

České vysoké učení technické v Praze
Fakulta jaderná a fyzikálně inženýrská

Katedra fyziky

Obor: Experimentální jaderná a částicová fyzika



Strojové učení pro jetovou fyziku

Machine Learning for Jet Physics

BAKALÁŘSKÁ PRÁCE

Vypracoval:	Georgij Ponimatin
Vedoucí práce:	RNDr. Jana Bielčíková, Ph.D.
Konzultant:	Dr. Ing. Josef Šivic
Rok:	2019



Katedra: fyziky

Akademický rok: 2018/2019

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Student: Georgij Ponimatkin

Studijní program: Aplikace přírodních věd

Obor: Experimentální jaderná a částicová fyzika

Název práce: Strojové učení pro jetovou fyziku
(česky)

Název práce: Machine Learning for Jet Physics
(anglicky)

Pokyny pro vypracování:

1. Jety a algoritmy pro jejich rekonstrukci.
2. Diskuse aktuálních výsledků týkajících se studia produkce jetů na urychlovačích RHIC a LHC.
3. Popis experimentu STAR na urychlovači RHIC a jeho jednotlivých detektorů.
4. Rešerše metod strojového učení a počítačového vidění pro jetovou fyziku.
5. Adaptace existující metody konvoluční neuronové sítě (Mask R-CNN) na detekci jetů na urychlovači RHIC.
6. Příprava simulovaných dat, evaluace modelu a diskuse získaných výsledků.

Doporučená literatura:

- 1] K. Yagi, et al.: Quark Gluon Plasma: From Big Bang to Little Bang, Cambridge University Press, 2008
- [2] R. Vogt: Ultrarelativistic Heavy-Ion Collisions, Elsevier Science, 2007
- [3] I. Goodfellow, et al.: Deep Learning, MIT Press, 2016
- [4] C. M. Bishop: Pattern Recognition and Machine Learning, Springer, 2011
- [5] L. de Oliveira, et al.: Jet Images - Deep Learning edition, JHEP 1607 069 (2016)
- [6] K. He, et al.: Mask R-CNN, 2017 IEEE International Conference on Computer Vision, DOI: 10.1109/ICCV.2017.322 (2017)

Jméno a pracoviště vedoucího bakalářské práce:

RNDr. Jana Bielčíková, Ph.D.

Ústav jaderné fyziky AV ČR, v.v.i.

odborný konzultant: Dr. Ing. Josef Šivic, CIIRC, ČVUT v Praze

Datum zadání bakalářské práce: 22.10.2018

Termín odevzdání bakalářské práce: 08.07.2019

Doba platnosti zadání je dva roky od data zadání.



.....
garant oboru



.....
vedoucí katedry



.....
děkan

V Praze dne 22.10.2018

Prohlášení

Prohlašuji, že jsem svou bakalářskou práci vypracoval samostatně a použil jsem pouze podklady (literaturu, projekty, SW atd.) uvedené v příloženém seznamu.

Nemám závažný důvod proti použití tohoto školního díla ve smyslu § 60 Zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon).

V Praze dne

.....
Georgij Ponimatkin

Poděkování

Chtěl bych poděkovat RNDr. Janě Bielčíkové, Ph.D. a Dr. Ing. Josefu Šivicovi za to, že mi umožnili vytvořit bakalářskou práci, která byla úzce propojená s mými zájmy. Poděkování dále patří UJF AV ČR a taktéž CIIRC ČVUT, který mi poskytl výpočetní kapacity na clusteru, které jsem použil v této práci. V neposlední řadě patří poděkování mé rodině, která mě po celou dobu mého studia podporovala.

Georgij Ponimatkin

Název práce:

Strojové učení pro jetovou fyziku

Autor: Georgij Ponimatkin

Studijní program: Aplikace přírodních věd

Obor: Experimentální jaderná a částicová fyzika

Druh práce: Bakalářská práce

Vedoucí práce: RNDr. Jana Bielčíková, Ph.D.

Ústav jaderné fyziky AV ČR, v.v.i.

Konzultant: Dr. Ing. Josef Šivic

Český institut informatiky, robotiky a kybernetiky, České vysoké učení technické v Praze

Abstrakt: Moderní metody strojového učení značným způsobem ovlivnily obor počítačového vidění, jehož cílem je naučit stroje chápat informaci proudící z vizuálních zdrojů (obrázky, videa atd.). Ukazuje se, že tyto metodiky nacházejí své využití i ve fyzice částic, když budeme chápat detektor jako vysokorychlostní 2D/3D kameru. Především výzkum týkající se aplikací metod strojového učení na jetovou fyziku řešil především otázku tzv. tagování jetů, který se dal chápat v kontextu klasifikace. Logickou evolucí této metody jsou algoritmy pro detekci objektů nebo segmentaci instancí. Tato bakalářská práce se zaměřuje na jeden z těchto algoritmů, Mask R-CNN, s jehož pomocí se provádí detekce jetů ve srážce pomocí jediného vstupního obrázku reprezentujícího rovinu detektoru v azimutálních (φ) a pseodorapiditních (η) souřadnicích. Získané výsledky jsou následně porovnány s klasickým klastrováním částic založeném na jetovém rekonstrukčním algoritmu anti- k_t .

Klíčová slova: strojové učení, počítačové vidění, jetová fyzika, jetové algoritmy, neuronové sítě

Title:

Machine Learning for Jet Physics

Author: Georgij Ponimatkin

Abstract: Modern methods of machine learning have revolutionized the field of Computer Vision, where the main goal is to learn the machine to understand information coming from visual sources (images, video, etc.). It has been shown that those methods can be applied in High Energy Physics as well, if we understand a detector as a high speed 2D/3D camera. Previous research in machine learning applied to jet physics was mainly dealing with a task of jet tagging, which could be understood in the context of a classification task. The logical evolution of those methods are algorithms for object detection and instance segmentation. This bachelor thesis deals with applications of one of such algorithms, Mask R-CNN, which classifies jets in a collision from a single input image, representing a plane of the detector in azimuthal (φ) and pseudorapidity (η) coordinates. Obtained results are compared with common jet clustering approach based on the anti- k_t algorithm.

Key words: machine learning, computer vision, jet physics, jet algorithms, neural networks

Obsah

Úvod	1
1 Částice, interakce a Standardní model	2
1.1 Standardní model	2
1.2 Silná interakce	2
1.2.1 Účinné průřezy v rámci QCD	4
2 Ultrarelativistické srážky těžkých iontů	6
2.1 Evoluce QGP	7
2.2 Kinematické a úhlové proměnné	8
2.3 Centralita a Glauberův model	8
2.4 Sondy kvark-gluonového plazmatu	10
2.4.1 Jaderné modifikační faktory	11
2.4.2 Dvoučásticové korelace	12
2.4.3 Zhášení jetů	14
3 Jetové algoritmy	16
3.1 Sekvenčně-rekombinační algoritmy	16
3.1.1 Jetové pozadí	18
3.2 Dekonvoluce fluktuujícího pozadí	18
4 Urychlovač RHIC a experiment STAR	22
4.1 Experiment STAR	23
4.1.1 TPC	23
4.1.2 BEMC	23
4.1.3 HFT	25
5 Strojové učení	26
5.1 Základní pojmy strojového učení	26
5.2 Formulace učícího problému s učitelem	26
5.3 Řešení učícího problému	27
5.4 Neuronové sítě	28
5.4.1 Konvoluční neuronová síť jako extraktor příznaků	29
5.5 Mask R-CNN	29
5.6 Nedávné výsledky v aplikacích strojového učení na jetovou fyziku . .	31
6 Detekce objektů pro jetovou fyziku	34
6.1 Množina dat	34

6.2	Příprava a trénování modelu	35
6.3	Evaluace modelu	39
Závěr		42

Seznam obrázků

1.1	Částice Standardního modelu. Převzato z [1].	3
1.2	Příklad partonových distribučních funkcí kvarků a gluonů v protonu získaných kolaborací CTEQ s pomocí globálního fitu na experimentálních datech pro $Q = 2$ GeV (vlevo) a $Q = 100$ GeV (vpravo). Převzato z [2].	4
1.3	Schéma evoluce jetu v p+p srážce.	5
2.1	Fázový diagram jaderné hmoty. Převzato z [3].	6
2.2	Evoluce jádro-jaderné srážky. Převzato z [4].	7
2.3	Reakční rovina pro srážku dvou těžkých iontů. Převzato z [5].	9
2.4	Geometrie srážky dvou jader. Převzato z [6].	9
2.5	Diferenciální účinný průřez v závislosti na počtu nabitých částic a následné stanovení centrality pomocí Glauberova modelu. Převzato z [6].	10
2.6	Faktor R_{CP} pro nabitě částice v závislosti na různých hodnotách $\sqrt{s_{NN}}$. Převzato z [7].	12
2.7	Jaderný modifikační faktor R_{AA} D^0 mezonů naměřený experimentem STAR pro různé centrality (vlevo) a porovnání s modifikačními faktory pro lehké částice změřenými experimenty STAR a ALICE (pravá dolní část). Převzato z [8].	12
2.8	Jaderný modifikační faktor R_{AA} pro průměr přes D^+ , D^0 a D^{*+} mezony naměřený experimentem ALICE na LHC v Pb+Pb srážkách při $\sqrt{s_{NN}} = 2,76$ TeV v centralitě 0-10 % a jeho porovnání s jaderným modifikačním faktorem naměřeným experimentem STAR. (vlevo) Stejný jaderný modifikační faktor a jeho porovnání s jaderným modifikačním faktorem pro lehké částice (vpravo). Převzato z [9].	13
2.9	Dvoučásticové korelace pro p+p, d+Au a Au+Au srážky pro $4 \text{ GeV}/c < p_T^{trigger} < 6 \text{ GeV}/c$ a $2 \text{ GeV}/c < p_T < p_T^{trigger}$ pro asociovanou částici (vlevo). Převzato z [10]. Analogické měření pro $8 \text{ GeV}/c < p_T^{trigger} < 15 \text{ GeV}/c$ v různých centralitách Au+Au srážek a hybnostech asociované částice (vpravo). Převzato z [11].	13
2.10	Jaderný modifikační faktor R_{AA} pro jety s rozlišovacím parametrem $R = 0,4$ naměřený experimentem ATLAS na LHC v Pb+Pb srážkách při $\sqrt{s_{NN}} = 2,76$ TeV pro různé intervaly rapidit v závislosti na p_T jetu. Převzato z [12].	14

2.11	Míra dvoujetové asymetrie změřená experimentem STAR pro Au+Au srážky při $\sqrt{s_{NN}} = 200$ GeV. Převzato z [13] (nahore). Analogické měření provedené experimentem CMS pro Pb+Pb srážky při $\sqrt{s_{NN}} = 2,76$ TeV (dole). Převzato z [14].	15
3.1	Vliv jetového algoritmu na výsledný tvar jetů a jejich plochy. Vlevo lze vidět příklad pro k_t algoritmus s $R = 1$, vpravo pro anti- k_t s $R = 1$. Převzato z [15].	18
3.2	Rozdělení δp_T pro $p_T^{emb} = 20,0$ GeV/ c . Červené a šedé spektrum představuje jednočásticový jet bez a s vlivem kolektivních jevů. Modré spektrum představuje jety generované modelem Pythia . Převzato z [16].	19
3.3	Příklad jetových spekter v Au+Au srážce při energii $\sqrt{s_{NN}} = 200$ GeV před dekonvolucí (černé body) a po (fialové, červené a modré body). Převzato z [17]	20
4.1	Schéma urychlovače RHIC. Převzato z [18].	22
4.2	Řez experimentem STAR. Převzato z [18].	23
4.3	Časově projekční komora (TPC) experimentu STAR. Detektor pokrývá rozsah pseudorapidity od -1 do 1. Převzato z [18].	24
4.4	Barelový elektromagnetický kalorimetr (BEMC) experimentu STAR. Analogicky jako TPC i tento detektor má pokrytí od -1 do 1 v pseudorapiditě. Převzato z [18].	24
4.5	(vlevo) Modul detektoru BEMC. (vpravo) Dělení modulu BEMC na jednotlivé věže. Převzato z [18].	25
4.6	Průřez detektorem HFT. Převzato z [19].	25
5.1	Příklady detekcí pomocí algoritmu Mask R-CNN. Převzato z [20]. . .	29
5.2	Architektura Mask R-CNN. Převzato z [20].	30
5.3	Příklad kotev. Převzato z [21].	30
5.4	Operace RoIAAlign . Převzato z [20].	31
5.5	Příklad jetového obrázku zprůměrovaného přes velký počet jetů. Osa Q_1 vznikla rotací a translací souřadnice φ , osa Q_2 rotací a translací souřadnice η . Převzato z [22].	32
5.6	Průměrná γ sprška v GEANT4 (nahore) a průměrná γ sprška v CALOGAN (dole). Převzato z [23].	33
6.1	Příklad masky, ve které se vyskytuje problém s nespojitostí ve φ . Vodorovná osa představuje diskretizované hodnoty φ v intervalu $[0, 2\pi]$, svislá osa hodnoty η v intervalu $[-1, 1]$	36
6.2	Prodloužení původní masky ze 128 na 192 pixelů dle vodorovné osy. Toto odpovídá prodloužení detektoru z $[0, 2\pi]$ do $[0, 3\pi]$	37
6.3	Maska po zpracování 20×20 morfologickou dilatací.	37
6.4	Separace dvou dilatovaných oblastí.	38
6.5	Konečná podoba masky, která jde na vstup do modelu Mask R-CNN. . .	38
6.6	Příklad detekce (červený rámeček) v porovnání s pravdivým jetem (zelený rámeček). Skóre 1.0/0.68 říká, že se jedná o jet s pravděpodobností 100% a IoU mezi predikcí a pravdivým objektem je 68%. . .	39

6.7	Efektivita detekce jetů v p+p srážkách pomocí modelu Mask R-CNN trénovaného na konstantním p_T spektru. Červené body ukazují výsledky modelu při testování na množině dat s konstantním p_T , zelené s realistickým p_T . Lze vidět, že model trénovaný na této trénovací množině dosahuje přibližně stejných efektivit napříč celým intervalem p_T , což je zapříčiněno rovnoměrným zastoupením jetů s různými hybnostmi. Některé chyby měření jsou menší než velikosti naměřených bodů.	40
6.8	Efektivita detekce jetů v p+p srážkách pomocí modelu Mask R-CNN trénovaného na realistickém p_T spektru. Červené body ukazují výsledky modelu při testování na množině dat s plochým p_T , zelené s realistickým p_T . Na rozdíl od modelu, který byl trénován na množině dat s konstantním p_T , dosahuje tento model horších výsledků pro jety s vysokým p_T . Toto je způsobeno nerovnoměrným zastoupením jetů v dané množině dat, kde dominují jety s nízkým p_T . Některé chyby měření jsou menší než velikosti naměřených bodů.	41
6.9	Efektivita detekce jetů v Au+Au srážkách pomocí modelu Mask R-CNN. Analogicky jako v předchozím případě je i výkon tohoto modelu ovlivněn nerovnoměrným zastoupením jetů v trénovací množině. Chyby měření jsou menší než velikosti naměřených bodů.	41

Úvod

Moderní metodiky strojového učení značným způsobem pozměnily obory počítačového vidění, zpracování přirozeného jazyka a jiné. Je proto zajímavou otázkou, jakým způsobem lze tyto moderní metodiky využít pro účely částicové a jaderné fyziky.

Strojové učení se v částicové fyzice používá již od 90. let minulého století. Tyto metodiky se ovšem soustředily převážně na klasifikaci (např. určení zda-li má trojice částic původ v rozpadu D -mesonu) nebo na učení bez učitele (redukce dimenzionality experimentálních dat). Takové metodiky ale představují jen malou část moderních algoritmů strojového učení. Cílem této práce je prozkoumat, zda-li nedávno vyvinuté metody detekce a segmentace instancí objektů mohou mít využití pro jetovou fyziku.

Práce je členěná následujícím způsobem: v první kapitole jsou uvedeny základní vlastnosti interakcí a částic, kde se zejména soustředíme na silnou interakci. Ve druhé kapitole jsou probrány základní vlastnosti fyziky ultrarelativistických srážek těžkých iontů, včetně několika experimentálních měření. Třetí kapitola se zaměřuje na popis jetových algoritmů. Čtvrtá kapitola se zabývá uspořádáním experimentu STAR. Pátá kapitola je krátkým úvodem do problematiky strojového učení a aplikací strojového učení na jetovou fyziku. Poslední, šestá kapitola, popisuje aplikaci modelu Mask R-CNN pro detekci jetů v $p+p$ a $Au+Au$ srážkách.

