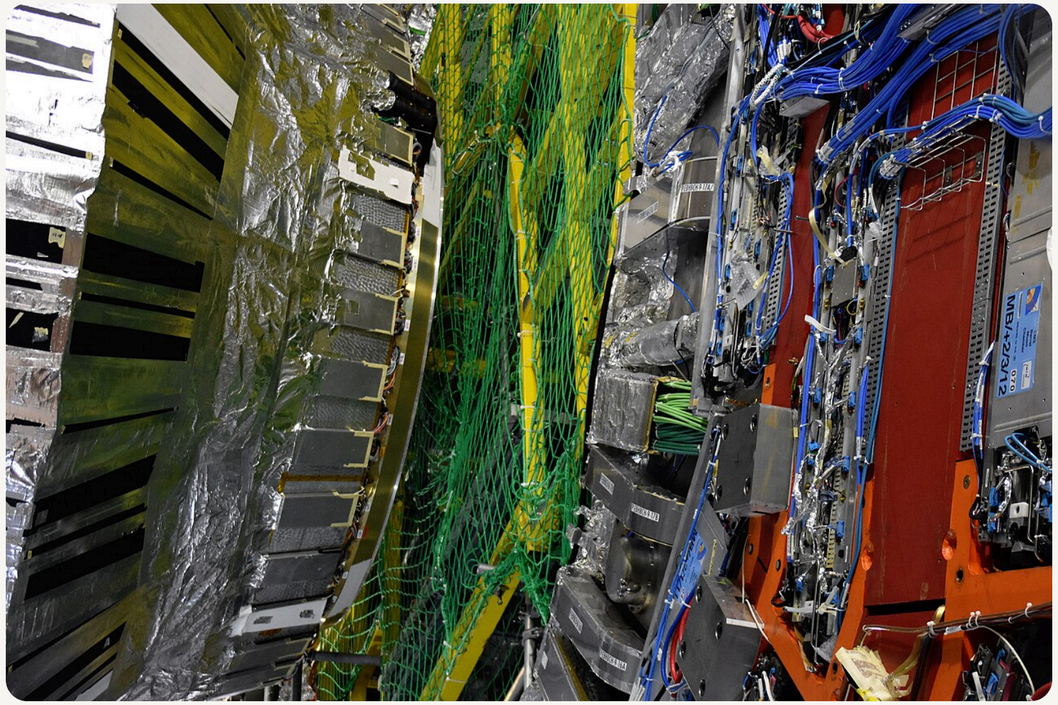
SIENA, NIPS-CERN, UFJF

Limiar

Redes neurais, do neurônio ao fractal

Aula 8: os dados do CERN, múons e massa invariante

04/12/2026. C∴ Chrysthofer A. A. Afonso



O CMS aberto, no LHC, em 2019. Foto de Simon Waldherr, Wikimedia Commons, CC BY 4.0.

1 · REVISÃO

O que ficou das aulas 6 e 7, de memória

1 O que o autograd faz?

2 Por que o PyTorch deu o mesmo resultado que o NumPy?

3 O que acontece com o erro final depois do limiar?

4 Para que servem treino, validação e teste?

1 · REVISÃO

O fio do curso, da aula 1 à 10

1	**O neurônio e o perceptron**	um neurônio desenha uma reta, e nenhuma reta separa o XOR	25/09, presencial
2	**Derivada e gradiente**	o passo contra a derivada; a taxa grande demais faz o treino divergir	02/10, presencial
3	**Matrizes e camadas**	duas camadas separam o XOR, com pesos escolhidos à mão	13/11, on-line
4	**O Jacobiano**	a regra da cadeia com vetores, uma camada de cada vez	13/11, on-line
5	**Retropropagação**	a regra da cadeia, de trás para a frente, acha os pesos do XOR	13/11, on-line
6	**O MNIST em NumPy**	um MLP em 70 mil dígitos escritos à mão, com treino, validação e teste	27/11, on-line
7	**PyTorch**	a biblioteca faz a retropropagação; o erro final contra a taxa	27/11, on-line
8	**Os dados do CERN**	um MLP estima a massa dos dois múons, conferida pela fórmula	04/12, on-line
9	**O fractal e o projeto**	a fronteira entre a taxa que converge e a que diverge	04/12, on-line
10	**A apresentação**	a apresentação do projeto e a revisão do curso inteiro	11/12, on-line

FAPEMIG CNPq Conselho Nacional de Desenvolvimento Científico e Tecnológico CAPES

O LHC e o CMS

27 km

a circunferência do LHC, até 175 m abaixo da fronteira entre a França e a Suíça

14 mil t

o CMS, um dos dois detectores de uso geral do LHC: 21 m de comprimento, 15 m de diâmetro

7 TeV

a energia das colisões de prótons de 2010, o ano dos 100 mil eventos desta aula

Acelerador, no curso: a máquina que dá energia às partículas e as faz colidir. Detector: o instrumento que registra o que sai da colisão.

O múon, um elétron 207 vezes mais pesado

105,66 MeV
a massa do múon

≈ 207 ×
a massa do elétron

−1
a carga, igual à do elétron; +1 no antimúon

eV, elétron-volt: a energia de um elétron numa diferença de potencial de 1 volt. MeV, um milhão de eV; GeV, um bilhão; TeV, um trilhão.

