

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

ESCOLA POLITÉCNICA

DEPARTAMENTO DE ELETRÔNICA E DE COMPUTAÇÃO

**DISCRIMINADORES NEURAIIS DE PARTÍCULAS PARA  
UM DETECTOR SUBMETIDO A UMA ALTA TAXA DE  
EVENTOS**

**Autor:**

---

Thiago Ciodaro Xavier

**Orientador:**

---

José Manoel de Seixas

**Examinador:**

---

Érica Polycarpo

**Examinador:**

---

José Gabriel Rodríguez Carneiro Gomes

**Examinador:**

---

Joarez Bastos Monteiro

**Examinador:**

---

Rodrigo Coura Torres

**DEL**

**Abril de 2008**

# Sumário

<b>Lista de Figuras</b>	<b>v</b>
<b>Lista de Tabelas</b>	<b>vii</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Motivação . . . . .	1
1.2 O que foi feito . . . . .	3
1.3 Organização do documento . . . . .	4
<b>2 O Detector ATLAS</b>	<b>5</b>
2.1 A história do átomo . . . . .	5
2.1.1 O átomo hoje . . . . .	7
2.2 Detectores e aceleradores de partículas . . . . .	9
2.3 O LHC . . . . .	10
2.4 ATLAS . . . . .	12
2.4.1 Sistema de coordenadas do ATLAS . . . . .	14
2.4.2 A calorimetria do ATLAS . . . . .	14
2.4.2.1 Calorímetro eletromagnético . . . . .	15
2.4.2.2 Calorímetro hadrônico . . . . .	17
2.4.3 O sistema de filtragem do ATLAS . . . . .	21

<b>3</b>	<b>O Segundo Nível de Filtragem do ATLAS</b>	<b>24</b>
3.1	Recursos e tempo de processamento . . . . .	24
3.2	Organização lógica do segundo nível . . . . .	25
3.2.1	O construtor de RoI . . . . .	25
3.2.2	O supervisor do segundo nível . . . . .	26
3.2.3	O decisor local . . . . .	27
3.2.4	Decisor global . . . . .	27
3.3	O ambiente ATHENA . . . . .	27
3.4	O processo de seleção de eventos . . . . .	28
3.4.1	O extração de características para os calorímetros . . . . .	28
3.4.2	Algoritmo do T2Calo . . . . .	30
3.4.3	O anelador . . . . .	31
3.4.3.1	Normalização . . . . .	32
3.4.4	Algoritmo de hipótese . . . . .	33
<b>4</b>	<b>Decisores Neurais</b>	<b>35</b>
4.1	Redes neurais artificiais . . . . .	36
4.2	Anelador neural . . . . .	37
4.2.1	Metodologia do treinamento neural . . . . .	38
4.3	Compactação da informação . . . . .	39
4.3.1	Relevância em discriminadores . . . . .	39
4.3.2	Análise de componentes principais . . . . .	42
<b>5</b>	<b>Resultados</b>	<b>49</b>

5.1	Resultados de classificação . . . . .	49
<b>6</b>	<b>Conclusões</b>	<b>53</b>
6.1	Trabalhos futuros . . . . .	54
	<b>Referências</b>	<b>56</b>

# Lista de Figuras

1	Modelo do átomo proposto por Thomson. . . . .	6
2	Experiência de Rutherford. . . . .	7
3	Modelo do átomo proposto por Rutherford. . . . .	8
4	Tipos de detectores de partículas. . . . .	10
5	LHC e seus detectores. Extraído de <a href="http://lh.web.cern.ch">http://lh.web.cern.ch</a> . . . . .	11
6	Subdetectores do ATLAS. Extraído de <a href="http://www.atlas.ch">http://www.atlas.ch</a> . . . . .	12
7	Decaimento e interação das partículas nos subdetectores do ATLAS. . . . .	13
8	Sistema de coordenadas do detector ATLAS. . . . .	14
9	Detalhes da calorimetria da seção e.m. 9(a) e da seção hadrônica 9(b) . . . .	16
10	Esquema da eletrônica de leitura do LAr, extraído de <a href="http://cds.cern.ch">http://cds.cern.ch</a> . . . . .	18
11	Esquema de uma <i>trigger tower</i> do LAr, extraído de <a href="http://cds.cern.ch">http://cds.cern.ch</a> . . . . .	19
12	Eletrônica de <i>front-end</i> do TileCal, extraído de <a href="http://cds.cern.ch">http://cds.cern.ch</a> . . . . .	20
13	Esquema de uma <i>trigger tower</i> do TileCal (região sombreada). . . . .	21
14	Diagrama do sistema de filtragem do ATLAS. . . . .	23
15	Contexto do supervisor do segundo nível. . . . .	26
16	RoI de um elétron típico interagindo com as camadas dos calorímetros do ATLAS. . . . .	29
17	RoI de um jato aceito pelo primeiro nível interagindo com as camadas dos calorímetros do ATLAS. Esse tipo de jato tem uma padrão de deposição de energia muito parecido com o do elétron. . . . .	30

18	Configuração genérica dos anéis para algumas camadas dos calorímetros do ATLAS. . . . .	32
19	Exemplo de uma rede neural artificial. . . . .	37
20	Relevância dos anéis pelo método MSE e SP. . . . .	41
21	Exemplo de componentes principais extraídos de uma base de dados genérica. . . . .	42
22	Curva de carga dos componente principais extraídos de forma não segmentada. . . . .	43
23	Relevância por SP dos componentes principais extraídos de forma não segmentada. . . . .	44
24	Histograma e representação no espaço de anéis do componente principal mais relevante. . . . .	45
25	Curva de carga de algumas das camadas dos calorímetros e.m. e hadrônico. . . . .	46
26	Relevância dos componentes principais extraídos de forma segmentada. . . . .	47
27	Curvas ROC dos discriminadores com anéis selecionados por relevância. . . . .	50
28	Curvas ROC dos discriminadores com anéis selecionados por relevância. . . . .	51
29	Curvas ROC dos melhores discriminadores, com os tipos de compactação empregados. . . . .	52

# Lista de Tabelas

1	Forças da natureza e suas partículas transportadoras. . . . .	9
2	Granularidade da seção e.m.. . . . .	17
3	Granularidade da seção hadrônica. . . . .	19
4	Descrição das variáveis do T2Calo. $E_{i \times j}$ representa a energia numa janela de $i$ células em $\eta$ e $j$ células em $\phi$ . $E_i$ representa a $i$ -ésima célula energética.	31
5	Número de anéis para cada camada dos calorímetros e.m. e hadrônico . . .	31
6	Esquema da normalização do tipo seqüencial. . . . .	33
7	Tipos de compactação e seus respectivos números de variáveis selecionadas.	45
8	Variáveis que foram ditas relevantes por todos os métodos. . . . .	48
9	Produtos SP dos melhores discriminadores, considerando um falso alarme de 5% e o SP máximo alcançado pelo discriminador. . . . .	52

# 1 Introdução

Nos dias de hoje, a capacidade de armazenamento e comunicação entre sistemas vem crescendo cada vez mais. Assim, a atividade de extrair informação relevante de uma massa de dados é, também, cada vez mais complexa. Podemos citar o reconhecimento e a classificação de padrões como uma técnica que vem sendo aplicada nesse problema, podendo-se destacar as redes neurais artificiais nesse contexto.

## 1.1 Motivação

Na fronteira entre a Suíça e a França está o laboratório científico CERN (do francês *Conseil Européen pour la Recherche Nucléaire*) [1]. Grandes experimentos de física são realizados no CERN, contando com a colaboração de institutos de pesquisa de todo o mundo. Através desta parceria, ferramentas para comprovação experimental de teorias são desenvolvidas por seus membros, fornecendo uma infra-estrutura de alta qualidade e indispensável para o processo científico. Os experimentos mais destacados no CERN são aqueles que se dedicam à física corpuscular, utilizando, principalmente, aceleradores de partículas.

Atualmente, o CERN e seus colaboradores estão construindo aquele que será o maior acelerador circular de partículas do mundo, o LHC (do inglês *Large Hadron Collider*). O LHC iniciará suas colisões em meados de 2008, e ficará operativo por mais 10 anos. Seu anel fica localizado a aproximadamente 175 m abaixo da terra e tem 27 km de circunferência. As partículas aceleradas pelo LHC podem alcançar cerca de 99,999999 % da velocidade da luz [2], colidindo frontalmente em pontos de colisão determinados. A aceleração das partículas em sentidos opostos permite que toda essa energia seja convertida

em novas partículas durante a colisão.

Um dos principais objetivos do LHC é a comprovação da existência de uma partícula (bóson de Higgs) que explicaria o mecanismo que dá massa às partículas elementares. Esta é a única partícula do modelo padrão de física de partículas que ainda não foi observada experimentalmente [3]. Por ser uma partícula extremamente instável, o bóson de Higgs só pode ser identificado através de seus decaimentos, o que chamamos de assinaturas. Existe um conjunto de assinaturas que levam ao Higgs, sendo o elétron o principal mensageiro de sua produção. Porém, jatos de partículas ordinárias mascaram elétrons, dificultando sua identificação. Com o auxílio do LHC, os físicos ainda esperam comprovar outras teorias físicas e, também, diminuir a incerteza de medidas realizadas em experimentos anteriores. Como a física de interesse do LHC é muito rara, é necessária uma alta taxa de colisões a fim de garantir estatística suficiente. As colisões do LHC ocorrerão a uma frequência de 40 MHz, podendo chegar a 1 GHz devido à probabilidade de colisões inelásticas acontecerem.

Nos pontos de colisão do LHC estão localizados detectores altamente especializados. Os detectores dão uma visão detalhada da colisão, indo desde características energéticas a imagens com alta resolução da trajetória das partículas resultantes. De posse das informações proporcionadas pelos detectores, podemos caracterizar cada sub-partícula criada após a colisão. Dentre os detectores do LHC está o ATLAS (do inglês *A Toroidal LHC ApparatuS*) [4]. Aliada à alta taxa de eventos do LHC, a informação gerada por cada colisão é de cerca de 1,5 MB [5], o que irá gerar uma elevada quantidade de informação a ser processada pelo sistema. Desta forma, a taxa de eventos deve ser reduzida por um sistema *online* de filtragem.

A fim de contornar o problema da enorme quantidade de informação gerada, foi concebido para o ATLAS um sistema de filtragem *online*, o Sistema de *Trigger*. Esse sistema de filtragem de eventos é composto por três níveis conectados em cascata, com complexidade e tempo de processamento crescentes. A cada nível há uma seleção de eventos, o que reduz a alta taxa de eventos original.

O primeiro nível de *Trigger* deve diminuir a taxa de eventos de 40 MHz para não mais

que 100 kHz. O segundo nível de *Trigger* refina a decisão do primeiro nível, entregando ao nível seguinte uma taxa de 1 kHz. O terceiro nível, por fim, analisará a informação proveniente de todos os sub-detectores e apresentará em sua saída uma taxa de 100 Hz, para armazenamento em mídia permanente. Posteriormente, uma análise *offline* mais aprofundada pode ser feita sobre os eventos selecionados.

## 1.2 O que foi feito

O presente trabalho se baseia na informação de calorimetria do ATLAS para otimizar o segundo nível de filtragem. Discriminadores baseados em redes neurais artificiais e técnicas de compactação da informação serão utilizados, a fim de aprimorar a separação elétron/jato, pois se trata da assinatura mais freqüente do bóson de Higgs (para cada 25 mil elétrons identificados pelo primeiro nível de filtragem, somente 1 é realmente um elétron [6]).

Um sistema baseado em redes neurais artificiais foi desenvolvido para a discriminação do canal elétron/jato [7]. Esse sistema utiliza as informações de calorimetria do evento, no segundo nível de filtragem, para realizar um mapeamento topológico em anéis concêntricos de deposição de energia, alcançando uma elevada capacidade de discriminação.

Estudos para aumentar ainda mais a compactação da informação, mas, ainda assim, manter uma eficiência de detecção elevada, foram realizados. Técnicas de pré-processamento estatístico como PCA (do inglês *Principal Component Analysis*) e estudos de relevância das variáveis de entrada do discriminador foram explorados.

Os resultados de discriminação alcançados mostraram que discriminadores neurais conseguem uma performance elevada, mesmo considerando as condições inóspitas de operação do ATLAS, onde a taxa de eventos é elevada (limitando o tempo de processamento) com milhares de canais de informação e demandando resultados elevados.

## 1.3 Organização do documento

No Capítulo 2, será apresentada uma introdução ao mundo da física de partículas, mostrando os principais conceitos. Também serão detalhadas as principais características do LHC e do ATLAS, bem como o sistema de filtragem *online* desenvolvido para o ATLAS.

No Capítulo 3, o foco será o segundo nível de filtragem do ATLAS, mostrando os sistemas de aquisição de dados dos calorímetros e os algoritmos de extração de características desse nível: o anelador e o T2Calo, que é o algoritmo de referência deste nível. O atual algoritmo de hipótese também será detalhado.

No Capítulo 4, será apresentada a proposta de substituição do algoritmo de hipótese existente por uma rede neural, bem como as técnicas de pré-processamento que serão feitas para diminuir a dimensão dos dados provenientes do anelador.

No Capítulo 5 serão mostrados os resultados e as análises dos discriminadores neurais desenvolvidos.

Por fim, o Capítulo 6 trará as conclusões do trabalho, bem como os trabalhos e estudos futuros que podem ser feitos em cima dos resultados aqui apresentados.

## 2 O Detector ATLAS

Física de partículas é o ramo da física que estuda as partículas elementares que compõem a matéria e a forma com que elas interagem. Neste capítulo, será apresentado o ambiente da física de partículas e o CERN, bem como os principais conceitos e instrumentos que levaram ao seu desenvolvimento.

### 2.1 A história do átomo

A primeira referência ao elemento básico da matéria foi feita pelo filósofo grego Demócrito, definindo-o como o átomo, que significa indivisível em grego. A idéia do átomo ficou esquecida durante séculos, até ser retomada pelos pensadores e cientistas do período renascentista. Pode-se dizer que o pai da teoria moderna do átomo foi o padre jesuíta chamado Rudjer Bosovich, baseando sua teoria principalmente na mecânica clássica newtoniana [8]. Posteriormente, sua teoria foi desenvolvida pelos irmãos Avogadro. Mas foi com os estudos de John Dalton que foi criado o primeiro modelo para o átomo, como uma minúscula esfera maciça e indivisível. Esta teoria perdurou por mais de um século [8].

Entre 1813 e 1834, Michael Faraday estudou a relação entre as quantidades de materiais em transformações químicas e a quantidade de eletricidade necessária para realizar essas transformações. Esses estudos evoluíram até que, em 1891, a unidade mais simples de eletricidade foi determinada e denominada elétron [8]. A descoberta de partículas com carga elétrica fez com que o modelo atômico de Dalton ficasse superado. Em 1897, Thomson idealizou um experimento para medir a carga elétrica do elétron [9]. Com base em seu experimento, e considerando o átomo eletricamente neutro (com quantidades iguais de partículas positivas e negativas), ele representou o átomo como uma esfera uniforme,

de carga positiva, incrustada de elétrons (Figura 1).

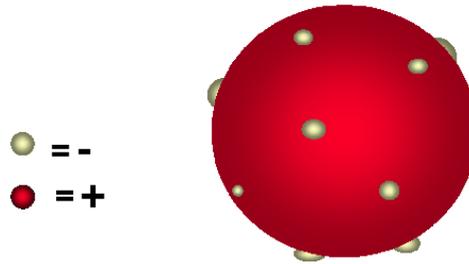


Figura 1: Modelo do átomo proposto por Thomson.

Em 1911, realizando experiências de bombardeio de lâminas de ouro com partículas alfa (partículas de carga positiva), Rutherford fez uma importante constatação: a grande maioria das partículas atravessava diretamente a lâmina, algumas sofriam pequenos desvios e outras, em número muito pequeno (uma em cem mil), sofriam grandes desvios em sentido contrário (detalhes na Figura 2) [9]. Com essas observações, chegou-se aos seguintes resultados:

- No átomo, existem grandes espaços vazios. A maioria das partículas o atravessava sem sofrer desvios
- No centro do átomo existe um núcleo muito pequeno e denso. Algumas partículas alfa colidiam com esse núcleo e voltavam, sem atravessar a lâmina.
- O núcleo tem carga elétrica positiva. As partículas alfa que passavam perto dele eram repelidas e, por isso, sofriam desvio em sua trajetória.

Pelo modelo atômico de Rutherford, o átomo é constituído por um núcleo central, dotado de cargas elétricas positivas (prótons), envolvido por uma nuvem de cargas elétricas negativas (elétrons). Rutherford demonstrou, ainda, que praticamente toda a massa do átomo fica concentrada na pequena região do núcleo. A Figura 3 mostra o modelo do átomo proposto por Rutherford.

Dois anos depois de Rutherford ter criado o seu modelo, o cientista dinamarquês Niels Bohr o completou, criando o que hoje é chamado modelo planetário [9]. Para Bohr, os

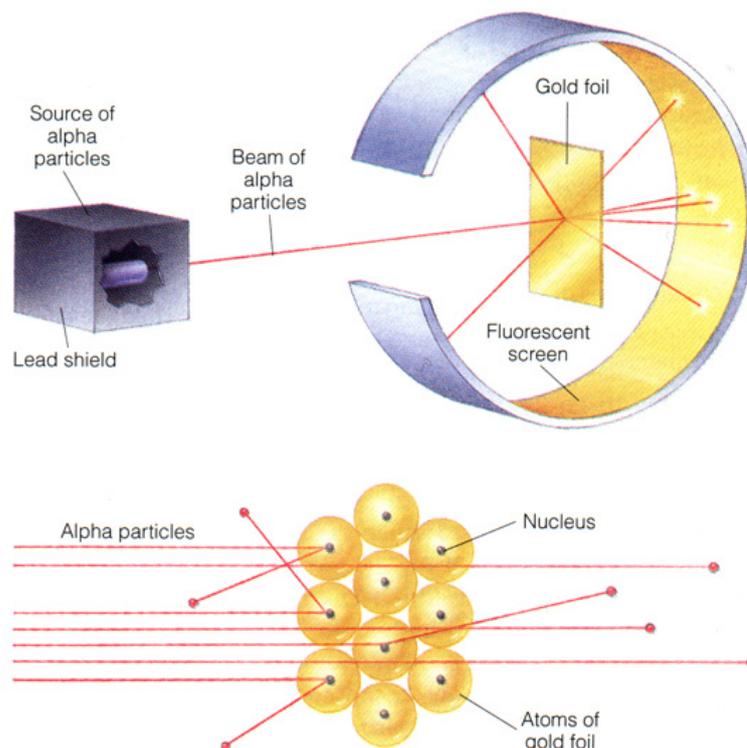


Figura 2: Experiência de Rutherford.

elétrons giravam em órbitas circulares, ao redor do núcleo. Depois desses, novos estudos foram feitos e novos modelos atômicos foram criados. O modelo que representa o átomo como tendo uma parte central chamado núcleo, contendo prótons e nêutrons, serve para explicar um grande número de observações sobre os materiais.