Kapitola 1

Částice, interakce a Standardní model

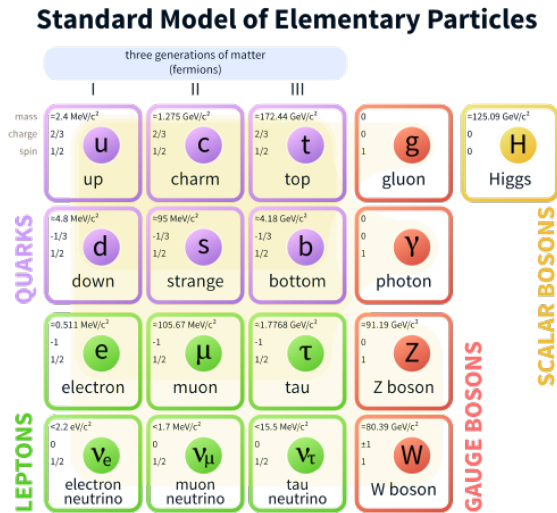
1.1 Standardní model

V rámci Standardního modelu jsou zahrnuty tři fundamentální interakce - elektromagnetická, slabá a silná, které jsou přenášeny kalibračními bosony. Elektromagnetickou interakci zprostředkovává foton (γ), který je částicí bez hmotnosti a náboje. Slabá interakce je zprostředkována masivními bosony Z^0 a $W^\pm$. Silná interakce je zprostředkována gluony, částicemi bez hmotnosti a s barevným nábojem. Posledním přírůstkem do rodiny bosonů je Higgsův boson (objevený v roce 2012), který odpovídá za mechanismus narušení elektroslabé symetrie.

Dále v rámci Standardního modelu máme částice hmoty, které jsou tvořeny fermiony. Ty se dělí se na dvě podkategorie - leptony a kvarky. Mezi leptony patří elektron (e), mion (μ) a tauon (τ) a jejich antičástice. Ke každé z těchto částic ještě existují jim odpovídající neutrina, ν_e, ν_μ a ν_τ , což jsou částice bez náboje, které interagují s pomocí slabé interakce. Mezi kvarky řadíme částice u, c a t s nábojem $+\frac{2}{3}e$ a d, s, b s nábojem $-\frac{1}{3}e$, kde e je elementární náboj. Z kvarků a gluonů se pak tvoří kompozitní částice zvané hadrony. Obr. 1.1 shrnuje přehled o elementárních částicích, které byly uvedeny v předešlých dvou odstavcích.

1.2 Silná interakce

Zbytek této kapitoly je věnován některým vlastnostem silné interakce, jelikož její projevy jsou stěžejní pro tuto práci. Silná interakce je popsána pomocí teorie zvané **kvantová chromodynamika** (QCD), která popisuje dynamiku interakcí mezi kvarky a gluony. Tato teorie zavádí nové kvantové číslo, kterému se říká barevný náboj, v analogii s nábojem elektrickým. Existují tři hodnoty barevných nábojů, nejčastěji jsou označovány jako červená (r), zelená (g) a modrá (b). Je nutné poznamenat, že



Obrázek 1.1: Částice Standardního modelu. Převzato z [1].

se nejedná o barvu v pravém slova smyslu, nýbrž o abstrakci. Barevný náboj kvarků je pak dán těmito náboji, resp. antináboji v případě antikvarků. Pro gluony pak existuje 8 různých nábojových stavů, které jsou kombinací základních nábojů.

Síla interakce je udávána pomocí tzv. vazebné konstanty. Pro QCD je tato vazebná konstanta proměnlivá v závislosti na kvadrátu přenesené čtyřhybnosti Q^2 jako

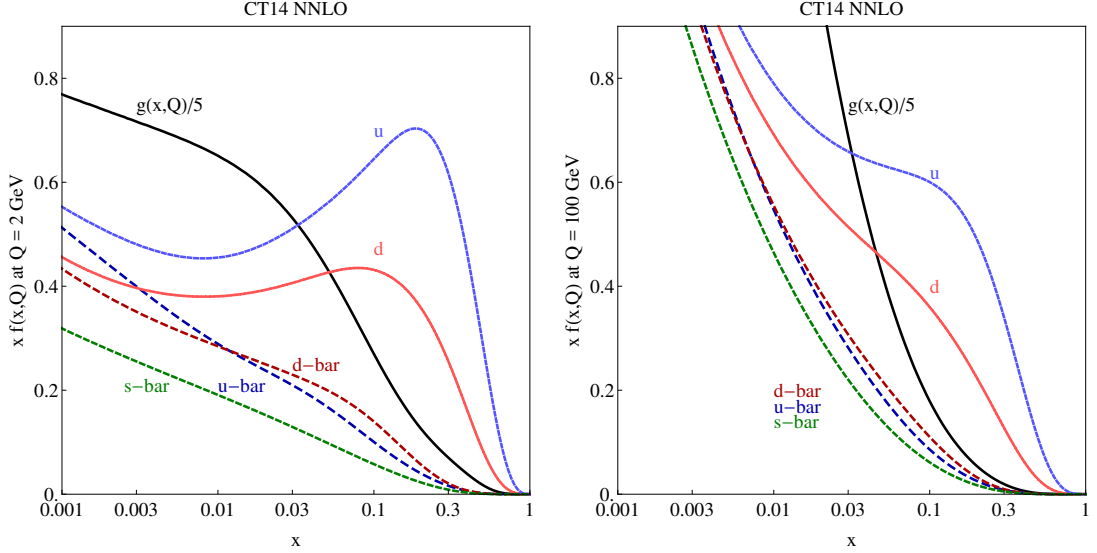
$$\alpha_s(Q^2) = \frac{12\pi}{(33 - 2N_f) \ln\left(\frac{Q^2}{\Lambda_{QCD}^2}\right)}, \quad (1.1)$$

kde N_f je počet kvarkových vůní, které se účastní interakce a $\Lambda_{QCD} \approx 220 \text{ MeV}$ je charakteristická energetická škála pro QCD [24].

Pro případ $\lim_{Q^2 \rightarrow +\infty} \alpha_s(Q^2) = 0$ dostáváme asymptotickou svobodu, která říká že čím vyšší energie, tím je menší interakce mezi částicemi účastnicími se silné interakce. S použitím této konstanty můžeme sestavit empirický potenciál jako

$$V(r) = -\frac{4}{3} \frac{\alpha_s}{r} + kr, \quad (1.2)$$

kde r je vzdálenost mezi kvarky a k je parametr udávající sílu lineárního členu. S pomocí tohoto potenciálu můžeme zjednodušeně vysvětlit jev barevného uvěznění, který způsobuje to, že se v přírodě nevyskytují volné kvarky. V případě, že by se vzdálenost mezi dvěma kvarky r zvětšovala, od jisté hodnoty vzdálenosti r_c se mezi těmito kvarky deponuje dostatečné množství energie pro tvorbu $q\bar{q}$ páru.



Obrázek 1.2: Příklad partonových distribučních funkcí kvarků a gluonů v protonu získaných kolaborací CTEQ s pomocí globálního fitu na experimentálních datech pro $Q = 2$ GeV (vlevo) a $Q = 100$ GeV (vpravo). Převzato z [2].

1.2.1 Účinné průřezy v rámci QCD

Obecně je možné účinný průřez pro srážku dvou hadronů při dostatečně vysokých energiích vyjádřit pomocí faktorizačního teoremu [25] jako

$$\sigma = \sum_{i,j,k,l} \int dx_1 dx_2 f_i(x_1, Q_1^2) f_j(x_2, Q_1^2) \hat{\sigma}_{ij \rightarrow kl} D(k \rightarrow h). \quad (1.3)$$

Rozeberme dále jednotlivé komponenty v tomto výrazu:

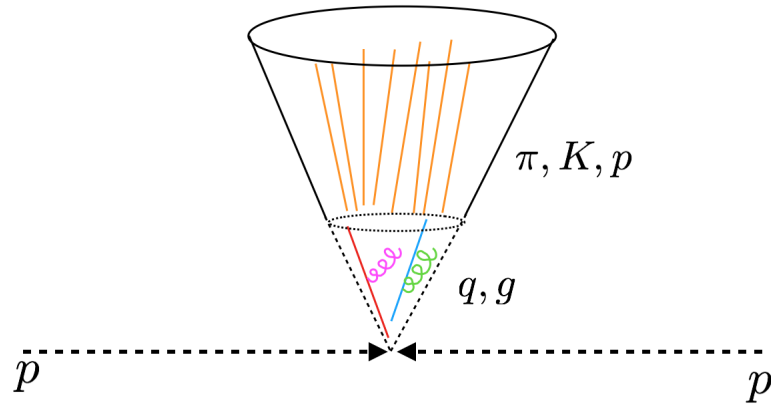
- $f_i(x, Q^2)$ je tzv. partonová distribuční funkce partonu i závislá na frakci hybnosti původního hadronu, který nese daný parton a kvadrátu přenesené čtyřhybnosti. Tato funkce udává rozdělení jednotlivých partonů ve sráženém hadronu. Příklad této funkce lze vidět na Obr. 1.2. Tyto funkce se však měří pro jednotlivé hodnoty Q^2 . Chceme-li získat hodnoty $f_i(x, Q^2)$ pro jiné hodnoty Q^2 , tak můžeme použít DGLAP [26, 27, 28] evoluční rovnici

$$\frac{\partial f(x, Q^2)}{\partial \ln(Q^2)} = \frac{\alpha_s}{2\pi} \int_x^1 \frac{dy}{y} f(y, Q^2) P_{q \rightarrow qg} \left(\frac{x}{y} \right) \quad (1.4)$$

s počáteční podmínkou $f(x, Q_0^2)$. Zde $P_{q \rightarrow qg}$ je větvičí funkce, která udává pravděpodobnost, že se kvark rozdělí na kvark a gluon.

- $\hat{\sigma}_{ij \rightarrow kl}$ je účinný průřez pro reakci $ij \rightarrow kl$ na partonové úrovni, daný výpočtem z poruchové kvantové teorie pole.
- $D(k \rightarrow h)$ je fragmentační funkce, udávající pravděpodobnost, že parton k přejde na pozorovanou částici h . Analogicky jako u partonové distribuční funkce lze spočítat i evoluci fragmentační funkce v závislosti na Q^2 s pomocí DGLAP rovnice (1.4).

V případě velmi energetických srážek dojde k interakci na partonové úrovni a partony vzniklé z této interakce budou následně vyzařovat gluony a hadronizovat. Výsledná kaskáda kolimovaných částic se nazývá jet. Jety hrají významnou roli ve fyzice částic, jelikož jejich vlastnosti úzce souvisí s vlastnostmi partonu, který je vyprodukoval. Příklad zjednodušené evoluce jetu v pp srážce lze vidět na Obr. 1.3.

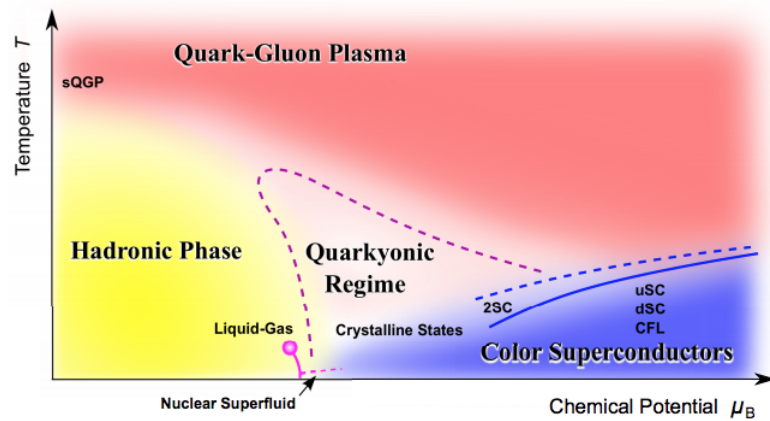


Obrázek 1.3: Schéma evoluce jetu v $p+p$ srážce.

Kapitola 2

Ultrarelativistické srážky těžkých iontů

Kvark-gluonové plazma (QGP) [29, 30] je extrémní stav hmoty, který měl existovat v prvních mikrosekundách existence našeho vesmíru. V tomto stavu hmoty jsou všechny kvarky a gluony kvazivolné. Pro vytvoření tohoto stavu hmoty je potřeba, aby vazebná konstanta α_s byla velmi malá (tj. aby Q^2 bylo velmi velké). Tohoto lze docílit např. srážkou dvou těžkých iontů. Tyto srážky umožňují dosáhnout potřebných hustot energie, které by vedly ke vzniku QGP dle fázového diagramu. Jednu z možných verzí fázového diagramu QGP lze vidět na Obr. 2.1. Zde oblast "Hadronic Phase" představuje jadernou hmotu, kterou normálně pozorujeme v přírodě, "Quark-Gluon Plasma", je oblast, ve které vzniká kvark-gluonové plazma. Oblasti "Quarkyonic Regime" a "Color Superconductors" představují hypotetické stavy jaderné hmoty, které by měly např. existovat v neutronových hvězdách.



Obrázek 2.1: Fázový diagram jaderné hmoty. Převzato z [3].

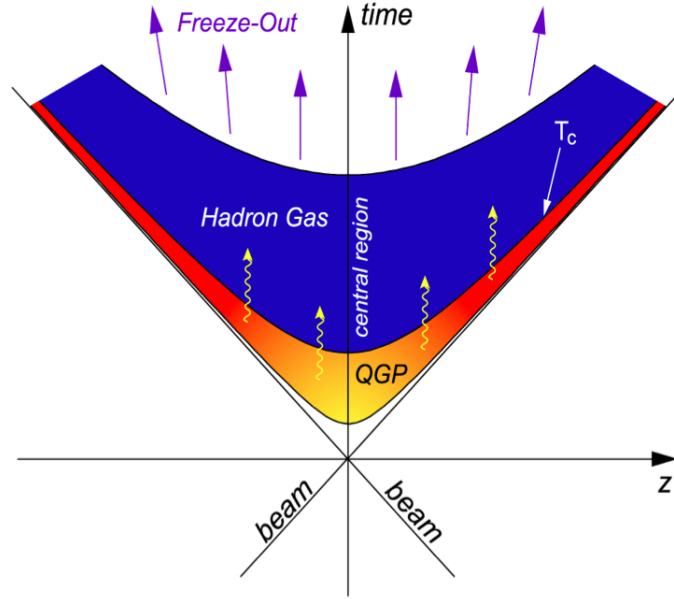
2.1 Evoluce QGP

Od srážky dvou těžkých iontů do svého konce prochází QGP velmi složitou evolucí. Platí-li předpoklad toho, že se médium chová jako ideální tekutina, lze jeho dynamiku popsat pomocí rovnic relativistické hydrodynamiky

$$\partial_\mu T^{\mu\nu} = 0, \quad (2.1)$$

$$\partial_\mu j_B^\mu = 0. \quad (2.2)$$

Zde $T^{\mu\nu}$ představuje energeticko-hybnostní tenzor a j_B^μ hustotu baryonového toku.



Obrázek 2.2: Evoluce jádro-jaderné srážky. Převzato z [4].

Zjednodušeně se dá evoluce jádro-jaderné srážky popsat v následujících krocích:

1. Srážce předchází počáteční fáze, která je ovlivněna geometrií srážky a počátečním rozložením energie v nukleonech.
2. Nastane samotná srážka. Během té může dojít k uvolnění kvarků a gluonů, obsažených v nukleonech. V této fázi vzniká nejvíc těžkých kvarků a též se vytváří majorita partonů které v pozdější fázi hadronizují na jet. Teplota, při které k tomuto ději dojde, se značí T_c a nazývá se kritickou teplotou.
3. Nastane hydrodynamická evoluce média, které se řídí rovnicemi relativistické hydrodynamiky. V důsledku interakce s rozpínajícím se médiem mohou uvolněné partony ztrácet svojí energii. V průběhu evoluce se teplota systému ochladí na teplotu T_{ch} , při které nastane chemická rovnováha systému, v důsledku toho že přestanou probíhat nepružné srážky mezi částicemi a systém přejde do stavu hadronového plynu. To má za následek, že se částicová kompozice systému přestane měnit během jeho další evoluce. V případě, že kritická teplota a počáteční hustota systému budou malé, může systém po srážce

ihned přejít do stavu hadronového plynu. Po jisté době systém dosáhne teploty T_{th} při které nastane kinetická rovnováha v systému. Pro teploty v rozmezí $T_{ch} > T > T_{th}$ se bude v systému měnit jenom rozdělení hybností, což je následkem pružných srážek mezi částicemi. Poté, co se teplota systému sníží pod T_{th} nastane vymrznutí systému, po kterém budou moci být detekovány částice, které z něj pocházely.

2.2 Kinematické a úhlové proměnné

Pro studium dynamiky QGP je třeba zvolit vhodné experimentálně měřitelné proměnné, které dokáží předat co nejvíce informací o systému. První základní proměnnou je tzv. příčná hybnost částice, definovaná jako

$$p_T = \sqrt{p_x^2 + p_y^2}. \quad (2.3)$$

Zde p_x a p_y jsou složky hybnostního vektoru dané částice. Dalšími dvěma proměnnými jsou rapidita (y) a pseudorapidita (η), definované jako

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z}, \quad (2.4)$$

$$\eta = \frac{1}{2} \ln \frac{|\mathbf{p}| + p_z}{|\mathbf{p}| - p_z}. \quad (2.5)$$

V rovnici (2.4) představuje E energii částice a p_z složku vektoru hybnosti částice podél osy z . V rovnici (2.5) představuje $|\mathbf{p}|$ normu hybnostního vektoru částice. Tyto proměnné určují odchylku částic vůči svazku v detektoru.

