

# Analiza svojstava čestičnih snopova pomoću neuralnih mreža

Ivan Jerčić

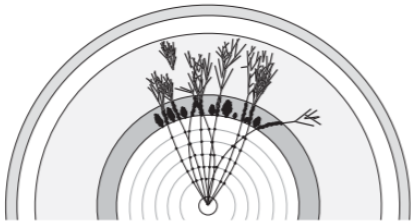
Fizički odsjek, Prirodoslovno-matematički fakultet, Bijenička 32, Zagreb

U ovom radu najprije smo simulirali sustav 10 čestica jasno definiranih masa i kanala raspada. Taj sustav je naš model stvarnosti koji postavljamo kao fizikalnu “istinu” i želimo na neki način reproducirati distribucije vjerojatnosti raspada čestica. Uz pretpostavku da imamo samo podatke koje sustav producira u detektoru, predlažemo metodu za nalaženje gustoća vjerojatnosti raspada čestica određenih masa. Metoda uključuje iterativno primjenjivanje neuralne mreže u ulozi klasifikatora između događaja koje u detektoru ostavi stvarna distribucija i događaja koje ostavi neka testna distribucija. Na predloženi način uspjeli smo rekonstruirati nepoznate kanale raspada u stvarnoj distribuciji, uključujući i skrivene rezonance koje se ne pojavljuju u konačnom stanju u detektoru. Dobiveni rezultati iz testne distribucije vrlo se dobro poklapaju sa stvarnim gustoćama vjerojatnosti. Na kraju dajemo prijedlog za primjenu metode na stvarnim podacima.

## I. UVOD

### I.1. Teorijska pozadina

Hipoteza zatočenja u QCD-u govori nam da su sva izolirana kvantna stanja bojno neutralna. Upravo zbog toga kvarkove, koji nose naboj boje, nikad ne nalazimo kao slobodne čestice, već uvijek zatvorene unutar hadrona. Međutim, u sudarima visokih energija, na primjer u procesu  $e^+e^- \rightarrow q\bar{q}$ , nastaju kvarkovi, a ne hadroni. Kao posljedica takvih interakcija, energija u polju jake interakcije dvaju kvarkova pretvara se u još parova kvarkova i antikvarkova kroz proces hadronizacije. Proces hadronizacije odvija se na putu duljine oko  $10^{-15}$  m te rezultira **snopom**, odnosno “**jetom**” hadrona. Dakle, pojedini kvark opažamo kao jet koji se uglavnom sastoji od nabitih čestica i fotona. U praksi često ne možemo razlučiti sve čestice jeta zbog toga što je vijek trajanja pojedinih čestica premalen.



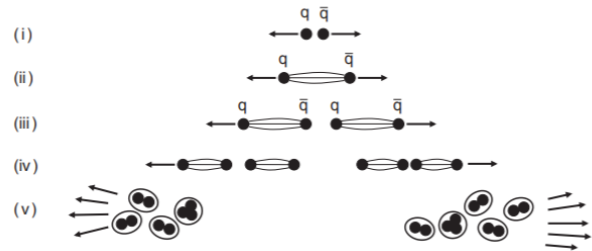
Slika 1. Prikaz “jeta” u detektoru.

Trenutno je razumijevanje detalja procesa hadronizacije poprilično oskudno, ali postoji mnogo fenomenoloških modela koje eksperimentalni podaci podržavaju. Pet glavnih stadija hadronizacije prikazano je na slici 2:

- (i) U interakciji nastane par kvark i antikvark  $q\bar{q}$ .
- (ii) Tijekom njihova razdvajanja polje boje sadržano je u cijevi gustoće energije oko 1 GeV/fm.
- (iii) Kako se kvarkovi nastavljaju odvajati, energija sadržana u polju boje dovoljna je za stvaranje no-

vih  $q\bar{q}$  parova te je razdvajanje polja boja u manje “cjevčice” energetski povoljnije.

- (iv) Isti proces se nastavlja i proizvodi se još parova  $q\bar{q}$  sve dok
- (v) svi kvarkovi i antikvarkovi imaju dovoljno nisku energiju i spajaju se u bezbojne hadrone. Rezultat ovog procesa su dva “jeta” hadrona, jedan usmjeren u početnom smjeru kvarka te drugi usmjeren u početnom smjeru antikvarka [1].



Slika 2. Pet glavnih stadija hadronizacije kvark – antikvark para  $q\bar{q}$ .

### I.2. Opis simuliranog sustava

U ovom odjeljku poblize opisujemo izmišljeni sustav čije ćemo ponašanje pokušati rekonstruirati. Sustav se smatra beskonačnim, izotropnim i izoliranim od svih vanjskih utjecaja. Sastoji se od 10 vrsta čestica, koje označavamo slovima A–J. Mase pojedinih čestica (u proizvoljnim jedinicama) te njihove vjerojatnosti raspada na druge čestice prikazani su na slici 3. Pod “vjerojatnosti raspada” mislimo na odnose amplituda pojedinih raspada.

