

# Ensemble Neural para Identificação de Partículas baseada na Informação de um Calorímetro Finamente Segmentado

Werner Spolidoro Freund<sup>1</sup>, Guilherme Souza Sobrinho<sup>1</sup>, e José Manoel de Seixas<sup>1</sup>

Universidade Federal do Rio de Janeiro, Rio de Janeiro RJ 68504, Brasil

**Resumo** A Ciência de Dados tem natureza multidisciplinar e possibilita estender soluções de engenharia para outras áreas. Uma aplicação em Física de Partículas reflete desafios comuns de detecção de eventos de interesse, como empilhamento de sinais, alta taxa, eventos raros e alteração dos padrões dos eventos conforme posição de incidência no detector e intensidade do sinal. No caso, um método de *ensemble* de redes neurais tornou-se, em 2017, parte da seleção *online* de eventos de interesse de um dos maiores experimentos da atualidade (ATLAS), onde o *ensemble* atua como um classificador de partículas. Neste trabalho, avalia-se a extensão da técnica para atuação *offline* através da comparação com a técnica vigente, deteriorada para ambas acessarem o mesmo nível de informação, em dados simulados. A proposta obtém redução de falso alarme a, no mínimo, 70 % (valor central) da referência para toda a faixa de energia para detecção similar.

**Palavras-chave:** rede neural, *ensemble*, classificação, *big data*, física de altas energias

## 1 Introdução

A Ciência de Dados e sua comunidade emerge do avanço da Era Digital para promover soluções multidisciplinares a problemas recorrentes em diversos campos. Muitas vezes, essas soluções envolvem técnicas de aprendizado de máquinas tendo as mais variadas aplicações: processamento de linguagem natural, reconhecimento de fala e caligrafia, detecção de fraudes, análises do mercado financeiro, diagnóstico médico.

Este trabalho é um exemplo de aplicação de solução multidisciplinar em Física de Partículas envolvendo aprendizado de máquinas para seleção (classificação) de eventos de interesse em uma grande massa de dados com raros eventos de interesse, o que aumenta a exigência do sistema dedicado à seleção. Os eventos estão sujeitos a distorções de sinais durante a coleta de dados devido à alta taxa de eventos e ocupação dos detectores que geram empilhamento de sinais nos sensores do experimento. Outra fonte de distorções nos padrões deve-se a não homogeneidade da disposição e resposta dos sensores conforme a posição em que o

sinal se desenvolve, bem como em relação à energia total do sinal original. Como uma forma de aliviar essa dificuldade, o experimento possui elevado número de canais. Em função disso, decorre alta dimensão nas regiões de interesse avaliadas, contendo cerca de 1k a 1,2k sensores, onde se aplica conhecimento especialista a fim de manter a interpretação física do processo, porém reduzindo a dimensão para 100 entradas. As mesmas são propagadas a um conjunto de redes neurais (*ensemble*) responsável pela classificação. O emprego do *ensemble* neural é realizado para sujeitar os classificadores a uma menor variação dos padrões conforme a posição e energia do sinal original, empregando o treinamento e operação das redes neurais por regiões do alcance dinâmico dessas variáveis.

O contexto é desenvolvido na Seção 2, onde também se descreve o experimento (Subseção 2.1). O evento de interesse é descrito na Seção 3, na qual há disponível um método que emprega cortes lineares e sequenciais em grandezas altamente descritivas e discriminantes, resultado de 50 anos de conhecimento especialista na área (Subseção 3.1), representando o paradigma anterior, bem como o método vigente que emprega uma estratégia multivariável através da técnica *Naïve Bayes* em uma seleção dessas grandezas discriminantes (Subseção 3.2) e, também, a proposta (*Ringer*) deste trabalho acima resumida (Subseção 3.3). Os resultados em eficiência para comparação entre as técnicas estão na Seção 4. Conclui-se o trabalho na Seção 5.

## 2 Contexto

A busca de eventos de interesse ocorre em um dos maiores dos experimentos atuais, o ATLAS (Subseção 2.1), que possibilita a análise das colisões do acelerador LHC para o estudo de Física de Partículas e que teve papel importante na descoberta do bóson de Higgs [1], com o posterior laureamento de Peter Higgs e François Englert do Nobel de Física de 2013.

A Física de Partículas, dedicada ao estudo dos constituintes básicos do Universo e de como regem suas interações, atravessa por um momento de grande atividade para a pesquisa incluindo abordagens por aprendizado de máquinas. O CERN, uma referência de avanços científicos que se estende além da Física de Partículas e onde está situado o LHC, é um exemplo disso, instaurando-se recentemente (2015) um grupo de pesquisa inter-experimentos do LHC dedicado à aplicação de técnicas de Aprendizado de Máquinas [2], assim como escolas dedicadas (2015) [3] e seminários (2017) [4] para difusão de conhecimento e emprego de aprendizado de máquinas entre sua comunidade.