O arquivo: 100.000 eventos com dois múons

coluna	o que é
`Type`	T, tracker muon; G, global muon
`Run, Event`	a corrida e o número do evento
`E`	a energia, em GeV
`px, py, pz`	as componentes do momento, em GeV
`pt`	o momento transversal, perpendicular ao feixe
`eta`	a pseudorrapidez, que mede o ângulo com o feixe
`phi`	o ângulo em volta do feixe, em radianos
`Q`	a carga
`M`	a massa invariante dos dois múons, em GeV

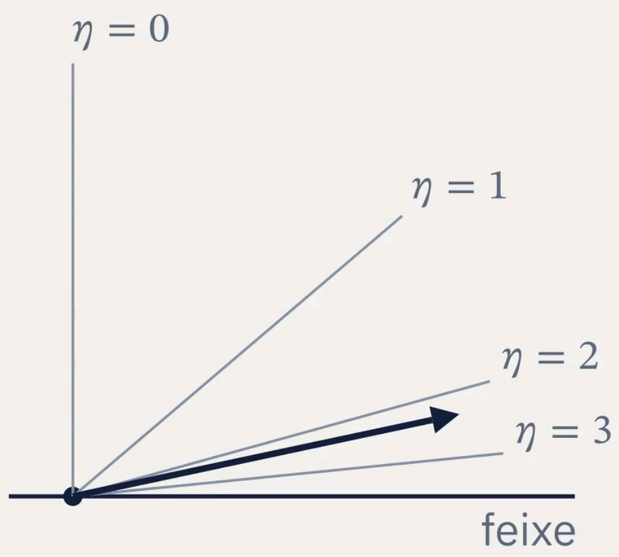
100.000

eventos, de 2 a 110 GeV

E, px, py, pz, pt, eta, phi e Q para cada múon, 1 e 2. Descrição do registro 303 do portal de dados abertos do CERN, CC0.

pt, eta e phi: a direção do múon

O tamanho p_T e os ângulos η e ϕ dão o momento do múon



de lado

$\theta = 12$ graus $\quad \eta = 2{,}25$



de frente

$\phi = 2{,}60$ rad

Da direção ao momento

$$p_x = p_T \cos\phi$$

$$p_y = p_T \operatorname{sen}\phi$$

$$p_z = p_T \operatorname{senh}\eta$$

senh, o seno hiperbólico: $(e^x - e^{-x})/2$

$$E = \sqrt{|\vec{p}\,|^2 + m^2}$$

O MÚON 1 DO EVENTO 3

$p_T = 4{,}091$ $\eta = 1{,}246$ $\phi = 1{,}632$

	pela conta	no arquivo
px	−0,2488	−0,24877
py	4,0834	4,08338
pz	6,5251	6,52511

Com a massa do múon, 0,10566 GeV, a fórmula da massa bate com a coluna M nos 100 mil eventos.

A massa invariante

$$M^2 = (E_1 + E_2)^2 - |\vec{p}_1 + \vec{p}_2|^2$$

|p|, o módulo: o tamanho do vetor momento

$$E^2 = |\vec{p}\,|^2 + m^2$$

Energia, momento e massa em GeV, com a velocidade da luz igual a 1.

Massa invariante: a massa da partícula que se desfez nos dois múons. Não depende de quem mede.

Um evento, à mão: o evento 3 do arquivo

	E	px	py	pz
múon 1	7,70222	−0,24877	4,08338	6,52511
múon 2	6,90202	−3,02439	3,75129	4,94041
soma	**14,60424**	**−3,27316**	**7,83467**	**11,46552**

Energia e momento em GeV. O momento se soma componente a componente.

$$(E_1 + E_2)^2 = 213{,}284$$

$$|\vec{p}_1 + \vec{p}_2|^2 = 203{,}554$$

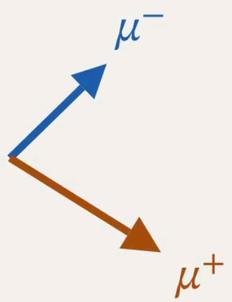
$$M^2 = 213{,}284 - 203{,}554 = 9{,}730$$

$$M = 3{,}119 \text{ GeV}$$

A coluna M diz 3,11929. A massa sai da diferença de dois números grandes.

