

## Estimação de Medidas de Altas Energias em Calorimetria utilizando Múltiplas Redes Neurais

Marcos da Rocha Vassali, Cássio Barboza Ribeiro, José Manoel de Seixas  
Laboratório de Processamento de Sinais - LPS  
COPPE/EE/Universidade Federal do Rio de Janeiro  
C.P. 68504, Rio de Janeiro, RJ - Brazil  
E-mails: {vassali, cassio, seixas}@lps.ufrj.br

### Abstract

*In high energy physics, calorimeters play an important role, as they are highly-segmented detectors that are able to measure the energy of incoming particles by combining hundreds of readout cells. In practical applications, calorimeters typically suffer from some degree of non-linearity, which must be compensated for to improve the final accuracy of their measurements. In this work, a neural technique for compensating for such non-linearities in calorimetry is presented. Combining the energy estimation of three artificial neural networks, each one dedicated to a specific energy range of the full dynamic range to be covered in the measurement, the energy of particles can be better estimated, when a comparison with classical linear combination of cells is performed. Experimental results show that such combined neural estimate achieves a linearity better than 4.5%.*

### 1. Introdução

Encontra-se em desenvolvimento no CERN (Laboratório Europeu para a Física de Partículas, Suíça) um novo colisionador de partículas, o LHC (Large Hadron Collider) [1]. Para a medição dos subprodutos destas colisões, são posicionados laboratórios de detecção ao redor dos pontos de colisão, dentre os quais destaca-se o laboratório ATLAS [2]. Diversos detectores são utilizados para estas medições, dentre os quais os calorímetros são de grande importância, uma vez que medem a energia das partículas resultantes das colisões [3]. Este trabalho concentra-se na melhoria da resposta de um calorímetro em particular, o TileCal, presente no laboratório ATLAS do LHC [2, 1].

O TileCal tem por objetivo a medição da energia de partículas hadrônicas, como *pions* e jatos de partículas [4], e esta medição é feita através da interação das partículas com o calorímetro, que acabam por absorver a energia destas partículas. Experimentalmente, confirma-se a alta acurácia destas medidas. No entanto, foi verificada a existência de alguma não-linearidade na resposta do calorímetro [4], o que conduziu a investigações de métodos clássicos, baseados em funções multi-paramétricas. Este enfoque se traduz nu-

ma tarefa demasiadamente complexa, uma vez que os parâmetros utilizados são dependentes da energia medida. Além disso, a compensação da não-linearidade no calorímetro tende a deteriorar sua resolução.

Foi implementada, neste trabalho, uma estimação neuronal da energia de uma partícula incidente, a partir do perfil completo de sua deposição no calorímetro TileCal. A abordagem que será detalhada ao longo deste trabalho utiliza três redes neurais, que cobrem toda a faixa de energia de 10 GeV a 400 GeV. Restringindo, assim, a região de atuação de cada rede neuronal (baixa, média e alta energia), obtém-se uma resposta com boa linearidade, acompanhada por uma aceitável deterioração da resolução do calorímetro.

Na próxima seção, são discutidas algumas das características básicas do calorímetro TileCal. Na Seção 3, os dados experimentais são analisados, e os métodos quantitativos para medição da não-linearidade e resolução são apresentados. A apresentação do estimador neuronal, bem como dos resultados obtidos com a abordagem proposta e com outros estimadores neurais (desenvolvidos em trabalhos anteriores), é feita na Seção 4. Por fim, a Seção 5 discute as conclusões observadas ao longo do desenvolvimento deste trabalho.

### 2. O Calorímetro TileCal

O TileCal é um calorímetro hadrônico de telhas cintilantes que, como já foi mencionado, realiza medidas de energia para o laboratório ATLAS do LHC. Quando uma partícula atinge o calorímetro, sua energia é totalmente absorvida por camadas de ferro, resultando em um conjunto de partículas de menor energia. As interações de partículas com o detector acabam por excitar as telhas cintilantes, que produzem luz proporcionalmente à energia da partícula incidente. A informação em forma de luz é então levada, através de fibras ópticas, à parte externa do detector, onde estão localizadas as foto-multiplicadoras, que convertem a luz em sinal elétrico. Por fim, este sinal é conduzido a um conversor analógico/digital, que produz a palavra digital correspondente à energia depositada na célula.

