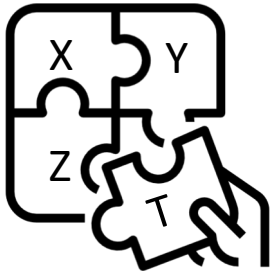
Sub-nanosecond Cherenkov photon detection for particle identification in high-occupancy conditions.

Floris Keizer, Chris Jones, Sajan Easo and Steve Wotton,

on behalf of the LHCb RICH Collaboration

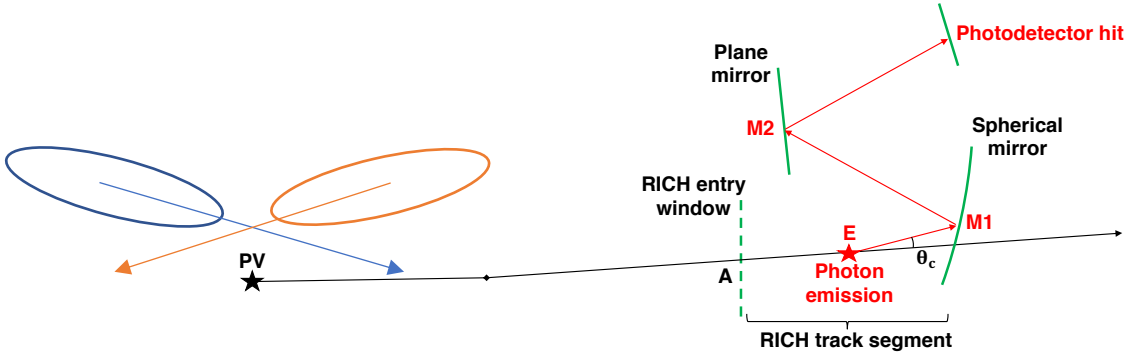
4th workshop on LHCb Upgrade II,
8-10 April 2019

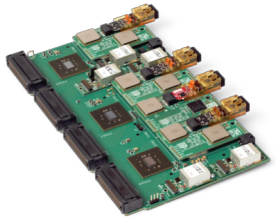
LHCb is a spatial experiment. The addition of the time coordinate is a powerful tool to reduce the sub-detector occupancies towards the HL-LHC.

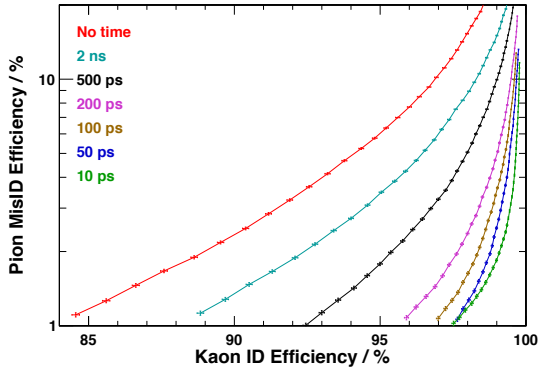


This study uses the Upgrade I simulation framework to explore the particle identification (PID) performance including the RICH photodetector hit time coordinates.

The intrinsic time resolution of the RICH detector is less than 10 ps.
How do we get to this resolution? *What* can we do with it?

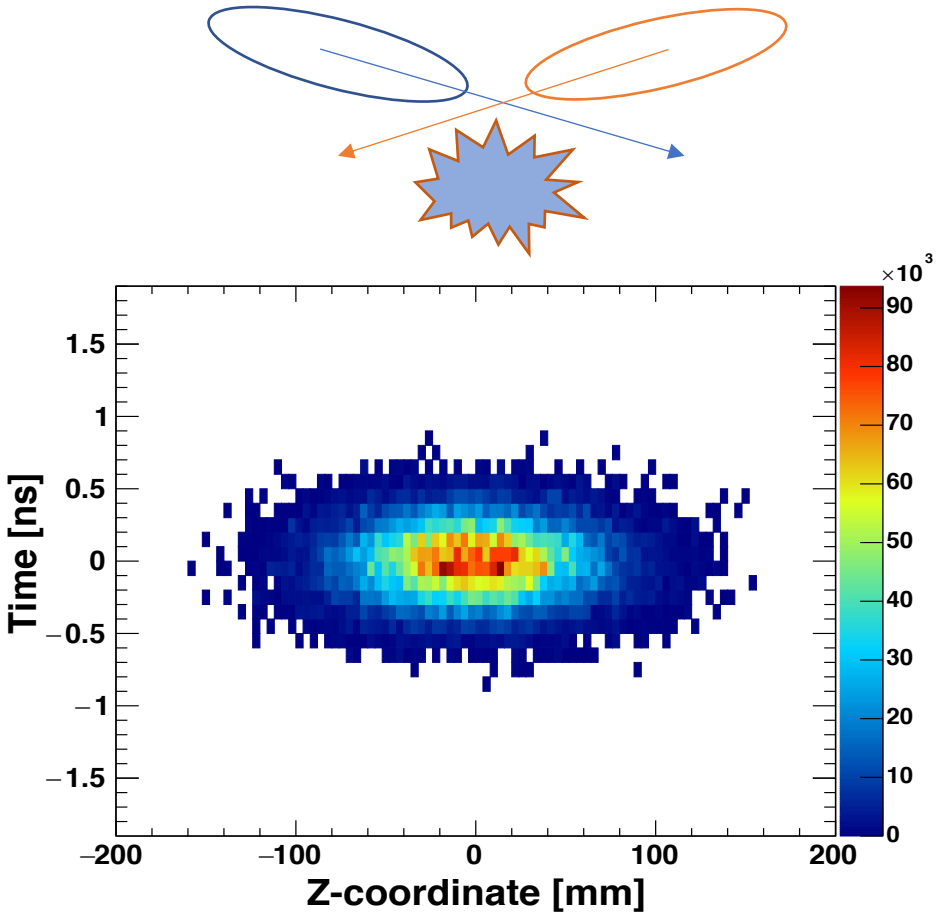






Note: quantitative results are preliminary, but the trends are clear.

The first step for time in the simulation is a new **primary vertex (PV) generation** tool.
The tool uses a Markov chain to get a 4D sample from the PDF of two colliding 3D Gaussian bunches.
More details are in my talk at the simulations meeting: https://indico.cern.ch/event/574360/ .

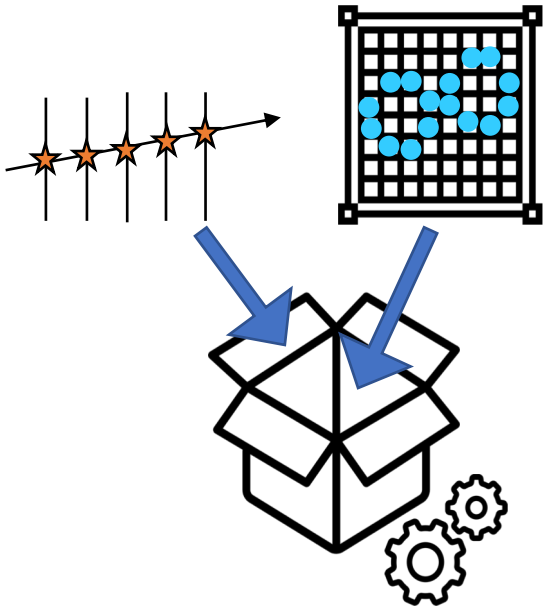


The PV distribution has a time spread of $\sigma \sim 210$ ps.

The time resolution of the RICH detector is about 2 orders of magnitude smaller than the PV time spread. The spread works to our benefit for pile-up mitigation.

The PV spread limits the size of a time gate applied at the front-end electronics to 1 to 2 ns.

The **PV time** is necessary to use sub-nanosecond time information in the RICH detector.



- ✓ Cherenkov angle
- ✓ Mirror reflection points

Tracks and photons are propagated in Gauss / Geant4.

The photodetector hits are stored in the RICH reconstruction **SmartID** with a hit time resolution of 0.2 ps. This allows us to study the finest time resolution, or to later ‘smear’ the resolution to match a realistic photodetector.

The RICH reconstruction creates one **photon object** for each RICH ‘hit + track’ combination. The decision to create a photon object includes spatial cuts.

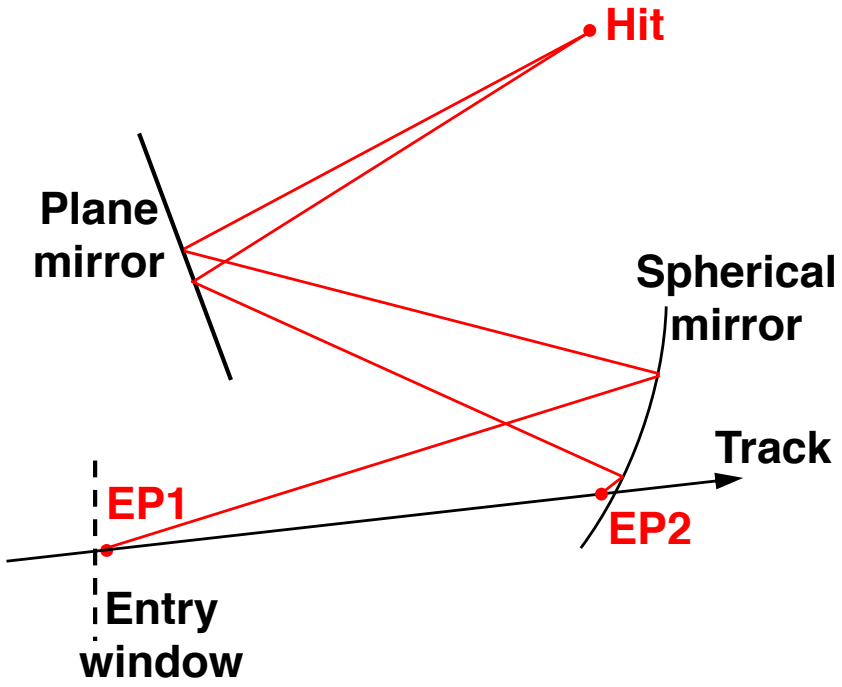
The RICH reconstruction cannot know where along the track and at what energy the photon was emitted:

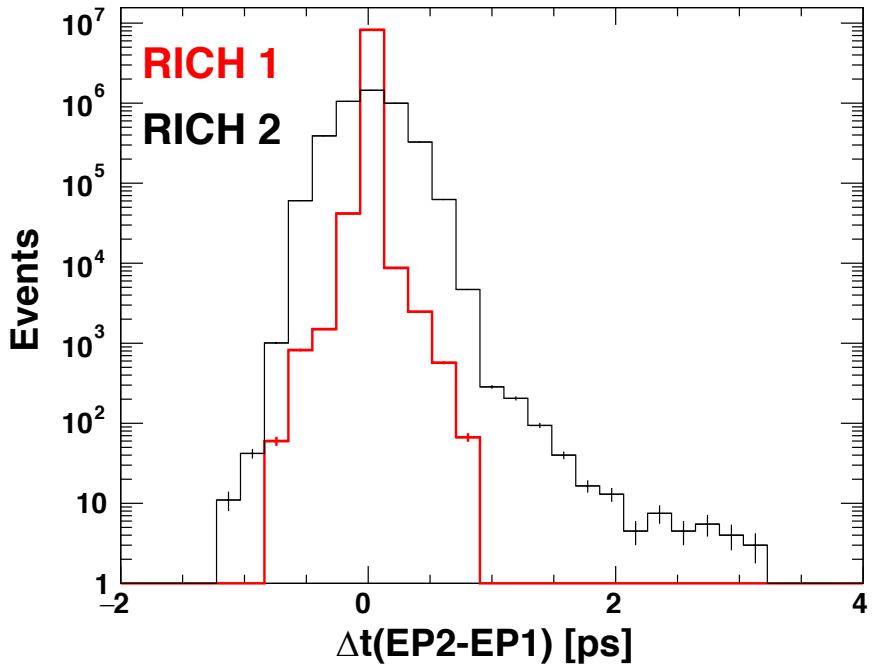
- **Emission-point error.**
- **Chromatic/dispersion error.**

The emission point error is the difference in time of arrival between Cherenkov photons emitted at the start (EP1) and at the end (EP2) of the RICH track segment.

