

Um Modelo Híbrido para a Otimização Neural do Desempenho de um Calorímetro

J.M. Seixas¹, P.V.M. da Silva¹, L.P. Calôba¹

¹COPPE/EE/UFRJ, CP 68504, Rio de Janeiro 21945-970, Brazil

E-mails: seixas@lps.ufrj.br, vitor@lacc.ufrj.br, caloba@lps.ufrj.br

Abstract

A neural mapping is established to improve the overall detector performance for a scintillating calorimeter, which is being designed to perform energy measurements in a next-generation high-energy collider experiment. Training a neural network with input vectors formed by the energy deposited on each cell of this granular detector; the original energy scale of the experimental particle beam is reconstructed and the linearity is significantly improved. In practice, the neural mapping corrects for nonlinearities that arise from the practical calorimeter design.

1. Introdução

Em física de altas energias, experimentos de colisão são realizados com frequência, como parte do estudo da estrutura da matéria. Partículas de alta energia são feitas colidir em pontos específicos e as reações obtidas são analisadas a partir de detectores colocados ao redor dos pontos de interação.

Para continuar a procura de novas descobertas, uma nova geração de colisionadores está sendo preparada. Neste caso encontra-se o Large Hadron Collider (LHC) que está sendo desenvolvido no CERN (Suíça) e deve entrar em operação no ano 2005. A importância do LHC é que ele irá basicamente reconstituir as condições presumidamente existentes apenas 10^{-12} segundos após o Big-Bang, a grande explosão que teria dado origem ao Universo [1].

O LHC possui dois laboratórios totalmente equipados com detectores, que irão medir os subprodutos resultantes das colisões. O ATLAS [2] é um destes laboratórios, e este trabalho enfoca um dos seus detectores, o calorímetro Tilecal.

Os calorímetros têm um papel muito importante, pois medem a energia das partículas incidentes que com ele interagem. Esses detectores possuem tipicamente um material pesado, que absorve a energia das partículas, e um material ativo, que amostra uma fração da energia que está sendo absorvida no detector. As partículas então perdem totalmente a sua energia por meio de um processo físico de decaimento, que produz uma cascata de partículas com valores de energia continuamente decrescentes [3].

O calorímetro não é apenas utilizado para a medição de energia, mas pode ser usado também para a identificação de classes diferentes de partículas, já que o perfil de deposição de energia é dependente da partícula incidente.

O Tilecal, no seu projeto prático, apresenta não-linearidades na sua resposta. Apesar desse fato não representar uma limitação dramática, seria interessante compensarem-se tais não-linearidades.

Normalmente para se fazer esse tipo de otimização são utilizadas funções multi-paramétricas, porém é bastante demorado e complexo obter-se o ajuste ótimo dos parâmetros utilizados por essas funções, que muitas vezes se apresentam dependentes da energia da partícula. A rede neural aparece como uma alternativa, devido a sua estrutura não linear e capacidade de realizar mapeamentos complexos com grande acuidade.

Nesse trabalho, visamos otimizar a resposta do calorímetro Tilecal através de uma rede neural que seja capaz de mapear, em uma escala linear, o perfil de deposição de energia produzido pelo calorímetro para partículas de diferentes energias. Com isto tentamos compensar os efeitos de não-linearidade, do projeto prático do detector.

Na Seção 2 discutiremos em detalhes a estrutura do calorímetro Tilecal e o tratamento empregado nos dados experimentais obtidos do calorímetro. Nas seções seguintes abordaremos a parte do mapeamento neural e seus resultados.

2. O Detector

Essa aplicação se refere ao Tilecal um calorímetro hadrônico de telhas cintilantes. O Tilecal fará parte do ATLAS, que possui ainda um detector interno de traços de partículas (para visualizar as interações próximas do ponto de impacto), um calorímetro eletromagnético (para medir a energia de partículas eletromagnéticas) e um detector de muons.