Čestice A–E mogu se detektirati u zamišljenom detektoru jer su očekivani put koji prijeđu za svog života i veličina detektora istog reda veličine. U tablici se to može očitati kao vjerojatnost da će se pojedina čestica

“raspasti na sebe samu”, što zapravo znači da je njen vijek trajanja dovoljno dugačak da bi se mogla detektirati u tom stanju. Od navedenih, čestice A i C su potpuno stabilne.

	A		B		C		D		E	
masa	0.1		0.6		1.3		1.9		4.4	
raspadi	1	A	0.7	B	1	C	0.3	AC	0.6	CC
			0.3	AA			0.3	AA	0.4	E
							0.4	D		
	F		G		H		I		J	
masa	6.1		8.4		14.2		18.1		25	
raspadi	0.5	AA	0.9	BB	0.6	DD	1	FG	0.5	FI
	0.5	BC	0.1	AF	0.25	DE			0.4	GH
					0.15	EF			0.1	EE

Slika 3. Tablica čestica u simuliranom sustavu.

Čestice F–I su skrivene rezonance: njihov očekivani životni vijek je vrlo malen te se one nikada neće detektirati. Čestica J je početna čestica (“majka”) od koje počinje raspad i koja se nakon toga ne pojavljuje u snopu.

Podaci se generiraju u navedenom sustavu tako da se kreće od početne čestice u mirovanju. Generiranje jeta od navedene čestice jednako je u slučaju onog što nazivamo “pravim” podacima i “testnim” podacima i opisano je u odjeljku II.2. Podatke koje generiramo pomoću masa i vjerojatnosti opisanih u ovom odjeljku nazivamo “pravim” podacima.

## II. EKSPERIMENTALNA METODA

U ovom seminaru predlažemo generalizaciju i nadogradnju postojeće metode u kojoj se pretpostavlja da su gustoće vjerojatnosti neovisne o masi te da su frakcije energije i impulsa u raspadu međusobno nekorelirane [2]. Nakon generiranja pravih podataka, želimo reproducirati skrivene rezonance te vjerojatnosti pojedinih raspada, gdje odsad pa nadalje držimo da su nam ti podaci nepoznati. Za reprodukciju navedenih parametara koristimo drugi set podataka koji nazivamo “testnim” setom i neuralnu mrežu u ulozi klasifikatora.

### II.1. Opis metode pomoću primjera

U ovom odjeljku pokušat ćemo pojasniti ideju korištenja klasifikatora za rekonstrukciju neke distribucije. Klasifikator treniran na dva skupa podataka kao izlaz daje vrijednost (“score”)  $s \in [0, 1]$ , pri čemu score 0 označava sigurnu pripadnost “testnom” skupu, a score 1 sigurnu pripadnost “pravom” skupu. Za podatak  $x \in \mathbb{R}^m$ , dvije distribucije određene parametrima  $\Theta_{\text{pravi}}, \Theta_{\text{test}} \in \mathbb{R}^k$ , vrijednost koju daje klasifikator, koju

označavamo s  $C_{\text{NN}}(x)$ , u idealnoj je situaciji jednaka:

$$C_{\text{NN}}(x) = \frac{\mathcal{L}(\Theta_{\text{pravi}}|x)}{\mathcal{L}(\Theta_{\text{pravi}}|x) + \mathcal{L}(\Theta_{\text{test}}|x)}, \quad (1)$$

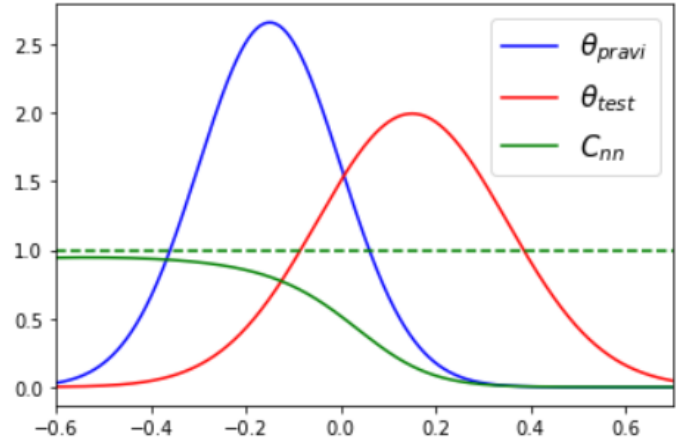
pri čemu je  $\mathcal{L}(\Theta_i)$  vjerojatnost (likelihood) da podatak  $x$  pripada distribuciji  $\Theta_i$ . Nadalje, ako parametri distribucija  $\Theta$  određuju neprekidnu vjerojatnosnu distribuciju, tad vrijedi:

