

ML meets HPC meets CERN

Problemas abiertos

Daniel Hugo Cámpora Pérez
dcampora@cern.ch

Universidad de Sevilla
CERN

Seminario (I+A)A, 12 junio, 2017

Section 1

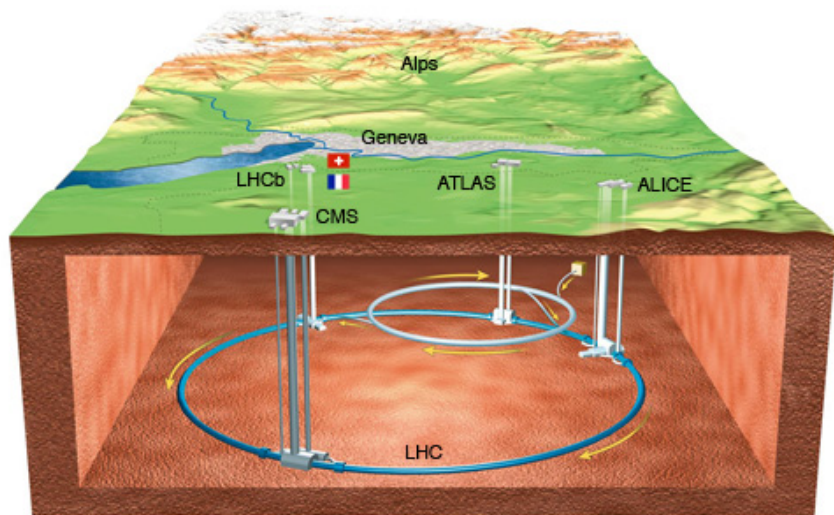
Introducción

Un poco sobre mí

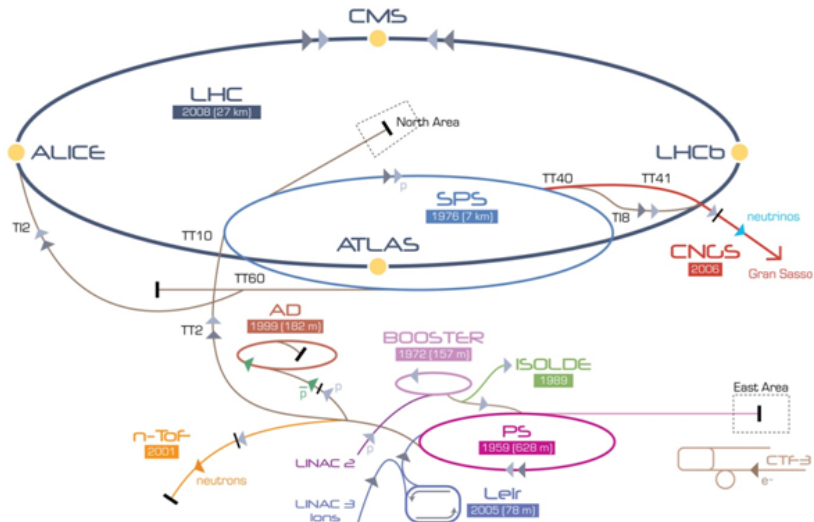
- He pasado por varios contratos en el CERN
- Actualmente Doctoral Student @ CERN, con la US
- Muchas preguntas abiertas, os presento algunas de ellas



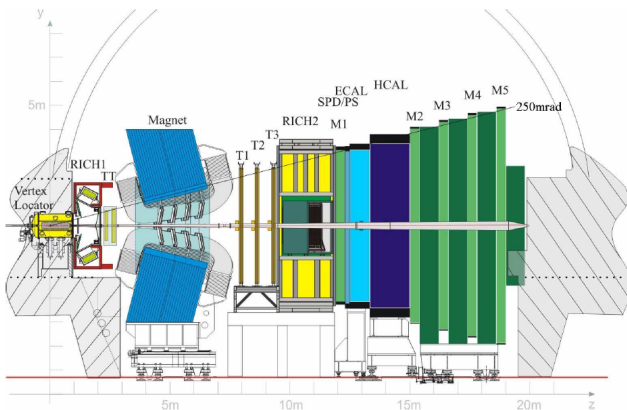
CERN



Un conjunto de aceleradores

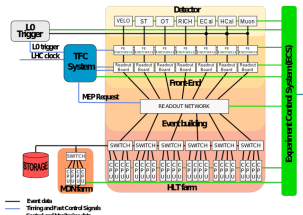
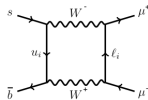
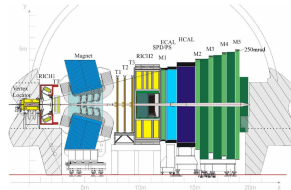
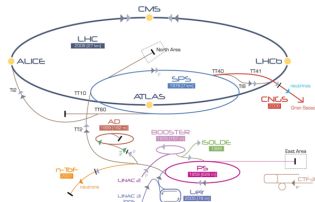


