Signal MC Study

MC matching / isSignal / TopoAna

Belle II

Cheolhun Kim
Hanyang University
Korea

HYU HANYANG UNIVERSITY SEOUL, KOREA

Contents

- About this talk
- The **motivation** for the Signal MC study
  - What we reconstruct from the data
  - Possible cases of mis-reconstruction
  - Mathematical Expression of the reconstructed cases
  - Ratios and Distributions of the reconstructed cases
  - Reconstruction Efficiency and its Uncertainty
- **Design flags** for True / WC / SCF
  - Is it possible to design flags for True/Wrong-Comb./Self-Crossfeed with basf2 variable related to "isSignal"?
- **Design "isSignalEvent"**
  - with MC-matching info. / PDG-code
  - Machine learning **label**
- Further: Consider γ

Contents

About this talk

About this talk

- Due to my incomplete understanding
  - Classifying my own questions
  - My own conclusions
    - this can be wrong
- **Conceptual**
- If you don’t understand, that might be due to my misunderstanding...
  - **So, let’s discuss**

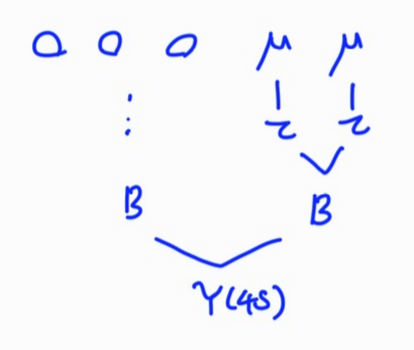
The motivation for the Signal MC study

The motivation for the Signal MC study

- Think about the final stage of the analysis first
  - Imagine there is no MC-matching study
    - Then,
      - we are not sure ...
      - we cannot distinguish ...

- Example
  - $\Upsilon(4S) \to \bar{B}^0 B^0$
    - $\bar{B}^0 \to generic$
    - $B^0 \to \tau^+ \tau^-$
      - $\tau^+ \to \mu^+ \nu_\mu \bar{\nu}_\tau, \tau^- \to \mu^- \bar{\nu}_\mu \nu_\tau$

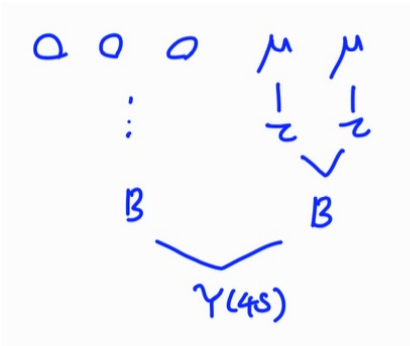
What we reconstruct from the data

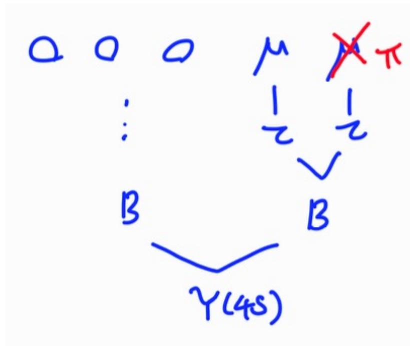


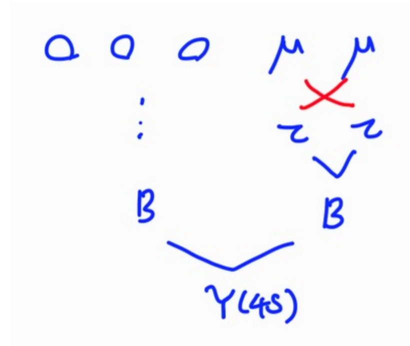
- Let us choose one of the final states, for example, $\mu\mu$
- $R \equiv R_{\mu\mu}$: Reconstruction of $B \rightarrow \tau\tau$ from $\mu\mu$ final state
- $P(R)$: Reconstruction Probability of the $B \rightarrow \tau\tau$ from $\mu\mu$ final state

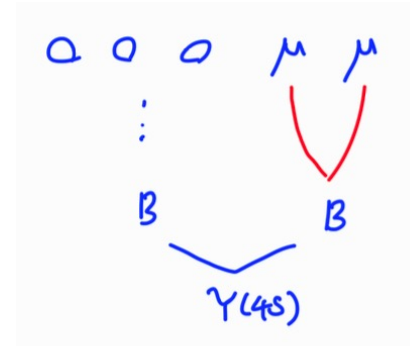
- For simplicity, let us assume
  - $P(R^C) = 0$, no reconstruction failure
    - (NaN case of isSignal)

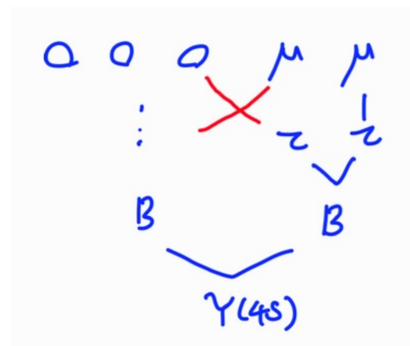
Possible cases of mis-reconstruction



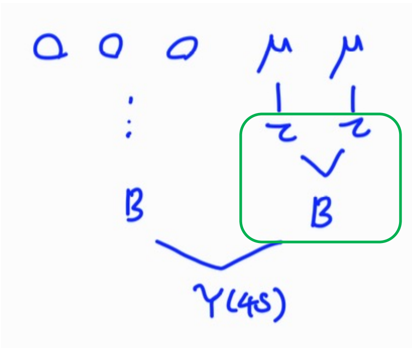


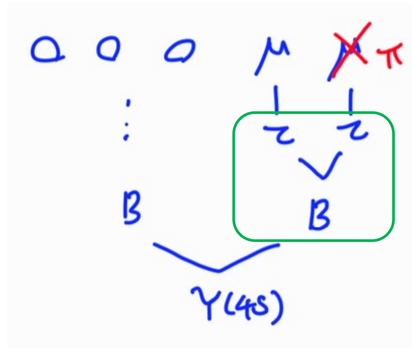


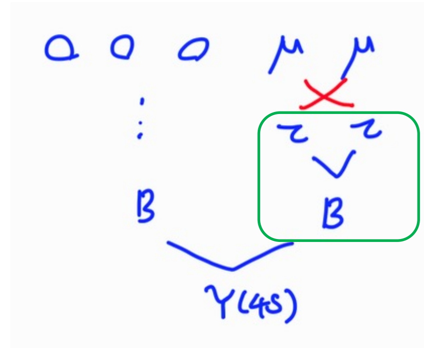


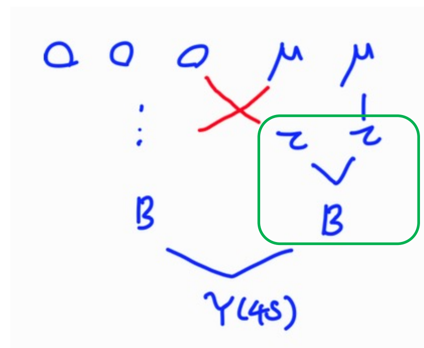


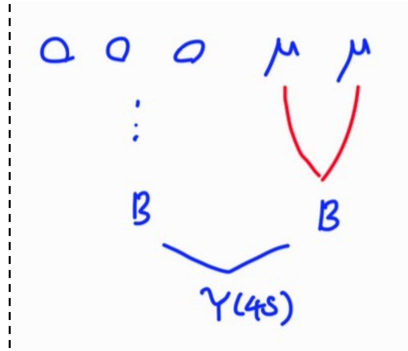
Possible cases of mis-reconstruction









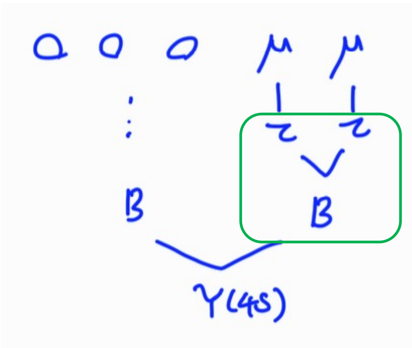


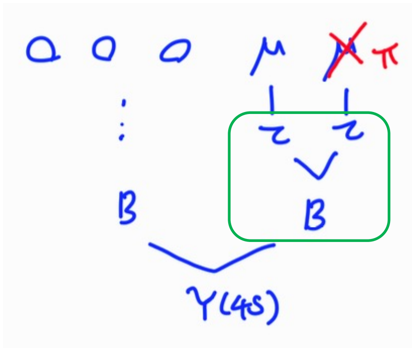
Signal O

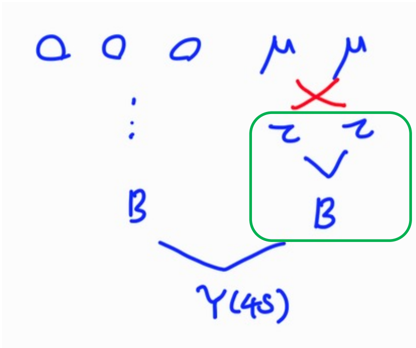
Signal O

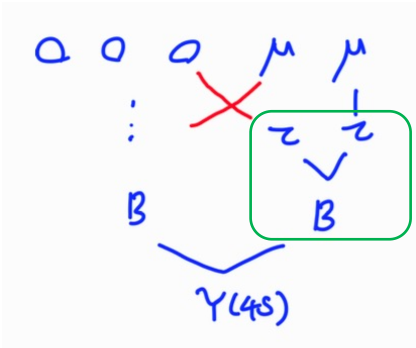
Signal O

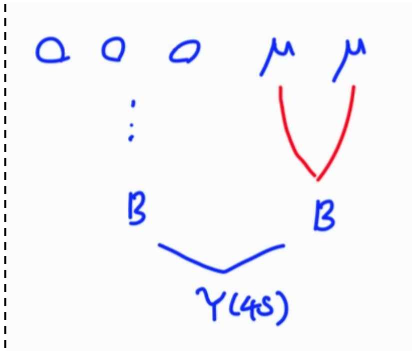
Signal X











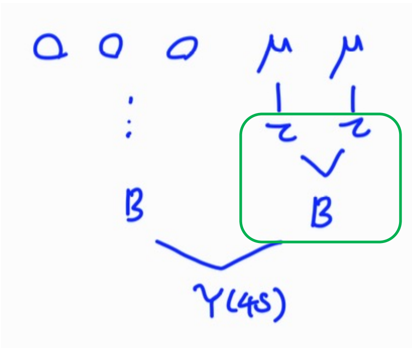
Signal O
Recon O
True

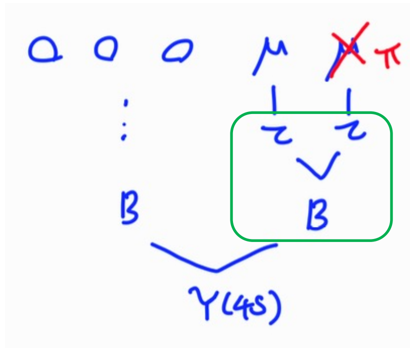
Signal O
Recon X
Wrong Comb.