$$\begin{aligned} \mathcal{L}(\Theta_{\text{pravi}}|x) &= \rho(x|\Theta_{\text{pravi}}), \\ \mathcal{L}(\Theta_{\text{test}}|x) &= \rho(x|\Theta_{\text{test}}), \end{aligned} \quad (2)$$

pri čemu su  $\rho(x|\Theta_i)$  gustoće vjerojatnosti pripadnost podatka  $x$  nekoj distribuciji  $\Theta_i$ . Inverzijom izraza (1) sad možemo eksplicitno odrediti  $\rho(x|\Theta_{\text{pravi}})$  uz poznate  $\rho(x|\Theta_{\text{test}})$  i izlaze klasifikatora  $C_{\text{NN}}$ :

$$\rho(x|\Theta_{\text{pravi}}) = \frac{C_{\text{NN}}(x)}{1 - C_{\text{NN}}(x)} \rho(x|\Theta_{\text{test}}) \quad (3)$$

Na slici 4 možemo vidjeti odnos distribucija vjerojatnosti i izlaza iz klasifikatora na primjeru dvije normalne razdiobe.



Slika 4. Primjer gustoća i rezultata pripadajućeg klasifikatora.

Iako smo ovdje ilustrirali način na koji izlaz klasifikatora i nama poznate testne distribucije možemo koristiti za rekonstrukciju prave razdiobe, u praksi, pa čak već i u ovom primjeru, susrećemo se s nekoliko problema: vrlo visoka dimenzionalnost prostora parametara i međuzavisnost dozvoljenih procesa (zbog ovisnosti distribucije o masi) onemogućuju eksplicitno rješavanje problema te za procjenu prave distribucije koristimo neuronsku mrežu.

### II.2. Generator jetova

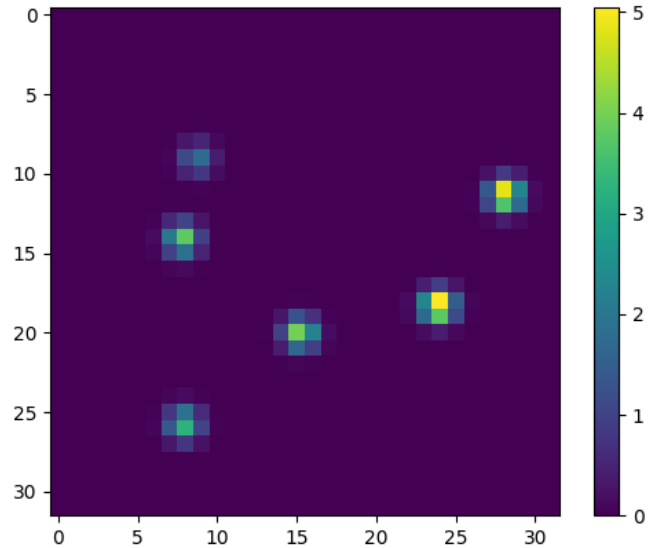
U ovom odjeljku dajemo opis generiranja jetova za slučaj “pravih” i “testnih” podataka.

1. Počinjemo od čestice u mirovanju mase  $m_0 = 25$ .
2. Čestica se dijeli na dvije nove čestice. Mase i momente novih čestica određujemo na sljedeći način:
  - (a) Promatramo familiju vjerojatnosnih gustoća  $\rho_M : \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $M \in \mathcal{M}$ . Treba naglasiti da je u teoriji skup brojeva koji mogu biti masa neke čestice konačan, ali mi ne znamo unaprijed koji su to brojevi. Također, ne očekujemo da će te distribucije biti bliske za bliske mase te zato nema smisla modelirati  $\rho_M$  neprekidnom funkcijom. Stoga, familiju vjerojatnosnih gustoća modeliramo na način da podijelimo interval  $[0, 25]$  na 50 jednakih podintervala, uz pretpostavku da će pojedina masa “upasti” u svoj podinterval, te postavljamo  $\rho_M = \text{const}$ . na svakom podintervalu.
  - (b) Za dani  $M$  iz distribucije određene s  $\rho_M$  određujemo par  $(a, b)$  uz  $a \in [0, 1]$ ,  $b \in [0, 1]$ . Gustoće  $\rho_M$  modeliramo kao po dijelovima konstantne funkcije na mreži dobivenoj dijeljenjem intervala  $[0, 1]$  na 30 podintervala. S time je ukupan broj parametara koji opisuju jednu familiju jednak 45 000. Sad računamo:

$$\begin{aligned}
 \frac{aM}{2} &= E_1, & \frac{abM}{2} &= p_1 \\
 p_2 &= -p_1, & \left(1 - \frac{a}{2}\right)M &= E_2 \\
 m_1^2 &= M^2 \left( \frac{a^2}{4} - \frac{a^2b^2}{4} \right) \\
 m_2^2 &= M^2 \left( \left(1 - \frac{a}{2}\right)^2 - \frac{a^2b^2}{4} \right) \quad (4)
 \end{aligned}$$

