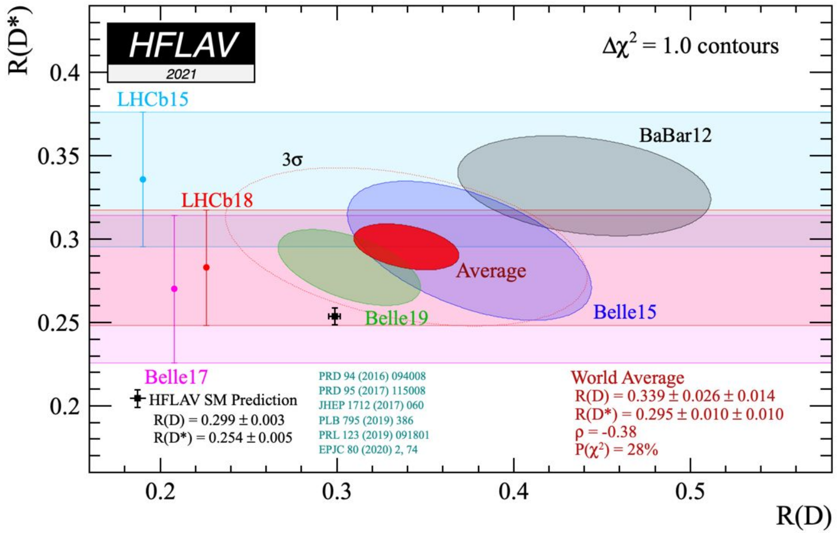
$$R(D) = \frac{BF(B \to D\,\tau\,\upsilon_\tau)}{BF(B \to D\,l\,\upsilon_l)} = \frac{\Gamma(B \to D\,\tau\,\upsilon_\tau)}{\Gamma(B \to D\,l\,\upsilon_l)} = \left(\frac{g_\tau}{g_{e,\,\mu}}\right)^2 \frac{F^\tau_{B\to D}(q^2)}{F^{e,\,\mu}_{B\to D}(q^2)}\,\frac{PS^\tau}{PS^{e,\,\mu}}$$

$l = e$ or μ

Measured and SM predicted R(D) & R(D*)

Obs.	Current World Av./Data	Current SM Prediction	Significance	
$\mathcal{R}(D)$	0.340 ± 0.030	0.299 ± 0.003	1.2σ	3.1σ
$\mathcal{R}(D^*)$	0.295 ± 0.014	0.258 ± 0.005	2.5σ	

Heavy FLavor AVeraging group has calculated world averages from all the available data



Semileptonic tagging R(D)(*)

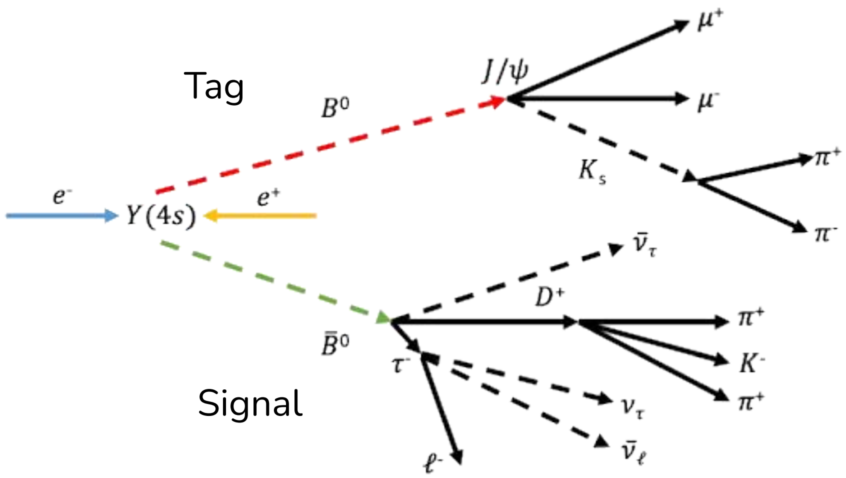
Result	Contribution	Uncertainty [%] Sys.	Stat.
$\mathcal{R}(D)$	$B \to D^{**}\ell\bar{\nu}_\ell$	0.8	
	PDF modeling	4.4	
	Other bkg.	2.0	
	$\epsilon_{\rm sig}/\epsilon_{\rm norm}$	1.9	
	Total systematic	**5.2**	
	Total statistical		**12.1**
	Total	**13.1**	
$\mathcal{R}(D^*)$	$B \to D^{**}\ell\bar{\nu}_\ell$	1.4	
	PDF modeling	2.3	
	Other bkg.	1.4	
	$\epsilon_{\rm sig}/\epsilon_{\rm norm}$	4.1	
	Total systematic	**4.9**	
	Total statistical		**6.4**
	Total	**8.1**	

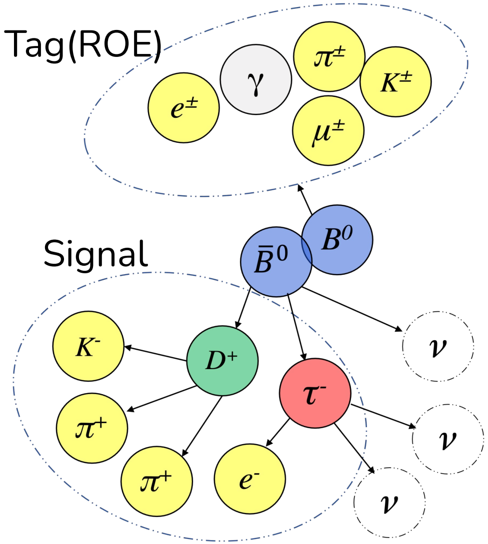
Belle19

More data or new method is needed
→ Belle II experiment and inclusive tagging method

Tagging methods

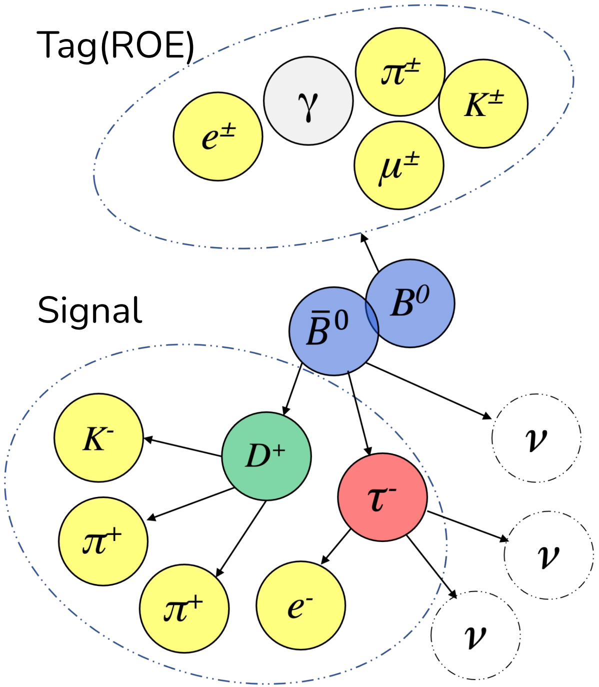
If one of the B mesons decays to a final state involving neutrinos, this B meson cannot be reconstructed completely.





