

# Dados para o trabalho

Bernardo Sotto-Maior Peralva

08 de junho de 2021

# Ambiente Experimental

CERN

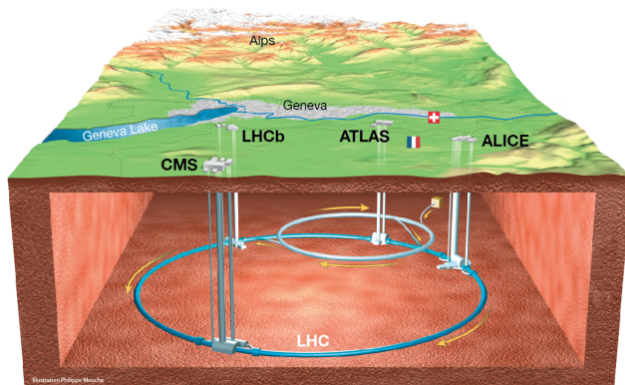
- O trabalho será desenvolvido no contexto do maior e mais energético acelerador de partículas até o momento, o LHC (do inglês, *Large Hadron Collider*), localizado no CERN (do francês *Centre Europeenne pour La Recherche Nucleaire*).



# Ambiente Experimental

## LHC

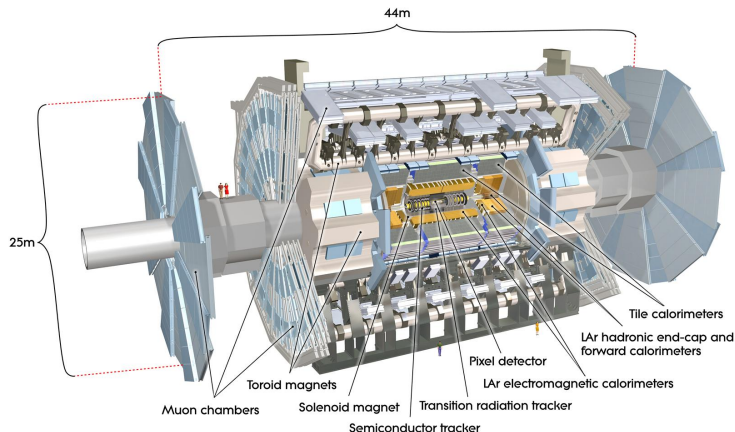
- O LHC é um acelerador circular de 27 km de extensão, acelerando feixes de prótons próximo a velocidade da luz.
- Colisões são realizadas em 4 experimentos: ATLAS, CMS, ALICE e LHCb.



# Ambiente Experimental

## ATLAS

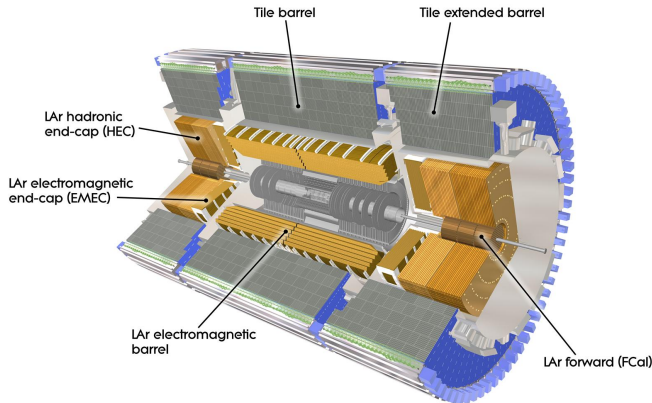
- Dentre os principais experimentos projetados para medir os subprodutos das colisões, encontra-se o detector de propósito geral ATLAS (do inglês, *A Toroidal LHC Apparatus*).



# Ambiente Experimental

## Sistema de calorimetria do ATLAS

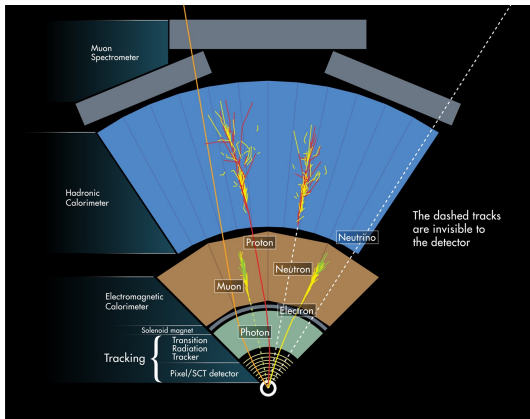
- No ATLAS, o sistema de calorimetria é composto por dois componentes: o Calorímetro de Argônio Líquido, ou LAr (do inglês *Liquid Argon Calorimeter*) e o Calorímetro Hadrônico de Telhas, ou TileCal (do inglês *Tile Hadronic Calorimeter*).