- (c) Time smo odredili mase i količine gibanja novonastalih čestica. Njihovu prostornu distribuciju postavljamo kao uniformnu u kutovima  $(\theta, \phi)$ ,  $\theta \in [0, \pi]$ ,  $\phi \in [0, 2\pi]$ .
3. Nakon prvog raspada, proces 2. ponavljamo za svaku od dvije dobivene čestice 3 puta, nakon čega dobivamo najviše 8 čestica. Situacija gdje se čestica nije nastavila dijeliti odgovara odabiru parametra  $a = 0$ . Treba naglasiti da prilikom ponavljanja procesa za svaku novu česticu parametre  $a$  i  $b$  biramo iz druge distribucije jer su distribucije ovisna o masi. Dodatno, parametri za pojedinu distribuciju odnose se na sustav centra mase nove čestice pa ih treba i Lorentz-transformirati u sustav centra mase početne čestice, koji se poklapa s referentnim laboratorijskim sustavom.
4. Konačnu listu podataka s energijama i količinama gibanja svake od  $n$  čestica nazivamo “jet”. Jet je određen sa  $1 + 2 + 4 = 7$  parova  $(a, b)$  te 7 parova  $(\theta, \phi)$ , tj. ukupno 28 parametara.

Jet vizualno prikazujemo histogramom (slikom) dimenzija 32x32 piksela. Osi histograma predstavljaju kutove  $(\theta, \phi)$  diskretizirane na 32 moguće vrijednosti, dok je boja svakog piksela određena energijom čestice koja se nalazi na tim koordinatama. Na energije čestice dodan je gaussijanski šum, što odgovara stvarnoj situaciji, u kojoj je u očitavanju sa senzora uvijek prisutna nesigurnost.



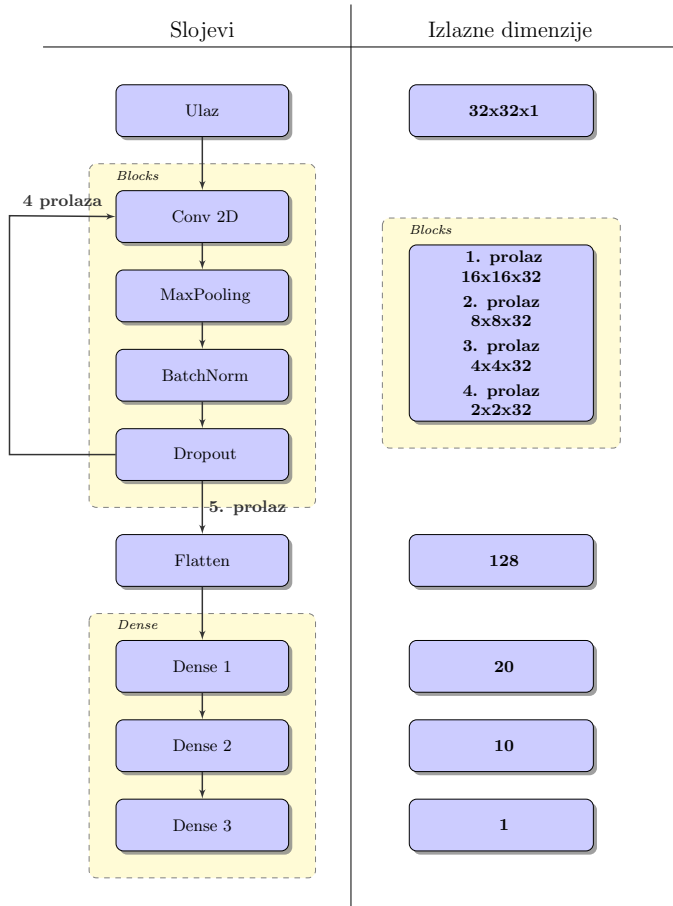
Slika 5. Primjer prikaza jeta, tj. histograma energija. Osi na slici predstavljaju kutove  $(\theta, \phi)$  dok je boja svakog piksela određena energijom čestice u proizvoljnim jedinicama.

### II.3. Klasifikator

Klasifikator koji koristimo za pronalaženje prave familije distribucija je “feed forward” konvolucijska neuronska mreža (CNN). Korištena arhitektura je prikazana na slici 6. Sastoji se od bloka slojeva u kojemu se nalaze dvodimenzionalni konvolucijski sloj (sa 32 filtera i (3,3) kernelom), “MaxPooling” sloja te “batch normalization” i “dropout” sloja. Taj blok se ponavlja 4 puta. Nakon njega slijedi “flatten” sloj te 3 “dense” sloja, gdje je u prva dva korištena ispravljena linearna aktivacijska funkcija (ReLU), dok je na zadnjem korištena sigmoid aktivacijska funkcija. Klasifikator treniramo tako da minimiziramo “binary cross entropy loss” [3], a za optimiziranje težina same konvolucijske neuronske mreže koristimo optimizator AdaM [4].

