

Métodos machine learning aplicados em estrelas de nêutrons

André G. da Silva

Orientador: Ricardo L. S. Farias

Departamento de Física
Universidade Federal de Santa Maria

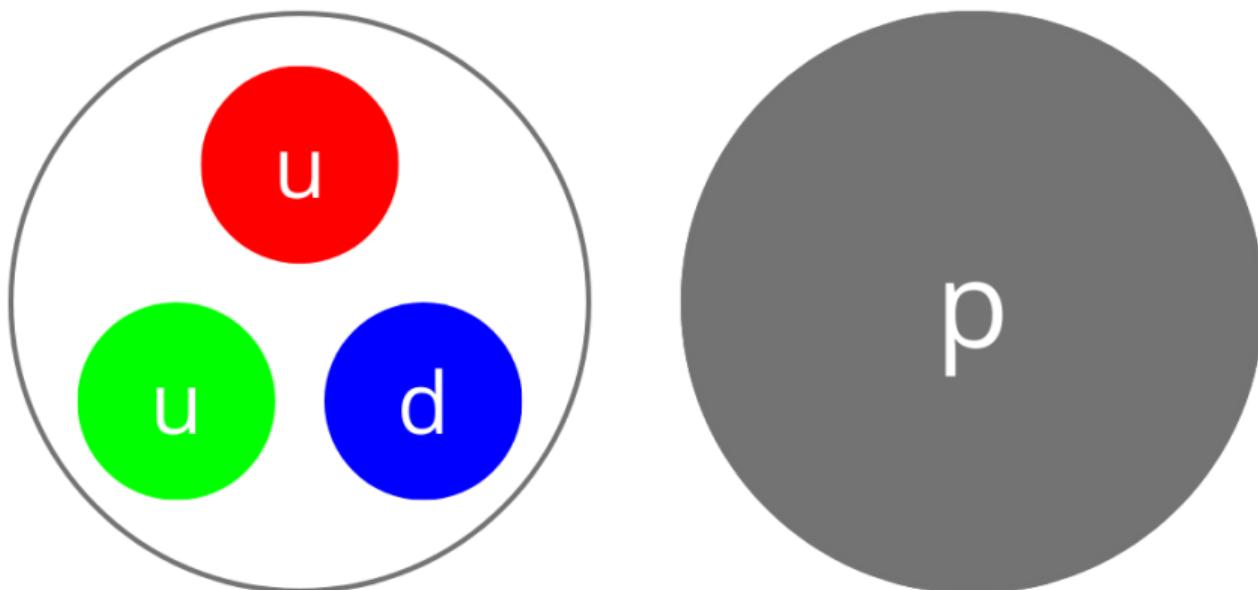
25 de setembro de 2025

Conteúdo

1. Introdução
2. Objetivos
3. Estrelas de nêutrons
4. Machine learning
5. Resultados
6. Conclusão e perspectivas

Introdução

- A cromodinâmica quântica (QCD) como a teoria da interação forte;
- Liberdade assintótica e simetria quiral;



- Regimes extremos;
- Diagrama de fases;
- Equação de estado da matéria densa;

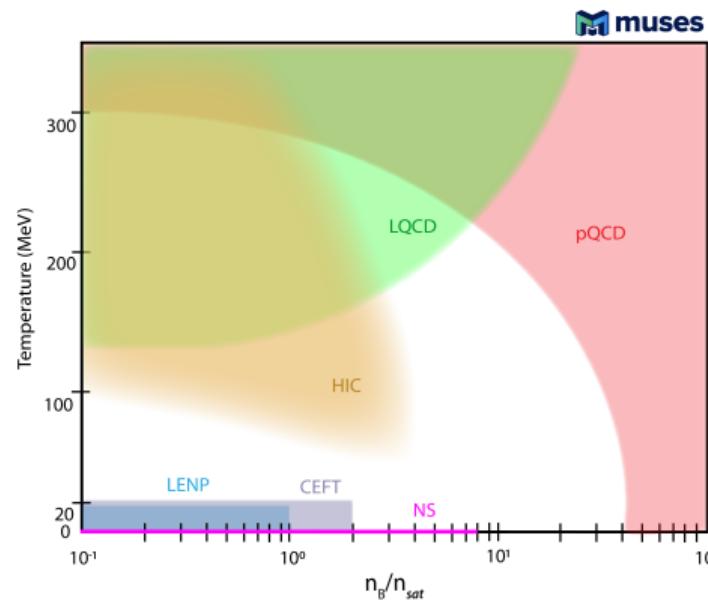


Figura: Regiões acessíveis por meio de experimentos e simulações. Imagem de R. Kumar et al., *Living Rev. Rel.*, 2024, 27, 3.

Por que obter a equação de estado de estrelas de nêutrons?



Por que obter a equação de estado de estrelas de nêutrons?