Este contexto está pareado com a atualização do programa de física do LHC [5] para novos objetivos — com a busca e caracterização de processos mais raros que o bóson de Higgs — e, também, da popularização da Ciência de Dados e de suas soluções multidisciplinares. Para possibilitar a acumulação de estatística e, com isso, a observação e caracterização desses processos extremamente raros, o LHC tem elevado gradualmente o número de colisões próton-próton ocorrendo por evento de cruzamento entre feixes, o que gera grande ocupação dos detectores dos experimentos, como no caso do ATLAS. Consequentemente, decorre uma

série de desafios para operação e análise da informação oriunda das colisões, ambos de extrema importância para o alcance dos objetivos do experimento, devido ao empilhamento de sinais.

Do ponto de vista de operação, um dos elementos essenciais para experimentos de Física de Partículas e bastante afetado pelo efeito de empilhamento é o Sistema de Filtragem. No ATLAS, tal sistema é responsável pela redução da alta taxa de eventos de 40 (33) MHz em pico (média) para uma taxa de saída média de 1 kHz. Realiza-se essa redução pela aplicação de cortes sequenciais em cadeias de filtragem com objetivos específicos para a análise de física. Como os processos físicos estudados são bastante instáveis, a sua observação é realizada através da reconstrução do processo original por mensageiros — partículas de interesse. Elétrons são mensageiros frequentes em diversas das análises físicas nos objetivos do ATLAS — como no caso do bóson de Higgs e de outros processos decorrentes de previsões de nova física. Assim, o ATLAS possui cerca de 100 cadeias de filtragem atreladas a coleta de eventos de colisões contendo elétrons como parte de seus resíduos que são armazenados para posterior análise *offline*.

Evitou-se o emprego de soluções multivariadas nessa área, em especial técnicas oriundas de aprendizado de máquinas, conforme sugerido na Ref. [6], por uma de questão introdução de sistemáticas de difícil análise nas amostras selecionadas, principalmente para atuação *online*. Nesse caso, perde-se acesso às amostras rejeitadas pelo classificador, o que dificulta a avaliação da sistemática introduzida pelo método. Porém, as condições mais severas introduzidas pelo empilhamento de sinais no detector abrem a oportunidade para a utilização de aprendizado de máquinas, uma vez que as técnicas convencionalmente utilizadas na área, com cortes lineares em grandezas discriminantes obtidas através de conhecimento especialista (Subseção 3.1), podem ser uma limitação para o experimento em manter os objetivos de estudo de Física de Partículas.

É neste contexto que se desenvolve este trabalho, como fruto de uma ideia de aplicação de redes neurais para classificação de elétrons [7] (sinal) e hádrons<sup>1</sup> (ruído) desenvolvida durante a década de 1990 para seleção *online* de eventos contendo ao menos um elétron no Experimento ATLAS. Somente recentemente (2017), com as condições mais severas de operação devido ao empilhamento, as cadeias de elétrons do ATLAS aderiram a essa abordagem, ainda que com uma série de atualizações [8, 9, 10] em relação a proposta inicial. Este trabalho, por sua vez, avalia a extensão da técnica (Subseção 3.3) para atuação na seleção de elétrons no ambiente de análise *offline*, dedicado a análise física, onde se espera eficiência máxima e o emprego de toda informação discriminante disponível.

## 2.1 Experimento ATLAS

O Experimento ATLAS [11] tem configuração cilíndrica<sup>2</sup>, medindo 45 m de comprimento e 25 m de altura, centrada no ponto de nominal de interação.

<sup>1</sup> Prótons, nêutrons são exemplos de hádrons comumente conhecidos.

<sup>2</sup> Refere-se a região central do cilindro por barril, enquanto as regiões remanescentes recebem o nome de tampas.

Compondo o experimento há conjunto de detectores, que se sobrepõem espacial e sequencialmente, sendo justamente o uso harmônico dessa coletânea de detectores que interagem e medem propriedades de natureza distinta que possibilita a busca e caracterização de diversos processos físicos. Este trabalho se refere às colisões próton-próton do segundo período de coleta de dados do LHC (*Run 2*), programadas para o período 2015–2018.

No ATLAS, utiliza-se um sistema de coordenada destro com origem no ponto de interação, onde a direção do eixo  $z$  segue a direção do feixe de partículas. O eixo  $x$  direciona-se do IP para o centro da circunferência do LHC e o eixo  $y$  direciona-se para cima. Coordenadas cilíndricas ( $r$ ,  $\phi$ ) são utilizadas no plano transversal, nas quais  $\phi$  é o ângulo azimutal em torno do eixo  $z$  e o ângulo polar  $\theta$  é o ângulo do eixo  $z$ . A pseudorapidez é definida em termos do ângulo polar  $\theta$  como  $\eta = -\ln \tan(\theta/2)$ .