A massa não depende de quem mede

Todo J/ψ dá a mesma massa: por isso o pico do histograma



$$E_1 = 2{,}18 \text{ GeV}$$

$$E_2 = 2{,}88 \text{ GeV}$$

$$|\vec{p}_1 + \vec{p}_2| = 4{,}00 \text{ GeV}$$

$$\boxed{M = 3{,}097 \text{ GeV}}$$

$$M^2 = (E_1 + E_2)^2 - |\vec{p}_1 + \vec{p}_2|^2$$

3 · OS DADOS E A MASSA INVARIANTE

A fórmula confere com o arquivo, quase sempre

99,6%

dos eventos: a fórmula, com E e p do arquivo, bate com a coluna M

413

eventos em que a energia de um múon não combina com o momento

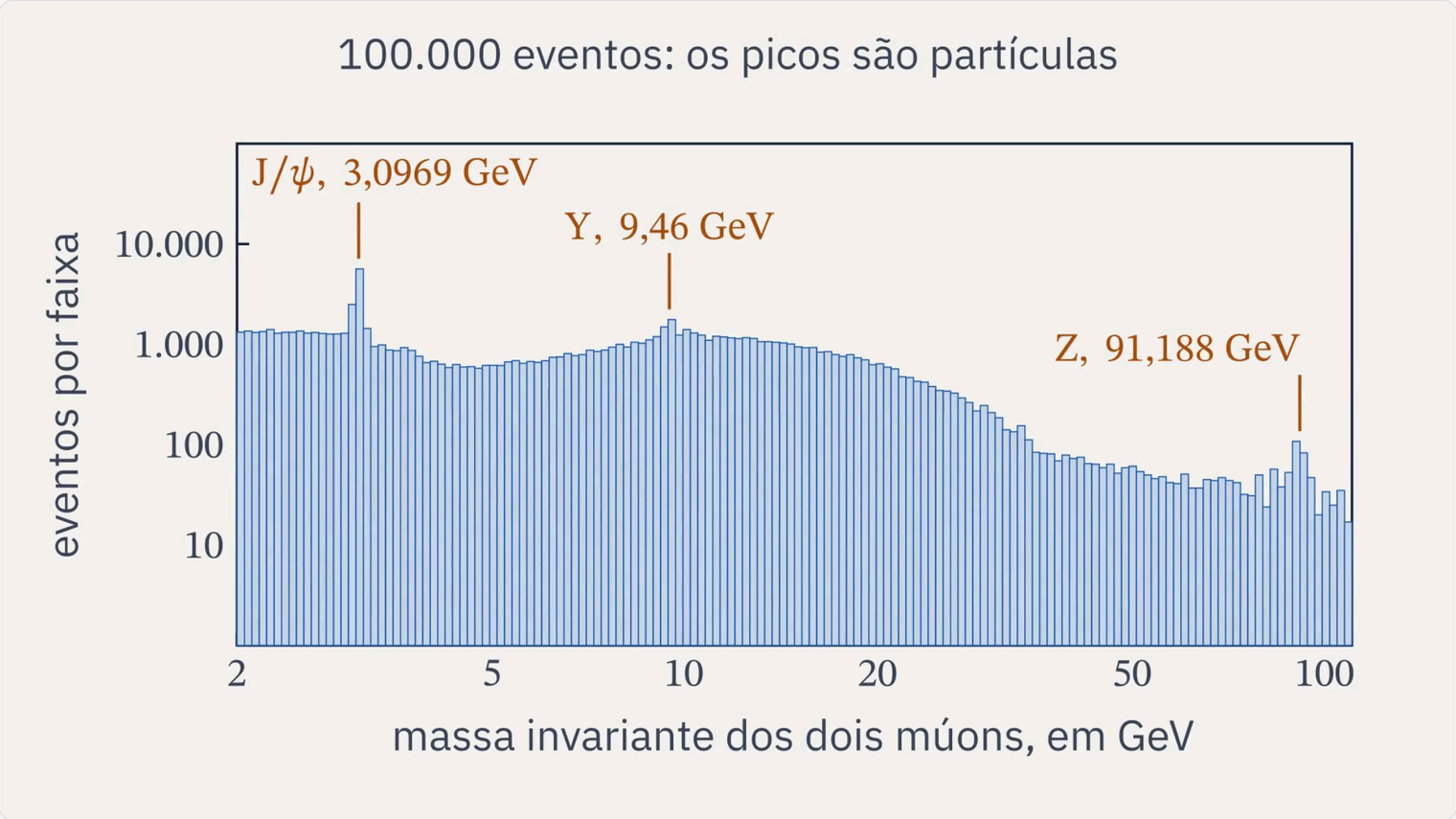
100%

com pt, eta e phi, a fórmula bate em todos

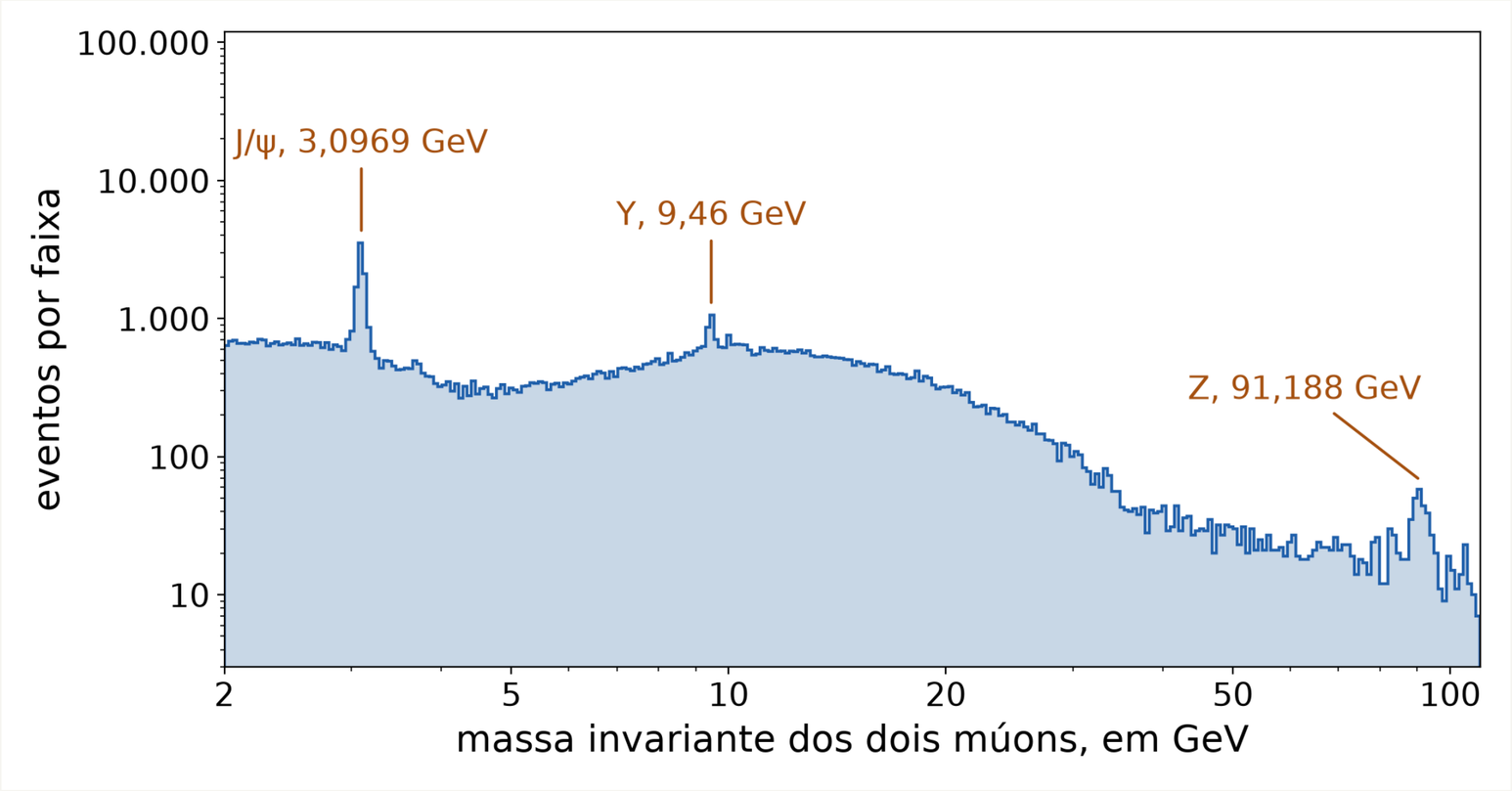
Num dos 413, um múon tem 184 mil GeV: 26 vezes os 7 mil GeV da colisão inteira.