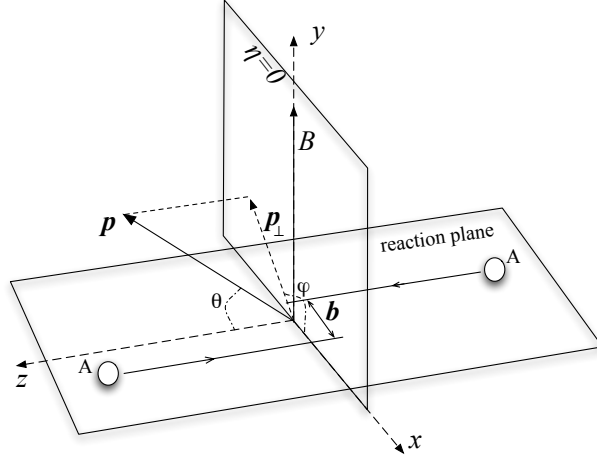
Pro měření některých jevů mohou být taktéž důležitými proměnnými azimutální úhel φ , mezi osami x a y v detektoru a radiální úhel θ mezi osou z a y . Všechny výše uvedené proměnné lze vidět na Obr. 2.3. Zde lze taktéž vidět i tzv. reakční rovinu, což je rovina mezi srážejícími se jádery, která je dána hlavně srážkovým parametrem.

2.3 Centralita a Glauberův model

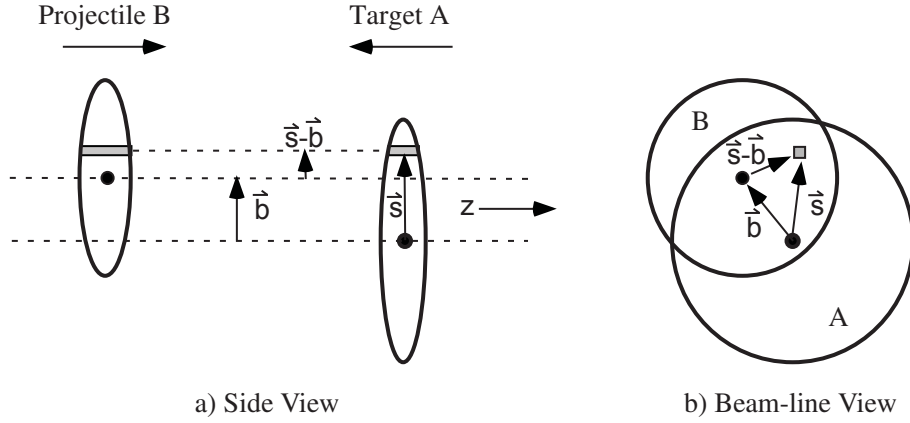
Další důležitou proměnnou v jádro-jaderných srážkách je srážkový parametr $\mathbf{b}$, který lze vidět na Obr. 2.3. To, jak moc se jádra budou při srážce překrývat, značným způsobem ovlivňuje výsledná pozorování. Nejčastěji se tento proces popisuje pomocí Glauberova optického modelu [6]. Nejprve se zavádí parametrizace hustoty jádra pomocí Woods-Saxonova rozdělení jako

$$\rho(r) = \frac{\rho_0}{1 + e^{\frac{r-R}{a}}}, \quad (2.6)$$

kde R je poloměr jádra s A nukleony daný parametrizací $R = 1,20 \cdot A^{\frac{1}{3}}$ fm, a je volný parametr a ρ_0 je hustota v centru jádra.



Obrázek 2.3: Reakční rovina pro srážku dvou těžkých iontů. Převzato z [5].



Obrázek 2.4: Geometrie srážky dvou jader. Převzato z [6].

Nechť se dvě jádra A a B srážejí se srážkovým parametrem $\mathbf{b}$ (viz. Obr. 2.4). Pokud jsou dány dva válce (tzv. flux-tubes) na vzdálenosti $\mathbf{s}$ od centra projektilu A a na vzdálenosti $\mathbf{s} - \mathbf{b}$ od centra projektilu B, bude pravděpodobnost toho, že se nukleon bude nacházet v daném válci

$$T(\mathbf{s}) = \int \rho(\mathbf{s}, z) dz, \quad (2.7)$$

kde $\rho(\mathbf{s}, z)$ je objemová pravděpodobnost pro nalezení nukleonu v $(\mathbf{s}, z)$. Pro pravděpodobnost, že se nukleony budou nacházet v průniku jader A a B platí

$$T_{AB}(\mathbf{b}) = \int T_A(\mathbf{b}) T_B(\mathbf{s} - \mathbf{b}) d^2 s, \quad (2.8)$$

kde T_{AB} je překryvová (tzv. overlap) funkce. Je-li znám celkový neelastický účinný průřez pro nukleon-nukleonové srážky (σ_{inel}^{NN}) můžeme určit celkovou pravděpodobnost n interakcí mezi jádry A a B jako

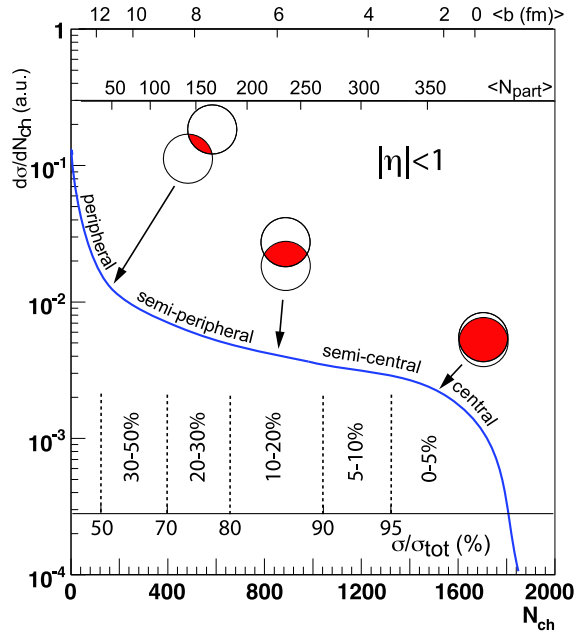
$$P(n, \mathbf{b}) = \binom{AB}{n} \left(T_{AB}(\mathbf{b}) \sigma_{inel}^{NN} \right)^n \left(1 - T_{AB}(\mathbf{b}) \sigma_{inel}^{NN} \right)^{AB-n}. \quad (2.9)$$

Celkový počet nukleon-nukleonových srážek je pak

$$N_{coll}(\mathbf{b}) = \sum_{n=1}^{AB} nP(n, \mathbf{b}) = AB T_{AB}(\mathbf{b}) \sigma_{inel}^{NN}, \quad (2.10)$$

a celkový počet nukleonů účastnících se interakce je

$$N_{part}(\mathbf{b}) = A \int T_A(\mathbf{s}) \left\{ 1 - \left(1 - T_B(\mathbf{s} - \mathbf{b}) \sigma_{inel}^{NN} \right)^B \right\} d^2s + \\ B \int T_B(\mathbf{s} - \mathbf{b}) \left\{ 1 - \left(1 - T_A(\mathbf{s}) \sigma_{inel}^{NN} \right)^A \right\} d^2s. \quad (2.11)$$



Obrázek 2.5: Diferenciální účinný průřez v závislosti na počtu nabitých částic a následné stanovení centrality pomocí Glauberova modelu. Převzato z [6].

Pomocí Glauberova modelu lze pak definovat centralitu, která také vyjadřuje geometrické překrytí jader. Centralita 0 % představuje nejvíce centrální srážky, 90 % představuje periferální atd. Experimentálně se centralita stanovuje např. pomocí závislosti diferenciálního účinného průřezu na počtu nabitých částic (N_{ch}). Příklad takového měření lze vidět na Obr. 2.5. Pro stanovení centrality se pak tato distribuce rozdělí tak, aby počet nabitých částic v jednom binu odpovídal jistému procentu geometrického překrytí při srážce.

2.4 Sondy kvark-gluonového plazmatu

Za pomoci dříve uvedených kinematických a úhlových proměnných je možné studovat různé pozorovatelné, které mohou být citlivé na projevy kvark-gluonového plazmatu. Příklady takových pozorovatelných jsou uvedeny níže.

2.4.1 Jaderné modifikační faktory

Hlavním projevem kvark-gluonového plazmatu je potlačení produkce jednotlivých částic v důsledku interakce partonů/hadronů s médiem. Ztráta energie probíhá pomocí kolizních a radiačních ztrát. Jedním ze způsobů, jak porovnat vliv média na produkci částic, je změření jaderného modifikačního faktoru R_{AA} . Tento faktor porovnává produkci částic v médiu vytvořeném v jádro-jaderné srážce s produkcí částic v p+p srážkách, které jsou škálovány počtem binárních nukleon-nukleonových srážek $\langle N_{coll} \rangle$ získaných z Glauberova modelu

$$R_{AA} = \frac{1}{\langle N_{coll} \rangle} \frac{d^2 N_{AA}/dp_T dy}{d^2 N_{pp}/dp_T dy}. \quad (2.12)$$

Hodnota $R_{AA} \approx 1$ bude vypovídat o tom, že se produkce částic v jádro-jaderných srážkách chová jako superpozice p+p srážek. Hodnota $R_{AA} > 1$ bude vypovídat o nadprodukci částic v jádro-jaderných srážkách. Nakonec, $R_{AA} < 1$ bude vypovídat o potlačení produkce částic v jádro-jaderných srážkách, která je pozorována u lehkých hadronů a půvabných částic [8, 9].

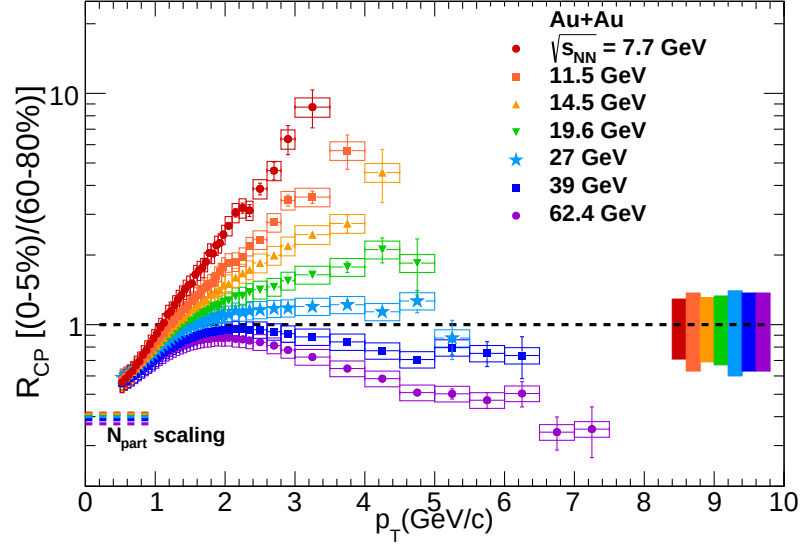
V případě, že není dostupná reference z p+p srážek, lze sestavit faktor R_{CP}

$$R_{CP} = \frac{\langle N_{coll} \rangle_{\text{Central}}}{\langle N_{coll} \rangle_{\text{Peripheral}}} \frac{d^2 N_{\text{Central}}/dp_T dy}{d^2 N_{\text{Peripheral}}/dp_T dy}, \quad (2.13)$$

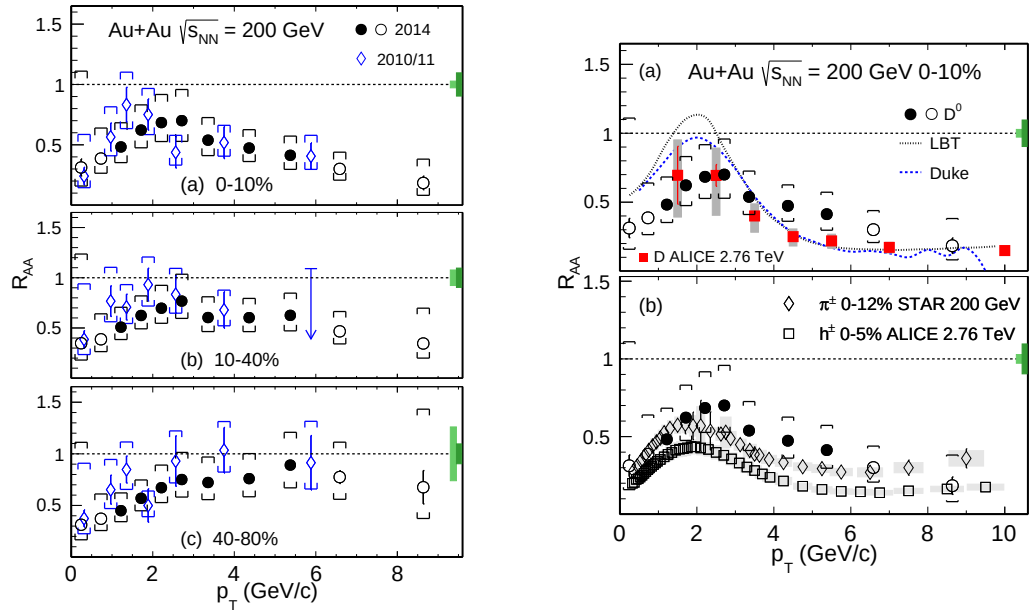
který porovnává produkci částic v centrálních srážkách vůči periferním srážkám. Zde $\langle N_{coll} \rangle_{\text{Central}}$ a $\langle N_{coll} \rangle_{\text{Peripheral}}$ představují počet binárních nukleon-nukleonových srážek v centrálních a periferních srážkách. Na experimentu STAR se pozoruje nadprodukce částic v Au+Au srážkách při energiích $\sqrt{s_{NN}}$ od 7,7 do 19,6 GeV, pro vyšší energie se už zase pozoruje potlačení produkce částic [7]. Toto měření lze vidět na Obr. 2.6.

Z teoretických předpovědí je známo, že by částice obsahující těžké kvarky měly ztrácet méně energie v médiu v důsledku jevu zvaného **dead-cone** efekt [31]. Ten spočívá v tom, že těžké kvarky nemohou vyzařovat gluony pod malými úhly během průchodu médiem a měli bychom pozorovat menší potlačení produkce částic obsahujících těžké kvarky než potlačení pro lehké částice. Experimentálně tento jev ovšem není dodnes pozorován a potlačení částic s těžkými kvarky je obdobné tomu u lehkých částic. Příklad jaderného modifikačního faktoru pro D^0 mezony změřeného pro různé centrality na experimentu STAR v Au+Au srážkách při $\sqrt{s_{NN}} = 200$ GeV lze vidět na Obr. 2.7 vlevo a porovnání tohoto faktoru s lehkými částicemi lze vidět na Obr. 2.7 vpravo v dolní části.

Podobné měření proběhlo i na experimentu ALICE na urychlovači LHC pro Pb+Pb srážky při těžiškové energii $\sqrt{s_{NN}} = 2,76$ TeV. Jaderný modifikační faktor zprůměrovaný přes D^+ , D^0 a D^{*+} mezony lze vidět na Obr. 2.8 vlevo. Vpravo lze vidět porovnání téhož jaderného modifikačního faktoru vzhledem k lehkým částicím. Tato měření jsou v relativní shodě s hodnotami obdrženy experimentem STAR a taktéž neukazují na projev dead-cone efektu.



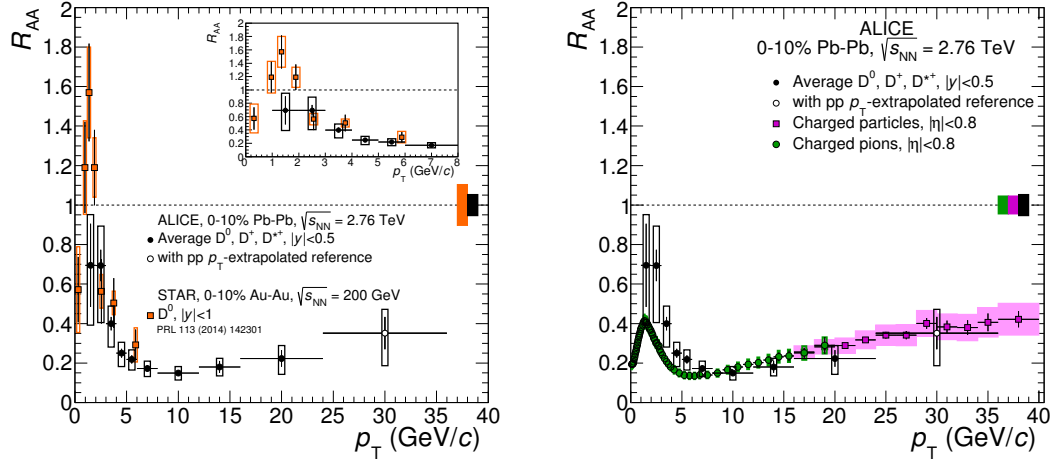
Obrázek 2.6: Faktor R_{CP} pro nabité částice v závislosti na různých hodnotách $\sqrt{s_{NN}}$. Převzato z [7].



Obrázek 2.7: Jaderný modifikační faktor R_{AA} D^0 mezonů naměřený experimentem STAR pro různé centrality (vlevo) a porovnání s modifikačními faktory pro lehké částice změřenými experimenty STAR a ALICE (pravá dolní část). Převzato z [8].