EP1 ‘early emission’ : the total path length is typically 1.4 mm *shorter*.
EP2 ‘late emission’ : the particle track travels *faster* than the photons.

The effects largely cancel out, resulting in a **negligible emission point error in time**.





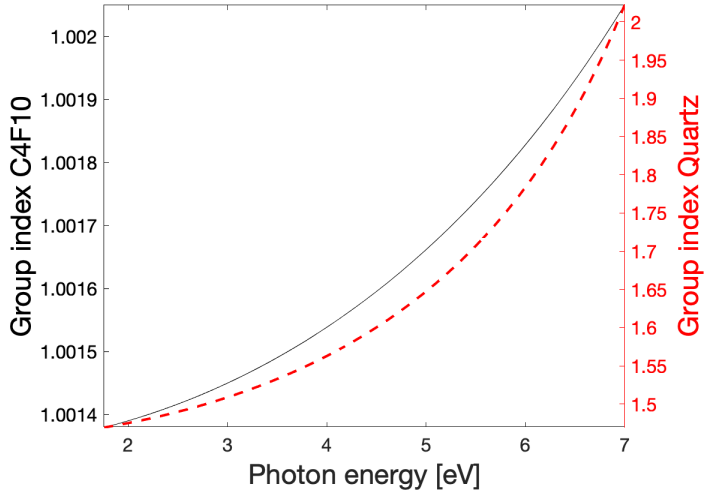
The chromatic error arises from the energy dependence of the photon group velocity.

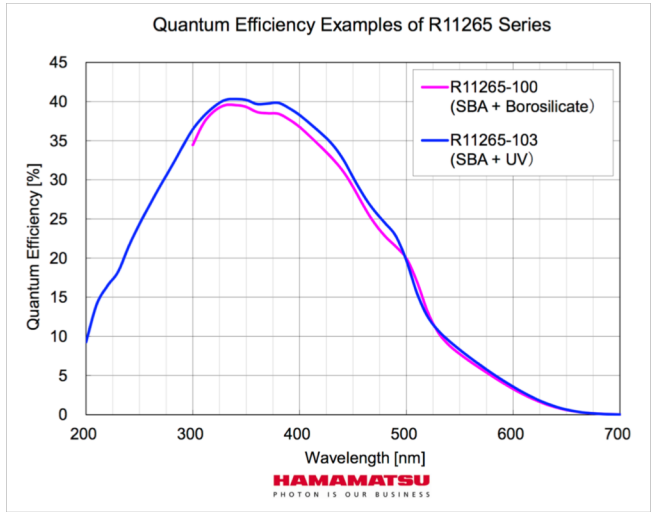
The main contributions are from the RICH gas radiators and the gas enclosure quartz window.

The chromatic error in time is on the picosecond scale.

Photodetectors with a high quantum efficiency in the longer-wavelength end of the visible spectrum can reduce the time spread due to dispersion.

Photon range	Δt : C4F10 gas	Δt : Gas enclosure quartz
180 – 710 nm / 7.0 – 1.75 eV	**5.4 ps**	**11.1 ps**
200 – 600 nm / 6.2 – 2.1 eV	**3.8 ps**	**6.9 ps**
400 – 800 nm / 3.1 – 1.55 eV	**0.7 ps**	**1.0 ps**



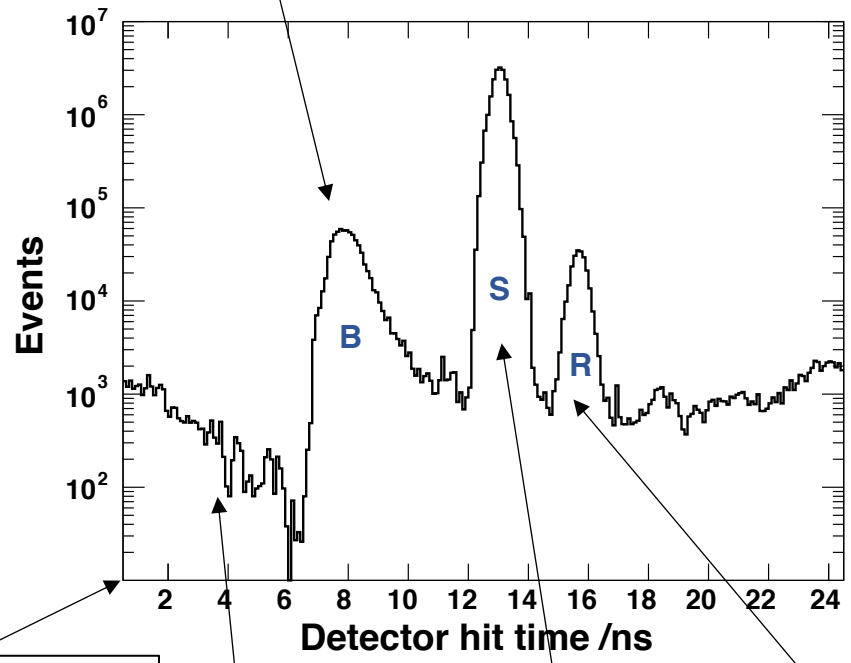


The Upgrade I simulation uses the MaPMT QE.

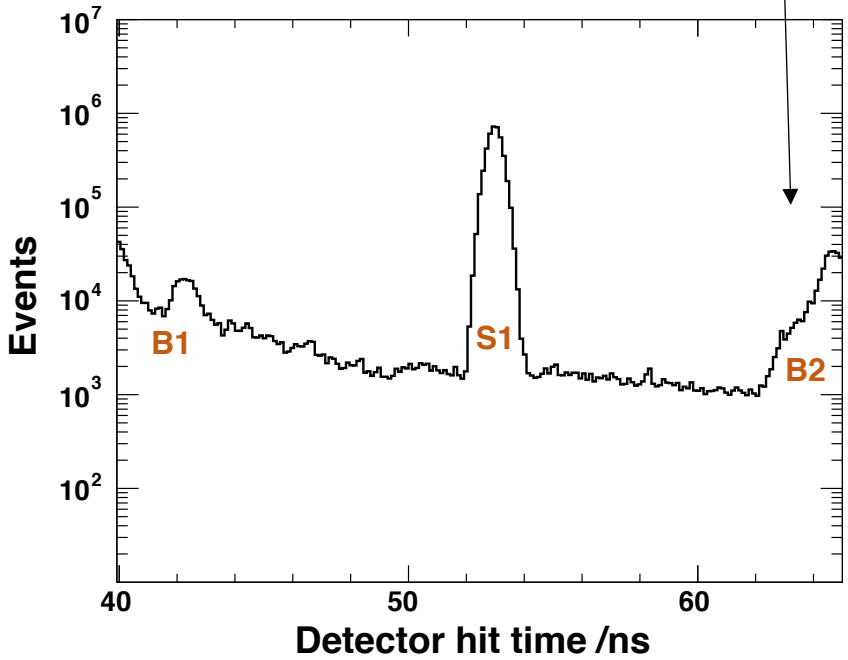
Background from tracks / photons travelling directly to the photodetector plane.

Background from the next bunch crossing.

RICH 1 photodetector hit time distribution



RICH 2 photodetector hit time distribution



Bunch crossing at $t = 0$.

Signal

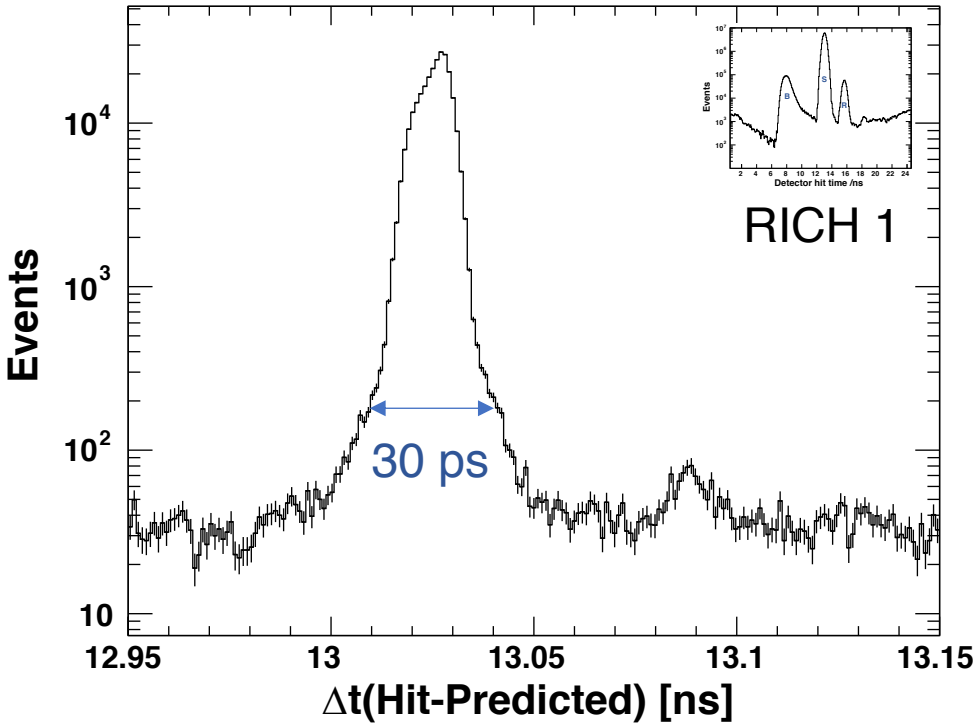
Reflected signal (treated as background).

