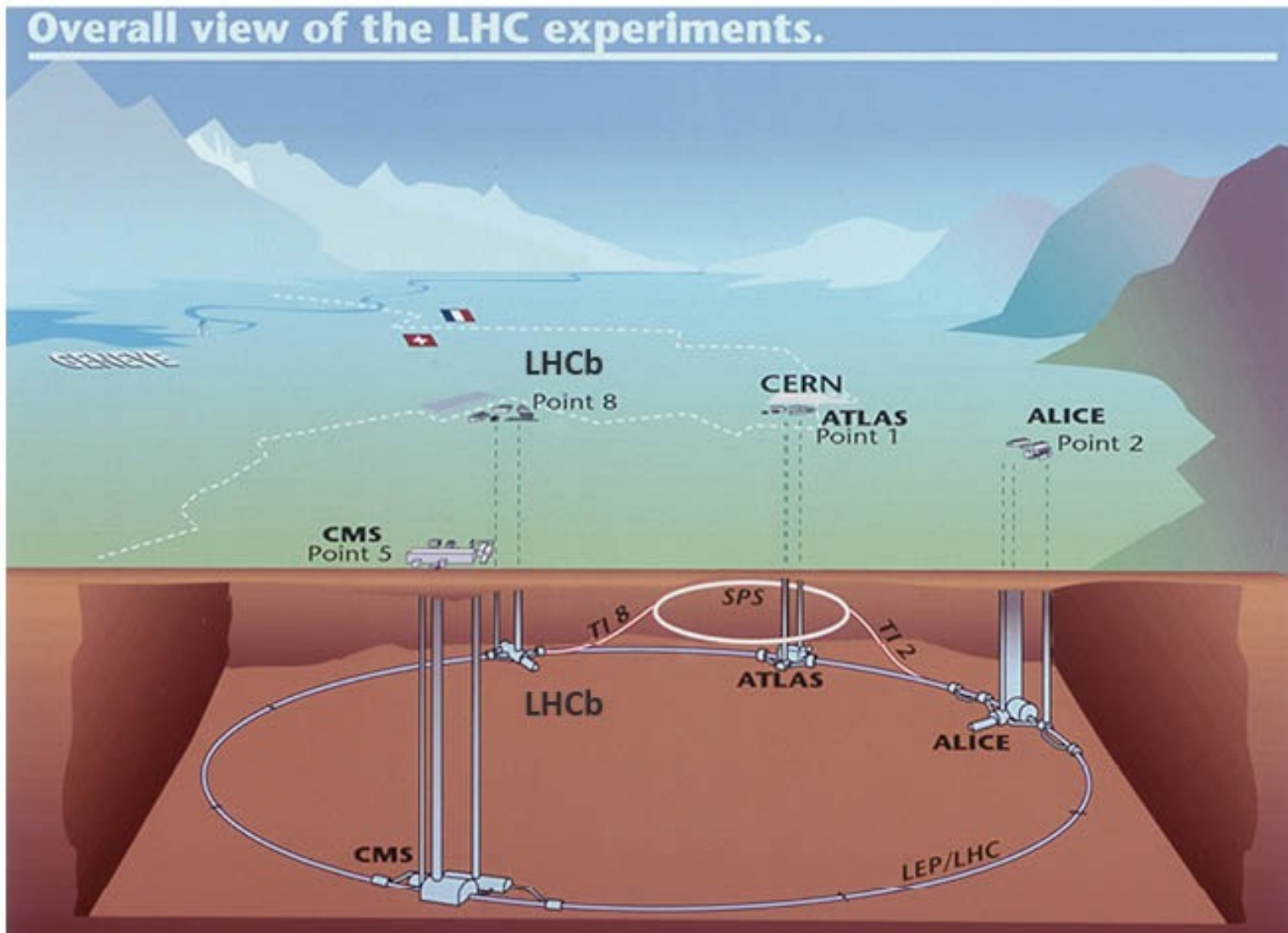# An automated tool to facilitate consistent test-driven development of trigger selections for LHCb's Run 3

Ross Hunter*, Olli Lupton*, Rosen Matev†, Sascha Stahl†, Mika Vesterinen*

*University of Warwick, U.K, †CERN, Switzerland

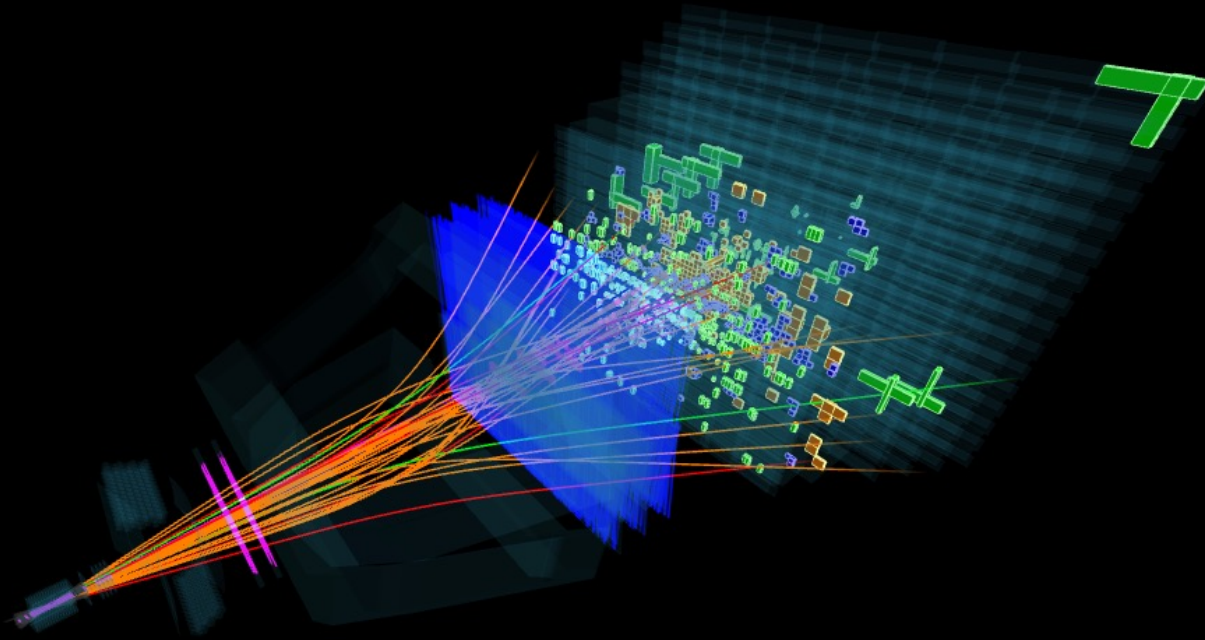*25th International Conference on Computing in High Energy and Nuclear Physics*

LHCb currently shutdown for upgrade. Restart for Run 3 in 2022.

Event 158826354
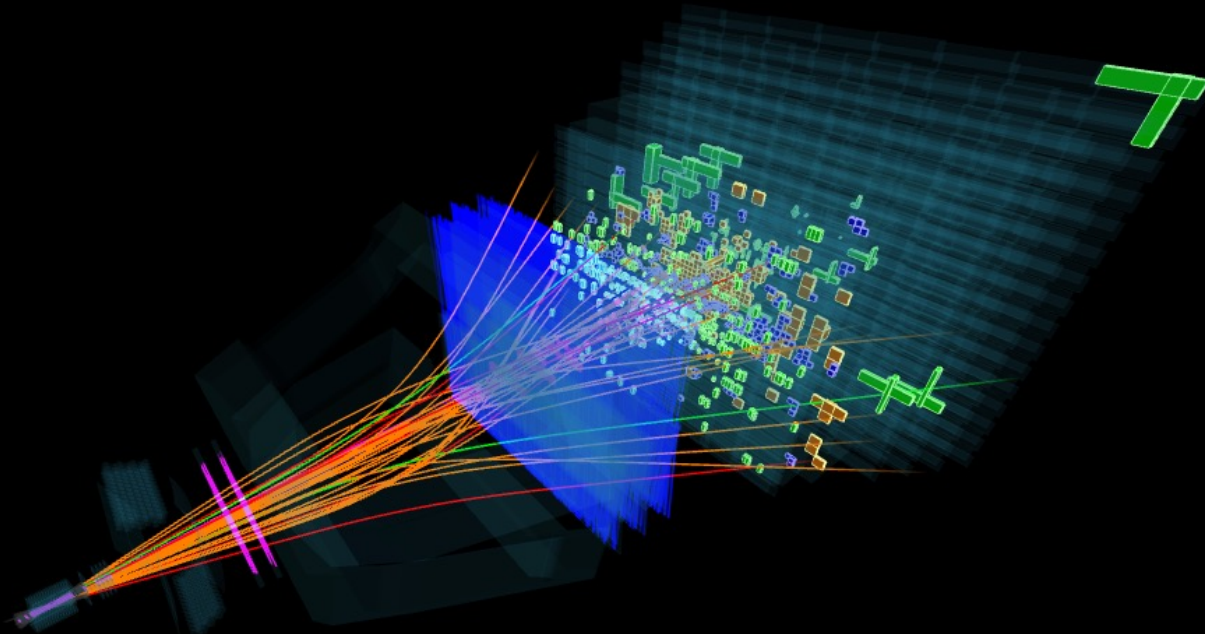Run 206854
Sat, 28 Apr 2018 21:48:17

