pfRICH event reconstruction considerations

Alexander Kiselev

ePIC Collaboration meeting, Argonne, January 11, 2024

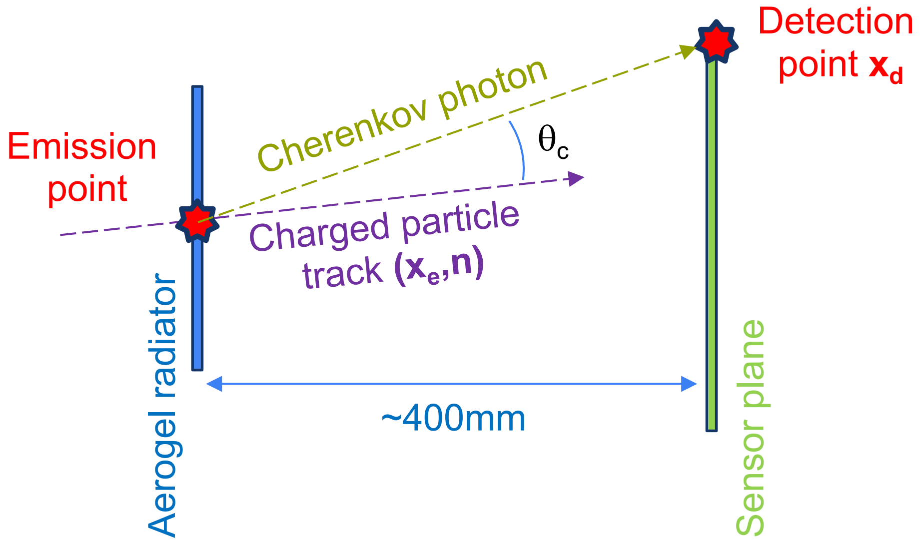
Disclaimer

- Talk will be focused on ring imaging rather than a “photon flash in the HRPPD window” timing
- Most part of the described algorithms was actually implemented for the March 2023 review as IRT 2.0
  - Will be indicated in case of the opposite

- Existing codes are algorithmic, combinatorial, χ^2 based
- This was all done in a standalone GEANT4 environment
- Porting to dd4hep is in the geometry description stage

Starting point

- Single track (and ideal tracking)
- Very thin radiator
  - No emission point uncertainty
- Single photon with a known λ
  - And therefore, a known $n(\lambda)$
- Ideal sensor plane
  - σ_{xy} ~ 0 (no detection point uncertainty)
- No mirrors & no refractive boundaries
  - Photon trajectory is a straight line
- No noise hits



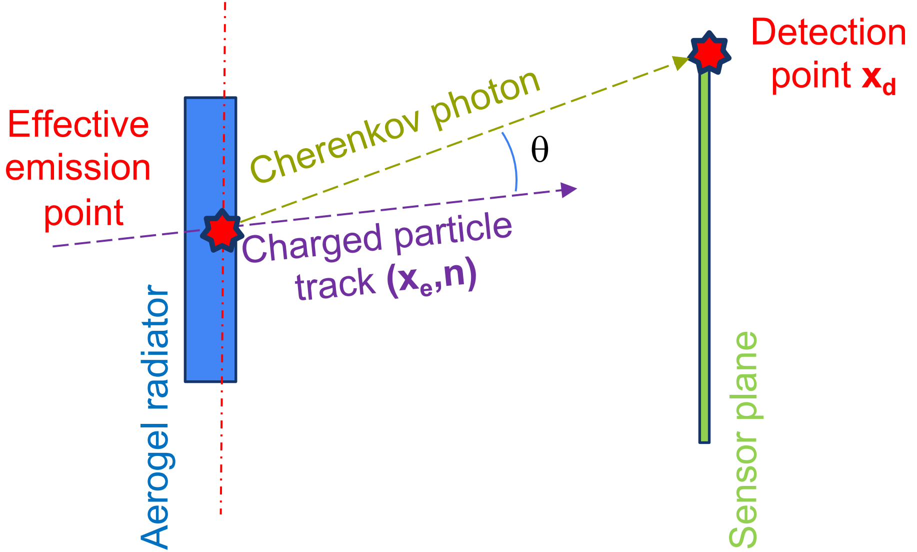
$$\cos\,\theta_c = \frac{\vec{n}(\vec{x}_d - \vec{x}_e)}{||\vec{x}_d - \vec{x}_e||} = \frac{1}{\beta n}$$