Por estarem envoltos por uma barreira de potencial (causada pelo movimento dos elétrons), prótons e nêutrons não podiam ser observados naturalmente. Para extraí-los seria necessário energia suficiente para romper essa barreira de potencial [10]. Os físicos atuais usam feixes de partículas aceleradas. Tal feixe pode ser acelerado contra um alvo fixo, ou contra outras partículas aceleradas em sentido contrário.

### 2.1.1 O átomo hoje

O advento de novas ferramentas para a observação física e novas teorias mudou a forma de se entender o átomo. O advento da teoria quântica, o princípio da incerteza de Heisenberg, os estudos de Louis de Broglie e Schrödinger, dentre outros físicos, moldaram o átomo como um núcleo composto de partículas pesadas (prótons e nêutrons) envolto por

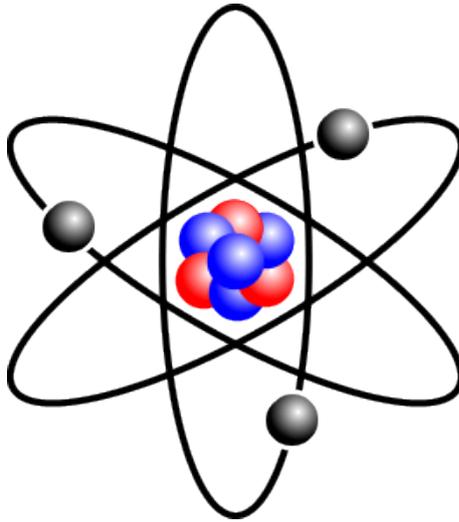


Figura 3: Modelo do átomo proposto por Rutherford.

uma nuvem de elétrons. Não se pode determinar com exatidão a posição do elétron em volta do núcleo: existe uma densidade de probabilidade associada à posição do elétron [10].

Já em relação ao núcleo, a diferença entre a massa entre átomos de um mesmo elemento foi explicada pela existência do nêutron: partícula eletricamente neutra com massa próxima do próton. Porém, ela não explica o porquê que prótons de um mesmo núcleo atômico não se repelem através da força eletromagnética, uma vez que todos os prótons têm carga positiva. Alguma força, hoje chamada de força forte, deve atuar entre essas partículas, quando a distância entre elas é muito pequena, de forma a manter a estabilidade do núcleo.

Os físicos procuravam chegar num modelo onde as partículas observáveis eram uma combinação de outras partículas, ditas fundamentais. Uma série de novas partículas e novas classificações foram observadas. Múons, mésons, píons, dentre outras, são exemplos dessas novas partículas observadas, enquanto hádrons e léptons são exemplos de classificações de partículas. Porém, não se conseguia chegar num modelo fechado que classificasse as partículas.

Em meados da década de 60, foi proposta por Murray Gell-Mann a existência de um novo tipo de partículas elementares, chamadas de quarks. Essas partículas iriam classificar, junto com léptons, todas as outras partículas. Inicialmente, a teoria previa a existência de 3 tipos de quarks (em inglês *up*, *down* e *strange*), mas aperfeiçoamentos

na teoria e observações posteriores aumentaram para 6 esse número (em inglês *charm*, *bottom*, ou *beauty*, e *top*, ou *truth*). De acordo com a teoria, os quarks não são encontrados sozinhos, mas em duplas (mésons) ou triplas (bárions). Vale ressaltar que cada quark tem a sua anti-partícula, ou antiquark, assim como qualquer outra partícula.

Hoje, o modelo padrão prevê a existência de partículas carregadoras de força. As 4 forças existentes podem ser vistas na Tabela 1, assim como suas partículas transportadoras. Destas, somente a partícula transportadora da força gravitacional ainda não foi comprovada.

Tabela 1: Forças da natureza e suas partículas transportadoras.

Partícula	Força
Fóton	Eletromagnética
Bósons W e Z	Nuclear fraca
Glúon	Nuclear forte
Gráviton	Gravitacional

## 2.2 Detectores e aceleradores de partículas

A aceleração de partículas é um processo conceitualmente simples: partículas eletricamente carregadas são submetidas a pulsos eletromagnéticos, no interior de tubos. Quando envolvidos por campos eletromagnéticos variantes, as partículas adquirem força (aceleração) [11].

Aceleradores de partículas podem ser circulares ou lineares. Em aceleradores circulares, as partículas são injetadas dentro do anel e ficam circulando até atingir a energia desejada. Seu formato circular permite que vários experimentos sejam realizados nos diversos pontos de colisão ao longo de sua circunferência. Porém, quanto maior a energia necessária na colisão, maior deve ser o anel do acelerador. Já em aceleradores lineares, as partículas são injetadas por uma extremidade e percorrem uma reta até colidirem com um alvo fixo ou com outras partículas.

Após a energia desejada ser alcançada, as partículas colidem com um alvo fixo ou com outras partículas aceleradas em sentido contrário. Para alvos fixos, temos detectores

com formas cônicas que ficam posicionados logo atrás do alvo. Já para colisão com outras partículas, detectores cilíndricos são os mais utilizados. A Figura 4 ilustra esses tipos de detectores.

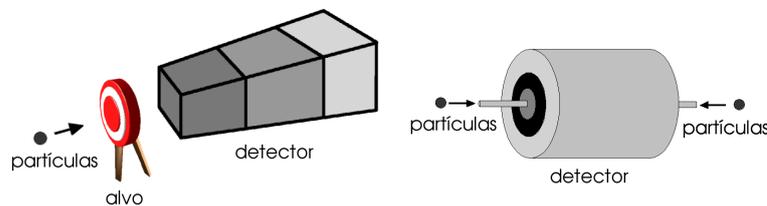


Figura 4: Tipos de detectores de partículas.

A informação gerada por cada colisão deve ser coletada e analisada por especialistas. O resultado é a identificação de partículas, fenômenos e a comprovação ou a eliminação de teorias físicas. Diversos detectores são montados para a observação e coleta dessas informações. Existem detectores de traços, para visualizar a trajetória percorrida pelas partículas resultantes da colisão; detectores de calorimetria, para a identificação da energia depositada por uma partícula ao atravessar o detector; e detectores de múons, para a absorção dessas partículas.

## 2.3 O LHC

O CERN [1] não é mais um laboratório apenas europeu, mas um centro internacional de estudos. Países não-europeus não podem se tornar países membros, mas podem contribuir financeiramente e intelectualmente em diversos experimentos físicos e áreas associadas, como as engenharias eletrônica, de computação e de materiais.

Neste contexto, o Brasil mantém uma colaboração com o CERN, principalmente nas áreas de física e engenharia eletrônica e de computação.

No CERN está sendo desenvolvido aquele que será o maior colisionador de partículas já construído, o LHC (do inglês *Large Hadron Collider*) [2]. O LHC é um acelerador de partículas circular com 27 km de perímetro. Nele, as partículas aceleradas poderão alcançar uma velocidade muito próxima à da luz.

Dentro do acelerador, dois feixes de partículas viajam em direção contrária, com elevada energia, até pontos de colisão definidos ao longo do LHC. Dois tubos mantidos em um vácuo quase ideal guiam os feixes, com ajuda de um sistema de poderosos ímãs, capazes de acelerar e desviar as partículas. Os campos magnéticos são tão intensos, que são precisos supercondutores operando a uma temperatura de  $-271^{\circ}\text{C}$ , resfriados através de uma sistema de distribuição de hélio líquido. Enquanto um sistema de ímãs desvia e acelera as partículas, outro comprime o feixe de forma a aumentar a densidade de partículas e, conseqüentemente, a probabilidade de uma colisão frontal com o outro feixe, acelerado em direção oposta.

No caso do LHC, pacotes de prótons serão utilizados na colisão. Em alguns pontos de colisão, ao longo do LHC, existem detectores de partículas (Figura 5), que extrairão as informações importantes da colisão. Dentre esses detectores temos o detector ATLAS.

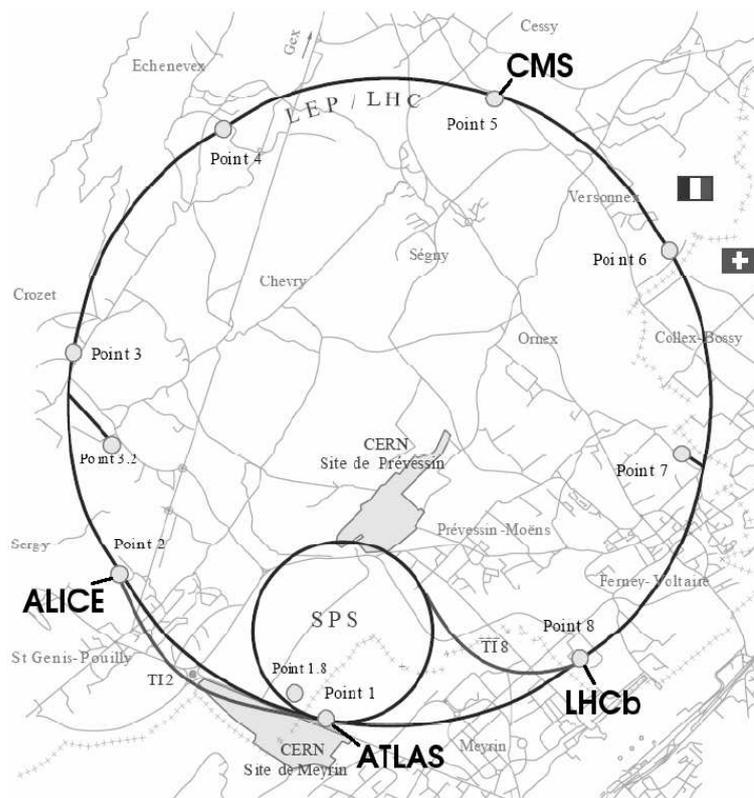


Figura 5: LHC e seus detectores. Extraído de <http://lhc.web.cern.ch>.

## 2.4 ATLAS

O ATLAS (do inglês *A Toroidal LHC Apparatus*) é um detector cilíndrico de partículas posicionado num dos pontos de colisão do LHC [4]. Seu objetivo é extrair o máximo de informação das colisões próton-próton que ocorrerão em seu interior. A extração dessa informação é feita por seus subdetectores (a partir do ponto de impacto): detector de traços, calorímetros eletromagnético e hadrônico e câmaras de múons (Figura 6).

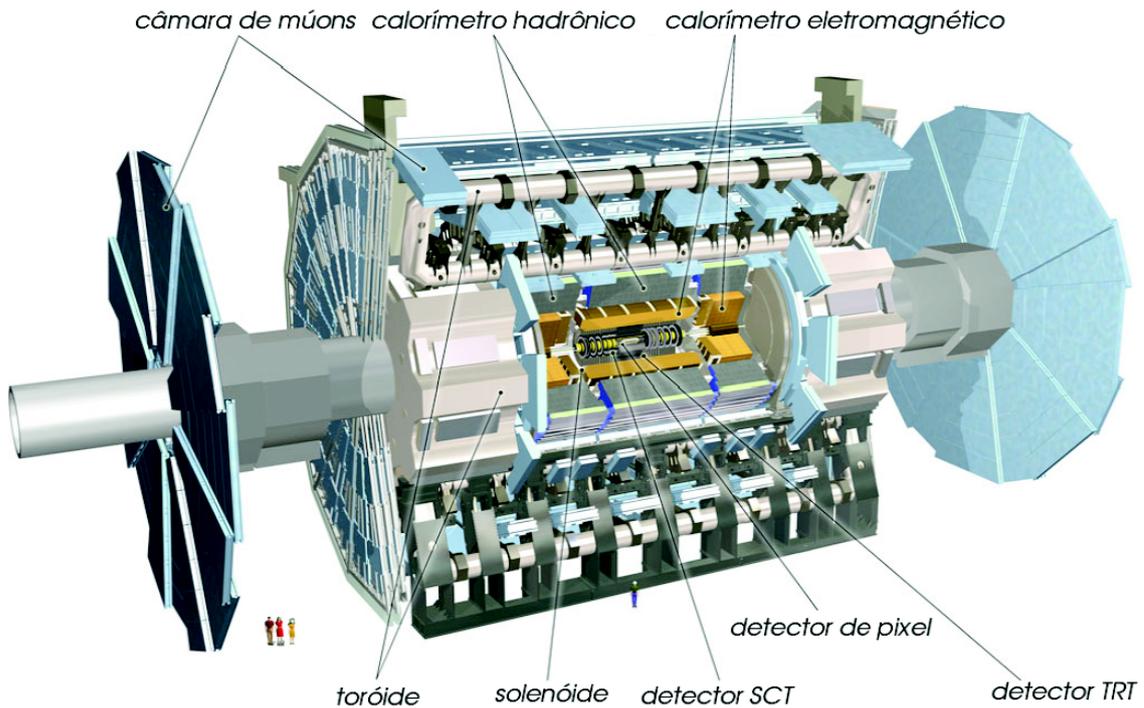


Figura 6: Subdetectores do ATLAS. Extraído de <http://www.atlas.ch>.

- Detector de traços: tem o objetivo de determinar a rota e curvatura das partículas. As partículas são expostas a um campo magnético que curva a sua trajetória, dando informações sobre a carga e o momento da partícula [12].
- Calorímetro: o calorímetro do ATLAS é dividido em duas seções. A seção eletromagnética é responsável pela determinação da energia de elétrons, pósitrons e fótons [13]. Estas partículas decaem formando chuviscos de outros elétrons, pósitrons e fótons, dentro desta seção. Já a seção hadrônica é responsável pela medição da energia total depositada por prótons, nêutrons e mésons. É composto por mate-

rial pesado, que força o decaimento dos hádrons num chuveiro hadrônico, bem mais largo e profundo que o eletromagnético [14].

- Câmaras de múon: é a parte mais externa do detector, tendo como objetivo a detecção de múons [15]. Essas partículas não são bem absorvidas pelos outros subdetectores.

A Figura 7 ilustra a interação entre algumas partículas com o detector.

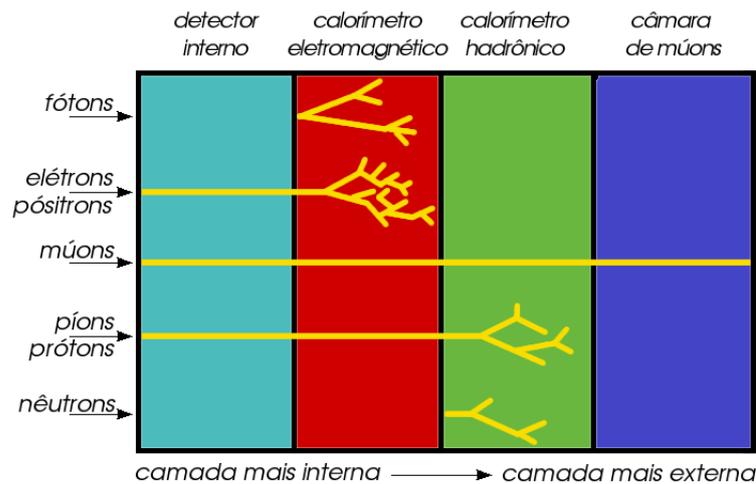


Figura 7: Decaimento e interação das partículas nos subdetectores do ATLAS.

Um dos objetivos do LHC (e do ATLAS) é a comprovação da existência do bóson de Higgs. Essa partícula teórica, prevista no modelo padrão, nunca foi comprovada experimentalmente. Segundo o modelo, o mecanismo de Higgs é o mecanismo que dá massa a todas as partículas elementares, e o bóson de Higgs é uma partícula escalar maciça e hipotética, prevista para validar o modelo padrão [16]. Sua detecção é difícil, uma vez que é necessária uma energia muito grande para sua observação. Ainda assim, por ser uma partícula muito energética, sua detecção se dá através de seus decaimentos, chamados de assinaturas do bóson de Higgs.

Numa assinatura, três fatores são importantes: o tipo de partícula, sua carga e sua energia. Dentre as assinaturas do bóson de Higgs, os físicos têm que escolher as mais fáceis de se detectar com a tecnologia atual de detectores. Assim, léptons (elétrons e múons) representam as principais assinaturas.

### 2.4.1 Sistema de coordenadas do ATLAS

O ATLAS é um detector cilíndrico, colocado ao longo do LHC. Sua estrutura é disposta ao redor do ponto de colisão. O sistema de coordenadas utilizado no experimento, então, é um sistema que acompanha a direção dos feixes de partículas gerados após a colisão [17]. O sistema escolhido é, então, o cilíndrico, mas com algumas modificações. As coordenadas são  $\eta$ ,  $\phi$  e  $z$ , em contraposição a  $x$ ,  $y$  e  $z$ .  $\eta$  e  $\phi$  se relacionam com  $x$  e  $y$  através das seguintes relações não-lineares:

$$\phi = \text{atan}\left(\frac{x}{y}\right)$$

$$\eta = -\log\left(\tan\left(\frac{\theta}{2}\right)\right)$$

A Figura 8 ilustra o sistema. Note que  $\phi$  representa a rotação do feixe em relação à direção de propagação dos pacotes de prótons ( $z$ ). Já  $\eta$  (chamada de pseudo-rapidez) representa a projeção do feixe de partícula após a colisão, num plano  $zx\phi$  específico.