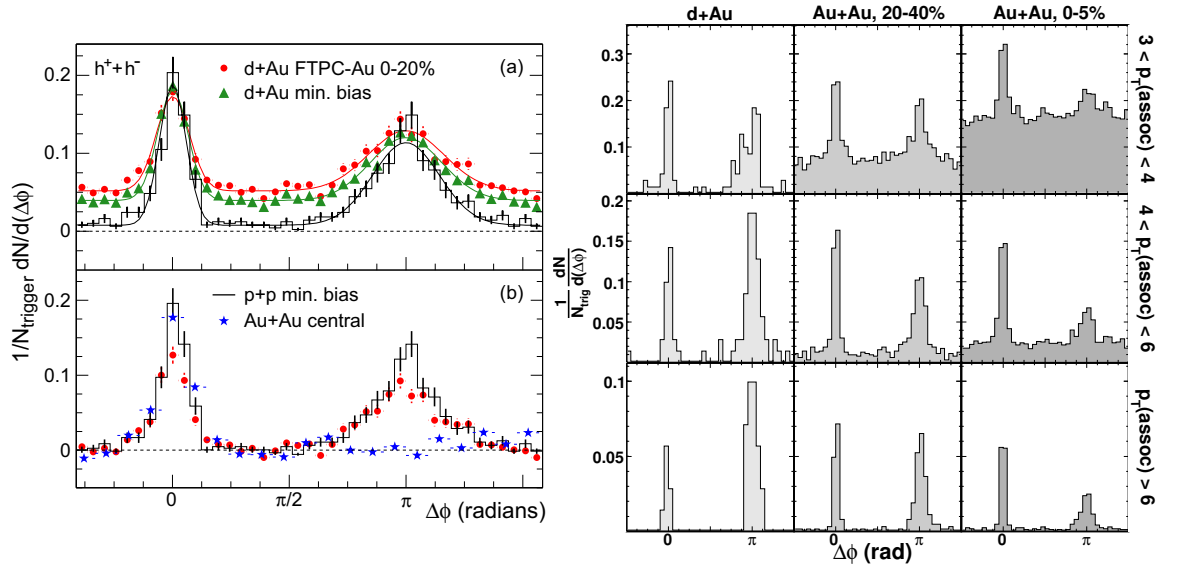
2.4.2 Dvoučásticové korelace

Příkladem dalšího měření, které může odhalit vliv média na produkováné částice, je měření dvoučásticových korelací. Tato metoda spočívá v tom, že se zvolí určitý interval p_T (např. $4 \text{ GeV}/c < p_T < 6 \text{ GeV}/c$) a nalezne se částice, která takové p_T má. Tato částice se označuje jako tzv. **trigger** částice. Pro každou takovou částici se měří azimutální úhlová korelace $\Delta\phi$ s jinými částicemi, které splňují $p_T^{min} < p_T <$



Obrázek 2.8: Jaderný modifikační faktor R_{AA} pro průměr přes D^+ , D^0 a D^{*+} mezony naměřený experimentem ALICE na LHC v Pb+Pb srážkách při $\sqrt{s_{NN}} = 2,76$ TeV v centralitě 0-10 % a jeho porovnání s jaderným modifikačním faktorem naměřeným experimentem STAR. (vlevo) Stejný jaderný modifikační faktor a jeho porovnání s jaderným modifikačním faktorem pro lehké částice (vpravo). Převzato z [9].

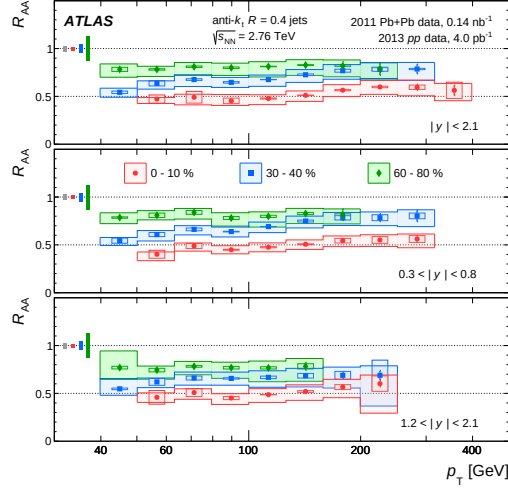
$p_T^{trigger}$ (tzv. asociované částice). Příklady takových měření lze vidět vlevo na Obr. 2.9. V případě p+p a d+Au srážek jsou podle očekávání vidět oba korelační píky, odpovídající tzv. near-side a away-side jetům. Na Obr. 2.9 vpravo lze vidět měření dvoučásticových korelací pro různé hodnoty $p_T^{trigger}$. Lze vidět, že při vysokých p_T u asociovaných částic se v Au+Au srážkách znovu objeví pík v oblasti $\Delta\phi = \pi$.



Obrázek 2.9: Dvoučásticové korelace pro p+p, d+Au a Au+Au srážky pro $4 \text{ GeV}/c < p_T^{trigger} < 6 \text{ GeV}/c$ a $2 \text{ GeV}/c < p_T < p_T^{trigger}$ pro asociovanou částici (vlevo). Převzato z [10]. Analogické měření pro $8 \text{ GeV}/c < p_T^{trigger} < 15 \text{ GeV}/c$ v různých centralitách Au+Au srážek a hybnostech asociované částice (vpravo). Převzato z [11].

2.4.3 Zhášení jetů

Efekty kvark-gluonového plazmatu se též projevují v modifikaci charakteristik jetů. Jedním z projevů je zhášení (částečné či kompletní) jednoho z jetů ve srážce. Jelikož jsou jety tvořeny hadrony, lze jejich potlačení kvantifikovat pomocí faktoru R_{AA} . Příklad takového měření, které provedla kolaborace ATLAS na urychlovači LHC v Pb+Pb srážkách při $\sqrt{s_{NN}} = 2,76$ TeV, lze vidět na Obr. 2.10. Zde naměřené potlačení jetů je prakticky konstantní.

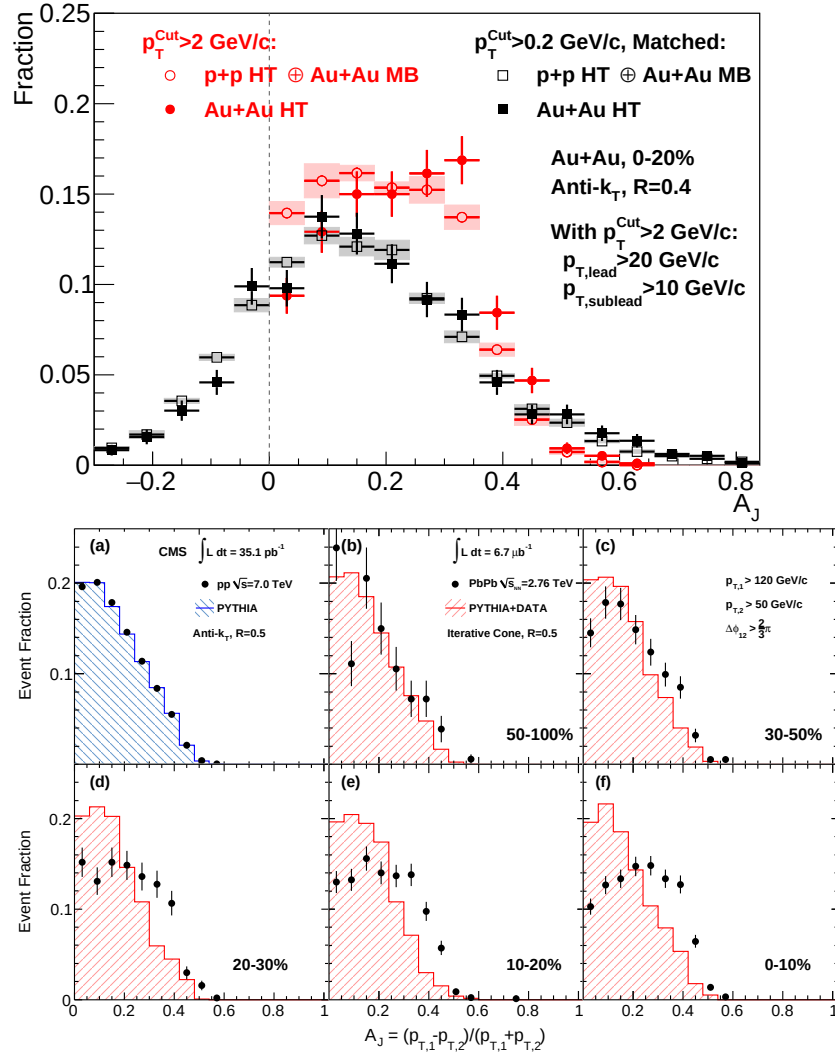


Obrázek 2.10: Jaderný modifikační faktor R_{AA} pro jety s rozlišovacím parametrem $R = 0,4$ naměřený experimentem ATLAS na LHC v Pb+Pb srážkách při $\sqrt{s_{NN}} = 2,76$ TeV pro různé intervaly rapidit v závislosti na p_T jetu. Převzato z [12].

V případě dvoujetových událostí lze potlačení kvantifikovat pomocí proměnné A_J , která je definována jako

$$A_J = \frac{p_{T,1} - p_{T,2}}{p_{T,1} + p_{T,2}}, \quad (2.14)$$

kde $p_{T,1}$ je hybnost jetu s nejvyšším p_T (tzv. leading jet) a $p_{T,2}$ je hybnost druhého jetu (tzv. sub-leading jet). V případě že $A_J \approx 0$, jsou pozorované jety málo ovlivněny médiem. Příklad $A_J \approx 1$ bude zase ukazovat na to, že se jeden z jetů prakticky zcela pohltí médiem. Příklad takových měření lze vidět na Obr. 2.11, na kterých je zobrazeno pravděpodobnostní rozdělení pro veličinu A_J . V těchto měřeních se porovnávají p+p srážky vložené do jádro-jaderného pozadí (p+p HT $\oplus$ Au+Au MB nahoře, či PYTHIA+DATA dole, kde PYTHIA představuje simulované srážky). V případě horního obrázku, při selekci konstituentů $p_T^{\text{Cut}} > 2$ GeV/c lze vidět, že se rozdělení A_J pro Au+Au HT data liší od p+p HT $\oplus$ Au+Au MB, pro omezení $p_T^{\text{Cut}} > 0,2$ GeV/c ovšem rozdíl mizí. Obdobně, v dolním obrázku, lze vidět že pro periferní srážky jsou naměřená data podobná PYTHIA+DATA. Se zvyšující se centralitou se ovšem naměřené rozdělení A_J mění. Toto může napovídat o tom, že se zvyšující se centralitou srážky roste interakce jetů s médiem.



Obrázek 2.11: Míra dvoujetové asymetrie změřená experimentem STAR pro Au+Au srážky při $\sqrt{s_{NN}} = 200$ GeV. Převzato z [13] (nahore). Analogické měření provedené experimentem CMS pro Pb+Pb srážky při $\sqrt{s_{NN}} = 2,76$ TeV (dole). Převzato z [14].

Kapitola 3

Jetové algoritmy

Výsledek částicových srážek, který je výstupem z experimentu, má nejčastěji složitou datovou strukturu založenou na tzv. eventech, což jsou digitalizované záznamy srážky, které obsahují informace o částicích (jejich počet, hybnost atd.), informace o detektorech, které byly aktivní atd. Pro samotné studium jetů je tedy nutné tyto jety nejprve zrekonstruovat. K tomuto účelu slouží algoritmy pro jetovou rekonstrukci, které prohledávají $\eta - \varphi$ prostor detektoru, kde η je pseudorapidita a φ je azimutální úhel částic a v něm rekonstruují jety. Dobře definovaný jetový algoritmus by měl splňovat následující podmínky:

- **Kolineární bezpečnost** - v případě, že se částice rozdělí vlivem rozpadu na dvě částice, které budou vůči sobě kolineární, neměl by se výsledný počet jetů změnit.
- **Infračervená bezpečnost** - jetový algoritmus musí být invariantní vůči částicím, které vznikají jako důsledek vyzařování gluonů s velmi nízkou hybností. Zase platí, že případné vyzařování měkkého gluonu by nemělo změnit počet jetů ve srážce.
- **Nezávislost na úrovni popisu** - algoritmus by neměl záviset na tom, zda-li popisujeme jet na partonové, hadronové či detektorové úrovni.

3.1 Sekvenčně-rekombinační algoritmy

Dále se zaměříme na třídu sekvenčně-rekombinačních algoritmů. Ty jsou v dnešní době nejpoužívanější, jelikož jsou infračerveně a kolineárně bezpečné. Tyto algoritmy fungují na principu, že zvolí počáteční částici a k ní pak iterativně přidávají jiné částice, které jsou blízké v $\eta - \varphi$ prostoru. V této třídě algoritmů jsou nejvýznamnějšími představiteli algoritmy k_t [32], anti- k_t [15] a Cambridge-Aachen [33, 34]. Tyto algoritmy lze generalizovat do následující podoby:

1. Pro množinu částic s indexem i , příčnou hybností $p_{T,i}$, a pozicí (η_i, φ_i) se spočte vzdálenost od svazku $d_i = p_{T,i}^{2k}$.

2. Pro každý pár částic s indexem i a j se spočte vzdálenost

$$d_{i,j} = \min\{p_{T,i}^{2k}, p_{T,j}^{2k}\} \frac{(\eta_i - \eta_j)^2 - (\varphi_i - \varphi_j)^2}{R^2}, \quad (3.1)$$

kde R je předem definované rozlišení algoritmu určující, kolik hybnosti z počátečního partonu dokážeme získat. Vyšší hodnoty R umožňují lepší rekonstrukci původní hybnosti, ovšem v tomto případě je rekonstruovaný jet více citlivý na přítomnost měkkého pozadí.

3. Nalezneme $d_{min} = \min\{d_{i,j}, d_i\}$.

4. Je-li $d_{min} = d_i$, pak je i přidáno do množiny finálních jetů. Pokud je $d_{min} = d_{i,j}$, pak jsou objekty i a j spojeny a postup opakován.

Volbou parametru k vznikají následující verze sekvenčně-rekombinačních algoritmů:

- pro $k = 1$ se jedná o k_t algoritmus,
- volba $k = 0$ odpovídá Cambridge-Aachen algoritmu,
- pro $k = -1$ hovoříme o anti- k_t algoritmu.

k_t algoritmus

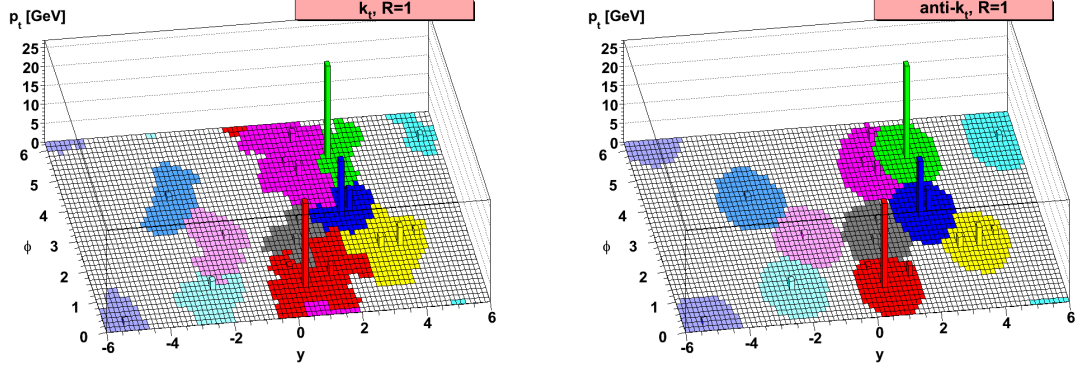
Algoritmus k_t začíná rekombinaci částicemi s nízkou příčnou hybností. Z toho plyne i nevýhoda tohoto algoritmu, jelikož je citlivý na přítomnost měkkého pozadí, nepocházejícího z procesů s vysokými hodnotami Q^2 , které je přítomné v jádro-jaderných srážkách. Toto má ovšem i využití při určení hustoty pozadí, která se používá pro korekci na jetové pozadí.

Anti- k_t algoritmus

Algoritmus anti- k_t je přesným opakem k_t algoritmu, kdy rekombinace částic do jetu začíná částicemi s vysokou příčnou hybností. V rámci jádro-jaderných srážek se tento algoritmus používá pro rekonstrukci primárních jetů, jelikož anti- k_t je méně citlivý na přítomnost měkkého pozadí než k_t .

Cambridge-Aachen algoritmus

Algoritmus Cambridge-Aachen ve své rekombinaci nezahrnuje příčné hybnosti, nýbrž jen prostorové rozložení částic v předpokládaném jetu. Tento algoritmus se dnes používá zejména pro analýzu jetové substruktury.



Obrázek 3.1: Vliv jetového algoritmu na výsledný tvar jetů a jejich plochy. Vlevo lze vidět příklad pro k_t algoritmus s $R = 1$, vpravo pro $\text{anti-}k_t$ s $R = 1$. Převzato z [15].