O primeiro detector do ATLAS com o qual partículas carregadas interagem (como os elétrons e hádrons carregados) é o Detector Interno (ID, na sigla em inglês). Empregam-se três tecnologias para obter pontos de interação no espaço que são utilizados para extrapolar o traço (trajetória) da partícula, sequencialmente, começando pelo mais interno e também de maior precisão: Detector de Pixel (92 Mi de canais), Detector de Silício (SCT, 6,3 Mi) e Detector por Transição de Radiação (TRT, 350 k). Os dois primeiros são responsáveis por pontos de alta precisão, até 3 (8) pontos no caso do Detector de Pixel (SCT). Cerca de 36 outros pontos são fornecidos pelo TRT, que possibilita, além da extrapolação da trajetória, obter informação discriminante elétron-hádron pela probabilidade de emissão de radiação de transição. Exceto o TRT ( $|\eta| < 2,0$ ), os detectores mencionados fornecem cobertura total para a região dedicada a estudos de precisão do ATLAS ( $|\eta| < 2,5$ ).

Em seguida, dispõe-se o Sistema de Calorimetria fornecendo outra parcela da informação discriminante para separação elétron-hádron. Esse processo de medição, por sua vez, é destrutivo, onde a energia da partícula incidente é absorvida através de uma cadeia contínua de eventos que gera um chuveiro de partículas com geometria (a grosso modo) cônica centrada na trajetória da partícula. As partículas do chuveiro excitam o material ativo do calorímetro, gerando um sinal em uma determinada célula que é proporcional a energia depositada. Assim, a soma da energia depositada em diversas células do calorímetro é proporcional à energia da partícula original que gerou o chuveiro. O perfil do chuveiro de partículas depende da natureza da partícula permitindo poder discriminante em dois grupos: eletromagnético (elétrons) e hadrônico (hádrons); e, também, da energia da partícula incidindo no calorímetro.

A disposição padrão, empregada no ATLAS, é um calorímetro (EMCal) dedicado a absorção de partículas eletromagnéticas, dentre eles os elétrons, e outro externo a esse (HCal), que absorve a energia remanescente de hádrons. Cada um desses calorímetros fornece três amostragens longitudinais (camadas) no ATLAS, ou seja, no eixo de propagação da partícula através do experimento (profundidade), para a região de precisão do ATLAS ( $\eta < 2,47$ ). No caso do Sistema de Calorimetria, exclui-se da região de precisão a transição entre a tampa e barril

( $1,37 < |\eta| < 1,54$ ), onde há deterioração de sua resposta devido à presença de uma fissura nessa região com menor número de sensores para possibilitar o acesso de cabos e serviços do ID. Para recuperar parte da energia depositada no ID e em outros materiais anteriores ao Sistema de Calorimetria, está disponível um calorímetro pré-amostrador (PS, na sigla em inglês) que fornece uma amostragem longitudinal extra prévia ao EMCal, totalizando, assim, 7 camadas de amostragem de calorimetria.

Há variação na granularidade (segmentação lateral das células no plano  $\eta \times \phi$ ) de acordo com a posição de incidência da partícula em  $\eta$  mas, também, conforme as amostragens longitudinais. Esses calorímetros contêm, em conjunto, um total de 190 k [12] canais de leitura que fornecem cobertura completa em  $\phi$ , boa parte dos canais (170 k) oriundos do EMCal, particularmente de sua primeira amostragem que possibilita grande poder de discriminação elétron-hádron.

### 3 Identificação de Elétrons

A principal tarefa para identificação de elétrons [13] pode ser resumida em um problema de classificação de duas classes: elétrons (sinal) e hádrons (ruído). Na Subseção 3.1, descreve-se a estratégia inicialmente empregada pelo ATLAS para seleção *offline* de elétrons que é inspirada no método tradicionalmente utilizado na Física de Partículas: extraem-se grandezas altamente descritivas e discriminantes através de conhecimento especialista para aplicação de cortes lineares. Com o aumento da ocorrência de empilhamento no detector, o ATLAS adotou uma estratégia multivariável para identificação de elétrons descrita na Subseção 3.2, que é atualmente vigente. O método proposto (Subseção 3.3), por sua vez, explora soma de energia em anéis como entradas discriminantes para alimentar um *ensemble* de redes neurais. Ademais, por questão da qualidade de reconstrução dos traços das partículas necessário para o estudo de física contendo elétrons, em todos os casos aplica-se um corte exigindo ao menos 1 ponto no Detector de Pixel e 7 pontos nos dois detectores de precisão do ID.

#### 3.1 Cortes Lineares

Os chuveiros de partículas possuem informação discriminante para separação elétron-hádron e comumente exploram-se dois aspectos dessa informação: a largura lateral e a penetração longitudinal. Utilizam-se cerca de 9 grandezas com grande potencial de síntese do processo de desenvolvimento do chuveiro, como razões em energia, formato e largura do chuveiro. Além disso, outra fonte de informação discriminante é proveniente do ID, de onde se geram outras 7 grandezas. Essas grandezas são utilizadas para aplicar um conjunto de cortes rígidos sequenciais em três critérios para atender necessidades de análises físicas distintas e lidar com o caso cego de detecção em dados de colisão. Os mesmos são, ordenados crescentemente conforme o falso alarme: *tight*, *medium*, *loose*.