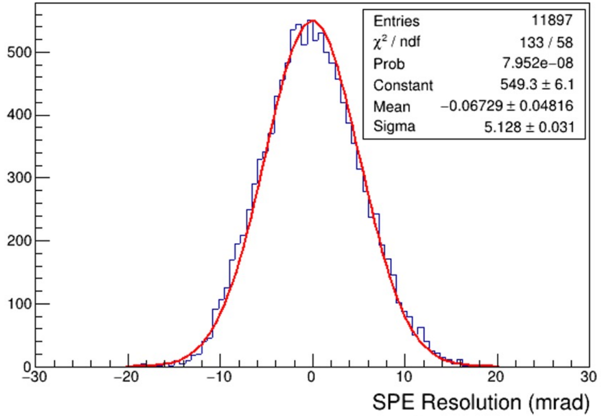
$$m = \frac{p}{c\beta\gamma}$$

➔ Substitute momentum & velocity in the formula and check PDG for this mass

Unrecoverable errors

- Emission point uncertainty
  - Aerogel radiator has a finite thickness
  - We have no means to measure the actual emission point of each photon
  - Wavelength-dependent absorbtion in the radiator
- Detection point uncertainty
  - Sensor plane has a finite resolution
- Chromatic effects
  - We do not know wavelengths on per-photon basis
  - Have to deal with <λ> and RMS
  - There is a n(λ) dependency …
  - … which translates into a spread of expected emission angles for a fixed particle momentum

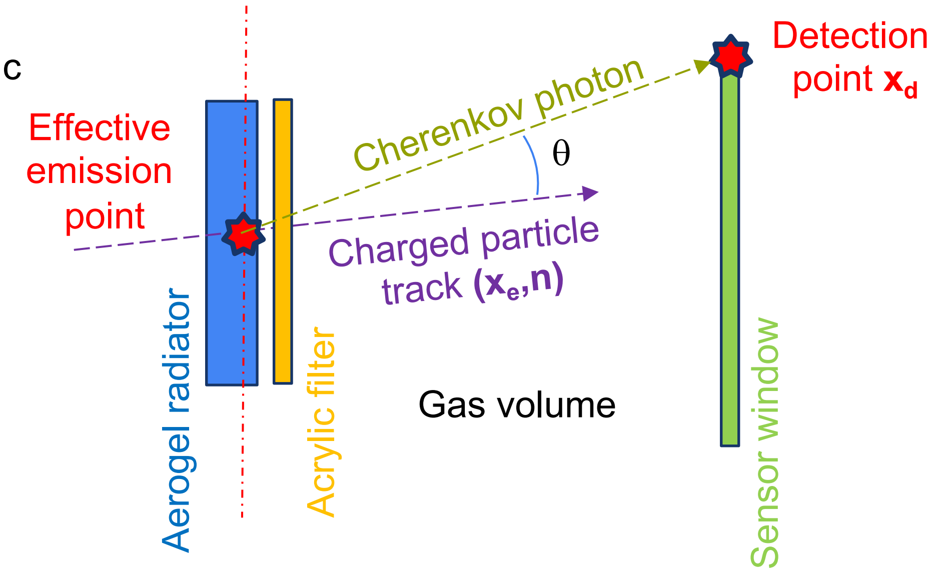




σ_{SPE} ~ 5 mrad

Refraction on optical media boundaries

- Photon path is not really a straight line
  - A sequence of *refractions* on aerogel, gas, acrylic filter and HRPPD window boundaries
- Taken care of by the IRT algorithm
  - No analytic solution exists in a generic case
  - Iterations (a 2D Newton-Gauss method)
  - Same technique used for optical paths involving *reflections* on mirror surfaces
  - At the end of the day what matters is the θ_c angle estimate (and to some extent ϕ_c) *at the anticipated emission vertex*