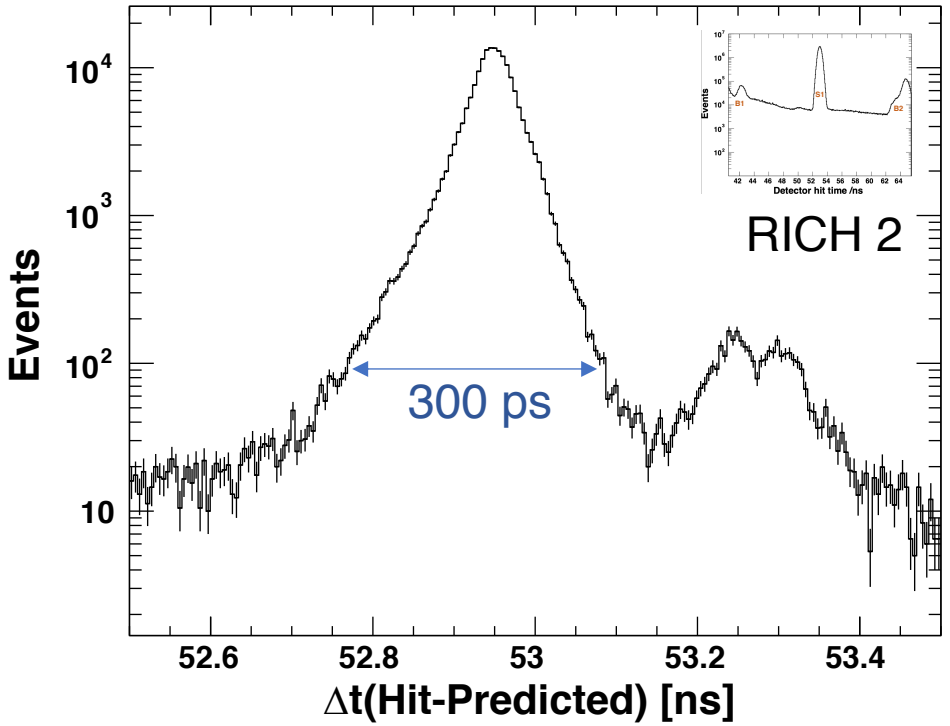
Spill-over hits from previous bunch crossing.

The signal peak widths reflect the PV spread of 1 to 2 ns.

To reveal the time spread in the RICH detectors, let us **remove the PV time spread** for now:

$$\Delta t_{\text{as plotted}} = t_{\text{hit}} - t_{\text{PV}} + \frac{z_{\text{PV}}}{c}$$



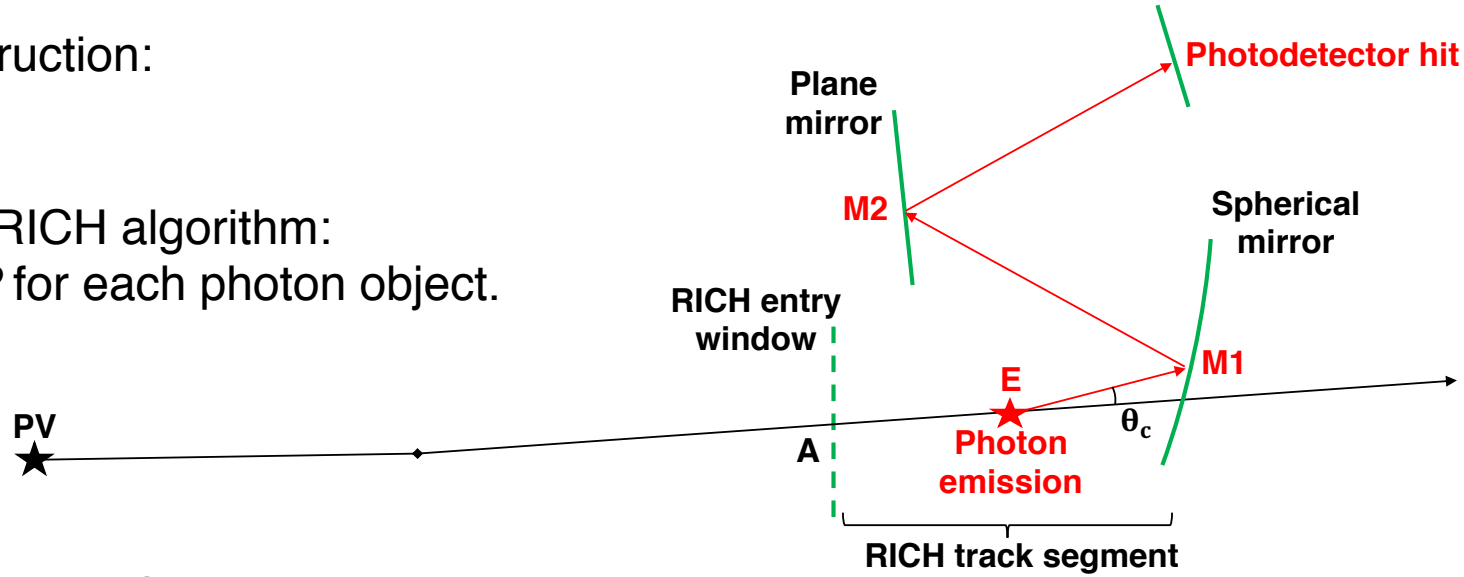


The spread is caused by the variation in track and photon trajectories and speeds.
The following slides will show that reconstructed quantities only are sufficient to further improve the prediction for the photon time of arrival. This allows more effective background rejection in reconstruction.

Input information:
- Track vector and momentum.
- Detector hits.

Input information or RICH reconstruction:
- PV time.

Reconstructed parameters in the RICH algorithm:
- Spatial points *A*, *E*, *M1* and *M2* for each photon object.
- The current mass hypothesis.



How can we construct an accurate RICH hit time prediction?
- PV spread.
- Track time-of-flight to the RICH entrance window.
- Segment time-of-flight in the gas radiator until photon emission.
- Photon time-of-flight.
- Low momentum curvature correction for RICH 2.

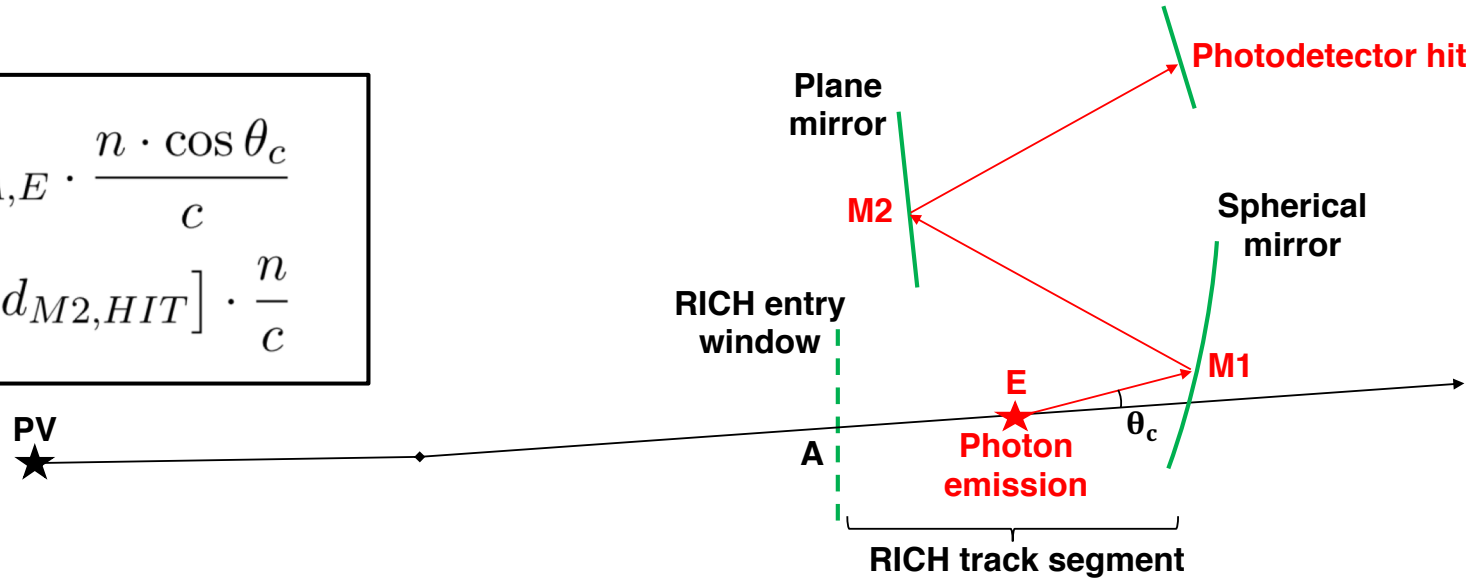
The first term is the **PV time**. It is obtained from Monte Carlo without error. The experimental error would simply propagate and spread the final hit time prediction.

The difference in predicted time choosing the **PV time or the individual track start time** is negligible. The PV timestamp is used because its resolution is $\sqrt{\mathrm{N}_{\mathrm{tracks}}}$ higher.

Analogous to TORCH, the **RICH could generate its own** PV timestamp using fast timing photodetectors. This may be useful if we install a small region of photodetectors with high time resolution in Upgrade I(b) for example.

Alternatively, a timestamp from LHCb could be used, for example from VELO or TORCH.

$$t_{pred} = t_{pv} + d_{pv,A} \cdot \frac{\sqrt{p^2 + m^2}}{p \cdot c} + d_{A,E} \cdot \frac{n \cdot \cos\theta_c}{c} + \left[d_{E,M1} + d_{M1,M2} + d_{M2,HIT} \right] \cdot \frac{n}{c}$$

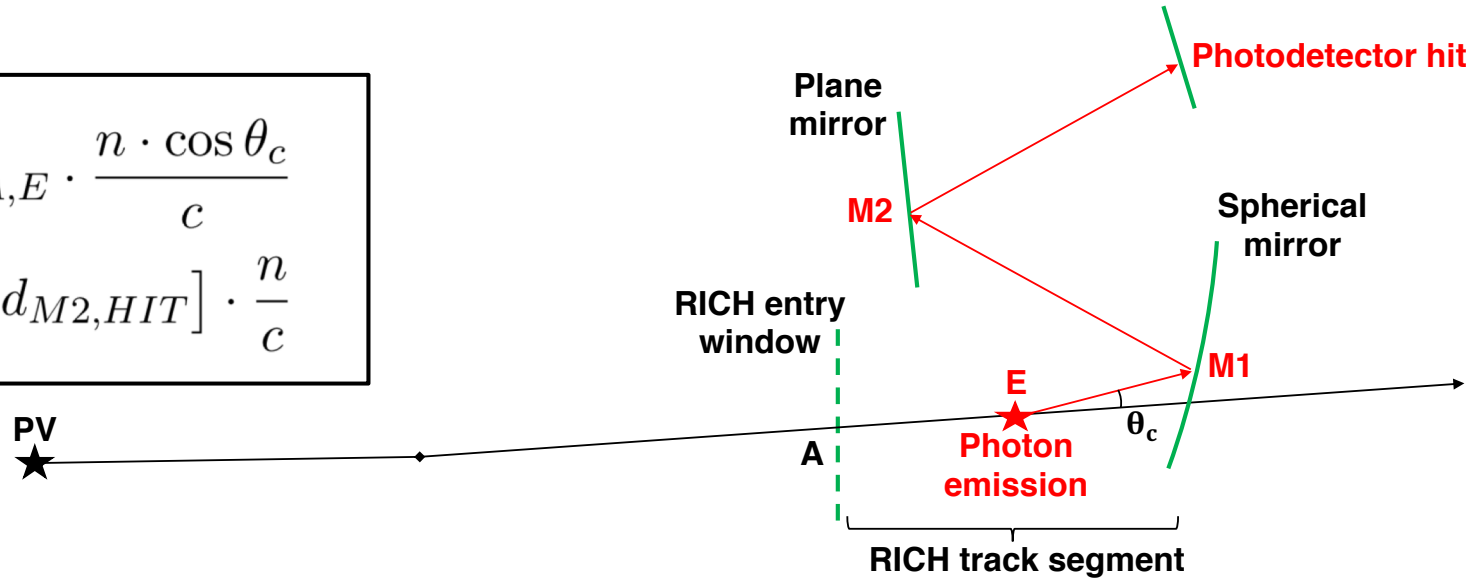


The second term is the **track time-of-flight**.