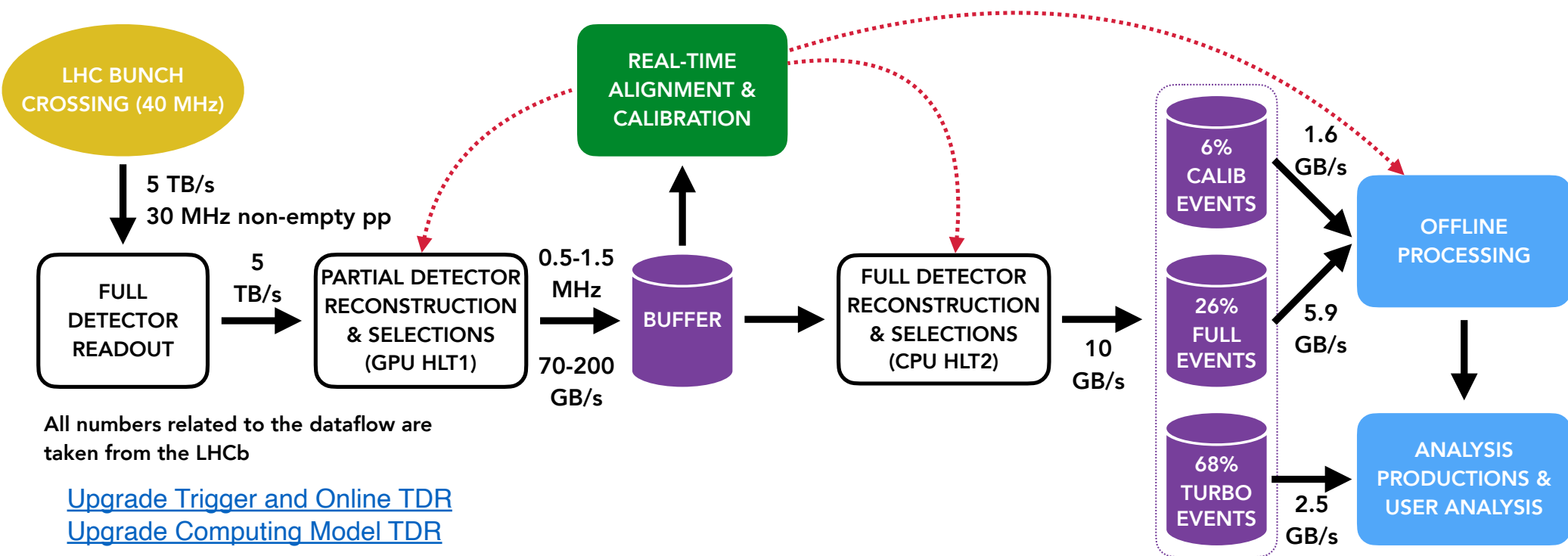30 million events per second! (x5 collisions per event in 2022)
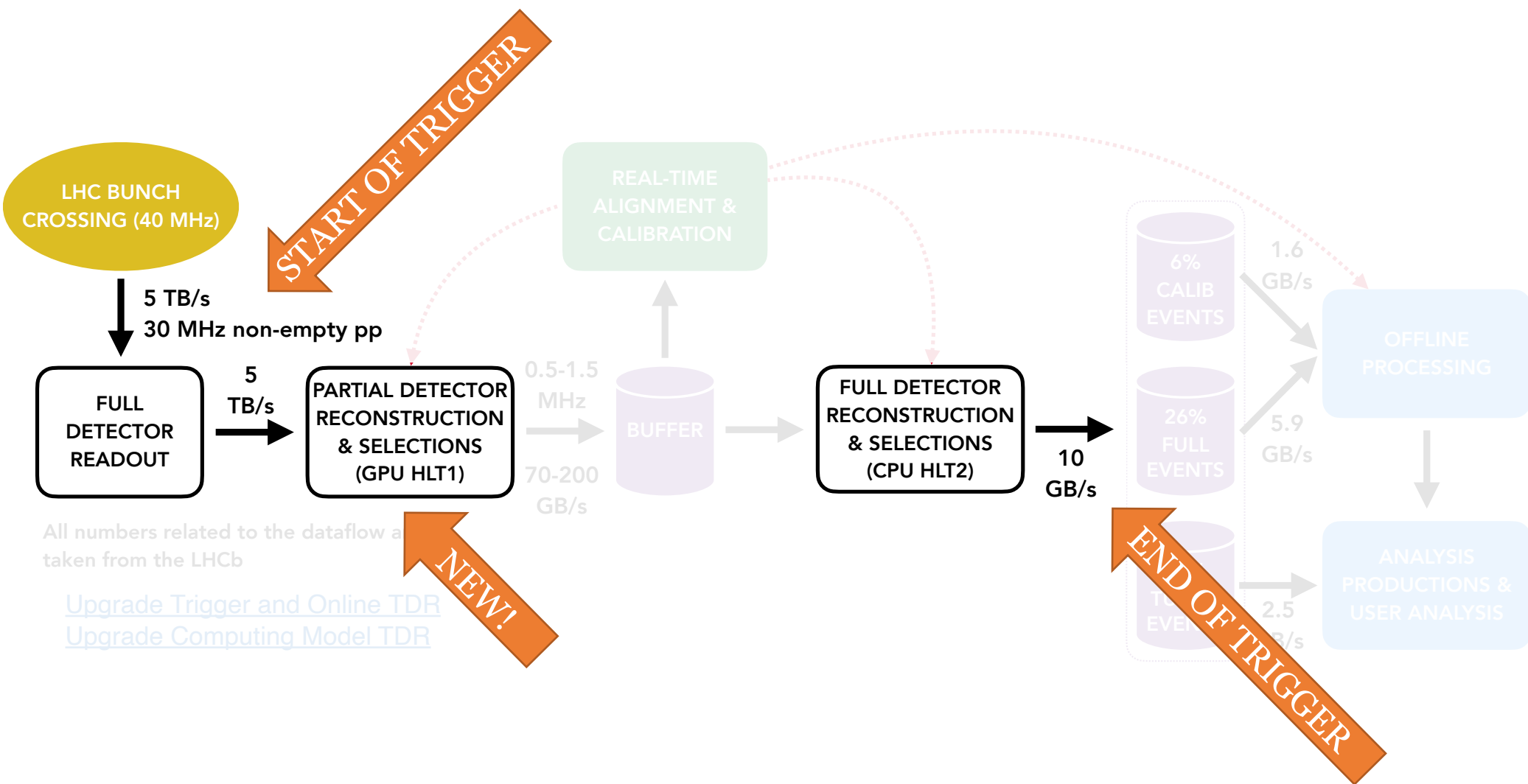
30 million events per second! (x5 collisions per event in 2022)
Interesting collisions are extremely rare.
We filter them out with a **trigger.**

The LHCb Dataflow in Run 3 (2022-2025)

LHC BUNCH CROSSING (40 MHz)

5 TB/s
30 MHz non-empty pp

FULL DETECTOR READOUT

5 TB/s

PARTIAL DETECTOR RECONSTRUCTION & SELECTIONS (GPU HLT1)

0.5-1.5 MHz

70-200 GB/s

BUFFER

REAL-TIME ALIGNMENT & CALIBRATION

FULL DETECTOR RECONSTRUCTION & SELECTIONS (CPU HLT2)

10 GB/s

6% CALIB EVENTS

26% FULL EVENTS

68% TURBO EVENTS

1.6 GB/s

5.9 GB/s

2.5 GB/s

OFFLINE PROCESSING

ANALYSIS PRODUCTIONS & USER ANALYSIS

All numbers related to the dataflow are taken from the LHCb

Upgrade Trigger and Online TDR
Upgrade Computing Model TDR

The LHCb Dataflow in Run 3 (2022-2025)

LHC BUNCH CROSSING (40 MHz)

5 TB/s
30 MHz non-empty pp

FULL DETECTOR READOUT

START OF TRIGGER

5 TB/s

PARTIAL DETECTOR RECONSTRUCTION & SELECTIONS (GPU HLT1)

NEW!