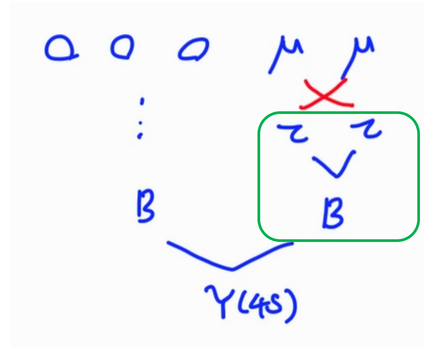
Signal O
Recon X
Self-Crossfeed

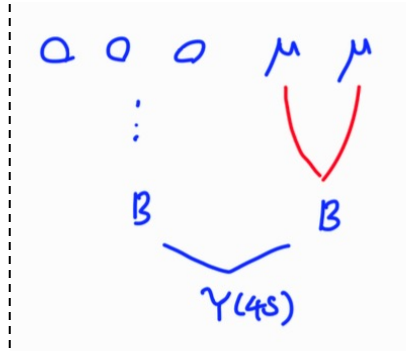
Signal X
Recon -
Background

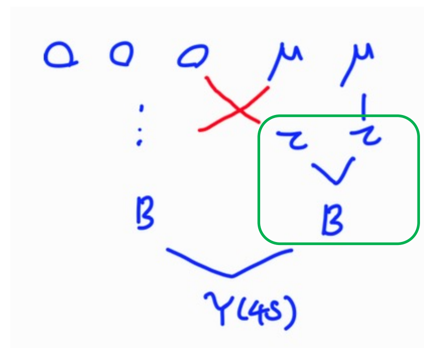
Possible cases of mis-reconstruction











Signal O
Recon O
True
$P(T|R)$

Signal O
Recon X
Wrong Comb.
$P(WC|R)$

Signal O
Recon X
Self-Crossfeed
$P(SCF|R)$

Signal X
Recon -
Background
$P(B|R)$

Mathematical Expression of the reconstructed cases

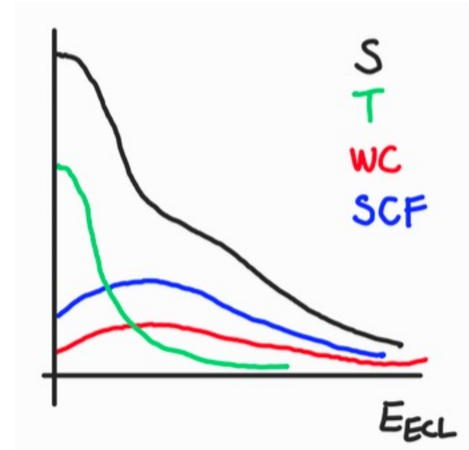
- $P(S|R) = P(T|R) + P(WC|R) + P(SCF|R)$
- $P(S|R) + P(B|R) = P(S \cup B|R) = P(R|R) = 1$
- where,
  - S: Signal
    - T: True
    - WC: Wrong-Combination
    - SCF: Self-Crossfeed
  - B: Background
  - $R \equiv R_{\mu\mu}$: Reconstruction of $B \to \tau\tau$ from $\mu\mu$ final state
- Assumption: No reconstruction failure (This also should be considered, but for simplicity)
  - $P(R^C) = 0$

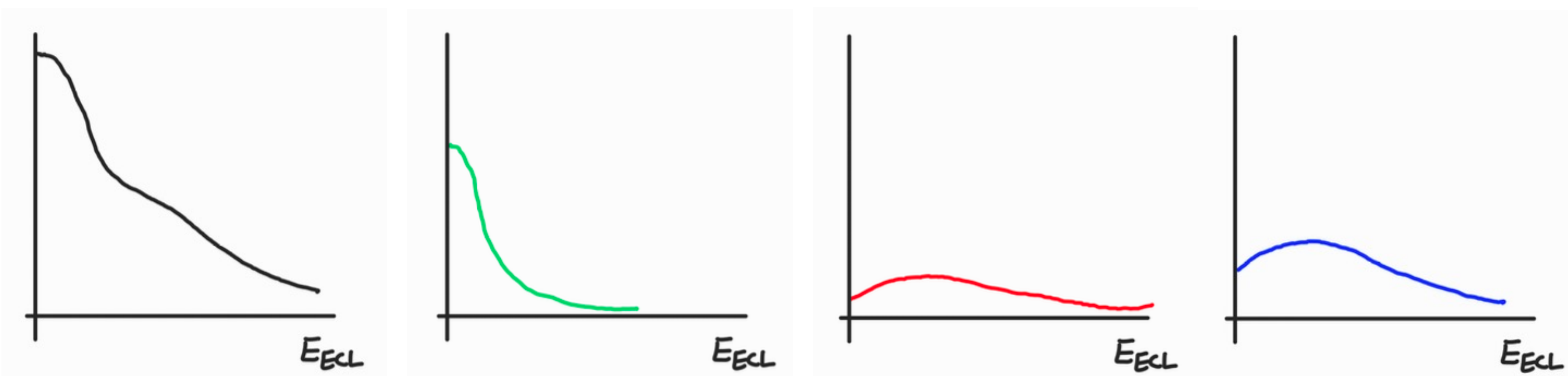
- $\boldsymbol{P(S|R)} = P(T|R) + P(WC|R) + P(SCF|R)$

- Some Analysis $P(S|R)$ can be approximated as,
  - $P(S|R) \approx P(T|R)$ with negligible WC and SC
  - $P(S|R) \approx P(T|R) + P(SCF|R)$ with negligible WC

- How can we know the ratios? ($r_T : r_{WC} : r_{SC}$)
  - $r_X = \frac{P(X|R)}{P(S|R)}$

- How can we know the distributions?

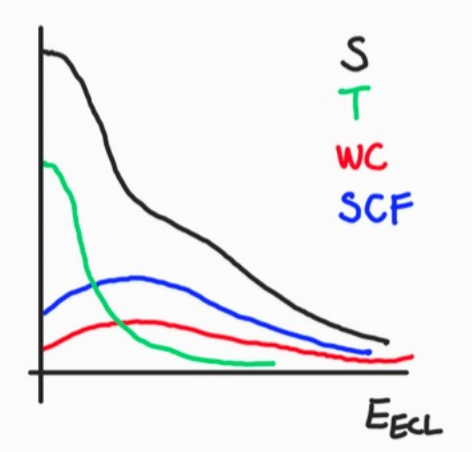




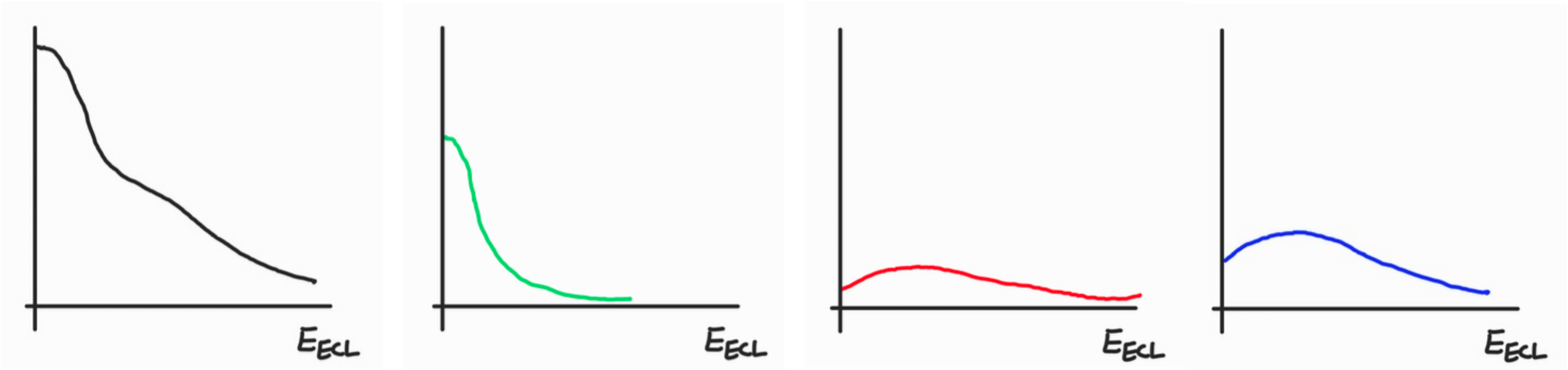
- $\boldsymbol{P(S|R)} = P(T|R) + P(WC|R) + P(SCF|R)$

- Some Analysis $P(S|R)$ can be approximated as,
  - $P(S|R) \approx P(T|R)$ with negligible WC and SC
  - $P(S|R) \approx P(T|R) + P(SCF|R)$ with negligible WC

- How can we know the ratios? ($r_T : r_{WC} : r_{SC}$)
  - $r_X = \frac{P(X|R)}{P(S|R)}$



- How can we know the distributions?



There is a way to estimate theses ratios and distributions!

Signal MC study

Reconstruction Efficiency and its Uncertainty

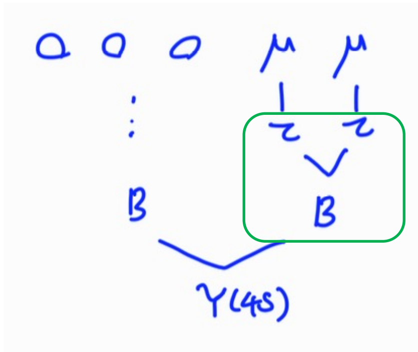
- For **example**,
  - all numbers here are arbitrary

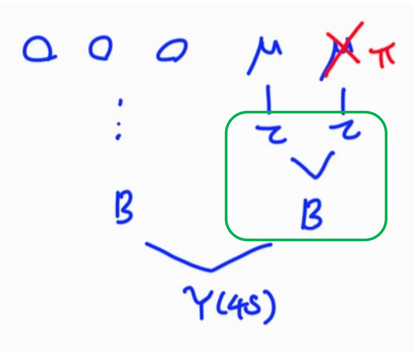
- Case 1) WC / SCF negligible
  - $\varepsilon(S = T) = 1 \pm 0.01\ \%$

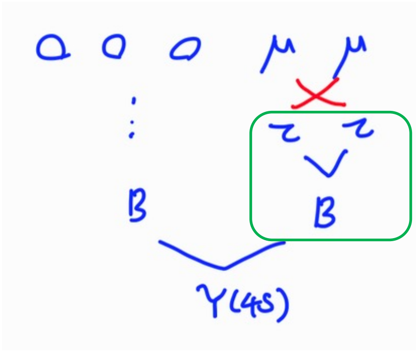
- Case 2) WC negligible, SCF significant
  - $\varepsilon(S = T + SCF) = 1 \pm 0.5\ \%$
    - SCF worsen the systematic uncertainty (Fit)
  - Fraction is the important factor which affect the uncertainty
    - $f(SCF) = \frac{N_{SCF}}{N_{sig}} = \frac{N_{SCF}}{N_{True}+N_{SCF}} = 23\ \%$
    - Thus, it is important to know the fraction and distribution.

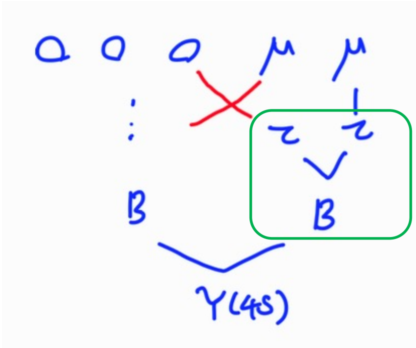
Design Flags
for True / WC / SCF

Design flags for True / WC / SCF









It might be possible as follows?

