

# Filtragem *Online* Segmentada Baseada em Redes Neurais Operando na Informação de um Calorímetro de Altas Energias de Fina Granularidade

João Victor da Fonseca Pinto

Orientador: José Manoel de Seixas

Universidade Federal do Rio de Janeiro (UFRJ), Programa de Engenharia Elétrica (PEE), COPPE

## I. INTRODUÇÃO

Ao longo das últimas décadas, a quantidade de dados gerados tem crescido de forma exponencial. O surgimento da internet aumentou de forma abrupta a quantidade de dados produzidos, e a popularização da internet das coisas fez sairmos da era dos terabytes para o petabytes.

Sistemas de filtragem de dados são ferramentas empregadas em diversas áreas da engenharia, medicina e física, com o objetivo de separar um conjunto de sinais de interesse daquelas que não agregam informação relevante ao problema. Atualmente, devido à complexidade de certos problemas, tem-se que o volume de dados a ser analisado torna-se bastante grande. Adicionalmente, o desbalanço entre a produção de eventos de interesse e ordinários pode tornar-se um fator limitante à aplicação, se não for introduzida a atuação de um sistema de filtragem.

Eventualmente, o descarte do ruído de fundo (dados não interessantes) adquirido por essa grande massa de dados, precisa ser feito de forma *online*, o que adiciona requisitos de velocidade (latência) e uso de memória ao problema. Nesse cenário, a latência para a realização da tomada de decisão geralmente se apresenta como um fator limitante ao emprego de técnicas mais sofisticadas de forma que, em geral, esses sistemas são projetados em níveis hierárquicos mais restritivos. Por fim, os eventos aprovados pelo sistema de filtragem *online* são armazenados em disco persistente e, posteriormente, podem ser analisados *offline*, onde se emprega plenamente a capacidade de processamento.

Esta tese de doutorado desenvolve um sistema *online* de seleção de elétrons para o ambiente de física experimental de altas energias, que estuda a estrutura da matéria e a origem do universo.

## II. APLICAÇÃO NA FÍSICA DE PARTÍCULAS

A física de partículas é o ramo da ciência que estuda as forças, ou interações, fundamentais da natureza e seus elementos, as partículas. Atualmente, no CERN, o acelerador principal *Large Hadron Collider* (LHC), podendo chegar até 14 TeV no centro de massa, é o mais potente acelerador de partículas construído até então pelo homem. Dentre os quatro

detectores instalados, dois deles são de propósito geral sendo um deles, o ATLAS, o maior experimento de partículas já construído. No ATLAS, o sistema de calorimetria, com seus mais de 200 mil, tem um papel fundamental de prover uma leitura rápida capaz de medir a energia da passagem das partículas pelo detector.

Em particular, os elétrons são partículas de extremo interesse para a rara física de interesse (bóson de Higgs, supersimetria, matéria escura, etc) e a utilização do calorímetro é imprescindível para a sua leitura. Com o objetivo de produzir estatística suficiente para obter outros processos ainda mais raros que o bóson de Higgs, o LHC vem aumentando a taxa de colisões próton-próton, estreitando cada vez mais a colimação dos feixes (luminosidade) de forma a aumentar cada vez mais o número de colisões por cruzamento (colisão) de pacotes.

Durante a primeira tomada de dados (*Run 1*), o LHC operou com 7 TeV de energia e com um espaçamento entre pacotes de 50 ns produzindo uma média de 20 interações por colisão. Todavia, durante a *Run 2* (2015-2018), o LHC obteve um recorde de 13 TeV de energia, operando com um espaçamento de 25 ns produzindo uma média de 60 interações por colisão. Com essa operação, o ATLAS produzia  $\sim 40$  TB/s de informação. Sendo aqui, indispensável o uso de um sistema de filtragem *online* de baixíssima latência e projetado de maneira eficiente capaz de selecionar as interações de interesse para a análise física (*offline*). Com o aumento da luminosidade e do número de interações esperados por colisão, o formato do sinal elétrico (pulso) extraído das células (sensor) do calorímetro é degradado.

