

Razotkrivanje strukture nukleona pomoću neuronskih mreža

Marko Cvitković

Mentor: prof.dr.sc. Krešimir Kumerički

Fizički odsjek, Prirodoslovno-matematički fakultet, Bijenička 32, Zagreb

28.01.2020.

Teorijski uvod

- ▶ motivacija: poznavanje generaliziranih partonskih distribucija (GPD) - daju detaljnu sliku raspodjele kvarkova i gluona u nukleonima
- ▶ 8 realnih GPD-ova: $H^q, E^q, H_T^q, E_T^q, H_T^q, E_T^q, H_T^q$ i E_T^q
- ▶ GPD-ovi povezani sa komtonskim form faktorima (CFF) preko konvolucija:

$$\mathcal{F}(\xi, t) = \sum_q e_q^2 \int_{-1}^1 dx \times \left[\frac{1}{\xi - x - i\epsilon} - \frac{1}{\xi + x - i\epsilon} \right] F^q(x, \xi, t) \quad (1)$$

$$\mathcal{F}(\xi, t) = \sum_q e_q^2 \int_{-1}^1 dx \times \left[\frac{1}{\xi - x - i\epsilon} - \frac{1}{\xi + x - i\epsilon} \right] F^q(x, \xi, t) \quad (2)$$

Teorijski uvod

- ▶ CFF-ovi - 8 kompleksnih funkcija: $\mathcal{F} = \mathcal{H}, \mathcal{E}, \dots$
- ▶ u najnižem redu doprinose $\mathcal{H}, \mathcal{E}, \mathcal{H}$ i \mathcal{E}
- ▶ CFF-ovi ovise o dvije varijable: bezdimenzionalnom ξ koji je blizak Bjorkenovoj x varijabli x_B i Mandelstamovoj t varijabli
- ▶ GPD funkcije imaju ovisnost o još jednoj varijabli
- ▶ glavni i najpogodniji eksperiment za određivanje CFF funkcija je duboko virtualno komptonско raspršenje (DVCS)

Teorijski uvod

- ▶ ulazni foton virtualan i dobija se zračenjem leptona - leptoprodukcija fotona $ep \rightarrow ep\gamma$

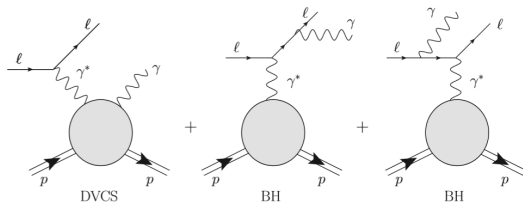


Figure: Leptoprodukcija realnog fotona kao koherentna superpozicija DVCS i Bethe-Heitler amplituda. Preuzeto iz [2].

- ▶ mjeri se udarni presjek za proces $ep \rightarrow ep\gamma$ i on je dan kao složena formula nepoznatih CFF-ova
- ▶ Kako iz tih mjerenja odrediti 4 kompleksne CFF funkcije ?

Teorijski uvod

- ▶ pomoglo bi kad bi imali snopove leptona različitih naboja (elektron i pozitron) i spinova, mogućnost različite polarizacije mete i kada bi za metu mogli koristiti i proton i neutron
- ▶ imali bi različite opservable kojima različiti CFF-ovi doprinose drugačije
- ▶ korištene opservable: *BSD*, *BSSw0*, *BSSw1*, *TSA* i *BCA*
- ▶ standardne metode izvlačenja (ekstrakcije) nepoznatih CFF-ova iz generiranih opservabli: lokalna i globalna ekstrakcija uz teorijske pretpostavke
- ▶ unos pristranosti u model → motivacija za neuronske mreže

Metoda: neuronske mreže

- ▶ pokušaj pronalaska matematičke reprezentacije obrade informacija u biološkim sustavima (McCulloch i Pitts 1943.)
- ▶ široki spektar algoritama koji se danas koristi za prepoznavanje uzoraka modeliranih po uzoru na ljudski mozak
- ▶ prvi korak: algoritam perceptrona za binarnu klasifikaciju

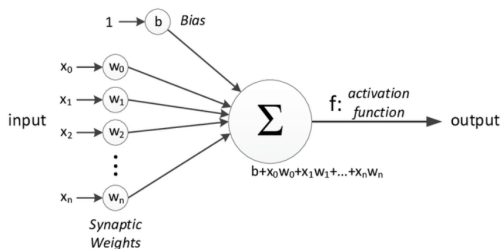


Figure: Perceptron. Preuzeto iz [9].