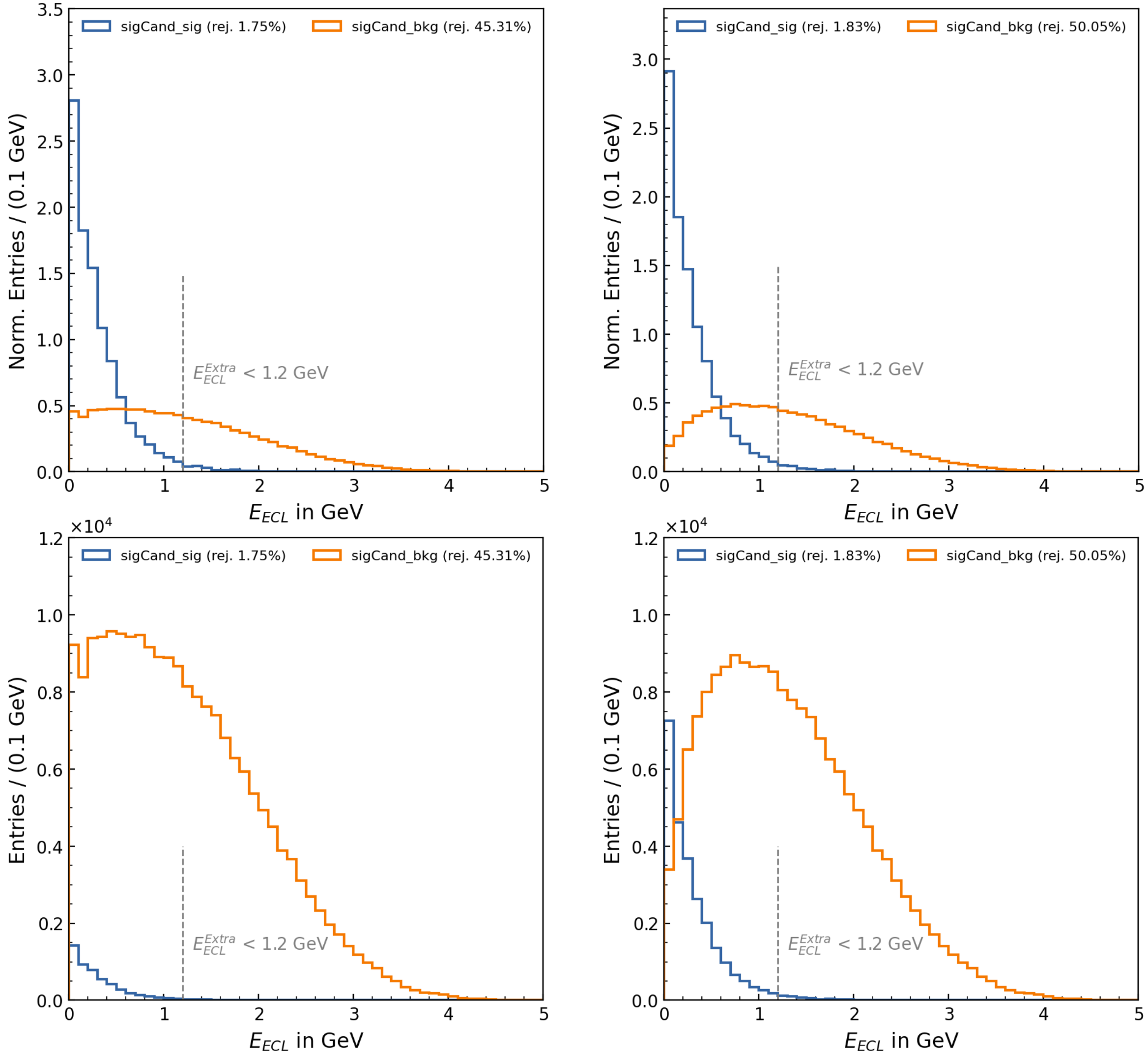
- WC and SCF with combinations of below flags?
  - misID
    - Particle Identification is not correct.
  - AddedWrongParticle
    - Check if the MC-matching is correct for each daughter. If not, the AddedWrongParticle flag is added.

※ [Monte Carlo matching in the Belle II software] https://doi.org/10.1051/epjconf/202125103021

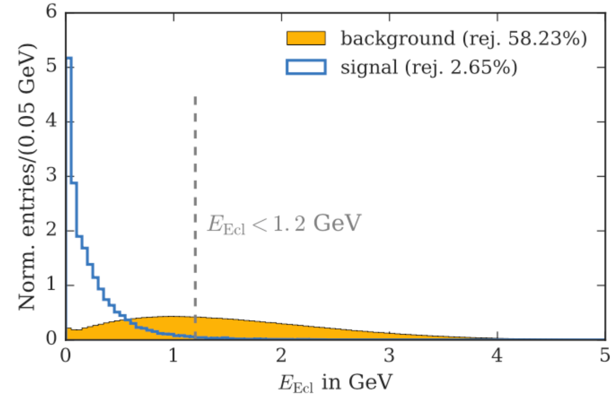
Design flags: isSignal Test

isSignalAccpetMissingNeutrino

isSignalAccpetMissingNeutrinoAndWrongFSPs



Belle



(a) Signal and background E_{ECL} distributions. Events with $E_{\mathrm{Ecl}} < 1.2$ GeV are selected.

Design flags: isSignal Test

- Pre-defined isSignal
  - isSignalAccpetMissingNeutrino
    - “mcErrors == 0” or “mcErrors == 8”
- User-defined isSignal
  - isSignalAccpetMissingNeutrinoAndWrongFSPs
    - “mcErrors == 0” or “mcErrors == 8” or “mcErrors == 128” or “mcErrors == 256”

```
vu.create_isSignal_alias("isSignalAcceptMissingNeutrinoAndWrongFSPs", [8, 128, 256])

evt_variables = [

    "isSignalAcceptMissingNeutrino",
    "isSignalAcceptWrongFSPs",
    "isSignalAcceptMissingNeutrinoAndWrongFSPs",
    "isContinuumEvent",
    "isNotContinuumEvent"

]
```

MC matching error codes

- calling **matchMCTruth(B+, path)** in one's steering file sets relations between **Particles** and **MCParticles**
  - recursively matches all daughter particles
  - bit-wise error flags indicate what went wrong in MC matching (variable **mcErrors**)

	c_Correct = 0	This Particle and all its daughters are perfectly reconstructed.
	c_MissFSR = 1	A Final State Radiation (FSR) photon is not reconstructed (based on MCParticle::c_IsFSRPhoton).
	c_MissingResonance = 2	The associated MCParticle decay contained additional non-final-state particles (e.g. a rho) that weren't reconstructed. This is probably O.K. in most cases.
	c_DecayInFlight = 4	A Particle was reconstructed from the secondary decay product of the actual particle. This means that a wrong hypothesis was used to reconstruct it, which e.g. for tracks might mean a pion hypothesis was used for a secondary electron.
✓	c_MissNeutrino = 8	A neutrino is missing (not reconstructed).
	c_MissGamma = 16	A photon (not FSR) is missing (not reconstructed).
	c_MissMassiveParticle = 32	A generated massive FSP is missing (not reconstructed).
	c_MissKlong = 64	A Klong is missing (not reconstructed).
✓	c_MisID = 128	One of the charged final state particles is mis-identified (wrong signed PDG code).
✓	c_AddedWrongParticle = 256	A non-FSP Particle has wrong PDG code, meaning one of the daughters (or their daughters) belongs to another Particle.
	c_InternalError = 512	No valid match was found. For tracks, it indicates that there is not a true track related to the reconstructed one. Might indicate fake or background track or cluster.
	c_MissPHOTOS = 1024	A photon created by PHOTOS was not reconstructed (based on MCParticle::c_IsPHOTOSPhoton).
	c_AddedRecoBremsPhoton = 2048	A photon added with the bremsstrahlung recovery tools (correctBrems or correctBremsBelle) has no MC particle assigned, or it doesn't belong to the decay chain of the corrected lepton mother.

Frank Meier (Duke University) Bremsstrahlung and MC matching 30.01.2023 11 / 17

※ [All you wanted to know about brems correction and MC matching and never dared asking, Belle II Physics Meeting]
https://indico.belle2.org/event/8674/#42-all-you-wanted-to-know-abou

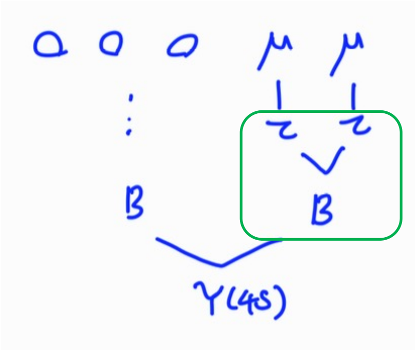
isSignal variables

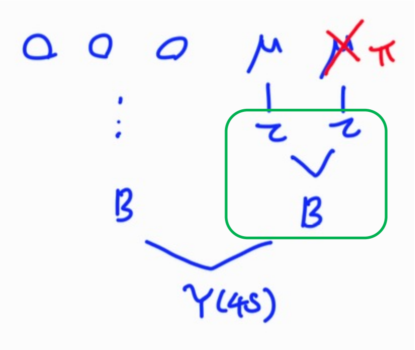
- convenience aliases for `mcErrors` variable
- `isSignal`: `mcErrors == 0`
- `isSignalAcceptMissingNeutrino`: `mcErrors == 0 or mcErrors == 8`
- `isSignalAcceptMissingGamma`: `mcErrors == 0 or mcErrors == 16`
- `isSignalAcceptBremsPhotons`: `mcErrors == 0 or mcErrors == 2048`
- `isSignalAcceptMissingMassive`: `mcErrors` ∈ [0, 32, 64, 96]
- `isSignalAcceptWrongFSPs`: `mcErrors` ∈ [0, 128, 256, 384]
- `isSignalAcceptMissing`: `mcErrors` ∈ [0, 8, 16, 24, 32, 40, 48, 56, 64, 72, 80, 88, 96, 104, 112, 120]
- `create_isSignal_alias(aliasName, flags)`
  - `create_isSignal_alias("isSignalAcceptMissingNeutrinoAndMissingGamma", [8, 16])`: `mcErrors` ∈ [0, 8, 16, 24]
- three possible return values: 0, 1, and NaN
  - isSignal variables can not be used as booleans
  - NaN if no MC partner found (running on data or `matchMCTruth` not run or fake tracks / cluster)

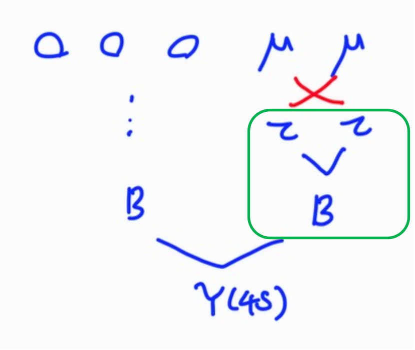
※ [All you wanted to know about brems correction and MC matching and never dared asking, Belle II Physics Meeting]
https://indico.belle2.org/event/8674/#42-all-you-wanted-to-know-abou

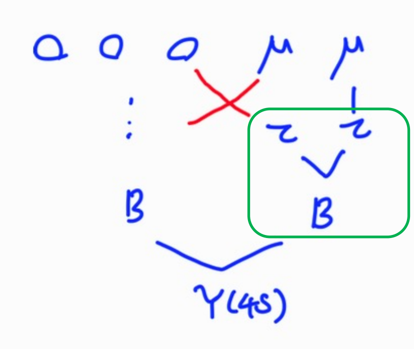
Design “isSignalEvent”

Design "isSignalEvent"







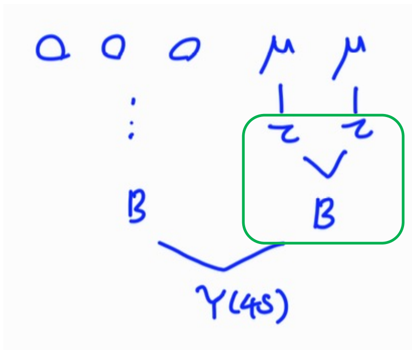


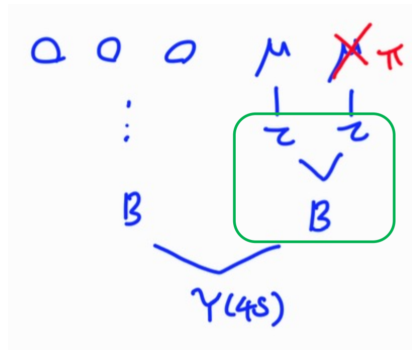
$$\boldsymbol{P(S|R)}$$

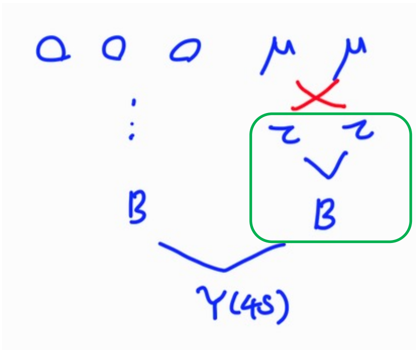
Make a flag as 1 (True), if this green box matches

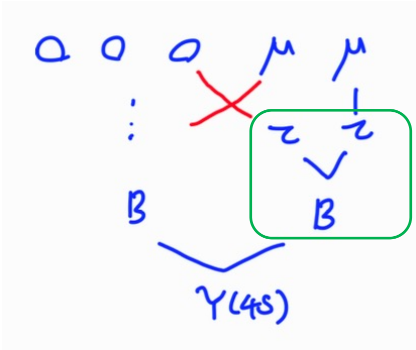
- Condition 0. Check whether the right PDG code
- Condition 1. Only sister τ (with the right PDG code)
- Condition 2. Mother B (with the right PDG code)

Design “isSignalEvent”









I think this flag can be the training flag.

$$P(S|R)$$

Make a flag as 1 (True), if this green box matches

- Condition 0. Check whether the right PDG code
- Condition 1. Only sister τ (with the right PDG code)
- Condition 2. Mother B (with the right PDG code)

- p15, line 210: "The signal is truth-matched [24]."
  - [24]: "***The truth-matching of the B+ candidate is as follows: the candidate should have a kaon daughter and 2 neutrinos***. Potential radiative photons are also accepted."