0.5-1.5 MHz

70-200 GB/s

BUFFER

REAL-TIME ALIGNMENT & CALIBRATION

FULL DETECTOR RECONSTRUCTION & SELECTIONS (CPU HLT2)

10 GB/s

END OF TRIGGER

6% CALIB EVENTS

26% FULL EVENTS

68% TURBO EVENTS

1.6 GB/s

5.9 GB/s

2.5 GB/s

OFFLINE PROCESSING

ANALYSIS PRODUCTIONS & USER ANALYSIS

All numbers related to the dataflow a
taken from the LHCb

Upgrade Trigger and Online TDR
Upgrade Computing Model TDR

The LHCb Dataflow in Run 3 (2022-2025)

START OF TRIGGER

LHC BUNCH CROSSING (40 MHz)

5 TB/s
30 MHz non-empty pp

FULL DETECTOR READOUT

5 TB/s

PARTIAL DETECTOR RECONSTRUCTION & SELECTIONS (GPU HLT1)

NEW!

0.5-1.5 MHz

70-200 GB/s

BUFFER

REAL-TIME ALIGNMENT & CALIBRATION

FULL DETECTOR RECONSTRUCTION & SELECTIONS (CPU HLT2)

10 GB/s

6% CALIB EVENTS

26% FULL EVENTS

68% TURBO EVENTS

1.6 GB/s

5.9 GB/s

2.5 GB/s

OFFLINE PROCESSING

ANALYSIS PRODUCTIONS & USER ANALYSIS

END OF TRIGGER

All numbers related to the dataflow are taken from the LHCb

Upgrade Trigger and Online TDR
Upgrade Computing Model TDR

Ideally, also: $\varepsilon_{interesting} = 1, \ \varepsilon_{boring} = 0...$

The LHCb Dataflow in Run 3 (2022-2025)

LHC BUNCH CROSSING (40 MHz)

5 TB/s
30 MHz non-empty pp

FULL DETECTOR READOUT

5 TB/s

PARTIAL DETECTOR RECONSTRUCTION & SELECTIONS (GPU HLT1)

START OF TRIGGER

NEW!

0.5-1.5 MHz

70-200 GB/s

BUFFER

REAL-TIME ALIGNMENT & CALIBRATION

FULL DETECTOR RECONSTRUCTION & SELECTIONS (CPU HLT2)

10 GB/s

6% CALIB EVENTS

26% FULL EVENTS

68% TURBO EVENTS

END OF TRIGGER

1.6 GB/s

OFFLINE PROCESSING

5.9 GB/s

2.5 GB/s

ANALYSIS PRODUCTIONS & USER ANALYSIS

All numbers related to the dataflow are taken from the LHCb

Upgrade Trigger and Online TDR
Upgrade Computing Model TDR

Ideally, also: $\varepsilon_{interesting} = 1$, $\varepsilon_{boring} = 0$...
However, expect signal rate 10 times higher! (5x L_{int}, 2x from no L0)

Test-driven trigger development
(with simulated data, before taking real data)

Bandwidth (MB/s)

Trigger Rate (kHz) x Event Size (kB)

Signal efficiency

triggered events /
we expect to trigger on

Test-driven trigger development
(with simulated data, before taking real data)

Bandwidth (MB/s)
Trigger Rate (kHz) x Event Size (kB)

Signal efficiency
triggered events /
we expect to trigger on

Previously done by independent developers!

Test-driven trigger development
(with simulated data, before taking real data)

Bandwidth (MB/s)
Trigger Rate (kHz) x Event Size (kB)

Signal efficiency
triggered events /
we expect to trigger on

Previously done by independent developers!

HltEfficiencyChecker was constructed to give these metrics in a consistent, transparent, easy-to-use and automated way.

Test-driven trigger development
(with simulated data, before taking real data)

Bandwidth (MB/s)
Trigger Rate (kHz) x Event Size (kB)

Signal efficiency
triggered events /
we expect to trigger on

Previously done by independent developers!

Everyone uses the same definitions

HltEfficiencyChecker was constructed to give these metrics in a consistent, transparent, easy-to-use and automated way.

Thoroughly-documented, accessible & centralized code

"Beginner" and "developer" modes, tutorials etc.

Results in 1 command

The problem with efficiencies

- The signal "decision" (DEC)* efficiency does not give the full picture.

* # triggered events /
# we expect to trigger on

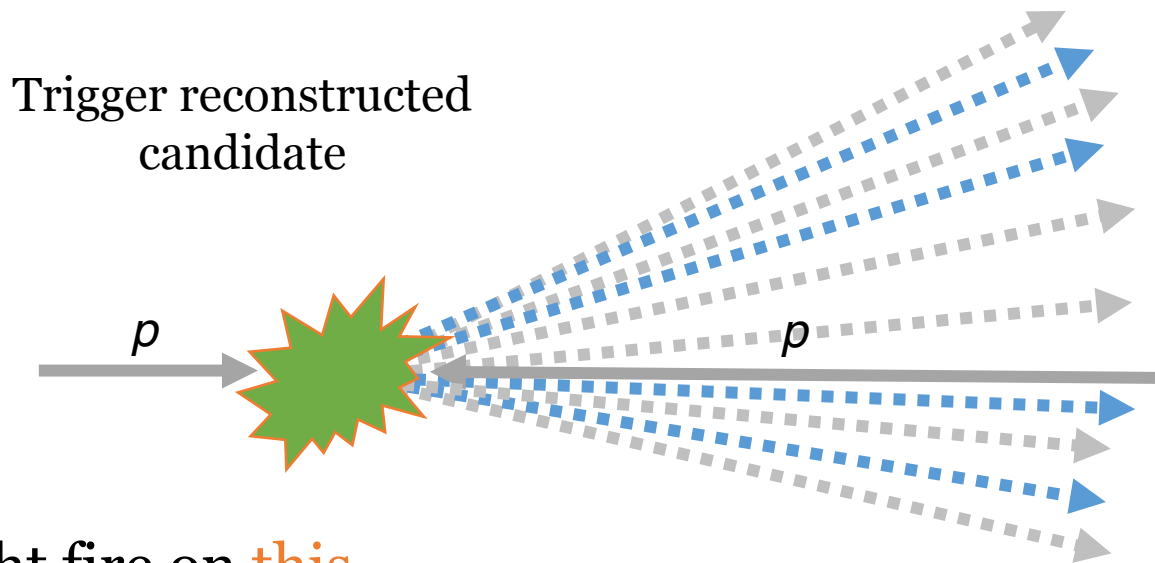Ross Hunter, University of Warwick

The problem with efficiencies

- The signal "decision" (DEC) efficiency does not give the full picture.
- Consider $B_s \to J/\Psi\, \phi$ with $(\phi \to K^+ K^-)$ and $(J/\Psi \to \mu^+ \mu^-)$...

True simulated
candidate

p

p

μ

K

K

μ
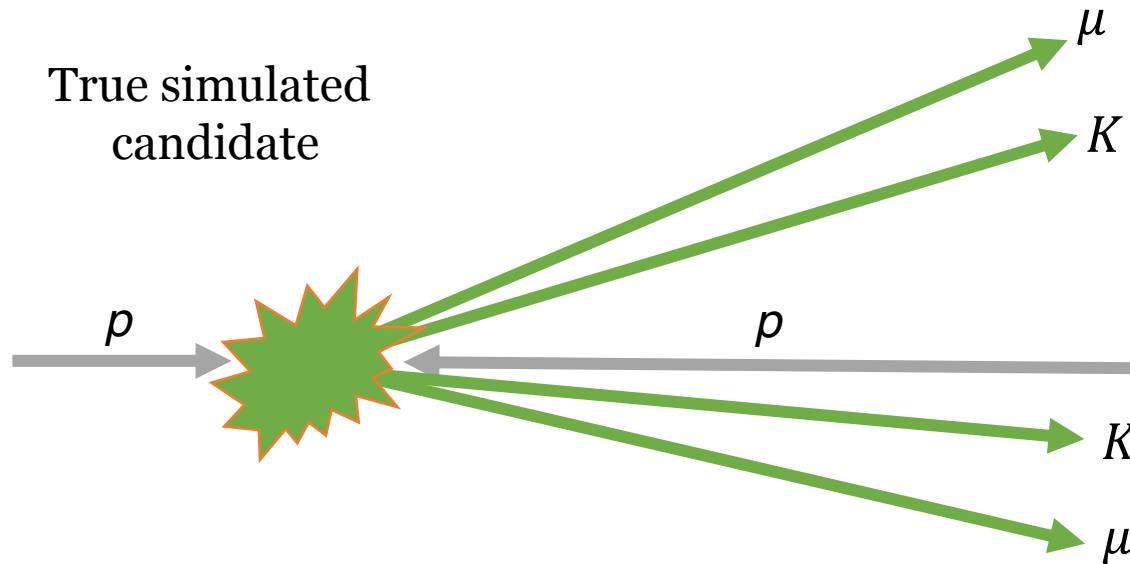
The problem with efficiencies

- The signal "decision" (DEC) efficiency does not give the full picture…
- Consider $B_s \to J/\Psi\,\phi$ with $(\phi \to K^+K^-)$ and $(J/\Psi \to \mu^+\mu^-)$…
- … in a hadron collider:

The problem with efficiencies

- The signal "decision" (DEC) efficiency does not give the full picture...
- Consider $B_s \to J/\Psi \, \phi$ with $(\phi \to K^+ K^-)$ and $(J/\Psi \to \mu^+ \mu^-)$...
- ... in a hadron collider:

Trigger reconstructed candidate
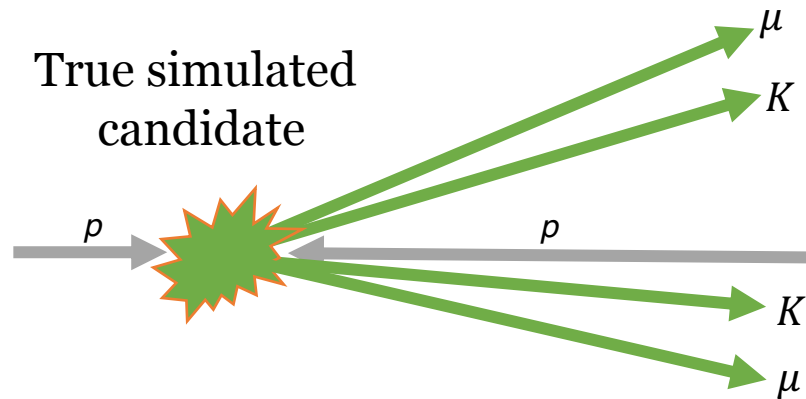
p

p
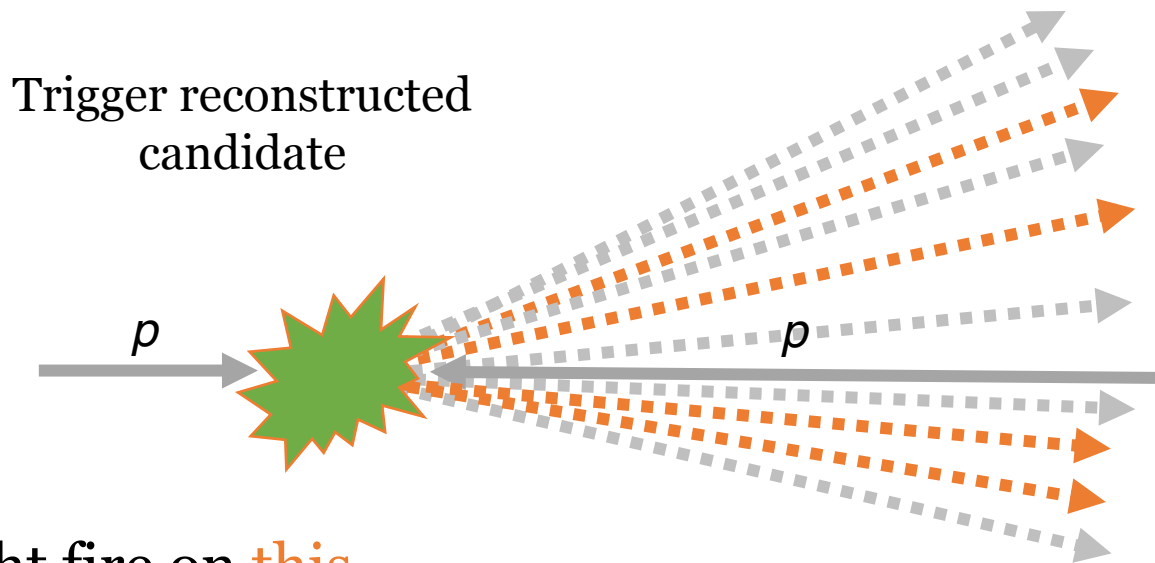
- Trigger might fire on this,

The problem with efficiencies

- The signal "decision" (DEC) efficiency does not give the full picture...
- Consider $B_s \to J/\Psi \, \phi$ with $(\phi \to K^+ K^-)$ and $(J/\Psi \to \mu^+ \mu^-)$...
- ... in a hadron collider:

Trigger reconstructed candidate

p

p

- Trigger might fire on this,
- Or this...

The problem with efficiencies

- The signal "decision" (DEC) efficiency does not give the full picture...
- Consider $B_s \rightarrow J/\Psi \, \phi$ with $(\phi \rightarrow K^+ K^-)$ and $(J/\Psi \rightarrow \mu^+ \mu^-)$...
- ... in a hadron collider:
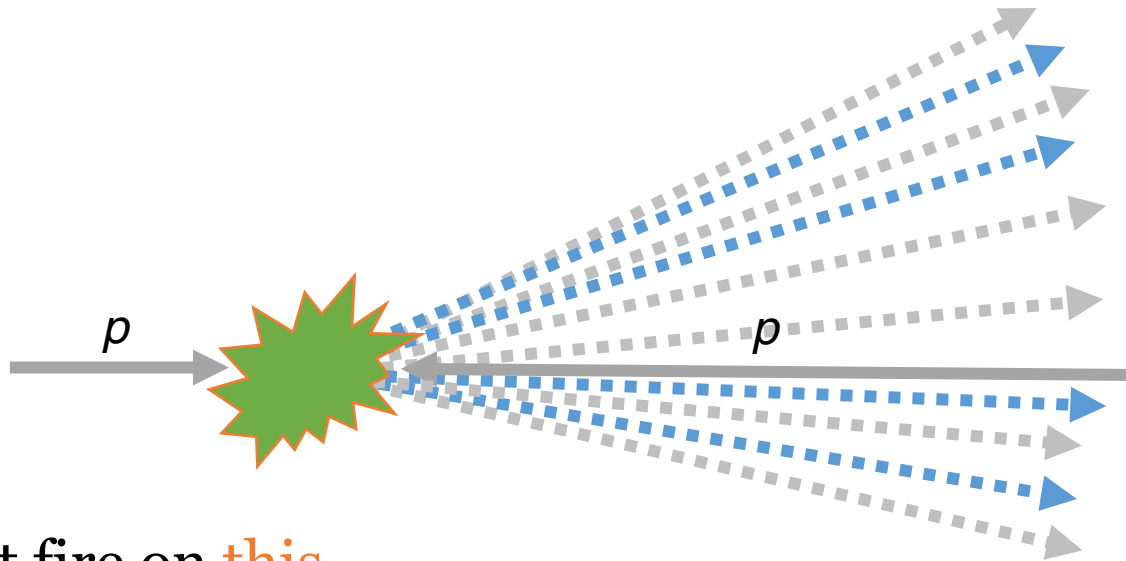
- Trigger might fire on this,
- Or this...
- ✗ Both of these count towards the DEC efficiency, signal or not.
- ✓ Need an efficiency that counts **only signal triggers**.

A better "trigger-on-signal" (TOS) efficiency

TOS efficiency = # triggered & **matched** events / # we expect to trigger on

A better "trigger-on-signal" (TOS) efficiency

TOS efficiency = # triggered & **matched** events / # we expect to trigger on

True simulated candidate

p p μ K K μ

simulated_hits =
    [hit1, hit2, ...]

Trigger reconstructed candidate

p p
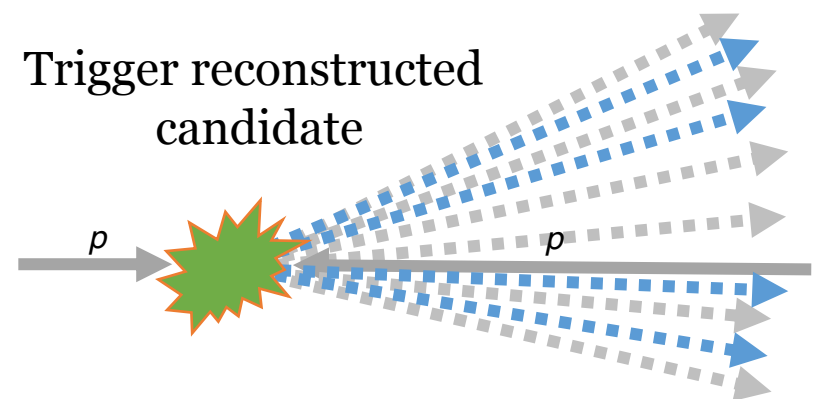
reconstructed_hits =
    [hit1', hit2', ...]