LHCb



- Detectores de trazas
- Detectores RICH
- Calorímetros

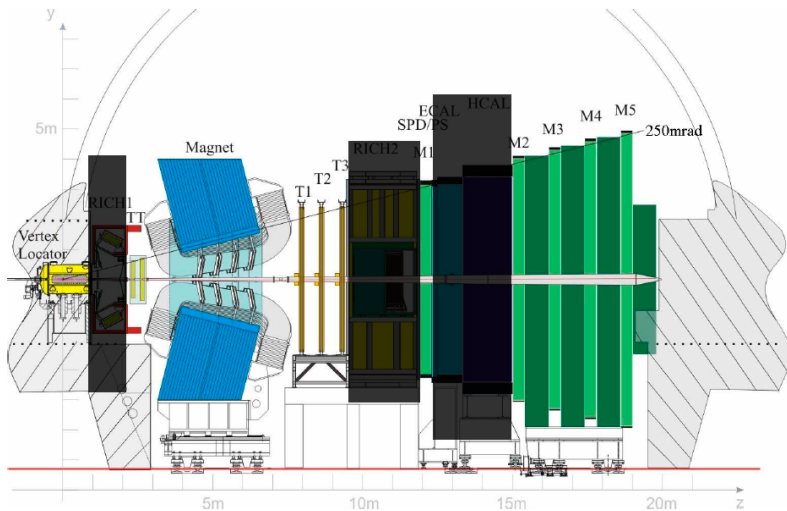
La cadena completa



Section 2

Detectores de tracks - Velopix

LHCb - Detectores de tracks

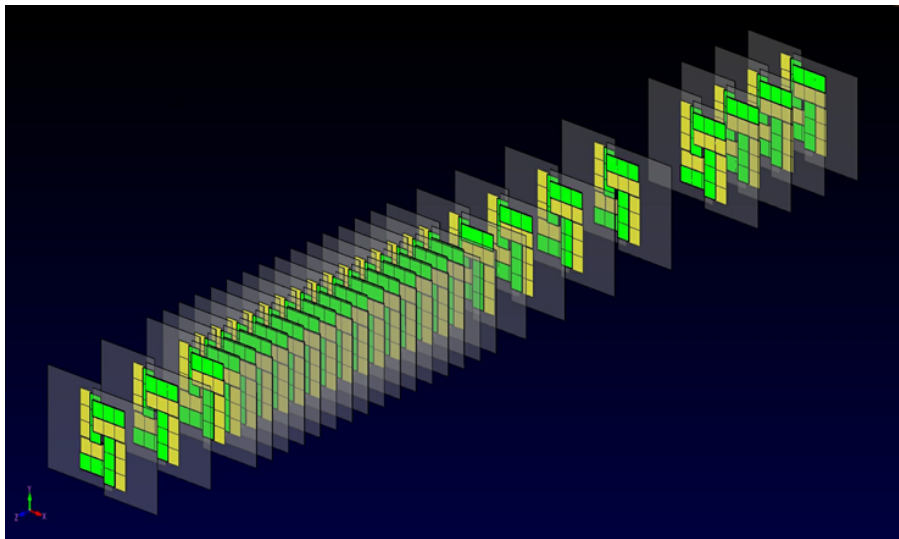


Algunas consideraciones

Los subdetectores de trazas de partículas (tracks) reconstruyen las trayectorias de las partículas a su paso por el detector.

- Las partículas se generan tras la colisión, dentro del subdetector VELO
- Viajan a través del VELO y TT en línea recta
- En el imán (magnet), su trayectoria se desvía acorde a la carga y momento de la partícula
- Más adelante, dejan trazas en los detectores T1, T2 y T3
- En caso de ser muones, es posible que dejen trazas en M1 - M5

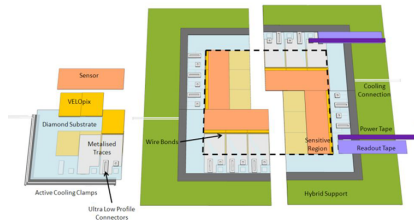
¿Qué es el VELO?