The RICH entrance point (A) and the track momentum are obtained from the LHCb tracking information.

In the reconstruction, the mass hypotheses for the tracks are changed in order to maximise the overall likelihood. The mass used in the prediction of the track time-of-flight is updated with each change in the mass hypothesis.

$$t_{pred} = t_{pv} + d_{pv,A} \cdot \frac{\sqrt{p^2 + m^2}}{p \cdot c} + d_{A,E} \cdot \frac{n \cdot \cos\theta_c}{c} + \left[d_{E,M1} + d_{M1,M2} + d_{M2,HIT} \right] \cdot \frac{n}{c}$$

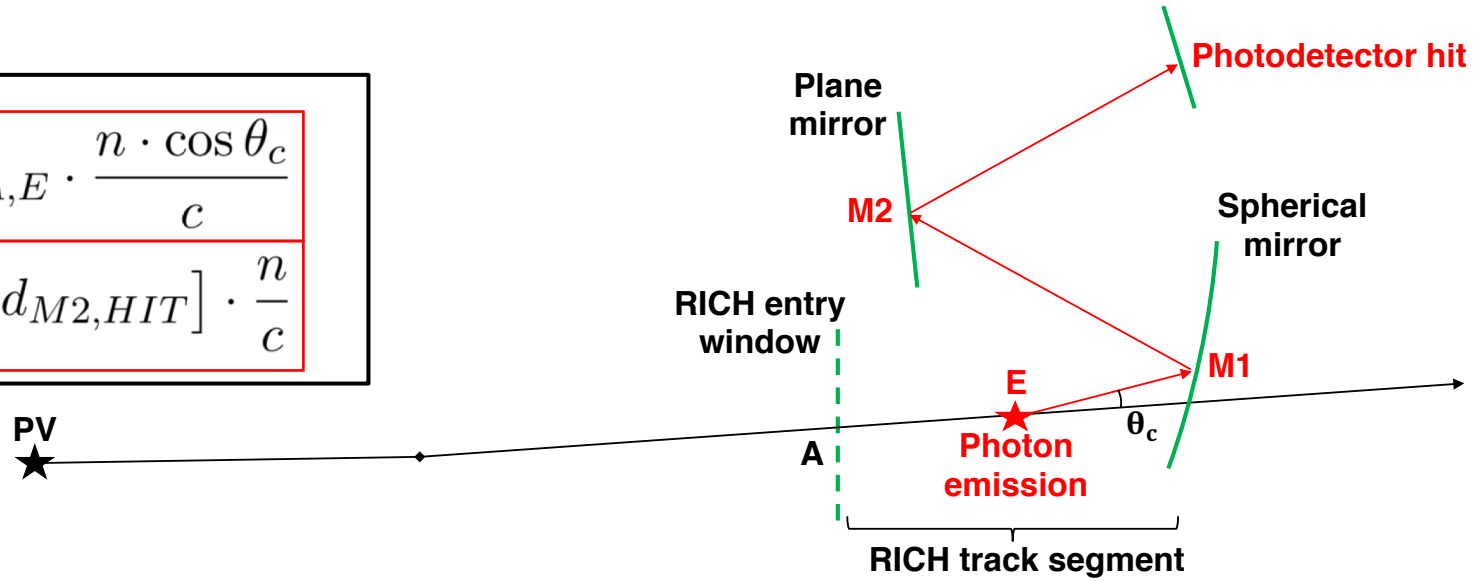


The third and fourth terms are the **RICH track segment and photon time-of-flight** in the gas radiator.

The prediction is based on the reconstructed Cherenkov angle, emission point (E) and the mirror reflection points (M1) and (M2).

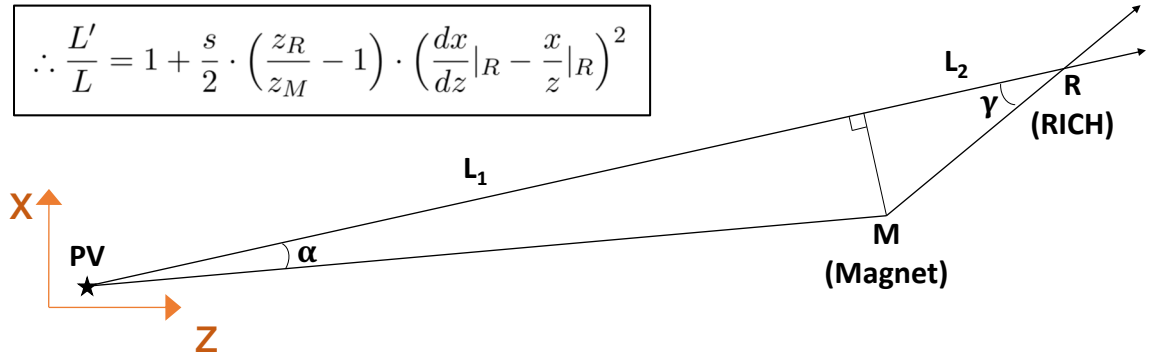
The photon speed is calculated using the (fixed) average refractive index.

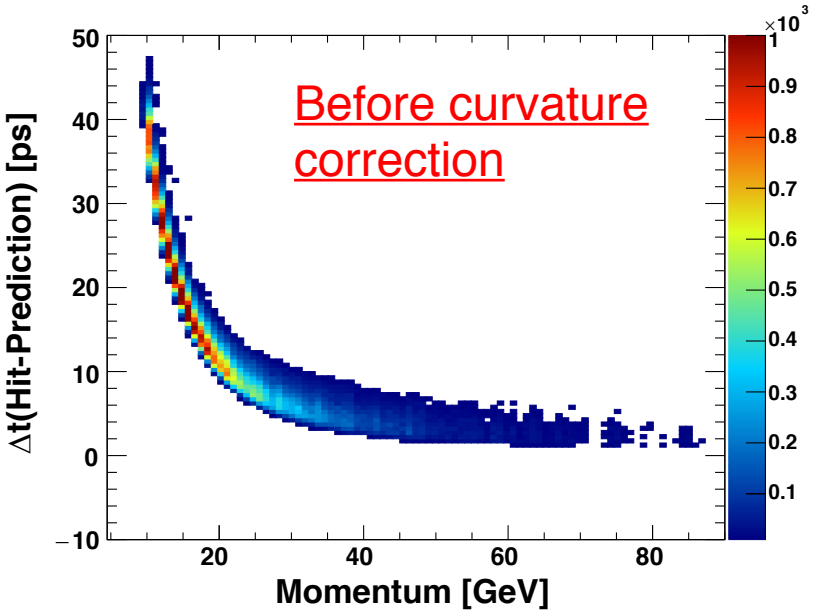
$$t_{pred} = t_{pv} + d_{pv,A} \cdot \frac{\sqrt{p^2 + m^2}}{p \cdot c} + d_{A,E} \cdot \frac{n \cdot \cos\theta_c}{c} + \left[d_{E,M1} + d_{M1,M2} + d_{M2,HIT} \right] \cdot \frac{n}{c}$$

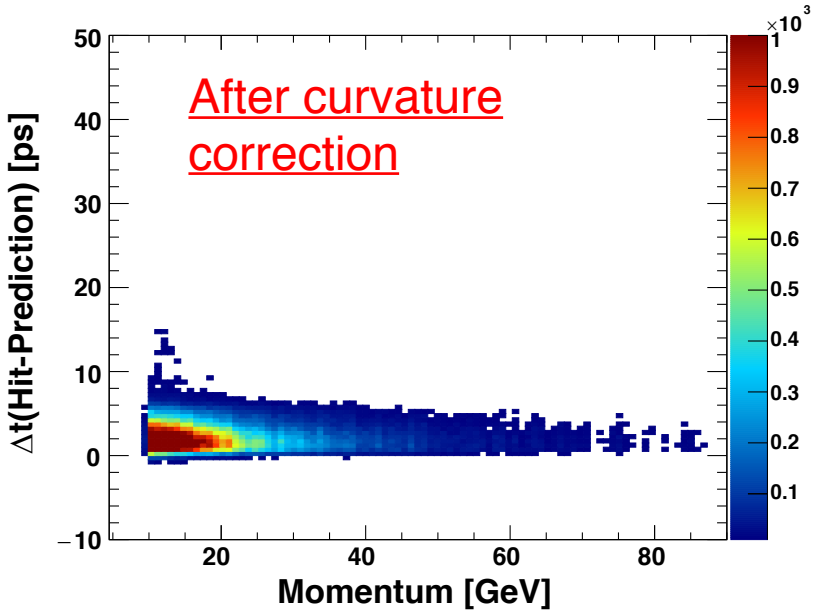


The **curvature of low momentum tracks** causes an error in the predicted time for RICH 2. Therefore, we apply a correction, inspired by the Outer Tracker (OT), which calculated the change in the track path length assuming a **"kick" from the LHCb magnet**.