- Exclusão de modelos que não descrevem a equação de estado;
- Ajuste de parâmetros de modelos;
- Fornecimento da EoS para simulações como, por exemplo, mergers de estrelas de nêutrons;

Quantificação de incertezas

- Incertezas com análises Bayesianas;

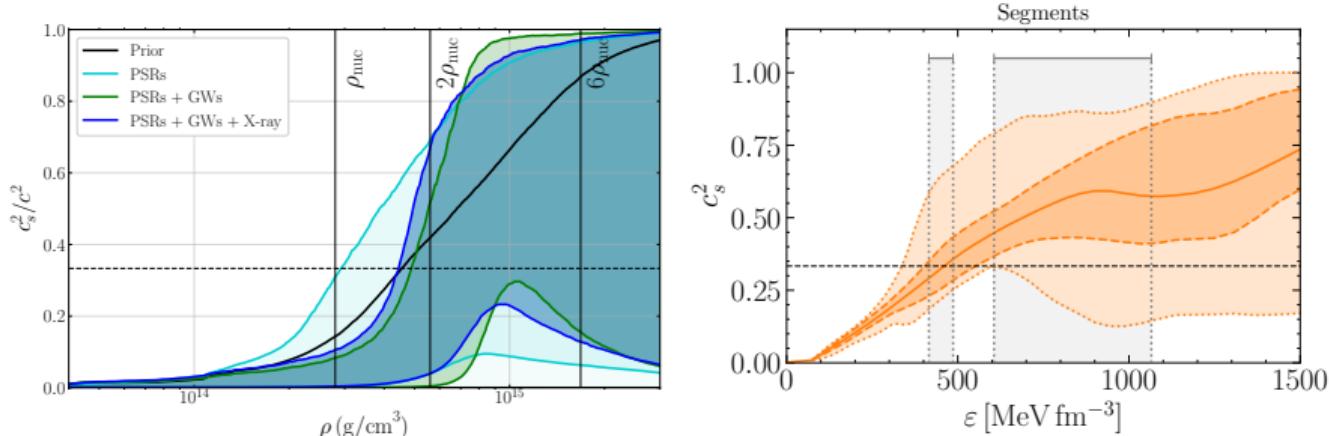


Figura: (Esquerda) resultados de P. Landry et al., *Phys. Rev. D*, 2020, **101**, 123007, CIs de 90%.
(Direita) resultados de L. Brandes et al., *Physical Review D*, 2023, **107**, DOI:
[10.1103/physrevd.107.014011](https://doi.org/10.1103/physrevd.107.014011), CIs de 68% (bandas escuras) e 95% (bandas claras).

■ Incertezas com redes neurais;

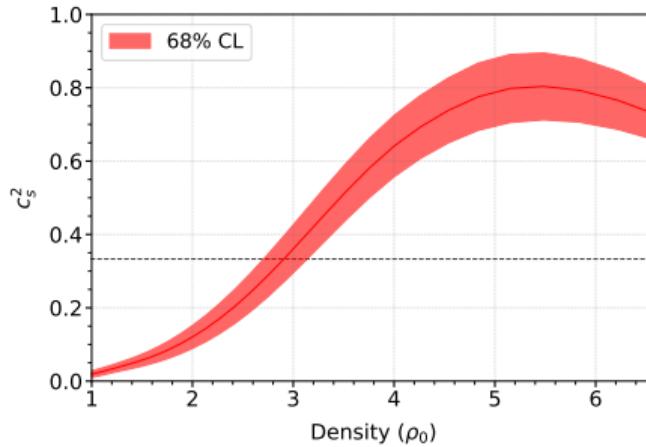
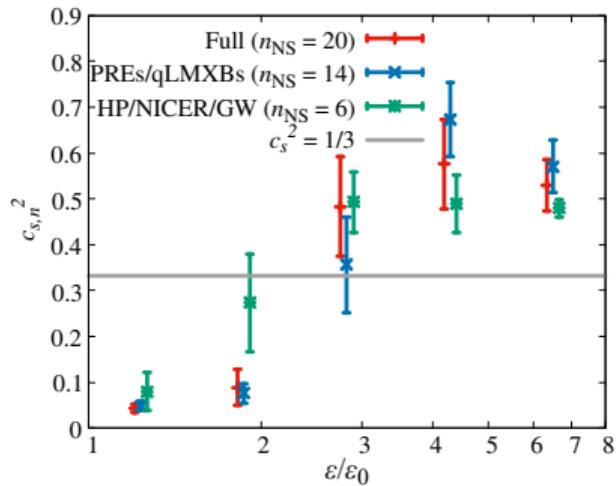


Figura: (Esquerda) resultados de Y. Fujimoto et al., *Phys. Rev. D*, 2024, **110**, 034035, utilizando médias e desvios padrão. (Direita) resultados de S. Soma et al., *Physical Review D*, 2023, **107**, DOI: [10.1103/physrevd.107.083028](https://doi.org/10.1103/physrevd.107.083028), CIs de 90%.

Objetivos

- Desenvolver um processo de inferência da equação de estado baseado em uma forma paramétrica;
- Desenvolver os modelos de inferência com base em rede neural e análise Bayesiana;
- Quantificar as incertezas relacionadas ao processo de inversão;
- Obtenção da equação de estado de matéria densa, juntamente com o diagrama massa-raio associado

Estrelas de Nêutrons

- Um dos objetos mais densos conhecidos;
- Rotações e campos magnéticos altos;
- Temperaturas baixas;