Subdetector VELO

Propiedades

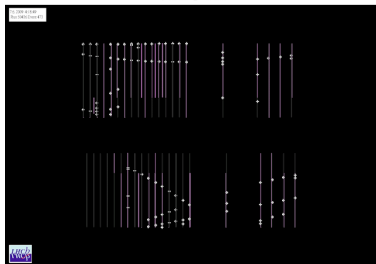
- 52 sensores, 26 a cada lado
- Sensor de silicio
- Cada sensor tiene 12 chips
- Cada chip tiene una resolución de 256x256 píxeles
- Visualización
(<https://lbevent.cern.ch/EventDisplay/index.html>)



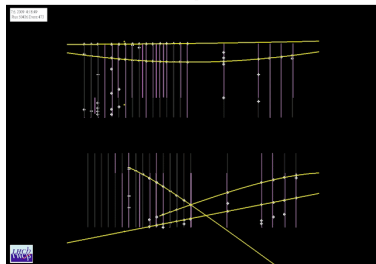
Formulación del problema

- Cluster (hit) - Excitación de un sensor
- Track - Conjunto de clusters

VELO RZ

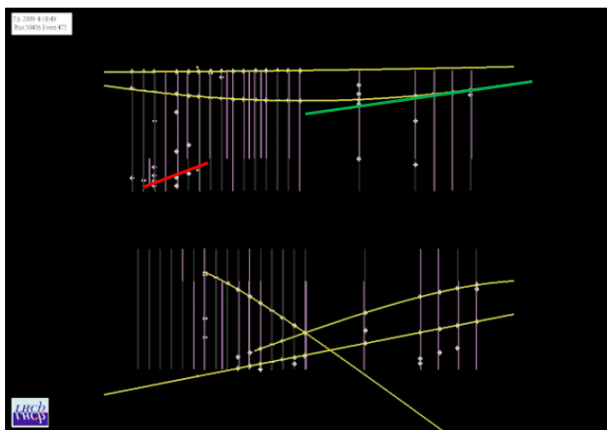


VELO RZ



Formulación del problema (2)

- Track fantasma - Generado por ruido
- Track clon - Copia de otro



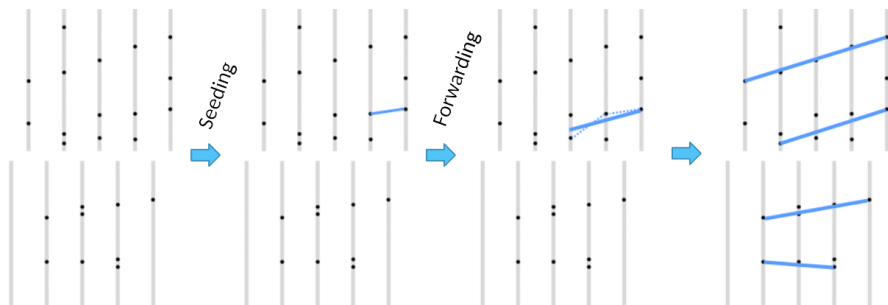
Características de los tracks

- Un track debe tener al menos tres clusters
- Los clusters compartidos entre tracks no deben superar un umbral (60%)
- El 96% de los tracks se produce en un solo lado del sensor
- Ante tracks similares, prima el track más largo
- Ante tracks de igual longitud, prima el que tenga menor χ^2

¿Cómo resolver el problema?

- Método clásico
- Transformada de Hough
- Clustering

Método clásico



- Se generan semillas (seeds) con pares de clusters
- Se extienden los pares mientras haya clusters compatibles en los sensores siguientes

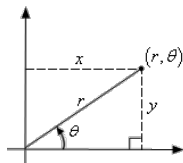
Método clásico (2)

Este método funciona bien, y es el que está implementado actualmente. No obstante, tiene algunas pegas:

- La búsqueda es dependiente del orden de búsqueda: Se queda con el primer cluster compatible que encuentra
- El algoritmo es local
- No es paralelizable

Transformada de Hough

Teniendo en cuenta la transformación a coordenadas polares:



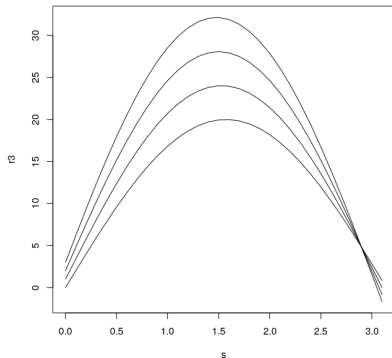
Todas las líneas que pueden pasar por un punto x, y pueden expresarse con la función

$$f(\rho) = x * \cos(\sigma) + y * \sin(\sigma)$$

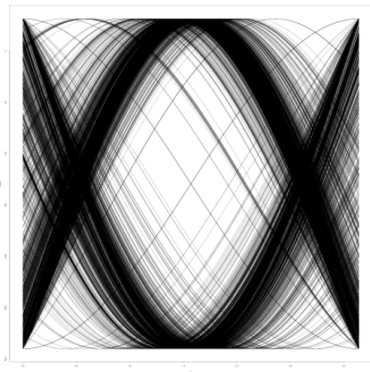
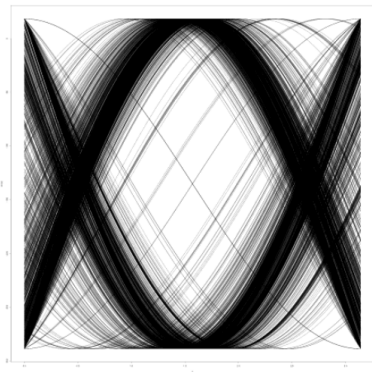
Transformada de Hough (2)

Por ejemplo, una sesión en R...

```
> s = seq(0,3.14,0.1)
> #y = 4x + 20
> r0 = 0*cos(s) + 20*sin(s)
> r1 = 1*cos(s) + 24*sin(s)
> r2 = 2*cos(s) + 28*sin(s)
> r3 = 3*cos(s) + 32*sin(s)
> plot(s, r3, type="l")
> lines(s, r2, type="l")
> lines(s, r1, type="l")
> lines(s, r0, type="l")
```



Pero la realidad no es tan fácil

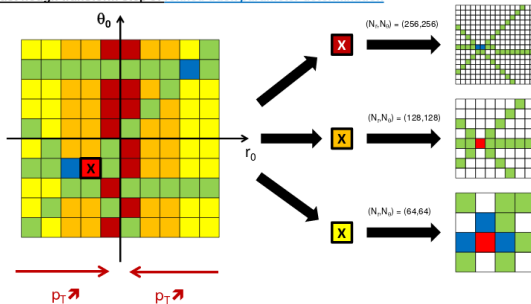


The Ring?

Ideas para la transformada de Hough

- Hacer un análisis más fino donde sea necesario

→ Fast hough transform step 3: refined track parameter determination



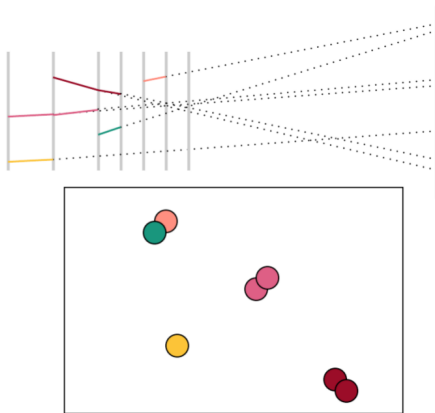
→ Second hough transform **p_T dependent** (higher p_T requires finer binning)

Ideas para la transformada de Hough (2)

- Hacer un análisis más fino donde sea necesario
- Elegir un threshold y un binning correcto
- Paralelizar

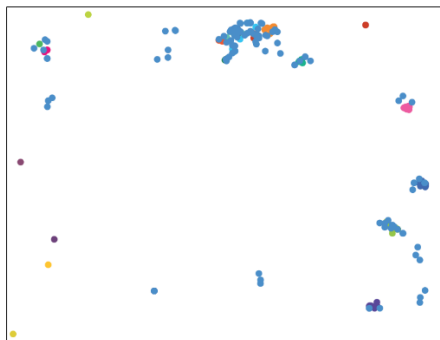
Clustering

Podemos transformar el problema en un problema de clustering



Clustering (2)

Esa idea no funcionó bien: *Mucho ruido y pocas nueces*



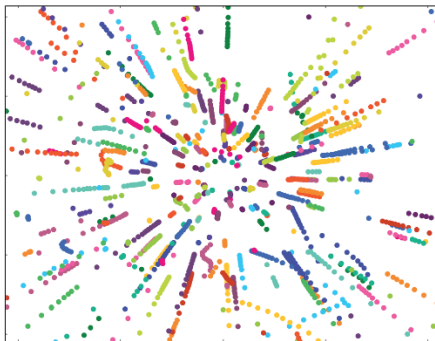
(a) Extrapolating pairs of hits(segments)