O Tilecal é um calorímetro especializado em medir a energia das partículas hadrônicas, como pions e jatos de partículas [4]. O Tilecal é segmentado em células o que permite que o calorímetro tenha acesso ao perfil detalhado de deposição de energia para partículas hadrônicas. Um módulo do calorímetro é constituído de ferro (material absorvedor), e de telhas cintilantes (material ativo). Fibras óticas são também utilizadas para o transporte dos

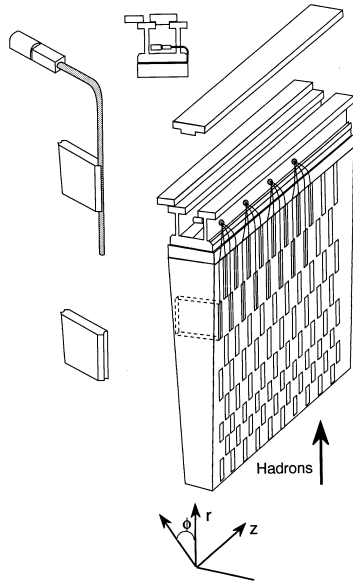


Figura 1: Um módulo do calorímetro.

sinais luminosos produzidos pelas telhas cintilantes (Veja Figura 1).

A partícula, ao incidir neste calorímetro, excita as telhas cintilantes, produzindo um sinal luminoso. Esse sinal é levado através das fibras ópticas (que se encontram acopladas nas laterais livres das telhas) até os tubos fotomultiplicadores, que irão converter o sinal luminoso produzido em sinal elétrico. Como para cada telha há 2 fibras ópticas acopladas, cada célula possui uma dupla leitura, correspondente à coleta do sinal de cada uma de suas laterais.

O sinal resultante é então amostrado, utilizando-se conversores analógico-digital de carga [5]. O detector possui segmentações radiais e longitudinais. Com isso toda a informação segmentada do calorímetro pode ser gravada para que mais tarde seja possível fazer-se uma análise *off-line* detalhada.

Recentemente, um protótipo de última geração do Tilecal foi submetido a um feixe de pions de diferentes energias. A Figura 2 mostra o arranjo experimental. O módulo central na figura refere-se a este protótipo avançado (Módulo 0) do Tilecal. Ele corresponde a um dos 64 módulos que irão cobrir toda a área central ao redor do ponto de colisão na montagem final. Para os testes realizados apenas metade do módulo foi instrumentado, de forma que esse protótipo produz 46 canais de leitura. Ao redor do Módulo 0 aparecem 5 outros módulos de uma geração anterior desta técnica de calorimetria, dois acima do Módulo 0 e três abaixo. Esses módulos têm como finalidade reproduzir (o melhor possível) a informação obtida pelos módulos vizinhos que irão existir no arranjo final do Tilecal.

A resolução em energia de um calorímetro é normalmente modelada por uma combinação linear de um fator

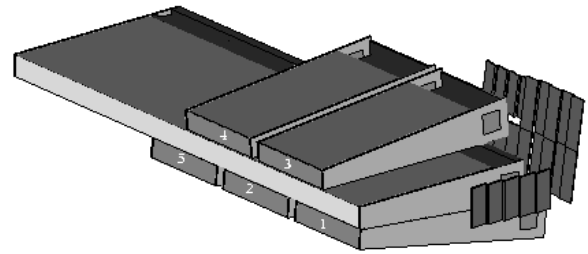


Figura 2: O arranjo experimental.

que escala com a energia (fator intrínscio) e um fator constante (que modela os efeitos instrumentais) [3]:

$$\frac{\sigma}{E} = \frac{A}{\sqrt{E}} + B \quad (1)$$

O fator de escala A representa a resolução intrínscia de um dado de calorímetro. O fator constante B representa as imperfeições da instrumentalização, levando em conta a montagem prática, ruído, erros sistemáticos, etc. O fator de escala é predominante para a faixa de mais baixa energia enquanto que o fator constante domina nas altas energias. O Tilecal tem apresentado em testes experimentais, quando combina-se linearmente as suas células, uma resolução de energia de $\frac{48,6\%}{\sqrt{E}} + 2,5\%$ [4].

Nos dados experimentais adquiridos de um feixe de pions, uma análise clássica de validação *off-line* de eventos foi realizada, produzindo cortes no conjunto de dados, com isso evita-se que partículas que na verdade não são pions sejam utilizadas no conjunto de treinamento da rede neural. Para pions de altas energias determinar tais cortes não é crítico, porque as partículas contaminadoras (muons) apresentam baixa energia e são facilmente identificadas. Porém, quando temos pions de baixa energia, eles se confundem com tais partículas que contaminam o feixe experimental de partículas. Para realizar a análise de validação de eventos utilizamos as informações provenientes de detectores adicionais colocados na linha do feixe (contadores de cintilação, câmaras de fios com alta resolução espacial e barreira de muons).