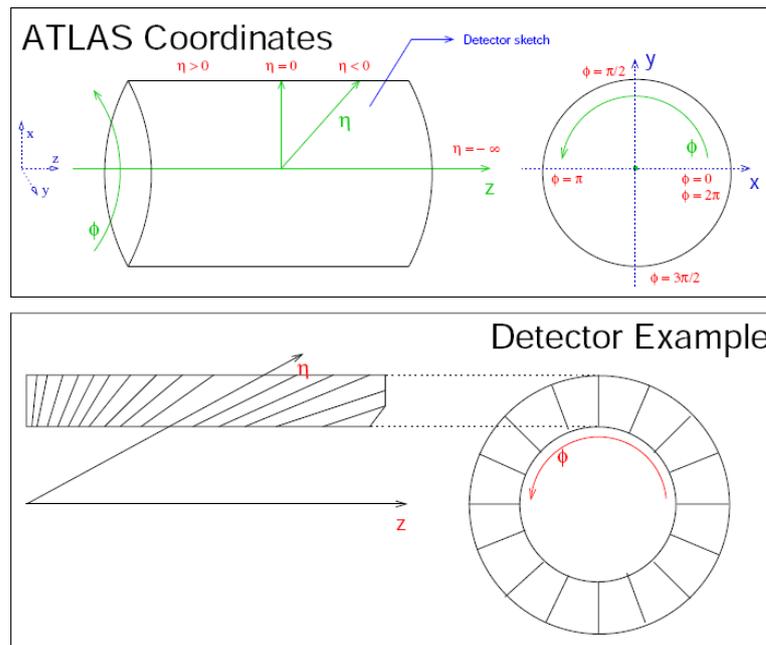


Figura 8: Sistema de coordenadas do detector ATLAS.

### 2.4.2 A calorimetria do ATLAS

Os calorímetros do ATLAS têm a função de medir a energia das partículas geradas após a colisão [18]. Um calorímetro é um bloco de matéria que interage com as partículas,

fazendo-as decair em outras partículas menos energéticas e mais estáveis. Ao decair, essas partículas liberam energia, medida pelo calorímetro [18].

Calorímetros são amplamente utilizados em pesquisas com física de partículas. Basicamente, calorímetros têm as seguintes vantagens:

- Calorímetros são sensíveis a partículas neutras e carregadas.
- Quanto maior a energia da partícula, mais preciso é o resultado. Essa característica não acontece em outros detectores.
- Para conter o desenvolvimento de cascatas dos objetos a serem medidos, a profundidade dos calorímetros aumenta logaritmicamente com a energia, permitindo o desenvolvimento de detectores compactos.
- Não necessitam de campos magnéticos.
- Podem ser segmentados, permitindo uma boa visualização do evento e uma medição de energia mais acurada.
- Resposta rápida, fator importante num ambiente de alta taxa de eventos.

O sistema de calorimetria do ATLAS é dividido em dois calorímetros (ou um calorímetro com duas seções): o calorímetro eletromagnético (e.m.) e o calorímetro hadrônico. A Figura 9 mostra detalhes da calorimetria do detector ATLAS, para as seções e.m. e hadrônica. As principais características desses calorímetros são:

#### **2.4.2.1 Calorímetro eletromagnético**

O calorímetro e.m. do ATLAS é composto de chumbo como material absorvedor das partículas, e chumbo imerso em argônio líquido como elemento amostrador. Os eletrodos amostram a energia da partícula através da ionização das células do calorímetro. Por ser composto por argônio líquido, o calorímetro e.m. também é conhecido como LAr (do inglês *Liquid Argon*).

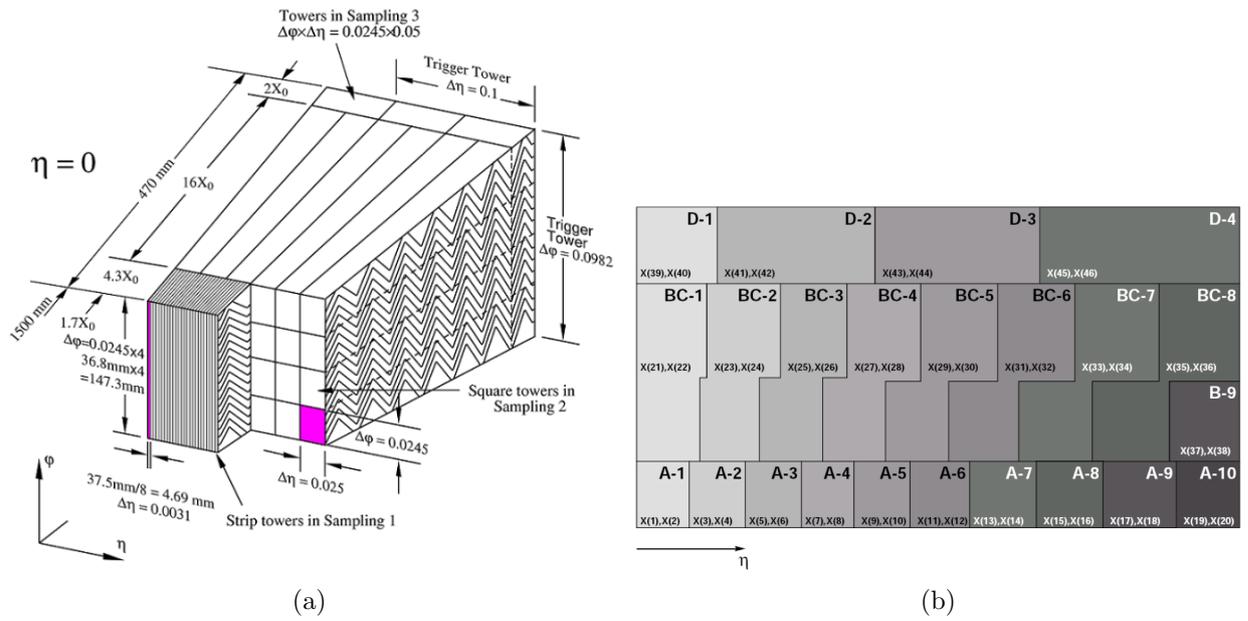


Figura 9: Detalhes da calorimetria da seção e.m. 9(a) e da seção hadrônica 9(b)

O LAr é dividido em 3 camadas, com diferentes características. Cada camada possui uma granularidade própria, ajudando a determinar detalhes dos objetos que interagem com o detector. A segunda camada e.m. é a maior e a que absorve mais energia, sendo a mais importante desse calorímetro. Há ainda, antes da primeira camada e.m., uma camada fina chamada de *pre-sampler*. Esta camada tem como objetivo a recuperação de parte da energia acumulada em material morto (cabos, encapamentos, etc.). O LAr é dividido em duas partes: o barril e a tampa. Juntos, eles completam o espaço ao redor da colisão até um valor de  $\eta = 3, 2$ .

A granularidade do calorímetro e.m. é constante em relação à rotação (constante em  $\phi$ ) mas variável em relação a  $\eta$ . Uma maior granularidade implica num maior número de células de deposição de energia numa mesma área do detector. A Tabela 2 mostra a granularidade do LAr por camada e por valor de  $\eta$ .

Ao todo, o LAr tem, aproximadamente, 190 mil canais eletrônicos de leitura da informação gerada [19]. Essa eletrônica de leitura do LAr pode ser dividida em 3 partes (Figura 10). A primeira parte se localiza dentro dos sistemas de amostragem, composta pela placa mãe. Pode-se modelar a resposta do detector à incidência de uma partícula por um pulso triangular de corrente. Esses pulsos são transportados até a *front end crate*, localizada na borda do detector. Esta segunda parte da eletrônica de leitura é responsável

Tabela 2: Granularidade da seção e.m..

Camada	Peça	$\eta_{\text{início}}$	$\eta_{\text{fim}}$	$\Delta\eta \times \Delta\phi$	$N_{\eta} \times N_{\phi}$
<i>Pre-sampler</i>	Barril	0	1,5	$0,025 \times 0,1$	$4 \times 1$
	Tampa	1,5	1,8	$0,025 \times 0,1$	$4 \times 1$
Camada 1	Barril	0	1,4	$0,003125 \times 1$	$32 \times 1$
		1,4	1,475	$0,025 \times 0,025$	$4 \times 4$
	Tampa	1,375	1,5	$0,025 \times 0,1$	$1 \times 4$
		1,5	1,8	$0,003125 \times 0,1$	$32 \times 1$
		1,8	2,0	$0,004167 \times 0,1$	$24 \times 1$
		2,0	2,4	$0,00625 \times 0,1$	$16 \times 1$
		2,4	2,5	$0,025 \times 0,1$	$4 \times 1$
2,5	3,2	$0,1 \times 0,1$	$1 \times 1$		
Camada 2	Barril	0	1,4	$0,025 \times 0,025$	$4 \times 4$
		1,4	1,475	$0,075 \times 0,025$	$1 \times 4$
	Tampa	1,375	2,5	$0,025 \times 0,025$	$4 \times 4$
		2,5	3,2	$0,1 \times 0,1$	$1 \times 1$
Camada 3	Barril	0	1,35	$0,05 \times 0,025$	$2 \times 4$
	Tampa	1,5	2,5	$0,05 \times 0,025$	$2 \times 4$

pela calibração do sistema, condicionamento do sinal, digitalização e armazenamento dos dados em memória do tipo *pipeline*.

Os sinais são, também, somados para formar uma *trigger tower*, formando um sinal que será utilizado pelo primeiro nível de filtragem, abordado na Seção 2.4.3. A Figura 11 mostra uma *trigger tower* para o LAr. Por fim, a terceira parte, já localizada fora do detector, é responsável pelo armazenamento dos dados para o segundo nível de filtragem, pela comunicação e interface com o sistema de aquisição de dados.

#### 2.4.2.2 Calorímetro hadrônico

O calorímetro hadrônico do ATLAS é composto por uma liga com aço e tem telhas cintiladoras (fibras óticas) como material amostrador. As telhas cintilam quando uma partícula interage com o aço, levando essa informação luminosa, através de fibras óticas, até células fotomultiplicadoras, que converterão o sinal luminoso em elétrico. Por ser composto de telhas, também é chamado de TileCal (do inglês *Tile Calorimeter*). O TileCal é posicionado logo após o LAr, sendo dividido em 3 partes: o barril, o barril estendido e a tampa. Assim como no LAr, a tampa é formada por argônio líquido.

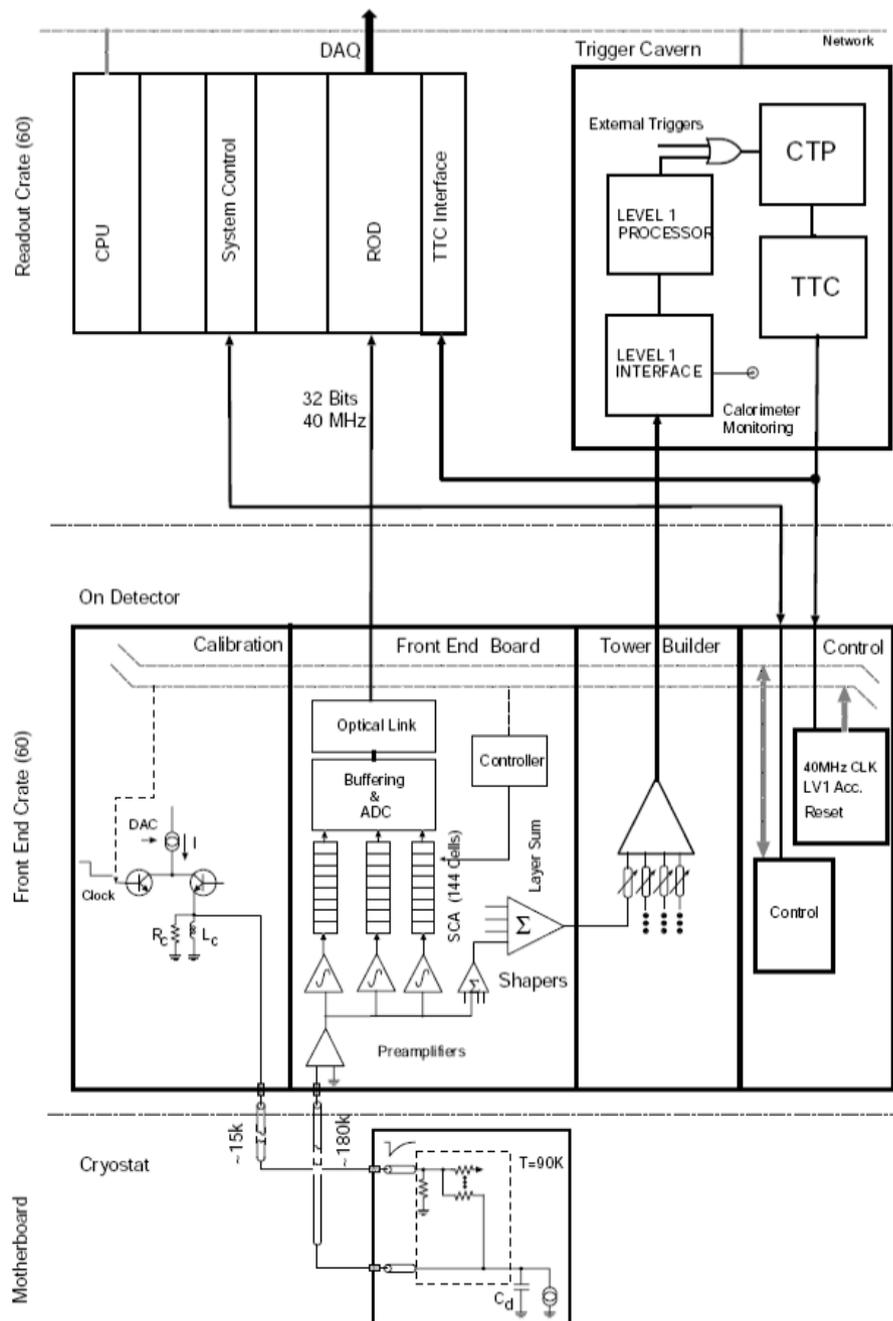


Figura 10: Esquema da eletrônica de leitura do LAr, extraído de <http://cds.cern.ch>.

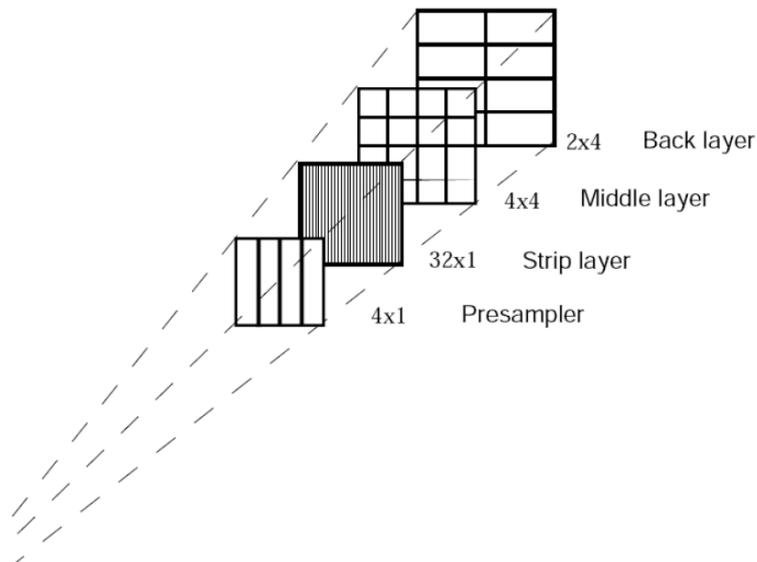


Figura 11: Esquema de uma *trigger tower* do LAr, extraído de <http://cds.cern.ch>.

O TileCal também é composto por 3 camadas, com diferentes granularidades e células com diferentes tamanhos em relação a  $\eta$ . A Tabela 3 mostra a granularidade das camadas do TileCal e sua variação com  $\eta$ .

Tabela 3: Granularidade da seção hadrônica.

Camada	Peça	$\eta_{\text{início}}$	$\eta_{\text{fim}}$	$\Delta\eta \times \Delta\phi$	$N_{\eta} \times N_{\phi}$
Camadas 1 e 2	Barril (TileCal)	0	1,0	$0,1 \times 0,1$	$2 \times 2$
	Barril Ext. (TileCal)	0,8	1,7	$0,1 \times 0,1$	$2 \times 2$
	Tampa (LAr)	1,5	2,5	$0,1 \times 0,1$	$2 \times 2$
		2,5	3,2	$0,2 \times 0,2$	$1 \times 1$
Camada 3	Barril (TileCal)	0	1,0	$0,2 \times 0,1$	$1 \times 2$
	Barril Ext. (TileCal)	0,8	1,7	$0,2 \times 0,1$	$1 \times 2$
	Tampa (LAr)	1,5	2,5	$0,1 \times 0,1$	$2 \times 2$
		2,5	3,2	$0,2 \times 0,2$	$1 \times 1$

Assim como no calorímetro e.m., o TileCal possui uma grande quantidade de canais eletrônicos de leitura. Aproximadamente 10 mil sinais com a informação de energia depositada pela partícula devem ser processados e disponibilizados ao sistema de aquisição de dados. A eletrônica de leitura do TileCal pode ser dividida em 2 partes: *front-end* e *back-end*. A eletrônica de *front-end* localiza-se no detector, recebendo os sinais luminosos das telhas cintilantes e convertendo-os em sinal elétrico. A amplitude do sinais elétricos é proporcional à energia depositada pela partícula ao atravessar o calorímetro hadrônico [20].

O sinal luminoso é coletado e guiado por fibras óticas até as células foto-multiplicadoras (PMT do inglês *photomultiplier tube*), que converterão o sinal luminoso em um sinal de corrente (uma PMT pode ser idealizada como um pulso de corrente).

Após a PMT, o sinal é conformado, de forma a manter uma largura de pulso específica e controlada. Dois ganhos são aplicados ao sinal, com uma relação de 1:64 e apresentados a um digitalizador. A utilização de dois ganhos, um alto e um baixo, fornece uma faixa dinâmica adequada ao sistema. Um sistema integrador também é utilizado para fins de calibração. O *shaper*, os ganhos e o integrador são implementados pelo circuito 3-in-1 [20], que pode ser visto na Figura 12.