Tagging Types		Advantage	basf2 algorithm	Reconstruction efficiency
Exclusive	Hadronic	Precise kinematics High purity	Full Event Interpretation (FEI)	<1%
	Semileptonic			~1%
Inclusive	-	High reco efficiency	Rest Of Event (ROE)	>10%

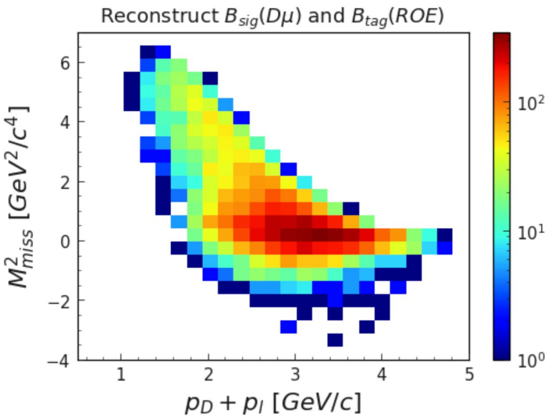
1. Introduction
2. Workflow
3. Event reconstruction
4. BDTs
5. Fitting and yield extraction
6. Summary

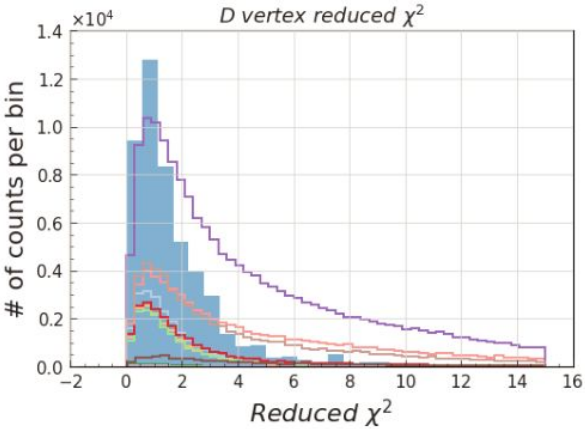


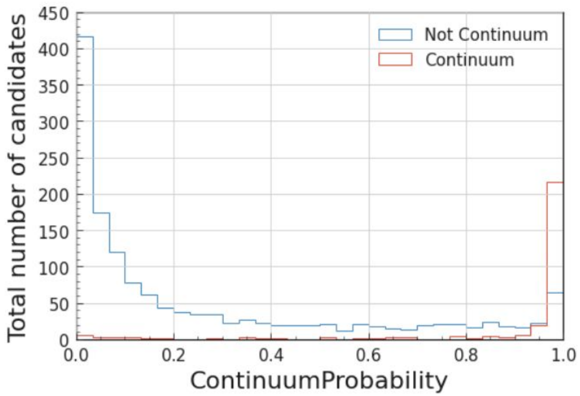
Workflow

Event reconstruction
Reconstruct signal and normalization modes simultaneously.

Additional bkg suppression (MVA)
1. Continuum suppression
2. Fake D suppression
3. Other BBbar physics bkg

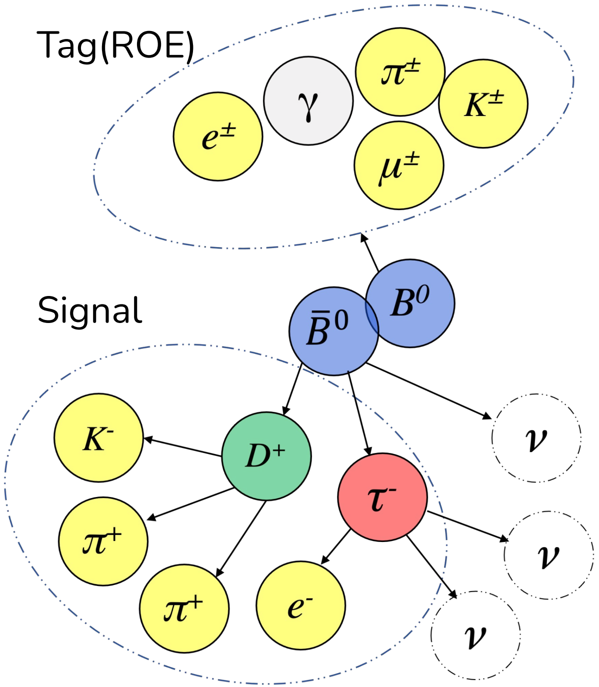






Yields extraction from 2D fit
2d fit: M^2_{miss} vs. ($|p_D| + |p_l|$)

1. Introduction
2. Workflow
3. Event reconstruction
4. BDTs
5. Fitting and yield extraction
6. Summary



D decay modes

$B^0 \to D^+ \tau^- \nu$	Boyang
D^+ decay modes	Branching fraction
$D^+ \to K^- \pi^+ \pi^+$	(9.38 ± 0.16) %

$B^- \to D^0 \tau^- \nu$	Tia
D^0 decay modes	Branching fraction
$D^0 \to K^- \pi^+$	(3.88 ± 0.05) %
$D^0 \to K^- \pi^+ \pi^0$	(13.9 ± 0.5) %
$D^0 \to K^- \pi^+ \pi^+ \pi^-$	(8.08 ± 0.20) %

D0 has more D* and D** feed down

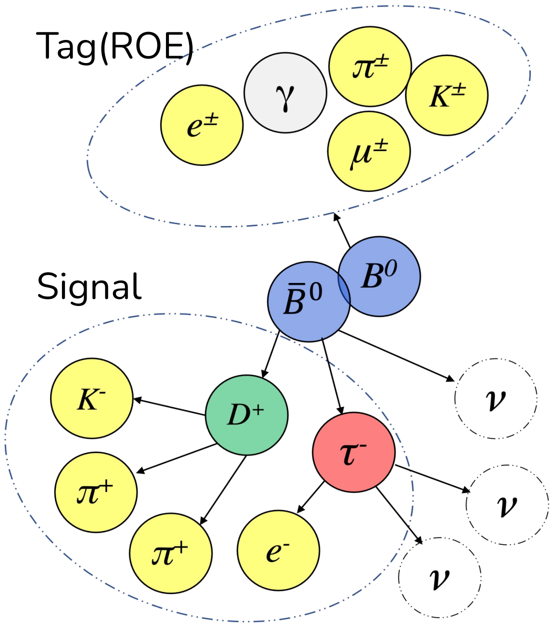
D+ has longer lifetime so that provides better vertex separation/fitting

D+ $c\tau$ = 311.8μm
D0 $c\tau$ = 122.9μm

τ decay modes	Branching fraction
$\tau \to e \nu \nu$	(17.83 ± 0.04) %
$\tau \to \mu \nu \nu$	(17.41 ± 0.04) %

Event reconstruction for B -> D* l ...

Signal Side (Efficiency ~20%)	
Reconstruct a vertex whose daughters are D + l (...)	Select good tracks + bremsstrahlung photons Final state Pid cuts + lepton momentum cuts D mass cut + vertex fit (D vertex and B vertex)
ROE side (Efficiency ~15% with BDTs)	
	Track masks + ECL masks
Best Candidate Selection (~1.2 candidates per event)	Signal side vertex reduced chi2
This will reconstruct all at once	(D* and D** decay to D with extra pi or photon)
R(D) signal: B $\rightarrow$ D^{+} τ ν	R(D) Normalization: B $\rightarrow$ D^{+} l ν
R(D*) signal: B $\rightarrow$ D^{*+} τ ν	R(D*) Normalization: B $\rightarrow$ D^{*+} l ν
B $\rightarrow$ D** τ ν background	B $\rightarrow$ D** l ν background



MM^2 calculation

To separate the signals and normalizations, calculated with ROE variables

$$
\begin{aligned}
MM^2 &= \left(p^*_{tot} - p^*_{ROE} - p^*_{Y}\right)^2 \\
&= \left(E^*_{tot} - E^*_{ROE} - E^*_{Y}\right)^2 - \left(\vec{P}^*_{tot} - \vec{P}^*_{ROE} - \vec{P}^*_{Y}\right)^2 \\
&\approx \left(\frac{1}{2}E^*_{tot} - E^*_{Y}\right)^2 - \left(\vec{P}^*_{ROE} + \vec{P}^*_{Y}\right)^2 \\
&= \left(\frac{1}{2}E^*_{tot}\right)^2 + \left(E^*_{Y}\right)^2 - E^*_{tot} \times E^*_{Y} - \left(\left|\vec{P}^*_{ROE}\right|^2 + \left|\vec{P}^*_{Y}\right|^2 + 2\,\vec{P}^*_{ROE} \cdot \vec{P}^*_{Y}\right) \\
&= \left(\frac{1}{2}E^*_{tot}\right)^2 - \left|\vec{P}^*_{ROE}\right|^2 + \left(E^*_{Y}\right)^2 - \left|\vec{P}^*_{Y}\right|^2 - E^*_{tot} \times E^*_{Y} - 2\,\vec{P}^*_{ROE} \cdot \vec{P}^*_{Y} \\
&= \boxed{\left(M^{ROE}_{bc}\right)^2} + M^2_{Y} - E^*_{tot} \times E^*_{Y} - \boxed{2\,\vec{P}^*_{ROE}} \cdot \vec{P}^*_{Y}
\end{aligned}
$$