As grandezas sofrem alterações em seus padrões conforme a região em  $\eta$  que incidem no detector, tanto para as grandezas empregando informação do

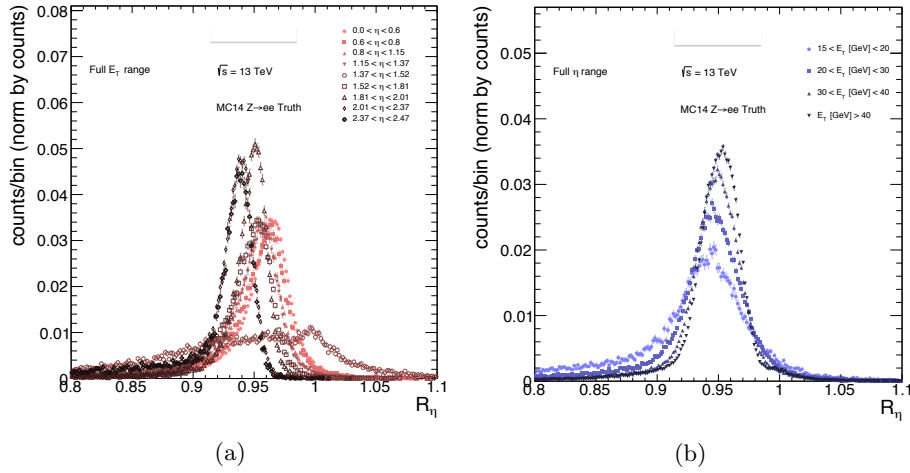


Figura 1: Perfil de alteração do padrão de elétrons em (a) quatro regiões em  $\eta$  e em (b) cinco regiões em  $E_T$  de uma das grandezas de calorimetria com o formato lateral do chuveiro na segunda amostragem do ECal para regiões empregadas pela LH.

calorímetro ou quanto do TRT. Citam-se estas fontes de distorção: geometria cilíndrica do detector, transição entre o barril e tampa, variação de densidade de material e granularidade de sensores disponíveis. Outra origem, para ambos detectores, de alteração dos padrões provém da energia da partícula incidente decorrente da dependência da resolução dos sensores nesse parâmetro (para calorimetria, *melhora* inversamente proporcional a raiz da energia, enquanto para o ID, *piora* diretamente proporcional a energia). Especificamente para o calorímetro, há alteração do desenvolvimento de chuveiros. Alguns exemplos dessas variações podem ser observados em  $\eta$  ( $E_T$ ) na Figura 1a (1b). Por isso, os cortes lineares variam em função de regiões de  $\eta \times E_T$ .

Empregou-se esse método para parte das análises da Run 1 (2010–2012), porém o ATLAS o substituiu pelo método LH descrito a seguir.

### 3.2 Verossimilhança (LH)

O ATLAS emprega uma estratégia multivariável baseada em *Naïve Bayes* para aumentar a eficiência devido às condições mais rígidas de operação. Isso permite o emprego de grandezas com maior sobreposição entre sinal e ruído e explorar a informação contida nas caudas das distribuições. Diversos outros métodos multivariáveis [14] foram avaliados pelo ATLAS, como redes neurais, *Boosted Decision Trees*, *kNN*, porém a escolha por *Naïve Bayes* se deu por sua simplicidade de construção e rapidez de execução. Otimizaram-se as grandezas empregadas pelo método  $n - 1$ , onde se mantiveram somente as grandezas cuja operação para  $n - 1$  é inferior a operação com  $n$ . Repetiu-se esse processo para

todos os critérios até obter a configuração de grandezas que permite a melhor performance para o ponto de operação desejado.

A fim de lidar com a alteração dos padrões, empregou-se a estimação das funções de densidade de probabilidade via KDE [14] gaussiano adaptativo (parâmetros *default* da referência citada) por regiões de  $\eta \times E_T$ , totalizando 9 (5) regiões em  $\eta$  ( $E_T$ ). As regiões são estas:  $E_T$  [GeV],  $15 \rightarrow 20 \rightarrow 30 \rightarrow 40 \rightarrow \infty$ ; e  $|\eta|$ ,  $0 \rightarrow 0,6 \rightarrow 0,8 \rightarrow 1,15 \rightarrow 1,37 \rightarrow 1,52 \rightarrow 1,81 \rightarrow 2,01 \rightarrow 2,37 \rightarrow 2,47$ . Os cortes também variam conforme regiões, porém, com uma grade mais fina em  $E_T$  (9 regiões), devido à menor necessidade de estatística quando em comparação com a extração de modelos via KDE. As regiões em  $E_T$  para limiares são  $15 \rightarrow 20 \rightarrow 25 \rightarrow 30 \rightarrow 35 \rightarrow 40 \rightarrow 45 \rightarrow 50 \rightarrow \infty$ .