(b) Extrapolating triplets of hits

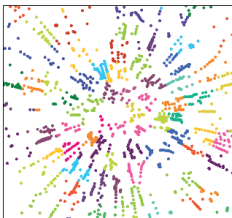
Clustering (3)

No obstante, algo mucho más sencillo da mejores resultados

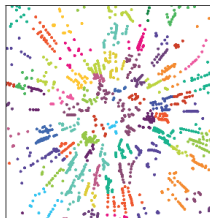


(c) Extrapolating segments with a virtual point in the origin. We can notice the lack of noise (dark blue) in this final picture, since the procedure inherently creates only one point in the plane for every hit.

Clustering (4)



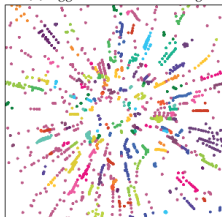
(a) KMeans



(b) Agglomerative Clustering



(c) DBSCAN



(d) HDBSCAN

Clustering (5)

Hasta ahora, HDBSCAN con extrapolación desde el origen es lo que mejor resultados nos ha dado, usando una técnica no convencional. ¡Pero los resultados están lejos de ser utilizables!

16621 tracks including	8784 ghosts (52.8%).	Event average	50.7%						
velo :	7623 from	28019 (27.2%, 28.3%)	55 clones (0.72%),	purity: (90.65%, 90.65%),	hitEff: (88.87%, 88.85%)				
long :	4171 from	9063 (46.0%, 47.7%)	37 clones (0.89%),	purity: (91.73%, 91.76%),	hitEff: (89.01%, 88.94%)				
long>5GeV :	3147 from	5950 (52.9%, 54.2%)	26 clones (0.83%),	purity: (92.38%, 92.47%),	hitEff: (90.40%, 90.43%)				
long_strange :	90 from	357 (25.2%, 22.9%)	0 clones (0.00%),	purity: (90.57%, 90.01%),	hitEff: (84.32%, 83.87%)				
long_strange>5GeV :	47 from	169 (27.8%, 23.7%)	0 clones (0.00%),	purity: (91.17%, 91.30%),	hitEff: (92.30%, 90.36%)				
long_fromb :	239 from	553 (43.2%, 44.7%)	2 clones (0.84%),	purity: (91.80%, 91.20%),	hitEff: (84.87%, 88.10%)				
long_fromb>5GeV :	197 from	441 (44.7%, 45.7%)	1 clones (0.51%),	purity: (92.04%, 91.29%),	hitEff: (86.42%, 88.39%)				

(g) HDBSCAN

8569 tracks including	49 ghosts (0.6%).	Event average	0.5%						
velo :	8287 from	28019 (29.6%, 30.8%)	89 clones (1.07%),	purity: (98.30%, 98.35%),	hitEff: (86.67%, 86.35%)				
long :	3792 from	9063 (41.8%, 43.4%)	67 clones (1.77%),	purity: (98.13%, 98.17%),	hitEff: (86.62%, 86.19%)				
long>5GeV :	2826 from	5950 (47.5%, 48.7%)	50 clones (1.77%),	purity: (98.31%, 98.40%),	hitEff: (88.21%, 87.91%)				
long_strange :	89 from	357 (24.9%, 23.9%)	0 clones (0.00%),	purity: (97.33%, 97.14%),	hitEff: (79.48%, 78.60%)				
long_strange>5GeV :	44 from	169 (26.0%, 22.9%)	0 clones (0.00%),	purity: (97.87%, 97.82%),	hitEff: (83.85%, 82.62%)				
long_fromb :	216 from	553 (39.1%, 40.1%)	3 clones (1.39%),	purity: (98.01%, 97.92%),	hitEff: (84.52%, 86.70%)				
long_fromb>5GeV :	176 from	441 (39.9%, 40.8%)	3 clones (1.70%),	purity: (98.07%, 97.93%),	hitEff: (85.85%, 87.04%)				

(h) HDBSCAN with postprocessing

¿Queréis probar?

Repositorio en Github

(https://github.com/dcampora/velopix_tracking)

Recapitulando

Reconstrucción Velopix

- Método clásico
- Transformada de Hough
- Clustering
- [Su anuncio aquí]

Section 3

RICH

Algunas preguntas

¿Cuál es la velocidad de la luz?