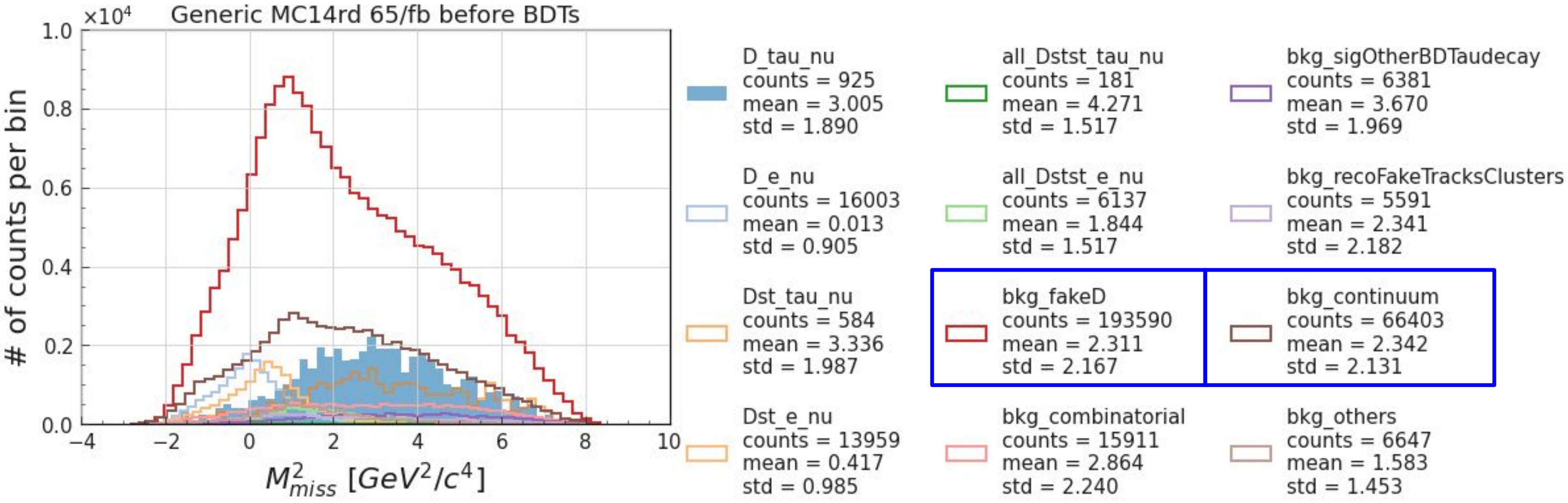
$$where\ p_Y = p_D + p_l$$

Reconstruction

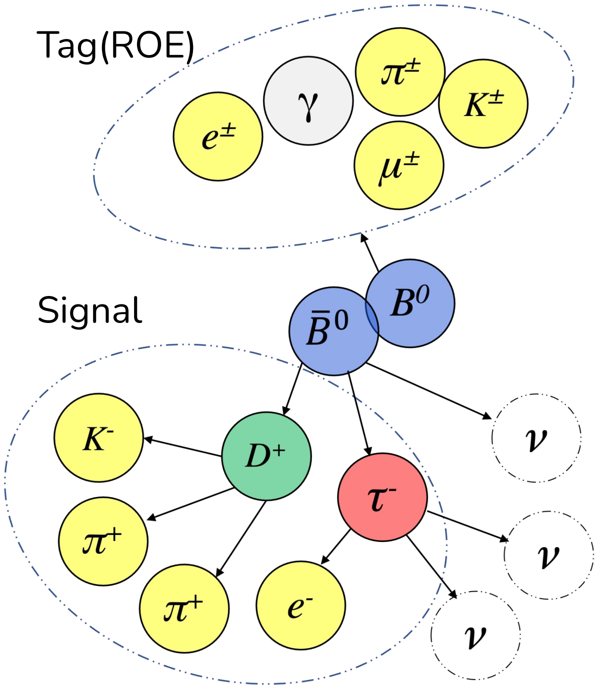
MC14rd generic (charged + mixed + qqbar + taupair)
65/fb
Reconstructed only in $\tau \rightarrow e\,\nu\,\nu$ mode;

Cut Flow	sigEff	bkgEff
No cut	100.00%	100.00%
D_vtxReChi2<13	95.22%	
B0_vtxReChi2<14	93.09%	
5.03<B0_roeMbc_my_mask	87.42%	
-3.5<B0_roeDeltae_my_mask<0.5	83.98%	
4.65<B0_CMS1_weMbc	81.13%	
-2.2<B0_CMS0_weDeltae<0.5	79.21%	
-3<B0_deltaE<-1	78.77%	
abs(B0_roeCharge_my_mask)<3	77.13%	
e_CMS_p > 0.2	77.12%	
B0_vetoeID 0.9 B0_vetomuID 0.9	53.65%	11.05%

Bottom plot: MM^2



1. Introduction
2. Workflow
3. Event reconstruction
4. BDTs
5. Fitting and yield extraction
6. Summary

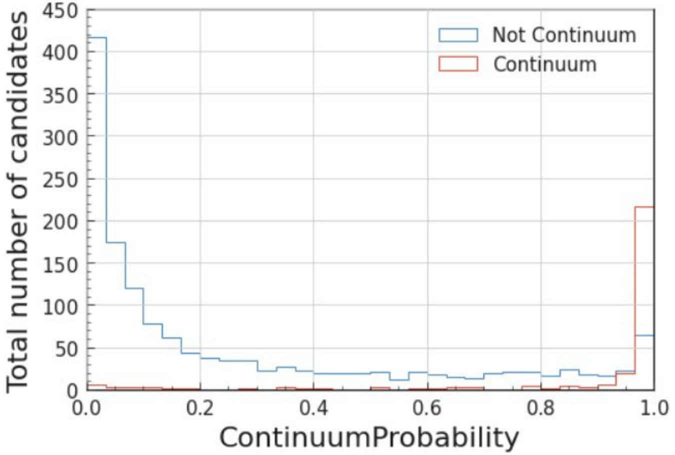


Fast BDTs:

1. Continuum suppression
2. Fake D
3. Other BBbar bkg

Hyperparameters:

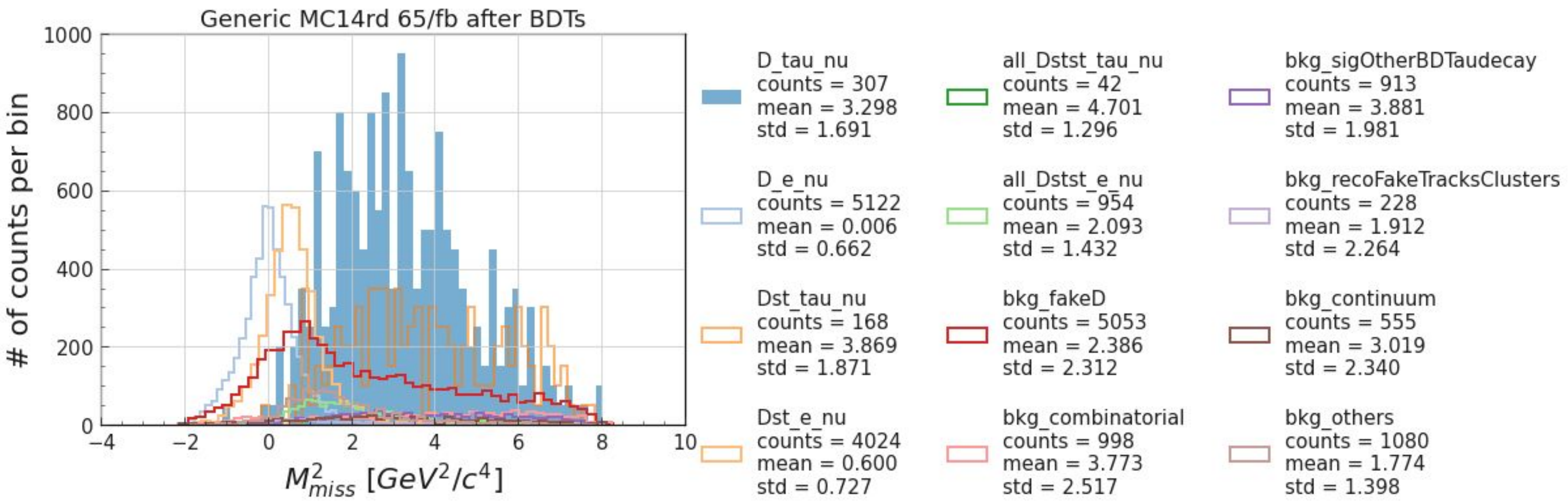
- nTrees = 1400
- Depth = 2
- Learning_rate = 0.05
- nCutLevels = 5
- Sub_sample_fraction = 0.5