A Figura 1 mostra um módulo deste calorímetro. Na Figura 2 é possível observar que as fibras são arrumadas de modo a definir as segmentações radiais e longitudinais do calorímetro, resultando em 46 canais de saída

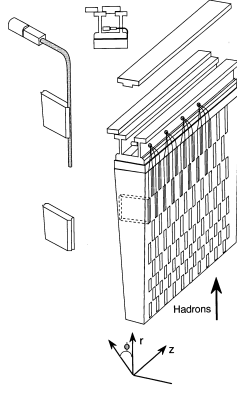


Figura 1: Um módulo do TileCal.

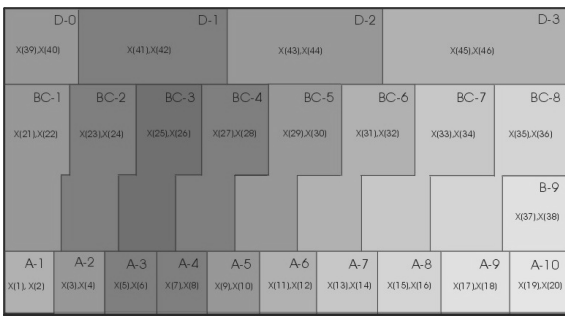


Figura 2: As camadas do TileCal.

por módulo do calorímetro. Para a completa absorção de energia no LHC, serão utilizados 192 módulos do calorímetro TileCal.

### 3. Análise dos Dados Experimentais

Foram feitos testes experimentais com o mais recente protótipo do TileCal com feixes de píons em diversas energias nominais: 10, 20, 50, 60, 80, 100, 150, 180, 300 e 400 GeV. Mais de 5000 eventos foram acumulados para cada energia nominal, permitindo, dessa forma, uma boa representação estatística das interações das partículas com o detector.

No entanto, não é possível a absorção de toda a energia de um feixe de píons por um único módulo do TileCal, resultando no que se chama de vazamento de energia. Este fenômeno torna ainda mais difícil a estimação da energia, principalmente para feixes de altas energias, nos quais o efeito de vazamento é ainda mais significativo. Isto pode ser confirmado através das distribuições de energia encontradas para cada energia nominal, quando as energias medidas em cada célula de calorímetro são somadas. Verifica-se que para energias até 180 GeV, as distribuições encontradas são aproximadamente gaussianas, enquanto que para as energias nominais de 300 GeV e 400 GeV, apesar da distribuição também assemelhar-se a uma distribuição gaussiana para valores próximos das energias nominais, é possível perceber uma calda estendendo-se para a região de mais baixa energia.

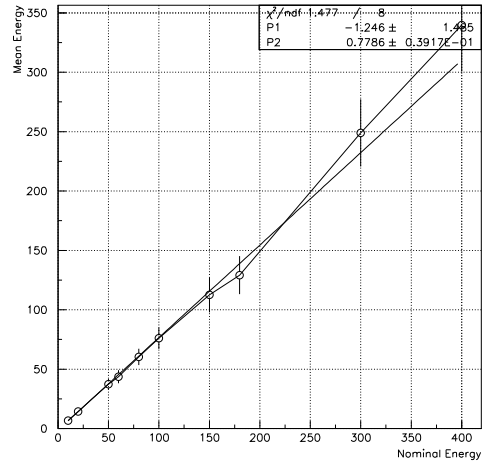


Figura 3: Análise de linearidade para a estimação linear.

A linearidade do calorímetro pode ser verificada a partir da regressão (ou *fitting*) linear sobre os valores médios de energia encontrados para cada energia nominal correspondente. A não-linearidade é percebida pelo fato de os pontos não se alinharem perfeitamente com a reta encontrada. A Figura 3 mostra o *fitting* linear obtido para os dados originais, quando a estimação da energia é dada pela combinação linear das energias nas células do calorímetro. A não-linearidade é medida quantitativamente pela máxima distância relativa entre pontos experimentais e aqueles obtidos através da reta encontrada como melhor solução segundo o método de mínimos quadrados. Para este método, verificou-se que a não-linearidade original do calorímetro era de 8,7%.