Building a χ^2 statistics in this simple case

- Prefer to work in a *measurement space*
  - Where "measurement" is a single photon emission angle θ_c with respect to the track

$$\chi^2_H = \frac{[\theta_H(p,n) - \theta_c]^2}{\sigma^2_\theta}$$

for a given PID hypothesis *H* (e.g. a pion)

- This quantity should be distributed as a tabulated χ^2 with *one* degree of freedom *for a correct PID hypothesis …*
  - In other words: a cumulative quantity (CCDF) should be a flat distribution between 0 and 1
- ... and systematically biased towards higher (less probable) values *for wrong hypotheses*
  - In other words: a CCDF plot would tend to produce a spike close to 0

All the rest is built on a simple basic principle: we construct *static* χ^2 estimates for various PID hypotheses (say π/K/p) for a given set of hit-to-track associations, and the smallest χ^2 wins

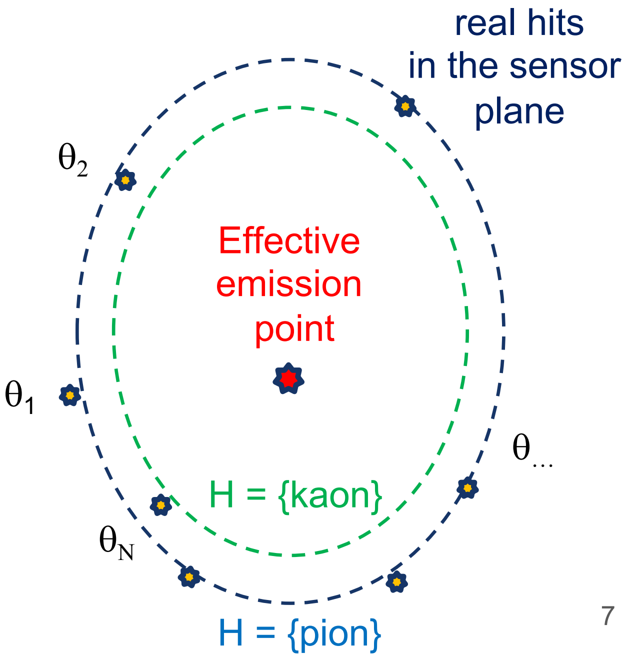
Multiple photons per track

- Assuming each photon is a single independent measurement, the χ^2 formulation is extended in a trivial way:

$$\chi_H^2 = \sum_{k=1}^{nhits} \frac{[\theta_H(p,n) - \theta_c^k]^2}{\sigma_\theta^2}$$

for a given PID hypothesis *H* (e.g. a pion)

- If there were no Rayleigh scattered photons, this quantity should be distributed as a tabulated χ^2 with *nhits* degrees of freedom *for a* correct PID hypothesis and be systematically biased towards higher (less probable) values *for wrong hypotheses*



In practical terms, this gives one a 1/√N factor in a *track-level* Cherenkov angle resolution, of an order of ~1.5 mrad for pfRICH

Accounting for a Poisson term

- Omitting certain technical issues, an expected number of "true" Cherenkov photons <$N^{expected}$> for a given combination of momentum and a mass hypothesis is known ($N_{\mathrm{pe}} = N_0 L \sin^2\theta_{c}$), and can be added to the overall χ^2 estimate as a so-called χ^2_λ term (Baker & Cousins notation), which accounts for a mismatch between the expected and the detected number of photons:

$$\chi^2_H = \sum_{k=1}^{nhits} \frac{[\theta_H(p,n) - \theta_c^k]^2}{\sigma_\theta^2} + \chi^2_{\lambda(H)}$$