3.1.1 Jetové pozadí

Po jetové rekonstrukci jsou výstupem objekty, které byly algoritmem určeny jako jety. Tyto objekty ovšem nemusí být jety ve fyzikálním smyslu. Tyto tzv. pseudojety se filtrují zavedením omezující podmínky na jejich p_T . Je-li $p_T > p_T^{\text{min}}$, pak je pseudojet považován za jet. V případě silných pozadí je dalším krokem odečtení tohoto pozadí. Toto je potřeba, jelikož cílem je získat jet co nejbližší teoretickému popisu - tj. pocházející z procesu s vysokým Q^2 . Silné pozadí ovšem může do jetu přidat částice pocházející z měkkých procesů. Prvním krokem je odečet střední hodnoty energie pozadí, který spočívá v tom, že do srážky je náhodně přidán velký počet měkkých částic s následnou rekonstrukcí pomocí algoritmu k_t . Díky tomu lze určit tzv. plochu jetu A , která je úměrná počtu měkkých částic, které v něm jsou nalezeny. Celková střední energie pozadí se spočte jako

$$\rho = \text{med} \left\{ \frac{p_{T,j}}{A_j} \right\}, \quad (3.2)$$

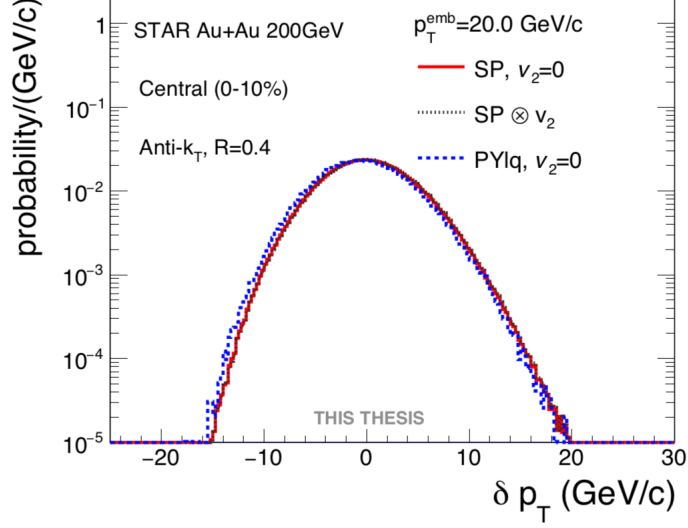
kde j prochází přes všechny jety v eventu. Korigovaná hodnota příčné hybnosti jetu i ($p_{T,i}^{\text{corr}}$) na střední hodnotu pozadí je pak určena jako

$$p_{T,i}^{\text{corr}} = p_{T,i} - \rho A_i. \quad (3.3)$$

3.2 Dekonvoluce fluktuujícího pozadí

Silné pozadí vznikající v jádro-jaderných srážkách značným způsobem ovlivňuje naměřená jetová spektra, která jsou konvolucí skutečného jetového spektra $p_{T,jet}^{\text{true}}$ a fluktuací pozadí. Pro získání skutečného jetového spektra z naměřených dat je potřeba provést operaci dekonvoluce, která odstraní efekty pozadí z naměřených dat. Necht pro naměřené distribuce platí vztah

$$R[p_{T,jet}^{\text{true}} \rightarrow p_{T,jet}^{\text{meas}}] p_{T,jet}^{\text{true}} = p_{T,jet}^{\text{meas}}, \quad (3.4)$$



Obrázek 3.2: Rozdělení δp_T pro $p_T^{emb} = 20,0$ GeV/c. Červené a šedé spektrum představuje jednočásticový jet bez a s vlivem kolektivních jevů. Modré spektrum představuje jety generované modelem Pythia. Převzato z [16].

kde $\mathbf{p}_{T,jet}^{meas}$ a $\mathbf{p}_{T,jet}^{true}$ představují vektory, jejichž prvky jsou dány biny jednotlivých p_T distribucí a $\mathbf{R}[p_{T,jet}^{true} \rightarrow p_{T,jet}^{meas}]$ je matice odezvy. Cílem je pak invertovat matici $\mathbf{R}[p_{T,jet}^{true} \rightarrow p_{T,jet}^{meas}]$ tak, aby byla získána pravá distribuce $\mathbf{p}_{T,jet}^{true}$. Tato úloha je velmi náročná, jelikož matice může být singulární, být špatně podmíněná atd. Nejprve je ale potřeba tuto matici sestavit, což se provede pomocí simulací, kdy se do jádro-jaderné srážky vloží simulovaný jet, celá srážka se rekonstruuje a pak se měří rozdíl δp_T který je daný jako

$$\delta p_T = p_{T,jet}^{rec} - \rho A - p_{T,jet}^{emb}, \quad (3.5)$$

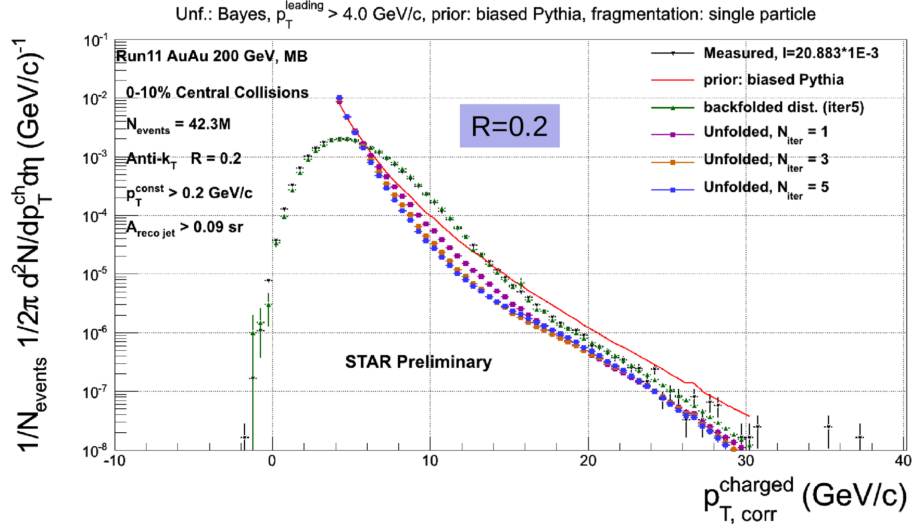
kde $p_{T,jet}^{rec}$ značí příčnou hybnost rekonstruovaného jetu a $p_{T,jet}^{emb}$ příčnou hybnost vloženého jetu. Rozdělení δp_T je následně normalizováno na pravděpodobnostní rozdělení a poté se matice sestaví tak, že se pro různé biny v $p_{T,jet}^{meas}$ na ose x zaplní odpovídající hodnoty δp_T pro určené hodnoty $p_{T,jet}^{emb}$ s tím, že $\delta p_T = 0$ je vždy na diagonále matice. Příklad rozdělení δp_T pro jety s $R = 0,4$ a $p_T^{emb} = 20,0$ GeV/c lze vidět na Obr. 3.2.

Jednou z metod, pomocí které lze provést dekonvoluci, je iterativní Bayesovská dekonvoluce [35], jejíž výsledek lze vidět na Obr. 3.3. Ta spočívá v několika následujících krocích:

1. Zvolí se apriorní distribuce pro p_T spektrum, ozn. $P_0(p_{T,j}^{true})$, kde $p_{T,j}$ je hybnost v j-tém binu.
2. Z Bayesovy věty se spočítá aposteriorní distribuce

$$P(p_{T,j}^{true} | p_{T,i}^{meas}) = \frac{P(p_{T,i}^{meas} | p_{T,j}^{true}) P_0(p_{T,j}^{true})}{\sum_{l=1}^{n_t} P(p_{T,i}^{meas} | p_{T,l}^{true}) P_0(p_{T,l}^{true})}, \quad (3.6)$$

kde n_t je počet binů.



Obrázek 3.3: Příklad jetových spekter v Au+Au srážce při energii $\sqrt{s_{NN}} = 200$ GeV před dekonvolucí (černé body) a po (fialové, červené a modré body). Převzato z [17]

3. Proveďte se převáhování jednotlivých binů jako

$$n(p_{T,j}^{true}) = n(p_{T,i}^{meas})P(p_{T,j}^{true}|p_{T,i}^{meas}), \quad (3.7)$$

kde $n(\cdot)$ je počet prvků v jednotlivém binu.

4. Zvolí se nová apriorní distribuce

$$P_0(p_{T,j}^{true}) = \frac{n(p_{T,j}^{true})}{N_{true}}, \quad (3.8)$$

kde N_{true} značí odhad na počet jetů v distribuci $\mathbf{p}_{T,jet}^{true}$ a zopakuje se iterace.

Další metodou, která se často využívá pro dekonvoluci, je metoda SVD (Singular Value Decomposition) rozkladu [36]. Ta spočívá v tom, že pro libovolnou matici $\mathbf{A} \in \mathbb{R}^{m \times n}$ existuje rozklad ve tvaru

$$\mathbf{A} = \mathbf{U} \mathbf{S} \mathbf{V}^T, \quad (3.9)$$

kde $\mathbf{U} \in \mathbb{R}^{m \times m}$ a $\mathbf{V} \in \mathbb{R}^{n \times n}$ jsou ortogonální matice a $\mathbf{S} \in \mathbb{R}^{m \times n}$ je diagonální matice s nenulovými diagonálními prvky. Použije-li se SVD rozklad pro matici $\mathbf{R}[p_{T,jet}^{true} \rightarrow p_{T,jet}^{meas}]$, přejde rovnice (3.4) na tvar

$$\mathbf{U} \mathbf{S} \mathbf{V}^T \mathbf{p}_{T,jet}^{true} = \mathbf{p}_{T,jet}^{meas}. \quad (3.10)$$

Z ortogonalit matice $\mathbf{U}$ pak plyne

$$\mathbf{S} \mathbf{V}^T \mathbf{p}_{T,jet}^{true} = \mathbf{U}^T \mathbf{p}_{T,jet}^{meas}. \quad (3.11)$$

Označí-li se $\mathbf{d} = \mathbf{U}^T \mathbf{p}_{T,jet}^{meas}$ a $\mathbf{z} = \mathbf{V}^T \mathbf{p}_{T,jet}^{true}$, pak má soustava tvar

$$\mathbf{S} \mathbf{z} = \mathbf{d}. \quad (3.12)$$

Inverze matice $\mathbf{S}$ je triviální, jelikož se jedná o diagonální matici a řešení pro vektor $\mathbf{z}$ je

$$\mathbf{z} = \mathbf{S}^{-1}\mathbf{d}. \quad (3.13)$$

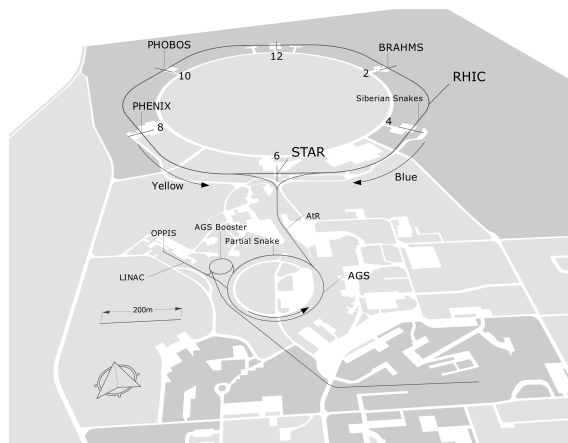
Využije-li se toho, že $\mathbf{z} = \mathbf{V}^T \mathbf{p}_{T,jet}^{true}$ a $\mathbf{V}^T$ je ortogonální matice, má celkové řešení tvar

$$\mathbf{p}_{T,jet}^{true} = \mathbf{V} \mathbf{S}^{-1} \mathbf{d}. \quad (3.14)$$

Kapitola 4

Urychlovač RHIC a experiment STAR

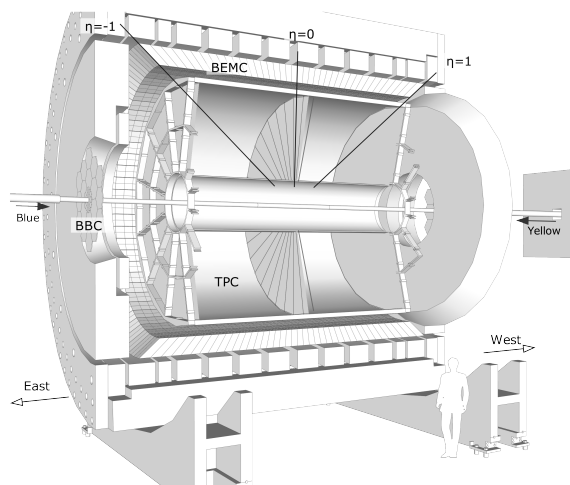
Urychlovač RHIC nacházející se v Brookhavenské Národní Laboratoři (BNL) je urychlovač nabitých iontů a protonů na vstřícných svazcích, jehož primárním cílem je studium dynamiky jaderné hmoty při vysokých energiích a spinové struktury protonu. Jeho předností je možnost srážet různorodé jádro-jaderné systémy (např. Au+Au, U+U, Cu+Cu, ale i asymetrické jako Au+Cu, d+Au) při variabilních energiích srážky v těžištové soustavě $\sqrt{s_{NN}}$ od 7,7 do 200 GeV na nukleonový pár. Je také jediným urychlovačem na světě, který dokáže srážet spinově polarizované protony. Na začátku byly na urychlovači přítomny 4 experimenty - BRAHMS [37], PHOBOS [38], PHENIX [39] a STAR [40]. V dnešní době je však aktivní pouze experiment STAR a v přípravě se nachází nový experiment sPHENIX, který nahradí experiment PHENIX.



Obrázek 4.1: Schéma urychlovače RHIC. Převzato z [18].

4.1 Experiment STAR

Následující část práce je zaměřená experiment STAR a budou proto blíže specifikovány informace o některých jeho detektorech, které lze vidět na Obr. 4.2. Celý experiment je obklopen solenoidem o magnetické indukci $B = 0,5 \text{ T}$, který pomáhá identifikovat náboj a hybnost částic v detektoru TPC.



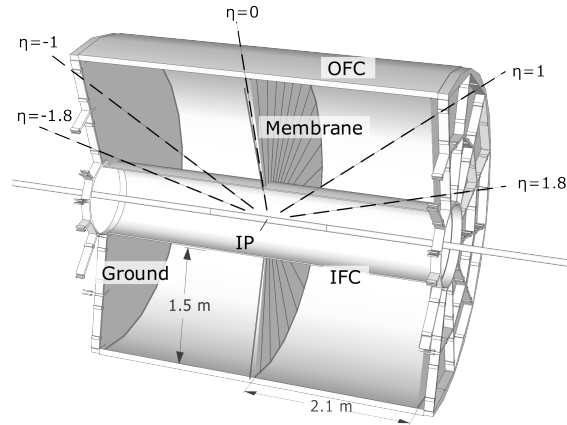
Obrázek 4.2: Řez experimentem STAR. Převzato z [18].

4.1.1 TPC

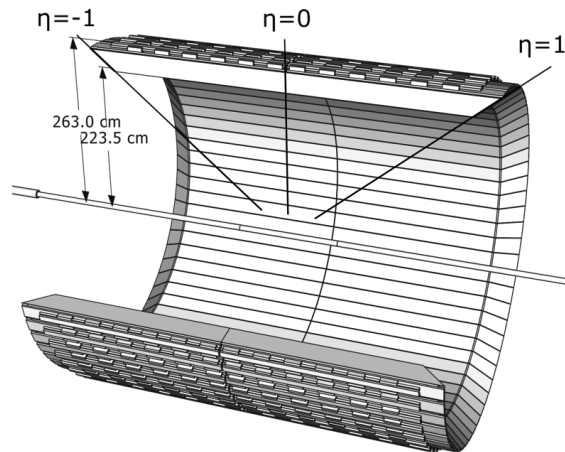
Časově projekční komora (TPC) [41] je plynový detektor, který slouží k identifikaci nabitých částic. Jeho konstrukci lze vidět na Obr. 4.3. Celý detektor je udržován pod vysokým napětím a je rozdělený na dva sektory. Během průchodů vytvoří částice v TPC elektron-iontové páry. Ionty budou driftovat k centrální katodě, elektrony ke koncům detektoru, které slouží jako anody. Díky tomu, že konce TPC jsou rozděleny do sektorů, lze jednotlivým zásahům přiřadit přesnou polohu v polárních souřadnicích, které se následně přepočítají do x a y složek v kartézské soustavě. Složka z se spočítá s pomocí času, který proběhne od spuštění záznamu srážky do zaznamenání zásahu dané částice na konci detektoru TPC. Identifikace nabitých částic probíhá na základě energetických ztrát částice při průchodu plynem. Lehké částice ztrácejí více energie a tedy produkují více elektron-iontových párů.