Para a análise de resolução do calorímetro, considera-se a distribuição obtida para cada energia nominal do feixe experimental. Teoricamente, a resolução de energia para o calorímetro segue o modelo dado pela equação 1.

$$\frac{\sigma}{E} = \frac{a}{\sqrt{E}} + b \quad (1)$$

Segundo este modelo, a resolução em energia depende inversamente da raiz quadrada da energia [3]. O parâmetro de escala  $a$  domina a resolução para baixas energias, e corresponde às características de resolução intrínsecas à técnica específica de calorimetria utilizada. Já o parâmetro  $b$  torna-se dominante em altas energias, e corresponde aos efeitos da instrumentação na leitura dos dados do calorímetro. Para o TileCal,  $a \approx 0,5$  é considerada uma boa aproximação, para uma faixa razoável de energia. Utilizando valores de  $\frac{\sigma}{E}$  calculados das distribuições encontradas, através de uma regressão (*fitting*) linear (Figura 4) foram obtidos os valores experimentais para estes parâmetros:  $a = 0,419$  e  $b = 0,062$ .

### 4 Mapeamento Neuronal

Para melhorar o desempenho do calorímetro, compensando sua não-linearidade observada, pode-se utili-

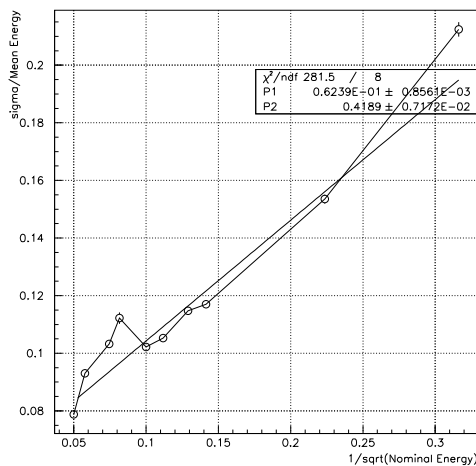


Figura 4: Análise de resolução para a estimação linear.

zar um mapeamento neuronal, de modo a obter uma combinação ótima das energias medidas em cada uma das 46 células do calorímetro, ao invés da simples combinação linear destas células [5].

Alguns estudos recentes vêm utilizando estimações através de redes neurais para as medições de energia do TileCal, utilizando, para isso, uma única rede neuronal para toda a faixa de atuação em energia (10-400 GeV) [6]. Verifica-se, experimentalmente, que a não-linearidade pode ser reduzida significativamente, degradando, no entanto, a resolução do calorímetro. Este fenômeno também é observado com a abordagem clássica através de funções multi-paramétricas, ou seja, qualquer melhoria na linearidade acaba por degradar a resolução.

Fisicamente, sabe-se que as fontes de não-linearidade dependem da energia das partículas que pretende-se medir. Dessa forma, surge a idéia de dividir a faixa de atuação do calorímetro em três regiões, e utilizar uma rede neuronal para aprender as peculiaridades de cada uma destas regiões. Em outras palavras, trata-se de utilizar múltiplas redes neurais para a compensação da não-linearidade do calorímetro, sendo que cada rede irá atuar em uma faixa de energia específica. Como este calorímetro opera em uma faixa consideravelmente extensa de energia (10-400 GeV), justifica-se a utilização de redes independentes atuando em diferentes regiões: baixa, média e alta energia, totalizando três redes neurais a serem utilizadas. A combinação linear das células foi usada para a estimativa inicial da energia da partícula incidente, permitindo direcionar a informação granular do calorímetro para a rede da respectiva faixa de energia e obter a estimativa neural da energia efetivamente depositada.

Dessa forma, a primeira rede neuronal (RN) concentrou-se em energias inferiores a 50 GeV. A segunda RN opera na faixa de energias que vai de 50 GeV a 180 GeV, enquanto que a terceira RN atua em ener-

gias superiores a 180 GeV. Os valores utilizados para divisão entre as regiões de atuação de cada rede neuronal (50 GeV e 180 GeV) foram escolhidos por razões físicas, uma vez que espera-se uma mudança de comportamento na deposição de energia dos feixes de píons entre cada uma das três regiões.