Além disso, há o impacto do empilhamento sobre a taxa de detecção do classificador (ver Figura 4), que é tratado através de um ajuste linear na saída do classificador para cada uma das regiões de forma a manter a detecção em um *plateau*. Por uma necessidade das análises físicas, realiza-se interpolação linear na saída do classificador com o objetivo de evitar descontinuidades em detecção pela mudança brusca de limiares em  $E_T$ . Esses dois ajustes foram desativados para possibilitar a comparação entre as técnicas, uma vez que a versão herdada do *Ringer* do Sistema de Filtragem não aplicava essas correções na época de implementação no ambiente *offline*.

### 3.3 Método Proposto (Ringer)

Sabendo-se que os chuveiros tem geometria aproximadamente cônica e que a informação lateral e longitudinal neles contidas é discriminante, outra maneira de caracterizar esse processo e possibilitar compressão da informação em comparação com a dimensão das células é através da formação de anéis concêntricos ao centro de deposição de energia nos calorímetros. O anelamento (Figura 2) é repetido para todas as camadas dos calorímetros do ATLAS até atingir uma janela de  $0,4 \times 0,4$  no plano  $\eta \times \phi$ , gerando 100 anéis de, tipicamente, 1000–1200 células presentes na janela. Esse processo permite compressão da informação mantendo a interpretação física e informação discriminante contida no desenvolvimento do chuveiro de partículas.

Por herança de sua aplicação no Sistema de Filtragem, apenas a informação de calorimetria é empregada neste trabalho, embora esteja em desenvolvimento uma estratégia para fusão da informação do ID com o Sistema de Calorimetria quando empregando os anéis para operação *online* que pode ser facilmente estendido para operação *offline*.

Da mesma forma que as grandezas físicas, os anéis estão sujeitos a variação de seus padrões conforme  $\eta$  e  $E_T$  (Figura 3), de modo que se empregou estratégia similar àquela da LH, porém realizando o treinamento das redes neurais por regiões, totalizando 4 (5) regiões em  $\eta$  ( $E_T$ ). As regiões são estas:  $E_T$  [GeV],  $15 \rightarrow 20 \rightarrow 30 \rightarrow 40 \rightarrow 50 \rightarrow \infty$ ; e  $|\eta|$ ,  $0 \rightarrow 0,8 \rightarrow 1,37 \rightarrow 1,54 \rightarrow 2,5$ .

Os anéis normalizados por norma-1 alimentam o *ensemble* a fim de ajustá-los ao alcance dinâmico da função de ativação das redes neurais. Emprega-se o método *jack-knife* com 10 subconjuntos para a escolha do número de neurônios

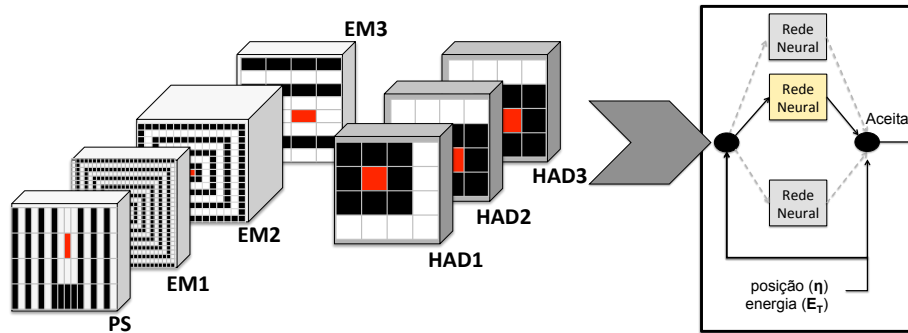


Figura 2: Esboço esquemático do método proposto (*Ringer*). Na esquerda, as células do calorímetro do ATLAS nas 7 camadas usadas para construir os anéis. O primeiro anel (centro de energia) está em vermelho, enquanto os anéis consecutivos se alternam em preto e branco. Na direita, a propagação dos anéis normalizados por norma-1 para a *ensemble*, onde a rede neural atuante depende da posição e energia da partícula, assim como o limiar de decisão aplicado.

(5–20) na camada escondida de cada rede neural MLP padrão do *ensemble* com uma camada escondida totalmente conectada, cuja função de ativação é  $\tanh(\cdot)$  para a camada escondida e neurônio de saída. Em cada sorteio, otimizam-se 100 novas redes, iniciadas conforme Ref. [15], pelo algoritmo da Ref. [16]. De um total de 320k redes por critério, mantêm-se somente aquelas 20 para compor o *ensemble* com menor falso alarme ( $P_F$ ) para a taxa de detecção de sinal ( $P_D$ ) mais próxima possível da referência para cada região em  $\eta$  e  $E_T$ . Caso ocorra empate entre arquiteturas, aplica-se parcimônia e emprega-se a arquitetura mais simples.