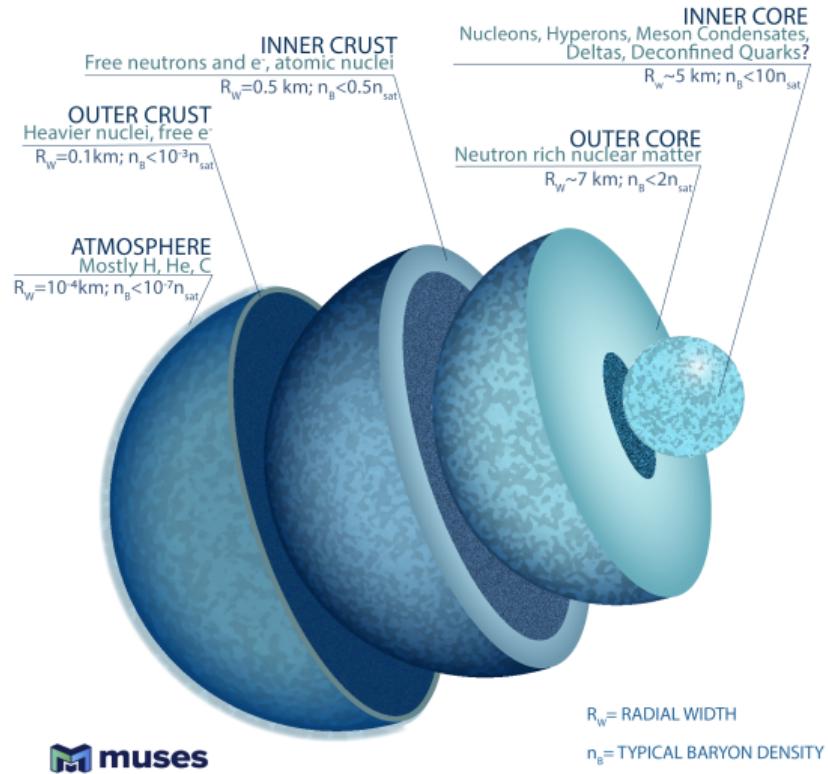


Figura: Seções de uma estrela de nêutrons. Imagem de R. Kumar et al., *Living Rev. Rel.*, 2024, **27**, 3.

- Dados mais recentes do NICER e LIGO/Virgo;

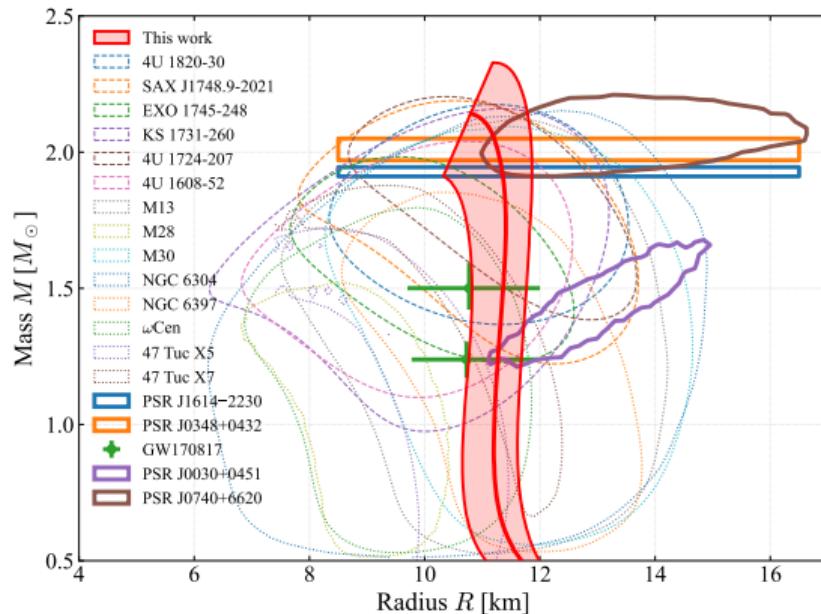


Figura: Compilação de medidas massa-raio medidas. Imagem de Y. Fujimoto et al., *Phys. Rev. D*, 2024, **110**, 034035.

Equilíbrio hidrostático

Observações de massa e raio podem ser ligadas à equação de estado (calculada por modelos de QCD) via as equações de Tolman-Oppenheimer-Volkoff¹:

$$\frac{dp}{dr} = -\frac{[p(r) + \epsilon(r)] [Gm(r) + 4\pi Gr^3 p(r)]}{r [r - 2Gm(r)]}, \quad (1)$$
$$\frac{dm}{dr} = 4\pi r^2 \epsilon(r),$$

onde resta uma incógnita $\epsilon(r)$ que pode ser colocada na forma $\epsilon(p)$ e é chamada de equação de estado (EoS). A EoS pode ser determinada por um modelo de física nuclear ou de forma paramétrica.

¹Aqui são utilizadas unidades em que $c = 1$

Na prática, as equações de TOV proporcionam soluções em que a pressão tende lentamente a zero. Para obter raios finitos, é necessário definir uma pressão de corte $p(r = R) \sim 1 \times 10^{-8}$ [u.p.], que pode depender fortemente da EoS utilizada. As equações de TOV podem ser colocadas na forma de pseudo-entalpia,

$$\begin{aligned}\frac{dr}{dh} &= -\frac{r(r-2m)}{m+4\pi r^3 p(h)}, \\ \frac{dm}{dh} &= -\frac{4\pi \epsilon(h) r^3 (r-2m)}{m+4\pi r^3 p(h)},\end{aligned}\tag{2}$$

onde a h é dado por

$$h(p) = \int_0^p \frac{dp'}{\epsilon(p') + p'}.\tag{3}$$

Essas equações foram apresentadas originalmente por^{2,3} e foram resolvidas para um gás de fermions degenerados. A equação de estado toma a forma⁴

$$\begin{aligned}\epsilon(x) &= \frac{\epsilon_n}{8} \left[(2x^3 + x)(1 + x^2)^{\frac{1}{2}} - \sinh^{-1}(x) \right], \\ p(x) &= \frac{\epsilon_n}{24} \left[(2x^3 - 3x)(1 + x^2)^{\frac{1}{2}} + 3 \sinh^{-1}(x) \right], \quad x = \frac{k_F}{m_n}\end{aligned}\tag{4}$$

²R. C. Tolman, *Phys. Rev.*, 1939, **55**, 364–373.