6. INPUT VARIABLES

In this section, variables that are used in the training of the main classifier are defined and shown. A sample of 8×10^5 events of each simulated category (signal, B^+B^-, $B^0\overline{B}^0$, $u\overline{u}$, $d\overline{d}$, $s\overline{s}$, $c\overline{c}$, $\tau^+\tau^-$) is reconstructed with the event selection described in the previous section apart from the requirement on the KaonPID (unless specified differently).

The signal is truth-matched [24]. Furthermore, additional weights are applied on the signal events to correct the phase space of the decay products as the MC was generated assuming on a three-body decay of the signal B. This procedure is described in detail in Appendix F The background samples are stacked with weights taking into account the cross-section of the process and the reconstruction efficiency. Both the signal and the stacked background histograms are normalized to unity and the signal-background separation power is estimated with the Jensen-Shannon distance [7], that gives 0 (1) for perfectly matching (diverging) histograms. Moreover, data taken during the Run 3123 from Experiment 8 is overlayed (randomly chosen from the list of 'good' runs).

All histograms are framed by colored boxes as some of the considered input variables are excluded from the training of the BDT:

※ ["Studies of $B^{\pm} \to K^{\pm}\nu\bar{\nu}$ decay using inclusive tagging method" by Filippo Dattola]
https://docs.belle2.org/record/2003

“isSignalEvent” Examples

```python
class IsSignalEvent(b2.Module):
    """
    define extraInfo isSignalEvent, that is set to true if, checking MC info:
    one of the B mesons has a son with the right PDG-ID, its mother is a B+ or B0 (depending of its charge)
    and its only sisters are a nu and an anti-nu)
    """
```

```python
class IsSignalEvent(b2.Module):
    """
    define extraInfo isSignalEvent, that is set to true if, checking MC info:
    one of the B mesons has a son with the right PDG-ID, its mother is a B+ or B0 (depending of its charge)
    and its only sisters are a tau+ and a tau-)
    """
```

※ [$B \to K\nu\nu$ Analysis Code by Giulio Dujany]
https://gitlab.desy.de/giulio.dujany/b2knn/-/blob/main/ntupliser/ntuple_variables.py
※ [$B \to K\tau\tau$ Analysis Code by Stefano Moneta]
https://gitlab.desy.de/stefano.moneta/btokst_tau_tau/-/blob/main/ntupliser/ntuple_variables.py

“isSignalEvent” Example ($B^+ \to K^+\nu\bar{\nu}$)

```python
class IsSignalEvent(b2.Module):
    """
    define extraInfo isSignalEvent, that is set to true if, checking MC info:
    one of the B mesons has a son with the right PDG-ID, its mother is a B+ or B0 (depending of its charge)
    and its only sisters are a nu and an anti-nu)
    """

    def __init__(self, listname, varname="isSignalEvent"):
        super().__init__()
        self.particleList = r.Belle2.PyStoreObj(listname)
        self.varname = varname

    def event(self):
        mcParticles = r.Belle2.PyStoreArray("MCParticles")
        KaonsPDGids = [
            321,
            310,
            323,
            313,
            130,
            311,
            10311,
            10321,
            100311,
            100321,
            10313,
            10323,
            20313,
            20323,
            100313,
            100323,
            30313,
            30323,
            315,
            325,
            10315,
            10325,
            20315,
            20325,
            317,
            327,
            319,
            329,
        ]
        Knunu_generic = []
        KJpsi_generic = []
        for a in KaonsPDGids:
            for b in [12, 14, 16]:  # here are the neutrinos!!!
                Knunu_generic.append(sorted([a, b, -b]))
                Knunu_generic.append(sorted([-a, b, -b]))
```

```python
    if sorted(
        list(
            filter(
                lambda x: x != 22,
                [abs(h.getPDG()) for h in j.getDaughters()],
            )
        )
    ) == [211, 310]:
        for k in j.getDaughters():
            if k.getPDG() == 310:
                if sorted(
                    list(
                        filter(
                            lambda x: x != 22,
                            [
                                m.getPDG()
                                for m in k.getDaughters()
                            ],
                        )
                    )
                ) == [-211, 211]:
                    particle.setExtraInfo(self.varname, 4)
    if sorted(
        list(
            filter(
                lambda x: x != 22,
                [abs(h.getPDG()) for h in j.getDaughters()],
            )
        )
    ) == [211, 311]:
        for k in j.getDaughters():
            if (
                k.getPDG() == 311
                and k.getDaughter(0).getPDG() == 310
            ):
                if sorted(
                    list(
                        filter(
                            lambda x: x != 22,
                            [
                                m.getPDG()
                                for m in k.getDaughter(
                                    0
                                ).getDaughters()
                            ],
                        )
                    )
                ) == [-211, 211]:
                    particle.setExtraInfo(self.varname, 4)

elif abs(j.getPDG()) == 313:
    if sorted(
        list(
            filter(
                lambda x: x != 22,
                [abs(h.getPDG()) for h in j.getDaughters()],
            )
        )
    ) == [211, 321]:
        particle.setExtraInfo(self.varname, 5)
```

※ [$B \to K\nu\nu$ Analysis Code by Giulio Dujany]
https://gitlab.desy.de/giulio.dujany/b2knn/-/blob/main/ntupliser/ntuple_variables.py

"isSignalEvent" Example ($B^0 \rightarrow K^{*0}\tau^+\tau^-$)

```python
class IsSignalEvent(b2.Module):
    """
    define extraInfo isSignalEvent, that is set to true if, checking MC info:
    one of the B mesons has a son with the right PDG-ID, its mother is a B+ or B0 (depending of its charge)
    and its only sisters are a tau+ and a tau-)
    """

    def __init__(self, listname, varname="isSignalEvent"):
        super().__init__()
        self.particleList = ROOT.Belle2.PyStoreObj(listname)
        self.varname = varname

    def event(self):
        mcParticles = ROOT.Belle2.PyStoreArray("MCParticles")
        KaonsPDGids = [323, 313]
        Ktautau_generic = []
        for KaonPDGid in KaonsPDGids:
            Ktautau_generic.append(sorted([KaonPDGid,15,-15]))
            Ktautau_generic.append(sorted([-KaonPDGid,15,-15]))
        for particle in self.particleList:
            if particle.hasExtraInfo(self.varname):
                continue
            particle.addExtraInfo(self.varname, 0)
            mcParticle = False
            for Upscand in mcParticles:
                if Upscand.getPDG() == 300553:
                    mcParticle = Upscand
                    break
            if not mcParticle:
                continue
            B_sons_PDGids = []
            for i in mcParticle.getDaughters():
                if abs(i.getPDG()) != 521 and abs(i.getPDG()) != 511 :
                    continue
                B_sons_PDGids.append(sorted(list(filter(lambda x: x!= 22,[j.getPDG() for j in i.getDaughters()]))))
            for B_son_PDGids in B_sons_PDGids :
                if B_son_PDGids in Ktautau_generic :
                    particle.setExtraInfo(self.varname, -1)
                else:
                    continue
            for i in mcParticle.getDaughters(): #loop on Bs
                sons_PDGids = sorted(list(filter(lambda x: x!= 22,[j.getPDG() for j in i.getDaughters()])))
                if sons_PDGids in Ktautau_generic: # then we know its K tau tau
                    for j in i.getDaughters():
                        if abs(j.getPDG()) == 313: # K*0 -> K- pi+
                            if sorted(list(filter(lambda x: x!= 22,[abs(h.getPDG()) for h in j.getDaughters()]))) == [211,321]:
                                particle.setExtraInfo(self.varname, 1)
```

※ [$B \rightarrow K\tau\tau$ Analysis Code by Stefano Moneta]
https://gitlab.desy.de/stefano.moneta/btokst_tau_tau/-/blob/main/ntupliser/ntuple_variables.py

Inclusive Use Case ($B^+ \rightarrow K^+\nu\bar{\nu}$)

```python
class IsB2Xnunu(b2.Module):
    """
    define extraInfo isB2Xnunu, that is set to true if, checking MC info:
    X has the right PDG-ID, its mother is a B+ or B0 (depending of its charge)
    and its only sisters are a nu_e and an anti-nu_e
    """

    def __init__(self, listname, varname="isB2Xnunu"):
        super().__init__()
        self.particleList = r.Belle2.PyStoreObj(listname)
        self.varname = varname

    def event(self):
        mcParticles = r.Belle2.PyStoreArray("MCParticles")
        for particle in self.particleList:
            if particle.hasExtraInfo(self.varname):
                continue
            particle.addExtraInfo(self.varname, 0)
            mcParticle = particle.getRelatedTo("MCParticles")
            if not mcParticle:
                continue
            mcMother = mcParticle.getMother()
            if not mcMother:
                continue
            B_PDG = 511 if particle.getCharge() == 0 else 521
            if abs(mcMother.getPDG()) != B_PDG:
                continue
            if mcMother.getCharge() != particle.getCharge():
                continue
            daughters_PDGs = sorted([i.getPDG() for i in mcMother.getDaughters()])
            if daughters_PDGs != sorted([particle.getPDGCode(), -12, 12]):
                continue
            particle.setExtraInfo(self.varname, 1)
```

※ [$B \rightarrow K\nu\nu$ Analysis Code by Giulio Dujany]
https://gitlab.desy.de/giulio.dujany/b2knn/-/blob/main/ntupliser/ntuple_variables.py

Inclusive Use Case ($B^0 \to K^{*0}\tau^+\tau^-$)

```python
class IsB2Xtautau(b2.Module):
    """
    define extraInfo isB2Xtautau, that is set to true if, checking MC info:
    X has the right PDG-ID, its mother is a B+ or B0 (depending of its charge)
    and its only sisters are a tau+ and a tau-
    """

    def __init__(self, listname, varname="isB2Xtautau"):
        super().__init__()
        self.particleList = ROOT.Belle2.PyStoreObj(listname)
        self.varname = varname

    def event(self):
        mcParticles = ROOT.Belle2.PyStoreArray("MCParticles")
        for particle in self.particleList:
            if particle.hasExtraInfo(self.varname):
                continue
            particle.addExtraInfo(self.varname, 0)
            mcParticle = particle.getRelatedTo("MCParticles")
            if not mcParticle:
                continue
            mcMother = mcParticle.getMother()
            if not mcMother:
                continue
            B_PDG = 511 if particle.getCharge() == 0 else 521
            if abs(mcMother.getPDG()) != B_PDG:
                continue
            if mcMother.getCharge() != particle.getCharge():
                continue
            daughters_PDGs = sorted([i.getPDG() for i in mcMother.getDaughters()])
            if daughters_PDGs != sorted([particle.getPDGCode(), -15, 15]):
                continue
            particle.setExtraInfo(self.varname, 1)
```

※ [$B \to K\tau\tau$ Analysis Code by Stefano Moneta]
https://gitlab.desy.de/stefano.moneta/btokst_tau_tau/-/blob/main/ntupliser/ntuple_variables.py

Further: Consider γ

Consider γ

- p15, line 210: "The signal is truth-matched [24]."
  - [24]: "The truth-matching of the B+ candidate is as follows: the candidate should have a kaon daughter and 2 neutrinos. ***Potential radiative photons are also accepted***."