Algunas preguntas

¿Cuál es la velocidad de la luz?

299 792 km/s.

¿En qué medio?

Algunas preguntas

¿Cuál es la velocidad de la luz?

299 792 km/s.

¿En qué medio?

En el vacío.

¿Qué le pasa a la luz cuando entra en el agua?

Algunas preguntas

¿Cuál es la velocidad de la luz?

299 792 km/s.

¿En qué medio?

En el vacío.

¿Qué le pasa a la luz cuando entra en el agua?

Se refracta.

¿Por qué se refracta?

Algunas preguntas

¿Cuál es la velocidad de la luz?

299 792 km/s.

¿En qué medio?

En el vacío.

¿Qué le pasa a la luz cuando entra en el agua?

Se refracta.

¿Por qué se refracta?

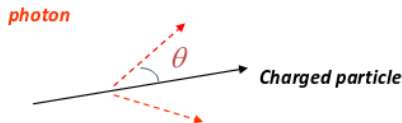
Porque cambia su velocidad.

Algunas preguntas (2)

¿Qué pasa si una partícula va más rápido que la velocidad de la luz?

La radiación de Cherenkov

La radiación de Cherenkov es un tipo de radiación electromagnética producida cuando una partícula va más rápido que la velocidad de la luz.



$$\cos(\theta) = 1 / (n \beta) \quad \text{where } n = \text{Refractive Index} = c/c_M = n(E_{ph})$$

$$\beta = v/c = p/E = p / (p^2 + m^2)^{0.5} = 1 / (1 + (m/p)^2)^{0.5}$$

β = velocity of the charged particle in units of speed of light (c) vacuum

p, E, m = momentum, Energy, mass of the charged particle.

c_M = Speed of light in the Medium (Phase velocity) ,

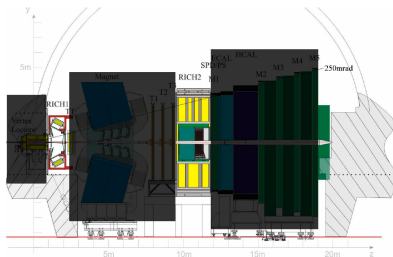
E_{ph} = Photon Energy, λ = Photon Wavelength.

➤ Theory of Cherenkov Radiation: Classical Electrodynamics by J.D.Jackson (Section 13.5)

La radiación de Cherenkov (2)

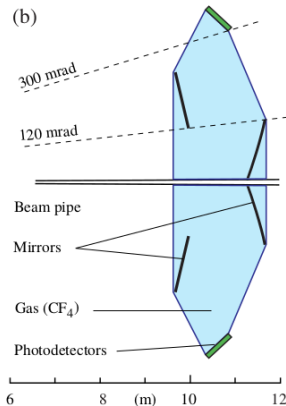
En concreto, el ángulo de los fotones producidos depende de la masa y el momento de la partícula, con lo que al detectar el diámetro de los círculos, estaremos más cerca de detectar qué partícula lo originó (protón, kaón, pión, etc.).

El cono de luz generado por la radiación de Cherenkov se detecta mediante detectores RICH (Ring Imaging Cherenkov detectors). En LHCb:



Formulación del problema

Dadas unos tracks de partículas, unos fotones y una geometría del detector, determinar la hipótesis más probable para cada partícula de entre: Pión, electrón, muón, kaón y protón.



Visualmente

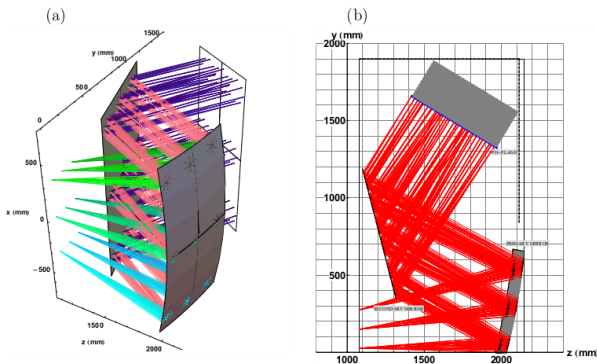
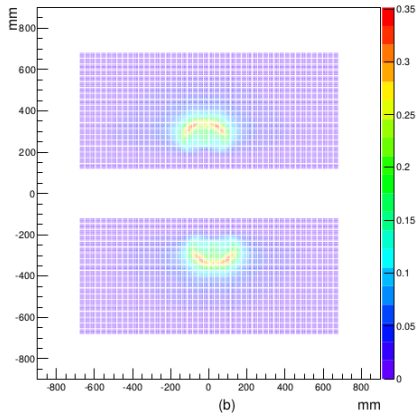
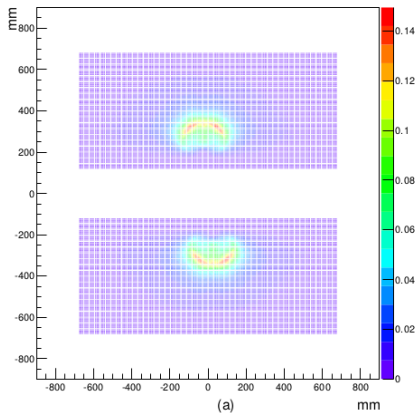