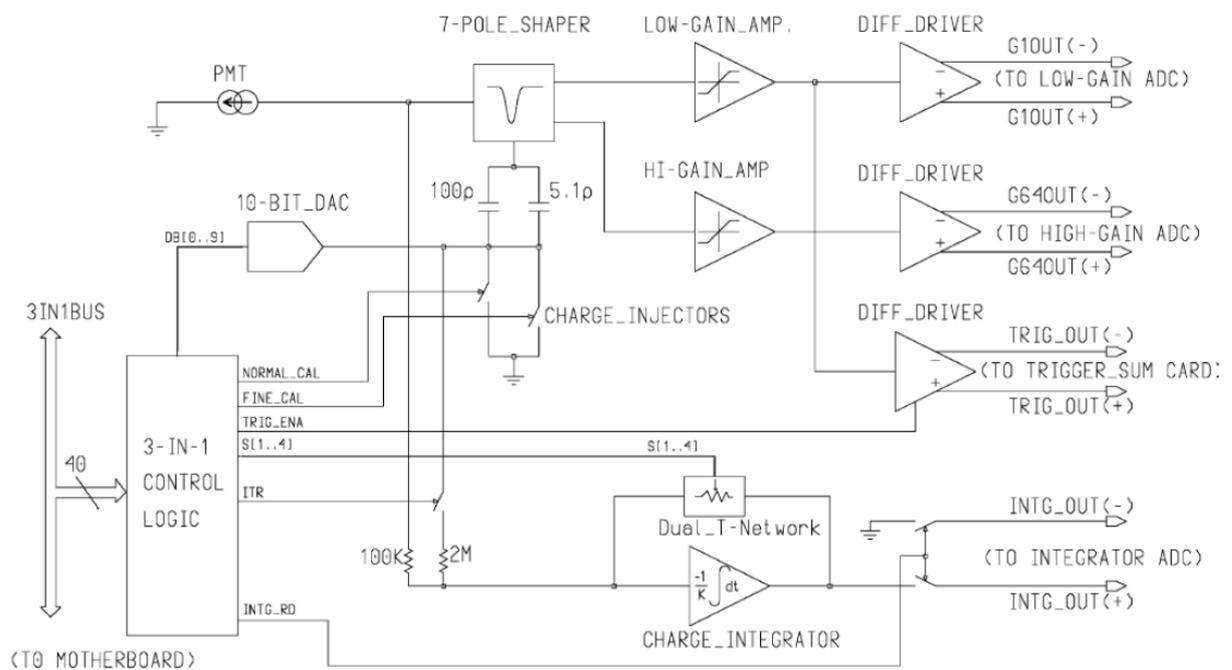


Figura 12: Eletrônica de *front-end* do TileCal, extraído de <http://cds.cern.ch>.

Uma cópia do sinal conformado é transmitido até placas somadoras, que irão formar um sinal de *trigger tower* [21]. Uma placa somadora soma até 6 sinais de diferentes PMT, formando um sinal por *trigger tower* que é transmitido até o primeiro nível de filtragem, assim como o calorímetro e.m. (a Figura 13 mostra uma *trigger tower* do TileCal). Um outro sinal, utilizado para a detecção de múons de baixa energia e raios cósmicos também é gerado [22]. Esse sinal é proporcional à energia depositada pela partícula numa célula da terceira camada hadrônica.

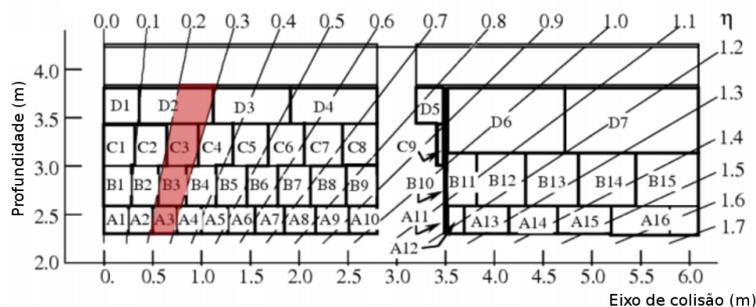


Figura 13: Esquema de uma *trigger tower* do TileCal (região sombreada).

Os sinais de saída do circuito 3-in-1, após terem sido digitalizados, são enviados, através de fibras óticas, para a eletrônica de *back-end* do TileCal. Essa eletrônica é composta, basicamente, pelas ROD (do inglês *Read-out Driver*), por sistemas de controle e comunicação entre o detector e os outros sistemas (aquisição de dados, controle do detector, etc).

### 2.4.3 O sistema de filtragem do ATLAS

Um evento gerado no ATLAS produz informações coletadas por todos os detectores, formando uma massa de dados da ordem de alguns megabytes [5]. Outro fator importante é a elevada taxa de eventos, que aumenta ainda mais a massa de dados a ser analisada. Estima-se que o ATLAS gerará uma massa de dados da ordem de  $TB/s$ , tornando o armazenamento da informação para posterior análise *offline* impraticável [5]. Assim, as informações coletadas pelos detectores serão processadas por um sistema de filtragem *online*, de forma a diminuir a alta taxa de eventos [23]. Para o ATLAS, foi concebido um sistema dividido em 3 níveis, ligados em cascata. Cada nível tem sua taxa de eventos e seu tempo de processamento. Esse sistema de filtragem *online* é conhecido como o Sistema de *Trigger* do ATLAS.

O primeiro nível de filtragem é responsável por diminuir a taxa de eventos de  $40 MHz$  para  $100 kHz$ , num tempo de processamento de  $1 \mu s$ . Para alcançar tal objetivo num tempo tão pequeno, esse nível é implementado totalmente em *hardware* de baixa programabilidade (FPGA, do inglês *Field Programmable Gate Array*) [24], fazendo uso de informação compactada do detector [23]. Por fim, o primeiro nível irá selecionar regiões

do calorímetro onde houve deposição significativa de energia, as chamadas regiões de interesse, ou RoI (do inglês *Region of Interest*). Essas regiões são os objetos vistos pelo segundo nível.

O segundo nível de trigger irá operar sobre as RoI selecionadas pelo primeiro nível. Por observar somente parte do detector e ser dotado de um tempo de processamento maior, de  $10\text{ ms}$ , este nível utilizará a resolução plena do detector. Refinando a decisão do nível anterior, o segundo nível deve diminuir a taxa de eventos de  $100\text{ kHz}$  para  $1\text{ kHz}$ , entregando a nova seleção de eventos para o nível seguinte.

O terceiro nível, então, processará os eventos selecionados pelo segundo nível, utilizando as informações provenientes de todos os subdetectores do ATLAS. Não mais que uma taxa de  $100\text{ Hz}$  deverá ser armazenada em mídia permanente. A Figura 14 ilustra as etapas do sistema de trigger.

Este projeto foi desenvolvido no segundo nível de trigger, utilizando somente as informações de calorimetria do experimento. Assim, os algoritmos de extração de características e os discriminadores devem operar com a resolução completa do detector, bem como com o tempo de processamento estipulado para este nível.

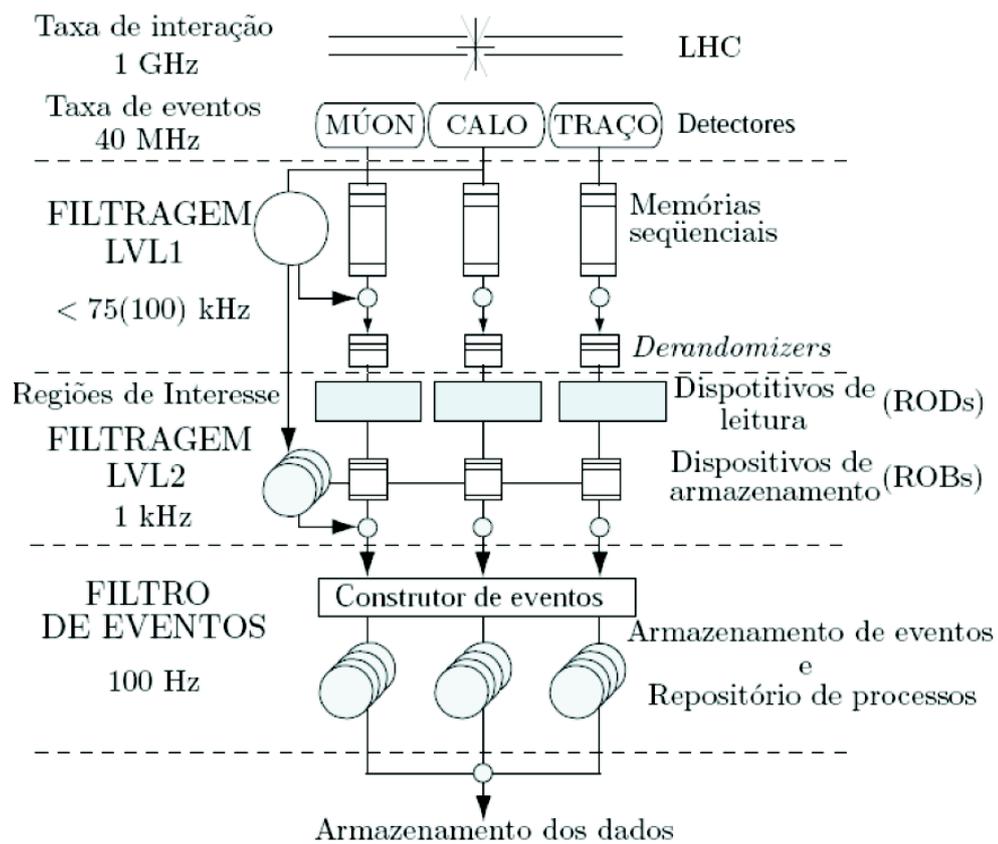


Figura 14: Diagrama do sistema de filtragem do ATLAS.

## 3 O Segundo Nível de Filtragem do ATLAS

As regiões selecionadas pelo primeiro nível de filtragem, RoI, serão analisadas com mais precisão pelo segundo nível. Através do pleno uso da fina granularidade dos detectores envolvidos, as características dos eventos podem ser melhor exploradas. O segundo nível irá, então, validar a decisão do primeiro nível. Os componentes envolvidos neste nível são: o construtor de RoI (RoIB, do inglês *RoI Builder*); o supervisor do segundo nível (L2SV, do inglês *Level 2 supervisor*); os processadores do segundo nível (L2P, do inglês *Level 2 processors*); os sub-sistemas de leitura (ROS, do inglês *Read-out sub-systems*) e os pseudo-subsistemas (pROS, do inglês *Pseudo-read-out sub-system*).

Este capítulo apresentará características específicas do segundo nível de filtragem do ATLAS, voltadas para a calorimetria do evento, seus algoritmos de extração de características e de hipótese.

### 3.1 Recursos e tempo de processamento

Espera-se que o segundo nível seja capaz de reduzir a taxa de eventos de  $100\text{ kHz}$  para não mais que  $1\text{ kHz}$ . Levando-se em conta a diversidade de operações que esse nível deve realizar, esta restrição é muito genérica. Alguns algoritmos serão mais rápidos que outros e, portanto, alguns canais (um canal representa uma assinatura do Higgs) tenderão a ser processados de forma mais rápida.

O tempo de processamento de cada canal não é somente um fator de sua taxa de entrada associada. Há também o poder de processamento associado a cada canal. Isso dependerá da carga de cada canal, ou da dificuldade de processá-lo.

Atualmente, o número de processadores é da ordem de milhares. Estima-se que, em média, 10 *ms* serão necessários para cada evento no segundo nível [5].

## 3.2 Organização lógica do segundo nível

Os RoIB's agrupam os fragmentos de informação gerados pelos diferentes sub-detectores do ATLAS e enviam os registros destes agrupamentos para um L2SV. O L2SV, por sua vez, seleciona um determinado L2P para este evento, com sua respectiva informação do primeiro nível. Após essa etapa, o L2SV fica aguardando a decisão do segundo nível ser retornada pelo L2P.

O L2P, ao receber o evento do L2SV, executa o *software* de seleção de eventos do segundo nível, requisitando os dados do evento aos ROS's. De posse dos dados, as características do evento são extraídas e o evento é classificado. Esta decisão retorna ao L2SV.

Caso o evento deva prosseguir ao terceiro nível de filtragem, o L2P envia as informações mais acuradas (no segundo nível opera-se com toda a resolução dos detectores) extraídas das RoI's aos pROS's, para que possam ser adicionadas ao evento a ser construído. Detalhes específicos sobre o ROS e aspectos do fluxo de dados podem ser vistos em [5].

Todo o segundo nível de filtragem foi implementado utilizando o máximo possível de tecnologias comerciais [25], visando fácil reposição de materiais. Todos os processadores são do tipo PC, conectados através de uma rede de *switches* Gigabit Ethernet, devido à velocidade, confiabilidade e padronização do protocolo de comunicação. Já as aplicações são todas desenvolvidas utilizando técnicas de orientação a objetos, implementadas em C++ [26].

### 3.2.1 O construtor de RoI

Após cada aceitação, o primeiro nível de filtragem envia informações como a posição da RoI e os limiares observados. O construtor de regiões de interesse (RoIB, do inglês *Region*

of *Interest Builder*) tem como objetivo combinar os fragmentos das RoI espalhados pelos *buffers* de leitura (ROB) em um único registro que é enviado ao supervisor do segundo nível (L2SV).

### 3.2.2 O supervisor do segundo nível

Os supervisores do segundo nível são um pequeno grupo de processadores, que têm a responsabilidade de supervisionar o fluxo de eventos no segundo nível, atuando como mediadores entre os sistemas do primeiro e segundo nível. Cada L2SV apresenta uma interface de comunicação com o RoIB, de forma a receber os dados dos eventos aceitos pelo primeiro nível.

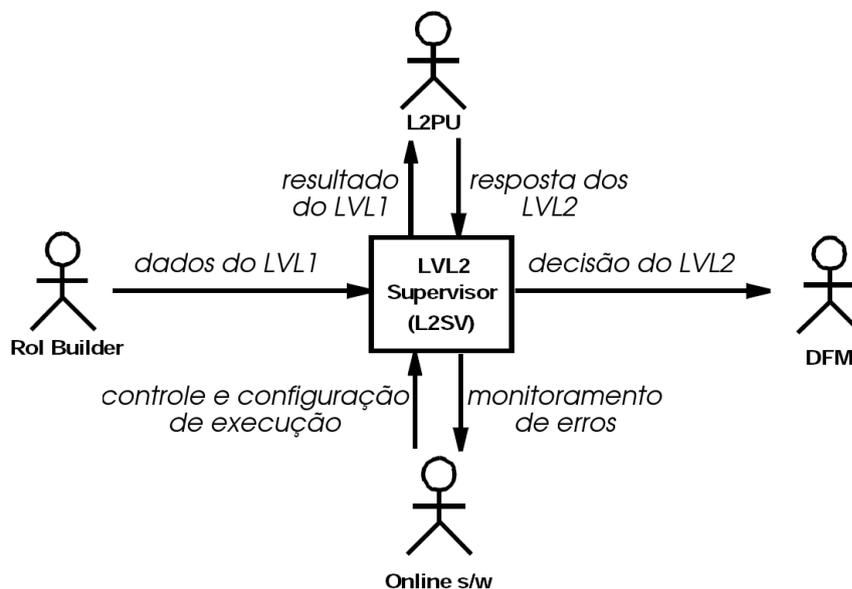


Figura 15: Contexto do supervisor do segundo nível. Extraído de <http://cds.cern.ch>.

O contexto do supervisor do segundo nível é indicado na Figura 15. O L2SV recebe as informações do primeiro nível em um único registro através do RoIB. O supervisor seleciona um L2P para o evento, que passa para uma de suas muitas unidades de processamento (L2PU, do inglês *LVL2 Processing Unit*). As L2PU's rodam algoritmos de seleção local e global, retornando uma decisão sobre o evento ao L2SV.

Se um determinado número de eventos for rejeitado, uma mensagem é enviada ao controlador de fluxo de dados (DFM, do inglês *Data Flow Manager*) do sistema, para que o mesmo coordene a remoção, dos *buffers* de leitura, dos dados correspondentes aos

eventos rejeitados.

Para a etapa de seleção de um L2P específico, o supervisor utiliza uma função de balanceamento de carga. Atualmente, o algoritmo executado para tal é o *round-robin* [27].

### 3.2.3 O decisor local

Decisores locais são executados pelos L2P's e são responsáveis por confirmar os objetos selecionados pelo primeiro nível. As principais tarefas de um decisor local são:

- Unidade de pré-processamento: extrai os dados das ROB, sendo estes dados organizados e pré-analisados.
- Extrator de característica: é onde ocorre a extração das características mais importantes dos objetos.
- Discriminador local: é onde ocorre a confirmação do objeto selecionado pelo primeiro nível. Qualquer método de classificação pode ser utilizado pelo discriminador local.

Existem decisores locais para os diversos tipos de objetos indicados pelo primeiro nível.

### 3.2.4 Decisor global

Os decisores globais reúnem os dados provenientes de todos os demais decisores locais, para um mesmo evento (colisão), comparando seus objetos com uma tabela de assinaturas que representem a física de interesse.

## 3.3 O ambiente ATHENA

O ambiente de desenvolvimento ATHENA [28] é o *framework* que permite desenvolver assinaturas, algoritmos e todo o funcionamento real do sistema de filtragem, permitindo eficiente elaboração de algoritmos de filtragem.

O objetivo do ATHENA é disponibilizar ferramentas que facilitem a escrita e análise de programas para os sistemas *online* e *offline* do ATLAS. Este ambiente possui ferramentas para realizar simulações de uma colisão próton-próton e simulação da resposta dos detectores, a partir de modelos precisos e em contínuo desenvolvimento dos detectores. Dessa forma, é possível realizar o desenvolvimento de todo o sistema de filtragem do segundo e terceiro níveis.

O desenvolvimento do ATHENA está sendo feito em C++. Todo o ambiente de desenvolvimento do sistema de filtragem do ATLAS está sendo desenvolvido utilizando esta linguagem.

Os dados utilizados neste trabalho são simulações realizadas no ambiente ATHENA, representando uma realidade muito próxima à esperada pelo sistema de filtragem. Como os discriminadores propostos neste trabalho operarão no segundo nível de filtragem, foram selecionados somente eventos aceitos pelo primeiro nível de filtragem. No total, aproximadamente 2400 RoI de elétrons e 800 de jatos são utilizadas.

## 3.4 O processo de seleção de eventos

Os algoritmos de filtragem do segundo nível são otimizados para rápida execução usando a granularidade e precisão completa das regiões de interesse. A redução na taxa de eventos é alcançada a partir de medidas mais precisas que as utilizadas pelo primeiro nível e pelo uso de algoritmos mais complexos. Os métodos de extração das características (FEX, do inglês *Feature Extraction*), executados pelos decisores locais, são desenvolvidos para diferentes tipos de objetos. Os métodos apresentados neste trabalho, e outros, podem ser encontrados em [29].

### 3.4.1 O extração de características para os calorímetros

Os algoritmos de extração de características para os dados de calorimetria buscam, principalmente, a confirmação da produção de cascatas eletromagnética e hadrônicas. O centro de impacto em  $\eta$  e  $\phi$ , a energia e propriedades transversal e longitudinal da cascata

são calculados, incluindo informações de isolamento. O formato da cascata no *presampler* também é obtido. Análises posteriores são realizadas em uma janela centralizada no ponto de impacto da cascata.

Por ser um canal de interesse muito freqüente, a busca por elétrons e fótons é bastante valorizada no segundo nível. Porém, devido a similar deposição de energia, jatos de partículas conseguem falsear elétrons. A Figura 16 mostra uma RoI típica de um elétron interagindo com as diversas camadas dos calorímetros. Cada bloco representa a energia depositada numa célula. Pode-se ver que elétrons formam uma cascata colimada, depositando sua energia quase que totalmente na *presampler* e nas 2 camadas do calorímetro eletromagnético. Nas outras camadas encontra-se, praticamente, ruído.