Bater, aqui: diferença relativa menor que 0,1%. Dado real se confere antes de usar.

Do evento ao histograma

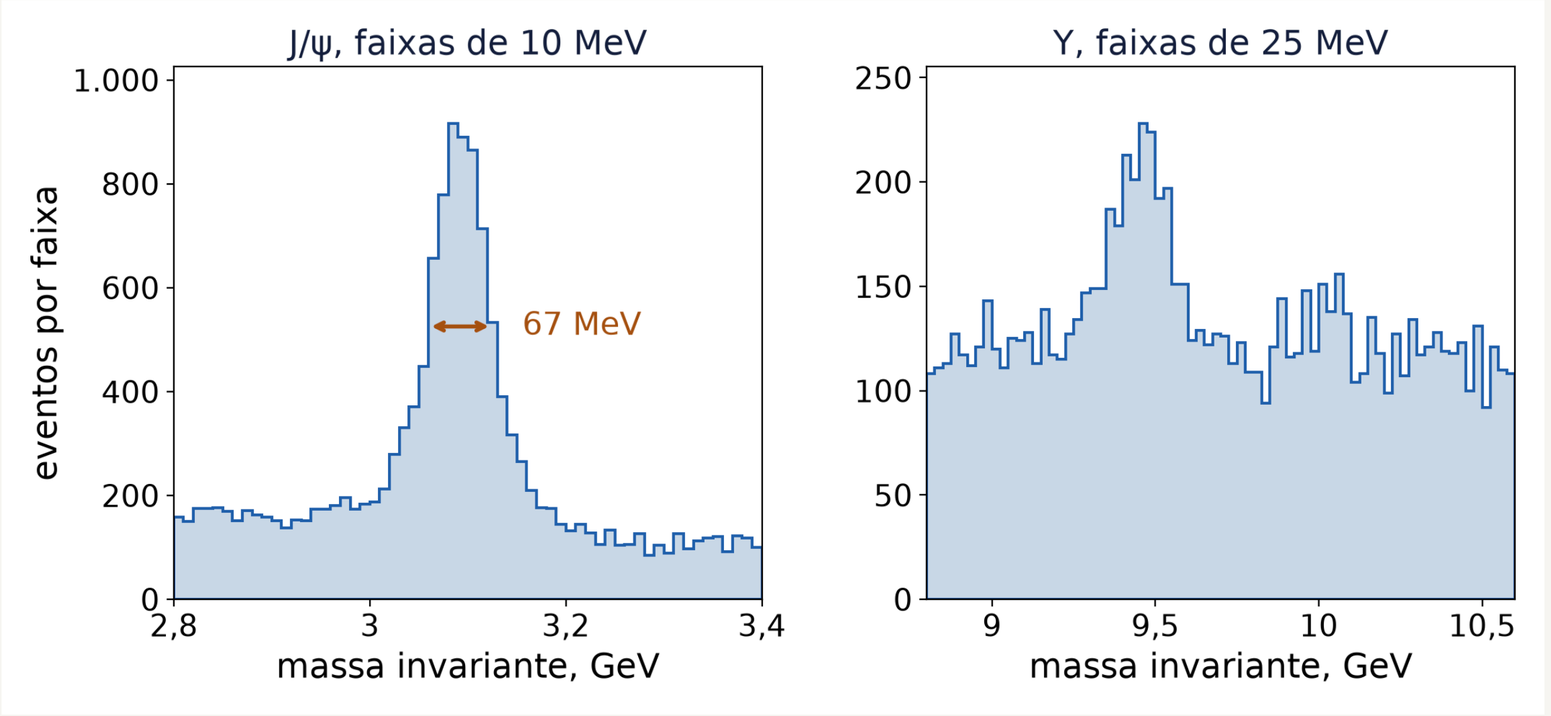


Os picos são partículas



Sobre o fundo liso, três picos: o J/ψ, o Y e o Z, cada um uma partícula que se desfez em dois múons.