Figure 2.21: Simulated Cherenkov photons in the upper half of the upgraded RICH 1; (a) 3D view, (b) 2D view in the vertical plane.

Visualmente (2)



¿Cómo resolver el problema?

- Solución analítica
- Reconstrucción de círculos
- Elastic Neural Network

Solución analítica

La solución analítica consiste en reconstruir las posibles trayectorias de los fotones para cada partícula, una a una, y asignar una probabilidad (likelihood) a la asignación. Después, se cambia la asignación track-partícula, maximizando la probabilidad.

Es decir, se compone de dos partes:

- Reconstrucción de fotones
- Maximización de probabilidad

Reconstrucción de fotones

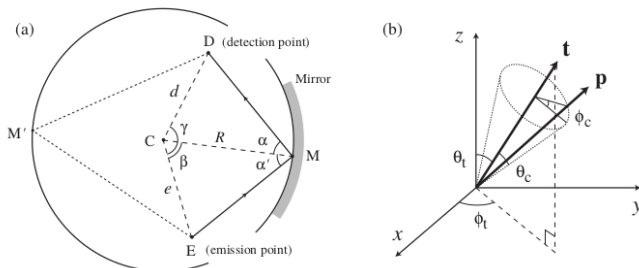


Figure 3: (a) Definition of parameters used in the reconstruction of the Cherenkov angles, for a photon emitted from point E , reflected in a spherical mirror at M and detected at D . (b) Definition of the angles describing the direction of a track $\mathbf{t}$ in the lab frame, and a photon $\mathbf{p}$ emitted by the track.

Maximización de probabilidad

Para cada track $t_i, i \in 0, \dots, n$, se debe probar con las hipótesis {pión, electrón, muón, kaón, protón}. Existen 5^n posibilidades.

No existe una heurística clara, así que se comienza con la hipótesis más probable para todas las partículas (pión), y se prueban partícula a partícula las otras cuatro posibilidades, asignando aquella que maximice la probabilidad.

Esto reduce la búsqueda a una búsqueda local, testeando únicamente $5n$ posibilidades.

Problemas de esta solución

Tanto la reconstrucción de fotones como la maximización de probabilidad son procesos costosos. En particular, para cada fotón, hay que calcular el reflejo en un espejo esférico, lo que implica resolver la ecuación:

$$4e^2 d^2 \sin^4 \beta - 4e^2 d_y R \sin^3 \beta + (d_y^2 R^2 + (e + d_x)^2 R^2 - 4e^2 d^2) \sin^2 \beta + 2e d_y (e - d_x) R \sin \beta + (e^2 - R^2) d_y^2 = 0 . \quad (3)$$

This can be solved, using for example a routine in the CERN library [6], and gives four solutions for $\sin \beta$, two complex and two real. Of the real solutions one is the “backward” reflection (that would exist if the mirror were a complete sphere, shown as M' in Fig. 3); the other is the desired solution, and can be selected from knowledge of the RICH detector geometry. The value of $\cos \beta$ can then be extracted using Eq. 2, and the coordinates of the reflection point M determined.

Problemas de esta solución

De hecho, es bastante costoso...