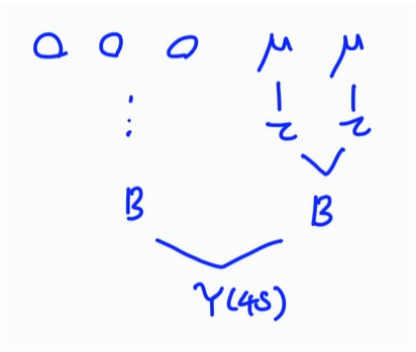
6. INPUT VARIABLES

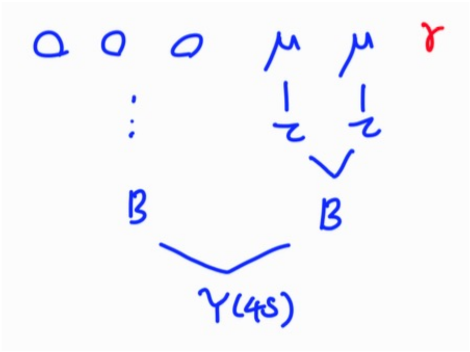
In this section, variables that are used in the training of the main classifier are defined and shown. A sample of 8×10^5 events of each simulated category (signal, B^+B^-, $B^0\overline{B}^0$, $u\overline{u}$, $d\overline{d}$, $s\overline{s}$, $c\overline{c}$, $\tau^+\tau^-$) is reconstructed with the event selection described in the previous section apart from the requirement on the KaonPID (unless specified differently).

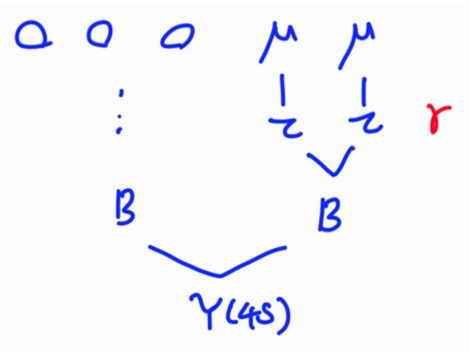
The signal is truth-matched [24]. Furthermore, additional weights are applied on the signal events to correct the phase space of the decay products as the MC was generated assuming on a three-body decay of the signal B. This procedure is described in detail in Appendix F The background samples are stacked with weights taking into account the cross-section of the process and the reconstruction efficiency. Both the signal and the stacked background histograms are normalized to unity and the signal-background separation power is estimated with the Jensen-Shannon distance [7], that gives 0 (1) for perfectly matching (diverging) histograms. Moreover, data taken during the Run 3123 from Experiment 8 is overlayed (randomly chosen from the list of 'good' runs).

All histograms are framed by colored boxes as some of the considered input variables are excluded from the training of the BDT:

※ ["Studies of $B^{\pm} \to K^{\pm}\nu\bar{\nu}$ decay using inclusive tagging method" by Filippo Dattola]
https://docs.belle2.org/record/2003


μ
μ
B
B
Υ(4S)


μ
μ
γ
B
B
Υ(4S)


μ
μ
γ
B
B
Υ(4S)

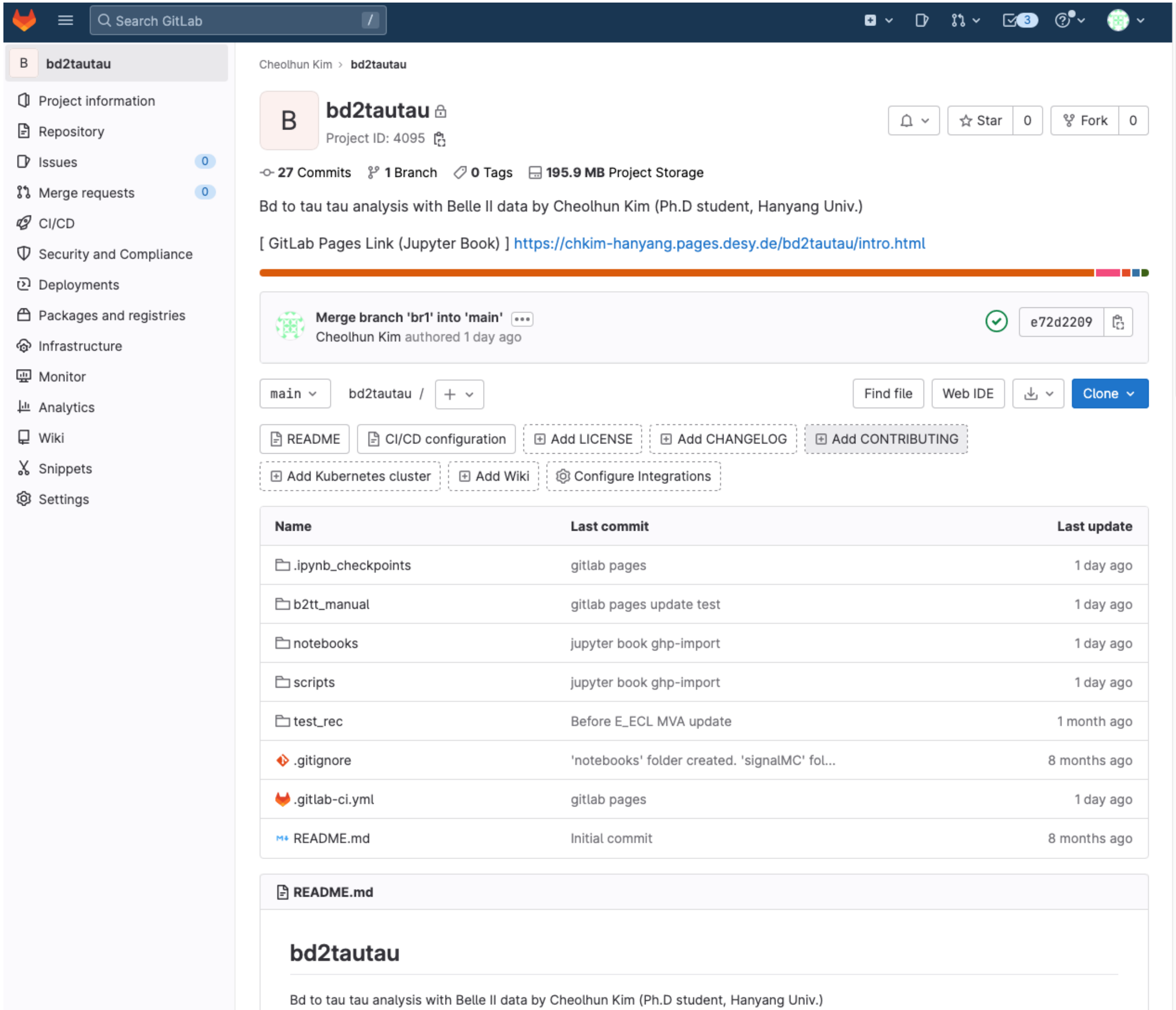
Others: Code Version Control and Documentation

Code version control / Documentation

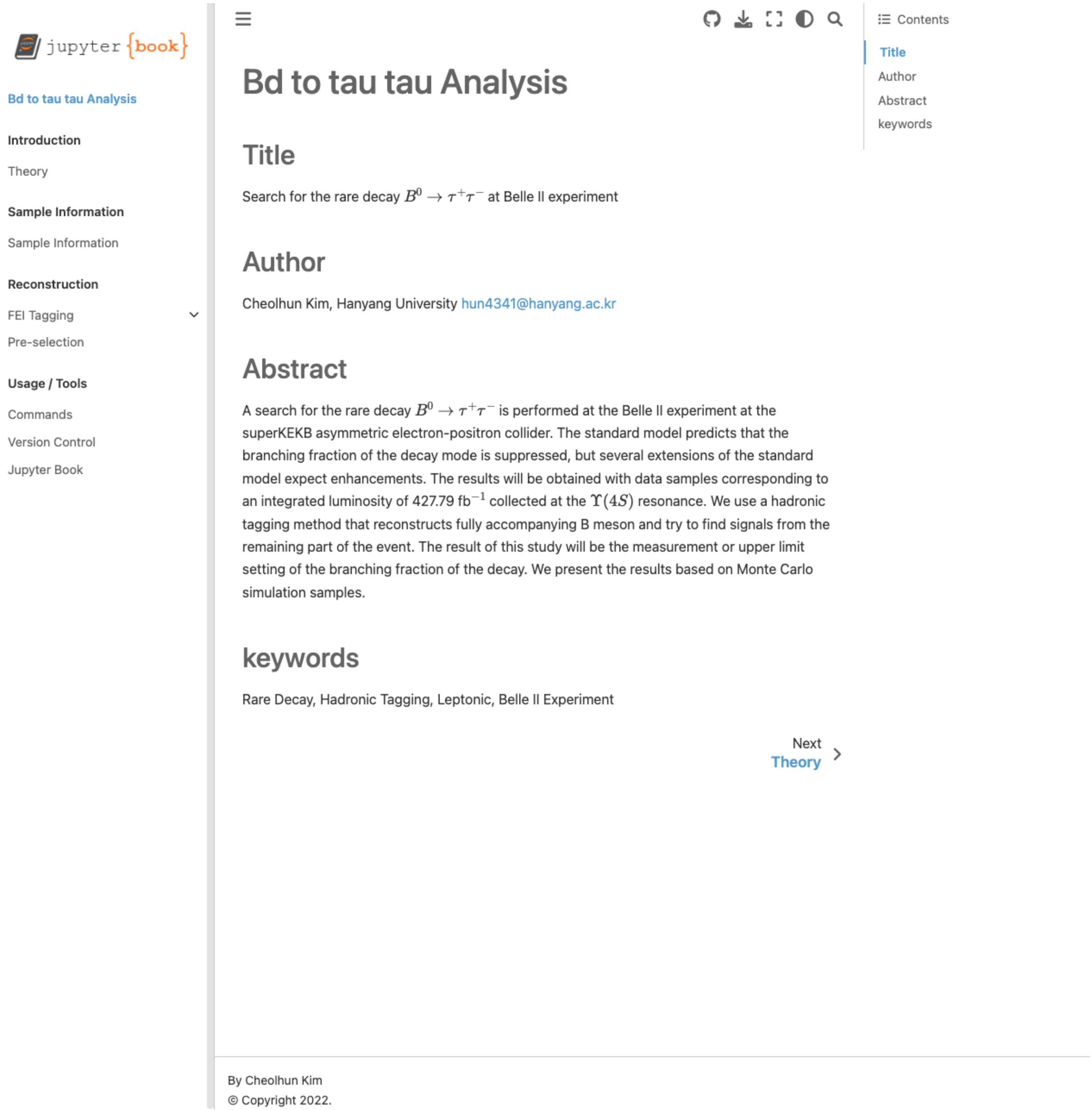
- Gitlab
  - Belle II uses “Gitlab” as git repository
  - My repository
    - https://gitlab.desy.de/chkim-hanyang/bd2tautau

- Jupyter book
  - Python-based documentation package
    - Markdown format
    - Comparability with Sphinx
    - HTML output
  - My documentation online page with “Gitlab Pages”
    - https://chkim-hanyang.pages.desy.de/bd2tautau/intro.html
  - This will be a supplementary of my Belle II note

My Gitlab Repository



My jupyter book Online



Backup

Backup

Backup Contents

- Tools for the MC study
- References

Tools for the MC study

MC matching

- MC matching?
  - Particle Candidates: Combinations of ***reconstructed*** objects
  - ***Simulated*** Particles: From the ***event generators***

- Purpose of MC matching?
  - To study **Detector effects**, **analysis backgrounds**, and **efficiencies**
  - To consider the **reconstruction effect** and **backgrounds**

- MC matching of tracks / MC matching of clusters
  - Track-based and Cluster-based particles (final-state particles) candidates

- "MC matching algorithm for Composite particles"
  - Goal: To establish the relation between the **reconstructed composite particle** and the **corresponding MC-particle**

- Evaluation algorithm for the "MC matching algorithm for Composite particles"
  - Categorizing failure cases: Several Error flags
  - **"Each analyst can CHOOSE to accept or not for his or her own analysis use case"**
  - Two major parts of the algorithm
    - Processes with **existing particles**
    - Processes with **missing particles**

※ [Monte Carlo matching in the Belle II software] https://doi.org/10.1051/epjconf/202125103021