Para termos uma base de comparação com os resultados obtidos das redes neurais discutidas nesse trabalho, foram produzidos histogramas que mostram as distribuições da combinação linear das células do calorímetro para cada energia usada. Foram usados dados de pions de 10 diferentes energias, cobrindo a faixa 10 GeV até 400 GeV. A Figura 3 mostra essas distribuições. Dessas distribuições foram feitos *fittings* gaussianos e desses *fittings* retirados os valores médios e variância para cada energia nominal do feixe. Com esses dados foram avaliadas a linearidade e a resolução em energia do detector.

A partir dos valores médios, foi traçada uma reta ótica utilizando-se o método dos mínimos quadrados e calculou-se o maior desvio dos pontos experimentais em relação a esta reta. Obtivemos então uma não lineari-

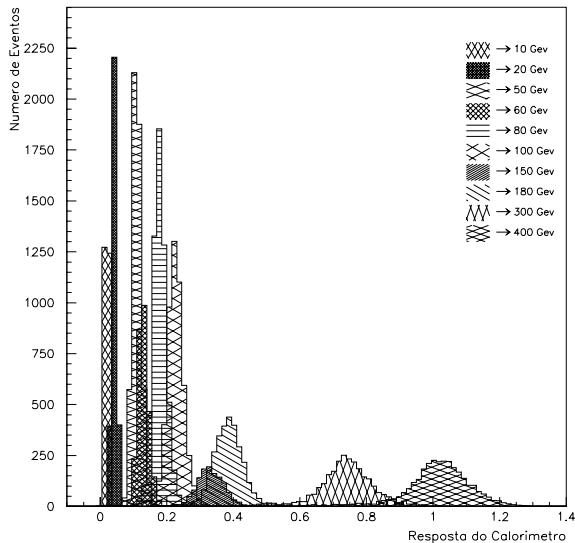


Figura 3: Distribuições normalizadas de energia quando se combina linearmente as células do calorímetro.

dade menor do que 10,6% e uma resolução em energia de $\frac{44\%}{\sqrt{E}} + 5,9\%$ para a resposta original do calorímetro, combinando-se linearmente a energia de cada célula. A Figura 4 mostra as curvas de linearidade e resolução.

3. Mapeamento Neural

A rede neural utilizada para realizar o mapeamento desejado tem como entrada os 46 sinais provenientes do protótipo Módulo 0. Esses vetores de entrada foram normalizados por um fator fixo igual a 0,1, levando-se em conta a distribuição dos valores de energia obtidos das células do calorímetro e a faixa dinâmica de operação da rede neural. A rede apresenta uma topologia com três camadas. A primeira camada possui 10 neurônios, a segunda 46 neurônios e a terceira camada (saída) tem apenas um neurônio. Uma rede neural com duas camadas internas apresentou uma melhor linearidade do que uma rede similar com uma camada escondida, diminuindo as não-linearidades para a faixa de energia mais baixa [6]. A função sigmóide foi escolhida para realizar a função de ativação das duas camadas intermediárias, sendo o neurônio de saída linear. Deste modo, a saída da rede reproduz a escala originalmente linear do feixe de partículas. Para realizar o treinamento, os dados experimentais foram divididos em dois conjuntos disjuntos, um de treinamento e outro de teste. Devida a boa estatística disponível, pode-se esperar que a rede venha a atingir uma boa generalização com o conjunto de teste, cujos eventos não são utilizados no processo de treinamento.

A rede foi treinada (por meio do pacote JETNET [7]) utilizando-se backpropagation como método de aprendizagem e como entrada o conjunto de treinamento composto por eventos com 10 diferentes níveis de energia nominal. A taxa de aprendizagem usada foi de 0,2 e de-