## 4 Resultados

Conjuntos de dados de simulação do ATLAS contendo condições similares às de colisões iniciais da Run 2 foram empregados para avaliar a possibilidade de aplicação da técnica para seleção elétron-hádron no ambiente de reconstrução da física *offline*. Nesses conjuntos estão candidatos a elétrons, dos quais 560k são elétrons verdadeiros e, outros 180k, hádrons. Calculam-se as eficiências através da chave de simulação, também referida por verdade de Monte Carlo.

Empregou-se uma comparação do *Ringer* com a LH modificada conforme mencionado na Subseção 3.2. Além disso, para possibilitar uma comparação justa, deteriorou-se a técnica LH para acessar somente informação de calorimetria (LHCalo). O discriminante da LHCalo foi obtido usando os mesmos *kernels* derivados pelo ATLAS para construir a LH, porém se calculou novos limiares para todas as regiões de operação da LH, de forma que o  $P_D$  da LHCalo fosse o mais próximo possível àquele da LH. Também se treinou o *ensemble* do *Ringer*



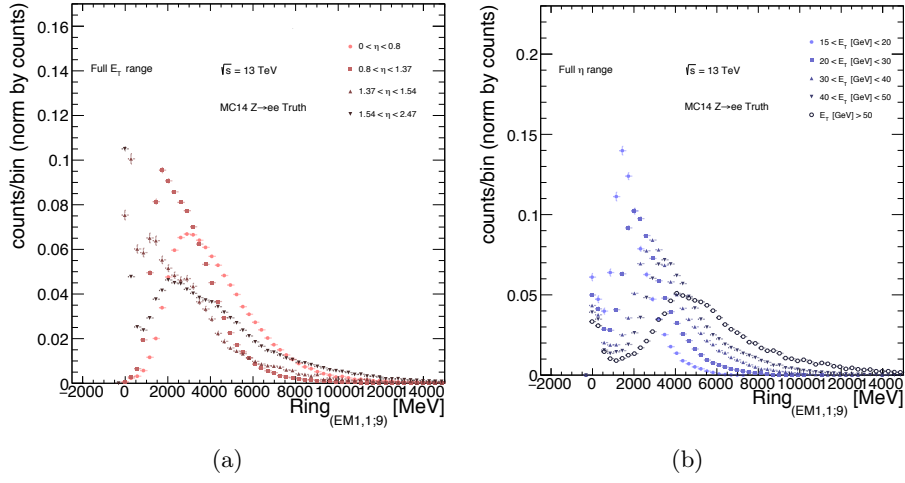


Figura 3: Perfil de alteração do padrão de elétrons em (a) quatro regiões em  $\eta$  e em (b) cinco regiões em  $E_T$  de um dos anéis mais relevantes, o segundo anel da primeira amostragem do ECal para regiões empregadas pelo *Ringer*.

de modo a obter, para cada um dos critérios de operação, o mesmo  $P_D$  que a LHCalo.

A Figura 4 mostra que o *Ringer* pode operar com eficiência similar a da LHCalo, porém se observa oscilação de cerca de 2 % em  $E_T$ , bem como uma maior oscilação para a região  $|\eta| > 1,37$ . No caso de  $E_T$ , essa oscilação deve-se da diferença da escolha das grades de regiões para o ajuste de limiares de operação. Assim, espera-se corrigir esse efeito ao empregar a mesma grade para limiares de decisão que a LH. O mesmo pode ser dito para  $|\eta| > 1,54$  (região da tampa), onde o *Ringer* emprega apenas uma região e a LHCalo, por sua vez, emprega 4, sendo o caso de  $|\eta| > 2,37$  o mais crítico. A região de transição do barril e da tampa do detector ( $1,37 < |\eta| < 1,54$ ) gerou uma eficiência menor em operação, mostrando que há menor capacidade das redes neurais nessa região de manterem o valor de  $P_D$  para os quais foram ajustados. Por sua vez, nessa região há menor disponibilidade de estatística para o ajuste dos discriminadores e maior variação dos padrões (ver figuras 1 e 3), esperando-se, assim, maior dificuldade de operação nesse caso. Ainda assim, essa região é descartada para análises de precisão. Em ambos casos em que o *Ringer* obtém detecção inferior, há espaço para recuperar o  $P_D$  com aumento de  $P_F$ .

Para todo o alcance de energia, exceto para o último *bin* onde não há certeza estatística suficiente, o *Ringer* opera com  $P_F$  inferior (razão inferior a 0,85) ao da LHCalo. Quando considerando apenas a região abaixo de 35 GeV, onde estão a maioria dos hádrons candidatos a elétrons, observa-se razão de  $P_F$  inferior a 0,65. Para a região do barril ( $|\eta| < 1,37$ ), também se observa aproximadamente esse valor na razão. Por outro lado, a tampa ( $|\eta| > 1,54$ ) possui maior oscilação na razão pelo mesmo motivo observado para o  $P_D$ . Em  $\langle \mu \rangle$ , observa-se o perfil