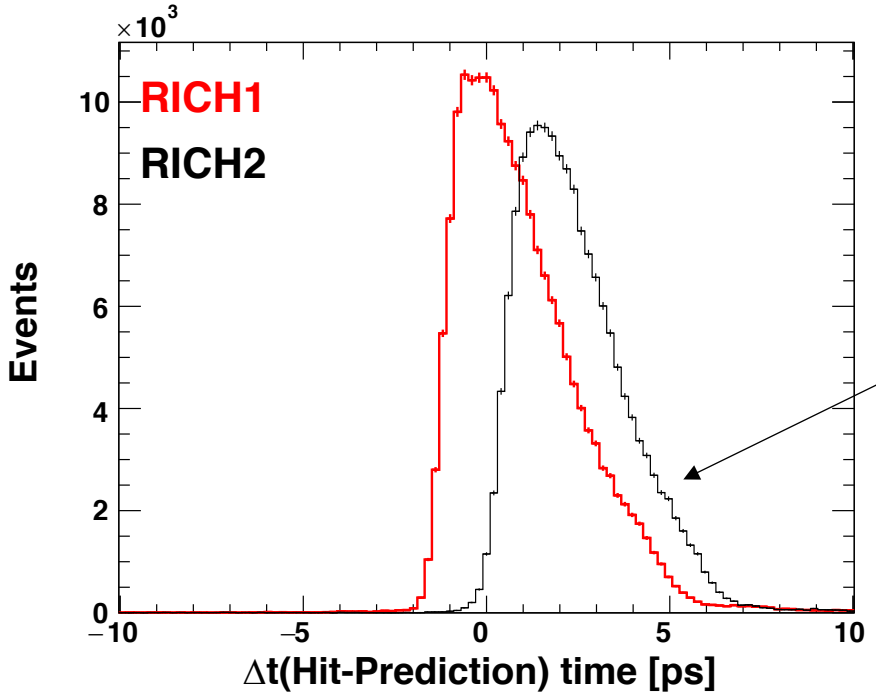
$$\therefore \frac{L'}{L} = 1 + \frac{s}{2} \cdot \left(\frac{z_R}{z_M} - 1\right) \cdot \left(\frac{dx}{dz}|_R - \frac{x}{z}|_R\right)^2$$







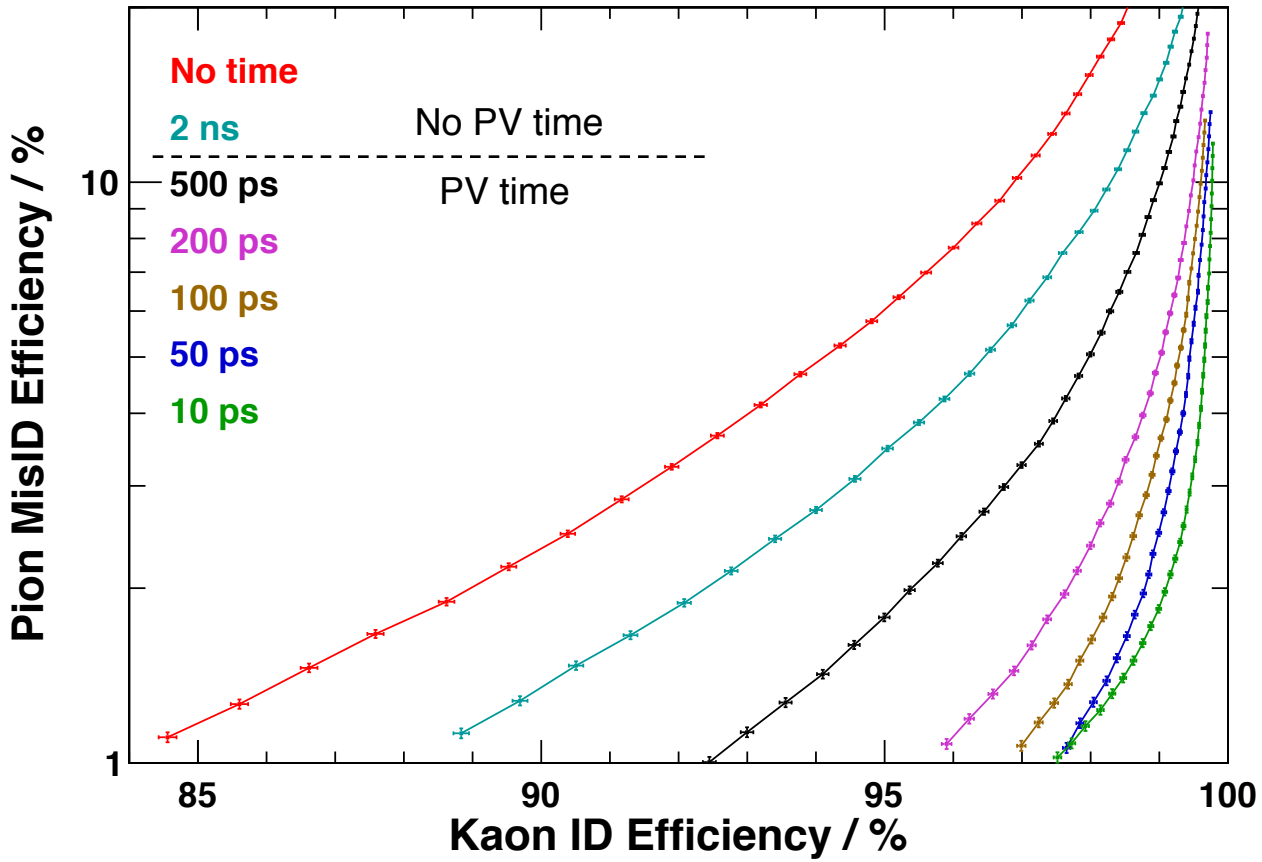
In summary:
*The prompt Cherenkov radiation and mirror geometry allow us to predict the RICH photon hit time to within **10 picoseconds** based on reconstructed parameters.*



The tail of the distribution arises from dispersion (hits arrive *later* than predicted), and would be even smaller for a green-shifted photosensor.

The high intrinsic time resolution is a powerful tool in high-occupancy conditions. It can be exploited by imposing a time gate for background rejection or feeding the information into the likelihood maximisation algorithm.

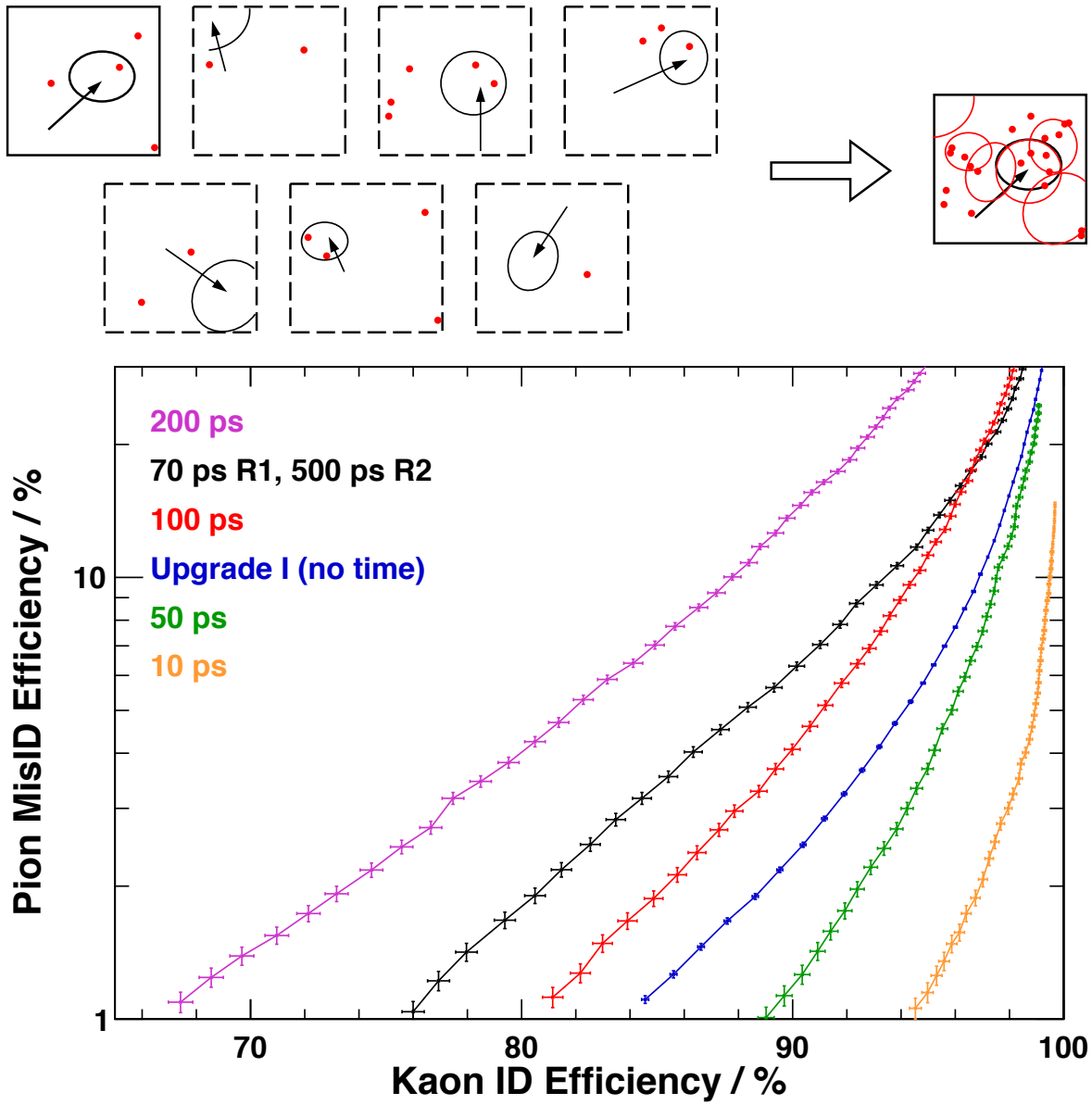
The curves show the **Upgrade I** particle identification efficiency for different time gates applied in the RICH reconstruction.



The results motivate “the faster the better” for the detector technology.

The final overall time resolution including the photodetector technology and PV time dictates which PID curve we can achieve.

The effect of pile-up in the HL-LHC.



The RICH hits from seven Upgrade I events are overlaid into one **Upgrade II** event. The MC information and tracking are available for one event. The six other events are piled up as background.

This is the best approximation to the HL-LHC (Luminosity increases by a factor 7.5 from $2\times10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ to $1.5\times10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$).

A gate between 50 and 100 ps can recover the Upgrade I performance in the HL-LHC.

The performance can even exceed Upgrade I using a hypothetical time resolution of 10 ps.

The 70 ps in RICH 1 + 500 ps in RICH 2 are gates around the PV time. These do not require RICH reconstructed parameters and can pre-select photon objects.