³J. R. Oppenheimer e G. M. Volkoff, *Phys. Rev.*, 1939, **55**, 374–381.

⁴I. Sagert et al., *European Journal of Physics*, 2006, **27**, 577–610.

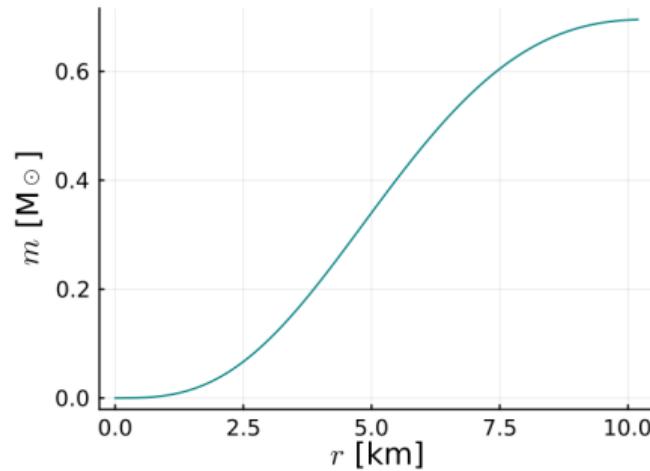
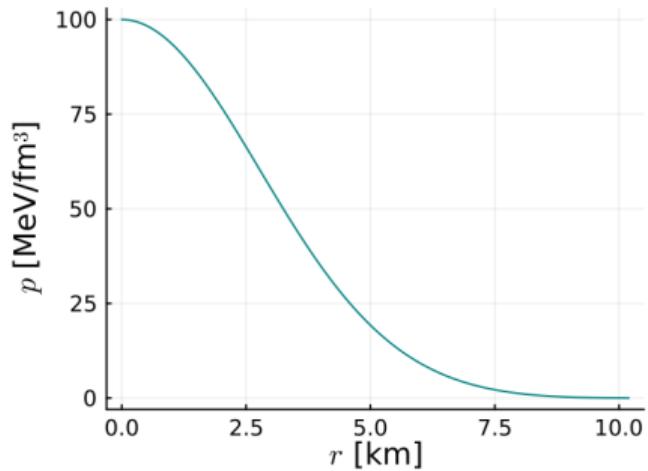


Figura: Solução das equações de TOV para $p(r = 0) = 100$ MeV fm $^{-3}$.

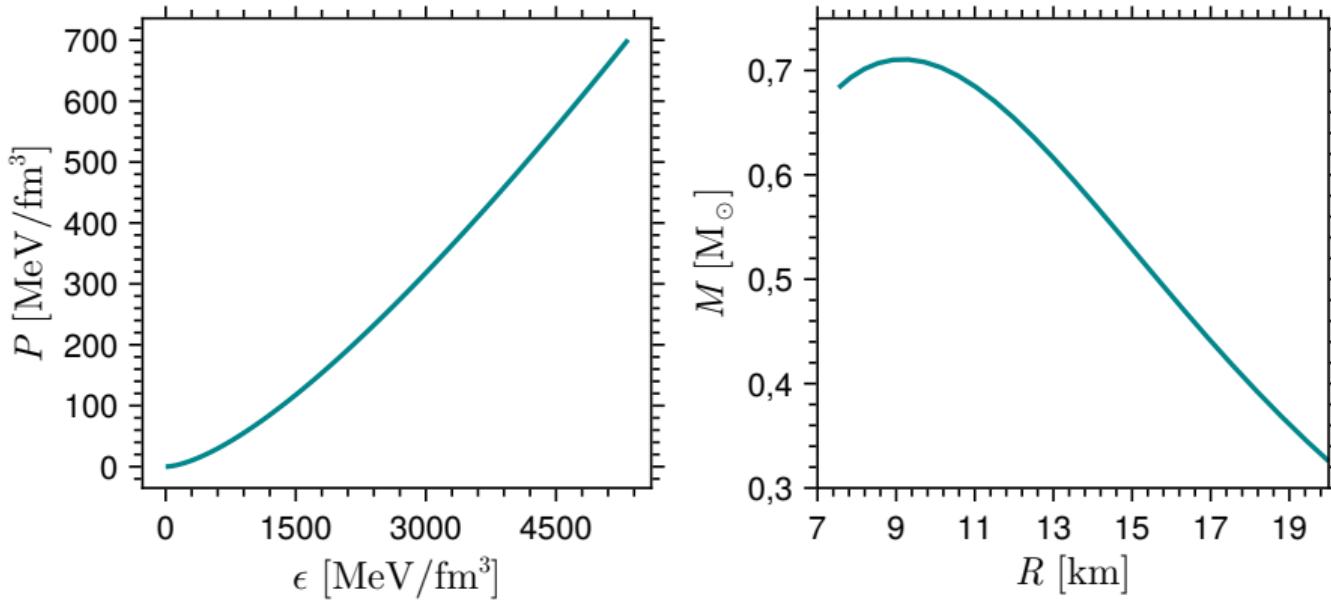


Figura: Diagrama massa-raio gerado pela equação de estado de gás de Fermi.