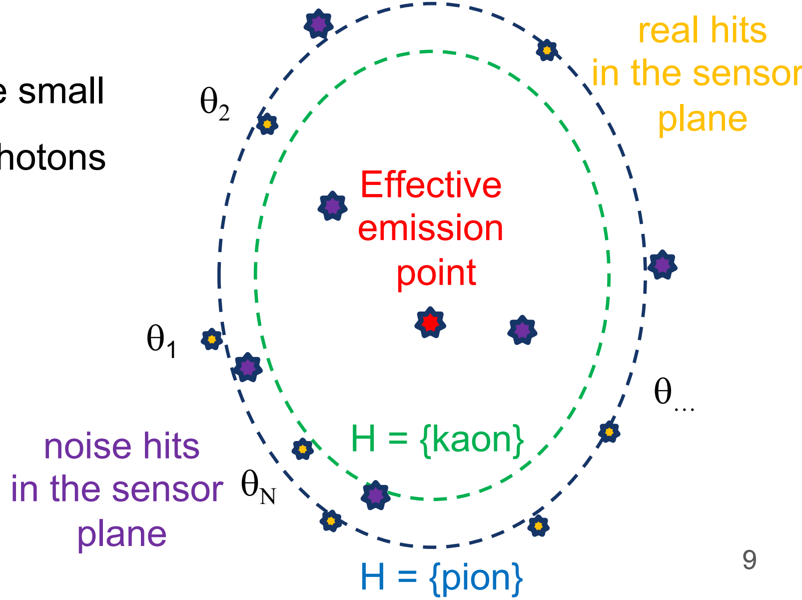
where

$$\chi^2_{\lambda(H)} = 2N^{nhits} ln(N^{nhits} / < N^{expected}(H) >\}$$

Accounting for noise hits

From this point on we assume that a hit-to-track association is *not* known

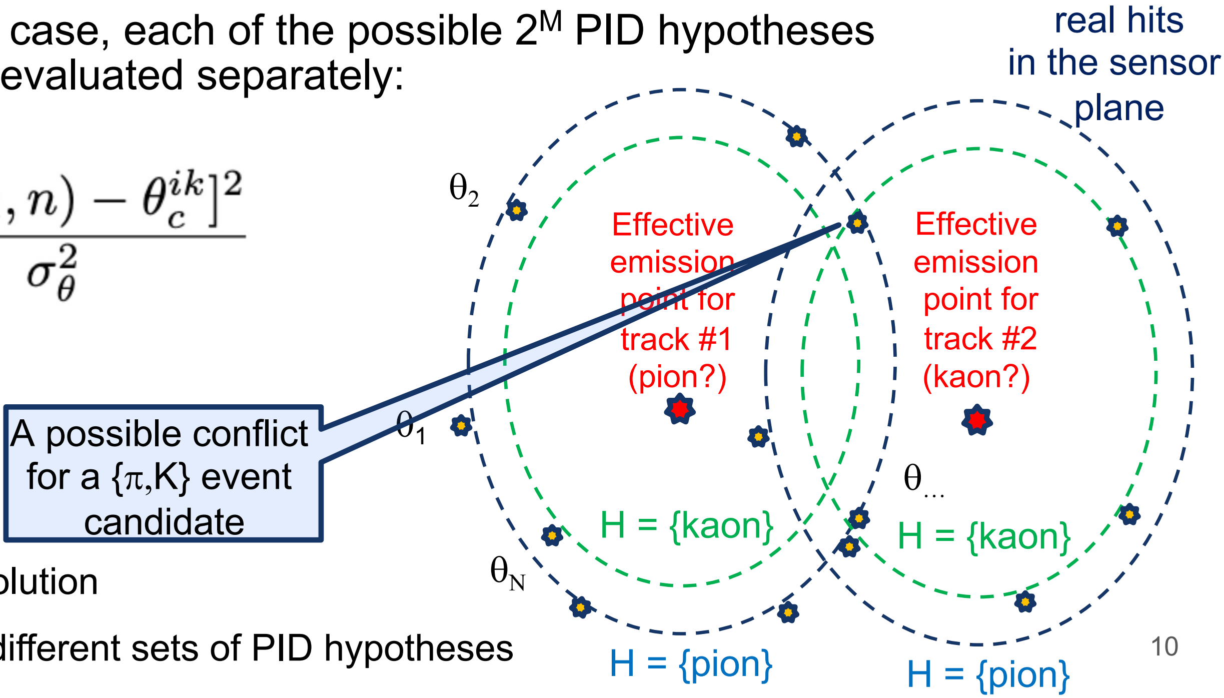
- We only deal with hits which are within a +/- 3σ band (think of +/-15 mrad) around a nominal Cherenkov θ angle *for at least one of the PID hypotheses*
- Omitting some technical complications, for a case of e.g. a π/K separation this defines two "circular" bands on the sensor plane, where such hits can be located
- A sheer area of these two bands, and a noise rate per cm^2 (as simulated!) defines an expected number of noise hits $<N^{bg}>$
  - HRPPDs: ~kHz/cm^2 & <50ps timing --> expect $<N^{bg}>$ to be small
  - Plus, another small contribution from Rayleigh scattered photons
- Each PID hypothesis can either account each of the hits as a "true" or as a "background" one, pushing highest χ^2 hits into a separate (background) χ^2_λ term one by one
  - In practical terms, a "wrong" hypothesis will incur an additional penalty because of existence of too many fake noise hits (and possibly too few true ones)