# Ambiente Experimental

## Sistema de calorimetria do ATLAS

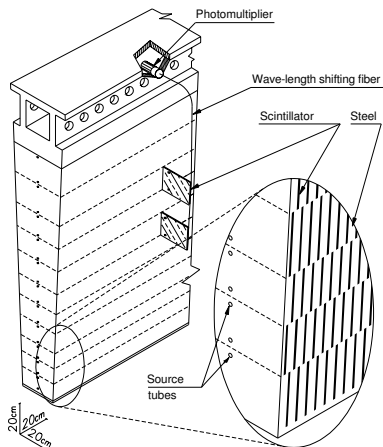
- Em calorimetria de altas energias, técnicas de estimação de parâmetros são aplicadas para o estimar propriedades importantes de partículas absorvidas e amostradas pela eletrônica de leitura de um calorímetro.



# Ambiente Experimental

## O Calorímetro de Telhas (TileCal)

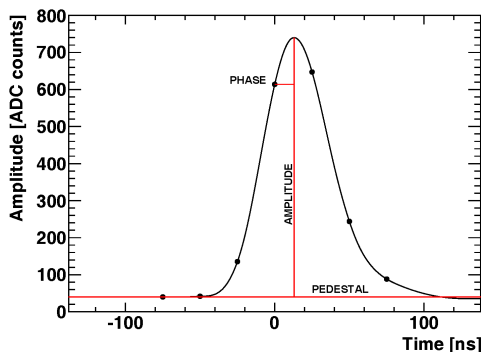
- O TileCal é baseado na técnica de amostragem que utiliza placas (ou telhas) de plástico cintilante como material ativo, intercalada com camadas de aço como material absorvente.



# Ambiente Experimental

## Calorimetria de Altas Energias

- Métodos de estimação são utilizados para estimar propriedades como **amplitude**, **fase** e **pedestal** de sinais gerados pela eletrônica de leitura de um calorímetro, sendo a amplitude a informação de maior importância.

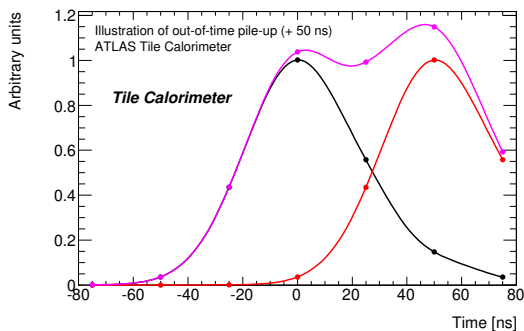




# Ambiente Experimental

## Empilhamento de sinais

- Em sistemas que operam em altas taxas de eventos, o sinal amostrado pode conter informações de colisões adjacentes, resultando no efeito de empilhamento de sinais e dificultando a estimação da amplitude do sinal de interesse.



- Dois conjuntos de dados são disponibilizados, contendo 100.000 realizações cada:
  - Sem sinal de interesse: ruído eletrônico + empilhamento
  - Com sinal de interesse cuja amplitude deve ser estimada pelos métodos propostos. Sinal imerso ao ruído eletrônico e empilhamento
- **Ruído eletrônico:** modelo gaussiano com média zero e 1,5 contagens de ADC de desvio padrão
- **Empilhamento de sinais:** modelo exponencial de média 30 contagens de ADC, variando de 0 a 100% de ocupação (probabilidade de ocorrer um sinal em uma dada colisão).
- **Sinal de interesse:** modelo exponencial de média 90 contagens de ADC
- **Desvio de fase:** modelo uniforme  $[-4,4]$  ns
- **Deformação do pulso:** amplitude da amostra temporal variando 1% uniformemente
- **Pedestal:** valor constante igual a 30 contagens de ADC

# Dados

Descrição do banco de dados e parâmetros utilizados

- Exemplos de ruído (ruído eletrônico+empilhamento)

var 1	var 2	var 3	var 4	var 5	var 6	var 7
30.6789	30.4616	30.5791	47.2937	100.1157	121.6271	78.5617
47.1437	35.8586	34.7989	44.4713	148.5912	256.1586	187.6970
95.3952	72.3061	49.1841	38.0381	29.2746	29.9340	31.8294