Parametrização da EoS

- Queremos encontrar uma forma da equação de estado que seja mais independente do modelo utilizado;

Uma forma paramétrica interessante é a que utiliza a velocidade do som definida por

$$c_s^2 = \left(\frac{\partial p}{\partial \epsilon} \right)_S. \quad (5)$$

Dividindo o intervalo de densidade de energia em 5 segmentos uniformemente espaçados no logaritmo no intervalo $(\epsilon_0, 8\epsilon_0]$ ($\epsilon_0 = 150 \text{ MeV fm}^{-3}$) podemos obter as pressões e densidades de energia nas interfaces desses segmentos a partir da velocidade do som média no intervalo⁵.

$$p_i = p_{i-1} + c_{s,i}^2 (\epsilon_i - \epsilon_{i-1}). \quad (6)$$

⁵Y. Fujimoto et al., *JHEP*, 2021, **03**, 273.

Utilizamos uma interpolação politrópica em que, no intervalo i a pressão é dada por parâmetros K_i e Γ_i

$$p = K_i \epsilon^{\Gamma_i}. \quad (7)$$

Resumindo:

- Podemos obter o diagrama massa-raio a partir da equação de estado utilizando as equações de TOV;
- A EoS pode ser aproximada por uma forma paramétrica. Aqui utiliza-se a velocidade do som média em intervalos de densidade de energia;
- Essa forma paramétrica depende de uma equação de estado inicial (crosta), aqui é utilizada a SLy4⁶;

⁶F. Douchin e P. Haensel, *Astronomy & Astrophysics*, 2001, **380**, 151–167.

■ Aprendizado de máquina;

Tendo uma forma paramétrica da equação de estado, podemos nos voltar ao problema de encontrar os parâmetros c_s^2 que melhor ajustam os dados observacionais. Focaremos em duas abordagens:

- Análise Bayesiana;
- Redes neurais;

Análise Bayesiana

Utilizamos o teorema de Bayes

$$P(\mathbf{c}_s^2, \mathbf{p}_c | \mathbf{D}, I) \propto P(\mathbf{D} | \mathbf{c}_s^2, \mathbf{p}_c, I) P(\mathbf{c}_s^2, \mathbf{p}_c | I), \quad (8)$$

em que as pressões centrais das estrelas de cada dado observacional foram adicionadas para a definição da *likelihood*. Ainda, assumimos independência dos dados e *priors*

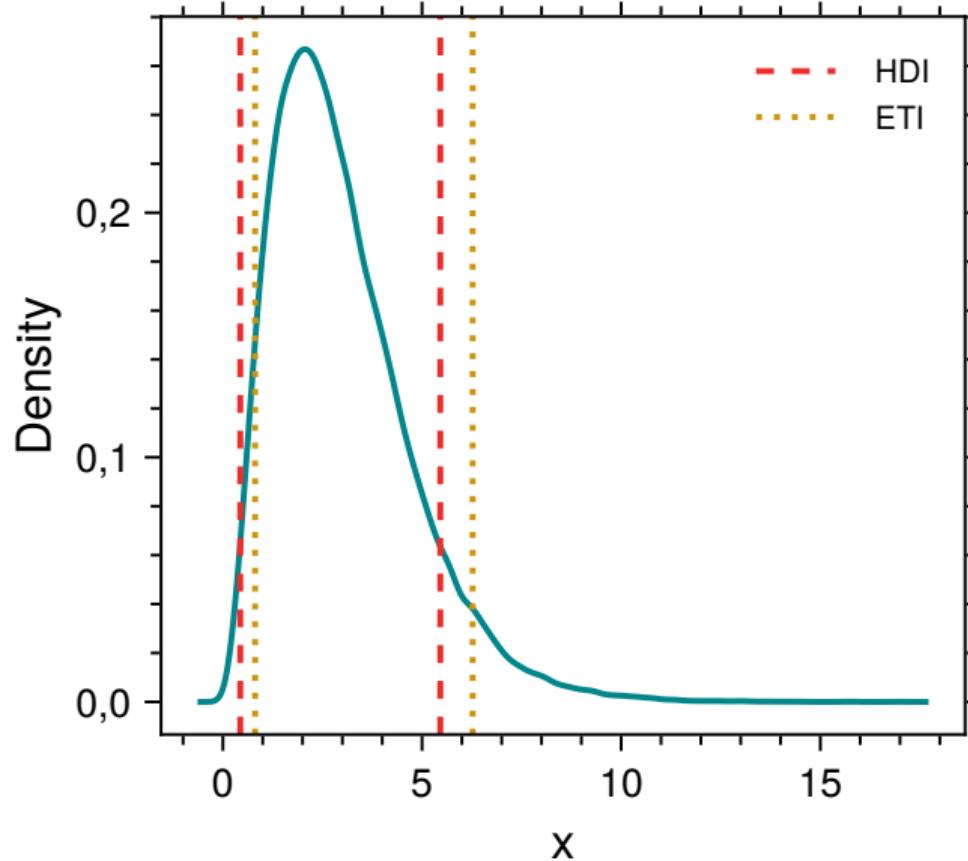
$$\begin{aligned} P(\mathbf{D} | \mathbf{c}_s^2, \mathbf{p}_c, I) &= \prod_{k=1}^N P(D_k | \mathbf{c}_s^2, p_{c,k}, I), \\ P(\mathbf{c}_s^2, \mathbf{p}_c | I) &= \left[\prod_{i=1}^n P(c_{s,i}^2 | I) \right] \left[\prod_{k=1}^N P(p_{c,k} | I) \right]. \end{aligned} \quad (9)$$