U svakoj iteraciji uzimamo jednak broj jetova iz prvog skupa podataka i isti broj generiranih pomoću testne familije distribucija dobivene u prethodnoj iteraciji. Za treniranje CNN-a koristimo 75% tih podataka, dok preostalih 25% koristimo za validaciju treniranog modela. Uvjet konvergencije je da klasifikator više nije u stanju

raspoznati iz kojeg skupa su podaci, tj. kad mu točnost bude 50%.



Slika 6. Arhitektura klasifikatora.

#### II.4. Ispravljanje vjerojatnosnih distribucija

Prema izrazu (3) vjerojatnost da se pojavi određeni jet  $z$  opisan parametrima  $z_i = (a_i, b_i)$ , za  $i \in (1, \dots, 7)$  za idealan klasifikator  $C$  iznosi

$$\rho(z) = \frac{C(z)}{1 - C(z)} q(z), \quad (5)$$

pri čemu je  $q(z)$  gustoća vjerojatnosti pojavljivanja mlaza  $z$  u skupu mlazova dobivenih predloženim generatorom. Uz pretpostavku da je svaki raspad međusobno nezavisan te da ovisi samo o masi čestice koja se raspada, jednadžba (5) može se zapisati na sljedeći način:

$$\prod_{i=1}^7 \rho(a_i, b_i | m_i) = \frac{C(z)}{1 - C(z)} \prod_{i=1}^7 q(a_i, b_i | m_i). \quad (6)$$

Logaritmiranjem tog izraza slijedi:

$$\sum_{i=1}^7 \ln \rho(a_i, b_i | m_i) = \ln C(z) - \ln(1 - C(z)) + \sum_{i=1}^7 \ln q(a_i, b_i | m_i). \quad (7)$$

Kako su distribucije  $q(a_i, b_i | m)$  poznate, uz pretpostavku idealnog klasifikatora  $C$ , distribucije  $\rho(a_i, b_i)$  mogu se odrediti rješavanjem linearnog sustava jednadžbi. No, zbog visoke dimenzionalnosti i činjenice da klasifikator nije idealan, rješavanje takvog sustava može se pokazati numerički zahtjevnim i nestabilnim. Umjesto toga, vjerojatnost  $\rho(a_i, b_i | m_i)$  je aproksimirana pomoću nove neuralne mreže  $f$  na sljedeći način:

$$\rho(a, b | m) = \frac{e^{f(a, b, m)}}{\sum_{j, k} e^{f(a_j, b_k, m)}}, \quad (8)$$

pri čemu indeksi  $j, k$  označavaju diskretni interval za vrijednosti parametara  $a$  i  $b$  te  $f(a, b, m)$  predstavlja izlaznu vrijednost neuralne mreže  $f$  za uređenu trojku  $(a, b, m)$ . Na ovaj način osigurava se da su gustoće vjerojatnosti glatke.

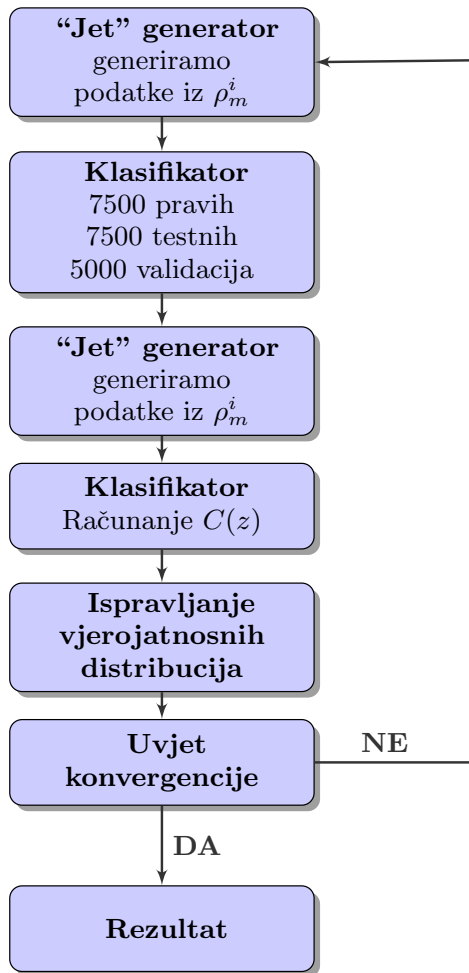
U ovom koraku predloženog algoritma generira se skup mlazova prema distribuciji  $q(z)$  te se za svaki mlaz odredi izlazna vrijednost klasifikatora. Nakon toga slijedi proces učenja neuralne mreže  $f$  u kojem se minimizira  $L_2$ -udaljenost između lijeve i desne strane jednadžbe (7).