MC matching algorithm for Composite particles

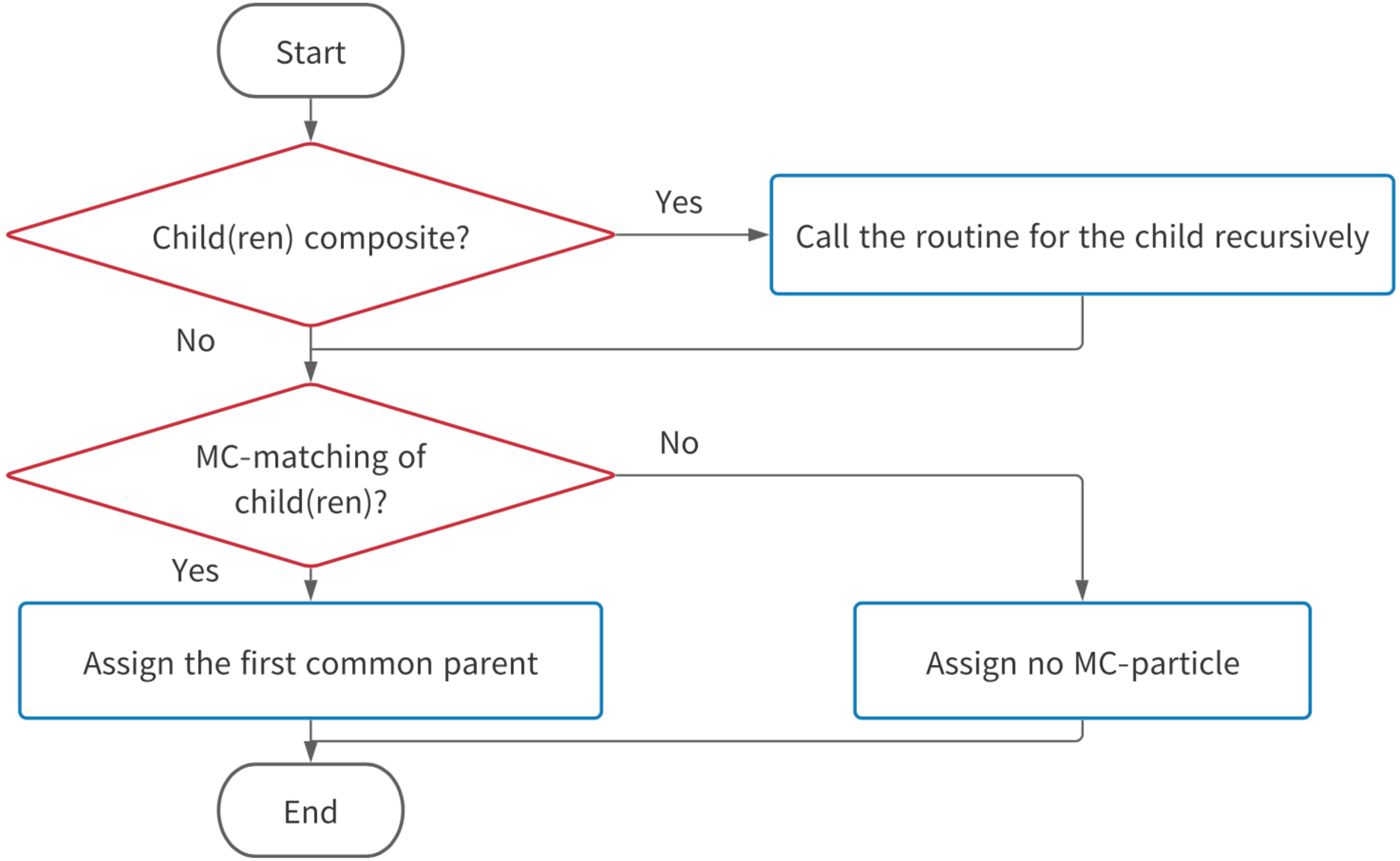


Figure 1. The algorithm flowchart of the MC-matching for a composite particle.

※ [Monte Carlo matching in the Belle II software] https://doi.org/10.1051/epjconf/202125103021

MC matching: The Evaluation Algorithm - Existing Particles

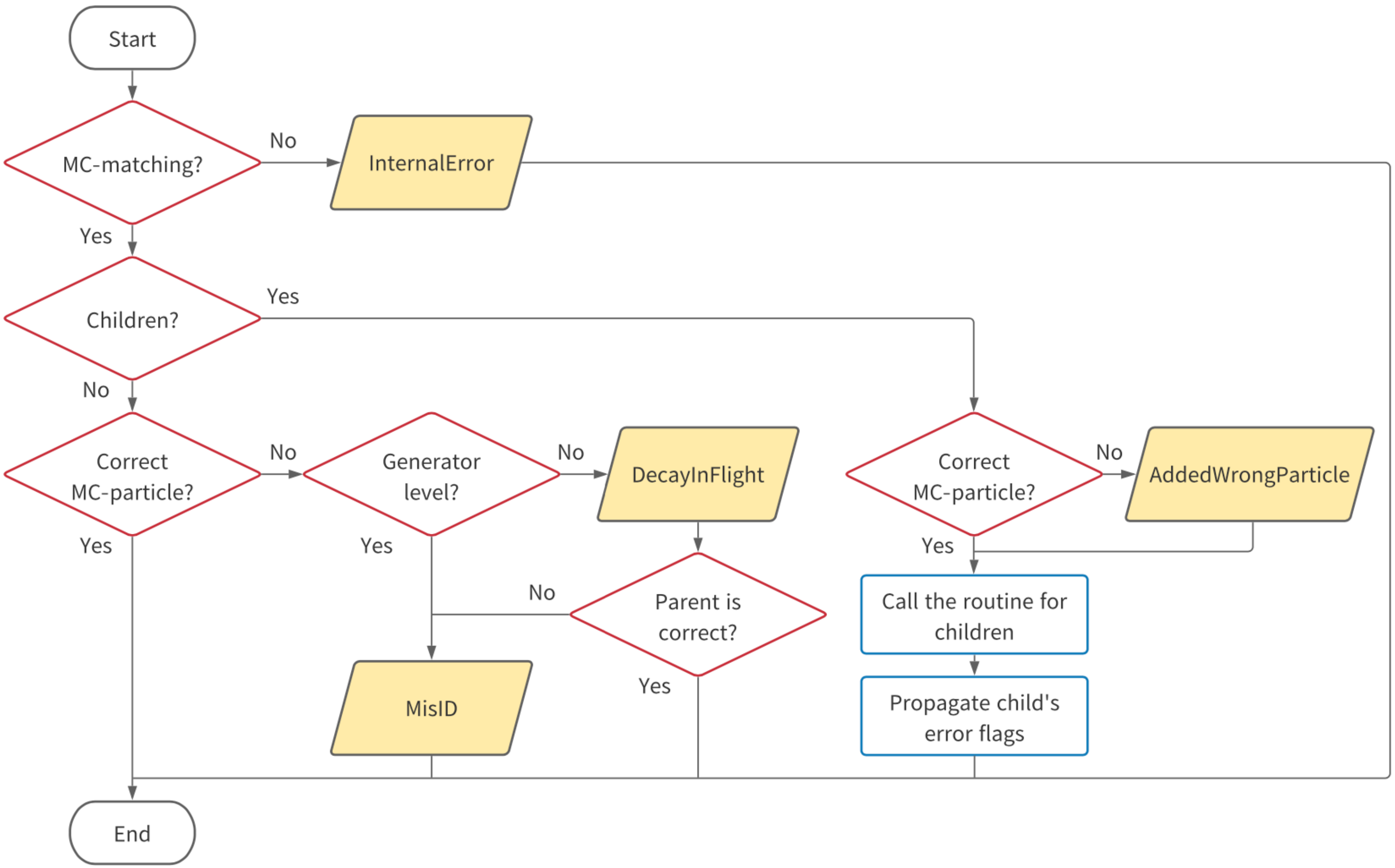


Figure 2. The algorithm flowchart of the evaluation for existing particles

※ [Monte Carlo matching in the Belle II software] https://doi.org/10.1051/epjconf/202125103021

MC matching: The Evaluation Algorithm - Missing Particles

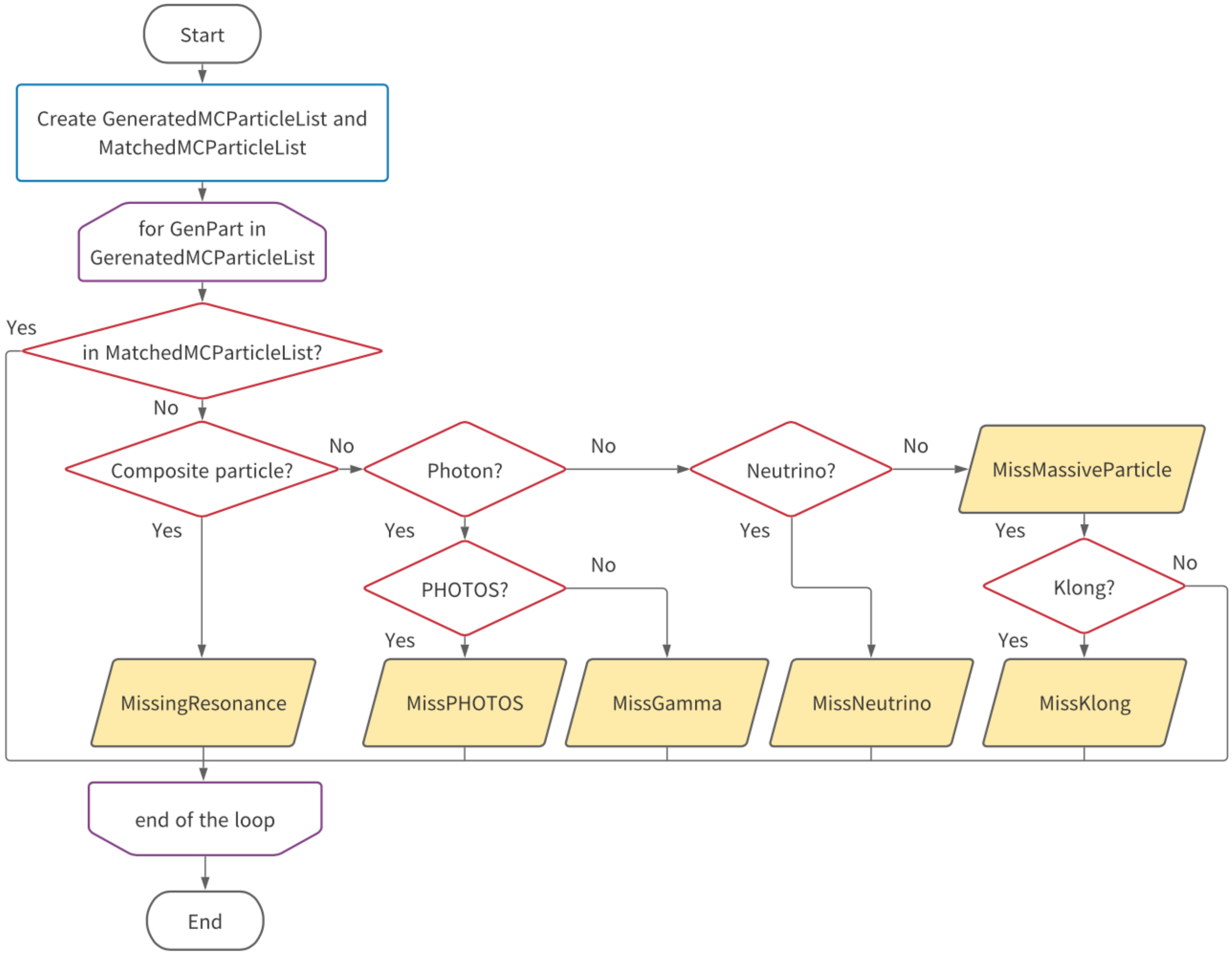


Figure 3. The algorithm flowchart of the evaluation for missing particles

※ [Monte Carlo matching in the Belle II software] https://doi.org/10.1051/epjconf/202125103021

※Chanho Kim will cover this part in detail.