Por fim, utilizamos *likelihoods* gaussianas

$$P(\mathbf{D}|\mathbf{c}_s^2, \mathbf{p}_c, I) \propto \exp\left(-\frac{\chi^2}{2}\right),$$
$$\chi^2 = \sum_{k=1}^N \frac{[M(\mathbf{c}_s^2, p_{c,k}) - M_k]^2}{\sigma_{M,k}^2} + \sum_{k=1}^N \frac{[R(\mathbf{c}_s^2, p_{c,k}) - R_k]^2}{\sigma_{R,k}^2}. \quad (10)$$

Os dados observacionais são ajustados por distribuições gaussianas com parâmetros $\{R_k, M_k, \sigma_{R,k}, \sigma_{M,k}\}$ e então é utilizado um *sampler* que retirará amostras da distribuição posterior $P(\mathbf{c}_s^2, \mathbf{p}_c | \mathbf{D}, I)$.

- Intervalos de credibilidade e convenções desse trabalho;



Faz parte da análise

- Parametrização utilizada;
- Escolha dos *likelihoods*;
- *Priors*;
- Algoritmo de amostragem, *sampler*;

Redes neurais

- Teorema de aproximação universal;
- Redes neurais profundas (DNNs);

O *output* de uma unidade de uma DNN é dado por

$$a_j^l = \sigma \left(\sum_k w_{jk}^l a_k^{l-1} + b_j^l \right) = \sigma(z_j^l). \quad (11)$$

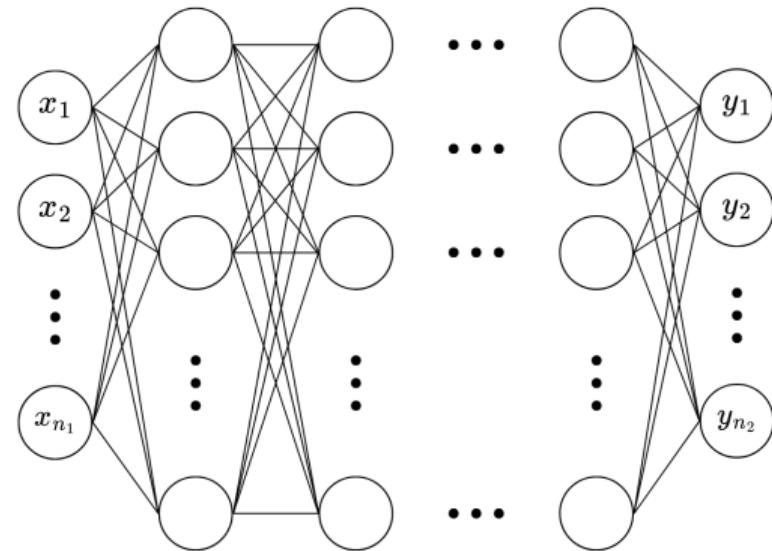


Figura: Representação de uma DNN.

- O teorema de aproximação universal não nos dá a arquitetura necessária (número de camadas, por exemplo) para aproximar uma função;
- Encontramos os parâmetros da rede utilizando um algoritmo de otimização de modo a minimizar uma função de erro;
- Para isso são necessários dados de treino que em alguns problemas podem ser gerados;

A prática geral é gerar (ou extrair) um conjunto de dados $\{\mathbf{x}_j, \mathbf{y}_k\}$ e separar em duas partes: treino ($\sim 80\%$) e validação ($\sim 20\%$). Então temos a atualização dos parâmetros

$$\theta_{i+1} = \theta_i - \eta \nabla_{\theta} E(\theta). \quad (12)$$

Alternativamente, isso é feito em *minibatches*

$$\theta_{i+1,k} = \theta_{i,k} - \eta \nabla_{\theta} E_k(\theta), \quad k = 1, \dots, n/M. \quad (13)$$

Todos os seguintes são ajustáveis:

- Número de camadas e unidades por camada;
- Funções de ativação;
- Tamanho do *minibatch* M ;
- Taxa de aprendizado η ;
- Função de erro $E(\theta)$;
- Algoritmo de otimização;
- Número de épocas;

Resultados da análise Bayesiana

Conclusão e perspectivas

Agradecimentos



Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul



cnpq-eps-converted-to.pdf



Referências I

- (1) R. Kumar et al., *Living Rev. Rel.*, 2024, **27**, 3.
- (2) P. Landry, R. Essick e K. Chatzioannou, *Phys. Rev. D*, 2020, **101**, 123007.
- (3) L. Brandes, W. Weise e N. Kaiser, *Physical Review D*, 2023, **107**, DOI: 10.1103/physrevd.107.014011.
- (4) Y. Fujimoto, K. Fukushima, S. Kamata e K. Murase, *Phys. Rev. D*, 2024, **110**, 034035.
- (5) S. Soma, L. Wang, S. Shi, H. Stöcker e K. Zhou, *Physical Review D*, 2023, **107**, DOI: 10.1103/physrevd.107.083028.
- (6) R. C. Tolman, *Phys. Rev.*, 1939, **55**, 364–373.
- (7) J. R. Oppenheimer e G. M. Volkoff, *Phys. Rev.*, 1939, **55**, 374–381.
- (8) I. Sagert, M. Hempel, C. Greiner e J. Schaffner-Bielich, *European Journal of Physics*, 2006, **27**, 577–610.