Metoda: neuronske mreže

- ▶ x_0, x_1, \dots i x_n predstavljaju značajke jednog primjera, dok je b pristranost
- ▶ sve značajke ulaze u model sa jednom od $n + 1$ težina $w_0 = 1, w_1, \dots, w_n$ koje su proporcionalne važnosti pripadne značajke:

$$a = b + \sum_{j=1}^n w_j \cdot x_j \quad (3)$$

- ▶ predikcija za dani primjer dobija se djelovanjem aktivacijske funkcije na skalarni produkt značajki sa težinama
- ▶ u slučaju perceptrona: step funkcija \rightarrow linearan model
- ▶ optimizacija: ažuriranje težina sve dok za konačni vektor težina sve predikcije postanu jednake stvarnim oznakama primjera

Metoda: neuronske mreže

- ▶ neuronske mreže \approx višeslojni perceptron; ulazni čvorovi kombiniraju se sa pripadnim težinama na opisani način u više čvorova u takozvanom skrivenom sloju koji sada služi kao ulaz za sljedeći sloj

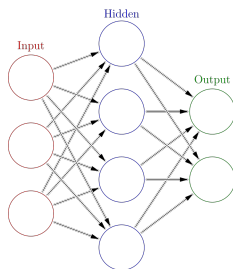


Figure: Umjetna neuronska mreža. Preuzeto iz [8].

- ▶ bitna razlika: aktivacijska funkcija

Metoda: neuronske mreže

- ▶ linearna aktivacijska funkcija kombinira slojeve u linearnu kombinaciju ulaza; mala složenost i kapacitet za učenje i modeliranje većine podataka

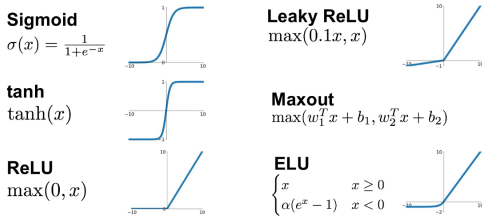


Figure: Nelinearne aktivacijske funkcije kod neuronskih mreža. Preuzeto iz [10].

- ▶ neuronske mreže - koriste se nelinearne aktivacijske funkcije koje se primjenjuju na skrivene slojeve → neuronske mreže kao univerzalni aproksimatori

Metoda: neuronske mreže

- ▶ prilagodba težina u svrhu što boljeg obavljanja zadaće
- ▶ učenje se odvija algoritmom propagacije unatrag (engl. backpropagation)
- ▶ gradijenti pogreške po svim težinama - ažuriranje težina + optimizacija hiperparametra
- ▶ ulančano deriviranje - puno efikasniji račun

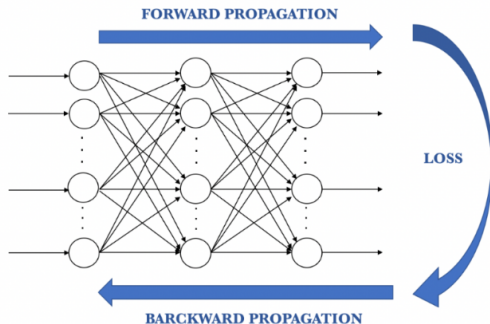


Figure: Ciklus u enja neuronske mreže. Preuzeto iz [9].

Rezultati i diskusija

- ▶ zamislimo da znamo sve CFF funkcije od interesa - pojednostavljeni Goloskokov-Kroll model
- ▶ pomoću poznatih BMK formula generiraju se vrijednosti opservabli za CFF-ove u odabranom kinematičkom rasponu sa simuliranom pogreškom - simulirani podaci

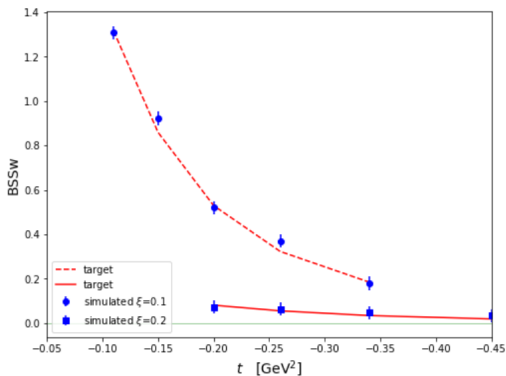
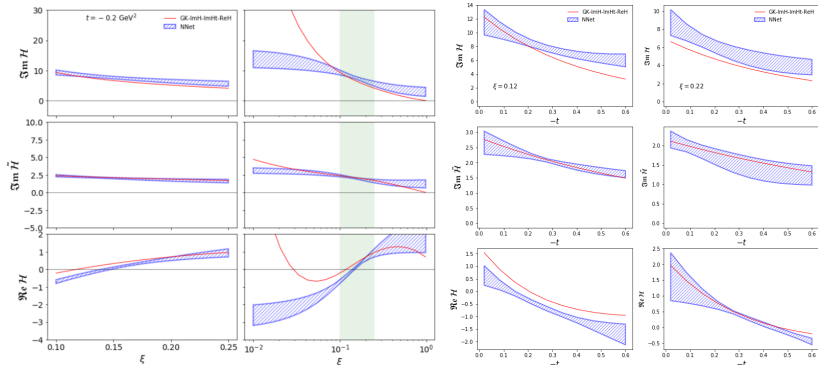