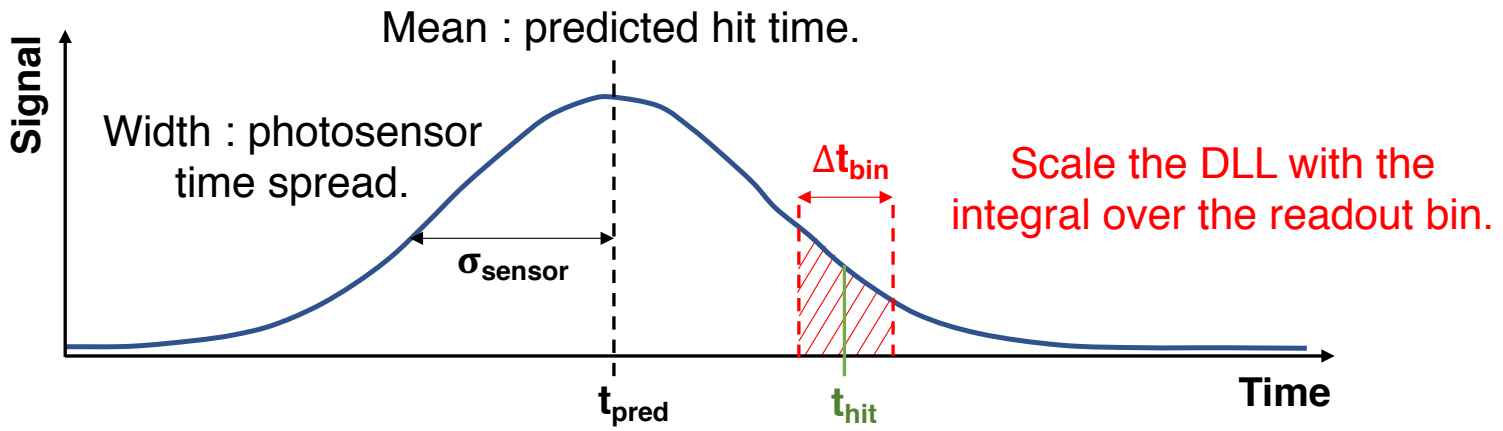
In addition to gating, the predicted photon time of arrival can be used to **scale the signal PDF in the Delta-Log-Likelihood (DLL) calculation** in the reconstruction.
The PDF of the signal photons is a Gaussian around the expected Cherenkov angle θ_c for particle type h_j:

$$f_{h_j}(\theta,\phi) = \frac{1}{(2\pi)^{3/2}\sigma(\theta)} \exp\left[-\frac{1}{2}\left(\frac{\theta-\theta_c(h_j)}{\sigma(\theta)}\right)^2\right]$$

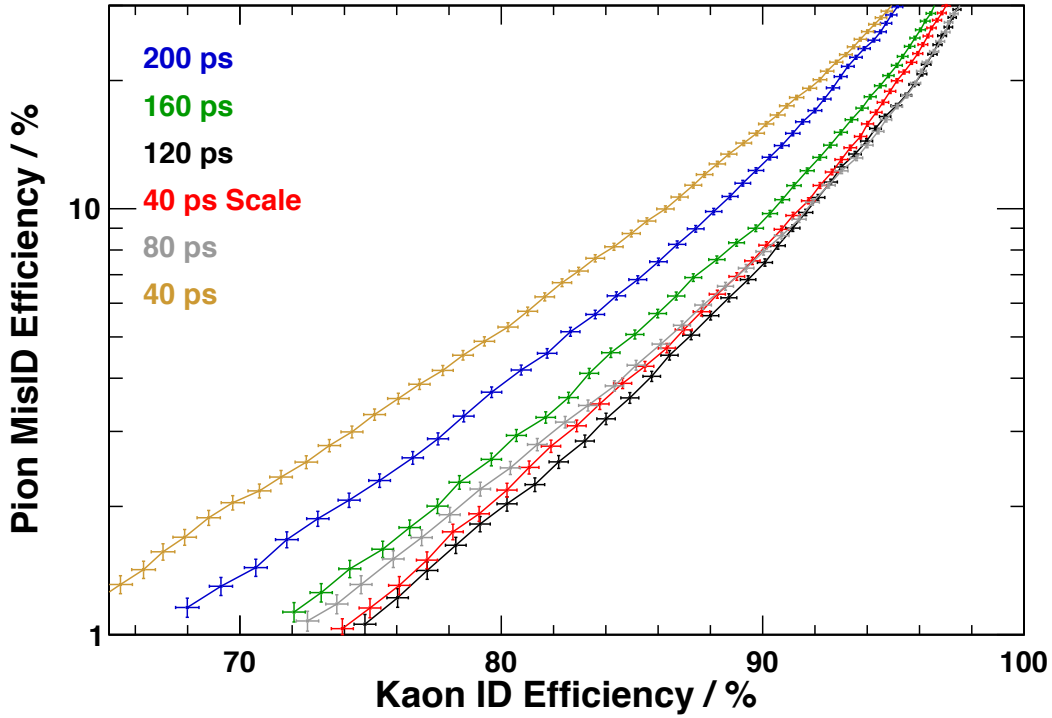
[LHCB/98-040 equation 18]

It assumes that the pixels are small compared to the Cherenkov ring.

In extending the PDF to the time domain, we cannot assume that the readout electronics bin size is much smaller than the photosensor time spread. Instead, the scaling takes into account the photosensor time spread, σ_{sensor}, and the electronic readout bin width, Δt_{bin}.



Take $\sigma_{sensor} = \mathbf{40\ ps}$,
for an MCP-PMT for example.
The hit times are smeared by a Gaussian.

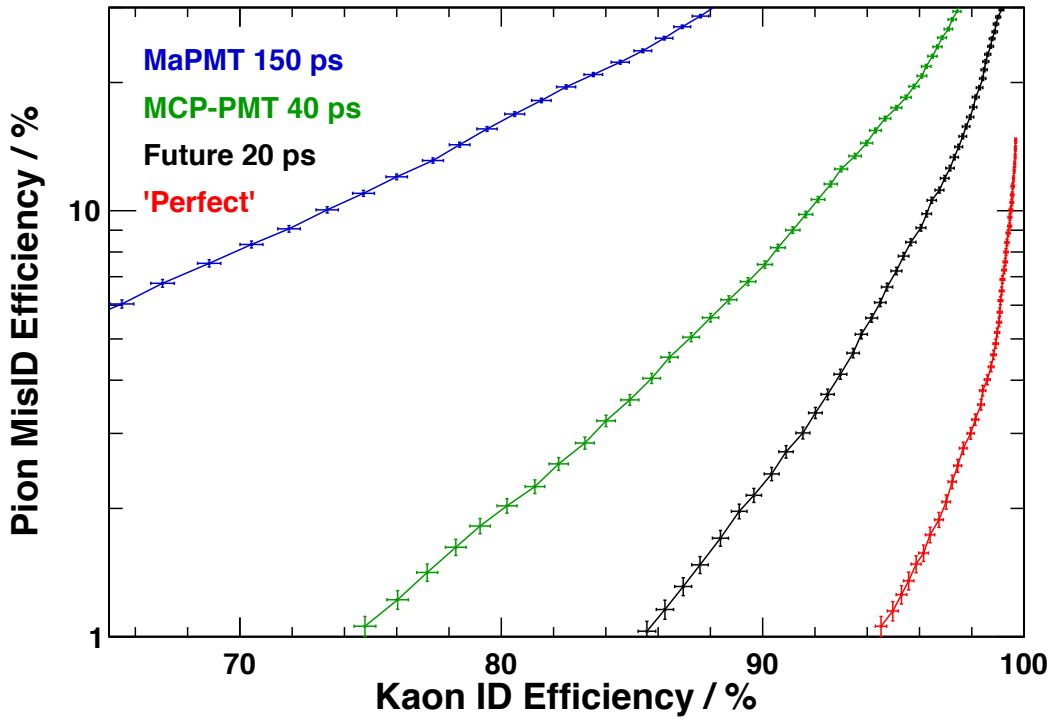


The plot shows the PID performance for a range of time *gates* from 200 ps to 40 ps. The red curve shows the *scaling* method using a 40 ps readout bin.

The optimal time gate is around 120 ps for the 40 ps sensor. The DLL scaling method performs similar to this optimised time gate.

Different photodetector time resolutions are compared on the next slide, where I apply a time gate of 3 σ_{sensor}.

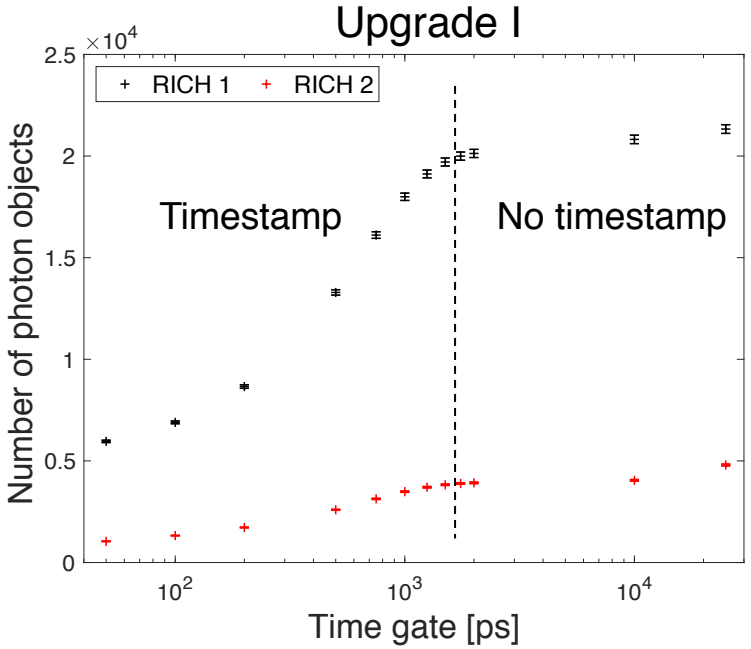
The PID performance for different photodetector time resolutions.

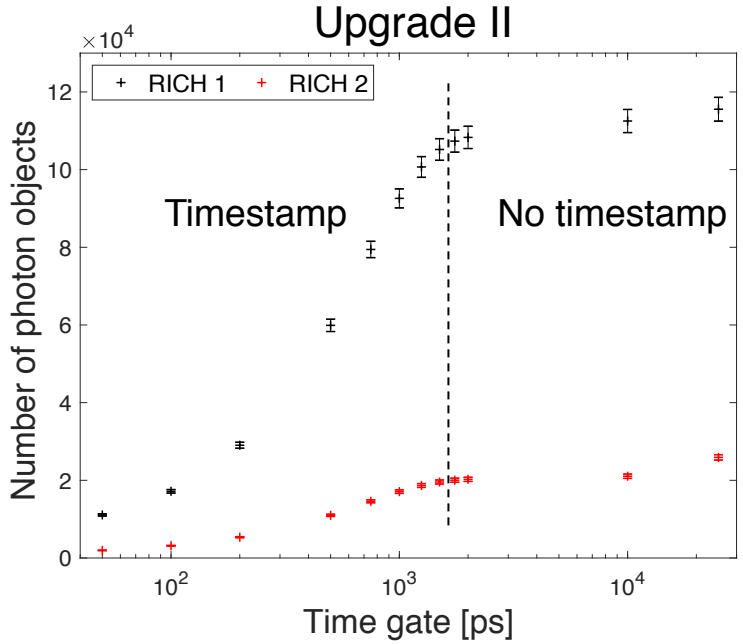


The trend indicates that a 20 ps sensor with 60 ps readout electronics can match the Upgrade I performance.

The improvements in PID are from time resolution only. We intend to also study the combined effect of the photodetector time resolution and spatial granularity for example.

The impact on reconstruction resources.





Reduction in background photon objects using a 20 ps time gate.