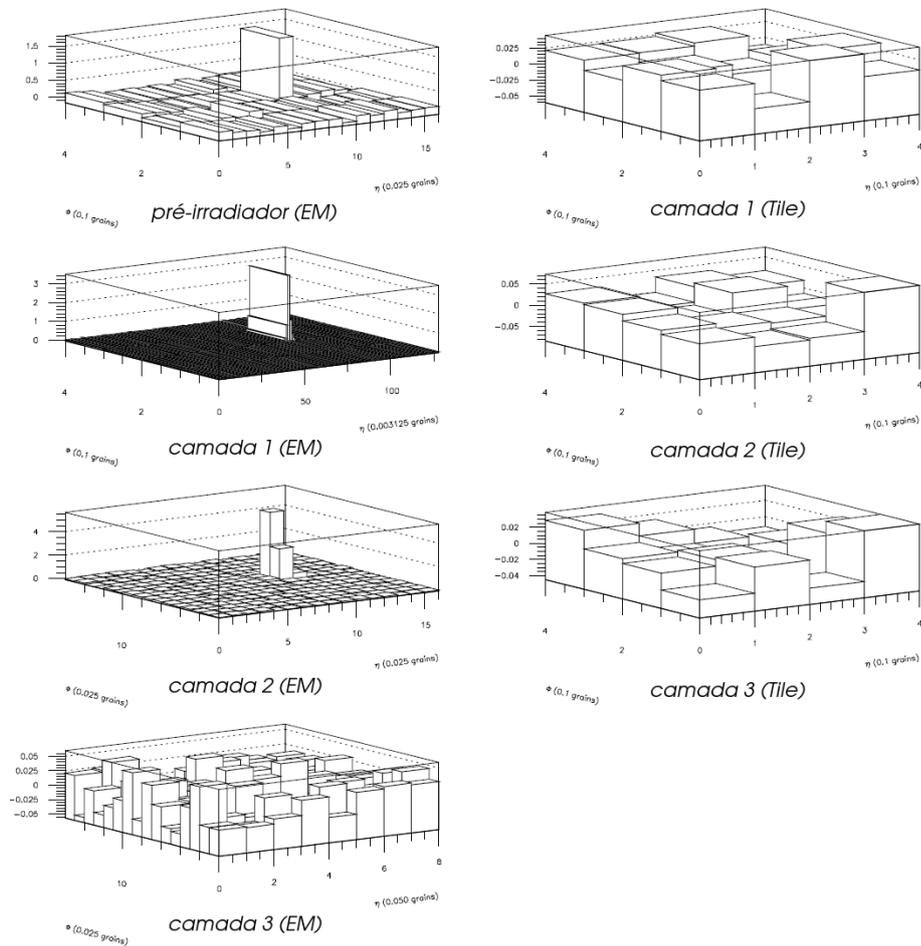


Figura 16: RoI de um elétron típico interagindo com as camadas dos calorímetros do ATLAS.

A Figura 17 mostra um jato com uma cascata de deposição de energia com padrão parecido com o de um elétron. A precisão do primeiro nível impossibilita a filtragem desse

tipo de jato. Como o segundo nível opera sobre a granularidade completa do detector, as sutis diferenças entre elétrons e jatos aceitos pelo primeiro nível podem ser observadas.

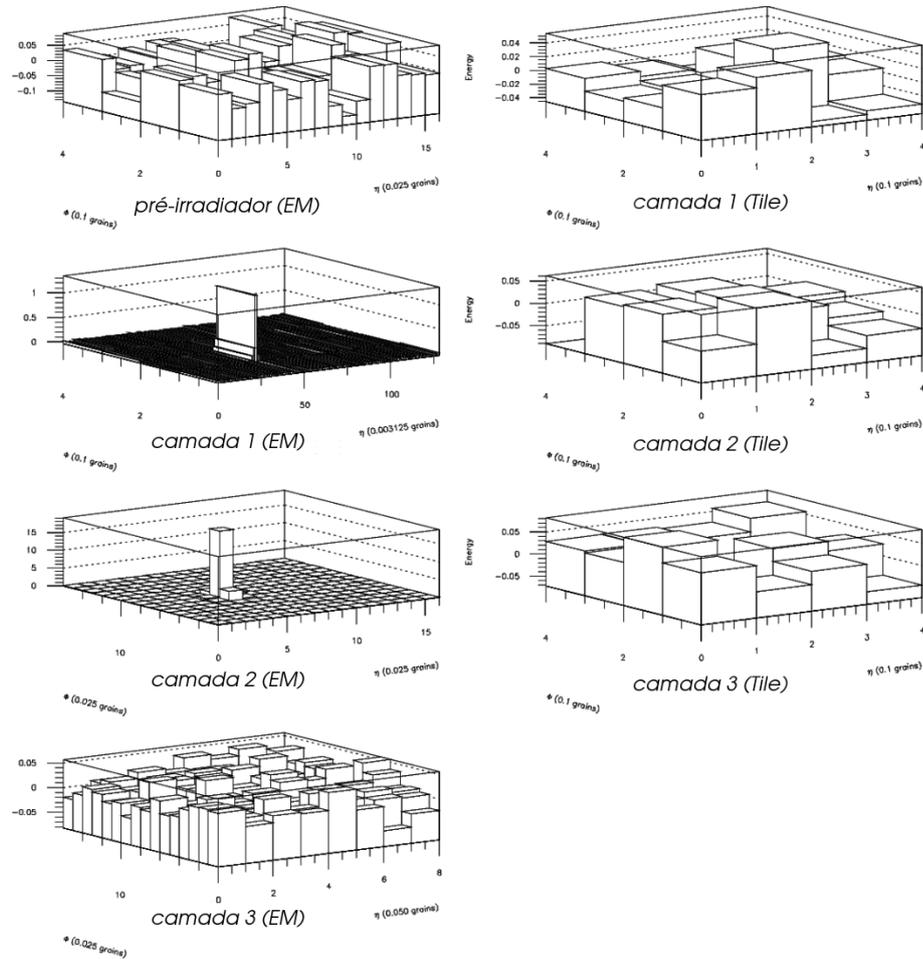


Figura 17: RoI de um jato aceito pelo primeiro nível interagindo com as camadas dos calorímetros do ATLAS. Esse tipo de jato tem uma padrão de deposição de energia muito parecido com o do elétron.

Para desempenhar essa função de identificação de chuveiros eletromagnéticos contaminados por jatos, a colaboração desenvolveu um algoritmo, chamado T2Calo, que será detalhado a seguir.

### 3.4.2 Algoritmo do T2Calo

O T2Calo é um algoritmo de classificação e extração de características do segundo nível, que utiliza somente as informações de calorimetria do evento [30]. Após sua execução, o algoritmo extrai um conjunto de variáveis altamente discriminantes.

O algoritmo do T2Calo se baseia no agrupamento de células do calorímetro para ex-

trair as informações relevantes. A primeira etapa é encontrar a célula com maior deposição de energia, na segunda camada e.m.. Depois, é construída uma janela em  $\eta$  e  $\phi$  ao redor da posição desta célula de maior energia, onde cada variável é calculada separadamente. Esse conjunto de variáveis é descrito a seguir:

Tabela 4: Descrição das variáveis do T2Calo.  $E_{i \times j}$  representa a energia numa janela de  $i$  células em  $\eta$  e  $j$  células em  $\phi$ .  $E_i$  representa a  $i$ -ésima célula energética.

Variável	Descrição	Variável	Descrição
T2Caeme	$E_{7 \times 7}$ e.m. total	T2Caemes0	$E_{7 \times 7}$ e.m. total na <i>presampler</i>
T2Caemes1	$E_{7 \times 7}$ e.m. total na camada 1	T2Caemes2	$E_{7 \times 7}$ e.m. total na camada 2
T2Caemes3	$E_{7 \times 7}$ e.m. total na camada 3	T2Cahade	$E_{7 \times 7}$ hadrônica total + Tampas
T2Cahades0	$E_{7 \times 7}$ hadrônica total na camada 1	T2Cahades1	$E_{7 \times 7}$ hadrônica total na camada 2
T2Cahades2	$E_{7 \times 7}$ hadrônica total na camada 3	T2Cahades3	Energia total no <i>scintillator gap</i>
T2Carcore	$E_{3 \times 7}/E_{7 \times 7}$ na camada 2 e.m.	T2Caeratio	$(E_1 - E_2)/(E_1 + E_2)$ na camada 1 e.m.
T2Cawidth	Largura ponderada em 3x5 na camada 2 e.m.	T2Caf73	$(E_{7 \times 7} - E_{3 \times 3})/E_{3 \times 3}$ na camada 1 e.m.

### 3.4.3 O anelador

Em média, cada RoI consiste em um conjunto de cerca de 1300 células espalhadas em diferentes camadas do calorímetro. Esta elevada dimensão do espaço de entrada requer a utilização de uma compactação eficiente. Propõe-se então, uma compactação em anéis concêntricos de deposição de energia [7].

Como cada camada do calorímetro tem granularidade diferente, é de se esperar que cada camada tenha o seu conjunto de anéis. Dessa forma, as características do evento têm o máximo de aproveitamento. A Tabela 5 mostra as camadas do detector e o número de anéis extraído.

Tabela 5: Número de anéis para cada camada dos calorímetros e.m. e hadrônico

	Camadas eletromagnéticas			Camadas hadrônicas			
	<i>presampler</i>	1 <sup>a</sup>	2 <sup>a</sup>	3 <sup>a</sup>	1 <sup>a</sup>	2 <sup>a</sup>	3 <sup>a</sup>
Número de anéis	8	64	8	8	4	4	4

A extração dos anéis ocorre da seguinte forma: a posição em  $\eta$  e  $\phi$  da célula de maior energia da segunda camada do calorímetro eletromagnético é considerada o centro de todos os outros anéis, de todas as outras camadas. A partir daí, soma-se o valor das

células concêntricas a este ponto. Cada conjunto de células adjacentes à parte externa de um anel é considerado um novo anel. Prossegue-se com esse algoritmo até o número de anéis da camada ser atingido.

Em alguns casos, como na camada *presampler* e na terceira camada do calorímetro e.m., os anéis têm uma forma alongada, devido à diferença no tamanho das células em  $\eta$  e  $\phi$ . Outros casos, quando a célula de maior energia se localiza próxima à borda do detector, podem ser criados anéis incompletos, mas a energia das células é somada para gerar uma única quantidade. A Figura 18 mostra uma configuração genérica dos anéis para algumas camadas dos calorímetros.

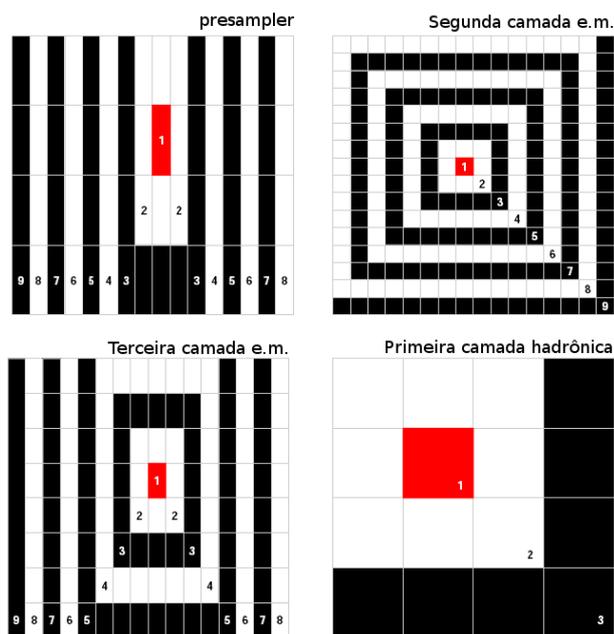


Figura 18: Configuração genérica dos anéis para algumas camadas dos calorímetros do ATLAS.

### 3.4.3.1 Normalização

Como a energia dos objetos observados tem uma variação elevada [5], uma normalização nos valores dos anéis deve ser utilizada. A normalização proporcionará uma variação mais suave no valor dos anéis. Normaliza-se, então, os anéis pela energia, permitindo que o discriminador se concentre nas diferenças nos padrões de deposição energética. Dentre vários tipos de normalização sugeridas em [7], a escolhida foi a do tipo seqüencial. Esta normalização consiste em (para cada camada) calcular a soma das energias de cada célula

e dividir o valor da energia do primeiro anel (antes da normalização) por esta soma. A energia do segundo anel é dividida pelo valor da soma menos a energia do primeiro anel. E assim procede-se até o término dos anéis. A Tabela 6 mostra esse algoritmo.

Tabela 6: Esquema da normalização do tipo seqüencial.

Anel	Normalização	Soma em Anel
1	$E(\text{camada})$	$E_1$
2	$E - E_1$	$E_2$
...	...	...
N-1	$E - \dots - E_{N-2}$	$E_{N-1}$
N	$E - \dots - E_{N-1}$	$E_N$

A normalização seqüencial propõe uma amplificação dos anéis periféricos, realçando a informação desses e, possivelmente, melhorando a eficiência do discriminador posterior. Porém, ocorre um problema: anéis com a menor relação sinal/ruído serão amplificados. Para contornar esse problema, quando a energia de um anel equipara-se a um patamar próximo ao nível de ruído, a energia passa a ser dividida pela energia total da camada.

### 3.4.4 Algoritmo de hipótese

O algoritmo de hipótese atual do segundo nível de *trigger* para calorimetria consiste em cortes lineares de energia. As distribuições energéticas observadas de 4 variáveis calculadas pelo T2Calo são vistas a partir de mapas bi-dimensionais e, incorporando-se o conhecimento especialista em calorimetria, etiquetará a partícula. Sendo o algoritmo em cascata, qualquer condição não aceita já classifica a partícula como rejeitada.

As variáveis utilizadas no algoritmo de hipótese são:

- $E_{et}$ : energia transversa total do calorímetro e.m., calculada em uma janela de 3 células em  $\eta$  e 7 em  $\phi$ .
- $E_{et_{had}}$ : energia transversa que vaza para o calorímetro hadrônico, calculada em uma janela de  $\Delta\eta \times \Delta\phi = 0,2 \times 0,2$ .
- $R_{shape}$ : é uma razão que caracteriza a deposição energética da cascata eletromagnética. Representa a variável T2Carcore do T2Calo.

- $R_{strip}$ : é uma razão que caracteriza colimação do feixe de partículas. Representa a variável T2Caeratio do T2Calo.

Cada algoritmo de extração de características do sistema de *trigger* tem o seu algoritmo de hipótese associado. Nada impede que a metodologia adotada para um algoritmo de hipótese não possa ser utilizada em outro. Essa independência garante uma maior flexibilidade na elaboração de algoritmos e discriminadores para o segundo nível de filtragem.

## 4 Decisores Neurais

Nos capítulos anteriores foram apresentados os algoritmos de extração de características (T2Calo e o anelador) e as principais características do sistema de filtragem do ATLAS, em especial o segundo nível de filtragem.

O anelador proposto extrai 100 anéis de deposição de energia das diversas camadas dos calorímetros e.m. e hadrônico. De forma a diminuir o número de anéis utilizados na classificação das partículas, pré-processamentos estatísticos são utilizados.

Uma das técnicas utilizadas é o estudo de relevância. Esta técnica consiste em selecionar somente as variáveis de entrada do discriminador que são importantes para a discriminação. Variáveis que não são discriminantes podem ser descartadas sem alterar significativamente o desempenho do discriminador. A análise de relevância pode, também, mostrar variáveis confundidoras. A presença dessas variáveis diminui o desempenho do discriminador.

Outra técnica utilizada é a análise de componentes principais (PCA, do inglês *Principal Component Analysis*), uma técnica estatística que descorrelaciona a informação contida numa base de dados. Os componentes principais extraídos podem ser escolhidos de duas maneiras: pela quantidade de energia (variância) ou pela relevância desses componentes ao discriminador. Os demais componentes não utilizados podem ser descartados, acarretando uma perda controlada de informação.

No contexto do segundo nível de filtragem do ATLAS, são necessários discriminadores que combinem velocidade e alta eficiência de detecção. Como a taxa de eventos é elevada e os objetos a serem discriminados são bastante complexos e apresentam sutis diferenças, o projeto de um discriminador para este nível é uma tarefa de dificuldade elevada, exigindo

a colaboração de especialistas de diversas áreas para alcançar a performance requerida. Em termos de eficiência do discriminador, a relação entre a eficiência de detecção de elétrons e a taxa de falso alarme (jatos classificados erroneamente como elétrons) deve ser a melhor possível. Já com respeito à velocidade do discriminador, deve-se lembrar que o segundo nível de filtragem tem um tempo pré-determinado de latência máxima, que deve ser respeitado uma vez que o discriminador proposto irá operar neste nível. Em [31], foi mostrado que o tempo de processamento de um discriminador neural, utilizando os anéis de deposição de energia, é compatível com as especificações do segundo nível. Qualquer compactação extra só diminuirá esse tempo de processamento.

Neste ambiente, estudos mostraram que decisores neurais podem apresentar grandes vantagens em ambientes de física de altas energias [32]. Decisores neurais foram testados no ambiente do segundo nível de filtragem do ATLAS [33, 31, 34], com dados simulados dentro do ambiente ATHENA, apresentando excelentes resultados.

## 4.1 Redes neurais artificiais

Decisores neurais são baseados em redes neurais artificiais (RNA). Esses discriminadores exploram as características dos eventos apresentados no seu treinamento e as generalizam para outros eventos nunca vistos. RNA podem trabalhar em espaços de elevada dimensão e aplicar cortes não lineares, alcançando bons resultados em aplicações de reconhecimento de padrões [35]. Já o atual algoritmo de hipótese se baseia em cortes lineares de informação especialista para classificar a partícula incidente.

Discriminadores neurais apresentam as seguintes vantagens:

**Robustez:** Em ambientes sujeitos a falhas, radiação e perda de alguma parte da cadeia de leitura, como é o caso do segundo nível de filtragem, redes neurais artificiais conseguem manter um elevado desempenho, mesmo quando parte dos dados de entrada é perdida ou corrompida.

**Generalização:** Mesmo com uma forte presença de ruído, redes neurais artificiais conseguem manter um desempenho elevado. Os objetos a serem discriminados possuem,

invariavelmente, forte ruído.