#### II.5. Opis algoritma

Na početku postavljamo testnu familiju distribucija  $\rho_M(a, b)$ . Konačni rezultat ne ovisi o testnoj familiji distribucija, ali što smo u početku udaljeniji od prave distribucije, trebat će nam treba više iteracija. Bitno je naglasiti da skup pravih jetova mora bit podskup svih jetova koji se mogu dobiti iz testne familije distribucija. Svaka iteracija algoritma se sastoji od tri koraka koji se zatim ponavljaju.

1. Pomoću jet generatora napravimo skup  $\mathcal{A}$  od 10 000 podataka (slika) koristeći trenutnu testnu distribuciju.
2. Pomoću 10 000 podataka iz testne distribucije te 10 000 pravih podataka (uz napomenu da se 25% svakog skupa koristi za validaciju) treniramo klasifikator kao što je opisano u odjeljku II.3.
3. Pomoću jet generatora napravimo novi skup  $\mathcal{B}$  od 10 000 podataka (slika) koristeći trenutnu testnu distribuciju.
4. Za svaki od podataka iz skupa  $\mathcal{B}$  izračunamo pripadajući score  $C_{NN}$  pomoću klasifikatora istreniranog u 2.

5. Podatke iz 4 koristimo za ispravljanje testne distribucije na način opisan u odjeljku II.4. Time dobivamo novu testnu distribuciju.
6. Provjeravamo uvjet konvergencije opisan u odjeljku II.3. Ako je zadovoljen, trenutna testna distribucija dobro opisuje pravu te je algoritam gotov. Ako uvjet nije zadovoljen, ponovimo sve prethodne korake za novu distribuciju dobivenu u 5.

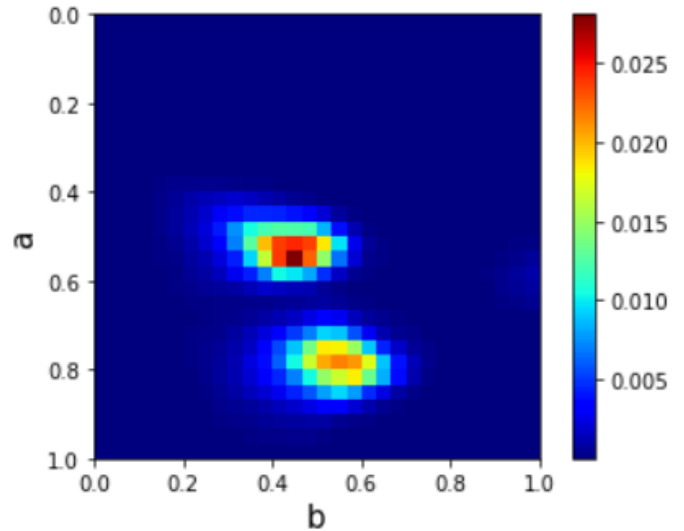


Slika 7. Shematski prikaz algoritma korištenog u radu.

### III. REZULTATI I RASPRAVA

Prvo ćemo promotriti raspad čestice “majke”, tj. raspad čestice mase 25. Sa slike 3. vidimo da se ona može raspasti na više načina: dvije čestice masa 6.1 i 18.1 sa vjerojatnosti 50%, dvije čestice masa 8.4 te 14.2 sa vjerojatnosti 40% ili na dvije jednake čestice masa 4.4 sa vjerojatnosti 10%. Kad to pomoću relacija (4) prevedemo u parametre  $a$  i  $b$  dobivamo  $(a, b) \in$

$\{(0.54, 0.41), (0.79, 0.53), (1.00, 0.94)\}$ . Na slici 8 prikazujemo dvodimenzionalnu raspodjelu parametara  $a$  i  $b$  dobivenu opisanim algoritmom prilikom raspada čestice mase 25 nakon 936 iteracija. Ovaj broj iteracija je odabran jer smo u toj fazi izrade rada više smo pažnje posvećivali prilagođavanju same mreže i dokazivanju predloženog koncepta te stoga još uvijek nismo imali jasno određen kriterij konvergencije između testne i prave raspodjele, a rezultati su izgledali obećavajuće.