Matched if #overlapping hits $\geq f \times$ #trigger candidate hits

# A better "trigger-on-signal" (TOS) efficiency

TOS efficiency = # triggered & **matched** events / # we expect to trigger on

True simulated candidate

p p

μ

K

K

μ

simulated_hits =
    [hit1, hit2, ...]

Trigger reconstructed candidate

p p
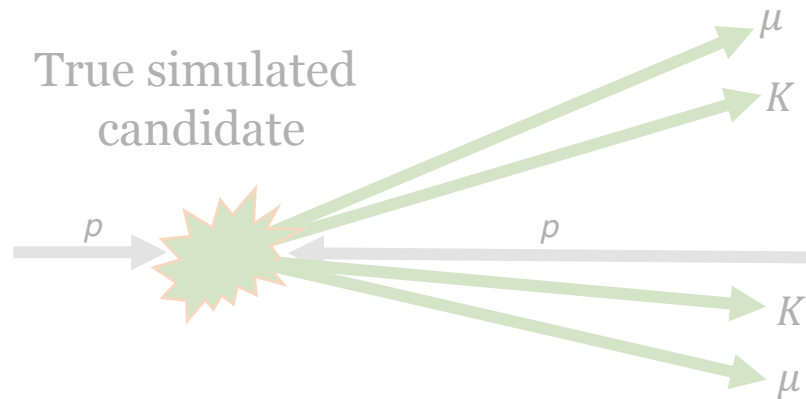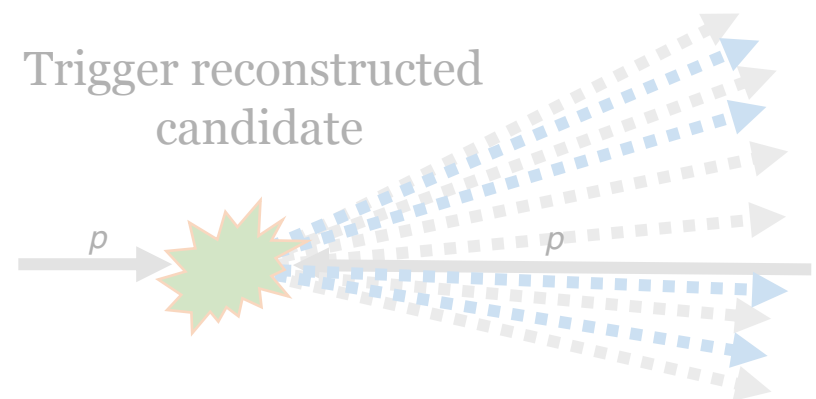
reconstructed_hits =
    [hit1', hit2', ...]

Matched if #overlapping hits ≥ f × #trigger candidate hits

What minimum matching fraction f should we choose?

Choice of minimum matching fraction

Reconstructed tracks from $B_s \to J/\Psi (\to \mu\mu)\phi(\to KK)$:
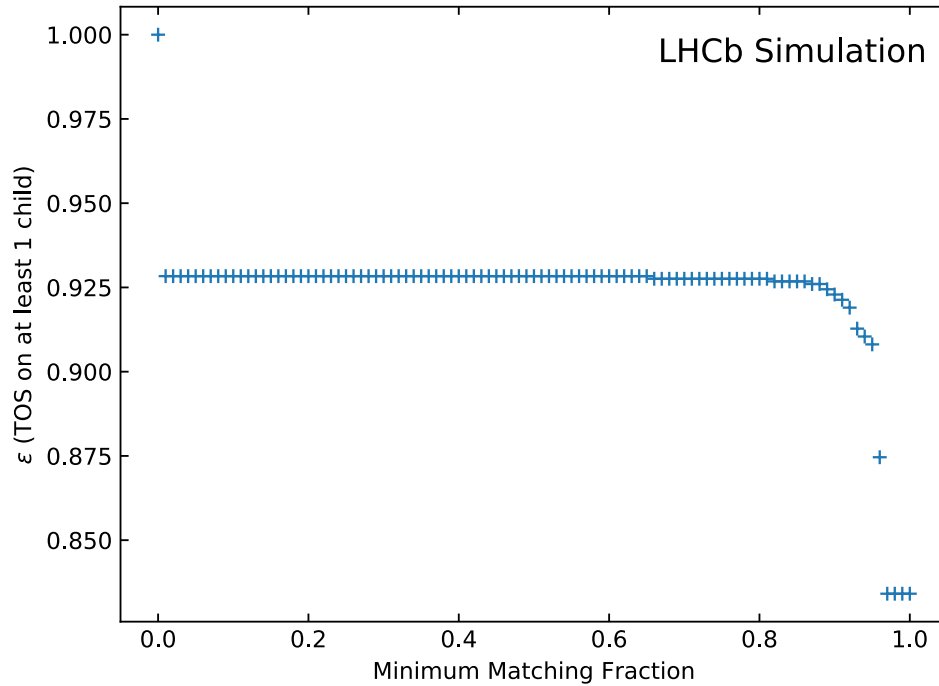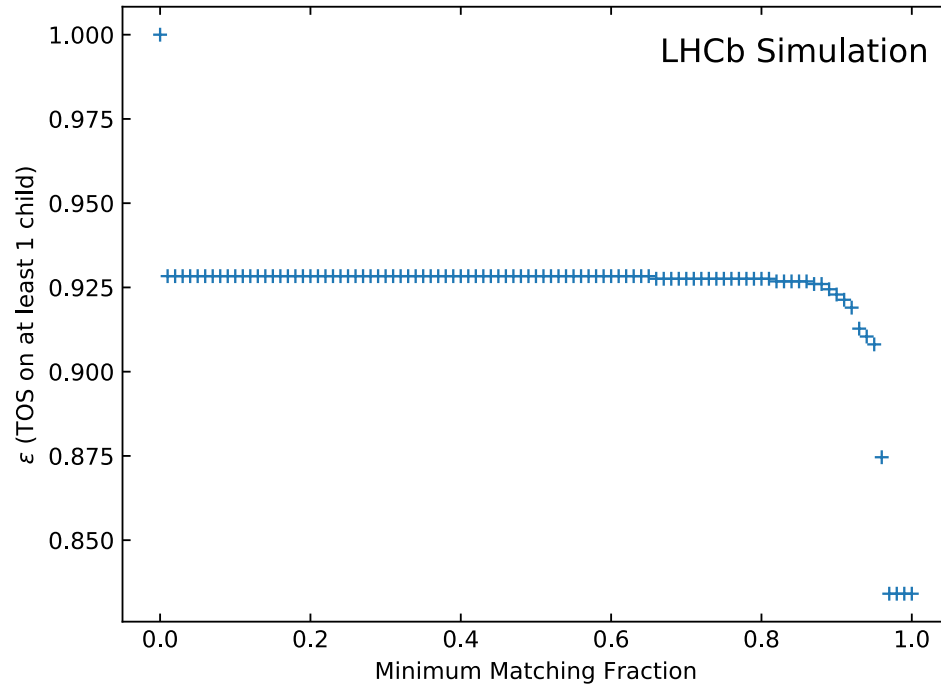

1-track selection (looks for $K^\pm, \mu^\pm$)
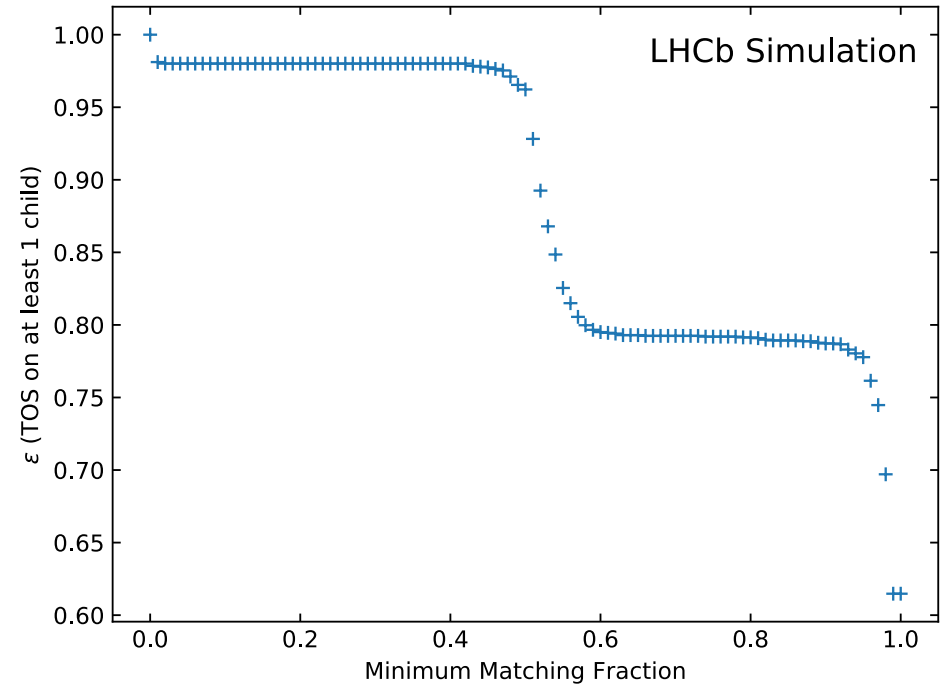
2-track selection (looks for $J/\Psi, \phi$)