Figure: Prikaz simuliranih podataka za opservablu BSSw1 u ovisnosti o kinemati kim varijablama

Rezultati i diskusija

- ▶ ekstrahiramo željene CFF-ove parametrizirane neuronskim mrežama sa kinematičkim varijablama na ulazu, uspoređujući tako dobijene udarne presjeke (preko izvučenih CFF-ova) sa referentnim modelom te tako od nasumičnih težina u mreži dobijemo one koje minimiziraju pogrešku
- ▶ dobijene CFF-ove uspoređujemo sa poznatim CFF-ovima (dobijenim pomoću referentnog modela) te promatramo je li čitava procedura konzistentna
- ▶ kao relevantne uzimaju se neuronske mreže za koje je vjerojatnost uz pripadajući χ^2 (tzv. "goodness of a fit") veća od praga 0.05
- ▶ parametar *nnets* određuje koliko takvih mreža želimo - propagiranje eksperimentalne nepouzdanosti
- ▶ dobiju se vrpce koje odgovaraju nepouzdanosti - standardne devijacije oko prosječne vrijednosti od CFF-ova ekstrahiranih *nnets* puta

Primjer: istovremeno uspješno ekstrahirano sva tri promatrana CFF-a iz svih pet opservabli



- ▶ Cilj: u tako stvorenom okruženju pokazati jesu li neke opservable pogodnije za ekstrakciju CFF-ova
- ▶ za razne kombinacije simuliranih opservabli pokušavamo uspješno ekstrahirati po jednu ili više CFF-ova te onda računati pogrešku u odnosu na referentni model

Rezultati i diskusija

Koristi se korijen iz kvadratne srednje pogreške (*engl. root mean square error*) koja je definirana kao:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_i - \hat{Y}_i)^2}, \quad (4)$$

- ▶ prvi korak: ekstrahirati iz svih 5 opservabli po 1 CFF, dominantni doprinos

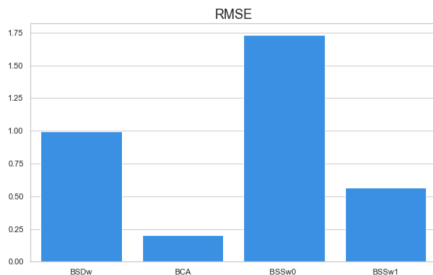


Figure: RMSE dominantnog doprinosa CFF-a pri ekstrakciji iz jedne opservable

Rezultati i diskusija

- ▶ drugi korak: RMSE pri ekstrakciji po 2 dominantna CFF-a uz sve kombinacije od po dvije opservable
- ▶ ukupni RMSE koji odgovara ekstrakciji uz kombinaciju dvije opservable izračunat je kao prosjek dvije pogreške uspješno ekstrahiranih dvaju CFF-ova

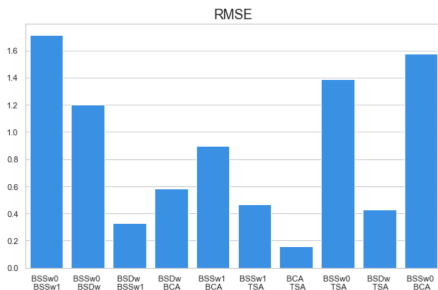


Figure: Prosje ni RMSE dvaju CFF-ova pri ekstrakciji iz dvije opservable

- ▶ kombinacija opservabli *BCA* i *TSA* daje najbolje rezultate

Rezultati i diskusija

- ▶ treći korak: u ovom slučaju su ekstrahirana sva tri CFF-a za svaku kombinaciju opservabli te je shodno tome prikazan prosjek RMS pogrešaka za po tri ekstrahirana CFF-a