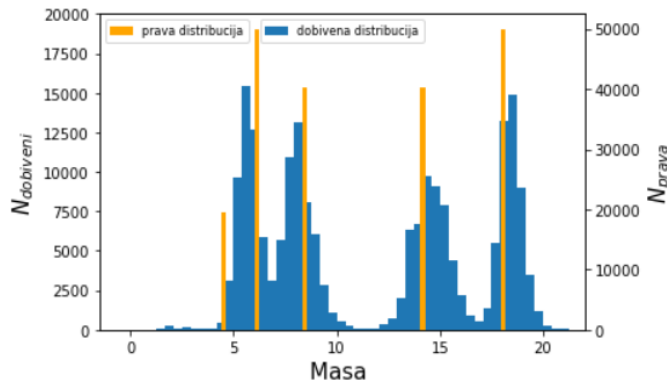
Slika 8. Dvodimenzionalna distribucija parametara  $a$  i  $b$  prilikom raspada čestice mase 25 nakon 936 iteracija.

Možemo primijetiti da su se pojavile dvije grupacije oko točaka  $(0.51, 0.41)$  i  $(0.79, 0.51)$  koje vrlo dobro odgovaraju parovima vrijednosti  $(a, b)$  za dva najvjerojatnija raspada u pravim podacima. Treći mogući raspad nije vidljiv u prikazanoj raspodjeli. Iz dobivenih testnih distribucija nakon 936 iteracija moguće je naći i mase čestica nastalih nakon raspada čestice “majke”, što prikazujemo na histogramu na slici 9. Za ovaj prikaz je dobivena testna distribucija primijenjena na 100 000 podataka. Narančastim linijama označene su mase čestica koje se pojavljuju u pravim podacima, dok su plavim stupcima označene pojavnosti masa čestica dobivenih testnim podacima. Vidi se da je poklapanje vrlo dobro, iako još uvijek ne postoje jasne naznake pojavljivanja čestica masa 4.4, što i ne čudi toliko, s obzirom da se radi o raspadu najmanje vjerojatnosti.

Kako bi kvantitativno odredili koliko dobro je algoritam rekonstruirao mase pojedinih čestica, na dobivene podatke napravili smo prilagodbe normalnih raspodjela, što prikazujemo na slici 10. Specifično, prilagodba koja je vidljiva na slici dana je s:

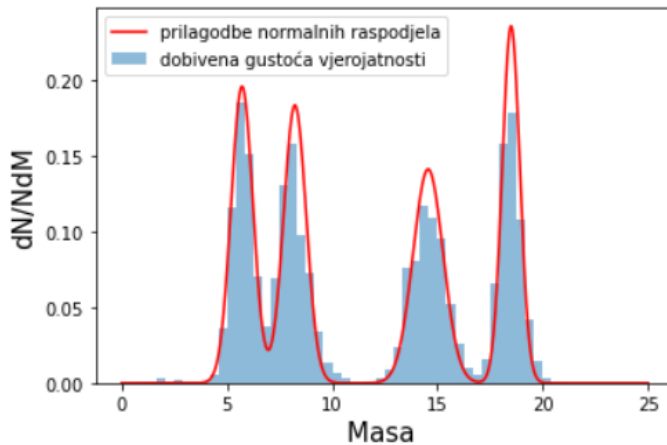
$$\frac{dN}{NdM} = C \sum_{i=1}^4 \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{1}{2} \left(\frac{M - M_{0i}}{\sigma_i}\right)^2\right), \quad (9)$$

pri čemu je  $M_{0i}$  centroid pojedine normalne raspodjele, a  $\sigma_i$  njegoova širina. Parametar  $C$  određen je tako da



Slika 9. Histogram masa nakon raspada čestice “majke” ( $m = 25$ ) na 100 000 podataka. Narančastim linijama označene su mase čestice koje se pojavljuju u pravim podacima, dok su plavim stupcima označene pojavnosti masa čestica dobivenih testnim podacima.

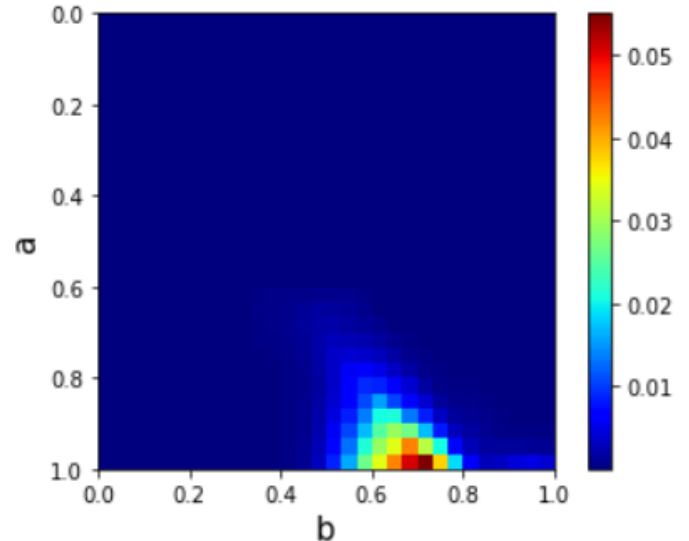
je sveukupna raspodjela normalizirana na 1. Dobiveni parametri prilagodbe ( $M_{0i}, \sigma_i$ ) koji najbolje opisuju podatke su (5.7,0.7), (8.2,0.7), (14.6,0.8) i (18.5,0.6). Usporedbom s parametrima stvarne distribucije vidimo da su sve reproducirane mase unutar jedne standardne devijacije od stvarnih masa, što dodatno potvrđuje uspješnost metode.