Essa degradação do sinal elétrico resulta do empilhamento de sinais advindo de uma maior ocupação do detector e prejudica a estimação das energias das partículas, bem como altera os perfis de deposição dos elétrons. Portanto, manter alta eficiência na seleção de elétrons é um enorme desafio para o sistema de filtragem *online*. Neste contexto, esse trabalho desenvolve uma solução original, através de um *ensemble* de modelos, capaz de lidar com as condições de alto empilhamento, ao mesmo tempo que já se prepara para os novos desafios de operação para as *Runs 3 e 4* e o *Future Circular Collider* (FCC).

### III. CONTRIBUIÇÕES PARA O SISTEMA DE FILRAGEM

#### Online DO ATLAS

O sistema de *trigger* do ATLAS faz uso de reconhecimento de padrões para realizar a identificação dos objetos físicos de interesse. A filtragem é feita em dois níveis. O primeiro nível (L1), totalmente implementado em *hardware*, tem o objetivo de reduzir a taxa de entrada de  $\sim 40$  MHz para para 100 kHz. O segundo nível, totalmente implementado em *software*, chamado de *High Level Trigger* (HLT)<sup>1</sup>, tem por objetivo manter a taxa média de saída em 1,5 kHz com uma latência média em torno de 200 ms. Para a identificação de elétrons, o HLT é particionado em dois grandes blocos: rápido (*Fast*) e preciso (*Precision*). Durante a execução do bloco rápido são utilizadas as informações dos sistemas que possuem menor latência, como é o caso o sistema de calorimetria (etapa *FastCalo*), com posterior adição de parte da informação do detector de traços (etapa *FastElectron*), que é extremamente custosa computacionalmente; enquanto que, no preciso, são utilizados algoritmos de maior latência para reconstruir a energia da região de interesse (etapa *PrecisionCalo*) e, posteriormente, toda a informação dos detectores de traço e calorimetria são combinadas para reconstruir, com maior precisão que o bloco rápido, o candidato à elétron (etapa *PrecisionElectron*).

Em cada etapa, um algoritmo de seleção é utilizado para tomar a decisão de aceitar ou não o candidato baseado nas informações disponibilizadas naquele nível. Por fim, uma sequência de algoritmos, que combinam os sensores dos detectores capazes de extrair determinadas características do evento, e testes de hipótese formam uma cadeia de seleção responsável por tomar a decisão final de aceitar ou não o evento. Os eventos aceitos são, então, gravados para posterior análise *offline*.

#### A. NeuralRinger no FastCalo

Durante as colisões, os elétrons terão suas energias absorvidas nos calorímetros, gerando um chuveiro de partículas. Estes chuveiros, por sua vez, são descritos por meio de grandezas físicas que levam em consideração a largura lateral e longitudinal do chuveiro medido nos calorímetros<sup>2</sup>. Uma forma diferente de obter estas características do chuveiro é através da construção de anéis concêntricos de deposição de energia. O *NeuralRinger* [4]–[6], localizado na etapa *FastCalo*, utiliza esta formatação por anéis, onde, para cada anel, temos a soma das energias depositadas nas células que compõem, numa ROI selecionada pelo L1. Este processo é repetido ao longo de todas as camadas de amostragem do sistema de calorimetria do ATLAS permitindo, assim, a compactação das informações da largura longitudinal dos chuveiros ( $\sim 1000$  células em 100 anéis). Por fim, esses padrões são utilizados para alimentar um *ensemble*<sup>3</sup> de redes neurais, com patamares de cortes

ajustáveis, em relação ao ponto de operação e ao efeito de empilhamento, capazes de selecionar de forma eficiente o evento produzido durante a colisão. O projeto do *ensemble* do modelo *NeuralRinger* para operar no ATLAS em condições de elevado empilhamento de sinais e as possibilidades de novas atualizações no sistema de filtragem de elétrons do ATLAS para *Run 3* foram os objetivos desta pesquisa, que fez com que o *NeuralRinger* operasse com sucesso, desde 2017 (*Run 2*), mesmo com o aumento contínuo da luminosidade do LHC.