Overlapping rings

- The algorithm works on the *event* level; let's assume there are M tracks
- We only deal with hits which are within +/- 3σ band (think of +/-15 mrad) around a nominal Cherenkov θ angle *for at least one of the PID hypotheses for at least one track*
- Assuming e.g. a π/K separation case, each of the possible 2^M PID hypotheses *event-level combinations {H}* is evaluated separately:

$$\chi^2_H = \sum_{i=1}^{mtracks} \sum_{k=1}^{nhits(i)} \frac{[\theta_H(p_i,n) - \theta_c^{ik}]^2}{\sigma_\theta^2}$$

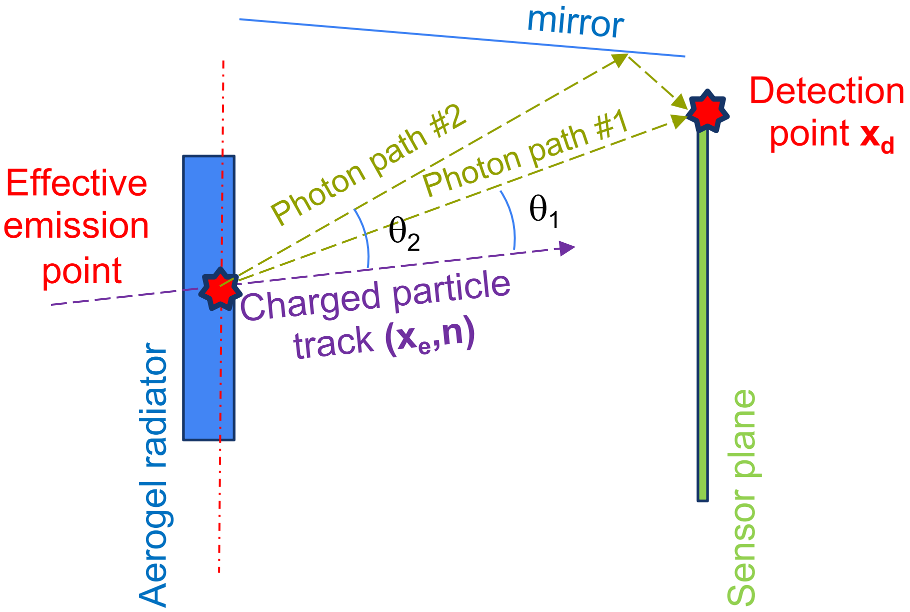


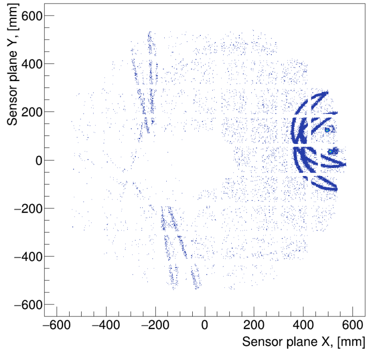
- Tracks compete for hits
  - Various options for a conflict resolution
  - Assignment can be different for different sets of PID hypotheses

Mirrors

- pfRICH has inner and outer conical mirrors, and (optionally) four small flat funneling mirrors per HRPPD sensor
  - This defines up to 20 possible optical paths for any pair of emission and detection points