**Detecção de novos fenômenos:** Um dos objetivos do ATLAS é a descoberta de nova física. Redes neurais artificiais têm a habilidade de detectar padrões nunca vistos, de forma bastante eficaz.

**Implementação simples:** Embora redes neurais artificiais apresentem uma melhor performance se implementadas em *hardware* dedicado, sua implementação em processadores genéricos não é determinante para a performance total do discriminador. Além do mais, essa opção é mais barata, flexível e compatível com os sistemas utilizados no segundo nível de filtragem do ATLAS.

A Figura 19 mostra uma rede neural artificial sem realimentação, com 3 camadas e totalmente conectada.

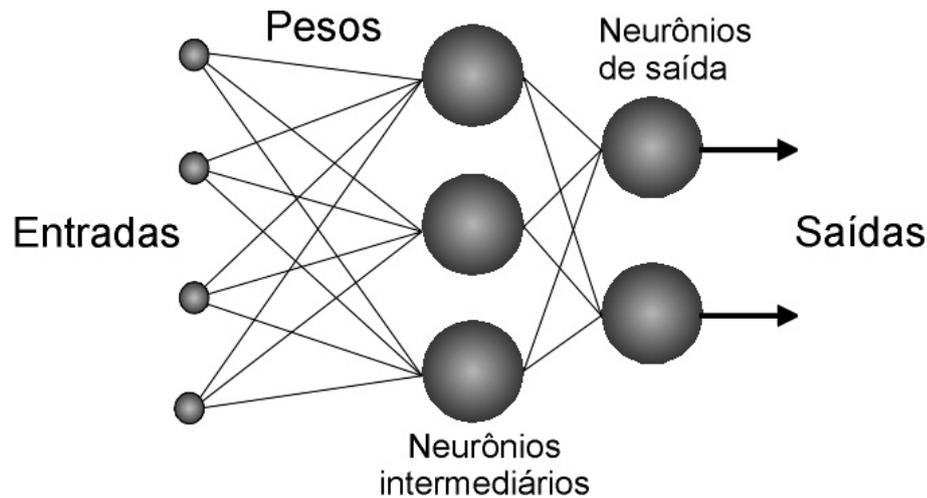


Figura 19: Exemplo de uma rede neural artificial.

## 4.2 Anelador neural

Uma alternativa para o segundo nível de filtragem do ATLAS é utilizar o anelador como o algoritmo de extração de características. Os anéis de deposição de energia extraídos das RoI são altamente discriminantes em relação ao T2Calo [31]. Como decisor, uma rede neural artificial é treinada e alimentada com os anéis. Este trabalho se concentra no canal elétron/jato, por ser a assinatura mais freqüente do bóson de Higgs. As redes neurais são treinadas para discriminar elétrons de jatos de outras partículas.

## 4.2.1 Metodologia do treinamento neural

O treinamento da rede neural artificial é uma parte essencial no projeto do discriminador. Idealmente, para o treinamento da rede neural, deve-se ter 3 grupos de dados: treino, teste e validação. O grupo de treino é utilizado, como o nome diz, no treinamento da rede, no ajuste dos pesos sinápticos. O grupo de validação é utilizado para validar o treinamento, impedindo que a rede neural perca sua capacidade de generalizar a resposta. Por fim, o grupo de teste serve para testar o desempenho do discriminador, apresentando na sua entrada eventos nunca antes vistos.

O conjunto de dados utilizados neste trabalho, com 2400 RoI de elétrons e 800 RoI de jatos, foi dividido em dois grupos: treino e teste. Devido à pequena estatística, o grupo de teste foi utilizado, também, como grupo de validação do treinamento. Para o grupo de treino, o conjunto de jatos foi replicado, diminuindo a diferença significativa de eventos entre jatos e elétrons. Dessa forma, consegue-se evitar uma possível tendência do treinamento em relação a elétrons.

Como se sabe a natureza dos objetos a serem discriminados, foi escolhido um algoritmo de treinamento supervisionado, chamado de *resilient backpropagation* [36]. Esse algoritmo utiliza somente a informação da direção do gradiente da curva de erro para o ajuste dos pesos, por bateladas.

A figura de mérito escolhida para a discriminação elétron/jato foi o produto SP [31], que relaciona as probabilidades de detecção de elétron e jato de acordo com a seguinte fórmula:

$$SP = (p_e p_j)(p_e + p_j)$$

Onde  $p_e$  indica a probabilidade de detecção de elétrons e  $p_j$  a probabilidade de detecção de jatos.

O produto SP foi escolhido pois, em aplicações como discriminação no segundo nível de filtragem, deseja-se que a eficiência de detecção seja elevada, mas que a taxa de falso alarme seja mantida a menor possível. O produto SP torna essa relação a mais homogênea possível. Inspeccionando a fórmula do produto SP, pode-se ver que ele assume valores entre

0 e 2.

Quanto à topologia do discriminador neural, foram treinadas diversas redes, variando-se o número de neurônios da camada escondida de 2 a 15. Para a camada de entrada, o número de nós acompanha o número de variáveis de entrada e, para a camada de saída, tem-se somente 1 neurônio. A resposta da rede neural será mapeada em 1 (elétrons) ou -1 (jatos). A função tangente hiperbólica foi utilizada como função de ativação dos neurônios, seja na camada escondida, seja na camada de saída. As variáveis de entrada são esferizadas (média nula e variância unitária) de acordo com o conjunto de treinamento. Assim, os valores de entrada são mantidos dentro da faixa dinâmica da rede neural.

### 4.3 Compactação da informação

Diminuir o espaço de entrada de um discriminador é um comprometimento entre eficiência de detecção e velocidade. Fica a critério do projetista encontrar um ponto de operação, o equilíbrio desejado entre compactação e performance. Pode-se, então, ajustar a relação entre compactação e performance, possibilitando um controle a mais sobre o discriminador.

As técnicas de compactação utilizadas neste trabalho são baseadas em informação estatística e serão apresentadas a seguir.

#### 4.3.1 Relevância em discriminadores

O estudo da relevância da entrada mostra quais são as variáveis mais importantes à discriminação [7]. Assim, as demais variáveis podem, a critério do projetista, ser descartadas do processamento. O mapeamento de relevância aqui utilizado se baseia no quanto o classificador desenvolvido altera seu desempenho quando se modificam estatisticamente os valores de uma dada variável explicativa. Dois mapeamentos foram utilizados para avaliar a relevância de cada componente: um baseado na variação média quadrática da saída da rede neural, e outro baseado na variação produzida no produto SP. Assim, temos:

$$R_{MSE_{x_i}} = \frac{1}{N} \sum_{j=1}^N \{O(\vec{x}_j) - O(\vec{x}_j|_{x_i=\bar{x}_i})\}^2$$

$$R_{SP_{x_i}} = SP(X) - SP(X|_{x_i=\bar{x}_i})$$

Onde  $R_{x_i}$  é a relevância da  $i$  –ésima variável de entrada,  $O(\vec{x})$  é a resposta neural ao estímulo  $\vec{x}$ ,  $X$  um conjunto de estímulos  $\vec{x}$  e  $\bar{x}_i$  representa o valor médio da variável  $x_i$ . As variáveis que produzem a maior variação na saída do discriminador serão ditas relevantes.

A Figura 20 mostra a relevância dos 100 anéis calculada tanto por MSE como por SP, utilizando tanto o grupo de treino quanto o grupo de teste. A linha vertical destacada em cada figura é o patamar de relevância, onde anéis com relevância maior que este valor (uma ordem de grandeza abaixo da variável de maior relevância) são considerados relevantes. Já as linhas horizontais destacadas marcam a passagem de uma camada do calorímetro para a outra, de forma crescente de dentro para fora do detector (*presampler*, 1ª camada e.m. e assim até a 3ª camada hadrônica). Assim, o número de anéis selecionados seria: MSE treino, 15 anéis; MSE teste, 19 anéis; SP treino, 23 anéis; SP teste, 35 anéis.

Mais que isso, pode-se notar que a maioria dos anéis ditos relevantes por um método, tanto para treino quanto para teste, são também ditos relevantes pelo outro. Isso reforça ainda mais a utilização desses anéis pelo discriminador. Outra característica interessante apontada pelo estudo da relevância é a importância que têm os anéis mais centrais da 1ª e da 2ª camada e.m. para a discriminação. A 1ª camada hadrônica também é considerada relevante, com quase todos os seus 4 anéis acima do patamar de relevância, tanto para os conjuntos de treino e teste, quanto para os métodos por MSE e por SP. Essa camada é importante na discriminação de jatos, uma vez que elétrons tendem a depositar toda a sua energia no calorímetro e.m..

Uma das vantagens do método de relevância por SP é a identificação de variáveis confundidoras. A presença dessas variáveis confundem o discriminador, ao invés de aumentar a eficiência de detecção. O gráfico de relevância por método SP da Figura 20 mostra que alguns anéis obtiveram valores negativos de relevância: o produto SP aumenta quando eles são substituídos pela sua média.

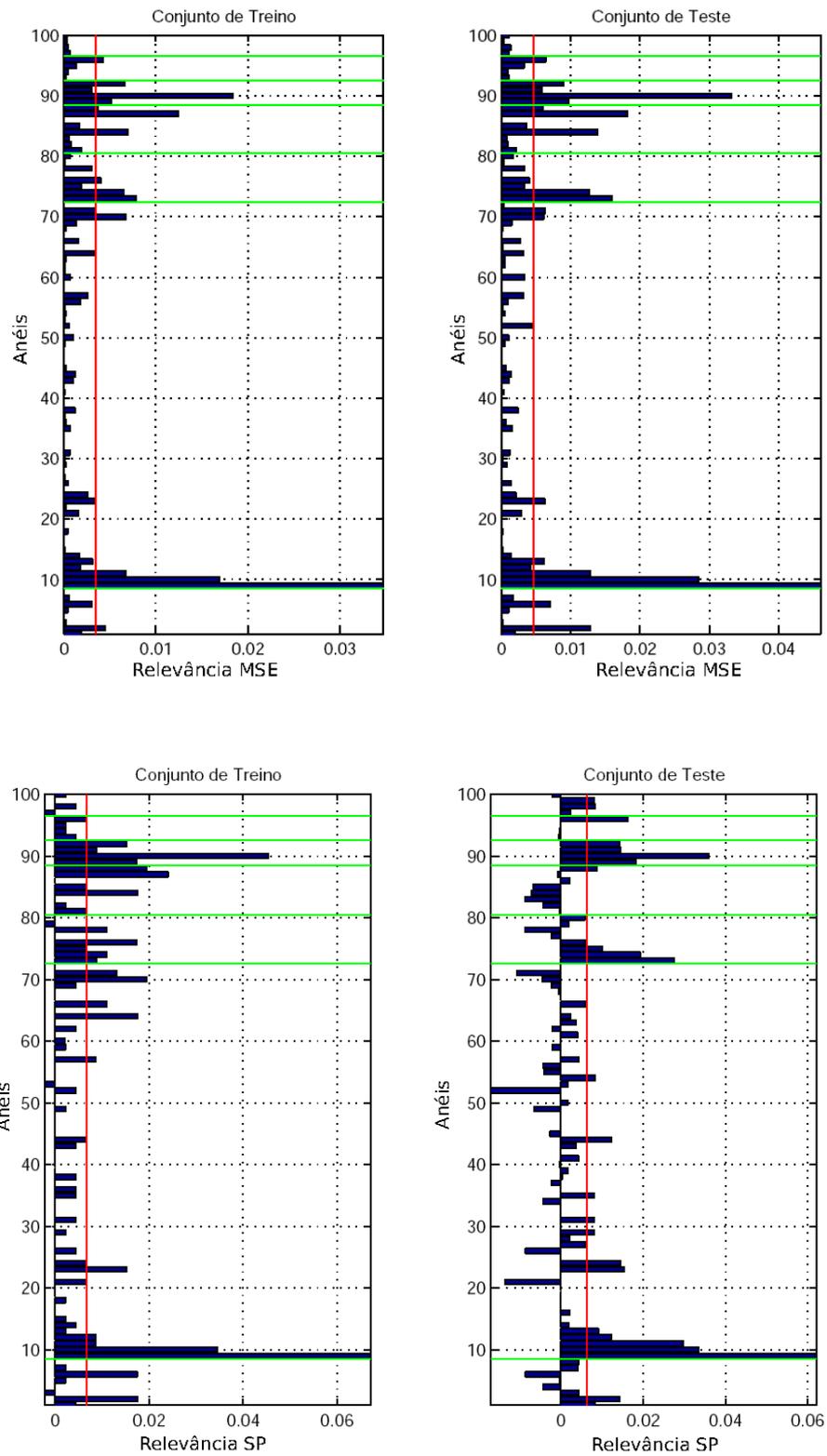


Figura 20: Relevância dos anéis pelo método MSE e SP.

### 4.3.2 Análise de componentes principais

Uma importante técnica estatística de compactação da informação é a análise de componentes principais. Esta técnica descorrelaciona a informação presente numa base de dados através de uma mudança de base [37]. Os componentes que formam esta nova base são chamados de componentes principais e podem ser ordenados de acordo com sua energia (variância). Dependendo do valor de sua energia, podemos simplesmente descartar componentes minoritários, acarretando uma perda controlada de informação.

A Figura 21 mostra um exemplo da extração de componentes principais de uma massa de dados. Pode-se ver que os componentes principais, representados pelos vetores  $\vec{v}_1$  e  $\vec{v}_2$ , são ortogonais, e a direção do primeiro componente principal ( $\vec{v}_1$ ) acompanha a direção de maior variância do conjunto de dados.

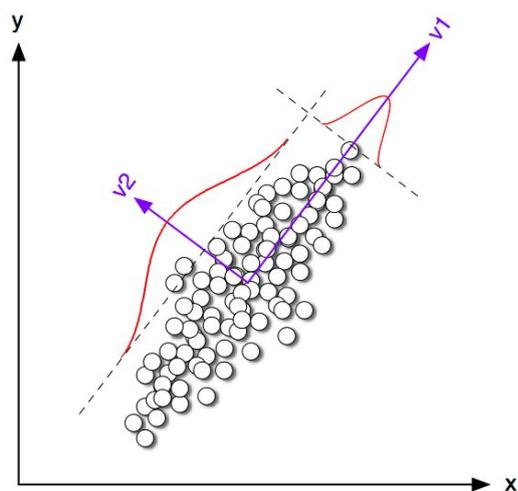


Figura 21: Exemplo de componentes principais extraídos de uma base de dados genérica.

Nesta aplicação, na qual o detector é segmentado, pode-se utilizar uma extração de componentes principais de forma segmentada [38]. Como o calorímetro é dividido em camadas de deposição de energia, os componentes principais podem, então, ser extraídos por camadas.

A escolha dos componentes principais utilizados no discriminador pode ser feita de duas formas: através da energia acumulada conforme a energia de cada componente é somada, ou pela relevância dos componentes principais para o discriminador.

A Figura 22 mostra a curva de carga para os componentes extraídos de forma não segmentada (os componentes principais são extraídos sem fazer discriminação por camadas). O eixo vertical mostra o percentual de energia acumulado conforme a energia de cada componente é somada. Pode-se ver que com 40 componentes temos 95% de energia acumulada.

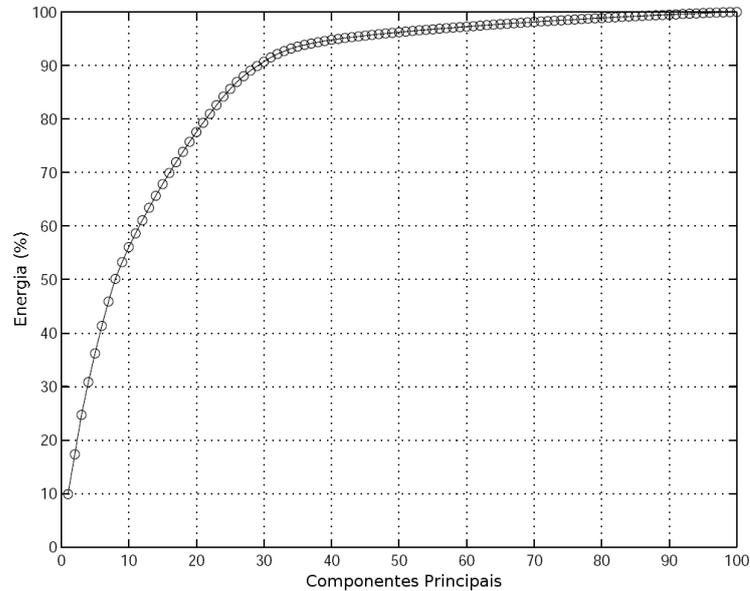


Figura 22: Curva de carga dos componente principais extraídos de forma não segmentada.

Já a Figura 23 mostra a relevância por SP dos componentes principais (a relevância por MSE mostrou resultado semelhante). Nota-se que somente um componente é considerado relevante.

A Figura 24 mostra o histograma dos valores desse componente principal mais relevante, para elétrons e jatos do conjunto de treino (o conjunto de teste mostrou resultado semelhante). Mesmo se tratando de somente um componente, pode-se ver que ele tem uma boa seletividade. A Figura 24 mostra, também, o componente principal. Pode-se interpretar seus valores como a contribuição que cada anel fornece ao componente principal. As linhas verticais mostram a passagem de uma camada para a outra do calorímetro, de forma crescente em relação ao eixo de colisão (*presampler*, 1ª camada e.m., ..., 3ª camada hadrônica). Ao cruzarmos as informações de relevância dos anéis com a contribuição que cada anel tem ao componente principal, pode-se ver que anéis considerados relevantes têm uma contribuição maior nesse componente principal.