Referências II

- (9) Y. Fujimoto, K. Fukushima e K. Murase, *JHEP*, 2021, **03**, 273.
- (10) F. Douchin e P. Haensel, *Astronomy & Astrophysics*, 2001, **380**, 151–167.
- (11) Y. Fujimoto, K. Fukushima e K. Murase, *Phys. Rev. D*, 2018, **98**, 023019.
- (12) D. P. Kingma e J. Ba, *Adam: A Method for Stochastic Optimization*, 2017.
- (13) M. Oertel, M. Hempel, T. Klähn e S. Typel, *Rev. Mod. Phys.*, 2017, **89**, 015007.
- (14) M. G. Alford, L. Brodie, A. Haber e I. Tews, *Phys. Rev. C*, 2022, **106**, 055804.
- (15) G. Baym, T. Hatsuda, T. Kojo, P. D. Powell, Y. Song e T. Takatsuka, *Reports on Progress in Physics*, 2018, **81**, 056902.
- (16) A. Clevinger, J. Corkish, K. Aryal e V. Dexheimer, *The European Physical Journal A*, 2022, **58**, DOI: 10.1140/epja/s10050-022-00745-3.
- (17) K. Otto, M. Oertel e B.-J. Schaefer, *Eur. Phys. J. ST*, 2020, **229**, 3629–3649.

Referências III

- (18) N. K. Glendenning, *Compact stars: Nuclear physics, particle physics, and general relativity*, Springer, 1997.
- (19) P. de Forcrand, *PoS*, 2009, **LAT2009**, ed. C. Liu e Y. Zhu, 010.
- (20) C. Gattringer e C. B. Lang, *Quantum chromodynamics on the lattice*, Springer, Berlin, 2010, vol. 788.

Massa máxima de estrelas de nêutrons

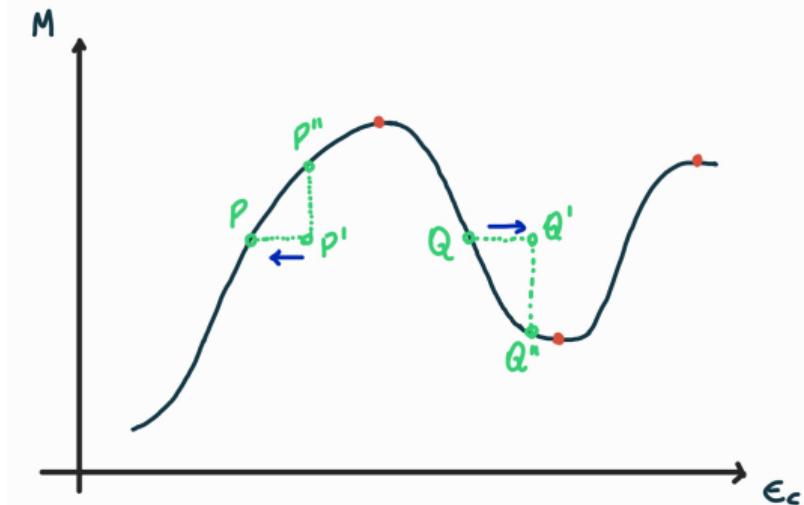


Figura: Ilustração de equilíbrio estável. Reproduzida de N. K. Glendenning, *Compact stars: Nuclear physics, particle physics, and general relativity*, Springer, 1997.

Markov Chain Monte Carlo



Distribuição desejada $P(x)$ e a distribuição estacionária (processo de Markov) $\pi(x)$. Do teorema de Bayes

$$P(x) = P(x|D) = \frac{P(D|x)P(x)}{P(D)} = Kf(x). \quad (14)$$

Exigimos um processo detalhado no equilíbrio

$$P(x'|x)P(x) = P(x|x')P(x'), \quad (15)$$

que temporariamente não vai ser satisfeito com

$$P(x'|x)P(x) > P(x|x')P(x'), \quad (16)$$

estão acontecendo mais transições $x \rightarrow x'$.