- Exemplos de sinal corrompido (sinal de interesse imerso em ruído eletrônico+empilhamento)

var 1	var 2	var 3	var 4	var 5	var 6	var 7	Amp
30.0214	29.9444	50.4430	66.9711	66.3779	65.7195	62.1902	18.0457
69.9513	47.8661	58.7896	113.0666	103.4931	61.8475	38.1190	69.1814
28.0913	32.8565	77.4947	133.6823	106.4655	62.6288	51.8289	2.3117

Dados disponíveis em:

<https://drive.google.com/drive/folders/1AGRuku7yjgJU8uuanUGm-TXy9ySohxsL?usp=sharing>

- Desconvolução de sinais, ou Multi-Amplitude Estimator (MAE):  
<https://github.com/ingoncalves/jcae-2021-analysis/blob/master/analysis/filters/cof.py>
- Optimal Filter (OF):  
<https://github.com/ingoncalves/jcae-2021-analysis/blob/master/analysis/filters/of2.py>
- Wiener Filter:  
<https://github.com/ingoncalves/jcae-2021-analysis/blob/master/analysis/filters/wiener.py>
- Sparse:  
[https://github.com/ingoncalves/jcae-2021-analysis/blob/master/analysis/filters/sparse\\_cof.py](https://github.com/ingoncalves/jcae-2021-analysis/blob/master/analysis/filters/sparse_cof.py)

# BACKUP

- Sistemas de calorimetria em geral utilizam estimadores lineares por fornecerem uma resposta rápida e serem de fácil implementação (hardware e software).

$$\hat{A} = \sum_{k=0}^{N-1} x[k]w[k].$$

- O *Optimal Filter* (OF) tem sido empregado atualmente em diversos sistemas de calorimetria.
- Além do OF, métodos alternativos como o *Constrained Optimal Filter* (COF), utilizam o conhecimento do formato do pulso em suas formulações matemáticas, bem como assumem o ruído eletrônico como sendo Gaussiano.

- Modela o sinal de saída das células do calorímetro como

$$x[k] = Ag[k - \tau] + n[k] + ped.$$

e é baseado na soma ponderada das amostras temporais

$$\hat{A} = \sum_{k=0}^{N-1} x[k]w[k].$$

- O vetor de coeficientes  $w$  é calculado de forma *offline* visando minimizar a variância da amplitude. Algumas restrições podem ser consideradas em seu projeto, para garantir a imparcialidade do estimador, imunidade à fase a à flutuações do pedestal (OF2)

$$\text{i) } \sum_{k=0}^{N-1} w[k]g[k] = 1 \quad \text{ii) } \sum_{k=0}^{N-1} w[k]\dot{g}[k] = 0 \quad \text{iii) } \sum_{k=0}^{N-1} w[k] = 0$$

em que os vetores  $g$  e  $\dot{g}$  correspondem ao pulso de referência normalizado e sua derivada, respectivamente.

- Os coeficientes  $w$  podem ser encontrados resolvendo o sistema matricial abaixo, utilizando os multiplicadores de Lagrange:

$$\begin{pmatrix} C[1,1] & C[1,2] & \cdots & C[1,N] & -g[1] & -\dot{g}[1] & -1 \\ C[2,1] & C[2,2] & \cdots & C[2,N] & -g[2] & -\dot{g}[2] & -1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ C[N,1] & C[N,2] & \cdots & C[N,N] & -g[N] & -\dot{g}[N] & -1 \\ g[1] & g[2] & \cdots & g[N] & 0 & 0 & 0 \\ \dot{g}[1] & \dot{g}[2] & \cdots & \dot{g}[N] & 0 & 0 & 0 \\ 1 & 1 & \cdots & 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} w[1] \\ w[2] \\ \vdots \\ w[N] \\ \lambda \\ \xi \\ \nu \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix},$$

em que a matriz  $C$  corresponde à matriz de covariância do ruído, e  $\lambda$ ,  $\xi$  e  $\nu$  os multiplicadores de Lagrange.

- Em geral  $C = I$  (matriz identidade), a qual descreve somente o ruído eletrônico. Porém, projetos otimizados podem usar a matriz de covariância do ruído de fundo resultante.