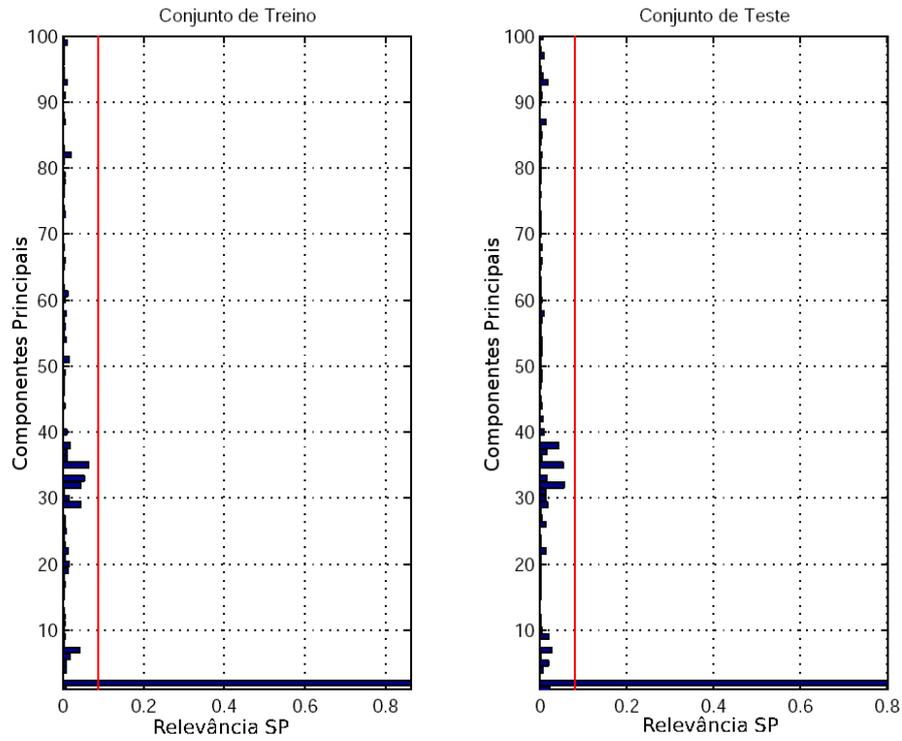


Figura 23: Relevância por SP dos componentes principais extraídos de forma não segmentada.

Analisando os resultados obtidos com a extração segmentada dos componentes principais, a Figura 25 mostra as curvas de carga de algumas das camadas do calorímetro e.m. e hadrônico. Selecionando os componentes principais através da energia acumulada, para cada camada, mostrou-se que 71 componentes mantêm entre 90% e 95% da energia.

Entretanto, utilizando a seleção de componentes principais por relevância, entre 11 a 15 componentes principais são selecionados. A Figura 26 mostra tanto a relevância pelo método MSE quanto pelo método SP.

Os métodos de compactação utilizados conseguiram diminuir o espaço de entrada do discriminador consideravelmente. A Tabela 7 sumariza a relação entre número de variáveis selecionadas e o tipo de compactação empregado.

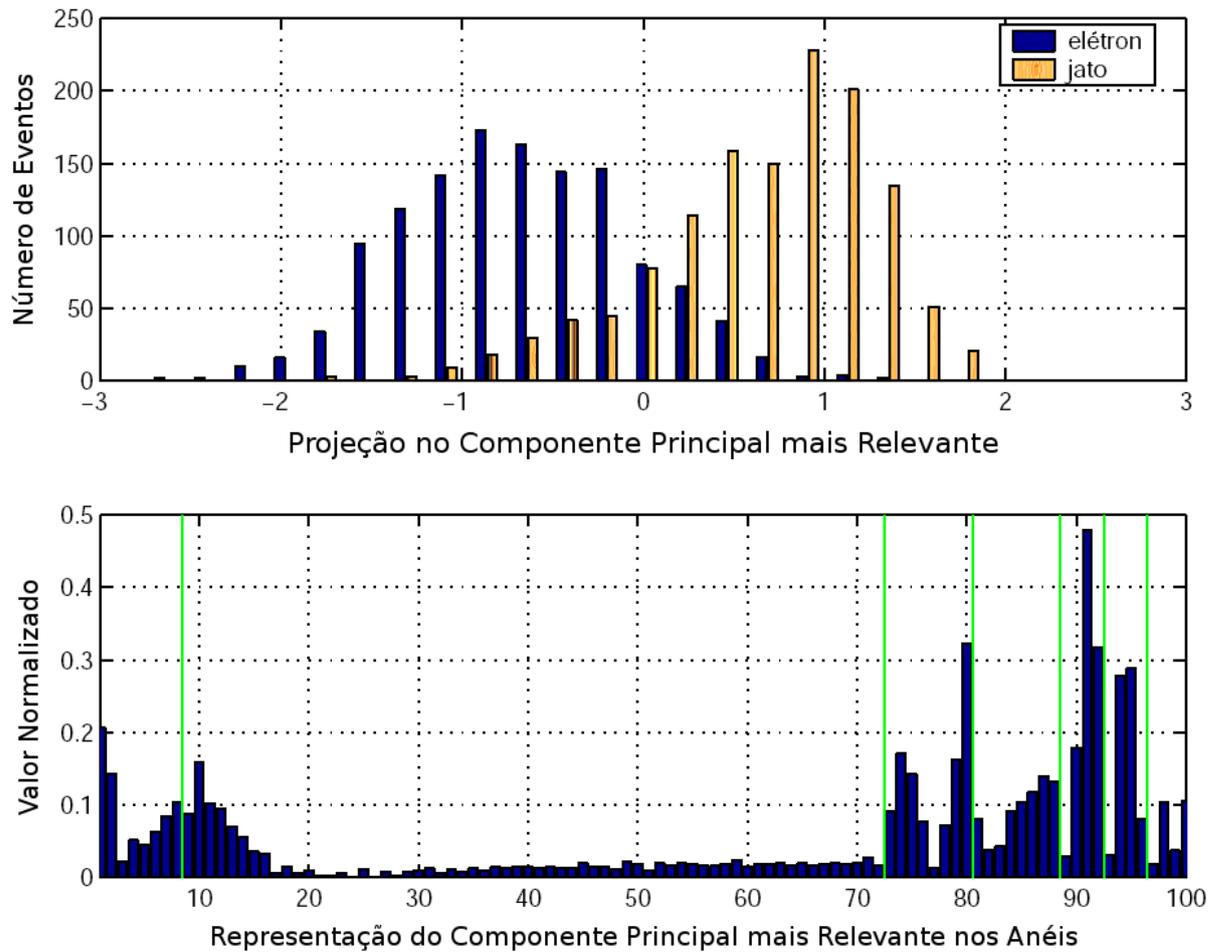
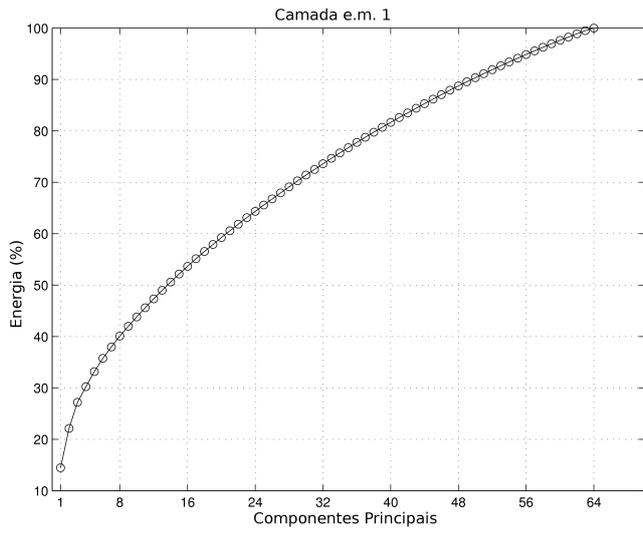


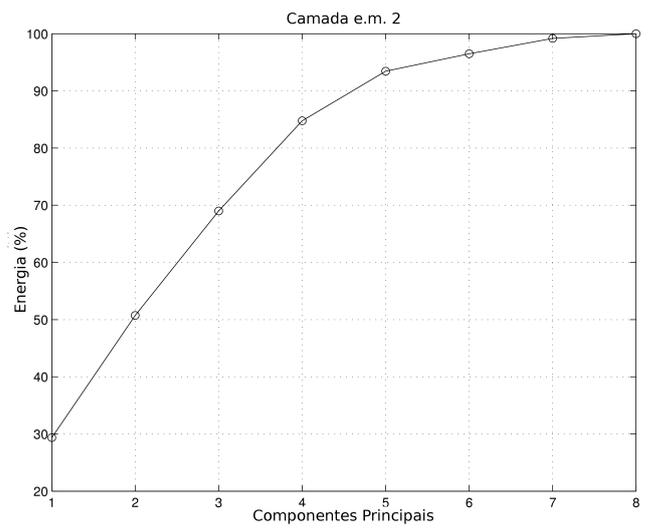
Figura 24: Histograma e representação no espaço de anéis do componente principal mais relevante.

Tabela 7: Tipos de compactação e seus respectivos números de variáveis selecionadas.

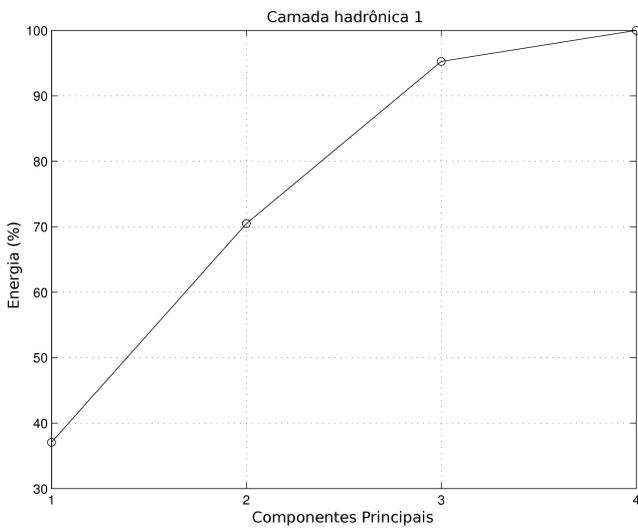
Tipo de compactação	Número de variáveis selecionadas
Relevância MSE dos anéis (treino)	15
Relevância MSE dos anéis (teste)	19
Relevância SP dos anéis (treino)	23
Relevância SP dos anéis (teste)	25
PCA não segmentada por energia	40
PCA não segmentada por relevância MSE e SP	1
PCA segmentada por energia	71
PCA segmentada por relevância MSE	11
PCA segmentada por relevância SP (treino)	14
PCA segmentada por relevância SP (teste)	15



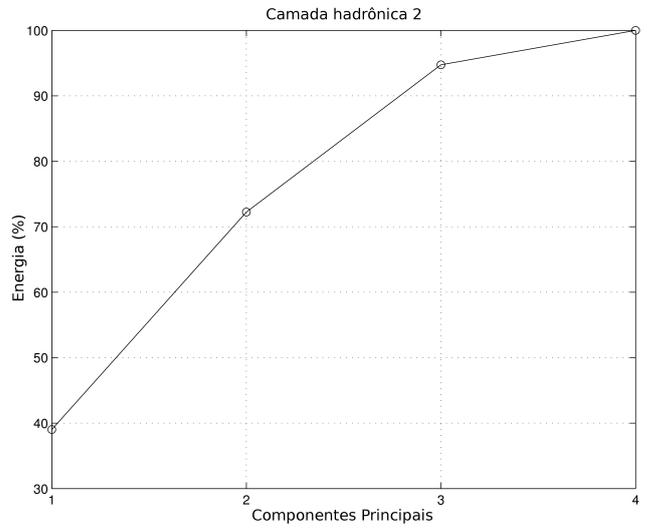
(a)



(b)



(c)



(d)

Figura 25: Curva de carga de algumas das camadas dos calorímetros e.m. e hadrônico.

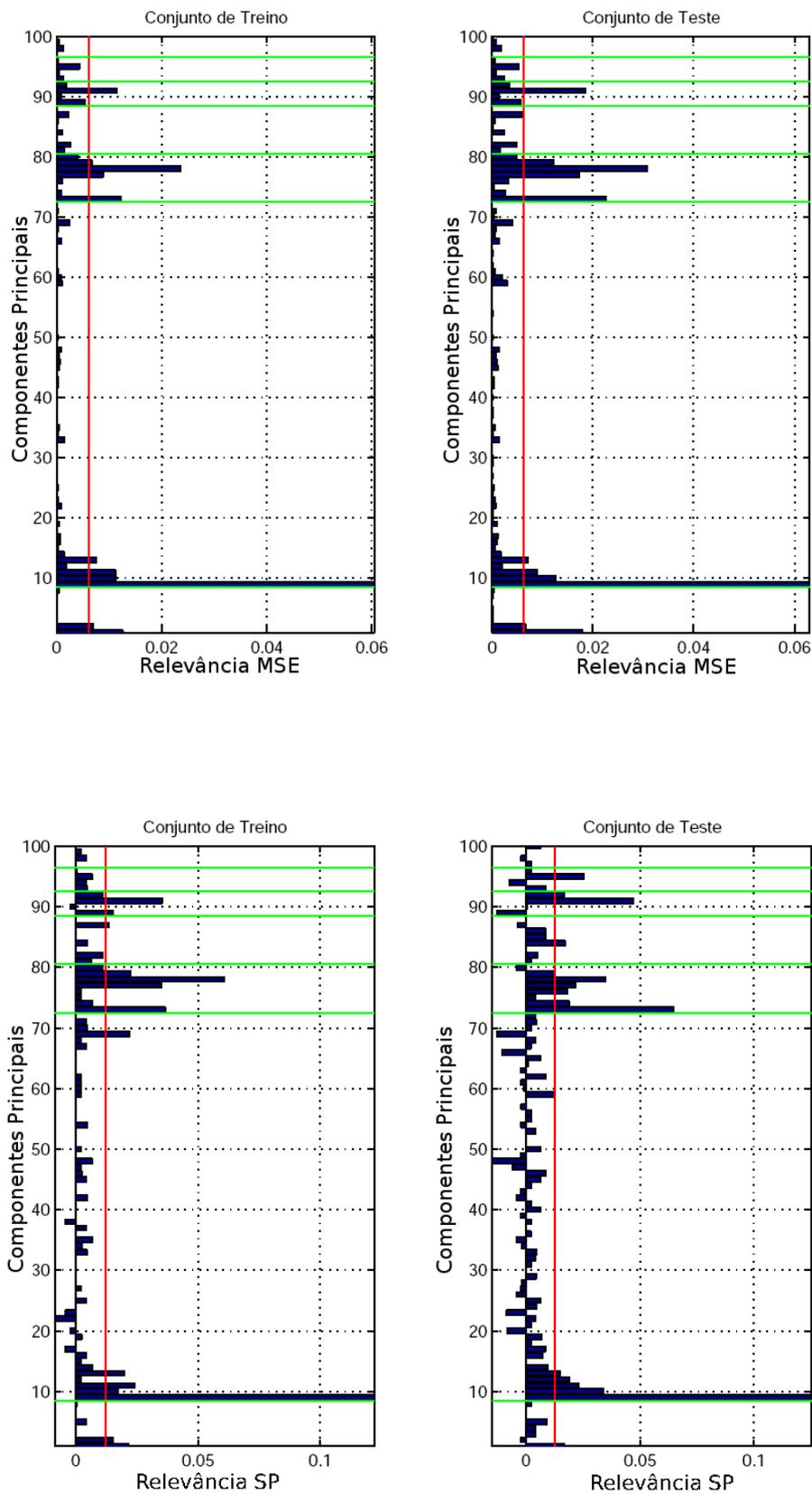


Figura 26: Relevância dos componentes principais extraídos de forma segmentada.

Comparando-se todas as compactações por relevância empregadas, pode-se destacar aquelas variáveis que foram ditas relevantes por todos os métodos, seja utilizando o conjunto de treino, seja o de teste. A Tabela 8 mostra os anéis e os componentes principais extraídos de forma segmentada ditos relevantes por todos os métodos (somente o segundo componente principal extraído de forma não segmentada foi considerado relevante por todos os métodos). Os números da Tabela 8 representam a ordem dos anéis e dos componentes principais (os anéis concêntricos são ordenados de dentro para fora, enquanto os componentes principais estão em ordem decrescente de energia).

Para os anéis ditos relevantes por todos os métodos, nem sempre os primeiros anéis das camadas, aqueles com maior energia, são os mais relevantes. Na 3ª camada e.m., por exemplo, o único anel considerado relevante é o oitavo anel, já distante do centro da RoI.

Já para os componentes principais extraídos de forma segmentada, o estudo de relevância mostrou que os componentes menos energéticos podem ser mais relevantes para a discriminação do que os componentes mais energéticos. Isso pode ser visto na extração de componentes principais na 3ª camada do calorímetro hadrônico.

Tabela 8: Variáveis que foram ditas relevantes por todos os métodos.

Camada do calorímetro	Anel	Componente Principal Segmentada
<i>presampler</i>	2	1
1ª camada e.m.	1, 2, 3	1, 2, 3, 5
2ª camada e.m.	1	1, 5, 6
3ª camada e.m.	8	nenhum
1ª camada hadrônica	1, 2	3
2ª camada hadrônica	nenhum	nenhum
3ª camada hadrônica	nenhum	nenhum

## 5 Resultados

Os discriminadores neurais projetados para as diferentes topologias e para as diferentes formas de compactação empregadas devem ser comparados. A comparação entre eles deve levar em conta não são só a eficiência de detecção, mas também a compactação utilizada e a arquitetura da rede neural artificial.

As redes neurais projetadas são do tipo *feedforward*, com 3 camadas (entrada, escondida e saída). O número de nós da camada de entrada é igual ao número de variáveis de entrada do discriminador. A camada de saída tem somente um neurônio, com sua resposta variando de -1 (jato) a 1 (elétron). Já para a camada escondida, o número de neurônios é variado de 2 a 15. A arquitetura que mostrar a melhor relação entre as eficiências de detecção de elétrons e jatos, para as 5 inicializações, será selecionada.

### 5.1 Resultados de classificação

Os diversos discriminadores neurais utilizados foram comparados de acordo com a curva ROC (do inglês *Receiver Operating Characteristics* [39]) de cada um e pelo produto SP alcançado. A curva ROC relaciona a eficiência de detecção (para elétrons) com o falso alarme correspondente. Conforme andamos na curva ROC, alteramos o valor do produto SP. Uma curva ROC ideal teria a forma de um quadrado, onde consegue-se 100% de eficiência de detecção com 0% de falso alarme, alcançando neste ponto, também, um produto SP máximo.

A Figura 27 mostra as curvas ROC das redes neurais treinadas compactando-se a informação por relevância. Nota-se que, se fixarmos um nível de 10% de falso alarme, todos os discriminadores mostram resultados semelhantes, apresentando 95% de eficiência

de detecção. Porém, se fixarmos em 5% o falso alarme, os cortes por relevância MSE apresentam uma eficiência de detecção levemente maior, com respeito àqueles adotados por relevância SP. Em todos os casos de compactação por relevância, a performance é cerca de 10 pontos percentuais melhor do que a alcançada com o uso da informação completa (100 anéis).