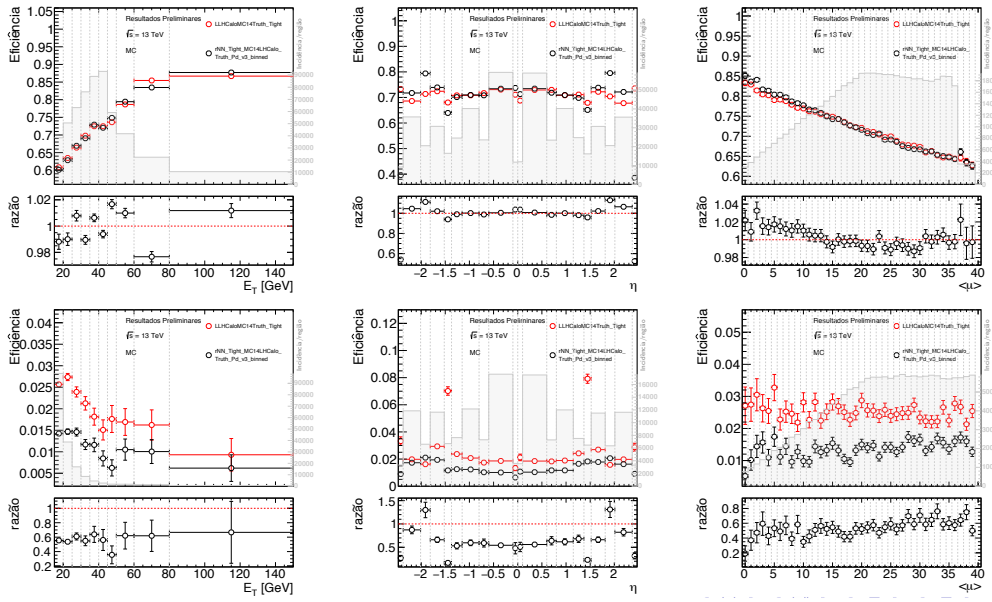


Figura 4: Detecção de elétrons (falso alarme) para o critério *tight* do método *Ringer* (preto) e LHCalo (vermelho) nas figuras superiores (inferiores) em função de, por ordem de coluna: energia ( $E_T$ ), posição ( $\eta$ ) e número de colisões inelásticas ( $\langle\mu\rangle$ ), um estimador do empilhamento de sinais. Na imagem inferior de cada figura está disponível a razão de eficiência do *Ringer* pela LHCalo. Em cinza, no fundo das imagens, está a contagem de candidatos a elétrons em cada *bin* do histograma. As barras de erro referem-se a propagação da incerteza estatística da divisão do número de candidatos que passaram o critério em relação ao total de candidatos. Calcula-se a barra de erro da razão assumindo desc correlação entre as técnicas.

oposto para a razão em  $P_D$  e em  $P_F$ , onde o *Ringer* apresenta uma queda ligeiramente maior na detecção conforme o aumento do empilhamento — ainda que continue operando com  $P_D$  aproximadamente dentro da faixa de 1 na razão, ambos com queda de  $\sim 20\%$  em  $P_D$  — e um aumento da razão em  $P_F$ . Tal efeito pode indicar uma ligeira maior dependência ao efeito de empilhamento do *Ringer* em relação a LHCalo, porém, com a aplicação da correção linear, mencionada na Subseção 3.2, é possível que esse pequeno efeito seja irrelevante. Além disso, o *Ringer* opera para toda a faixa do estimador de empilhamento com  $P_F$  inferior ao do LHCalo, enquanto o  $P_D$  se mantém aproximadamente no mesmo valor que a referência (razão  $> 0,98$ ).

Os critérios *medium* e *loose* também foram avaliados, onde se observou comportamento similar à Figura 4, porém com menor dependência da queda de detecção ao estimador de empilhamento (queda de 16% no *medium* e 10% no *loose*). A razão em  $P_F$  para a região abaixo de 35 GeV nesses casos é inferior a

0,70 em ambos critérios e também se observa o valor central sempre inferior ao da referência para toda a faixa dinâmica dessa variável.

## 5 Conclusão

Estendeu-se o *Ringer*, proposta que entrou em operação recentemente para filtragem *online* no Experimento ATLAS, para o ambiente de análise *offline* desse experimento. Uma comparação com a técnica vigente, deteriorada para as mesmas condições que a versão atual disponível do *Ringer*, mostrou que há viabilidade de aplicação do método nesse ambiente, onde o mesmo conseguiu obter detecção de elétrons similar àquela da LHCalo, porém com razão em  $P_F$  inferior a 0,65 (0,70) para o critério *tight* (*medium* e *loose*). Esse trabalho se aproveita de um ambiente propício para o seu desenvolvimento com a proliferação do aprendizado de máquinas em Física de Partículas, o que possibilita maior aceitação e emprego da proposta para análise física pela colaboração científica do experimento. Para que o método tenha adesão no ambiente *offline*, necessitou-se atenuar a influência do empilhamento de sinais na taxa de detecção, o que se espera obter ao aplicar o ajuste linear do limiar de decisão em função de estimadores desse efeito, estratégia que já obteve sucesso para a aplicação *online*. Além disso, é necessário incorporar ao método a informação do ID, obtida de sensores com natureza distinta de calorimetria. Para isso, pretende-se aplicar redes especializadas para possibilitar a fusão da informação de calorimetria e do ID.