Slika 10. Normalizirani histogram masa nakon raspada čestice “majke” ( $m = 25$ ) na 100 000 podataka. Crvena linija je prilagodba sume normalnih raspodjela na dobivene podatke. Parametri prilagodbe i njihova usporedba sa stvarnim parametrima dani su u tekstu.

U mogućnosti smo razmotriti i daljnje raspade čestica dobivenih pomoću testne distribucije nakon 936. iteracije. Kao primjer, na slici 11 je prikazana dobivena dvodimenzionalna raspodjela parametara  $a$  i  $b$  za raspad čestice mase 18.1. Na dobivenoj raspodjeli može se uočiti samo jedna grupacija oko točke (0.90,0.63), što se isto vrlo dobro poklapa s jedinim mogućim raspadom u pravim podacima, u kojima se čestica mase 18.1 može raspasti samo na čestice mase 6.1 i 8.4. Za taj raspad vrijed-

nost parametra  $a$  iznosi 0.90, dok vrijednost parametra  $b$  iznosi 0.66. Rezultati analiza za daljnje čestice uglavnom pokazuju istu uspješnost metode kao i za česticu mase 25, te ih nismo uključili u ovaj rad.



Slika 11. Dvodimenzionalna distribucija parametara  $a$  i  $b$  za raspad čestice mase 18.1. U ovom slučaju dozvoljen je samo jedan raspad u pravim podacima, što se i ovdje očituje kao jedna jasna nakupina centrirana oko (0.9,0.63).

#### IV. ZAKLJUČAK

U ovom seminaru predstavljena je metoda za nalaženje skrivenih gustoća vjerojatnosti fizikalnih sustava koji uključuju kaskadne raspade čestica. U tu svrhu razvijen je generator događaja fizikalnog sustava koji uključuje samo diskretne raspade određenih vjerojatnosti. Cilj metode je odrediti gustoće vjerojatnosti za raspade čestice određene mase. Na taj se način mogu pronaći čak i skrivene rezonance koje se ne pojavljuju u konačnom stanju. Dobiveni rezultati iz testne distribucije vrlo se dobro poklapaju s pravim gustoćama vjerojatnosti te su skrivene rezonantne čestice uglavnom uspješno identificirane, pogotovo u raspadima čestica veće mase.

Ovaj rad služi kao dokaz koncepta tako da hiperparametri korištenih neuralnih mreža još uvijek nisu optimizirani, što uzrokuje nepreciznosti u konačnim rezultatima. Također, u ovom trenutku uvjet konvergencije nije dovoljno jasno definiran. U budućem istraživanju potrebno je jasno odrediti mjeru pomoću koje će se definirati uvjet konvergencije. Jedna moguća mjera je diskriminatorska moć klasifikatora.

Prednost predložene metode je to što način na koji podaci trebaju biti reprezentirani nije određen pa se može koristiti bilo kakva neuralna mreža kao klasifikator. To omogućava primjenu metode na realne podatke prikupljene nekim od postojećih visokoenergetskih sudarivača



kao što je LHC. Također, metoda se može primijeniti i na kontinuirane gustoće vjerojatnosti kakve se pojavljuju u realnim fizikalnim sustavima te vjerujemo da bi se konvergencija postigla brže jer bi u takvim sustavima bilo manje vrlo nevjerojatnih događaja što bi uzrokovalo da *loss* funkcije neuralnih mreža budu “glade”.

Naravno, navedena metoda u ovom trenutku ima i neke nedostatke. Npr. trenutno je ograničena razlučivost u gustoćama vjerojatnosti zbog diskretizacije prostora parametara  $(a, b)$ . Povećanje razlučivosti moguće je postići finijom diskretizacijom tog prostora parametara, no to bi uvelike otežalo numeričke proračune.

## V. LITERATURA

- [1] Modern Particle Physics - Thomson. Jack Prail.
- [2] Jercic, M.; Poljak, N. Exploring the Possibility of a Recovery of Physics Process Properties from a Neural Network Model. *Entropy* 2020, 22, 994
- [3] Nielsen, M.A. Neural Networks and Deep Learning. Available online: [neuralnetworksanddeeplearning.com/](http://neuralnetworksanddeeplearning.com/)
- [4] Kingma, D.; Ba, J. A Method for Stochastic Optimization. In Proceedings of the International Conference on Learning Representations, Banff, AB, Canada, 14–16 April 2014.