4.1.2 BEMC

Barelový elektromagnetický kalorimetr (BEMC) [42] je dalším detektorem, který je součástí experimentu STAR. Jeho průřez lze vidět na Obr. 4.4. Hraje velmi důležitou roli při měřeních, která se týkají jetové fyziky, jelikož umožňuje detekovat neutrální částice, které jsou v těchto jetech přítomné. Funguje na principu, že částice (nabitá i nenabitá) vletí do oloveného absorbátoru, ve kterém vznikne elektromagnetická sprška. Ta následně vletí do scintilátoru (experiment STAR se používá kombinaci

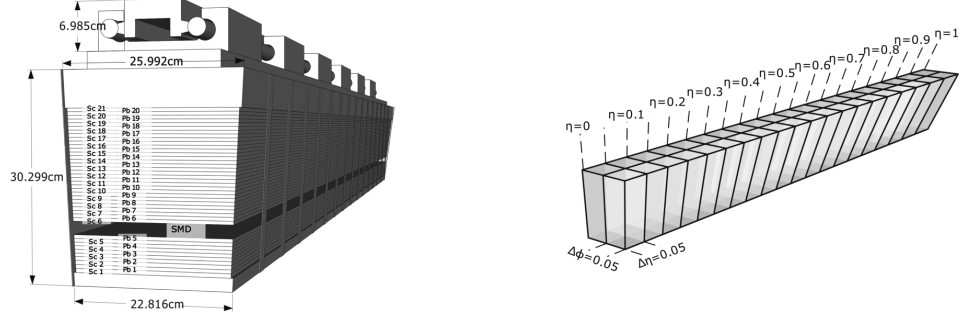


Obrázek 4.3: Časově projekční komora (TPC) experimentu STAR. Detektor pokrývá rozsah pseudorapiditě od -1 do 1. Převzato z [18].



Obrázek 4.4: Barelový elektromagnetický kalorimetr (BEMC) experimentu STAR. Analogicky jako TPC i tento detektor má pokrytí od -1 do 1 v pseudorapiditě. Převzato z [18].

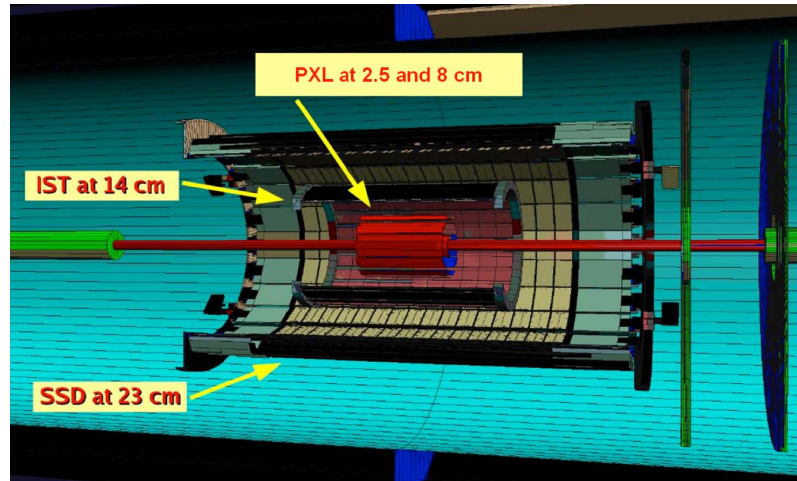
plastových a olověných scintilátorů), které se následně vyčtou. Energie, která je deponována částicí, je pak úměrná intenzitě vyčteného signálu. Detektor je tvořen jednotlivými věžemi (tzv. "tower"), kde každá věž má rozměr $\Delta\varphi = 0,05$ a $\Delta\eta = 0,05$. Příklad BEMC modulu lze vidět na Obr. 4.5 vlevo, dělení na jednotlivé věže je na Obr. 4.5 vpravo.



Obrázek 4.5: (vlevo) Modul detektoru BEMC. (vpravo) Dělení modulu BEMC na jednotlivé věže. Převzato z [18].

4.1.3 HFT

Posledním detektorem, který může mít využití pro měření týkající se jetové fyziky, je tracker těžkých vůní (Heavy Flavour Tracker, HFT) [43]. Jedná se o vícevrstvý křemíkový detektor, jehož schematický náčrtek kterého lze vidět na Obr. 4.6. Zde PXL představuje detektor PiXeL [44], který je založen na technologii MAPS, IST je vnitřní křemíkový tracker a SSD [45] je křemíkový stripový detektor. Díky své konstrukci a použitým technologiím umožňuje HFT detekovat sekundární vertexy těžkých částic (obsahujících c a b kvarky) s vysokou přesností ($\approx 40\mu m$).



Obrázek 4.6: Průřez detektorem HFT. Převzato z [19].

V rámci jetové fyziky je tento detektor zajímavý pro měření jetů, které vznikly v následku fragmentace c nebo b kvarku.

Kapitola 5

Strojové učení

Strojové učení [46, 47] je vědní obor nacházející se na pomezí aplikované matematiky, informatiky a statistiky. Ústřední úlohu v tomto oboru hrají algoritmy, které se učí z dat. Díky tomu dokážou řešit úlohy, které by se těžko řešily pomocí algoritmů založených na pevně daných pravidlech.

5.1 Základní pojmy strojového učení

Před dalším výkladem o strojovém učení je třeba zavést několik základních pojmů a notací užívaných v tomto textu. V dalším textu bude pomocí $\mathcal{X}$ značená množina tzv. příznakových vektorů a $\mathcal{Y}$ bude značit jim příslušnou množinu značek. Případ, kdy jsou přítomny obě dvě množiny a cílem je se naučit zobrazení $f : \mathcal{X} \rightarrow \mathcal{Y}$, se nazývá učení s učitelem. V tomto případě je uspořádaná dvojice $\{(\mathbf{x}_i, y_i)\}_{i=1}^N \in \mathcal{X} \times \mathcal{Y}$ tzv. trénovací množina.

Úloha s učitelem se nazývá regrese, pokud bude snaha se naučit zobrazení $f : \mathcal{X} \rightarrow \mathcal{Y} \subseteq \mathbb{R}^n$. Úloha je klasifikací, pokud hledaným zobrazením je $f : \mathcal{X} \rightarrow \mathcal{Y} \subseteq \{0, 1\}^M$. Zde M je počet klasifikačních tříd. Tato úloha je velmi často modifikovaná a namísto 0-1 klasifikace se provádí určení pravděpodobnosti, že daný příklad náleží dané třídě. Hledaným zobrazením je pak zobrazení $f : \mathcal{X} \rightarrow [0, 1]^M$.

Je nutné poznamenat, že existuje nespočet jiných učících přístupů, kde některé z nich např. leží na pomezí učení s učitelem a učení bez učitele.

5.2 Formulace učícího problému s učitelem

Jelikož je učení s učitelem stěžejní pro tuto práci, následující text bude zaměřen na jeho podrobnou formulaci. Ostatní učící případy (bez učitele atd.) budou podobné níže uvedenému konstruktu ovšem s některými úpravami. Nechť jsou dány předpoklady:

- Množina příznakových vektorů $\mathcal{X}$ se řídí pravděpodobnostním rozdělením $p(\mathbf{x})$.

- Množina značek $\mathcal{Y}$, pro které existuje (neznáme) pravděpodobnostní rozdělení $p(y|\mathbf{x}) : \mathcal{X} \rightarrow \mathcal{Y}$.
- Trénovací množina $\{(\mathbf{x}_i, y_i)\}_{i=1}^N \in \mathcal{X} \times \mathcal{Y}$ je pak generována rozdělením $p(\mathbf{x}, y)$, u kterého existuje předpoklad nezávislosti a stejného rozdělení prvků (**iid**).
- Je zvolen parametrický model $f(\mathbf{x}; \boldsymbol{\theta})$.
- Je zvolená ztrátová funkce $L(f(\mathbf{x}; \boldsymbol{\theta}), y)$, která bude kvantifikovat výkon našeho modelu.

Cílem je pak najít takovou parametrizaci modelu, která minimalizuje ztrátovou funkci, tj.

$$\boldsymbol{\theta}^* = \arg \min_{\boldsymbol{\theta}} \mathbb{E}_{\mathbf{x}, y \sim p(\mathbf{x}, y)} L(f(\mathbf{x}; \boldsymbol{\theta}), y). \quad (5.1)$$

V rámci hlubokého učení je nejčastější volbou ztrátové funkce křížová entropie mezi dvěma distribucemi $p_{model}(y|\mathbf{x}; \boldsymbol{\theta})$ a $p(\mathbf{x}, y)$. Úloha je pak formulována jako

$$\boldsymbol{\theta}^* = \arg \min_{\boldsymbol{\theta}} \mathbb{E}_{\mathbf{x}, y \sim p(\mathbf{x}, y)} \ln p_{model}(y|\mathbf{x}; \boldsymbol{\theta}) \quad (5.2)$$

a model musí být schopen reprezentovat pravděpodobnostní rozdělení. V případě regrese se nejčastěji předpokládá, že $p_{model}(y|\mathbf{x})$ je Gausovskou distribucí s průměrem y a jednotkovou kovarianční maticí, v případě klasifikace o N třídách pak platí předpoklad tzv. multinomického rozdělení.

5.3 Řešení učícího problému

Přesná podoba modelu bude blíže specifikována později a tato část se bude zabývat řešením optimalizačního problému (5.2). Základní metodou ve strojovém učení je gradientní sestup [48, 49], který je dán jako

$$\boldsymbol{\theta}^{(t+1)} = \boldsymbol{\theta}^{(t)} - \lambda \nabla \mathbb{E}_{\mathbf{x}, y \sim p(\mathbf{x}, y)} \ln p_{model}(y|\mathbf{x}; \boldsymbol{\theta}), \quad (5.3)$$

kde λ je parametr řídící intenzitu optimalizace, kterému se ve strojovém učení říká míra učení. V praxi však není podoba pravého pravděpodobnostního rozdělení $p(\mathbf{x}, y)$ známá, jelikož existuje jen konečný počet bodů z dané trénovací množiny. Proto se přechází k minimalizaci empirického risku ve tvaru

$$\boldsymbol{\theta}^{(t+1)} = \boldsymbol{\theta}^{(t)} - \lambda \frac{1}{N} \sum_{i=1}^N \nabla \ln p_{model}(y^{(i)}|\mathbf{x}^{(i)}; \boldsymbol{\theta}). \quad (5.4)$$

Tento přístup ovšem selhává u velkých trénovacích množin, jelikož vyžaduje vyčíslení gradientu přes všechny její prvky, což je výpočetně velmi náročné. Proto se zavádí stochastický gradientní sestup [50, 51, 52], který nejprve vybere M náhodných prvků z trénovací množiny a pak provede optimalizační krok analogicky jako v předchozím případě, tj.

$$\boldsymbol{\theta}^{(t+1)} = \boldsymbol{\theta}^{(t)} - \lambda \frac{1}{M} \sum_{i=1}^M \nabla \ln p_{model}(y^{(i)}|\mathbf{x}^{(i)}; \boldsymbol{\theta}) \quad (5.5)$$

V tomto případě ovšem není použitý gradient v pravém slova smyslu, nýbrž jeho statistická aproximace. Existují i pokročilejší varianty stochastického gradientního sestupu, např. stochastický gradientní sestup s hybností, kde jednotlivé iterace jsou dány jako

$$\mathbf{v}^{(t+1)} = \alpha \mathbf{v}^{(t)} - \lambda \nabla \mathbb{E}_{\mathbf{x}, y \sim p(\mathbf{x}, y)} \ln p_{\text{model}}(y | \mathbf{x}; \boldsymbol{\theta}), \quad (5.6)$$

$$\boldsymbol{\theta}^{(t+1)} = \boldsymbol{\theta}^{(t)} - \mathbf{v}^{(t+1)}, \quad (5.7)$$

kde α udává sílu příspěvku "hybnostního" členu. Existují ovšem i jiné, mnohem sofistikovanější metody, např. Adam [53], AdaDelta [54], které se snaží přiblížit metodám druhého řádu. Podrobněji se o tom lze dočíst např. v 8. kapitole v [47]. Parametr míry učení λ je jedním z nejdůležitějších parametrů v optimalizační části učícího problému. Bohužel dodnes nejsou známy přesné postupy, jak ho správně zvolit. Existuje ovšem velký počet heuristik, jak tento parametr zvolit, např. volba takové největší počáteční hodnoty λ , při které optimalizační problém ještě nediverguje. S průběhem trénování je taktéž potřeba měnit hodnoty λ . Lze tak učinit např. vydělením počáteční hodnoty jistým faktorem po splnění daných podmínek (např. přesnost modelu přestala růst), předdefinovaným schématem (např. lineární snížení k určité hodnotě) nebo i cyklicky [55].

5.4 Neuronové sítě

Dále bude rozebrána problematika neuronových sítí, jelikož jsou stěžejní součástí této práce. Ve své podstatě je dopředná neuronová síť složenou parametrickou funkcí

$$F(\mathbf{x}; \mathbf{W}, \mathbf{b}) = f_n(\mathbf{h}_{n-1}; \mathbf{W}_n, \mathbf{b}_n) \circ \dots \circ f_1(\mathbf{x}; \mathbf{W}_1, \mathbf{b}_1), \quad (5.8)$$

kde f_i je tzv. vrstva. Každá tato vrstva koná lineární operaci (konvoluce nebo afinní transformace), kterou následuje nelinearita. Dnes je nejčastější volbou nelinearity funkce ReLU, která má tvar

$$\text{ReLU}(x) = \max(0, x). \quad (5.9)$$

Trénování neuronové sítě probíhá pomocí algoritmu zpětné propagace [56, 57, 58], kdy se počítá derivace ztrátové funkce vzhledem ke všem parametrům v neuronové síti a pro výpočet gradientu v n -té vrstvě se používá gradient z předešlých vrstev.

Důležitou podmínkou pro dobrou konvergenci algoritmu je správná inicializace parametrů. Existuje mnoho metod inicializace, zde jsou však nastíněny tři nejpoužívanější. První metoda, tzv. Xavier inicializace [59], inicializuje jednotlivé parametry $\mathbf{W}_i$ vrstvy pomocí uniformní distribuce $\mathcal{U}(-a, a)$ s volbou $a = \sqrt{\frac{6}{m+n}}$, kde m je počet vstupů do vrstvy a n je počet výstupů z vrstvy. Analogicky lze inicializovat z Gausovské distribuce $\mathcal{N}(0, a)$, kde $a = \sqrt{\frac{2}{m+n}}$. Modifikací parametru a jako $a = \sqrt{\frac{6}{m}}$ pro uniformní distribuci či $a = \sqrt{\frac{2}{m}}$ pro normální distribuci dostaneme tzv. Kaiming inicializaci [60]. Poslední populární metodou je inicializace parametrů pomocí náhodných ortogonálních matic [61]. Volba vhodné inicializace pro hluboké neuronové sítě je velmi aktivním zdrojem výzkumu a neexistuje spolehlivá metoda, jak ji zvolit. Je tedy nutné volit inicializaci na základě empirických experimentů na zvolené množině dat.

5.4.1 Konvoluční neuronová síť jako extraktor příznaků

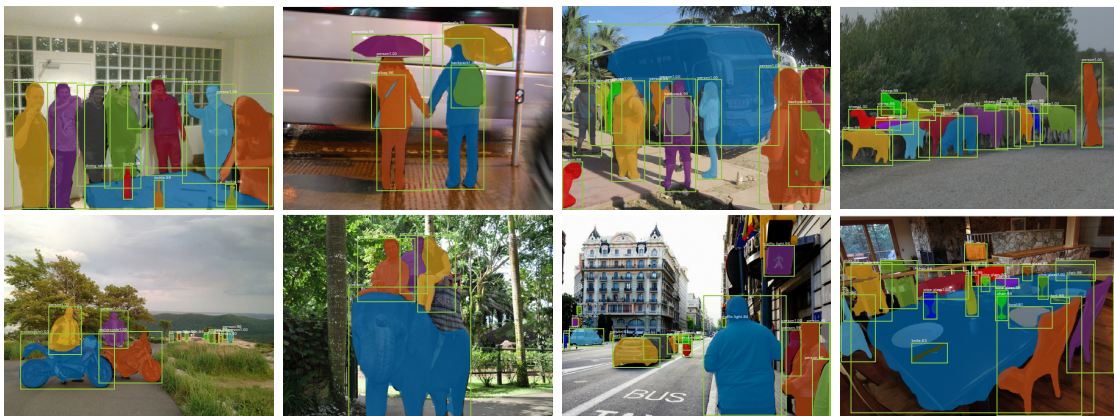
Jednou z velmi důležitých vlastností konvolučních neuronových sítí, na kterých jsou založeny další aplikace, je extrakce příznaků. Konvoluční neuronová síť může být přepsaná jako

$$F(\mathbf{x}; \mathbf{W}, \mathbf{b}) = f_{\text{head}}(\mathbf{h}_e; \mathbf{W}_h, \mathbf{b}_h) \circ f_{\text{embedding}}(\mathbf{x}; \mathbf{W}_e, \mathbf{b}_e), \quad (5.10)$$

kde f_{head} je výstupní část, která může být tvořená jak jednoduchým lineárním modelem, tak dopřednou neuronovou sítí a která provádí regresní či klasifikační úlohu. Tato část bere jako vstup vektory z $f_{\text{embedding}}$, které mají oproti vstupním datům nižší dimenzionalitu a vyšší informační obsah. Na základě této vlastnosti funguje algoritmus Mask R-CNN [20], který byl také použit v této práci.