	Upgrade I	Upgrade II
RICH 1	72 %	93 %
RICH 2	91 %	97 %

One **photon object** is created for each 'hit + track' combination.
A cut on the number of background photon objects significantly enhances the reconstruction speed.

The number of photon objects decreases most for time gates between 1 ns and 100 ps.

- Upgrade I : about 72 % of the background photon objects is removed by a 20 ps time gate. The remainder is combinatorial background of tracks and photons arriving from the same PV.
- Upgrade II: the background photon objects are reduced by 93%, owing to the PV time spread.
- The RICH 2 detector has larger dimensions allowing a better background reduction.

To conclude,

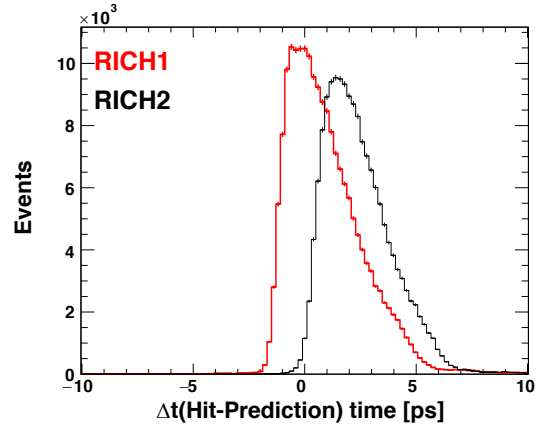
The RICH detector has a high intrinsic time resolution of less than 10 ps.

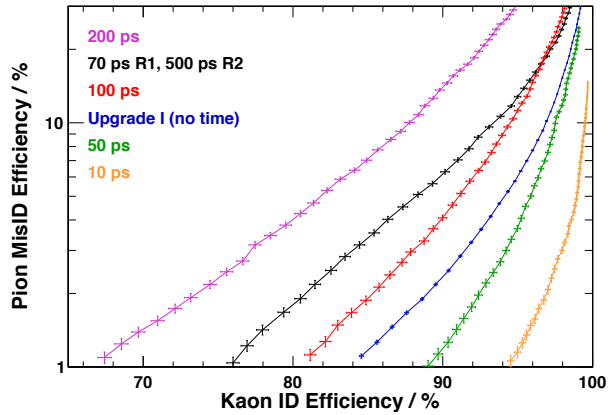
Timing is a powerful tool for background rejection towards the HL-LHC.

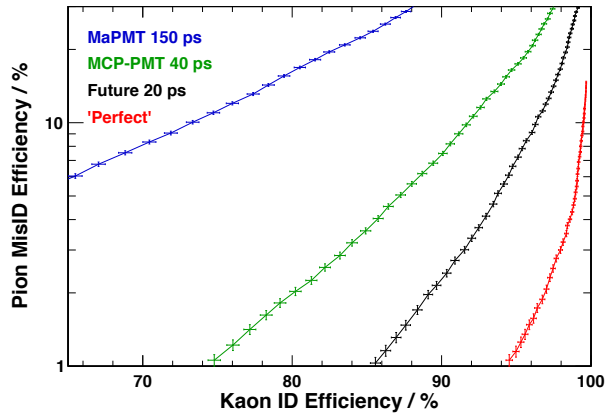
The reconstruction benefits from timing in both PID performance and reconstruction resources.

The PID performance will ultimately be limited by the photodetector technology. The higher the time resolution, the better the performance.

Thank you for your time!







Backup

The first step for time in the simulation is a new **primary vertex (PV) generation** tool:

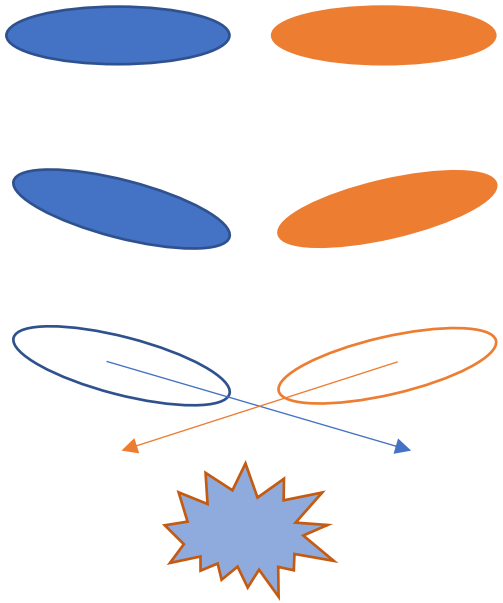
1) Take two Gaussian bunches.
   (write down the spatial Probability Density Function, PDF)

2) Rotate each bunch about the bunch crossing angles.
   (rotate the covariance matrices)

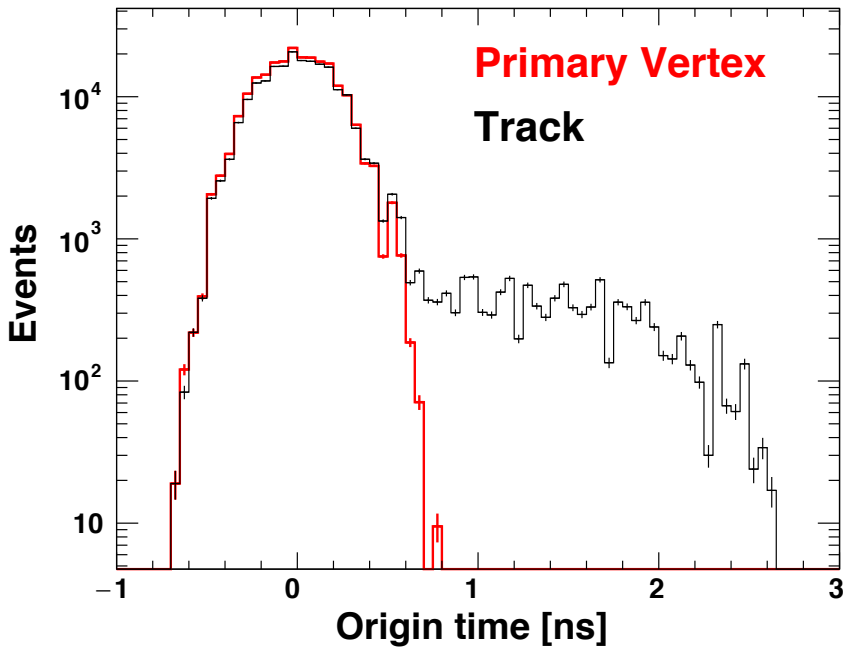
3) Set the centre of mass of each bunch into motion.
   (this step introduces time)

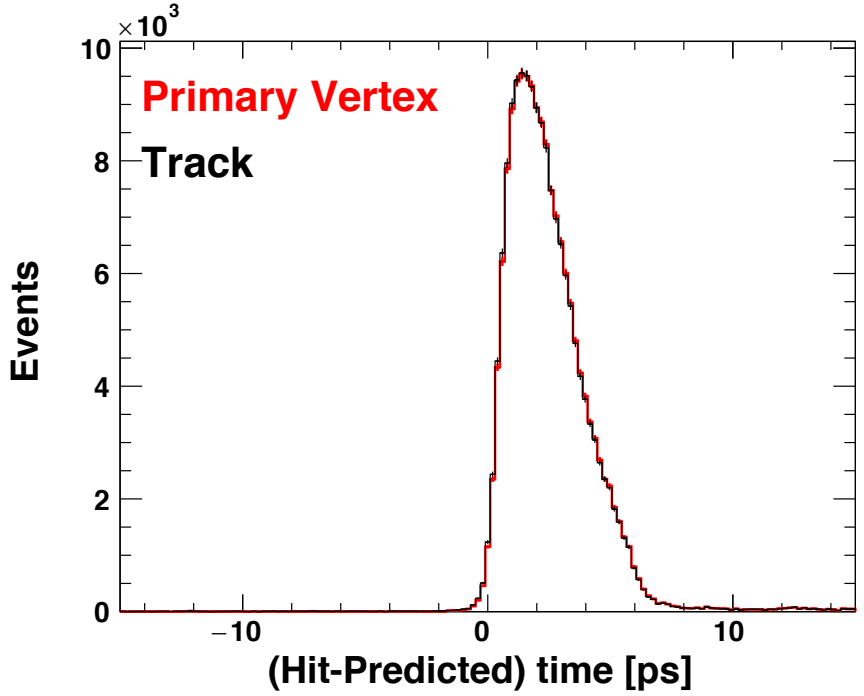
4) Collide the bunches.
   (take the product of the PDFs)

5) Sample from the 4-dimensional distribution of primary vertices.
   (using a Markov chain)



The difference in predicted time choosing the **PV time or the individual track start time** is negligible.
The PV timestamp is used because its resolution is $\sqrt{N_{tracks}}$ higher.



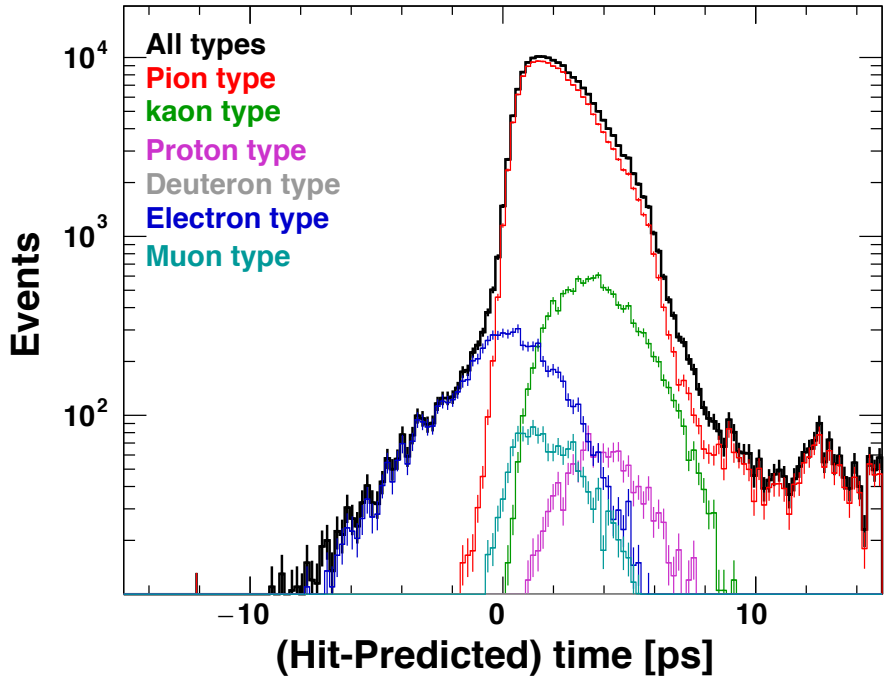


Using the **mass in the DLL** requires the track path and the predicted time excluding the track to be separately stored in the photon object.