Choice of minimum matching fraction
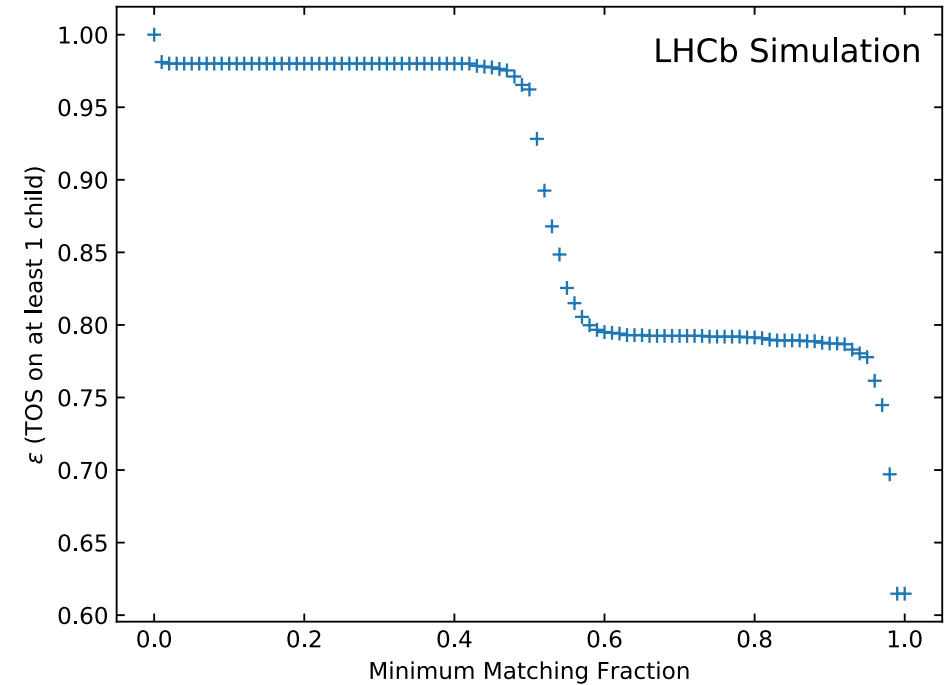
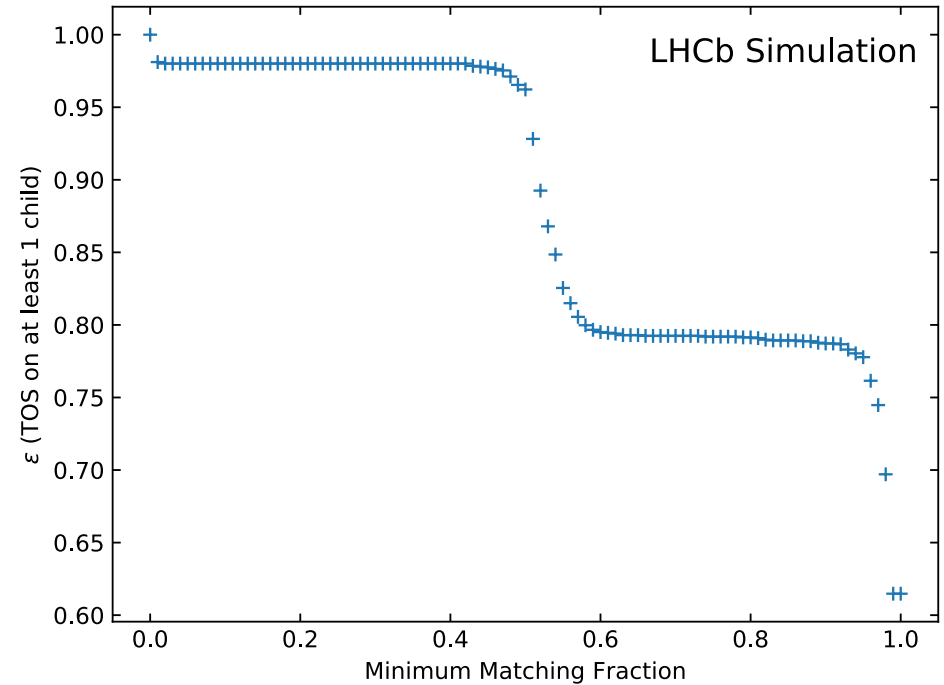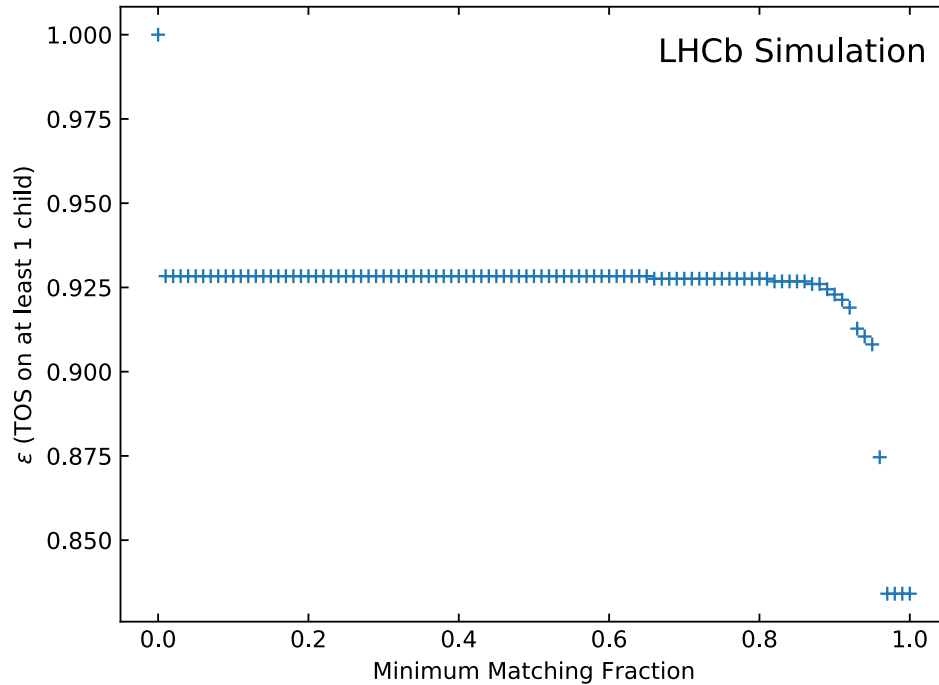## Reconstructed tracks from $B_s \to J/\Psi\,(\to \mu\mu)\phi(\to KK)$:

1-track selection (looks for $K^\pm, \mu^\pm$)   2-track selection (looks for $J/\Psi, \phi$)

- Reject good (accept bad) matches at very high (very low) f,

Choice of minimum matching fraction

Reconstructed tracks from $B_s \to J/\Psi (\to \mu\mu)\phi(\to KK)$:
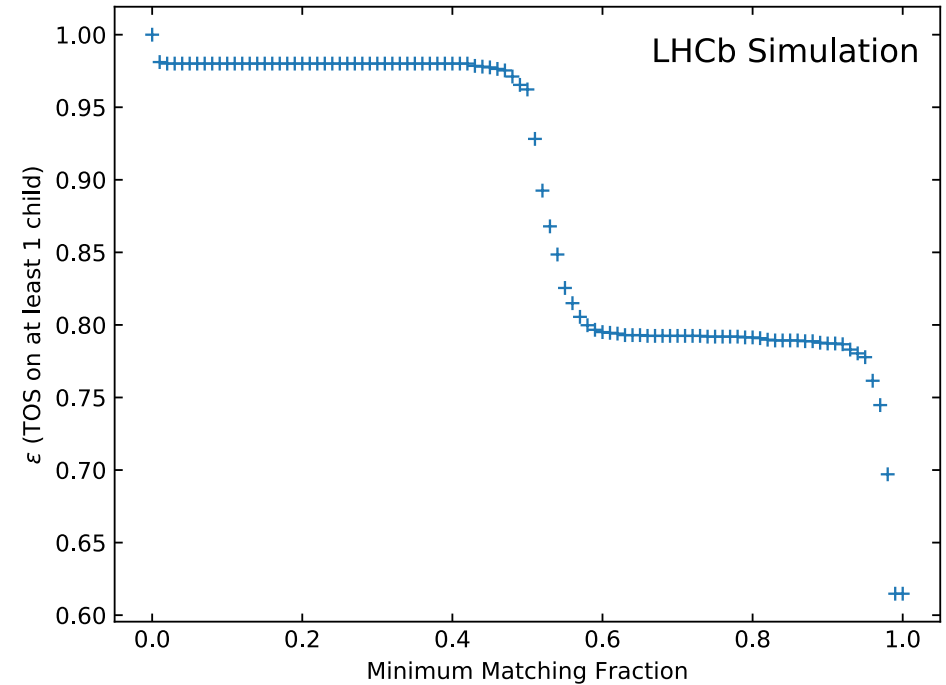