Os picos de perto



O pico do J/ψ tem 67 MeV à meia altura. O J/ψ em si tem 93 keV: o resto é o erro de medida do detector.

4 · DE ONDE VEM

O que é cada pico

J/ψ

3,0969 GeV

- um quark charm e um antiquark charm

vive cerca de

$7{,}2 \times 10^{-21}$ s

Υ

9,46 GeV

- um quark bottom e o seu antiquark

vive cerca de

$1{,}21 \times 10^{-20}$ s

Z

91,188 GeV

- carrega a força fraca
- sem carga elétrica

vive cerca de

3×10^{-25} s

Méson: um quark e um antiquark. Nenhuma das três chega ao detector: o CMS vê os múons, e a massa invariante mostra a partícula.

4 · DE ONDE VEM

Três picos, três descobertas

1974	O J/ψ, descoberto ao mesmo tempo no SLAC, por Burton Richter, e em Brookhaven, por Samuel Ting.
1977	O Y, pelo experimento E288 de Leon Lederman, no Fermilab, em pares de múons.
1983	O Z, no CERN, pelos experimentos UA1 e UA2, no colisor SPS.

Fontes: Wikipedia, verbetes “J/psi meson”, “Upsilon meson” e “W and Z bosons”.

Uma rede aprende a fórmula?

Oito números entram, a energia e o momento de cada múon. Sai a massa. A rede não conhece a fórmula.

5 · UMA REDE APRENDE A FÓRMULA?

A rede: oito entradas, uma saída

$$8 \xrightarrow{\tanh} 256 \xrightarrow{\tanh} 256 \longrightarrow \ln M$$

$$x' = \operatorname{asinh}(x/5)$$

asinh, o arco seno hiperbólico: comprime os números grandes como um logaritmo, sem perder o sinal

O TREINO

- 80 mil eventos no treino, 10 mil na validação, 10 mil no teste
- o Adam, taxa 0,001, lotes de 256, 100 épocas

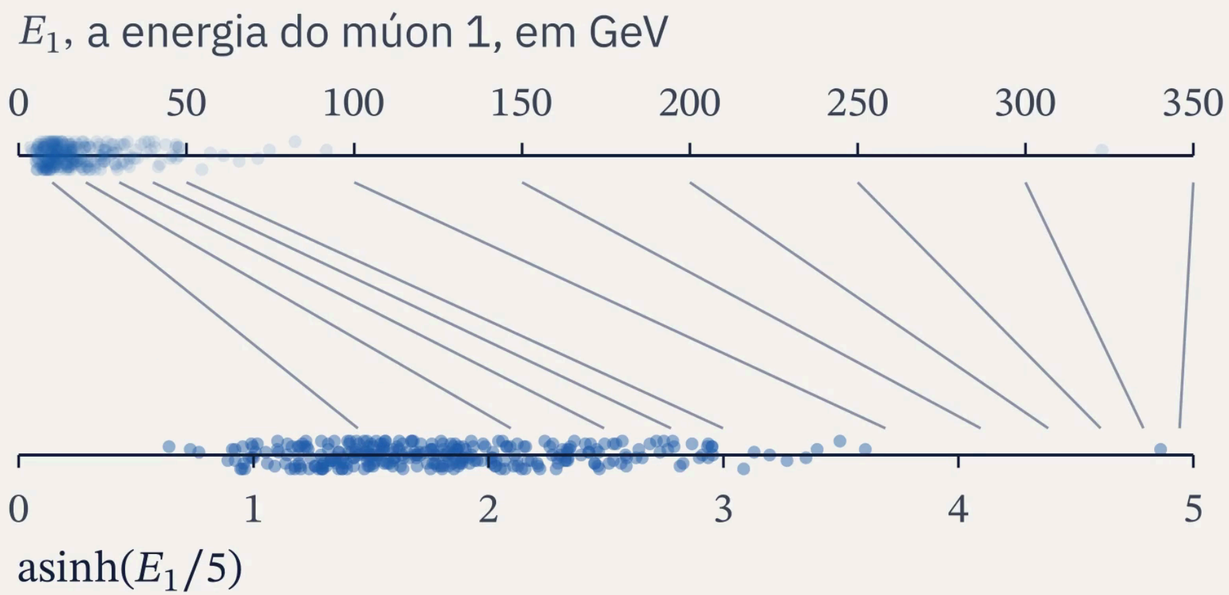
A COMPRESSÃO, EM 30 ÉPOCAS

10,6% → 3,4%

o erro mediano na validação, sem e com o asinh

A compressão das entradas

Os primeiros 50 GeV ganham 60% da linha; os outros 300, o resto



5 · UMA REDE APRENDE A FÓRMULA?