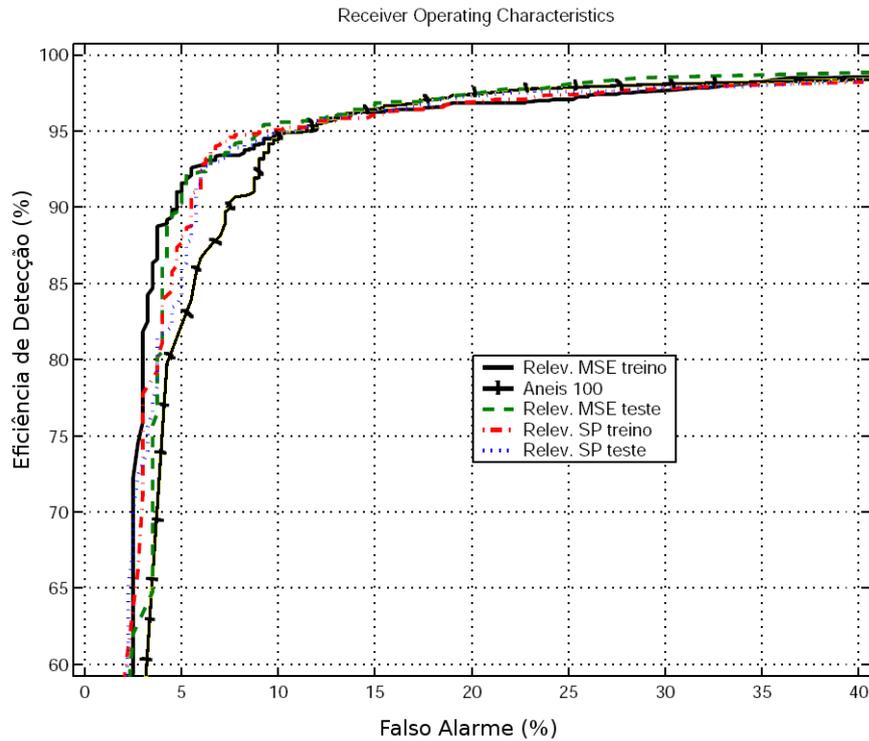
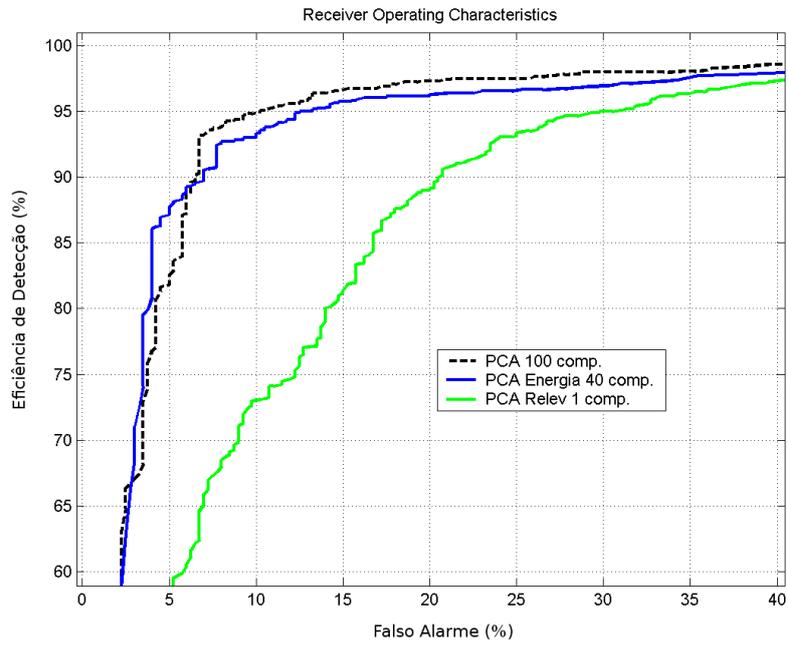


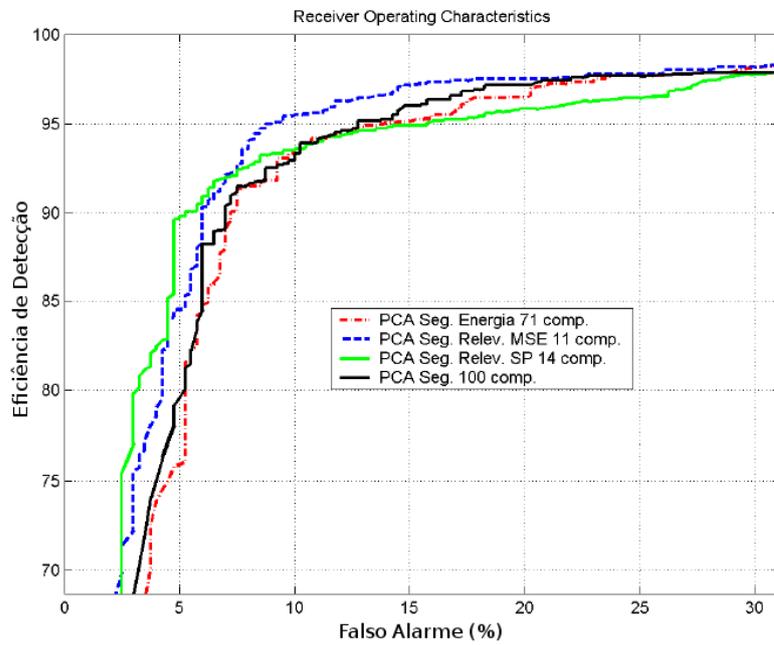
Figura 27: Curvas ROC dos discriminadores com anéis selecionados por relevância.

A Figura 28(a) mostra as curvas ROC para as análises com PCA não segmentada, enquanto a Figura 28(b) mostra as curvas ROC para as análises com PCA segmentada. As redes neurais adotando cortes por relevância do conjunto de treino e teste, por MSE e SP, tiveram desempenho similar e estão representadas somente pela curva ROC referente à seleção pelo conjunto de treino. Pode-se ver que o componente mais relevante da PCA não segmentada consegue resultados inferiores, porém significativos por se tratar de apenas um componente

Se fixarmos, também, um nível de falso alarme de 5%, o discriminador com corte por PCA segmentada e relevância por SP consegue uma eficiência de detecção entre 2,5 e 15 pontos percentuais melhor que os outros métodos. Já se fixarmos em 10% de falso alarme, os discriminadores têm praticamente a mesma eficiência de detecção.



(a)



(b)

Figura 28: Curvas ROC dos discriminadores com anéis selecionados por relevância.

Assim, considerando o nível de falso alarme e levando em conta o número de variáveis de entrada do discriminador em cada análise, a Figura 29 mostra as curvas ROC para os melhores discriminadores, nas abordagens por relevância e por análise de componente principais.

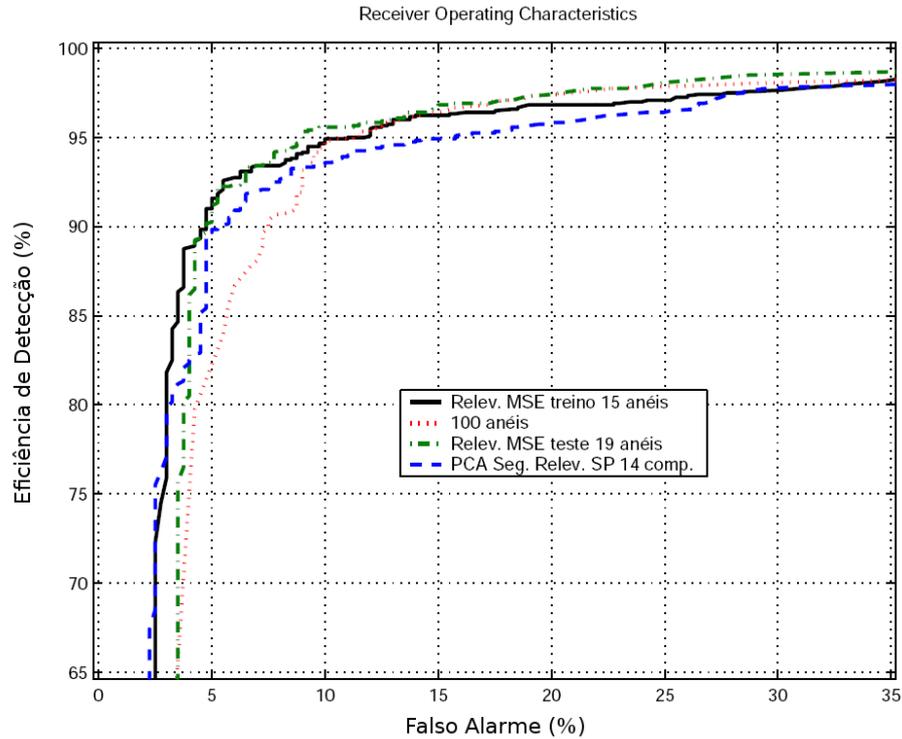


Figura 29: Curvas ROC dos melhores discriminadores, com os tipos de compactação empregados.

A Tabela 9 mostra o SP para um nível fixo de 5% de falso alarme, o SP máximo alcançado por cada discriminador e a arquitetura utilizada (número de nós da camada de entrada : número de neurônios da camada escondida : número de neurônios na camada de saída).

Tabela 9: Produtos SP dos melhores discriminadores, considerando um falso alarme de 5% e o SP máximo alcançado pelo discriminador.

	Anel 100	PCA Seg. Relev. SP	Relev. MSE treino	Relev. MSE teste
SP Falso alarme 5%	1,410	1,577	1,623	1,612
SP Máximo	1,578	1,591	1,629	1,612
Arquitetura / FLOP	100:2:1	14:7:1	15:4:1	19:2:1

## 6 Conclusões

Esse trabalho desenvolve discriminadores neurais que atuarão no segundo nível de filtragem do detector ATLAS, localizado no CERN. No ATLAS, inúmeros estudos sobre a origem da matéria serão feitos, principalmente a comprovação da partícula teórica bosón de Higgs. Por ser uma partícula rara, uma alta taxa de eventos é necessária para sua observação, gerando uma quantidade enorme de informação que não pode ser armazenada em mídia. Um sistema de filtragem *online* deve extrair as características necessárias para se discriminar as partículas de interesse. No caso desse trabalho, foi escolhido o canal elétron/jato, por ser o canal de maior frequência do Higgs.

Os discriminadores desenvolvidos mostraram um desempenho elevado com relação à classificação de partículas, mesmo operando no complicado e desafiante ambiente do LHC. O mapeamento das regiões de interesse em anéis concêntricos de deposição de energia consegue extrair características (físicas) importantes.

As técnicas posteriores de compactação utilizando informação estatística mostraram resultados interessantes. O estudo de relevância dos anéis mostrou que, dos 100 anéis extraídos originalmente, somente 8 são considerados relevantes, seja pelo método MSE, seja pelo SP. Basicamente, pode-se destacar alguns anéis das camadas do calorímetro e.m. e da primeira camada do calorímetro hadrônico. Nota-se que não só os anéis mais interiores, aqueles onde mais energia é depositada, são considerados relevantes.

A análise de componentes principais descorrelacionou a informação, porém sua seleção por energia (variância) mostrou uma performance pior do que sua seleção por relevância. Para os componentes extraídos de forma não segmentada, a seleção de somente um componente principal relevante deve ainda ser melhor estudada. O segundo componente

principal possui uma relevância muito elevada em relação aos outros componentes, mas seus resultados são inferiores aos resultados obtidos se utilizarmos, também, os componentes não considerados relevantes.

Para a extração de componentes principais de forma segmentada, a seleção de componentes por relevância mostrou uma performance mais elevada quando comparada com a performance da seleção feita por energia. O estudo de relevância dos componentes principais extraídos de forma segmentada, mostrou que somente os componentes principais de algumas camadas do calorímetro são relevantes à discriminação. Esse estudo também mostrou que, em algumas camadas do calorímetro, componentes principais com menor energia são mais relevantes que componentes mais energéticos.

## 6.1 Trabalhos futuros

Os estudos apresentados neste trabalho abrem espaço para muitos outros estudos. A performance dos discriminadores desenvolvidos deve ser comparada, também, com os resultados alcançados pelo T2Calo, assim como outras técnicas de compactação estatística podem ser utilizadas

Estudos anteriores mostraram que as variáveis extraídas pelo T2Calo podem servir de entrada para discriminadores neurais, alcançando resultados elevados [34]. Assim como para o espaço de anéis, as técnicas de compactação utilizada (relevância e análise de componentes principais) podem ser aplicadas ao espaço de variáveis do T2Calo.

Os resultados do estudo de relevância pelo método SP, seja dos anéis, seja dos componentes principais, mostraram a existência de variáveis confundidoras. Estudos mais profundos devem ser feitos sobre essas variáveis, como, por exemplo, a exclusão de somente essas variáveis para o projeto dos discriminadores.

Outras formas de compactação estatística devem ser estudadas: análise de componentes principais discriminantes e análise de componentes independentes. Como o detector opera a uma elevada taxa de eventos, acontece o efeito de empilhamento (do inglês *pile-up*), onde a informação do evento anterior ainda pode ser percebida no evento atual.

Essa informação, se possível, deve ser separada da informação atual, de forma a evitar a discriminação errônea das partículas.

# Referências

- [1] The European Laboratory for Particle Physics. Acessado em julho de 2007. <http://www.cern.ch>.
- [2] The Large Hadron Collider. Acessado em julho de 2007. <http://lhc.web.cern.ch>.
- [3] D. Griffiths. *Introduction to Elementary Particles*. John Wiley & Sons, 1987.
- [4] ATLAS Collaboration. “ATLAS: Technical Proposal for a General-Purpose Experiment at the Large Hadron Collider at CERN”, 1994. CERN/LHCC 94–43.
- [5] ATLAS Trigger and Data Acquisition Collaboration. “ATLAS High-Level Triggers, DAQ and DCS Technical Design Report”, 2000. CERN/LHCC 2000–017.
- [6] S. Gonzalez, T. Hansl-Kozanecka e M. Wieters. CERN preprint 002, CERN, 2000.
- [7] A. R. dos Anjos. “Sistema Neuronal Rápido de Decisão Baseado em Calorimetria de Altas Energias”. Dissertação de mestrado, COPPE/UFRJ, abril 2001.
- [8] T. M. Peruzzo e E. L. Canto. *Química na abordagem do cotidiano*. Editora Moderna, 1 edition, 1999.
- [9] B. Pullman. *The atom in the history of human thought*. Oxford Univ. Press, 2002.
- [10] R. Resnick e R. Eisberg. *Física Quântica – Átomos, Moléculas, Sólidos, Núcleos e Partículas*. Editora Campus, 1979.
- [11] D. Griffiths. *Introduction to Electrodynamics*. Prentice Hall, 3 edition, 1999.
- [12] ATLAS Collaboration. “ATLAS Inner Detector Technical Design Report”, 1997. CERN/LHCC 97–17.
- [13] ATLAS LArg Unit. “Liquid Argon Calorimeter Technical Design Report”, 1996. CERN/LHCC 96–41.
- [14] ATLAS/Tile Calorimeter Collaboration. “Tile Calorimeter Calorimeter Technical Design Report”, 1996. CERN/LHCC 96–42.
- [15] ATLAS Muon Collaboration. “ATLAS Muon Spectrometer Technical Design Report”, 1997. CERN/LHCC 97–22.
- [16] P. W. Higgs. “Broken symmetries, massless particles and gauge fields”. *Physics Letters*, 12:132–133, 1964.
- [17] B. Nicquevert et al. “The ATLAS coordinate system”, 1996. ATL–GE–CERN–QAP–0204–01.

- [18] R. Wigmans. “Advances in Hadron Calorimetry”. *Annual Review of Nuclear and Particle Science*, 1991.
- [19] J. Colas. “Overview of the ATLAS LArg Electronics”. Technical report, CERN, 1999.
- [20] K. Anderson, S. Berglund, G. Blanchot et al. “Front-end Electronics for the ATLAS Tile Calorimeter”. Technical report, CERN, setembro 1998.
- [21] A. S. Cerqueira e T. Davidek. “TileCalTrigger Signal Specifications and Testbeam Performance”. Technical report, CERN, fevereiro 2006.
- [22] A. S. Cerqueira, T. Davidek e G. L. Usai. “Tile Calorimeter Muon Trigger Signal”. Technical report, CERN, abril 2002.
- [23] ATLAS Trigger Performance Group. “ATLAS Trigger Performance Status Report”, 1998. CERN/LHCC 98–15.
- [24] S. Kilts. *Advanced FPGA Design: Architecture, Implementation, and Optimization*. Wiley-IEEE Press, 2007.
- [25] M. Abolins, S. Armstrong, J. T. Baines et al. “The Second Level Trigger of the ATLAS Experiment at CERN’s LHC”. *IEEE Transaction on Nuclear Science*, 51(3), 2004.
- [26] H. Schildt. *C++: The Complete Reference*. McGraw-Hill, 2003.
- [27] A. S. Tanenbaum. *Modern operating systems*. Prentice-Hall, Inc., 1992.
- [28] The ATLAS Project. “Athena User Guide and Tutorial”. Manual, CERN, agosto 2001.
- [29] ATLAS Trigger Performance Group. “ATLAS Trigger Performance Status Report”, 1998. CERN/LHCC 98–15.
- [30] The PESA Core Algorithms Group. “Algorithms for the ATLAS High-Level Trigger”. Technical report, CERN, junho 2003.
- [31] A. R. dos Anjos, R. C. Torres, J. M. de Seixas, B. C. Ferreira e T. C. Xavier. “Neural Triggering System Operating in High Resolution Calorimetry Information”. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 559(1):134–138, abril 2006.
- [32] C. Lindsey, B. Denby e T. Lindblad. “Artificial neural networks in high energy physics”. Acessado em julho de 2007. <http://neuralnets.web.cern.ch/NeuralNets/nnwInHep.html>.
- [33] J. M. de Seixas et al. “Neural Second Level Trigger System Based On Calorimetry”. *Computer Physics Communications*, 85:143–157, 1996.
- [34] T. C. Xavier e J. M. de Seixas. “Otimização do Segundo Nível de Trigger do ATLAS Baseado em Calorimetria”. *Encontro Nacional de Física de Partículas e Campos*, 1:1–5, 2006.

- [35] S. Haykin. *Neural Networks. A Comprehensive Foundation*. Prentice-Hall, Inc., 2 edition, 1999.
- [36] P. D. Wasserman. *Neural computing: theory and practice*. Van Nostrand Reinhold, 1989.
- [37] I. T. Jolliffe. *Principal Component Analysis*. Springer-Verlag, 2002.
- [38] J. M. de Seixas e H. P. Lima Júnior. “A segmented principal component analysis applied to calorimetry information at ATLAS”. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 559(1):129–133, 2006.
- [39] H. Van Trees. *Detection, Estimation and Modulation*. John Wiley & Sons, 1966.