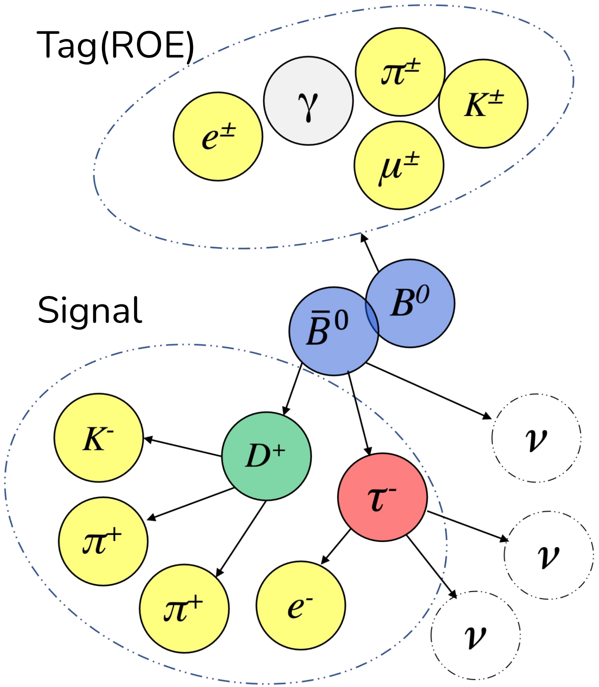
MM^2 after BDTs + roeMbc>5.26

MC14rd generic 65/fb

Reconstructed only in $\tau \rightarrow e \nu \nu$ mode;



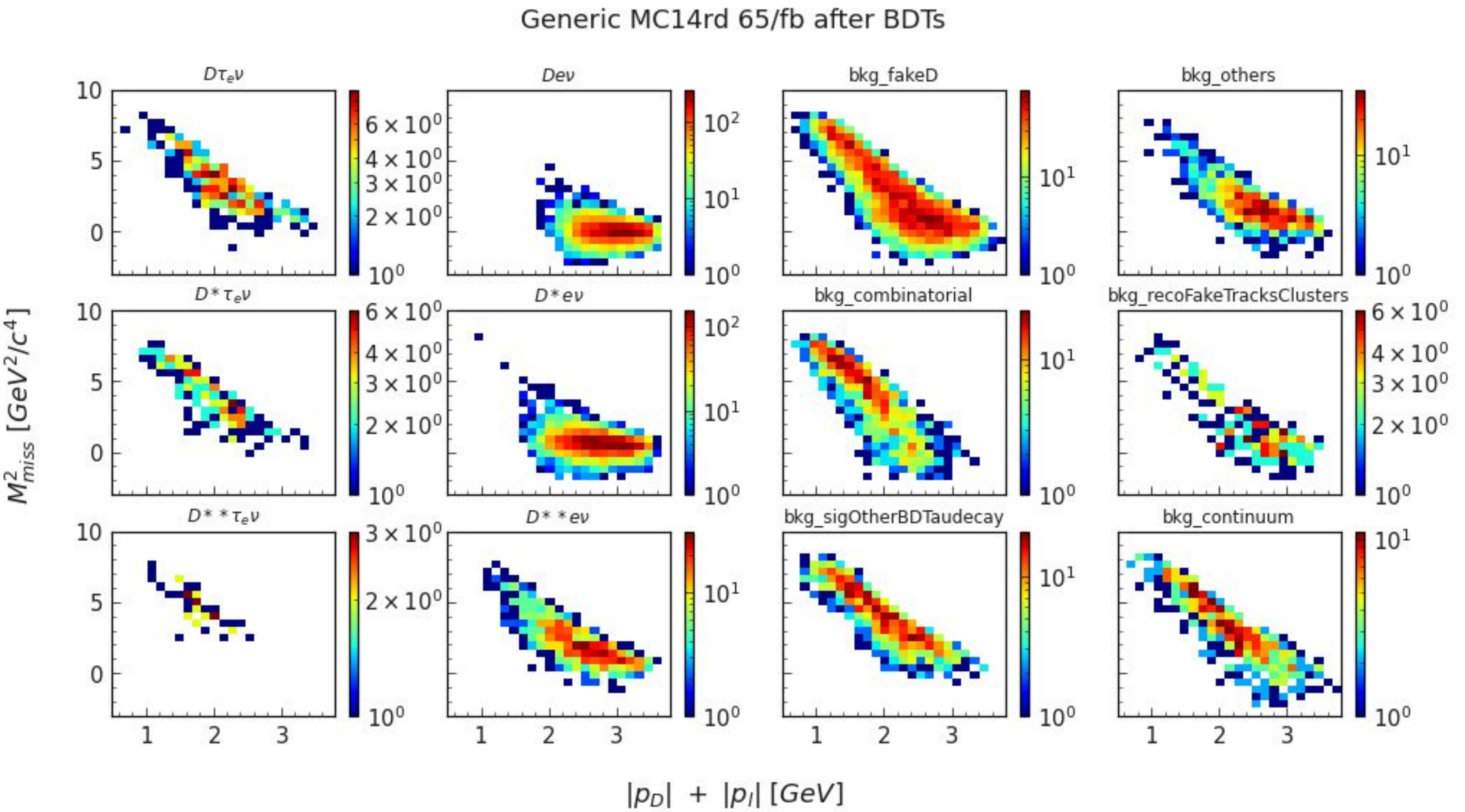
1. Introduction
2. Workflow
3. Event reconstruction
4. BDTs
5. Fitting and yield extraction
6. Summary



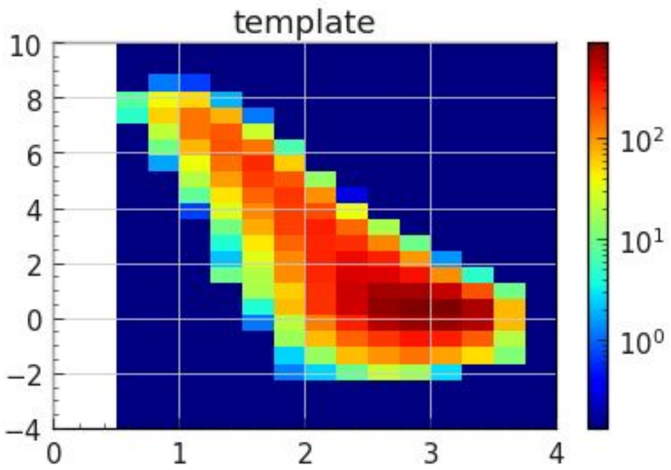
2D template fitting with pyhf:

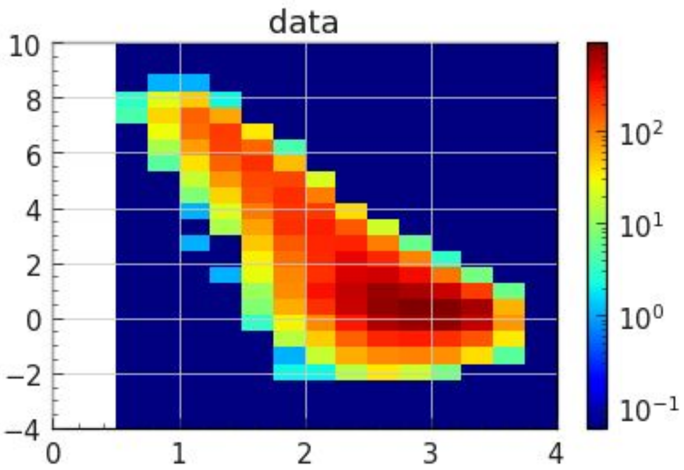
84/fb template
65/fb data(MC)