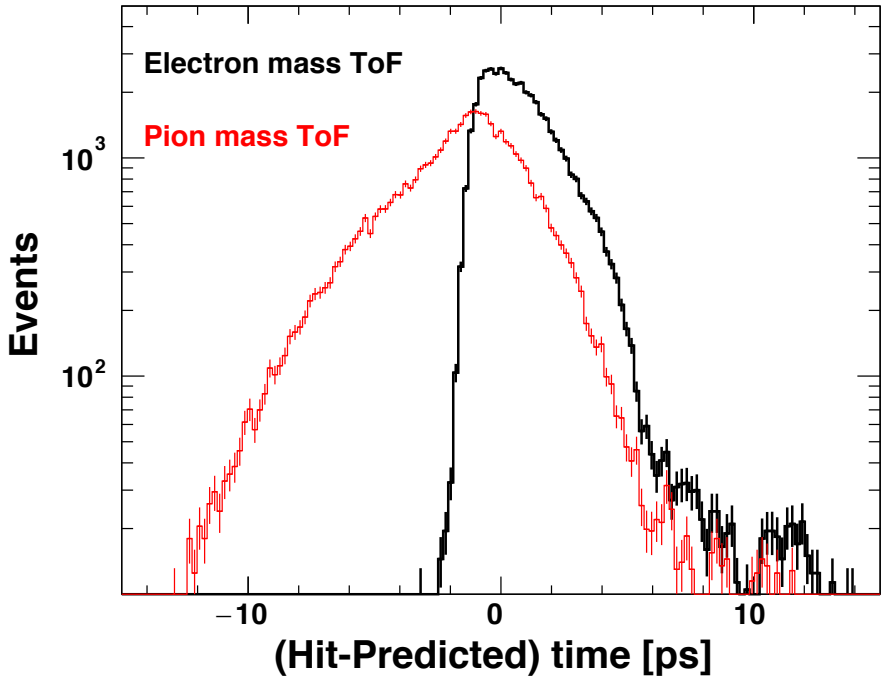
- Adds to the computational demand.
- Increases the time resolution from 20 ps to 10 ps.
- Only advisable if the photodetector resolution is small.

$$t_{pred} = t_{pv} + \boxed{d_{pv,A} \cdot \frac{\sqrt{p^2 + m^2}}{p \cdot c}} + d_{A,E} \cdot \frac{n \cdot \cos\theta_c}{c} + \left[d_{E,M1} + d_{M1,M2} + d_{M2,HIT}\right] \cdot \frac{n}{c}$$

A prediction based on the **pion mass** for the track ToF is accurate up to 20 ps.

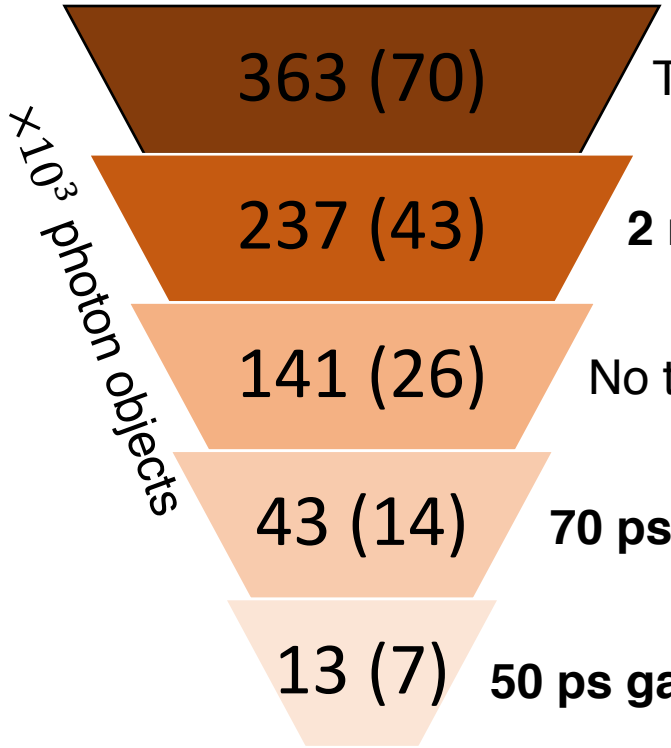


The effect is largest for the **electron type**. Using the correct mass in the DLL, the resolution of less than 10 ps is recovered.



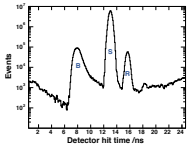
The number of photon objects is an important parameter for the **speed of the reconstruction**. The advantage of a cut is largest if it can be made **prior to** the full reconstruction of the photon path and the calculation of the photon time of arrival. These cuts are marked by ✓

Upgrade II (Upgrade I)



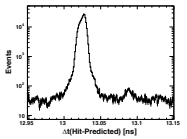
The initial number of photon objects within spatial constraints.

2 ns gate. This can be done in hardware. ✓

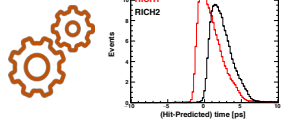


No time gate but full reconstruction spatial cut. *Current baseline.*

70 ps (RICH 1) and 500 ps (RICH 2) gates around the PV timestamp. ✓



50 ps gate and full reconstruction spatial cut.

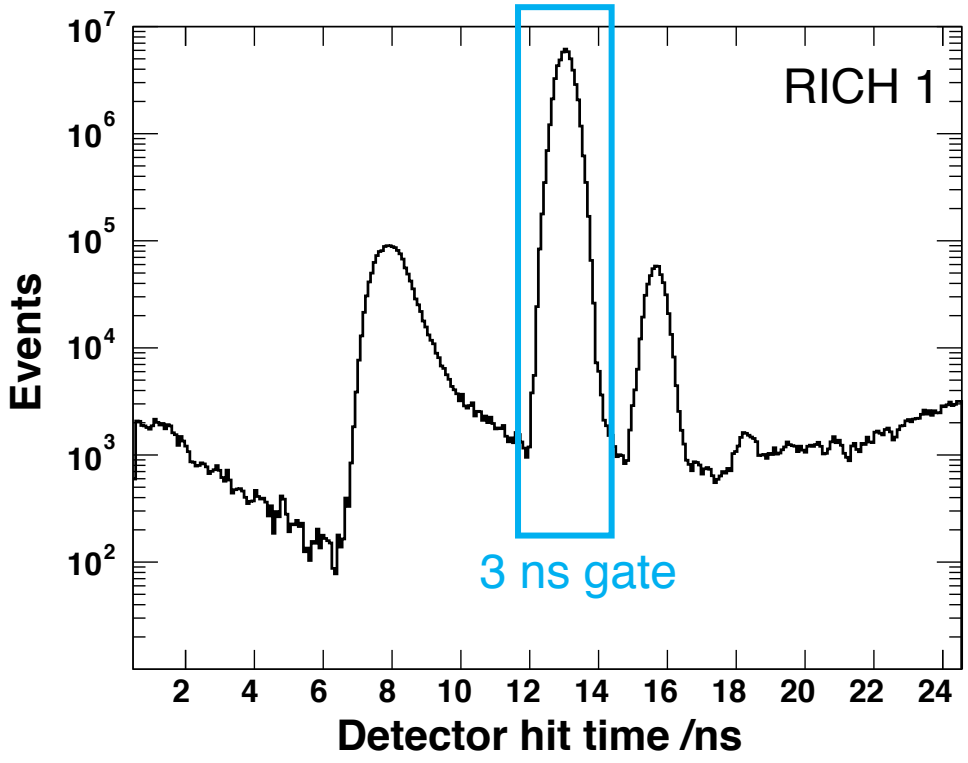


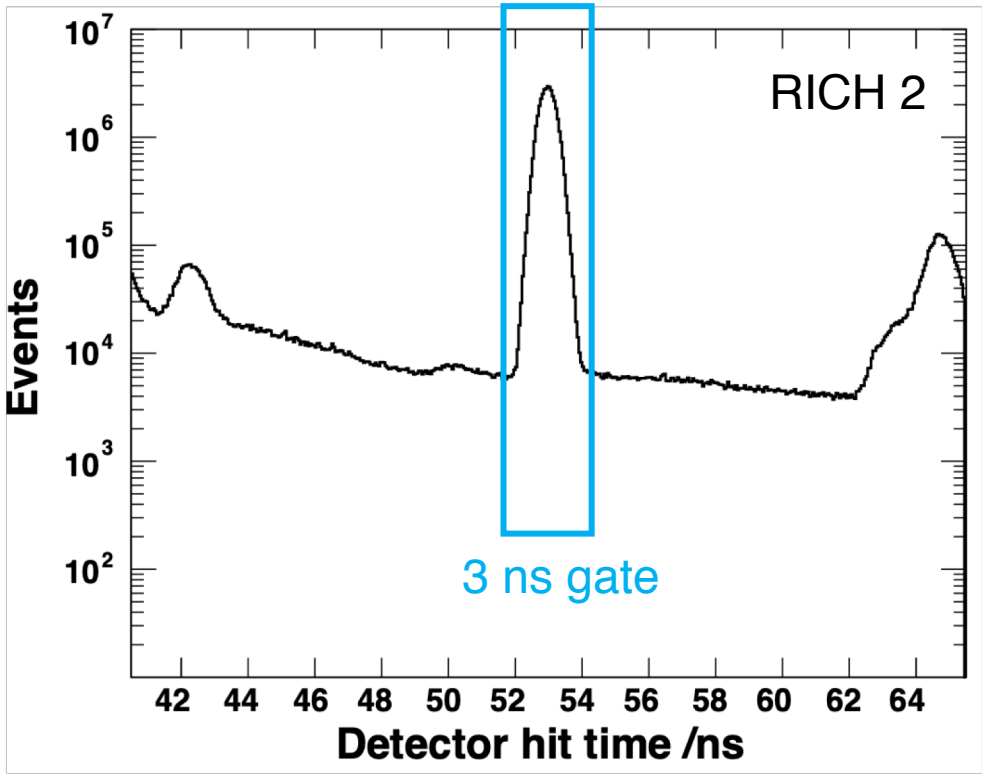
Possible strategy:

- Improve speed: Apply hardware gate + gate around the PV timestamp. ✓
- Further improve PID: Predict the photon time of arrival and apply high-resolution gate.

Time is already being implemented in the RICH Upgrade I hardware.

- A “super-latching” readout scheme applies a 3 to 6 ns time gate in the flexible FPGA-based digital readout board.
- The scheme has been tested in the October 2018 beam tests of the Upgrade I system (https://indico.cern.ch/event/774892/).
- The main motivation is to reduce sensor noise (“SIN”).
- Additionally, expect to remove around 5-10% of background photons.





The implemented **deserialisation in the FPGA** is limited by the high-speed clock.

- The method can be applied up to around 1 ns gating.

To use more precise time information we have to move towards a picosecond-capable TDC in a custom ASIC or implemented in an FPGA (using e.g. a tapped delay line).

- Since we are still dealing with the PV spread at the front-end, the amount of data significantly increases due to picosecond readout over a few nanoseconds capturing period.
- Possibilities for zero-suppression, clustering in time, etc. will be explored.