MC matching

- MC matching useful to
  - optimize selection requirements, *e.g.* label training data for MVAs
  - calculate signal efficiencies
  - study background sources
- two steps
  1. relate mdst dataobjects (`Track`, `ECLCluster`, `KLMCluster`) with `MCParticle` (with weights)
  2. relate reconstructed particles with MC particles
- interpretation of MC matches necessary

※ [All you wanted to know about brems correction and MC matching and never dared asking, Belle II Physics Meeting]
https://indico.belle2.org/event/8674/#42-all-you-wanted-to-know-abou

MC matching error codes

- calling **matchMCTruth(B+, path)** in one's steering file sets relations between **Particles** and **MCParticles**
  - recursively matches all daughter particles
  - bit-wise error flags indicate what went wrong in MC matching (variable **mcErrors**)

c_Correct = 0	This Particle and all its daughters are perfectly reconstructed.
c_MissFSR = 1	A Final State Radiation (FSR) photon is not reconstructed (based on MCParticle::c_IsFSRPhoton).
c_MissingResonance = 2	The associated MCParticle decay contained additional non-final-state particles (e.g. a rho) that weren't reconstructed. This is probably O.K. in most cases.
c_DecayInFlight = 4	A Particle was reconstructed from the secondary decay product of the actual particle. This means that a wrong hypothesis was used to reconstruct it, which e.g. for tracks might mean a pion hypothesis was used for a secondary electron.
c_MissNeutrino = 8	A neutrino is missing (not reconstructed).
c_MissGamma = 16	A photon (not FSR) is missing (not reconstructed).
c_MissMassiveParticle = 32	A generated massive FSP is missing (not reconstructed).
c_MissKlong = 64	A Klong is missing (not reconstructed).
c_MisID = 128	One of the charged final state particles is mis-identified (wrong signed PDG code).
c_AddedWrongParticle = 256	A non-FSP Particle has wrong PDG code, meaning one of the daughters (or their daughters) belongs to another Particle.
c_InternalError = 512	No valid match was found. For tracks, it indicates that there is not a true track related to the reconstructed one. Might indicate fake or background track or cluster.
c_MissPHOTOS = 1024	A photon created by PHOTOS was not reconstructed (based on MCParticle::c_IsPHOTOSPhoton).
c_AddedRecoBremsPhoton = 2048	A photon added with the bremsstrahlung recovery tools (correctBrems or correctBremsBelle) has no MC particle assigned, or it doesn't belong to the decay chain of the corrected lepton mother.

Frank Meier (Duke University) Bremsstrahlung and MC matching 30.01.2023 11 / 17

※ [All you wanted to know about brems correction and MC matching and never dared asking, Belle II Physics Meeting]
https://indico.belle2.org/event/8674/#42-all-you-wanted-to-know-abou

MC matching: IsSignal

isSignal variables

- convenience aliases for `mcErrors` variable
- `isSignal`: `mcErrors == 0`
- `isSignalAcceptMissingNeutrino`: `mcErrors == 0 or mcErrors == 8`
- `isSignalAcceptMissingGamma`: `mcErrors == 0 or mcErrors == 16`
- `isSignalAcceptBremsPhotons`: `mcErrors == 0 or mcErrors == 2048`
- `isSignalAcceptMissingMassive`: `mcErrors` ∈ [0, 32, 64, 96]
- `isSignalAcceptWrongFSPs`: `mcErrors` ∈ [0, 128, 256, 384]
- `isSignalAcceptMissing`: `mcErrors` ∈ [0, 8, 16, 24, 32, 40, 48, 56, 64, 72, 80, 88, 96, 104, 112, 120]
- `create_isSignal_alias(aliasName, flags)`
  - `create_isSignal_alias("isSignalAcceptMissingNeutrinoAndMissingGamma", [8, 16])`: `mcErrors` ∈ [0, 8, 16, 24]
- three possible return values: 0, 1, and NaN
  - isSignal variables can not be used as booleans
  - NaN if no MC partner found (running on data or `matchMCTruth` not run or fake tracks / cluster)

Frank Meier (Duke University) Bremsstrahlung and MC matching 30.01.2023 12 / 17

※ [All you wanted to know about brems correction and MC matching and never dared asking, Belle II Physics Meeting]
https://indico.belle2.org/event/8674/#42-all-you-wanted-to-know-abou

Decay strings

- MC matching interpretation configured via decay arrow, markers, and keywords
- four different allowed arrow types
  - `->` decays via intermediate resonances and / or with radiative photons are counted as signal
  - `=norad=>` if in actual decay photon was radiated it is **not** counted as signal; intermediate resonances do not have to be specified
  - `=direct=>` if actual decay proceeds via intermediate resonance it is **not** counted as signal; decays with radiative photons are counted as signal
  - `=exact=>` intermediate resonances and radiative photons have to be specified explicitly, otherwise decay is not counted as signal
- different arrows allowed in same decay string `D*+ -> [D0 =direct=> K- pi+ pi0] pi+`
- example:
  - generated decay: $B^+ \to J/\psi\, K^+$ with $J/\psi \to e^+e^-$ and final state photon
  - `B+:default -> K+ e+ e-` $\Rightarrow$ `isSignal == 1`
  - `B+:noFSR =norad=> K+ e+ e-` $\Rightarrow$ `isSignal == 0` and `mcErrors == 1`
  - `B+:nores =direct=> K+ e+ e-` $\Rightarrow$ `isSignal == 0` and `mcErrors == 2`
  - `B+:exact =exact=> K+ e+ e-` $\Rightarrow$ `isSignal == 0` and `mcErrors == 3`

Frank Meier (Duke University) Bremsstrahlung and MC matching 30.01.2023 13 / 17

※ [All you wanted to know about brems correction and MC matching and never dared asking, Belle II Physics Meeting]
https://indico.belle2.org/event/8674/#42-all-you-wanted-to-know-abou

Decay string marker and keywords

- `@` marked (composite) particle can be any particle and decay is still counted as signal
- `(misID)` marked (final state) particle can be other particle type and decay is still counted as signal
- `(decay)` decay in flight of marked particle counted as signal, *e.g.* $\pi \to \mu\nu_\mu$
- `@`, `(misID)`, and `(decay)` have to be placed in front of particle
- `...` even if there are more massive final state particles than specified, decay is counted as signal (**isSignal** behaves like **isSignalAcceptMissingMassive**)
  - missing π^0 are not counted as signal
- `?nu` neutrinos are ignored in MC matching (**isSignal** works like **isSignalAcceptMissingNeutrino**)
- `?addbrems` Bremsstrahlung correction photons are ignored in MC matching (**isSignal** works like **isSignalAcceptBremsPhotons**)
- `?gamma` missing photon(s) are ignored in MC matching (**isSignal** works like **isSignalAcceptMissingGamma**)
- `...`, `?nu`, `?addbrems`, and `?gamma` have to be placed at the end of the decay string

Frank Meier (Duke University) Bremsstrahlung and MC matching 30.01.2023 14 / 17

※ [All you wanted to know about brems correction and MC matching and never dared asking, Belle II Physics Meeting]
https://indico.belle2.org/event/8674/#42-all-you-wanted-to-know-abou

TopoAna: Topological Analysis?

※Special thanks to Jaeyoung Kim. He teaches me how to use the TopoAna package.
※He will cover "TopoAna".

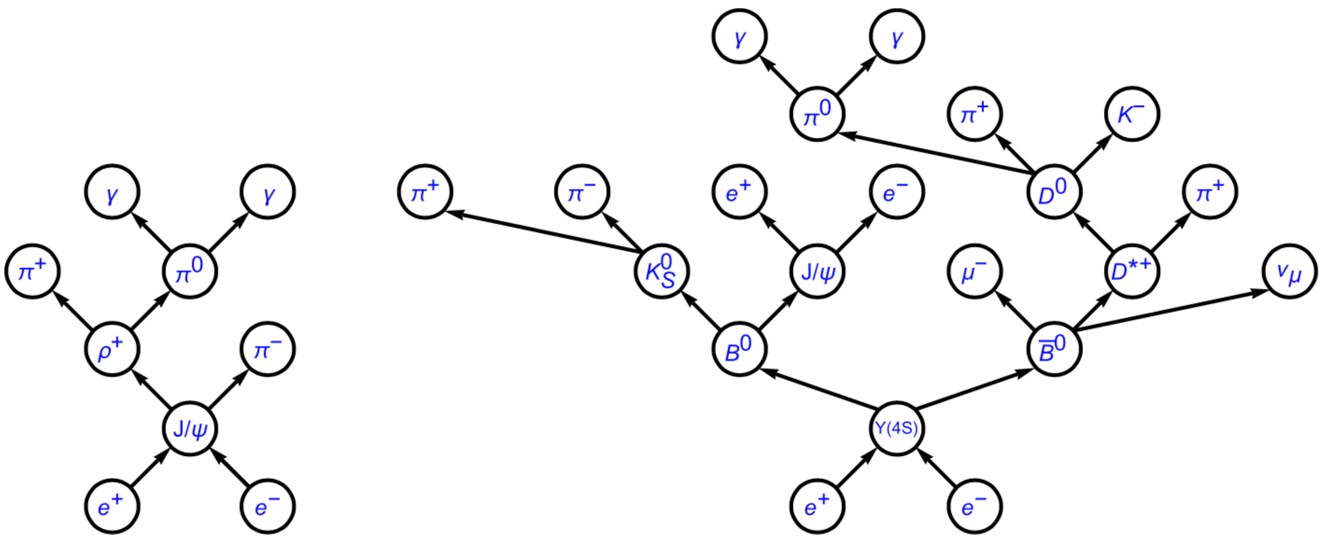
- Before TopoAna package (ex: Belle)
  - Analysts usually wrote some **private codes** to match a **few** signals and/or backgrounds for their own studies.
  - The limited functions of these codes do not satisfy the increasing demands for topology analysis.

- Motivation
  - A generic, powerful, and easy-to-use program is required for the topology analysis

- History
  - TopoAna is developed for **BESIII** (τ-Charm energy region)
  - Extended to Belle II (already applied to some analysis in Belle)

- Main usage
  - One can figure out the main backgrounds (especially peaking ones)
  - Optimize the selection criteria

※ [TopoAna: A generic tool for the event type analysis of inclusive Monte-Carlo samples in high energy physics experiments]
https://www.sciencedirect.com/science/article/pii/S0010465520302575?via%3Dihub



(a) (b)

Fig. 1. Topology diagrams of (a) $e^+e^- \to J/\psi$, $J/\psi \to \rho^+\pi^-$, $\rho^+ \to \pi^+\pi^0$, $\pi^0 \to \gamma\gamma$ and (b) $e^+e^- \to \Upsilon(4S)$, $\Upsilon(4S) \to B^0\bar{B}^0$, $B^0 \to K_S^0 J/\psi$, $\bar{B}^0 \to \mu^- D^{*+}\nu_\mu$, $K_S^0 \to \pi^+\pi^-$, $J/\psi \to e^+e^-$, $D^{*+} \to D^0\pi^+$, $D^0 \to \pi^0\pi^+K^-$, $\pi^0 \to \gamma\gamma$. As if trees grow, the diagrams are plotted from bottom to top.