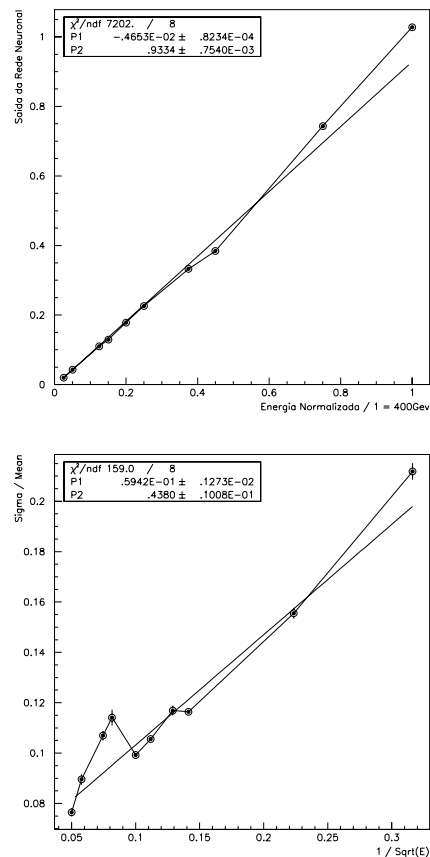


Figura 4: Linearidade (topo) e resolução (abaixo) originais obtidas do calorímetro.

crescia ao longo do treinamento. A época do treinamento foi de 10 eventos, que eram selecionados aleatoriamente dos 10 conjuntos de energias. Com isso um elemento de cada nível de energia nominal foi selecionado para o treinamento da rede a cada época.

Primeiramente foram utilizados como alvo os valores nominais de energia normalizados pelo valor máximo de energia (400 GeV). O critério de parada do treinamento da rede foi a máxima eficiência na estimativa da energia da partícula incidente, considerando-se uma resolução alvo de $50\%/\sqrt{E}$ (Veja equação (1)). Uma figura de mérito baseada no produto destas eficiências de estimativa para cada energia permitiu determinar a melhor rede mapeadora.

A Figura 5 mostra as distribuições de energia produzidas na saída da rede neural. Podemos observar que a rede foi capaz de identificar de forma bastante eficiente os valores nominais de energia. Em geral, o mapeamento neural produziu distribuições gaussianas para cada energia, excetuando-se a faixa de maior energia (acima de 100 GeV). Em verdade, parte do desvio observado do comportamento gaussiano, que se apresenta na forma de uma cauda acentuada na parte esquerda (menor energia) das distribuições, se deve ao efeito de vazamento de energia, provocado pelas dimensões reduzidas do protótipo Módulo 0. Portanto esta cauda também é ob-

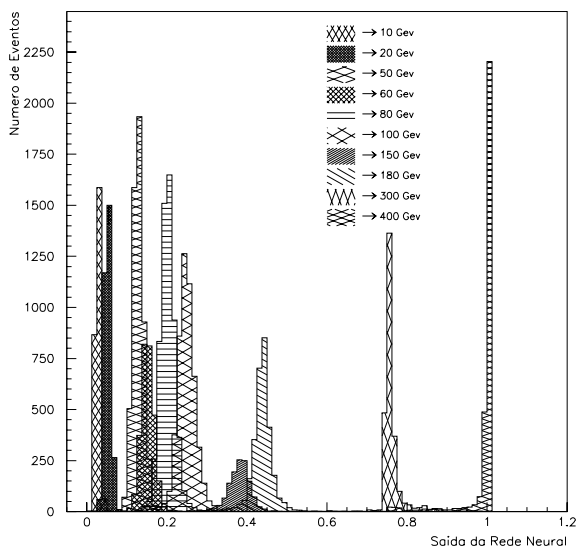


Figura 5: Distribuição da saída para a faixa (normalizada) de 10-400 GeV, quando o alvo no treinamento é o valor nominal da energia do feixe.

servada na distribuição original do calorímetro (Veja Figura 6). Para diminuir os seus efeitos de distorção, as figuras de resolução e linearidade do detector foram obtidas a partir dos valores médios e variâncias, extraídos de aproximações (*fittings*) gaussianas que se prolongaram até o canal para o qual o efeito de vazamento era notável.

Para a linearidade medimos o maior desvio dos pontos experimentais em relação a curva traçada, de onde tiramos que a não-linearidade ficou menor do que 2,9%. A resolução em energia foi obtida utilizando-se o modelo mostrado pela equação (1), para o qual obtivemos um fator de escala de 98% e um fator constante de 4,80%. A Figura 7 mostra as curvas de linearidade e resolução. Como podemos observar, o mapeamento neural produz melhora acentuada na figura de linearidade, em relação a combinação linear das células do calorímetro. Entretanto este mapeamento ao priorizar a linearidade não consegue manter a mesma resolução original do calorímetro. Este compromisso entre linearidade e resolução também é observado na solução clássica de estimação de energia baseada em funções multiparamétricas.