Podemos separar a probabilidade de transição em duas partes

$$g(x'|x)\alpha(x'|x)P(x) > g(x|x')\alpha(x|x')P(x'), \quad (17)$$

se queremos que esse equilíbrio aconteça, podemos maximizar o lado direito escolhendo $\alpha(x|x') = 1$ (estamos deixando todos as transições $x' \rightarrow x$ acontecer). Então impomos novamente o equilíbrio

$$\alpha(x'|x) = \frac{P(x)g(x|x')}{P(x')g(x'|x)}, \quad (18)$$

Esse expressão ainda pode ser maior que 1 (caso em que acontecem mais transições $x \rightarrow x'$), finalmente chegamos à *acceptance function* de Metropolis

$$\alpha(x'|x) = \min \left(1, \frac{P(x)g(x|x')}{P(x')g(x'|x)} \right). \quad (19)$$

Algorítmo:

1. Escolha um ponto inicial x_0 ;
2. Gere um candidato x' a partir de uma distribuição $g(x'|x_i)$;
3. Calcule a taxa de aceitação $\alpha(x'|x_i)$;
4. Gere um número aleatório u uniformemente distribuído em $[0, 1]$;
5. Se $u < \alpha(x'|x_i)$ aceite o candidato e defina $x_{i+1} = x'$, caso contrário rejeite o candidato e defina $x_{i+1} = x_i$;

Problema de sinal

Em LQCD geralmente são calculados valores médios utilizando integrais de caminho da forma

$$\langle A \rangle_\rho = \frac{\int \mathcal{D}\sigma A[\sigma] \rho[\sigma]}{\int \mathcal{D}\sigma \rho[\sigma]}, \quad (20)$$

quando a parte fermiônica já é integrada analiticamente

$$\rho = \det M[\sigma] e^{-S[\sigma]}, \quad (21)$$

e temos algo do tipo

$$M = \not{D} + m + \mu \gamma^0. \quad (22)$$

O operador \not{D} satisfaz

$$\gamma_5 \not{D} \gamma_5 = \not{D}^\dagger, \quad (23)$$

enquanto que $\gamma_5 \gamma^0 \gamma_5 = -\gamma^0$. Então

$$\det(\not{D} + m + \mu \gamma^0) = \det(\not{D} + m - \mu^* \gamma^0)^\dagger, \quad (24)$$

ou seja, para termos um determinante real precisamos de μ zero ou puramente imaginário⁷.

⁷P. de Forcrand, *PoS*, 2009, **LAT2009**, ed. C. Liu e Y. Zhu, 010.

Voltando ao valor médio, precisamos ter um determinante real e positivo para podermos utilizar o método de Monte Carlo utilizado para calcular as integrais de caminho com dimensionalidade muito alta. Temos algo do tipo

$$\int \mathcal{D}\sigma e^{-S} \sim \int d\sigma_1 d\sigma_2 \dots d\sigma_N e^{-S}. \quad (25)$$

Podemos reescrever integrais difíceis como valores médios utilizando uma distribuição uniforme⁸

$$\frac{1}{b-a} \int_a^b dx f(x) = \langle f \rangle_\rho = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N f(x_i), \quad (26)$$

em que x_i são amostras retiradas da distribuição uniforme $\rho(x) = 1/(b-a)$.

⁸C. Gattringer e C. B. Lang, *Quantum chromodynamics on the lattice*, Springer, Berlin, 2010, vol. 788.

Porém, pontos que não contribuem tanto ainda são amostrados e com a mesma importância, aumentando bastante a variância do resultado. Vemos que no caso da integral de caminho o fator e^{-S} da importância diferente para diferentes configurações de campo. Podemos fazer uma amostragem com *importance sampling*, considere

$$\langle f \rangle_\rho = \frac{\int_a^b dx f(x) \rho(x)}{\int_a^b dx \rho(x)} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N f(x_i), \quad (27)$$

em que x_i agora são amostrados de acordo com a probabilidade

$$dP(x) = \frac{\rho(x)dx}{\int_a^b dx \rho(x)}. \quad (28)$$

Em particular

$$\langle A \rangle = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N A[\sigma_n], \quad (29)$$

em que as configurações de campo U_n são amostradas de acordo com a probabilidade

$$dP[\sigma] = \frac{e^{-S[\sigma]} \mathcal{D}\sigma}{\int \mathcal{D}\sigma e^{-S[\sigma]}}. \quad (30)$$

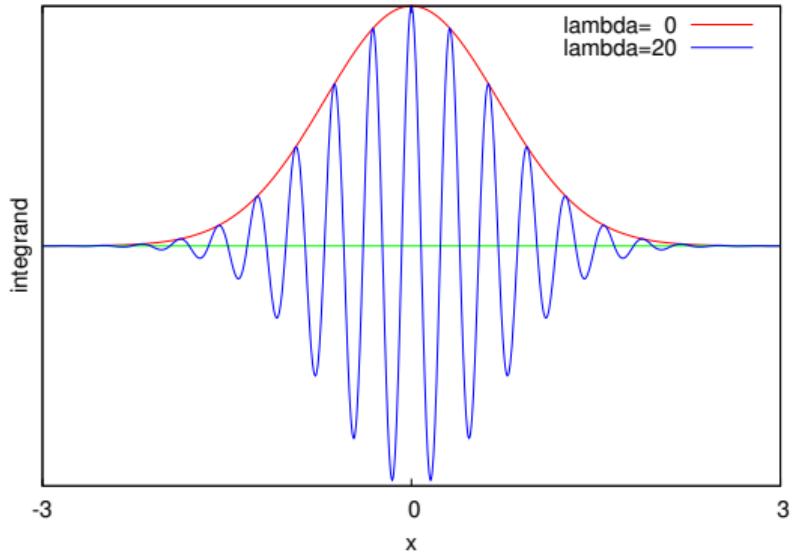


Figura: Integração de uma função oscilatória $f(x) = \exp(-x^2 + i\lambda x)$. Imagem de P. de Forcrand, PoS, 2009, LAT2009, ed. C. Liu e Y. Zhu, 010