5.5 Mask R-CNN

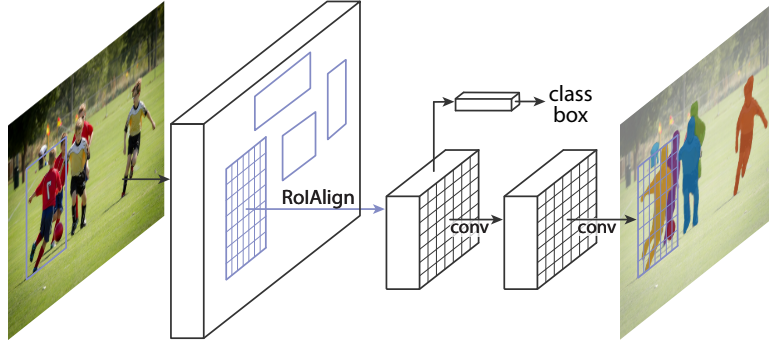
V této práci je dále užíván algoritmus Mask R-CNN pro účely detekce jetů v $\eta-\varphi$ rovině. Dále se tedy práce věnuje podrobnému popisu tohoto algoritmu. Mask R-CNN je algoritmem který provádí detekci objektů a instanční segmentaci na vstupních obrazech. Příklady detekcí lze vidět na Obr. 5.1.



Obrázek 5.1: Příklady detekcí pomocí algoritmu Mask R-CNN. Převzato z [20].

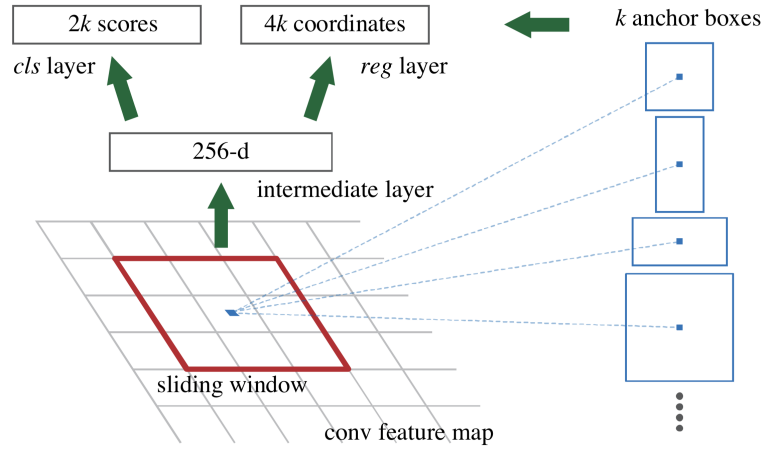
Na rozdíl od klasifikace, kde algoritmus očekává obrázek a k němu příslušející značku, očekává během trénování Mask R-CNN navíc rámeček a pixelovou masku kolem objektů k detekci. Toto v praxi komplikuje přípravu dat pro tento algoritmus, jelikož je potřeba nejen vyznačit co za objekt je na obrázku, ale je také potřeba vyznačit jeho přesnou polohu a okraje.

Dopředný směr fungování algoritmu lze vidět na Obr. 5.2. Obrázek se tedy nejprve zpropaguje přes konvoluční neuronovou síť, která funguje jako extraktor příznaků. Poté je vytvořena příznaková mapa. Na ní pak musíme najít oblasti zájmu (Region of Interest, RoI), které budou použity v *RoIAlign* vrstvě pro extrakci fixních kousků příznakové mapy s obsahem objektů, které jsou určeny k detekci. Model Mask R-CNN vychází z modelu Faster R-CNN [21], který tento problém řeší tím, že po celém



Obrázek 5.2: Architektura Mask R-CNN. Převzato z [20].

obrázku nejprve rozdistribuuje tzv. kotvy, které mají sloužit jako možní kandidáti na RoI. Příklad toho, jak takový přístup funguje, lze vidět na Obr. 5.3.



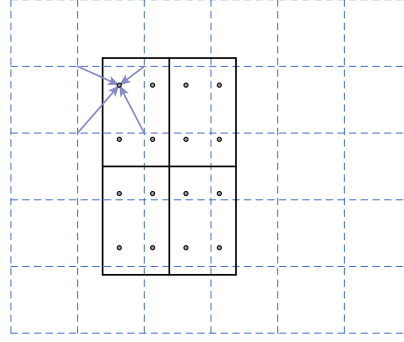
Obrázek 5.3: Příklad kotev. Převzato z [21].

Tito kandidáti jsou pak zpracováni pomocí Region Proposal Network (RPN), která má dva úkoly - určit, zda-li je v oblasti vyznačené kotvou přítomen objekt zájmu, a dále korigovat polohu jednotlivých kotev tak, aby lépe popisovaly skutečnou polohu objektu. K jejímu trénování se užívají rámečky, které jsou v trénovací množině. Optimalizace probíhá přes ztrátovou funkci ve tvaru

$$L_{RPN}(p, t, p^*, t^*) = \frac{1}{N_{cls}} L_{cls}(p, p^*) + \frac{\lambda}{N_{reg}} p^* L_{reg}(t, t^*), \quad (5.11)$$

kde p je předpovězená pravděpodobnost, že RoI obsahuje objekt, $p^* \in \{0, 1\}$ určuje, zda-li RoI obsahuje objekt, t je předpovězený vektor rámečků a t^* je správný vektor rámečků. Faktor p^* se zasluhuje o to, aby optimalizace predikce rámečků probíhala pouze v případě pozitivních objektů. Člen L_{cls} odpovídá za přesnost klasifikace objektů zájmu a člen L_{reg} řídí přesnost korekce kotev.

Poté jsou RoI extrahovány pomocí operace **RoIAlign**, která z příznakové mapy extrahuje RoI a následně ho rozdělí na fixní velikost (např. 7×7 binů) s tím, že použije bilineární interpolaci pro dopočet hodnot mezi pixely. Fungování této operace lze vidět na Obr. 5.4.



Obrázek 5.4: Operace RoIAlign. Převzato z [20].

Následně se z této mřížky provede agregace (např. zvolením největšího prvku v binu), čímž se získá fixní tenzor o velikosti $7 \times 7 \times C$, kde C je počet kanálů v poslední konvoluční vrstvě. Tento tenzor se pak dále propaguje do výstupních vrstev - čtyřrozměrné lineární regrese pro detekci rámečků, softmax klasifikátor pro detekci tříd a malou konvoluční síť, která emituje pixelové masky kolem detekovaných objektů. Trénování této části modelu se provádí přes ztrátovou funkci

$$L_{heads}(p, t^u, m, u, v, w) = L_{cls}(p, u) + L_{loc}(t^u, v) + L_{mask}(m, w), \quad (5.12)$$

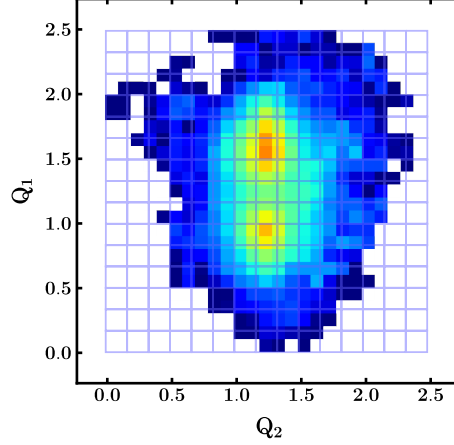
kde p je predikce třídy detekovaného objektu, t^u predikovaný rámeček, m predikovaná instanční maska, u, v, w pravdivá třída, pravdivý rámeček a pravdivá maska detekovaného objektu. Jak pro RPN, tak pro výstupy, je L_{cls} kategorickou křížovou entropií, L_{loc}/L_{reg} Huberová ztrátová funkce a L_{mask} binární křížová entropie mezi pixely. Člen L_{cls} zde ovlivňuje přesnost s jakou bude model klasifikovat jednotlivé třídy, L_{loc} odpovídá za polohu emitovaného rámečku a L_{mask} odpovídá za přesnost emitovaných masek.

Pro Mask R-CNN existuje více postupů trénování, které lze nalézt v [20, 21]. Zde je popsán jeden z nich, který je označen jako "approximate joint training". Ten spočívá v tom, že se přes model provede dopředná propagace, při které se vygenerují RoI kandidáti. Tyto RoI se pak dále předpokládají za fixní. Zpětná propagace následně modifikuje váhy RPN a zbytku sítě současně. Problém tohoto přístupu je, že nebere v úvahu derivace vzhledem k výstupům RPN sítě. Proto se nejedná o plnohodnotný přístup, ovšem s jeho pomocí lze efektivně trénovat model od začátku do konce.

5.6 Nedávné výsledky v aplikacích strojového učení na jetovou fyziku

Závěr této kapitoly je věnován hlavním výsledkům v oblasti aplikací strojového učení na jetovou fyziku. Majorita výzkumu se dnes soustředí na otázku tagování jetů - tj. na identifikaci částice, která vedla ke vzniku daného jetu. Jedny z prvních prací, které aplikovaly moderní strojové učení na tagování jetů, jsou [22, 62]. V nich zavedli autoři pojem jetového obrázku, což umožnilo použít aparát konvolučních neuronových sítí

na tagování jetů pocházejících z rozpadů W bozonů. Příklad takové reprezentace lze vidět na Obr. 5.5.

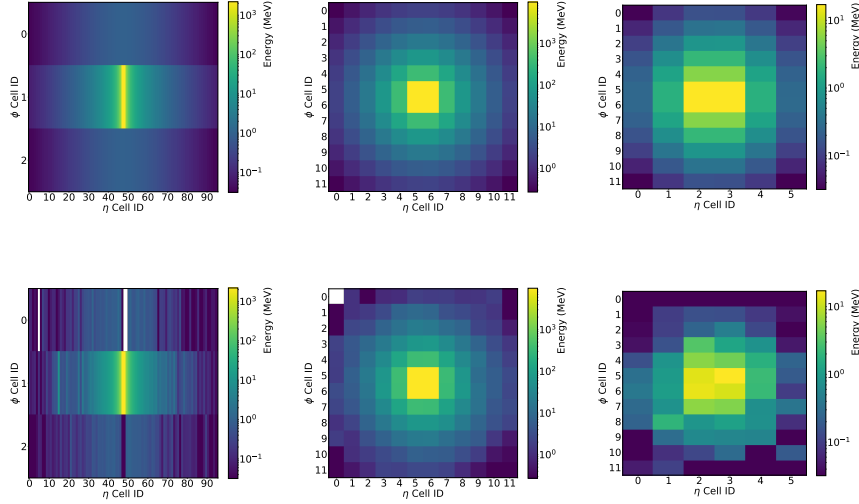


Obrázek 5.5: Příklad jetového obrázku zprůměrovaného přes velký počet jetů. Osa Q_1 vznikla rotací a translací souřadnice φ , osa Q_2 rotací a translací souřadnice η . Převzato z [22].

Tento přístup umožnil provádět tagování jetů na základě nízkoúrovňové informace bez potřeby pracného odvozování vysokoúrovňových proměnných. Dalším krokem bylo přidání barevných kanálů do těchto obrázků v [63], kde jednotlivý kanál reprezentoval nabitě nebo neutrální částice. V tomto článku autoři zkoumají, jak rozlišovat jety pocházející z kvarků od jetů pocházejících z gluonů. Hlavním přínosem bylo potvrzení intuitivního faktu, že přidání vícevrstvé informace vede ke zvýšení klasifikačního výkonu.

Tagování založené na jetových obrázcích ovšem selhává v případě potřeby tagovat jety pocházející z b a c kvarků, jelikož v jejich případě je potřeba znát informaci o vertexech částic. V tomto případě existuje několik možností - extrahovat vysokoúrovňovou informaci z jetu, použít nízkoúrovňovou informaci anebo jejich kombinaci, která se bude zpracovávat pomocí klasických dopředných sítí nebo rekurentních modelů (tyto přístupy lze např. najít v [64]). Tyto přístupy ovšem mají i své nedostatky. V případě prvního modelu se jedná o pracné odvozování těchto proměnných. Ve druhém případě jde o konceptuální problém - rekurentní modely vyžadují přítomnost posloupnosti v datech. Neexistuje ovšem žádný přirozený způsob, jak vyznačit posloupnost v jetu po výstupu z anti- k_t klastrování. Proto autoři zavádějí umělá uspořádání - např. dle p_T drah částic, dle vzrůstajících vzdáleností vertexů drah částic od primárního vertexu. Lze tedy říct, že otázka finální reprezentace v tagování jetů pocházejících z rozpadů b a c kvarků je stále otevřená.

Jiným velmi zajímavým směrem je simulace elektromagnetických spršek, které vznikají následkem průchodu částic jetu detektorem. Správný a rigorózní způsob vyžaduje simulaci průchodu částic pomocí výpočetního kódu GEANT [65, 66, 67], který je však výpočetně velmi náročný. Naproti tomu se nabízí možnost použití generativních adversiálních sítí [68] pro simulaci těchto spršek [23]. Ačkoliv trénování těchto sítí je komplikované, výsledné predikce na GPU jsou mnohonásobně rychlejší než



Obrázek 5.6: Průměrná γ sprška v GEANT4 (nahore) a průměrná γ sprška v CALOGAN (dole). Převzato z [23].

to, co se získá použitím simulátoru GEANT. I kvalitativně jsou takto generované spršky podobné těm, které se získají pomocí kódu GEANT. Srovnání obou přístupů lze vidět na Obr. 5.6.

V rámci této práce se čtenář dočte o aplikaci modelu Mask R-CNN pro detekci jetů v $\eta - \varphi$ rovině detektoru. Oproti dřívějším pracím, které se primárně soustředily na generativní modelování či na klasifikaci jednotlivých objektů, bude tento přístup používat celý obrázek srážky, která má simulovat výčet srážky z kalorimetru. Cílem je pak najít všechny jety v této srážce bez potřeby provádět rekonstrukci srážky a následné klastrování pomocí algoritmu anti- k_t .

Kapitola 6

Detekce objektů pro jetovou fyziku

Tato kapitola se soustředí na aplikaci modelu Mask R-CNN pro detekci jetů z jediného obrázku reprezentujícího $\eta - \varphi$ rovinu detektoru. Nejprve bude popsán proces přípravy dat, pak bude následovat popis postupu trénování a nakonec budou představeny výsledky týkající se analýzy modelu.

6.1 Množina dat

Jako inspirace pro postup generace dat posloužily [62, 23]. Kalorimetr se modeluje jako 2D histogram v (η, φ, p_T) prostoru, ve kterém se zaznamenávají jednotlivé zásahy částic v binech s tím, že se p_T těchto částic sčítá. V této práci byl zvolen detektor s $\eta \in [-1, 1]$ a $\varphi \in [0, 2\pi]$, který se diskretizuje do 128×128 binového histogramu. Částice jsou generovány pomocí Monte-Carlo generátoru *Pythia8* [69, 70] pro p+p srážky, resp. pomocí generátoru *HYDJET++* [71, 72, 73, 74, 75] pro Au+Au srážky.

Pythia8 je Monte-Carlo generátor srážek, který dokáže simulovat široké spektrum jevů, vyskytujících se v p+p a $e^- + e^+$ srážkách. V dnešní době se jedná o nejpopulárnější generátor pro simulaci p+p srážek a proto je schopný už i v základní konfiguraci dobře simulovat vlastnosti jetů v nich. *Pythia8* ovšem neumí popisovat jádro-jaderné srážky, proto pro tyto simulace byl zvolen generátor *HYDJET++*. Tento generátor v sobě kombinuje dvě části - měkkou část a tvrdou část. Měkká část je založena na parametrizaci relativistické hydrodynamiky, která umožňuje simulovat hydrodynamické vlastnosti jádro-jaderných srážek bez potřeby náročných výpočtů v relativistické hydrodynamice. Tvrdá část, která slouží pro simulace jetů v jaderném médiu, vychází z modifikovaného generátoru *Pythia6* do kterého byly přidány mechanismy pro energetické ztráty partonů. Základní konfigurační soubory, které jsou součástí generátoru, umožňují provádět simulace různých aspektů jetové fyziky v jádro-jaderných srážkách s dobrou přesností.

Generátor *Pythia8* se inicializuje na p+p srážky o energii $\sqrt{s} = 200$ GeV. Dále se zvolí všechny tvrdé procesy s pomocí selektoru `HardQCD:all = on`. U těchto procesů je nutné zvolit parametry $\hat{p}_{T,min}$ a $\hat{p}_{T,max}$. Tyto parametry ovlivňují celkovou

podobu p_T distribuce rekonstruovaných jetů, proto jsou voleny dvě sady parametrů. Jedna sada se zvolí jako $\hat{p}_{T,min} = 9,0 \text{ GeV}/c$ a hodnota $\hat{p}_{T,max}$ se neomezuje. Tato sada má generovat realisticky vypadající jetová p_T spektra. V druhé sadě se volí $\hat{p}_{T,min}, \hat{p}_{T,max}$ po $10 \text{ GeV}/c$ intervalech od 10 do $60 \text{ GeV}/c$. Tato volba zajistí uniformní p_T spektrum. Pro generátor HYDJET++ se užívá konfigurace ze souboru RunInputHydjetRHIC200, jediná změna, kterou provádíme, spočívá v tom, že je zvolen maximální rozsah centralit srážky. Během generace jsou přijaty pouze ty částice, jejichž $p_T > 0,2 \text{ GeV}/c$ a $|\eta| < 1$. Tato volba odpovídá jetovým analýzám v experimentu STAR.