Todas as redes neurais utilizadas neste trabalho foram do tipo *feedforward*, completamente conectadas, e treinadas com o método de *backpropagation* [7]. A entrada da rede neuronal foi composta pelas energias medidas em cada uma das 46 células do calorímetro. A arquitetura das redes também foi a mesma, independente da região de atuação da rede, ou seja, foram utilizadas duas camadas escondidas, com 10 e 46 neurônio, respectivamente, e um único neurônio na camada de saída, que forneceu a estimação da rede neuronal para o evento. A utilização de duas camadas escondidas é capaz de prover uma melhor compensação de linearidade, principalmente para baixas energias [5]. A função de ativação para os neurônios das camadas escondidas foi uma sigmóide, enquanto que, para o neurônio de saída, a função de ativação utilizada foi linear, uma vez que desejou-se uma resposta linear para toda a faixa de atuação do calorímetro. Para o treinamento destas redes, utilizaram-se os valores nominais de energia dos feixes como alvo, unitariamente normalizados para a energia de 400 GeV. Todas as redes neurais foram simuladas com o pacote Jetnet 2.0 [8].

As Figuras 5, 6 e 7 mostram as distribuições obtidas através deste método de estimação neuronal. De uma forma genérica, a qualidade da estimação é muito boa, muito embora alguns fenômenos adversos tenham sido observados para as energias próximas aos valores de transição em 50 GeV e 180 GeV. Consideremos a distribuição original de energia (soma linear) para um feixe de 50 GeV. A estimação neuronal para eventos que tenham energia abaixo de 50 GeV, por exemplo, pode constituir um mapeamento diverso da estimação para eventos com energia maior do que 50 GeV, uma vez que estes mapeamentos serão feitos por redes neurais diferentes. Assim, para uma energia nominal próxima a estes pontos de transição, é possível que boa parte dos eventos seja compensada por uma rede neuronal, e outra parte por outra rede neuronal, extraindo diferentes características para o mapeamento. Evidentemente, isto levará a uma espécie de indecisão por parte do sistema neuronal responsável pela estimação, como um todo, o que é caracterizado pelo aparecimento de dois picos na distribuição na saída do estimador neuronal. Nas Figuras 5 e 6, isto pode ser observado para energias nominais de 60 GeV e 80 GeV<sup>1</sup>.

Para resolver este problema, propõe-se que a definição das regiões de atuação para as redes neurais admitam uma superposição, existindo, portanto, faixas de energia que podem ser compensadas por duas redes neurais diferentes. Este procedimento minimiza o proble-

<sup>1</sup>Vale lembrar que estamos lidando com apenas um módulo de calorímetro que, portanto, não é capaz de absorver toda a energia da partícula. Daí a explicação para o fato de os eventos correspondentes a energias nominais de 60 GeV e 80 GeV possuírem energias próximas ao valor de transição: 50 GeV.

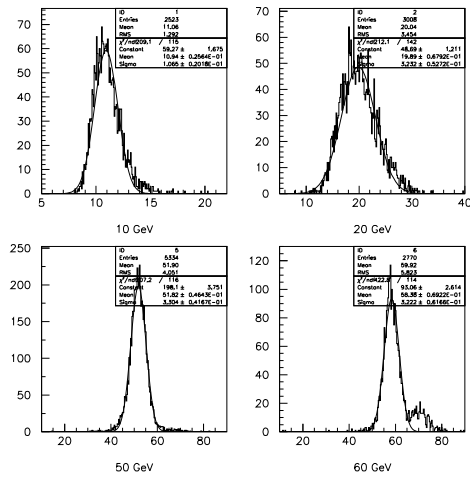


Figura 5: Distribuição dos eventos após o mapeamento neuronal e *fitting* gaussiano para energias nominais de 10, 20, 50 e 60 GeV.

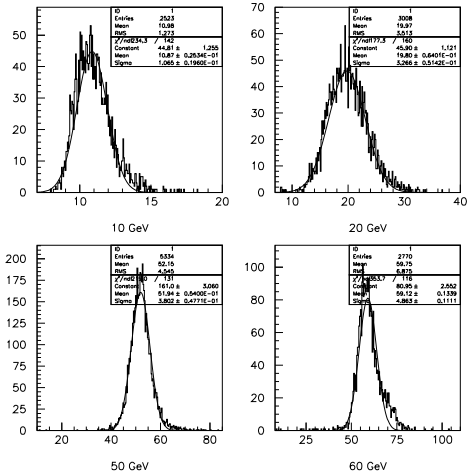


Figura 8: Distribuição dos eventos após o mapeamento neuronal com superposição e *fitting* gaussiano para energias nominais de 10, 20, 50 e 60 GeV.