1-track selection (looks for $K^{\pm}, \mu^{\pm}$) 2-track selection (looks for $J/\Psi, \phi$)

- Reject good (accept bad) matches at very high (very low) f,
  - Step at $f = 0.5$ where match to 1 of the 2 tracks well,

Choice of minimum matching fraction

Reconstructed tracks from $B_s \rightarrow J/\Psi (\rightarrow \mu\mu)\phi(\rightarrow KK)$:

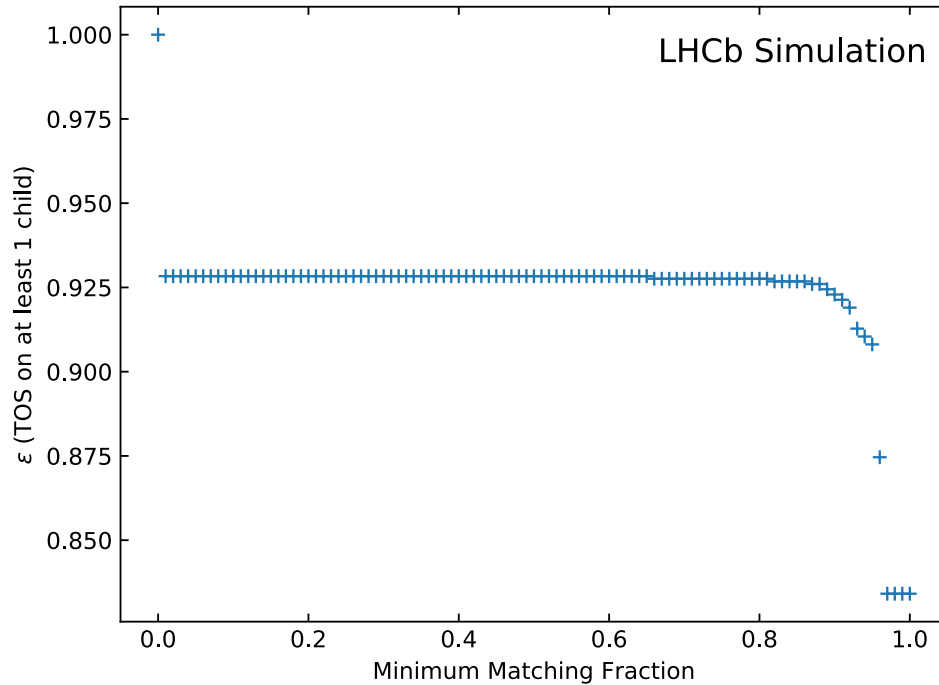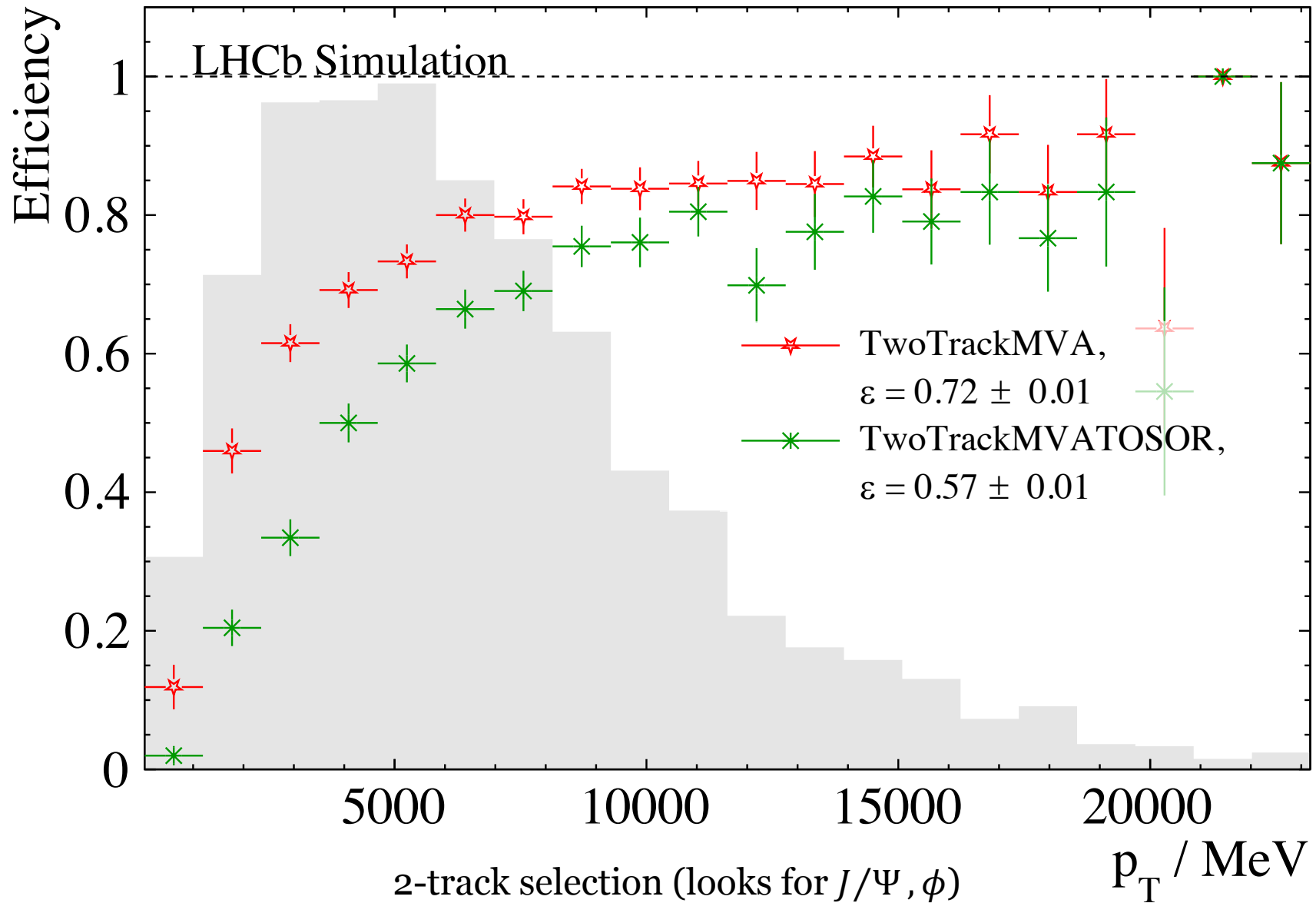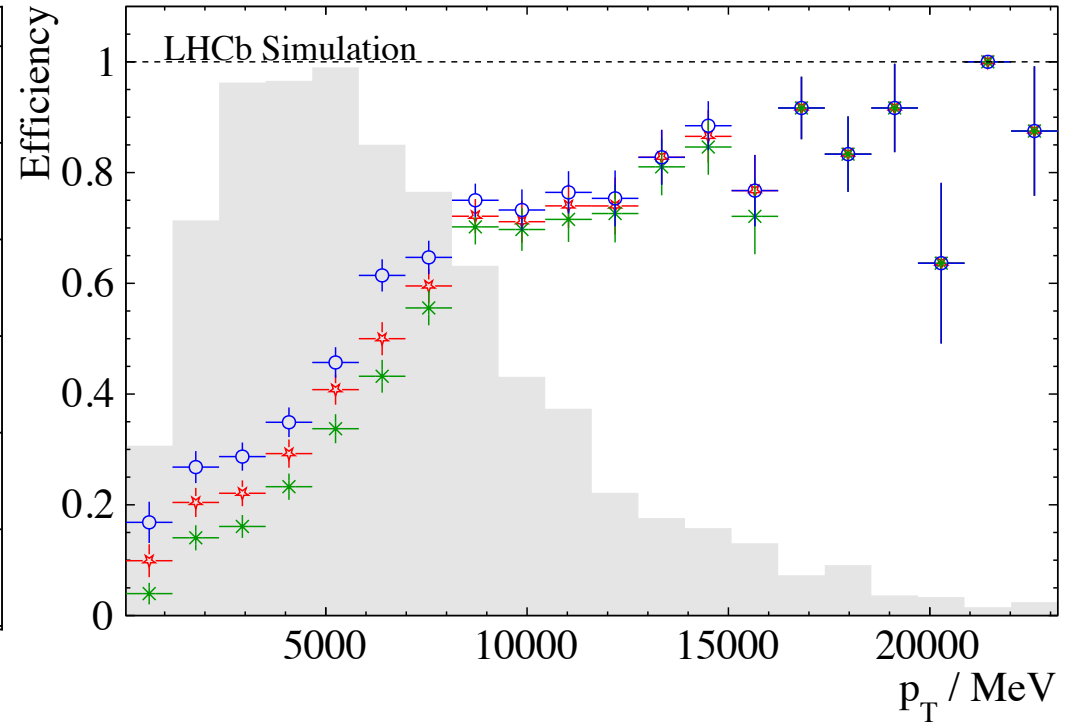
1-track selection (looks for $K^{\pm}, \mu^{\pm}$)   2-track selection (looks for $J/\Psi, \phi$)