#### B. Resultados

Durante o primeiro comissionamento do *NeuralRinger* (2017) [1]–[3], a atuação do *ensemble* de redes neurais, do tipo *Multi Layer Perceptron* (MLP), alimentadas pela informação anelada (*ringer*) obteve um ganho em rejeição de falsos candidatos de  $1.5\times$  até  $3\times$  na primeira etapa (*FastCalo*), dependendo da cadeia de elétrons utilizada, em relação a antiga estratégia de seleção (*noringer*) adotada pela colaboração (cortes rígidos) desde o início da *Run 1* (2010). Após o sucesso do comissionamento, o *NeuralRinger* passou a ser o algoritmo de seleção padrão das cadeias de elétrons no *FastCalo*. Após concluída a *Run 2*, este trabalho focou nos avanços para a nova tomada de dados da *Run 3*.

Contudo, ainda durante a *Run 2* observou-se que as cadeias *ringer* possuíam uma perda em eficiência na detecção de elétrons para processos físicos em que era possível produzir léptons colimados<sup>4</sup>, chamados de *boosted*.

Para o comissionamento da *Run 3* (2023), este trabalho propôs uma redução no número de anéis na entrada dos discriminadores com o objetivo de reduzir artificialmente a largura lateral da informação processada pelos discriminadores, eliminando, por tanto, a interferência do segundo léptons na periferia da região de interesse. Os resultados mostraram que a redução no número de anéis (pela metade) não afetou o poder de discriminação dos classificadores e, ainda, foi capaz de recuperar as perdas, em detecção, observadas durante o último ano da *Run 2*. Assim, por esse motivo, essa estratégia foi escolhida, pela colaboração, para operar no início da *Run 3*.

Por fim, estratégias de fusão da informação e modelos profundos, como redes de convolução unidimensional, foram avaliadas com o objetivo de otimizar os ganhos em rejeição do sistema. As estratégias que combinavam, através de uma rede de fusão, os modelos neurais especialistas em calorimetria (*FastCalo*) e redes especialistas na informação do detector de traço, disponibilizadas pelo *FastElectron*, apresentaram uma redução de pelo menos  $2\times$ , em relação ao *NeuralRinger* da *Run 2*, quando todo o bloco rápido era utilizado para tomar a decisão baseada em uma estratégia multivariada, combinando os dois sistemas (calorimetria e traço).

<sup>1</sup>Composto atualmente por mais de 10 mil núcleos de processamento, em computadores do tipo PC.

<sup>2</sup>Aqui são utilizados os calorímetros eletromagnéticos (4 camadas) e os calorímetros hadrônicos (3 camadas).

<sup>3</sup>Composto por redes neurais especialistas treinadas para cada região do calorímetro e faixa de energia, sendo utilizadas, ao todo, 25 redes neurais no teste de hipótese do *FastCalo*.

<sup>4</sup>Em outras palavras, a energia do par elétron-pósitron, por exemplo, é tão alta que o ângulo entre as trajetórias dos léptons é pequeno a ponto de ambas as partículas atingirem o calorímetro em regiões próximas ou, no pior caso, sobrepostas.

#### IV. CONTRIBUIÇÕES PARA A SIMULAÇÃO DE EVENTOS EM CALORIMETRIA

Com os avanços da inteligência computacional em diversas áreas, em especial na física de partículas, iniciou-se uma busca por novas soluções que se beneficiem das informações mais elementares do problema. Especificamente no problema de classificação utilizando a informação da calorimetria, diversos trabalhos têm explorado modelos de computação visual a partir da informação das células para tomar uma decisão mais acurada. Todavia, na grande maioria dos experimentos, em especial no ATLAS, esse tipo de informação não é disponibilizada para os grupos de análise devido ao alto custo de armazenamento exigido. Ainda, a indisponibilidade de dados com diferentes perfis de empilhamento ou condições de operação não avaliadas pela colaboração, dificulta a disseminação de novas propostas e estudos, em especial no âmbito acadêmico.