De modo a verificar a capacidade de generalização do mapeamento neural, testes foram realizados a partir de um conjunto de treinamento reduzido, onde algumas energias não se encontravam representadas. Assim, treinando-se com metade dos eventos com energia nominal igual a (10, 50, 60, 100, 150, 300, 400 GeV) e testando-se com a outra metade dos eventos com todas as energias disponíveis (incluindo-se ainda, os pions de 20, 80 e 180 GeV) obtivemos uma não linearidade menor do que 7,5% e uma resolução de $\frac{66,7\%}{\sqrt{E}} + 0,67\%$ (Veja Figura 8). Isto mostra uma extração robusta por parte do processamento neural de características do perfil de deposição de energia para pions.

Um outro teste significativo foi realizado, utilizando-

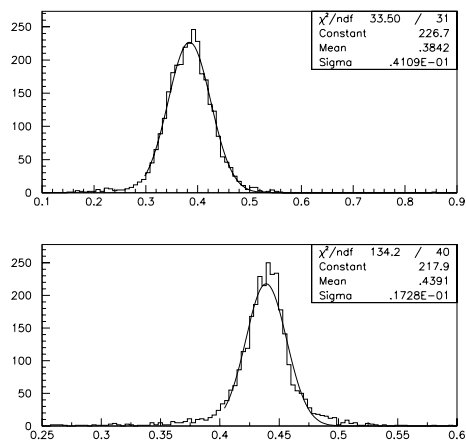


Figura 6: Efeito de vazamento para partículas de 180 GeV: distribuição original da resposta do calorímetro (topo) e do mapeamento neural (abaixo).

se como alvo no treinamento a energia medida pela soma das energias depositadas em cada célula do detector (resposta original do calorímetro). Deste modo pudemos verificar a capacidade de aprendizado da rede neural para realizar o mapeamento original do calorímetro. Em verdade, a rede reproduziu fielmente esta resposta, obtendo uma não linearidade menor do que 10,6% e uma resolução de $\frac{42\%}{\sqrt{E}} + 6,4\%$.

4. Modelo Híbrido

Com base nesses resultados foi desenvolvido um modelo híbrido de treinamento. Neste caso, a rede é treinada por um certo número de ciclos com alvos definidos pelos valores nominais de energia do feixe de partículas. Em seguida, o processo de aprendizagem prossegue por um determinado número de ciclos em que os alvos são obtidos da soma das energias depositadas nas células do detector (mapeamento linear da resposta original do calorímetro). Assim o processo de treinamento prossegue de modo alternado. Com essa abordagem procuramos combinar a resposta linear do mapeamento neural baseado na informação da energia nominal do feixe com a característica de resolução do calorímetro, quando se combina linearmente a informação das suas células.

Para se determinar o número de ciclos de treinamento realizados com alvo de energia nominal, a rede neural mapeadora da seção anterior foi inicializada com os pesos obtidos com o treinamento realizado a partir dos alvos definidos pela resposta original do calorímetro (combinação linear das energias das células). A partir desta inicialização, observou-se que são necessários cerca de 30000 passos para que o mapeamento neural corrija a forte não linearidade do detector e realize o mapeamento que privilegia a linearidade. Da mesma forma, inicializando-se a rede neural mapeadora com os pesos obtidos ao final do treinamento baseado em alvos relativos a energia nominal do feixe de partículas, treinou-se a