- Reject good (accept bad) matches at very high (very low) f,
  - Step at $f = 0.5$ where match to 1 of the 2 tracks well,
- "Goldilocks" region around $f = 0.7$ where ε is stable & reliable.

Comparison of DEC and TOS

LHCb Simulation

TwoTrackMVA, $\varepsilon = 0.72 \pm 0.01$

TwoTrackMVATOSOR, $\varepsilon = 0.57 \pm 0.01$

Efficiency

2-track selection (looks for $J/\Psi, \phi$)

p_T / MeV

Selection tweaking in action
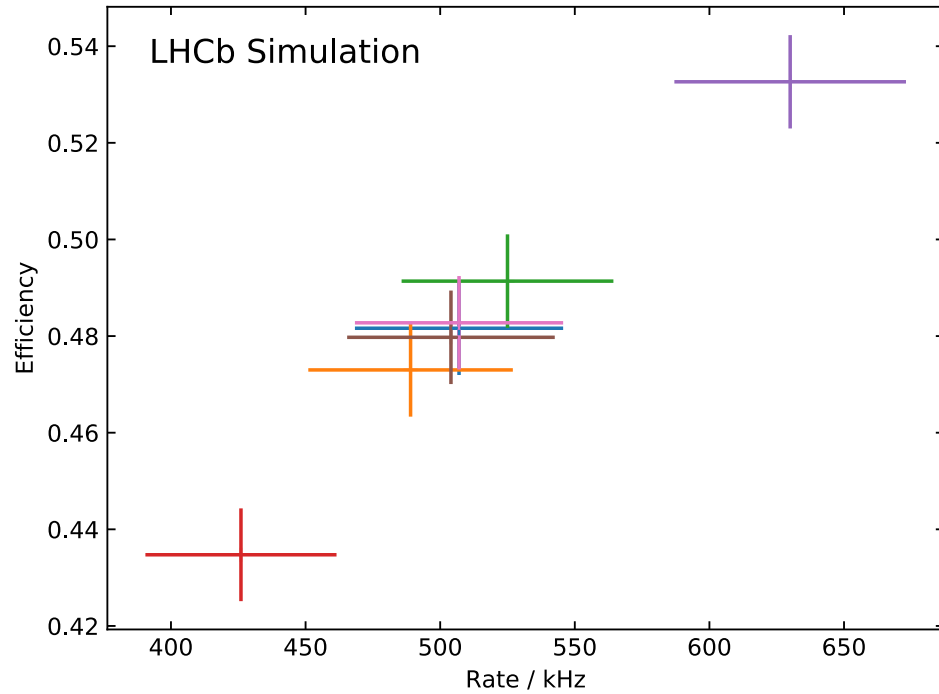
⬆️ Running many tweaked versions of the same selection ⬆️

Summary

- Balancing high signal efficiency with availability of data storage requires a highly optimised LHCb trigger.

Summary

- Balancing high signal efficiency with availability of data storage requires a highly optimised LHCb trigger.

- Developed a centralised, automated, easy-to-use and consistent tool to facilitate this optimisation in a test-driven way.

Summary

- Balancing high signal efficiency with availability of data storage requires a highly optimised LHCb trigger.

- Developed a centralised, automated, easy-to-use and consistent tool to facilitate this optimisation in a test-driven way.

- This tool also calculates a "trigger-on-signal" efficiency, which is more characteristic of selection performance.

Summary

- Balancing high signal efficiency with availability of data storage requires a highly optimised LHCb trigger.

- Developed a centralised, automated, easy-to-use and consistent tool to facilitate this optimisation in a test-driven way.

- This tool also calculates a "trigger-on-signal" efficiency, which is more characteristic of selection performance.

- HltEfficiencyChecker is in use today helping LHCb prepare for the start of Run 3 in 2022 and will evolve into an important part of commissioning & quality assurance.

Thank you for your attention. Any questions?

Backup: Why TOS is important

- Default in Run 3 will be to persist through the TURBO stream, which persists only the trigger candidate to disk.
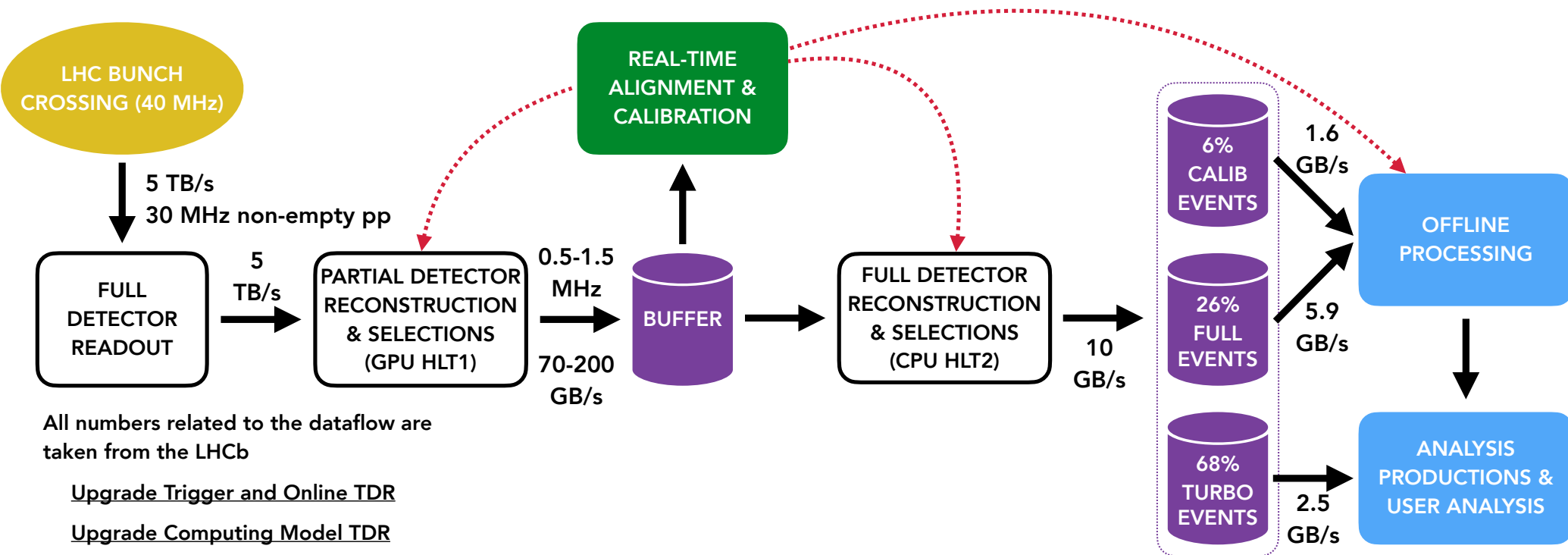
- In most analyses, the only candidates that can be reliably understood are those that are TOS.

Workflow of the tool

Get some MC

↓

Run it through Moore

↓

Save decisions and kinematics to a tuple

- Configured with a single options file.

- 2 ways to give this configuration: "wizard" and "by hand".

↓

Work out some rates/efficiencies

↓

Plot/tabulate them!

- 2 python scripts, called by the "wizard" or "by hand" via the command line.