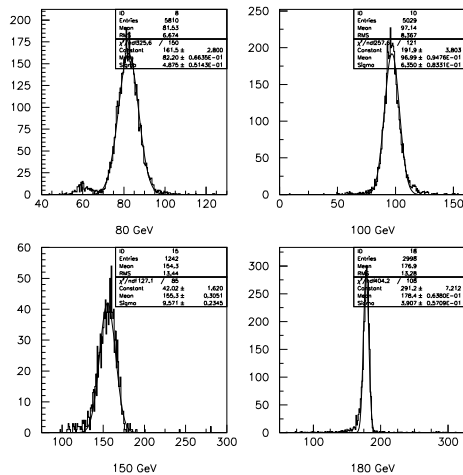


Figura 6: Distribuição dos eventos após o mapeamento neuronal e *fitting* gaussiano para energias nominais de 80, 100, 150 e 180 GeV.

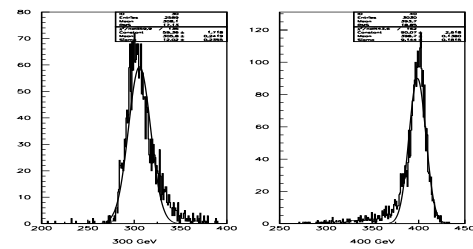


Figura 7: Distribuição dos eventos após o mapeamento neuronal e *fitting* gaussiano para energias nominais de 300 e 400 GeV.

ma das transições abruptas entre mapeamentos diferentes (devido a atuação de redes neuronais diferentes). Para isso, admite-se agora que a primeira rede neuronal irá operar em energias inferiores a 80 GeV (ao invés de 50 GeV). A segunda RN irá cobrir energias de 50 GeV até 180 GeV (como antes), enquanto que a terceira RN irá operar em energias superiores a 150 GeV (ao invés de 180 GeV). Não há mudanças na arquitetura das três redes neuronais. Para as regiões em que a estimativa é feita por duas redes neuronais (50-80 GeV e 150-180 GeV), a resposta final é dada pela média aritmética da saída das duas redes. Espera-se, portanto, uma transição bem mais suave entre regiões de operação de duas redes neuronais diferentes. Como resultado deste procedimento, é possível ver nas Figuras 8, 9 e 10, que exibem as distribuições após a estimativa neuronal com regiões de superposição, que não mais ocorrem dois picos na mesma distribuição, como ocorria no caso anterior, antes da introdução das regiões de superposição.

Os resultados obtidos são exibidos na Tabela 1, que compara as estimativas obtidas através do método linear (clássico) e não-linear (neuronal, com três redes e regiões de superposição). Os valores de média ( $\mu$ ) e desvio padrão ( $RMS$ ) calculados através dos histogramas são expostos em coluna própria, diferenciando-se da média e desvio padrão obtidos através de *fitting* gaussiano. Percebe-se claramente, através da análise desta tabela, que a estimativa neuronal apresenta resultados bastante superiores àqueles obtidos através da combinação linear, uma vez que, em geral, as médias aproximam-se mais dos valores nominais desejados, com um menor desvio padrão. De fato, o mapeamento neuronal não só compensa as não-linearidades, como também o efeito do vazamento de energia, e portanto, espera-se que os resultados de linearidade sejam bem superiores aos obtidos com a combinação linear.

Tabela 1: Resultados de Histogramas e *Fittings*

Energia	Estimador Linear				Estimador Neuronal			
	Histograma		Fitting		Histograma		Fitting	
	$\mu$	RMS	E	$\sigma$	$\mu$	RMS	Em	$\sigma$
10	6,8	1,5	6,8	1,4	11,0	1,3	10,9	1,1
20	14,3	2,3	14,3	2,2	20,0	3,5	19,8	3,3
50	37,3	4,6	37,3	4,4	52,2	4,6	51,9	3,8
60	43,7	5,5	43,7	5,0	59,8	6,9	59,1	4,9
80	60,4	6,9	60,3	6,4	81,6	6,8	81,9	6,0
100	76,1	9,1	76,3	7,8	97,9	8,4	97,9	6,4
150	112,6	14,8	112,3	12,6	154,1	16,1	154,3	12,0
180	129,2	16,0	130,0	13,4	177,8	14,5	178,7	7,3
300	249,0	28,3	250,8	23,3	309,3	21,9	310,7	17,3
400	339,8	39,2	345,7	27,3	390,3	22,3	394,2	13,5