- During a hit-to-track association process, the algorithm loops through all optical paths, calculates a Cherenkov emission angle, and picks up a path which gives a best match *for a PID hypothesis presently considered for this track*





Except for an obvious computational overhead, this does not change anything in the rest of the χ^2 evaluation process

Tracking information

- Small bending in the magnetic field is accounted in a trivial way (a track parameterization at the location of effective emission point is taken)

- Finite tracking system resolution can technically be accounted via replacing a static χ^2 evaluation (with no free parameters) by a full MINUIT-like pass where a track state vector at the effective emission point is added as a separate term to the χ^2 ansatz, with its inverse covariance matrix *C* as a metric:

$$\chi^2_H = \sum_{k=1}^{nhits} \frac{[\theta_H(p,n) - \theta_c^k]^2}{\sigma_\theta^2} + < X_{track}|C^{-1}|X^T_{track} >$$

where

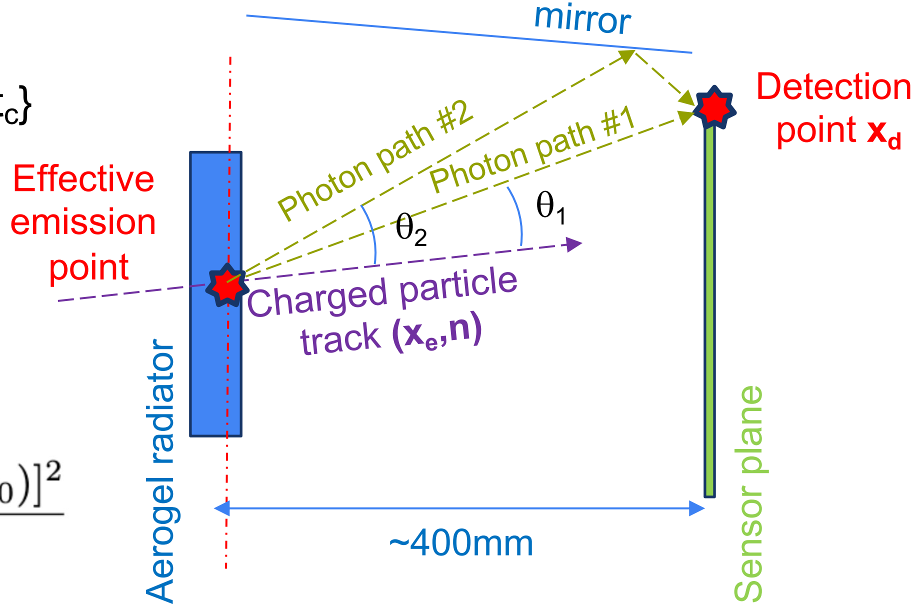
$$X_{track} = \{x, y, sx, sy, 1/p\}_{@aerogel}$$