Assim, com o objetivo de promover, mais facilmente, a disseminação de dados de simulação de colisões com diferentes condições de empilhamento e acesso as informações elementares do calorímetro, como as células e os pulsos resultantes da amostragem da energia, foi criado o *framework* de reconstrução *Lorenzetti* [7]–[9]. Embora que o ATLAS possua um sofisticado sistema de simulação, entendeu-se que o desenvolvimento desta ferramenta seria de extrema importância para o desenvolvimento de novos trabalhos na área que acabam sendo incapacitados devido a falta de dados desta natureza específica, inclusive para as condições previstas para *Run 4* e o FCC.

##### A. *Lorenzetti* : Um Framework de Reconstrução e Simulação de Partículas em Um Calorímetro Genérico

O sistema de calorimetria do *Lorenzetti*, baseado nos esquemáticos técnicos do experimento ATLAS, consiste em calorímetros simulados de amostragem envoltos por um campo magnético. A cadeia de simulação do *Lorenzetti* segue a mesma lógica de geração de eventos dos *frameworks* mais modernos restritos a colaboração como geração de eventos, simulação, digitalização e reconstrução dos objetos físicos.

A etapa de geração de eventos utiliza como núcleo a ferramenta *Pythia*, desenvolvida pela comunidade de física de partículas, para simular um evento de colisão e seus produtos. Aqui, o *framework* permite que o usuário realize a geração de determinados decaimentos ou processos físicos, chamados de evento principal, através de um conjunto de filtros específicos já disponibilizados. Para realizar a simulação de Monte Carlo é utilizado o *Geant4*, outra ferramenta também desenvolvida pela comunidade, para a simulação da passagem das partículas pela matéria do detector. Após a fase de propagação do evento pelo *Geant4*, inicia-se a etapa de digitalização composta pela emulação do pulso eletrônico e a estimação de energia das células do calorímetro. Ainda, esta etapa é responsável por simular os efeitos de empilhamento do sinal. Por fim, a etapa de reconstrução é responsável por acessar as informações das células do calorímetro e reconstruir as informações de alto nível referentes às regiões de interesse fornecidas pelo evento principal, como por exemplo, os anéis concêntricos de energia.

#### V. CONCLUSÕES

Os avanços discutidos neste trabalho permitiram o comissionamento do *NeuralRinger*, após quase 20 anos de sua concepção original pela COPPE/UFRJ, durante a tomada de dados da *Run 2* e, num segundo momento, sua atualização e manutenção para o início da *Run 3*. Ainda, foi explorado que o uso de toda a informação disponibilizada pelo sistema de reconstrução do bloco rápido, até então não utilizada pelo *NeuralRinger*, através da fusão da informação, poderá reduzir drasticamente o número de falsos candidatos antes do bloco de precisão.

Os testes realizados no simulador mostraram que a construção do calorímetro genérico e a montagem da sequência de reconstrução do evento foram capazes de produzir dados simulados (elétrons provenientes do bóson Z ou jatos hadrônicos) cujo perfis das grandezas de calorimetria estavam de acordo com o esperado pela Física. Em relação aos testes de simulação de empilhamento, observou-se a capacidade do simulador de produzir eventos de simulação em diferentes condições de empilhamento. Em termos de desenvolvimento de *software*, a construção do *framework*, inspirado majoritariamente no *framework* de reconstrução do ATLAS, permitiu uma rápida curva de aprendizado dos pesquisadores envolvidos devido a simplicidade de sua organização e abstração das ferramentas de simulação comumente utilizadas na física de partículas (*Geant4* e *Pythia*). Hoje, esta proposta conta com a colaboração de diversos pesquisadores (UFBA, UFJF, UERJ, COPPE/UFRJ e *Sorbonne/Paris*) e está em constante evolução para permitir os diversos estudos em calorimetria, além dos já cobertos pelo *trigger* do ATLAS.