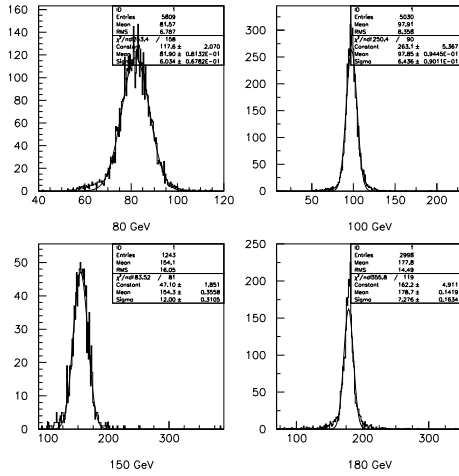


Figura 9: Distribuição dos eventos após o mapeamento neuronal com superposição e *fitting* gaussiano para energias nominais de 80, 100, 150 e 180 GeV.

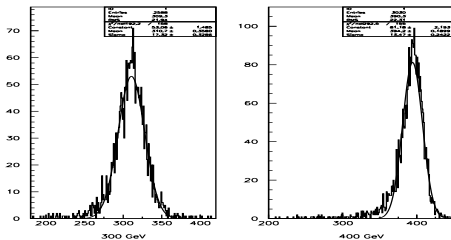


Figura 10: Distribuição dos eventos após o mapeamento neuronal com superposição e *fitting* gaussiano para energias nominais de 300 e 400 GeV.

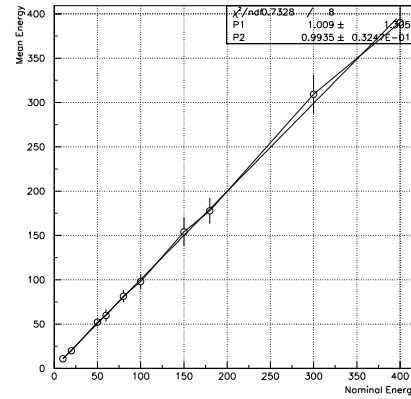


Figura 11: Compensação neuronal da não-linearidade na resposta do calorímetro.

As Figuras 11 e 12 mostram as regressões (*fitings*) lineares, obtidas através do método de estimação de energia com três redes neuronais, permitindo regiões de superposição. A não-linearidade encontrada foi de apenas 4,5%, bem menor do que a não-linearidade obtida pela combinação linear das células (8,7%). Em termos de resolução, a estimação neuronal obteve  $\frac{\sigma}{E} = \frac{0,35}{\sqrt{E}} + 0,025$ , o que também significa uma melhoria na resolução anteriormente encontrada, muito embora a regressão linear seja bastante pobre para este caso. Na tentativa de encontrar uma melhor descrição para a resolução da estimação, encontra-se em desenvolvimento um treinamento neuronal híbrido [9].

A Tabela 2 faz uma comparação dos desempenhos obtidos para diferentes métodos de estimação de energia aplicados ao calorímetro TileCal. Como já foi mencionado anteriormente, a combinação linear das células do calorímetro tende a obter boa resolução, porém com uma não-linearidade relativamente alta. Com uma rede neuronal apenas, obtém-se excelente correção da não-linearidade, às custas de uma maior deterioração da resolução. Resultados já obtidos com o treinamento

Tabela 2: Resumo do desempenho de diferentes estimadores de energia aplicados ao TileCal.

	Regressão Linear	Não-linearidade	Análise de Resolução
Combinação Linear das Células	$y = 0,779x - 1,246$	8,71%	$R = \frac{0,42}{\sqrt{E_m}} + 0,062$
Múltiplas Redes Neurais	$y = 0,994x + 1,009$	4,5%	$R = \frac{0,35}{\sqrt{E_m}} + 0,025$
Uma Única Rede Neuronal [9]	$y = 1,001x + 0,002$	2,9%	$R = \frac{0,98}{\sqrt{E_m}} + 0,048$
Treinamento Híbrido [9]	$y = 0,993x + 0,005$	4,0%	$R = \frac{0,46}{\sqrt{E_m}} + 0,040$