Though unlikely, this may actually improve the tracking estimate itself

Timing information

- Measurements become 2D vectors: $\{\theta_c\}$ -> $\{\theta_c, t_c\}$
  - Where t_c is a time measured at the sensor plane
- Timing is used in both hit-to-track association for a given mass hypothesis, and in the χ^2 ansatz:

$$\chi_H^2 = \sum_{k=1}^{nhits} \frac{[\theta_H(p,n) - \theta_c^k]^2}{\sigma_\theta^2} + \sum_{k=1}^{nhits} \frac{[t_H(p) - (t_c^k - t_0)]^2}{\sigma_t^2}$$

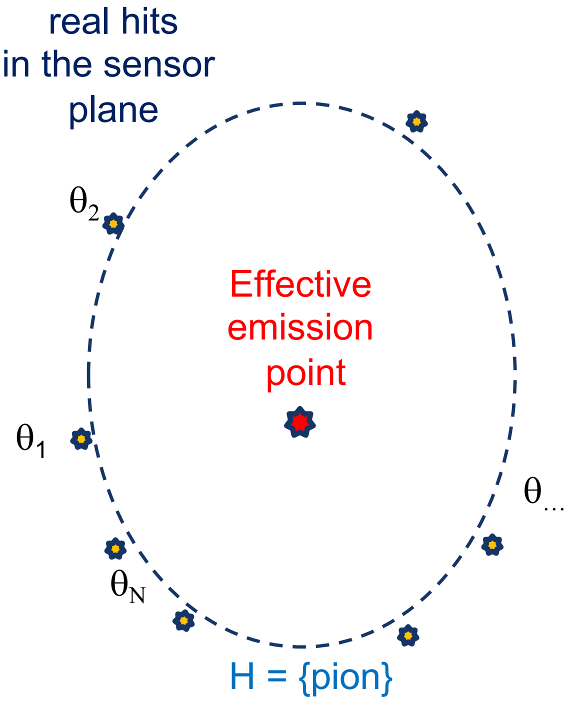


Path#1 and path#2 will not only have a different θ_c, but a substantially different length (and therefore, a very different - compared to a ~50ps resolution - flight time between the emission and detection points)

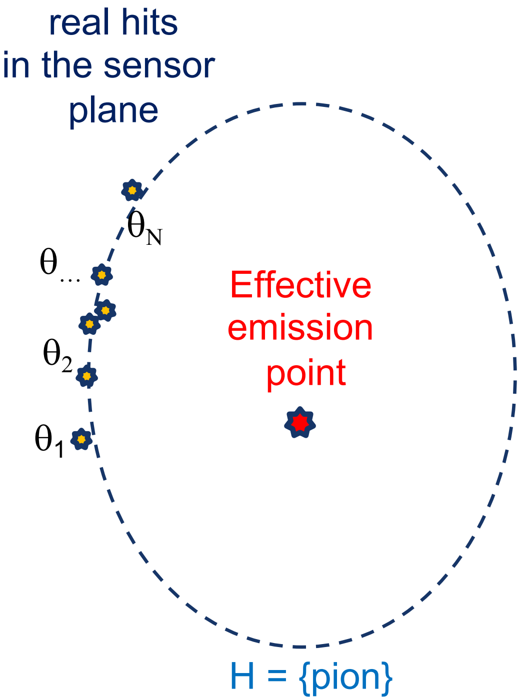
- Presently only a *static* χ^2 evaluation
  - Assume t_0 is known within few dozens of ps, resulting in an “effective” hit timing resolution σ_t ~ 50ps
- Apparently one can add a t_0 *estimate* in a (linearized) MINUIT-like fashion

Other considerations

- Make use of a signature χ^2?



- Make use of binning in ϕ_c?



Both configurations are "unlikely", but the presently available algorithm does not capture this