Ainda é importante mencionar que as contribuições realizadas pelos mais de 10 anos de pesquisa do autor nesta área permitiram sua aprovação no programa de *authorship A* do ATLAS, em 2018, oficializando-o como autor do experimento; na sua aprovação no processo de doutorado direto sem a necessidade de defesa de mestrado pelo PEE/COPPE e na participação, como coordenador, através de eleição fechada, do grupo de pesquisa responsável pela filtragem *online* de elétrons e fótons do ATLAS (01 de abril de 2021 até 31 de maio de 2022). Adicionalmente, um artigo referente ao desenvolvimento do *NeuralRinger* durante a *Run 2* (44 páginas), e que será assinado pelos mais de 4 mil autores do experimento encontra-se em revisão pela colaboração. Por fim, mesmo que indiretamente, as contribuições produzidas por esse trabalho catalisaram diversas outras propostas de pesquisa brasileiras na linha de estimação de energia (dois professores da UFBA aprovados em processos de *authorship*, em conjunto com a universidade de *Sorbonne*, Paris); Na extensão da proposta deste trabalho para classificação de fótons (um aluno de doutorado da UFBA); na manutenção do grupo brasileiro na coordenação do grupo de pesquisa do ATLAS (atualmente gerenciado por um aluno de pós-doutorado da UFBA) e diversas outras propostas originadas deste trabalho.

## REFERENCES

- [1] Georges Aad et al. “Performance of electron and photon triggers in ATLAS during LHC Run 2”. Em: Eur. Phys. J. C 80.1 (2020), p. 47. doi: 10.1140/epjc/s10052019-7500-2. arXiv: <https://arxiv.org/abs/1909.00761>
- [2] Pinto, João. “Ring-shaped Calorimetry Information for a Neural Egamma Identification with ATLAS Detector”. Em: J. Phys. Conf. Ser. 762.1 (2016). Ed. por Luis Salinas e Claudio Torres, p. 012049. doi: 10.1088/17426596/762/1/012049 <https://iopscience.iop.org/article/10.1088/1742-6596/762/1/012049>
- [3] Pinto, João. “An Ensemble of Neural Networks for Online Filtering Implemented in the ATLAS Trigger System”. Em: J. Phys. Conf. Ser. 1162.1 (2019), p. 012039. doi: 10.1088/17426596/1162/1/012039 <https://iopscience.iop.org/article/10.1088/1742-6596/1162/1/012039>
- [4] Torres, Rodrigo. Sistema *Online* de Filtragem em um Ambiente com Alta Taxa de Eventos e Fina Granularidade. Tese de doutorado em engenharia elétrica, COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2010.
- [5] Freund, Werner. Identificação de Elétrons Baseada em um Calorímetro de Altas Energias Finamente Segmentado. Tese de doutorado em engenharia elétrica, COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2018.
- [6] Pinto, João. Filtragem *Online* Segmentada Baseada em Redes Neurais Operando na Informação de um Calorímetro de Altas Energias de Fina Granularidade. Tese de doutorado em engenharia elétrica, COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2022.
- [7] João Victor da Fonseca Pinto, Jlieberm, Werner Freund, Gabriel Gazola Milan, Micael Veríssimo de Araújo, and Guilherme Gonçalves. (2022). *lorenzetti-hep/lorenzetti*: 2.0.0 (2.0.0). Zenodo. <https://doi.org/10.5281/zenodo.7494463>
- [8] M. V. Araújo et al. “Lorenzetti Showers - A general-purpose framework for supporting signal reconstruction and triggering with calorimeters”. Em: Comput. Phys. Commun. 286 (2023), p. 108671. doi: 10.1016/j.cpc.2023.108671.
- [9] Pinto, João, “lorenzetti”, <https://github.com/lorenzetti-hep/lorenzetti>

## APPENDIX

### A. ATLAS Authorship

Para qualificar-se como um membro oficial da colaboração ATLAS o novo membro deve ter possuir um trabalho, ou proposta, relevante previamente aprovada pelo comitê. Após aprovada, o autor terá o prazo máximo de um ano para concluir a proposta e apresentar os resultados para todos os grupos envolvidos na área do experimento escolhida. Caso o trabalho seja reconhecido e aceito pela comunidade, o autor passa a integrar a lista de membros oficiais da colaboração ATLAS.