Templates


Generic MC14rd 65/fb after BDTs
Dτeν
Deν
bkg_fakeD
bkg_others
D*τeν
D*eν
bkg_combinatorial
bkg_recoFakeTracksClusters
D**τeν
D**eν
bkg_sigOtherBDTaudecay
bkg_continuum
M²miss [GeV²/c⁴]
|pD| + |pl| [GeV]

Fitting Results





D_tau_nu counts: 500
D_tau_nu counts uncertainty: 278

D_l_nu counts: 5139
D_l_nu counts uncertainty: 151

Dst_tau_nu counts: 2e-09
Dst_tau_nu counts uncertainty:84

Dst_l_nu counts: 3096
Dst_l_nu counts uncertainty: 202

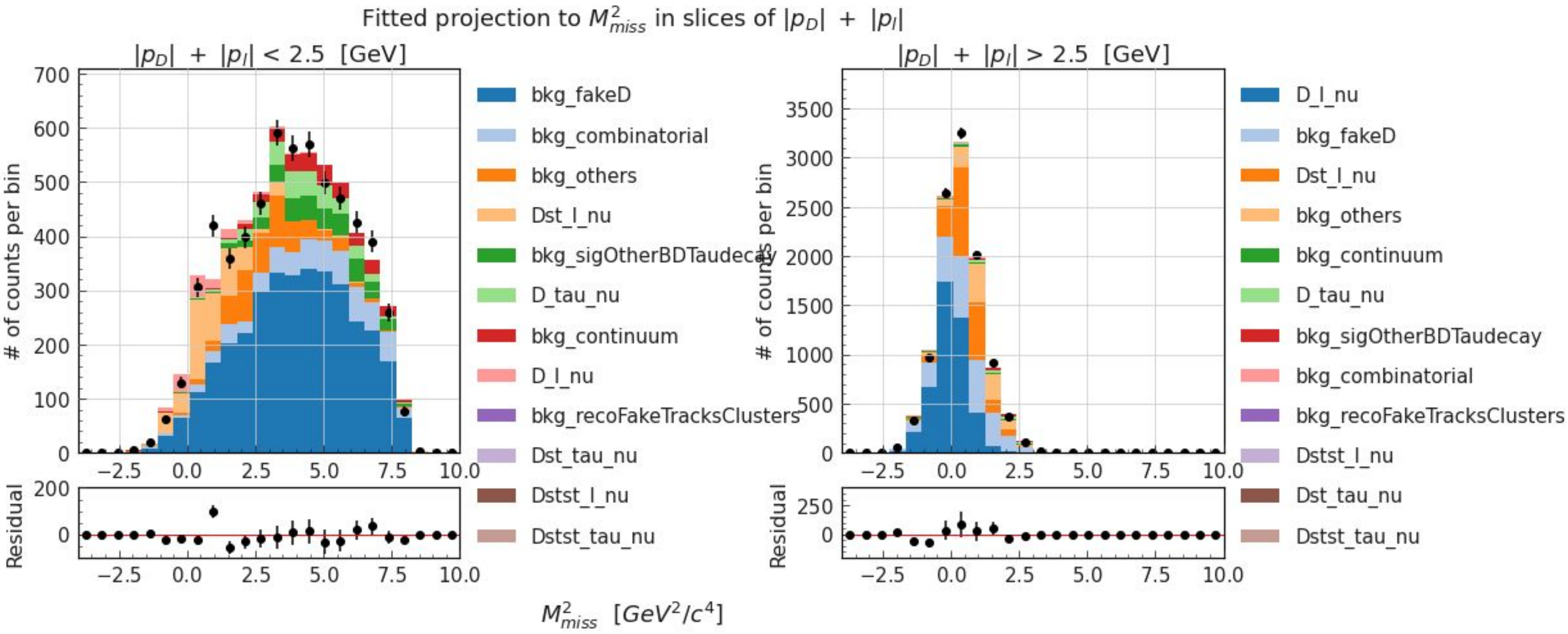
MC truth:
D_tau_nu: 307

D_l_nu: 5122

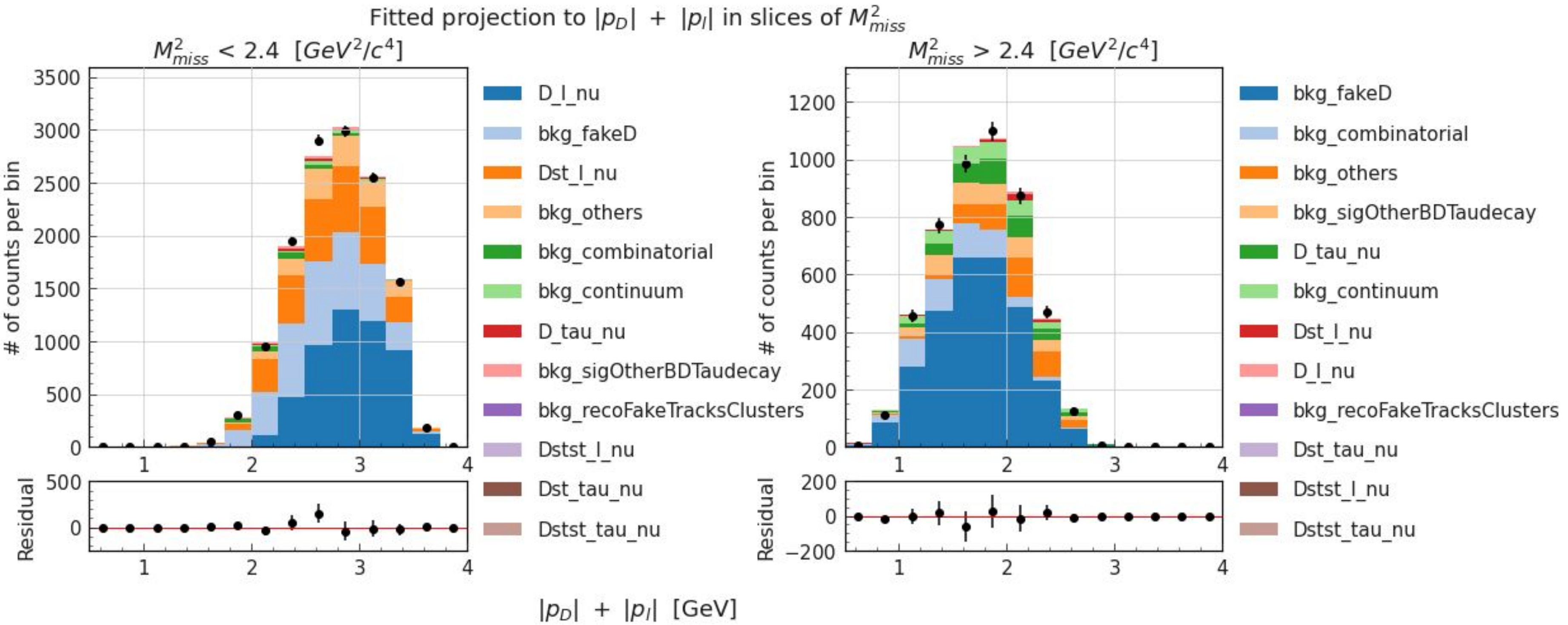
Dst_tau_nu: 168

Dst_l_nu: 4024

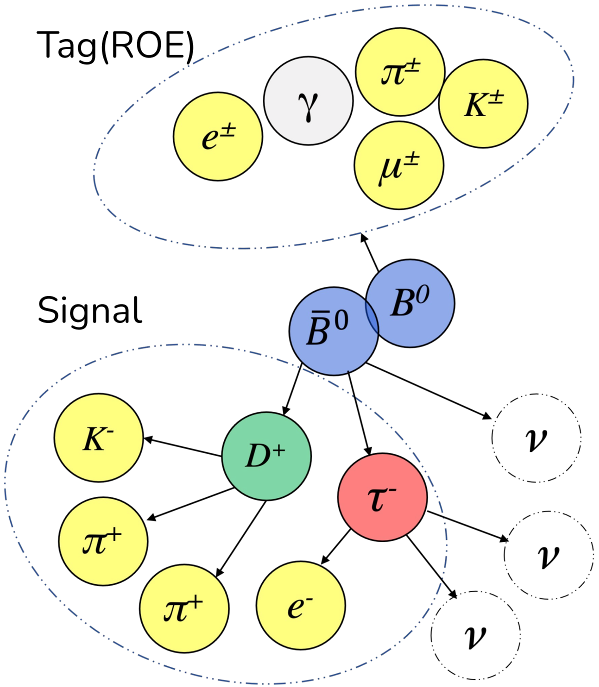
Projections of fitted templates



Projections of fitted templates



1. Introduction
2. Workflow
3. Event reconstruction
4. BDTs
5. Fitting and yield extraction
6. Summary



Summary

This analysis tests the lepton universality

Get R(D) and R(D*) with one reconstruction [optimized based on R(D)]