Padronizar, e o logaritmo da massa

$$\ln M_{\text{rede}} - \ln M \approx \frac{M_{\text{rede}} - M}{M}$$

≈, aproximadamente igual

$$x'' = \frac{x' - \mu}{\sigma}$$

μ e σ, a média e o desvio padrão, só do treino

$$\ln 3{,}16 - \ln 3{,}10 = 0{,}0192$$

$$(3{,}16 - 3{,}10) \,/\, 3{,}10 = 0{,}0194$$

PADRONIZAR

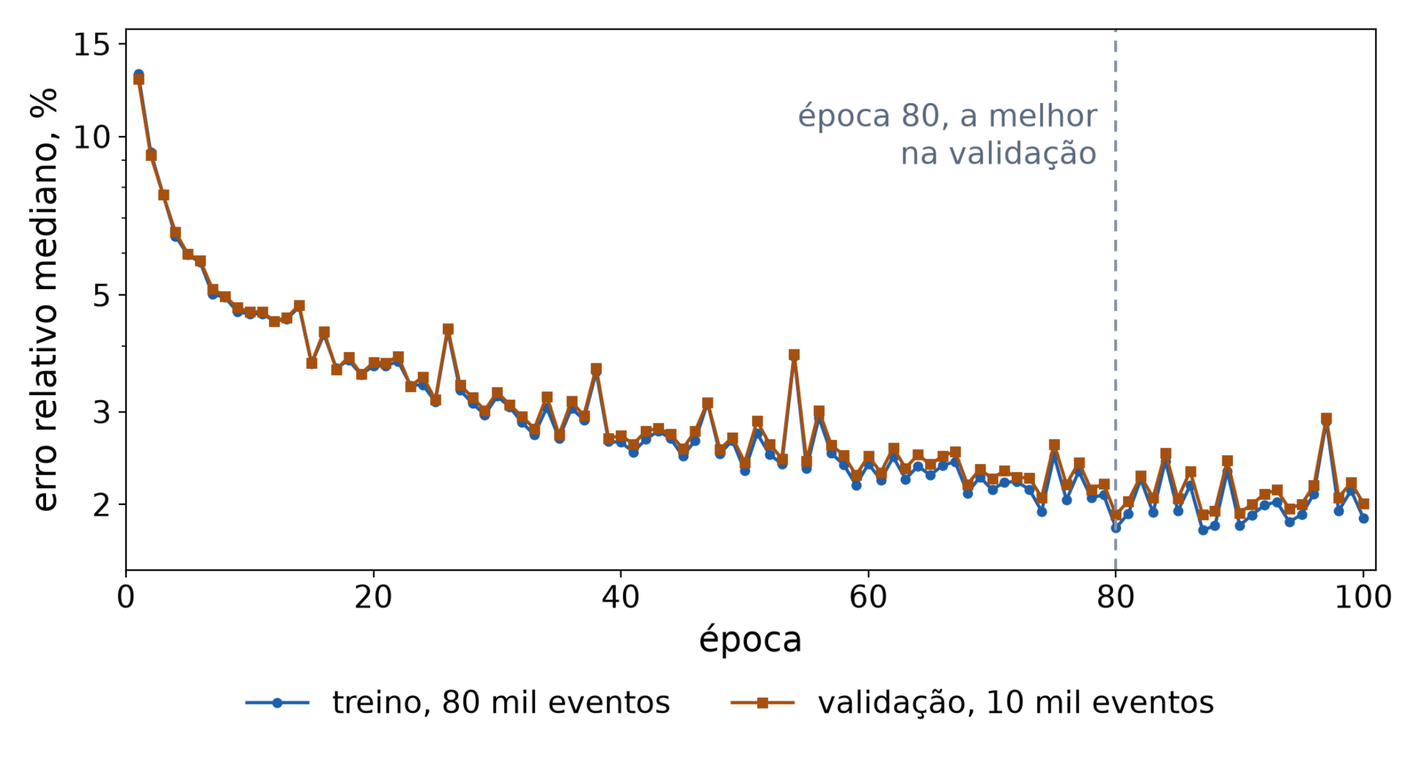
- cada entrada com média 0 e desvio 1
- a média e o desvio vêm só do treino
- o teste não ensina nada à rede

POR QUE LN M

- M vai de 2 a 110; ln M, de 0,69 a 4,70
- o erro quadrático em ln M cobra o erro em porcentagem

FAPEMIG CNPq Conselho Nacional de Desenvolvimento Científico e Tecnológico CAPES

A curva do treino

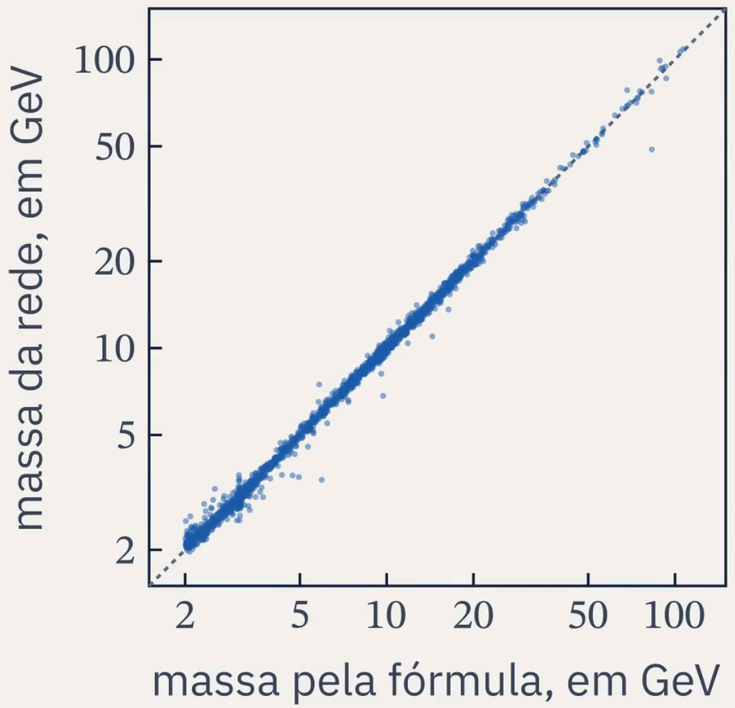


Treino e validação descem juntos: a rede não decora. A melhor época pela validação é a 80, com 1,91%.