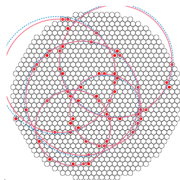
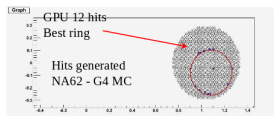
Visualización coste en tiempo (https://lbevent.cern.ch/online/dcamporupgrade_visualization/index.html)

Reconstruir círculos

Una idea completamente diferente sería partir de la imagen con los fotones, y tratar de reconstruir círculos. Esto funciona en otros detectores, por ejemplo:



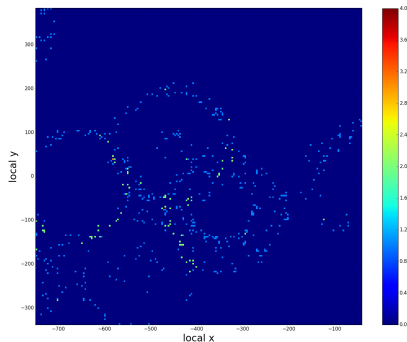
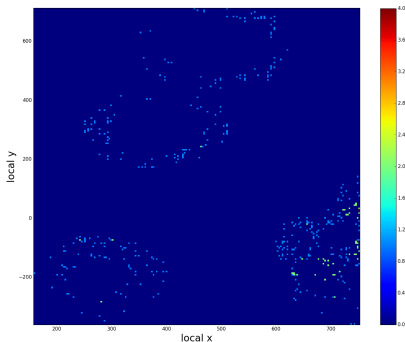
- Natively built for pattern recognition problems
- **First attempt:** ring reconstruction in RICH detector.



It's a **pilot project**, very promising R&D!

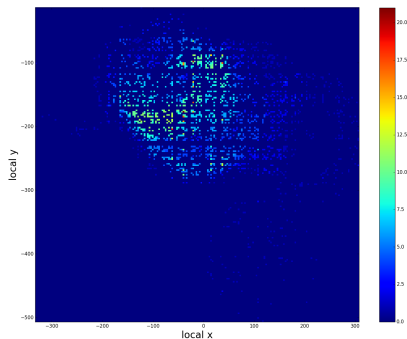
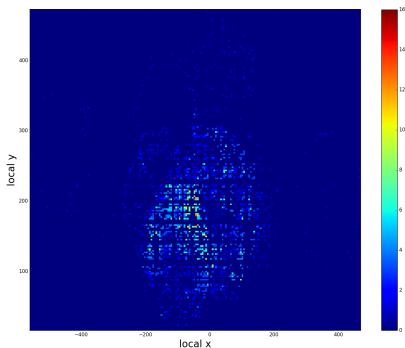
Encuentra los círculos

En LHCb, los círculos no son tan fáciles de ver:



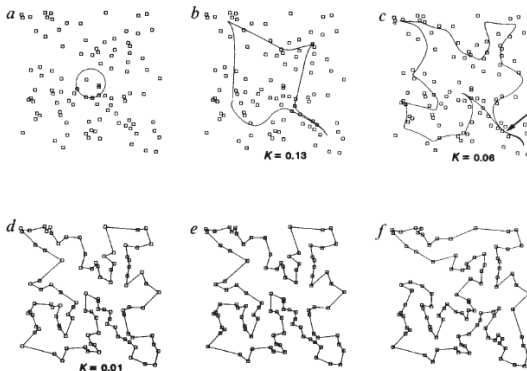
Encuentra los círculos (2)

En LHCb, los círculos no son tan fáciles de ver:



Elastic Neural Networks

En 1987, se publicó una heurística para el problema del viajante de comercio (TSP), consistente en empezar con una figura, y modelar cada ciudad como un punto que ejerce una fuerza atrayente sobre la figura, deformándola.



Elastic Neural Networks

Basándose en esta idea, el problema de encontrar círculos en el RICH puede también partir de una figura que se pueda deformar, permitiendo solo círculos y solo que crezca.

At the recognition step a special energy function is used to separate ring hits from others.

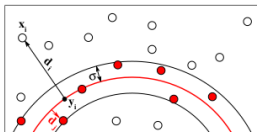


Figure 3: Selection of ring hits at the recognition step

Let us define the following variables (figure 3):

$\mathbf{x}_i$ — the i -th hit;

$\mathbf{y}_i$ — the elastic net point closest to the i -th hit;

$\mathbf{d}_i = (\mathbf{x}_i - \mathbf{y}_i)$ — distance between the hit and the net.

With these variables:

- the external forces are defined by minimizing the sum of (not squared) distances to hits

$$E = \sum |\mathbf{d}_i|, \quad (4)$$

Recapitulando

Reconstrucción Velopix

- Método clásico
- Transformada de Hough
- Clustering
- [Su anuncio aquí]

Reconstrucción RICH

- Solución analítica
- Reconstrucción de círculos
- Elastic Neural Network
- [Su idea aquí]

¡Gracias! ¿Preguntas?