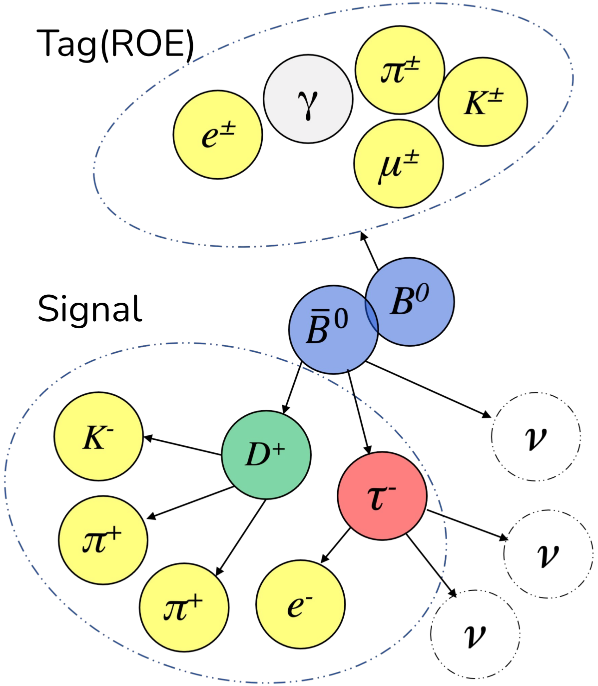
Many systematic uncertainties cancel due to the double ratio

Use BDTs to maximize the separating power between different components

Tagging efficiency is high, ~15%:

- Everything is done with only 84/fb (BDT training, fitting). Performance can be improved with bigger sample 700/fb

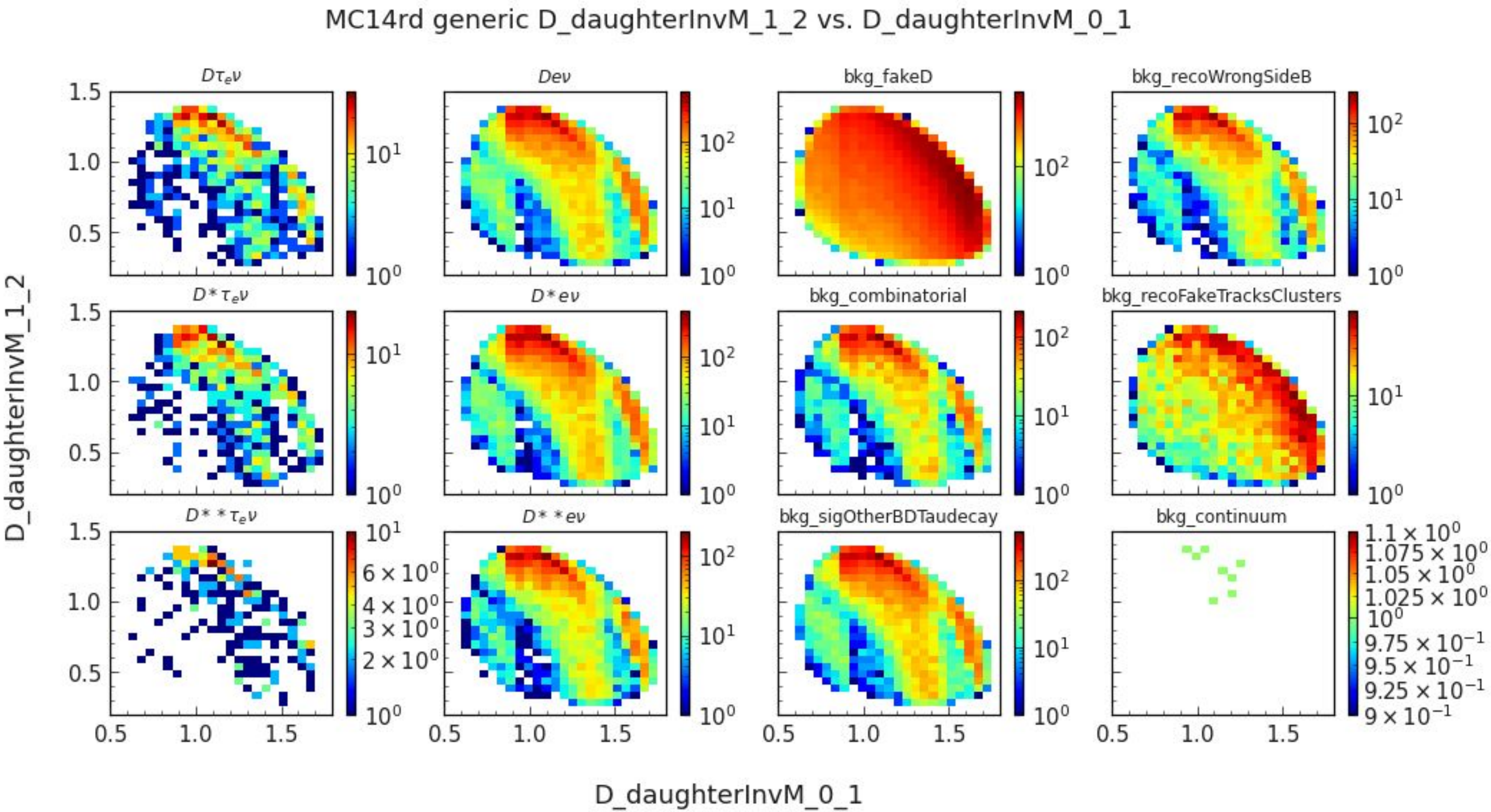
B tagging	Experiment	Algorithm	$B^{\pm}$	B^0
Hadronic	Belle II	FEI	0.76%	0.46%
	Belle II	FEI (FR channels)	0.53%	0.33%
	Belle	FR	0.28%	0.18%
	BABAR	SER	0.4%	0.2%
Semileptonic	Belle II	FEI	1.80%	2.04%
	Belle	FR	0.31%	0.34%
	BABAR	SER	0.3%	0.6%



Thank you

Backup

Dalitz plot used in BDT



Fitting Results sig

Fitted:

D_tau_nu counts: 500
D_tau_nu counts uncertainty: 278

D_l_nu counts: 5139
D_l_nu counts uncertainty: 151

Dst_tau_nu counts: 2e-09
Dst_tau_nu counts uncertainty: 84

Dst_l_nu counts: 3096
Dst_l_nu counts uncertainty: 202

Dstst_tau_nu counts: 8e-13
Dstst_tau_nu counts uncertainty: 41

Dstst_l_nu counts: 4e-09
Dstst_l_nu counts uncertainty: 72

MC truth:

D_tau_nu: 307

D_l_nu: 5122

Dst_tau_nu: 168

Dst_l_nu: 4024

Dstst_tau_nu: 42

Dstst_l_nu: 952

Fitting Results bkg

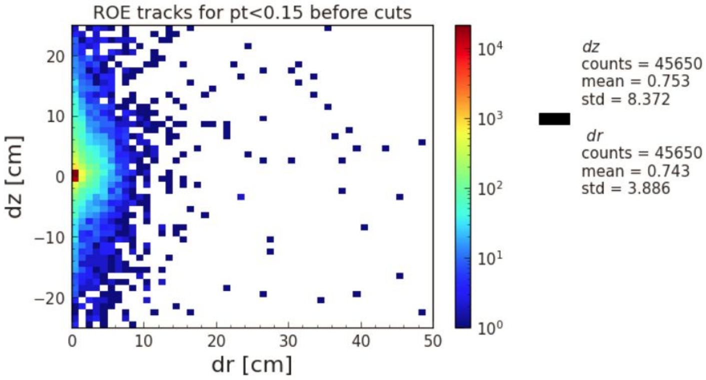
Fitted:

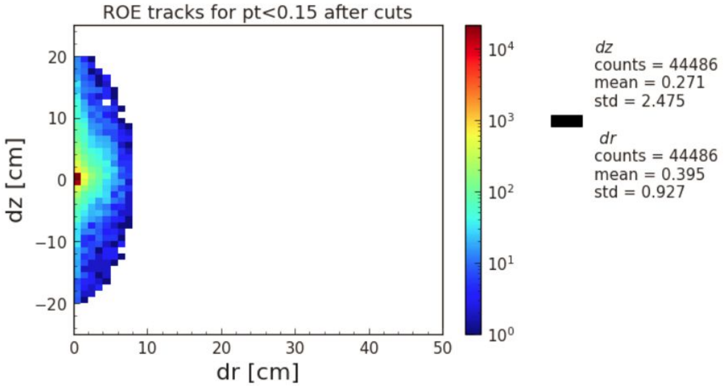
bkg_fakeD counts:	7104
bkg_fakeD counts uncertainty:	522
bkg_combinatorial counts:	764
bkg_combinatorial counts uncertainty:	187
bkg_sigOtherBDTaudecay counts:	521
bkg_sigOtherBDTaudecay counts uncertainty:	236
bkg_recoFakeTracksClusters counts:	3e-07
bkg_recoFakeTracksClusters counts uncertainty:	212
bkg_continuum counts:	437
bkg_continuum counts uncertainty:	202
bkg_others counts:	1902
bkg_others counts uncertainty:	188

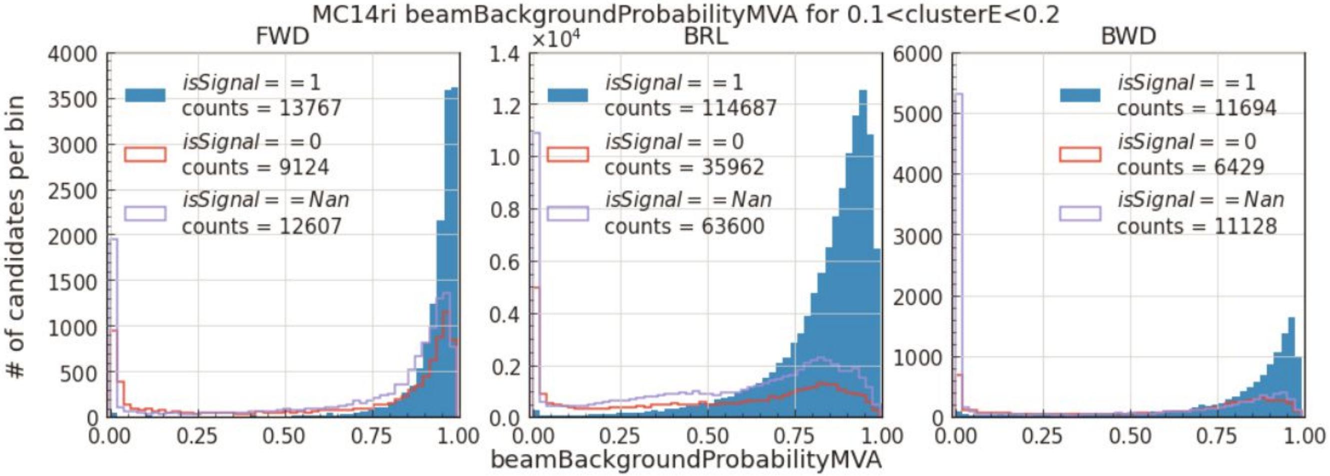
MC truth:

fake_D:	5053
combinatorial:	998
OBDTau:	913
FakeT/C:	228
Continuum:	555
Others:	1080

Optimization of ROE mask







Optimization of ROE mask (tracks)

Tracks	ECL Clusters
nCDCHits>0 and thetaInCDCAcceptance and pt>0.075 and [pt<0.15 and (dr^2/64+dz^2/400)<1] or [0.15<pt<0.25 and (dr^2/64+dz^2/225)<1] or [0.25<pt<0.5 and (dr^2/36+dz^2/100)<1] or [0.5<pt<1 and (dr^2/9+dz^2/16)<1] or [pt>1 and (dr^2/0.64+dz^2)<1]	goodGamma and abs(clusterTiming)<clusterErrorTiming and [E<0.1 and beamBkgMVA>0.2 and minC2TDist>25] [0.1<E<0.2 and beamBkgMVA>0.4 and minC2TDist>25 and clusterZernikeMVA>0.05] or [0.2<E<0.5 and beamBkgMVA>0.4 and minC2TDist>20 and clusterZernikeMVA>0.05] or [E>0.5 and beamBkgMVA>0.5]

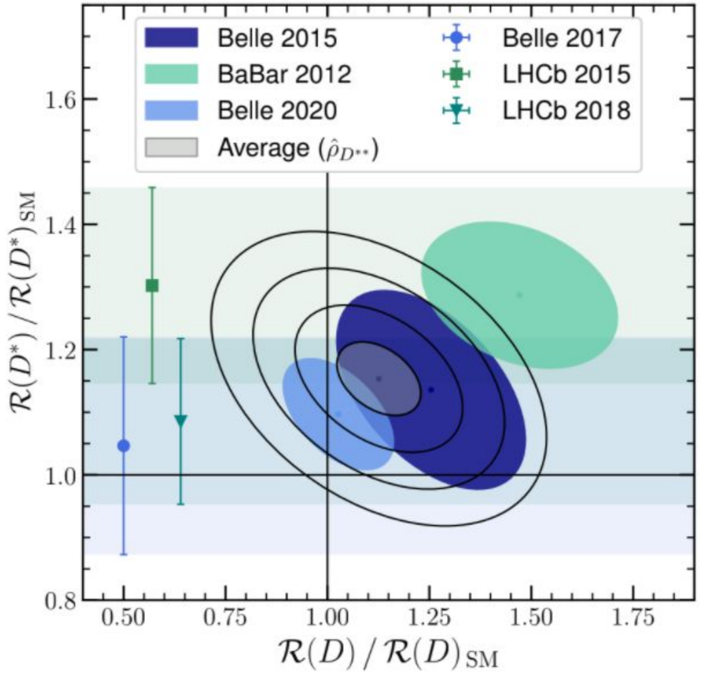
Uncertainties

Result	Experiment	τ decay	Tag	Systematic uncertainty [%]					Total uncert. [%]		
				MC stats	$D^{(*)}l\nu$	$D^{**}l\nu$	Other bkg.	Other sources	**Syst.**	**Stat.**	**Total**
$\mathcal{R}(D)$	*BABAR* [a]	$\ell\nu\nu$	Had.	5.7	2.5	5.8	3.9	0.9	**9.6**	**13.1**	**16.2**
	Belle[b]	$\ell\nu\nu$	Semil.	4.4	0.7	0.8	1.7	3.4	**5.2**	**12.1**	**13.1**
	Belle[c]	$\ell\nu\nu$	Had.	4.4	3.3	4.4	0.7	0.5	**7.1**	**17.1**	**18.5**
$\mathcal{R}(D^*)$	*BABAR* [a]	$\ell\nu\nu$	Had.	2.8	1.0	3.7	2.3	0.9	**5.6**	**7.1**	**9.0**
	Belle[b]	$\ell\nu\nu$	Semil.	2.3	0.3	1.4	0.5	4.7	**4.9**	**6.4**	**8.1**
	Belle[c]	$\ell\nu\nu$	Had.	3.6	1.3	3.4	0.7	0.5	**5.2**	**13.0**	**14.0**
	Belle[d]	$\pi\nu$, $\rho\nu$	Had.	3.5	2.3	2.4	8.1	2.9	**9.9**	**13.0**	**16.3**
	LHCb[e]	$\pi\pi\pi(\pi^0)\nu$	—	4.9	4.0	2.7	5.4	4.8	**10.2**	**6.5**	**12.0**
	LHCb[f]	$\mu\nu\nu$	—	6.3	2.2	2.1	5.1	2.0	**8.9**	**8.0**	**12.0**