※ [TopoAna: A generic tool for the event type analysis of inclusive Monte-Carlo samples in high energy physics experiments]
https://www.sciencedirect.com/science/article/pii/S0010465520302575?via%3Dihub

TopoAna: Input

```
% Names of input root files
{
  /gpfs/group/belle/users/chkim/b2ana/output/bd2tautau/signal/local/bd2tautau_local_sig_MC15_BGx1_230321_par_TopoAna_000.root
}

% TTree name
{
  tree
}

# Here, AOI, VOI, MSI, and MSD are short for array of integers, vector of integers, multiple scalar integers, and multiple scalar doubles.
% Storage type of input raw topology truth information (Six options: AOI, VOI, MSI, MSF, MSD, and MSID. Default: AOI)
{
  MSID
}

% Component analysis --- decay trees
{
  Y     100
}

% Component analysis --- cascade decay branches of particles
{
  Upsilon(4S)  Y_4S  150
  B0           B0    150
  tau+         taup  100
  tau-         taum  100
}

% Component analysis --- inclusive decay branches
{
  B0 --> tau+ tau-                  & B0Xtt    & 100
}

% Common name of output files (Default: Name of the card file)
{
  topoana
}
```

※ [Topology Analysis: Sphinx Manual]
https://b2-master.belle2.org/software/sphinx/light-2212-foldex/online_book/analysis/topoana.html

TopoAna: Output: $B^0 \to \tau^+\tau^-$

List of Tables

※ [Topology Analysis: Sphinx Manual]
https://b2-master.belle2.org/software/sphinx/light-2212-foldex/online_book/analysis/topoana.html

TopoAna: Output (1): $B^0 \to \tau^+\tau^-$

Table 1: Decay trees and their respective initial-final states.

rowNo	decay tree (decay initial-final states)	iDcyTr	nEtr	nCEtr
1	$\Upsilon(4S) \to B^0\bar{B}^0, B^0 \to \tau^+\tau^-, \bar{B}^0 \to \rho^- D^+, \tau^+ \to \mu^+\nu_\mu\bar{\nu}_\tau, \tau^- \to \mu^-\bar{\nu}_\mu\nu_\tau, \rho^- \to \pi^0\pi^-,$ $D^+ \to \pi^+\pi^+K^-$ $(\Upsilon(4S) \dashrightarrow \mu^+\mu^-\nu_\mu\bar{\nu}_\mu\nu_\tau\bar{\nu}_\tau\pi^0\pi^+\pi^+\pi^-K^-)$	491	4	4
2	$\Upsilon(4S) \to B^0\bar{B}^0, B^0 \to \rho^+ D^-, \bar{B}^0 \to \tau^+\tau^-, \rho^+ \to \pi^0\pi^+, D^- \to \pi^-\pi^-K^+, \tau^+ \to \bar{\nu}_\tau\pi^0\pi^+,$ $\tau^- \to \mu^-\bar{\nu}_\mu\nu_\tau$ $(\Upsilon(4S) \dashrightarrow \mu^-\bar{\nu}_\mu\nu_\tau\bar{\nu}_\tau\pi^0\pi^0\pi^+\pi^+\pi^-\pi^-K^+)$	493	4	8
3	$\Upsilon(4S) \to B^0\bar{B}^0, B^0 \to \pi^+\pi^+\pi^- D^-, \bar{B}^0 \to \tau^+\tau^-, D^- \to \pi^-\pi^-K^+, \tau^+ \to \mu^+\nu_\mu\bar{\nu}_\tau, \tau^- \to \mu^-\bar{\nu}_\mu\nu_\tau$ $(\Upsilon(4S) \dashrightarrow \mu^+\mu^-\nu_\mu\bar{\nu}_\mu\nu_\tau\bar{\nu}_\tau\pi^+\pi^+\pi^-\pi^-\pi^-K^+)$	893	4	12
4	$\Upsilon(4S) \to B^0\bar{B}^0, B^0 \to \tau^+\tau^-, \bar{B}^0 \to \pi^+\pi^-\pi^- D^+, \tau^+ \to e^+\nu_e\bar{\nu}_\tau\gamma^F\gamma^F, \tau^- \to e^-\bar{\nu}_e\nu_\tau, D^+ \to \pi^+\pi^+K^-$ $(\Upsilon(4S) \dashrightarrow e^+e^-\nu_e\bar{\nu}_e\nu_\tau\bar{\nu}_\tau\pi^+\pi^+\pi^+\pi^-\pi^-K^-\gamma^F\gamma^F)$	16	3	15
5	$\Upsilon(4S) \to B^0\bar{B}^0, B^0 \to \tau^+\tau^-, \bar{B}^0 \to D^{*+}a_1^-, \tau^+ \to \mu^+\nu_\mu\bar{\nu}_\tau\gamma^F, \tau^- \to \mu^-\bar{\nu}_\mu\nu_\tau, D^{*+} \to \pi^+D^0,$ $a_1^- \to \pi^- f_0(600), D^0 \to \pi^0\pi^+K^-, f_0(600) \to \pi^+\pi^-$ $(\Upsilon(4S) \dashrightarrow \mu^+\mu^-\nu_\mu\bar{\nu}_\mu\nu_\tau\bar{\nu}_\tau\pi^0\pi^+\pi^+\pi^+\pi^-\pi^-K^-\gamma^F)$	23	3	18

⋮

rowNo	decay tree (decay initial-final states)	iDcyTr	nEtr	nCEtr
97	$\Upsilon(4S) \to B^0\bar{B}^0, B^0 \to \tau^+\tau^-, \bar{B}^0 \to \pi^+\pi^-\rho^- D^{*+}, \tau^+ \to \bar{\nu}_\tau\pi^0\pi^0\pi^+, \tau^- \to e^-\bar{\nu}_e\nu_\tau\gamma^F, \rho^- \to \pi^0\pi^-,$ $D^{*+} \to \pi^+D^0, D^0 \to \pi^0\pi^+K^-$ $(\Upsilon(4S) \dashrightarrow e^-\bar{\nu}_e\nu_\tau\bar{\nu}_\tau\pi^0\pi^0\pi^0\pi^0\pi^+\pi^+\pi^+\pi^+\pi^-\pi^-K^-\gamma^F)$	71	2	279
98	$\Upsilon(4S) \to B^0\bar{B}^0, B^0 \to \pi^+\pi^+\rho^- D^{*-}\gamma^F, \bar{B}^0 \to \tau^+\tau^-, \rho^- \to \pi^0\pi^-, D^{*-} \to \pi^0 D^-, \tau^+ \to \bar{\nu}_\tau\pi^0\pi^+,$ $\tau^- \to e^-\bar{\nu}_e\nu_\tau\gamma^F, D^- \to \pi^-K^+K^-$ $(\Upsilon(4S) \dashrightarrow e^-\bar{\nu}_e\nu_\tau\bar{\nu}_\tau\pi^0\pi^0\pi^0\pi^+\pi^+\pi^+\pi^-\pi^-K^+K^-\gamma^F\gamma^F)$	74	2	281
99	$\Upsilon(4S) \to B^0\bar{B}^0, B^0 \to \tau^+\tau^-, \bar{B}^0 \to \pi^0\rho^- D^{*+}, \tau^+ \to \bar{\nu}_\tau\pi^0\pi^+, \tau^- \to e^-\bar{\nu}_e\nu_\tau, \rho^- \to \pi^0\pi^-,$ $D^{*+} \to \pi^+D^0, D^0 \to K^-a_1^+, a_1^+ \to \pi^+\pi^+\pi^-\gamma^F$ $(\Upsilon(4S) \dashrightarrow e^-\bar{\nu}_e\nu_\tau\bar{\nu}_\tau\pi^0\pi^0\pi^0\pi^+\pi^+\pi^+\pi^+\pi^-\pi^-K^-\gamma^F)$	76	2	283
100	$\Upsilon(4S) \to B^0\bar{B}^0, B^0 \to D^{*-}a_1^+, \bar{B}^0 \to \tau^+\tau^-, D^{*-} \to \pi^0 D^-, a_1^+ \to \pi^+\pi^+\pi^-\gamma^F, \tau^+ \to \bar{\nu}_\tau\pi^0\pi^+,$ $\tau^- \to \nu_\tau\pi^0\pi^+\pi^-\pi^-, D^- \to \pi^0\pi^-K_S^0, K_S^0 \to \pi^+\pi^-$ $(\Upsilon(4S) \dashrightarrow \nu_\tau\bar{\nu}_\tau\pi^0\pi^0\pi^0\pi^0\pi^+\pi^+\pi^+\pi^+\pi^+\pi^-\pi^-\pi^-\pi^-\pi^-\gamma^F)$	82	2	285
rest	$\Upsilon(4S) \to$ others (1696 in total) $(\Upsilon(4S) \dashrightarrow$ corresponding to others)	—	2030	2315

※ [Topology Analysis: Sphinx Manual] https://b2-master.belle2.org/software/sphinx/light-2212-foldex/online_book/analysis/topoana.html

TopoAna: Output (3): $B^0 \to \tau^+\tau^-$

Table 3: Cascade decay branches of B^0.

rowNo	cascade decay branch of B^0	iCascDcyBrP	nCase	nCCase
1	$B^0 \to \tau^+\tau^-, \tau^+ \to \bar{\nu}_\tau\pi^0\pi^+, \tau^- \to \nu_\tau\pi^0\pi^-$	6	73	73
2	$B^0 \to \tau^+\tau^-, \tau^+ \to \mu^+\nu_\mu\bar{\nu}_\tau, \tau^- \to \nu_\tau\pi^0\pi^-$	39	46	119
3	$B^0 \to \tau^+\tau^-, \tau^+ \to \mu^+\nu_\mu\bar{\nu}_\tau, \tau^- \to \mu^-\bar{\nu}_\mu\nu_\tau$	30	43	162
4	$B^0 \to \tau^+\tau^-, \tau^+ \to \bar{\nu}_\tau\pi^0\pi^+, \tau^- \to \mu^-\bar{\nu}_\mu\nu_\tau$	124	40	202
5	$B^0 \to \tau^+\tau^-, \tau^+ \to \bar{\nu}_\tau\pi^0\pi^+, \tau^- \to \nu_\tau\pi^-$	102	35	237
6	$B^0 \to \tau^+\tau^-, \tau^+ \to \mu^+\nu_\mu\bar{\nu}_\tau, \tau^- \to \nu_\tau\pi^0\pi^0\pi^-$	139	30	267
7	$B^0 \to \tau^+\tau^-, \tau^+ \to \bar{\nu}_\tau\pi^+, \tau^- \to \nu_\tau\pi^0\pi^-$	109	29	296
8	$B^0 \to \tau^+\tau^-, \tau^+ \to \bar{\nu}_\tau\pi^0\pi^+, \tau^- \to \nu_\tau\pi^0\pi^0\pi^-$	99	24	320
9	$B^0 \to \tau^+\tau^-, \tau^+ \to e^+\nu_e\bar{\nu}_\tau, \tau^- \to \mu^-\bar{\nu}_\mu\nu_\tau$	122	24	344
10	$B^0 \to \tau^+\tau^-, \tau^+ \to e^+\nu_e\bar{\nu}_\tau, \tau^- \to \nu_\tau\pi^0\pi^-$	180	24	368
	⋮			
145	$B^0 \to \tau^+\tau^-, \tau^+ \to \bar{\nu}_\tau\pi^0\pi^+, \tau^- \to \nu_\tau\pi^-\gamma^F$	732	3	1173
146	$B^0 \to \pi^+\pi^-p\bar{\Lambda}_c^-, \bar{\Lambda}_c^- \to \pi^-K^+\bar{p}$	741	3	1176
147	$B^0 \to \pi^0\pi^+\eta D^{*-}, \eta \to \gamma\gamma, D^{*-} \to \pi^-\bar{D}^0, \bar{D}^0 \to \pi^0\pi^-K^+$	755	3	1179
148	$B^0 \to D^-D_s^+, D^- \to \pi^-\rho^-K^+, D_s^+ \to \pi^+\pi^-K^+, \rho^- \to \pi^0\pi^-\gamma^F$	764	3	1182
149	$B^0 \to \pi^0\pi^0\rho^+D^{*-}, \rho^+ \to \pi^0\pi^+, D^{*-} \to \pi^-\bar{D}^0, \bar{D}^0 \to K^+K^-$	765	3	1185
150	$B^0 \to \pi^0K^{*+}D^-, K^{*+} \to \pi^0K^+, D^- \to \pi^-K^+K^-$	840	3	1188
rest	$B^0 \to$ others (898 in total)	—	1127	2315

※ [Topology Analysis: Sphinx Manual] https://b2-master.belle2.org/software/sphinx/light-2212-foldex/online_book/analysis/topoana.html

References

References

- [Monte Carlo matching in the Belle II software]
  - https://doi.org/10.1051/epjconf/202125103021
- [All you wanted to know about brems correction and MC matching and never dared asking, Belle II Physics Meeting (2023.02.06 MON)]
  - https://indico.belle2.org/event/8674/#42-all-you-wanted-to-know-abou
- [TopoAna: A generic tool for the event type analysis of inclusive Monte-Carlo samples in high energy physics experiments]
  - https://www.sciencedirect.com/science/article/pii/S0010465520302575?via%3Dihub
- [Topology Analysis: Sphinx Manual]
  - https://b2-master.belle2.org/software/sphinx/light-2212-foldex/online_book/analysis/topoana.html