## Constrained Optimal Filter

- Foi desenvolvido no intuito de prover um estimador onde não é necessário conhecer as características *a priori* do empilhamento. Modela o sinal de saída das células do calorímetro como um conjunto  $a[k]$  de deposições de energia

$$x[k] = \sum_i (g[i]a[n - i]) + n[k]$$

- O valor do pedestal é subtraído das amostras temporais  $x[k]$  antes de se aplicar o filtro. O valor do pedestal pode ser obtido através de um banco de dados, por exemplo.
- A estimação da energia implica em desconvoluir a sequência  $x[k]$  da resposta ao impulso  $g[k]$ . Aplicando um procedimento similar ao OF, a amplitude pode ser dada por

$$\hat{a}_p = W_p^T x$$

em que

$$W_p = C_p^{-1} G_p (G_p^T C_p^{-1} G_p)^{-1}.$$

## Constrained Optimal Filter

- A matriz  $G_p$  é dada por

$$G_p = \begin{pmatrix} g[3] & g[4] & g[5] & g[6] & 0 & 0 & 0 \\ g[2] & g[3] & g[4] & g[5] & g[6] & 0 & 0 \\ g[1] & g[2] & g[3] & g[4] & g[5] & g[6] & 0 \\ g[0] & g[1] & g[2] & g[3] & g[4] & g[5] & g[6] \\ 0 & g[0] & g[1] & g[2] & g[3] & g[4] & g[5] \\ 0 & 0 & g[0] & g[1] & g[2] & g[3] & g[4] \\ 0 & 0 & 0 & g[0] & g[1] & g[2] & g[3] \end{pmatrix},$$

- Caso o ruído de fundo seja considerado Gaussiano, a amplitude estimada pode ser dada por

$$\hat{a} = G^{-1}x$$

- Esta expressão não depende da matriz de covariância do ruído  $C$ , sendo esta uma das vantagens do método COF com relação ao método OF.
- De posse do vetor de amplitudes, é aplicado um corte linear para selecionar somente amplitudes acima de um patamar definido no projeto do filtro.

## Filtro de Wiener

- O filtro de Wiener não utiliza um modelo do sinal recebido e do ruído, utilizando apenas a estatística do conjunto de dados.
- É um filtro linear com 8 coeficientes (7 pesos e 1 bias), projetados para diferentes condições de empilhamento de sinais.
- Utiliza um processo de simulação onde se conhece o valor desejado do parâmetro a ser estimado.
- Busca minimizar o erro médio quadrático do erro de estimação.

$$J = E \left[ (\hat{A} - A)^2 \right]$$

- Para que  $J$  seja minimizado, é necessário calcular sua derivada em função dos coeficientes  $w[k]$ , onde

$$\frac{\partial}{\partial w[k]} \mathbb{E}\{e[n]^2\} = 2 \left( \sum_{i=0}^{N-1} w[i] \mathbb{E}\{x[n-k]x[n-i]\} \right) - 2 \mathbb{E}\{x[n-k]d[n]\} .$$

- Pelo princípio da ortogonalidade, para que a função de custo  $J$  atinja o seu mínimo, sua derivada deve ser igual a zero. Logo,

$$\sum_{i=0}^{N-1} w[i] \mathbb{E} \{x[n-k]x[n-i]\} = \mathbb{E} \{x[n-k]d[n]\} \quad k = 0, 1, \dots, N-1.$$

- Utilizando as Equações de Wiener-Hopf  $R[i, k]$  e  $p[k]$ , representando autocorrelação da entrada do filtro a correlação cruzada entre a entrada do filtro e a saída desejada, respectivamente

$$R[i, k] = \frac{1}{N} \sum_{i=0}^{N-1} x[n-k]x[n-i] \quad p[k] = \frac{1}{N} \sum_{i=0}^{N-1} x[n-k]d[n]$$

obtém-se o sistema de equações lineares

$$\sum_{i=0}^{N-1} w[i] R[i, k] = p[k] \quad k = 0, 1, \dots, N-1.$$

- Em forma matricial, o sistema é dado por

$$\mathbf{w}\mathbf{R} = \mathbf{p}.$$

onde os pesos do Filtro de Wiener podem ser obtidos pela solução

$$\mathbf{w} = \mathbf{R}^{-1}\mathbf{p},$$

- Para absorver o valor médio do ruído e compensar tal contribuição na estimação, este trabalho propõe a inclusão de um elemento adicional constante igual a 1 em cada sinal de entrada como último elemento, resultando num conjunto de pesos com  $N + 1$  elementos. Ao final do processo de filtragem, o elemento  $w[N]$  é somado ao resultado, compensando o valor médio do ruído

$$\hat{A}_{FW} = \left( \sum_{i=0}^{N-1} w[i]x[i] \right) + w[N]$$