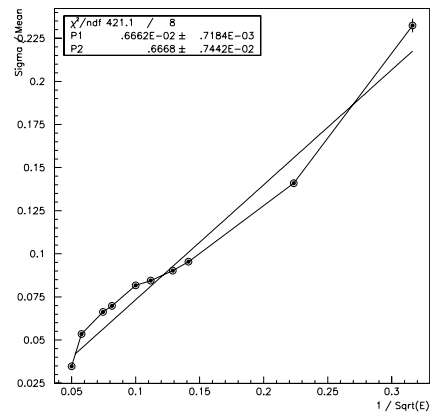
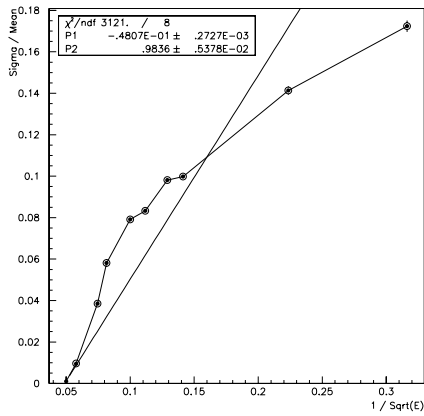
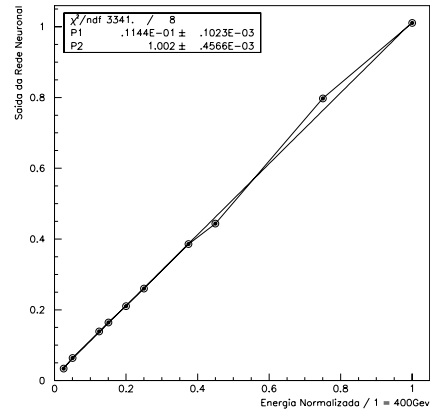
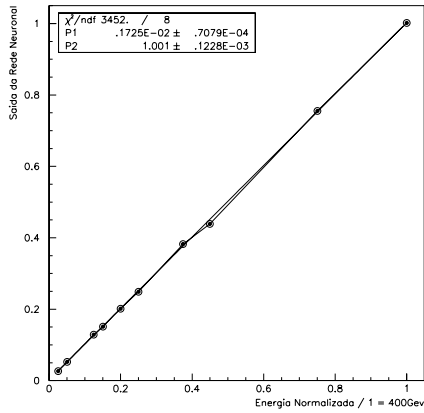


Figura 7: Linearidade (topo) e resolução (abaixo) obtidas pelo mapeamento neural, baseado na energia nominal do feixe.

Figura 8: Linearidade (topo) e resolução (abaixo) obtidas pelo mapeamento neural, com um número restrito de níveis de energia (Vide texto).

rede neural para que aprendesse o comportamento original do calorímetro, alcançando a resolução desejada. Para tal, foram necessários 7500 ciclos de treinamento.

Utilizando-se a mesma topologia para a rede neural (Seção 3), o modelo híbrido de treinamento foi implementado, realizando-se ciclos alternados de 30000 passos de treinamento com alvos de energia nominal e 7500 passos com o alvo definido pela energia medida originalmente pelo detector. A Figura 9 mostra a saída da rede ao final do treinamento e a Figura 10 mostra a linearidade e resolução alcançadas. Com este modelo, o mapeamento neural de energia apresenta uma não-linearidade máxima de 4% e uma resolução de $\frac{46,5\%}{\sqrt{E}} + 4,0\%$. Pode-se notar que a rede melhorou a característica não linear do calorímetro (10,6%), embora tenha deteriorado levemente a característica de linearidade do mapeamento neural puro, baseado na energia nominal do feixe (2,9%). No que se refere à resolução, também se atingiu o objetivo, melhorando-se o comportamento original do mapeamento neural, pois alcançamos uma resolução semelhante ao resultado obtido pela combinação linear das energias depositadas em cada célula.

5. Conclusões

Um mapeamento neural foi apresentado como alternativa para compensar as não-linearidades de um calorímetro em relação aos métodos clássicos (funções multi-paramétricas) tipicamente utilizados. Para um calorímetro hadrônico uma rede de três camadas de neurônios treinada pelo método de backpropagation foi capaz de mapear numa escala linear os 10 valores de energia da faixa de 10 GeV até 400 GeV, com uma não-linearidade máxima de 2,9%, abaixo da obtida pela resposta original do calorímetro. De modo a melhorar a resolução em energia deste mapeamento, um método híbrido de treinamento foi desenvolvido, alternando-se ciclos de treinamento em que se fez a rede neural se comportar como o calorímetro original com ciclos de mapeamento neural puro. Apesar da rede híbrida apresentar uma linearidade levemente inferior àquela obtida inicialmente com o treinamento supervisionado anterior, a resolução em energia foi aprimorada, atingindo-se $\frac{46,5\%}{\sqrt{E}} + 4,0\%$.