5 · UMA REDE APRENDE A FÓRMULA?

O treino, época a época

Época 80, a melhor pela validação: os picos viram manchas



rede = fórmula

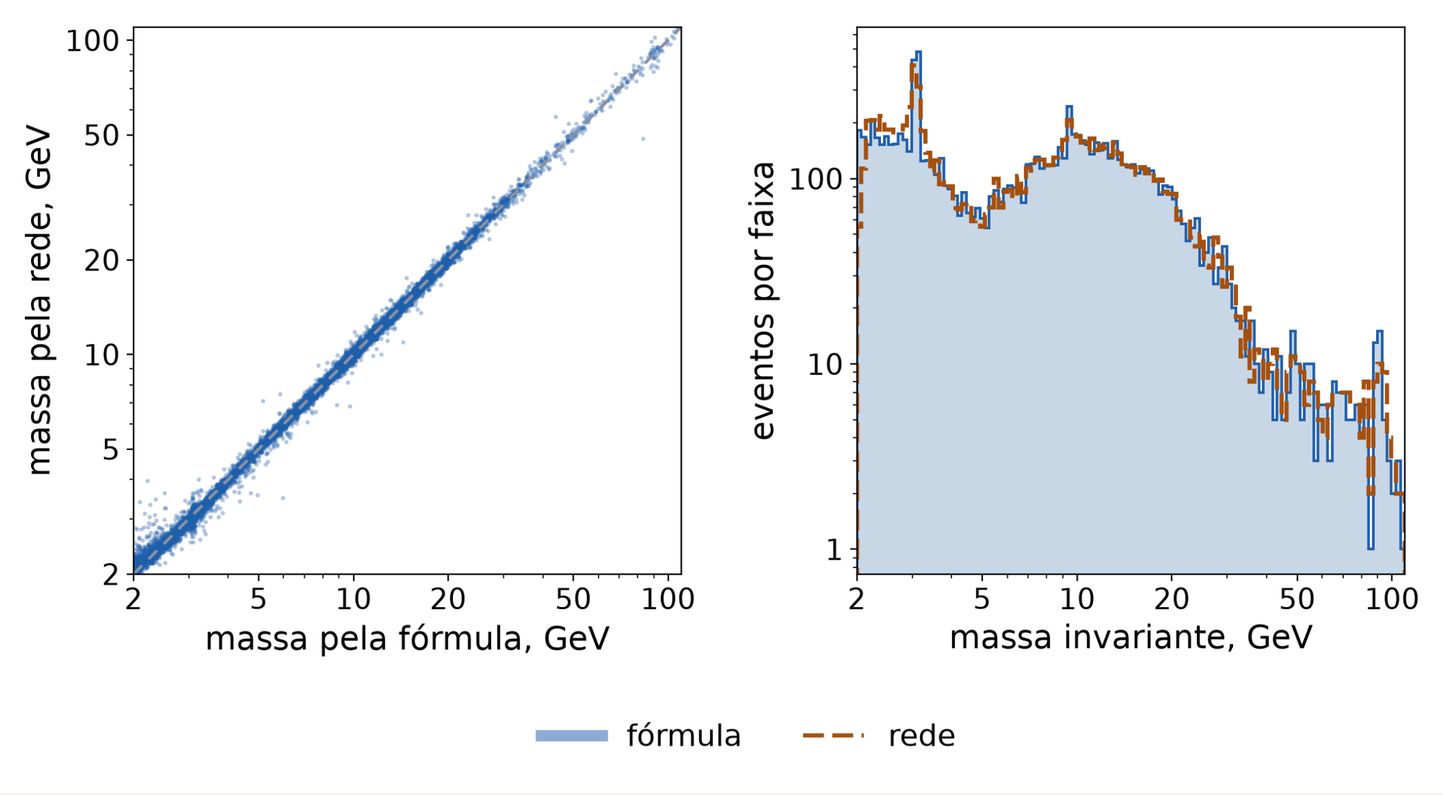
época

80

erro mediano

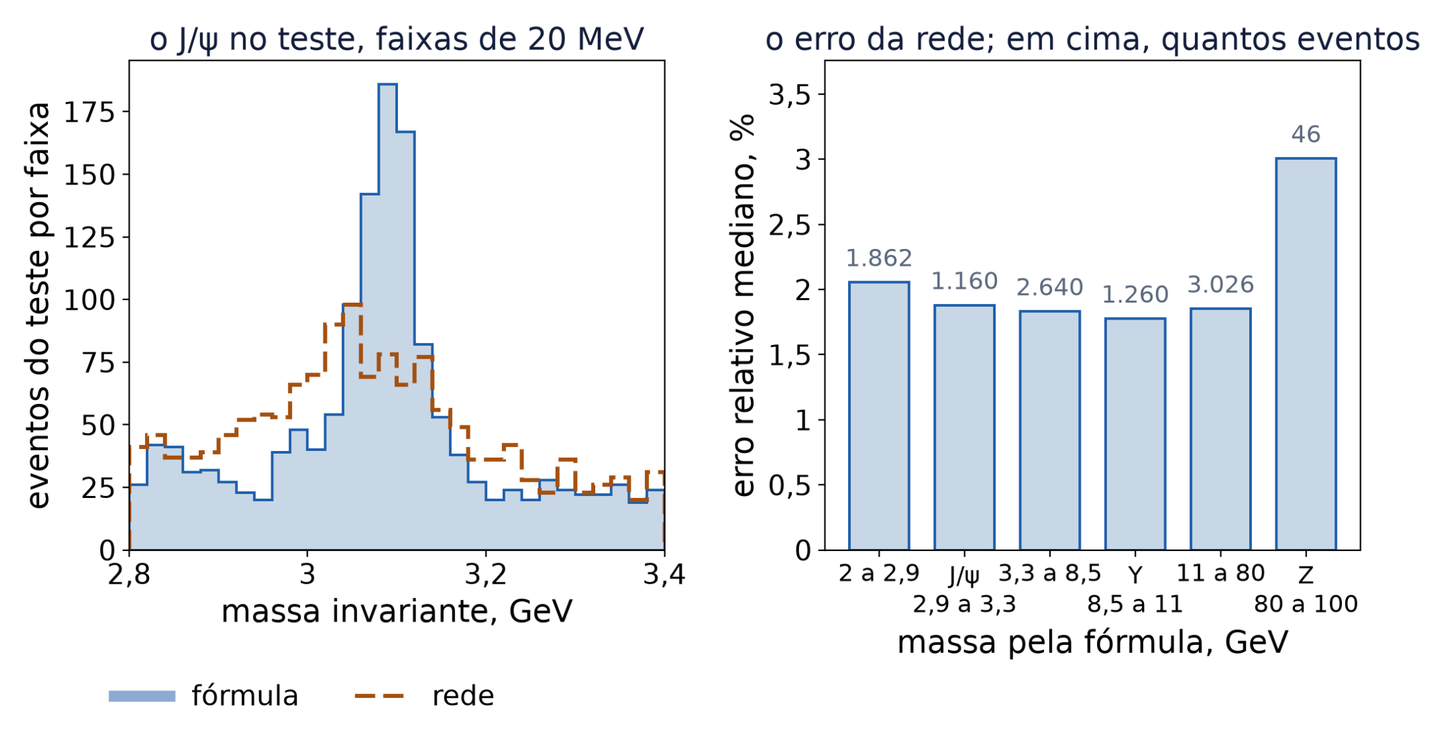
1,9%

A rede contra a fórmula



No teste: erro relativo mediano de 1,88%, e 90% dos eventos abaixo de 5,26%. A fórmula erra em milionésimos.

Onde a rede erra



A rede borra o pico do J/ψ e erra mais no Z, onde há poucos exemplos: 364 no treino.

6 · O CÓDIGO

Os dados do CMS, passo a passo

1. dimuon.py lê o arquivo e calcula a fórmula pelos dois caminhos
2. O histograma, com faixas iguais em escala logarítmica
3. mlp_massa.py comprime as entradas e treina a rede
4. A melhor época pela validação, e o teste uma vez
5. Os números em resultados/massa.json, para as figuras

O NUMPY QUE RESOLVE

```
np.sqrt((E1 + E2)**2
    - (px1 + px2)**2
    - (py1 + py2)**2
    - (pz1 + pz2)**2)
np.histogram(M, bins=bordas)
np.arcsinh(X / 5)
```

6 · O CÓDIGO

O treino em PyTorch

MLP_MASSA.PY, O LAÇO DO TREINO

```
modelo = rede()
opt = torch.optim.Adam(modelo.parameters(), lr=1e-3)
perda_fn = torch.nn.MSELoss()
for epoca in range(1, 101):
    ordem = tre[rng.permutation(len(tre))]