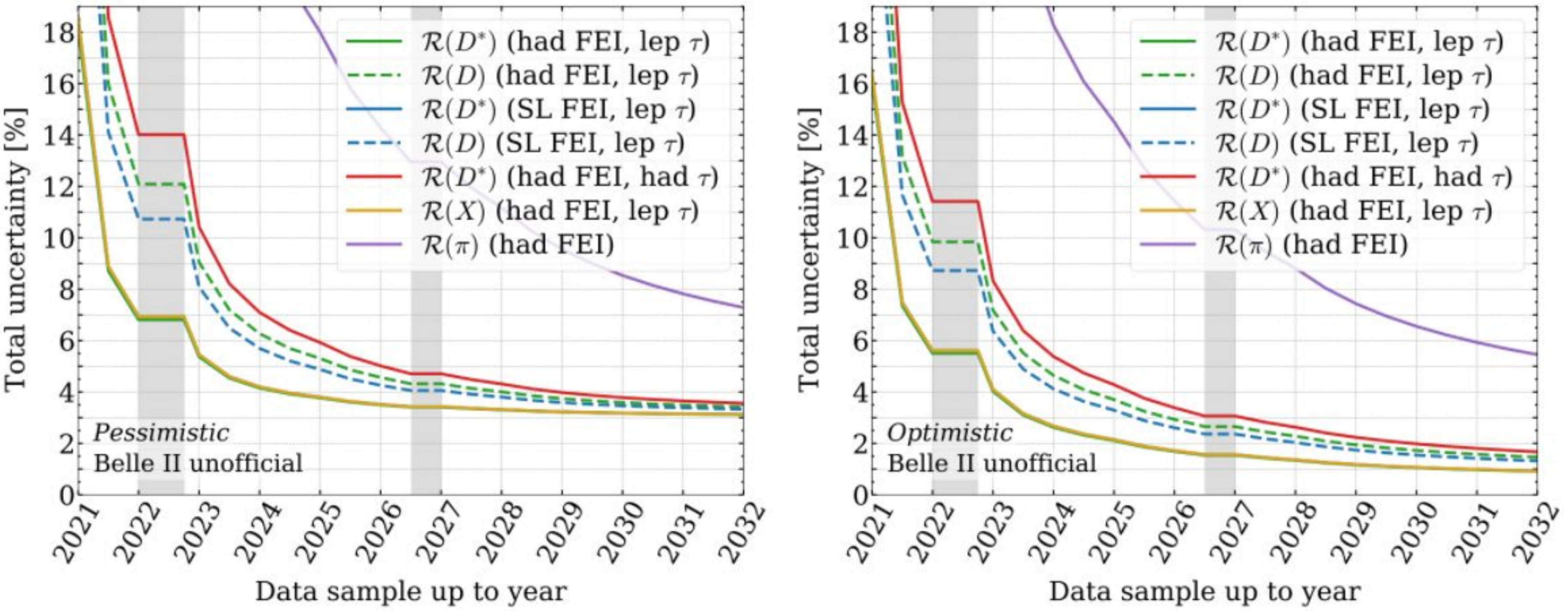
Significance

0.2σ and 1.1σ are observed in the semileptonic paper for R(D) and R(D*) respectively

Obs.	Current World Av./Data	Current SM Prediction	Significance	
$\mathcal{R}(D)$	0.340 ± 0.030	0.299 ± 0.003	1.2σ	3.1σ
$\mathcal{R}(D^*)$	0.295 ± 0.014	0.258 ± 0.005	2.5σ	
$P_\tau(D^*)$	$-0.38 \pm 0.51^{+0.21}_{-0.16}$	-0.501 ± 0.011	0.2σ	
$F_{L,\tau}(D^*)$	$0.60 \pm 0.08 \pm 0.04$	0.455 ± 0.006	1.6σ	
$\mathcal{R}(J/\psi)$	$0.71 \pm 0.17 \pm 0.18$	0.2582 ± 0.0038	1.8σ	
$\mathcal{R}(\pi)$	1.05 ± 0.51	0.641 ± 0.016	0.8σ	
$\mathcal{R}(D)$	$\mathbf{0.337 \pm 0.030}$	0.299 ± 0.003	1.3σ	$\mathbf{3.6\sigma}$
$\mathcal{R}(D^*)$	$\mathbf{0.298 \pm 0.014}$	0.258 ± 0.005	2.5σ	



More data expected from Belle II

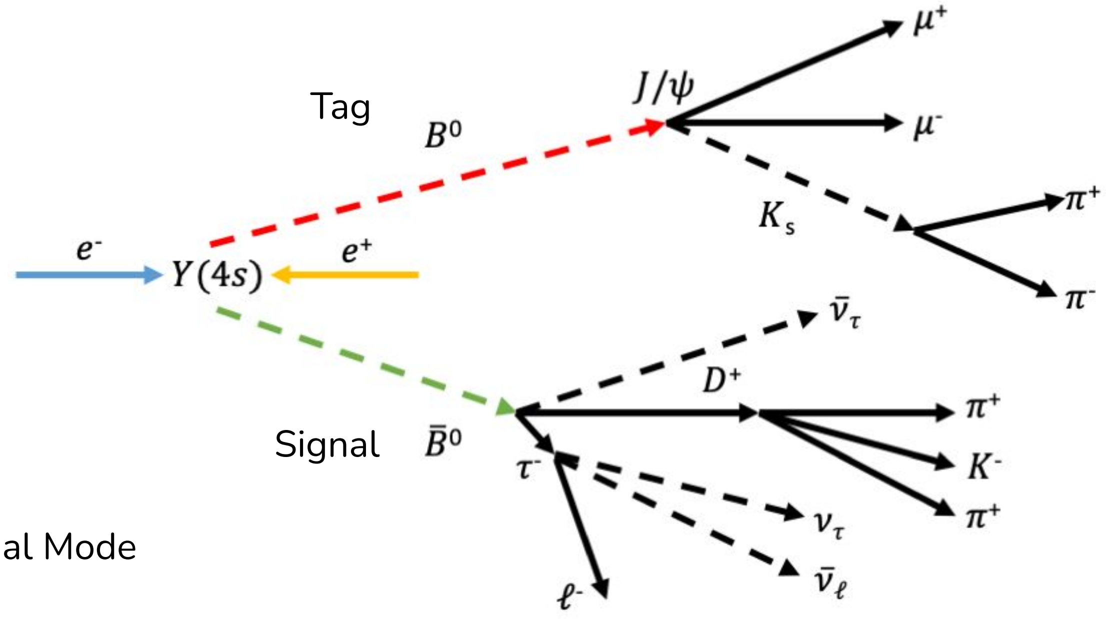


R(D) & R(D*)

My interest:

1. I will study the modes:
   $B \rightarrow D(*)\,\tau\,\nu$ and
   $B \rightarrow D(*)\,l\,\nu$

2. I will measure R(D) and R(D*) (ratio, but not absolute branching fractions)



$$R(D) = \frac{\mathcal{B}(\bar{B} \rightarrow D\tau^-\bar{\nu}_\tau)}{\mathcal{B}(\bar{B} \rightarrow D\ell^-\bar{\nu}_\ell)}$$

Signal Mode (numerator)

Normalization Mode (denominator)

$$R(D^*) = \frac{\mathcal{B}(\bar{B} \rightarrow D^*\tau^-\bar{\nu}_\tau)}{\mathcal{B}(\bar{B} \rightarrow D^*\ell^-\bar{\nu}_\ell)}$$