Com esses resultados mostra-se que uma rede neural pode construir uma escala linear de energia e melhorar a resolução em energia do detector de tal forma que pode-se estimar com bastante acurácia a energia de uma

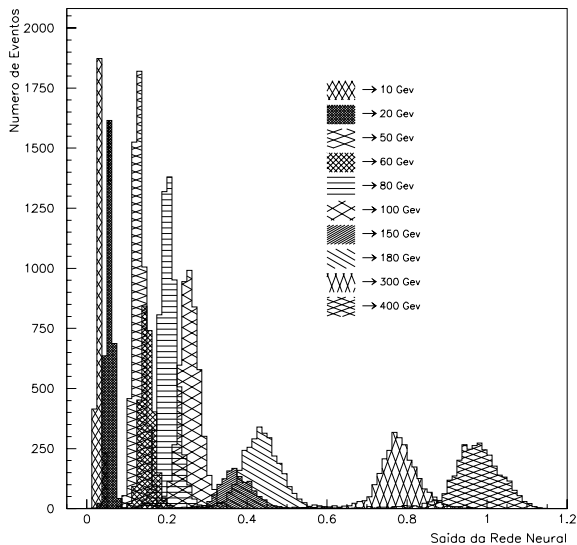


Figura 9: Distribuição da saída da rede para a faixa (normalizada) de 10-400 Gev, quando se utiliza o modelo híbrido de treinamento.

partícula incidente de alta energia em experimentos com colisionadores de partículas.

Em estudos futuros pretende-se utilizar as informações dos cinco módulos adjacentes ao protótipo de calorímetro estudado (200 canais a mais de informação), de modo a diminuir o efeito de vazamento observado no protótipo e aprimorar a estimativa de energia.

Agradecimentos

Gostaríamos de agradecer ao apoio dado a esse trabalho pela FUJB, CNPq (Brasil) e CERN(Suiça). Em especial queremos expressar nossa gratidão aos nossos colegas da colaboração do Tilecal no CERN, e particularmente, o grupo português de Lisboa, por nos fornecerem os conjuntos de dados e contribuirem com valiosas idéias nesse trabalho.

Referências

[1] Atlas - *Technical Proposal for a General Purpose pp Experiment at the Large Hadron Collider at CERN*. CERN/LHCC/94-43, 1994.

[2] The Tilecal Collaboration. *ATLAS - Tile Calorimeter Technical Design Report*. 332 páginas, CERN/LHCC/96-42, 1996.

[3] R. Wigmans. *Advances in Hadron Calorimetry*. *Rev. Nucl. Part. Sci.*, 41:133-148, 1991.

[4] F. Ariztizabal et al. *Construction and Performance of an Iron-Scintillator Hadron Calorimeter with Longitudinal Tile Configuration*. *Nuclear Instruments and Methods A349:384-397*, 1994.

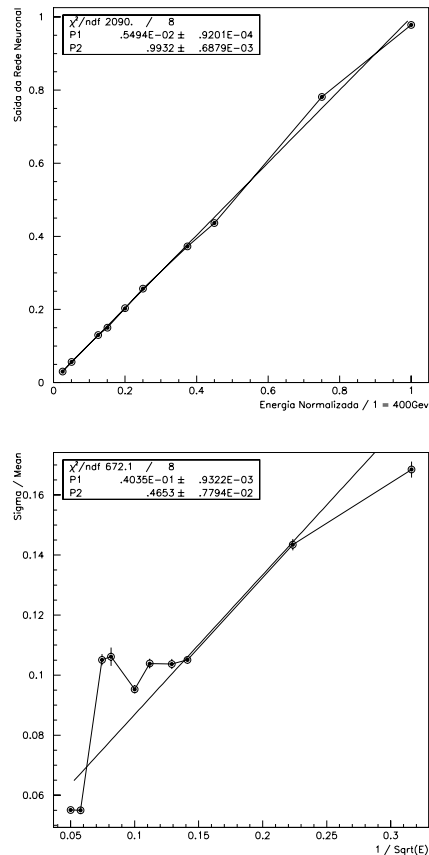


Figura 10: Linearidade (topo) e resolução (abaixo) alcançadas pelo modelo neural híbrido.

[5] P.W. Nicholson. *Nuclear Electronics*. John Wiley & Sons, 1974.

[6] M. Höppner, D. Wegener. *Applications of Neural Networks to Shower Analysis in a Highly Segmented LAr Calorimeter*. *Nuclear Instruments and Methods A389:154-156*, 1997.

[7] L. Lönnblad et al. *Pattern Recognition in High Energy Physics with Artificial Neural Networks - JETNET 2.0*. *Computer Physics Communications*, 70:167-185, 1992.