    for k in range(0, len(ordem), 256):
        i = torch.from_numpy(ordem[k:k + 256])
        opt.zero_grad()
        perda = perda_fn(modelo(Xp[i]), yp[i])
        perda.backward()
        opt.step()
```

1. zero_grad() zera as derivadas do lote anterior
2. A passada para a frente e a perda, o erro quadrático em ln M
3. backward(): a retropropagação da aula 5, sozinha
4. step(): o passo do Adam em cada peso

6 · O CÓDIGO

O que a entrega pede

O HISTOGRAMA

- a massa dos 100 mil eventos
- os picos com o nome e a massa de cada partícula

A REDE

- um teste separado antes do treino
- o erro relativo mediano contra a fórmula
- uma figura que compare as duas

ENTREGA

- o histograma e o erro da rede
- uma pergunta sobre o estudo

Extra, opcional: outras entradas, como pt, eta e phi, ou um erro menor no Z.

7 · FECHAMENTO

Para estudar e entregar

Estudo

- 3Blue1Brown, “Fractals are typically not self-similar”, 2017
- Sohl-Dickstein, “Neural network training makes beautiful fractals”, 2024

Entrega

- o histograma com os picos identificados
- o erro da rede contra a fórmula
- uma pergunta sobre o estudo

A seguir

- aula 9: o fractal da taxa e o projeto
- aula 10: a apresentação e a revisão