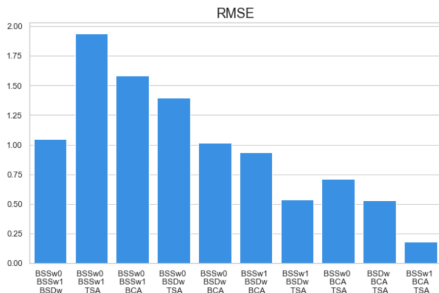


Figure: Prosje ni RMSE tri CFF-a pri ekstrakciji iz tri opservable

- ▶ ponovno primjećujemo da se među kombinacijama s najmanjom prosječnom pogreškom najčešće pojavljuju kombinacija *BCA* i *TSA* uz moguće treće varijable

Rezultati i diskusija

- ▶ četvrti korak: iz svih mogućih kombinacija od po 4 opservable i iz svih pet opservabli, račune se prosječna RMS pogreška pri ekstrakciji sva 3 promatrana CFF-a

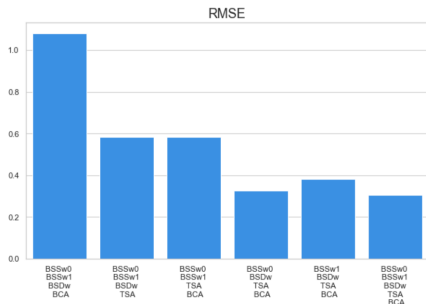


Figure: Prosje ni RMSE tri CFF-ova pri ekstrakciji iz etiri i pet opservabli

- ▶ po treći put se vidi da se kombinacija opservabli *BCA* i *TSA* pojavljuje u sve tri kombinacije sa najmanjom prosječnom pogreškom

Zaključak

- ▶ rezultati ukazuju da su *BCA* i *TSA* u kombinaciji najpogodnije za ekstrakciju tri promatrana Comptonska form faktora
- ▶ time se potvrđuju neki od rezultata (I. Ćorić u svom diplomskom radu [11]), gdje se za to koriste metode dubokog učenja
- ▶ napomena je da su u ovom radu korištene neuronske mreže sa jednim skrivenim slojem sa najčešće 7 neurona nasuprot arhitektura u modernom paketu *Tensorflow* sa puno većim brojem skrivenih slojeva i neurona u njima u [11]

Literatura

- [1] Accardi, A., Albacete, J.L., Anselmino, M. et al, *Electron Ion Collider: The Next QCD Frontier*, Eur. Phys. J. A (2016) 52: 268.
- [2] Krešimir Kumerički, Simonetta Liuti, and Herve Moutarde, *GPD phenomenology and DVCS fitting. Entering the high-precision era.*, European Physical Journal A, 52:157, June 2016.
- [3] Krešimir Kumerički, Dieter Mueller, Andreas Schafer, *Neural network generated parametrizations of deeply virtual Compton form factors*, A. J. High Energ. Phys. (2011) 2011: 73.
[https://doi.org/10.1007/JHEP07\(2011\)073](https://doi.org/10.1007/JHEP07(2011)073)
- [4] S. V. Goloskokov and P. Kroll, *The role of the quark and gluon GPDs in hard vector-meson electroproduction*, Eur. Phys. J. C53 , 2008, 367–384
- [5] Burkert, V.D., Elouadrhiri, L. Girod, F.X., *The pressure distribution inside the proton.*, Nature 557, 396–399 (2018)
[doi:10.1038/s41586-018-0060-z](https://doi.org/10.1038/s41586-018-0060-z)
- [6] A.V. Belitsky, D. Müller, A. Kirchner, *Theory of deeply virtual Compton scattering on the nucleon*, Nuclear Physics B, 629, May 2002, 323-392

Literatura

- [7] Kresimir Kumerički, Dieter Mueller, Andreas Schafer, *Parametrizing Compton form factors with neural networks*, Nuclear Physics B, 222-224, January 2012, 199-203
- [8] https://en.wikipedia.org/wiki/Artificial_neural_network
- [9] <https://towardsdatascience.com/introduction-to-artificial-neural-networks-ann-1aea15775ef9>
- [10] <https://towardsdatascience.com/complete-guide-of-activation-functions-34076e95d044>
- [11] I. Ćorić, *Istraživanje kvarkovsko-gluonske strukture protona pomoću strojnog učenja*, 2019 (Diplomski rad). Dostupno na: <https://urn.nsk.hr/urn:nbn:hr:217:856388>

HVALA NA PAŽNJI !