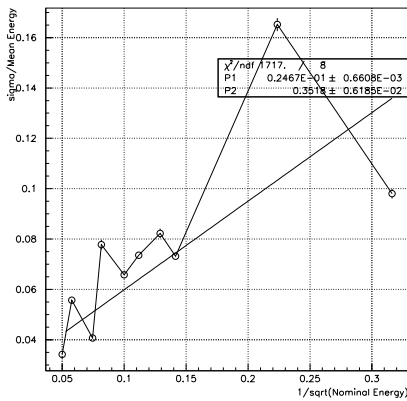


Figura 12: Compensação neuronal da resolução em energia na resposta do calorímetro.

híbrido mostram uma boa linearidade com a manutenção da resolução a valores comparáveis aos obtidos com a estimação linear. Percebe-se, ainda, que a utilização de múltiplas redes neurais, com superposição das regiões de atuação, fornece resultados de alta qualidade, tanto de linearidade quanto de resolução, ainda que a regressão linear obtida pela Figura 12 sofra de baixo nível de confiança.

## 5 Conclusões

Foi apresentada uma abordagem para a compensação das não-linearidades da resposta do calorímetro hadrônico TileCal. Trata-se da utilização de múltiplas redes neurais, cada uma treinada para características específicas de uma determinada faixa de energia, sendo que é permitida a superposição de parte das faixas de atuação de cada rede, quando, então, a saída estimada é dada pela média das respostas obtidas com cada rede neuronal. Foram obtidos resultados significativos na redução da não-linearidade, aliados a uma resolução bem superior àquela obtida quando apenas uma única rede neuronal é utilizada.

Este trabalho encontra-se em expansão, visando incluir métodos mais sofisticados para a combinação das respostas das redes neurais, nas regiões em que há

superposição nas regiões de atuação destas redes. A utilização de treinamentos híbridos para múltiplas redes neurais também encontra-se em fase de estudos[10].

## Agradecimentos

Agradecemos ao apoio financeiro fornecido por CNPq, CAPES e FAPERJ (Brasil). Agradecemos ainda aos colegas da colaboração TileCal (CERN - Suíça), pela disponibilidade dos dados experimentais. Finalmente, agradecemos em especial ao Dr. João Seixas (IST - Portugal) e Paulo V. M. da Silva (UFRJ), pelas frutíferas discussões pertinentes a este trabalho.

## Referências

- [1] The large hadron collider project. <http://lhc.web.cern.ch/lhc/>.
- [2] ATLAS/TileCal Collaboration. Atlas tile calorimeter technical design report (TDR) 3. CERN/LHCC/96-42, 1996.
- [3] R. Wigmans. Advances in hadron calorimetry. *Review on Nuclear Science*, pages 41–133, 1991.
- [4] F. Ariztizabal, et al. Calorimeter with longitudinal configuration. *Nuclear Instruments and Methods*, pages 349–384, 1994.
- [5] M. Höppner and D. Wegener. Applications of neural networks to shower analysis in a highly segmented lar calorimeter. *Nuclear Instruments and Methods*, pages 154–156, 1997.
- [6] J. M. de Seixas, P. V. M. da Silva, and L. P. Calôba. A neural mapping for improving the performance of a high-energy calorimeter. *IEEE Computer Society Edition for the Brazilian Symposium on Neural Networks*, pages 204–209, 1998.
- [7] S. Haykin. *Neural Networks: a Comprehensive Foundation*. Prentice-Hall, 2nd edition, 1999.
- [8] L. Lönnblad, C. Peterson, and T. Rönvaldsson. In *Computer Physics Communications*, volume 70, 167. 1992.
- [9] J. M. de Seixas, P. V. M. da Silva, and L. P. Calôba. Um modelo híbrido para a otimização neural do desempenho de um calorímetro. pages 332–337, ITA, São José dos Campos - SP - Brasil, 1999. IV Congresso Brasileiro de Redes Neurais.
- [10] M. R. Vassali, C. B. Ribeiro, P. V. M. da Silva, J. M. Seixas, and J. Seixas. Estimating high-energy measurements in calorimetry with multiple neural networks. Rio de Janeiro - RJ - Brasil, 2000. SAWCAS 2000 1<sup>st</sup> South American Workshop on Circuits and Systems.