## Agradecimentos

Desenvolveu-se essa pesquisa com o apoio de CNPq, CAPES, RENAFEA (MCTI), FAPERJ e do Núcleo Avançado de Computação de Alto Desempenho (NACAD) da COPPE, Universidade Federal do Rio de Janeiro (UFRJ); bem como da União Européia, que apoiou o trabalho através do projeto E-PLANET. Os autores agradecem também à colaboração do Experimento ATLAS pelas frutíferas discussões a respeito deste trabalho.

## Referências

1. ATLAS Collaboration. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. Em: *Phys. Lett. B* 716 (2012), p. 1. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].
2. *Inter-Experimental LHC Machine Learning Working Group (IML LHC)*. Acessado em 06 de Agosto de 2017. URL: <https://iml.web.cern.ch/>.
3. *Third Machine Learning in High Energy Physics Summer School 2017*. Acessado em 06 de Agosto de 2017. URL: <https://indico.cern.ch/event/613571/overview>.

4. *Data science seminars (CERN)*. Acessado em 06 de Agosto de 2017. URL: <https://indico.cern.ch/category/9320/>.
5. A Liss e J Nielsen. *Physics at a High-Luminosity LHC with ATLAS*. Rel. téc. ATL-PHYS-PUB-2013-007. Geneva: CERN, jul. de 2013. URL: <https://cds.cern.ch/record/1564937>.
6. Bruce Denby. “A bat’s ear view of neural nets in physics”. Em: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 389.1 (1997). New Computing Techniques in Physics Research V, pp. 8–11. ISSN: 0168-9002. DOI: [http://dx.doi.org/10.1016/S0168-9002\(97\)00030-2](http://dx.doi.org/10.1016/S0168-9002(97)00030-2). URL: <http://www.sciencedirect.com/science/article/pii/S0168900297000302>.
7. J.M. Seixas et al. “Neural second-level trigger system based on calorimetry”. Em: *Computer Physics Communications* 95.2–3 (1996), pp. 143–157. ISSN: 0010-4655. DOI: [http://dx.doi.org/10.1016/0010-4655\(96\)00012-4](http://dx.doi.org/10.1016/0010-4655(96)00012-4). URL: <http://www.sciencedirect.com/science/article/pii/0010465596000124>.
8. J. V. F. Pinto, W. S. Freund e J. M. de Seixas. “Sistema de filtragem de Online Utilizando um Ensemble de Redes Neurais e Informação de Calorimetria para Operar em Altas Taxas de Eventos”. Trabalho submetido para 13th Brazilian Congress on Computational Intelligence. 2017.
9. W. S. Freund, J. V. F. Pinto e J. M. de Seixas. “Redes neurais aplicadas na identificação de partículas em um detector finamente segmentado”. Em: *2016 XXI Congresso Brasileiro de Automática*. Out. de 2016, pp. 3512–3517.
10. ATLAS Collaboration. “Ring-shaped Calorimetry Information for a Neural Egamma Identification with ATLAS Detector”. Em: *Journal of Physics: Conference Series* 762.1 (2017), p. 012049. URL: <http://stacks.iop.org/1742-6596/762/i=1/a=012049>.
11. ATLAS Collaboration. “The ATLAS Experiment at the CERN Large Hadron Collider”. Em: *Journal of Instrumentation* 3.08 (2008), S08003. URL: <http://stacks.iop.org/1748-0221/3/i=08/a=S08003>.
12. ATLAS Collaboration. *ATLAS Run-2 Detector Status*. Disponível em: <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ApprovedPlotsATLASDetector>. Acessado em 12 de agosto 2017.
13. *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*. Rel. téc. ATLAS-CONF-2016-024. Geneva: CERN, jun. de 2016. URL: <https://cds.cern.ch/record/2157687>.
14. Andreas Hocker et al. “TMVA - Toolkit for Multivariate Data Analysis”. Em: *PoS ACAT* (2007), p. 040. arXiv: [physics/0703039](https://arxiv.org/abs/physics/0703039) [PHYSICS].
15. Derrick Nguyen e Bernard Widrow. “Improving the Learning Speed of 2-Layer Neural Networks by Choosing Initial Values of the Adaptive Weights”. Em: *Proceeding of the International Joint Conference on Neural Networks*. Vol. 3. San Diego, USA, 1990, pp. 21–26.
16. Martin Riedmiller e Heinrich Braun. “A Direct Adaptive Method for Faster Backpropagation Learning: The RPROP Algorithm”. Em: *Proceedings of the IEEE International Conference on Neural Networks*. San Francisco, USA, 1993, pp. 586–591.