Po zaznamenání srážky se postupně extrahují (η, φ, p_T) každého binu a předpokládá se, že hmotnost "částice" je nulová. Toto pak slouží jako informace na vstup do anti- k_t algoritmu s rozlišovacím parametrem $R = 0,4$. Následně jsou jako jety zvoleny ty objekty, jejichž $p_T^{jet} > 10 \text{ GeV}/c$. Pro Au+Au srážky se taktéž provádí odhad pozadí a jsou přijímány takové jety, pro které platí $p_{T,raw}^{jet} - \rho A > 10 \text{ GeV}/c$. V obou případech se také vyžaduje tzv. fiduciální akceptance, tj. aby $|\eta_{jet}| < 0,6$. Pro zjednodušení trénovacího procesu se vyžaduje, aby počet konstituentů v jetu $N_{const} > 1$.

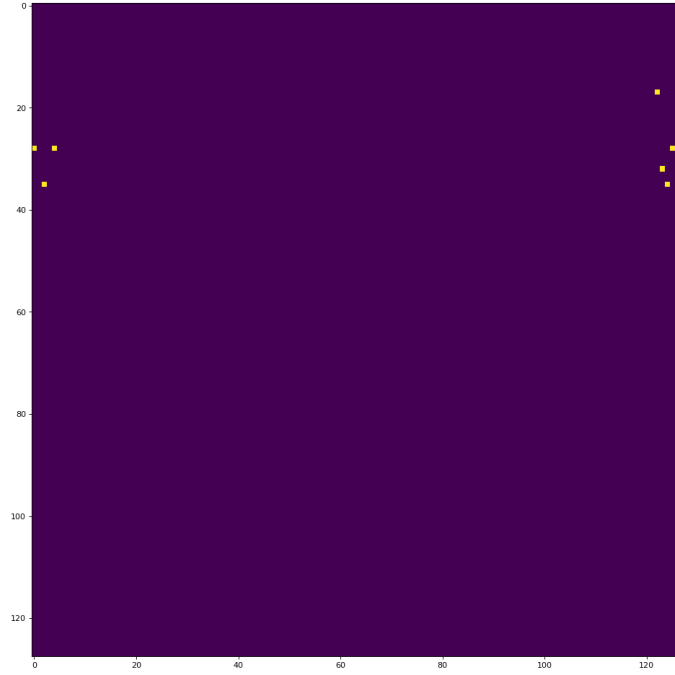
Po klastrování se pak provádí ukládání jetů a srážek, ve kterých se našly. Jednotlivý jet je reprezentován jako binární matice o rozměru 128×128 . Srážka se pak ukládá jako 128×128 matice, kde jednotlivé prvky odpovídají p_T v daném binu. Tyto masky a srážky uložené do souborů ve formátu ROOT [76] jsou pak užity jako vstup do modelu Mask R-CNN.

Pro potřeby této práce jsou tedy generovány 3 množiny dat - p+p srážky s realistickým p_T spektrem, s konstantním p_T spektrem a Au+Au srážky. Každá množina dat obsahuje 10^6 vygenerovaných srážek pro trénování a 10^5 srážek pro testování a validaci.

6.2 Příprava a trénování modelu

V této práci používáme implementaci modelu Mask R-CNN [77], který je implementován v knihovně Keras [78]. Pro potřeby této práce jsme provedli některé modifikace tohoto modelu, zejména:

- Byl přidán model ResNet-18 [79] jako jednodušší příznakový extraktor. Toto je umožněno tím, že data, se kterými se pracuje, jsou vizuálně "jednoduchá". Přínosem je pak rychlejší a jednodušší trénování modelu.
- Počáteční inicializace je zvolena na ortogonální matice [61]. Tato modifikace byla vybrána na základě empirických experimentů, ve kterých se projevila jako nejlepší.
- Byla přidána konfigurace modelu přes YAML soubory, což umožňuje jednodušší automatizaci experimentů.

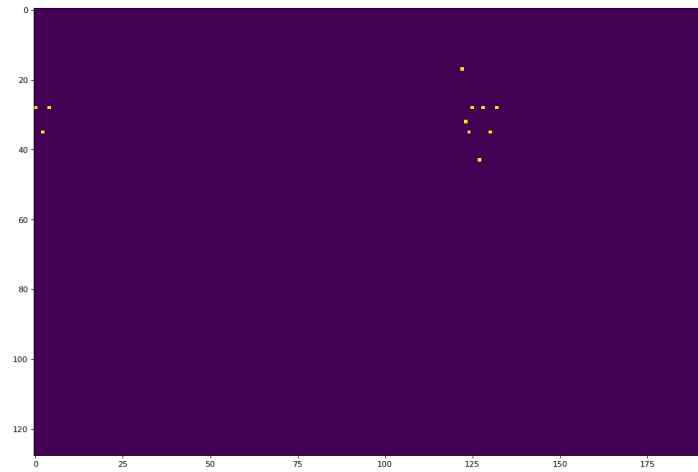


Obrázek 6.1: Příklad masky, ve které se vyskytuje problém s nespojitostí ve φ . Vodorovná osa představuje diskretizované hodnoty φ v intervalu $[0, 2\pi]$, svislá osa hodnoty η v intervalu $[-1, 1]$.

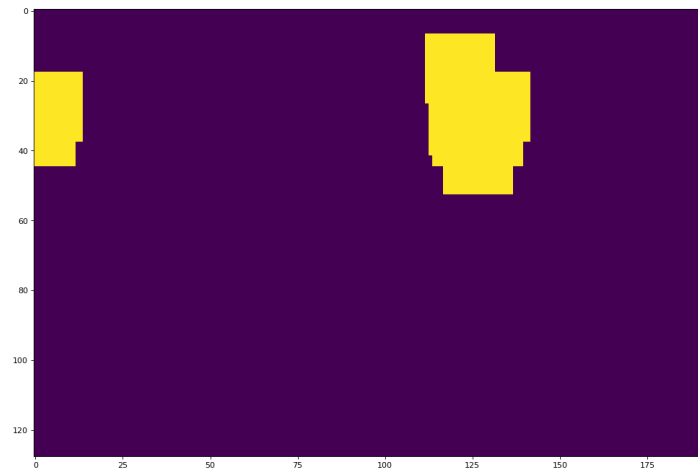
Při naivním přístupu k trénování vzniká problém, který je důsledkem vnesení umělé nespojitosti v souřadnici φ při diskretizaci. Příklad takové nespojitosti lze vidět na Obr. 6.1. Její řešení spočívá v prodloužení zdrojového obrázku a masky tak, aby na něm byl přítomen celý jet. Pak se pomocí morfologické dilatace jednotlivé jety zvětší a rozdělí na dvě souvislé oblasti. Plnohodnotný jet je pak dán logickým AND s původní prodlouženou maskou.

Celkově lze tedy tento postup zformulovat do následujícího algoritmu:

1. Nejprve se prodlouží detektor ve φ souřadnici od $[0, 2\pi]$ do $[0, 3\pi]$, což bude odpovídat zvětšení obrázku z 128 na 192 pixelů. Toto lze vidět na Obr. 6.2. Výskyt duplikovaných částic není během trénování problematický, jelikož tento postup eliminuje duplikátní jety během trénování. Během testování modelu bude ovšem potřeba zvolit vhodný postprocessing, aby se zamezilo duplicitním detekcím.
2. Jako další krok se aplikuje tzv. binární morfologická dilatace. Ta mnohonásobně zvětší jednotlivé pixely v masce, což lze vidět na Obr. 6.3.

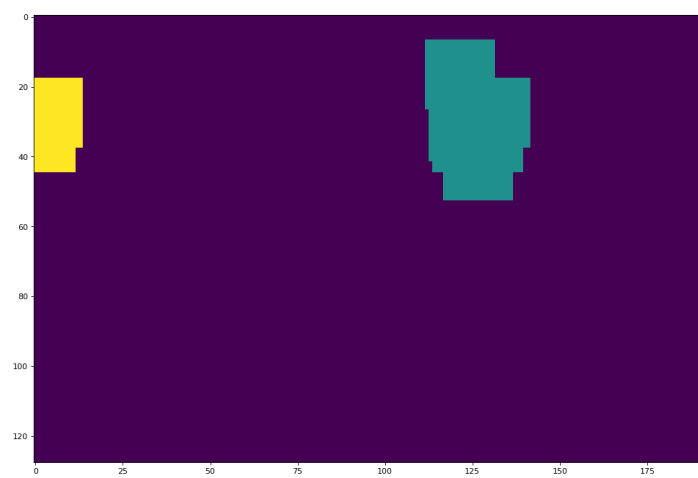


Obrázek 6.2: Prodloužení původní masky ze 128 na 192 pixelů dle vodorovné osy. Toto odpovídá prodloužení detektoru z $[0, 2\pi]$ do $[0, 3\pi]$.

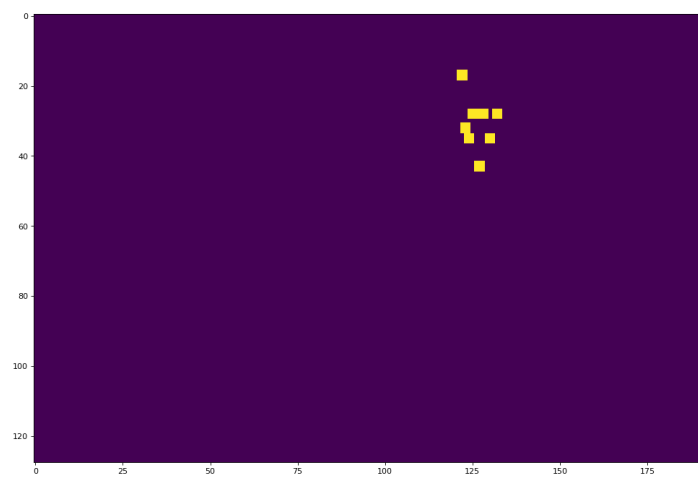


Obrázek 6.3: Maska po zpracování 20×20 morfologickou dilatací.

3. Výstupem této operace jsou dvě nezávislé oblasti (toto je zaručené, jelikož každá maska reprezentuje jen jeden jet), ve které jedna představuje původní jet a druhá jeho polovinu. Jelikož tyto oblasti jsou souvislé, tak je možné je vyseparovat. Tento krok je zobrazen na Obr. 6.4.
4. Po vyseparování těchto dvou oblastí pak stačí vzít tu s největší plochou a provést logický AND mezi zvětšenou plochou a původní prodlouženou maskou. Následně se pak pixely zvětšují pomocí 3×3 dilatace pro usnadnění učícího procesu. Výsledek lze vidět na Obr. 6.5.



Obrázek 6.4: Separace dvou dilatovaných oblastí.

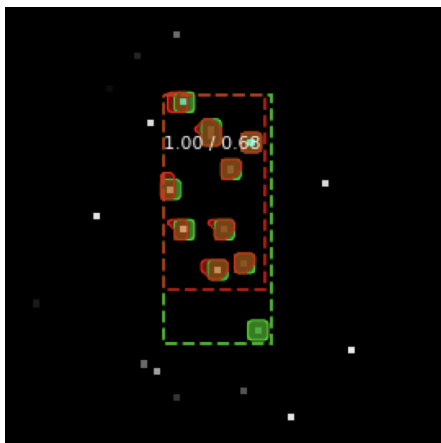


Obrázek 6.5: Konečná podoba masky, která jde na vstup do modelu Mask R-CNN.

Po přípravě modelu se přistoupilo k jeho trénování. Pro obě množiny dat byly zvoleny velikosti dávek pro estimaci gradientu jako 8 obrázků a následně byl model trénován po 1500 iterací, kde každá iterace obsahovala 200 dávek. Na konci každé iterace bylo provedeno vyhodnocování modelu na 10 dávkách z validační množiny dat. Počáteční míra učení byla zvolena jako $\lambda = 2,5 \cdot 10^{-3}$ s tím, že se ve 2/3 iterací snížila na $2,5 \cdot 10^{-4}$ a v 8/10 iterací na $2,5 \cdot 10^{-5}$. Po skončení trénování byl jako finální model zvolen ten, který dosáhl nejmenší hodnoty ztrátové funkce na validační množině dat.

6.3 Evaluace modelu

Před evaluací modelu na testovací množině dat je nejprve potřebné provést "kalibraci" modelu, tj. najít takové kombinace parametrů pro inferenci, které budou dávat nejlepší předpovědi. Příklad dobré předpovědi lze vidět na Obr. 6.6.



Obrázek 6.6: Příklad detekce (červený rámeček) v porovnání s pravdivým jetem (zelený rámeček). Skóre 1.00/0.68 říká, že se jedná o jet s pravděpodobností 100% a IoU mezi predikcí a pravdivým objektem je 68%.

V první řadě byla zvolena mez na klasifikační skóre a to na 98%. Tj. všechny předpovědi modelu, u nichž model má jistotu menší než 98%, jsou zahozeny. Dále je potřeba provést postprocessing, který spočívá v následujících krocích:

1. V první řadě se pomocí rámečků predikovaných Mask R-CNN vyčte p_T predikovaného jetu. V případě, že $p_T > 10 \text{ GeV}/c$ je rámeček přijat, v opačném případě je zahozen.
2. Dále je potřeba eliminovat mnohonásobné detekce v jedné oblasti. Pro tento účel se využívá metrika IoU (Intersection over Union), která je dána jako

$$\text{IoU} = \frac{B_1 \cap B_2}{B_1 \cup B_2}, \quad (6.1)$$

kde B_1 a B_2 jsou predikované rámečky. Tato metrika udává, nakolik se rámečky překrývají procentuálně. Je předpokládáno, že pokud $\text{IoU} > 0,2$, pak se jedná

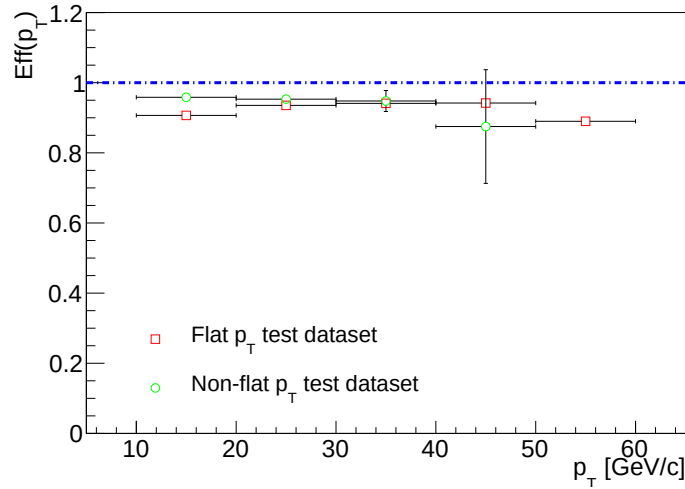
o případ mnohonásobné detekce a je zahozen ten rámeček, jehož jet má nižší p_T .

3. Posledním krokem je odstranění možných duplikátních detekcí. Ty vznikají v případě, že φ souřadnice jetu je plně obsažena v $[0, 2\pi]$ mimo okraje. Při prodloužení na $[0, 3\pi]$ pak takový jet bude obsažen na obrázku dvakrát a tedy je možné, že i model provede detekci dvakrát. Pro eliminaci tohoto efektu je v první řadě potřeba vzít polohu nejvzdálenějšího rámečku, posunout ho o 128 pixelů doleva a poté spočítat jejich IoU. Je-li $\text{IoU} > 0,8$, zcela jistě se jedná o shodné objekty a predikované rámečky lze taktéž zahodit.

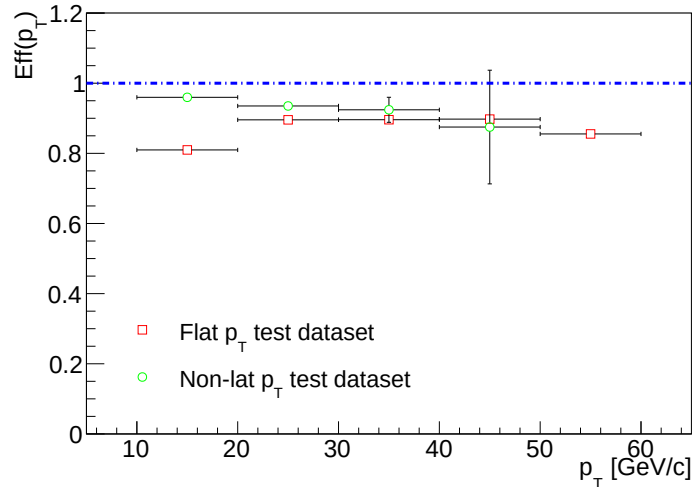
Po kalibraci modelu lze přistoupit k jeho vyhodnocení. Cílem je porovnat, jak moc je navržená metoda účinná oproti standardnímu klastrování pomocí algoritmu anti- k_t . Tato účinnost se bude určovat pomocí proměnné zvané efektivita (Eff) v pevně daném p_T binu jako

$$\text{Eff}(p_T) = \frac{\# \text{ jetů detekovaných Mask R-CNN}}{\# \text{ jetů rekonstruovaných anti-}k_t}. \quad (6.2)$$

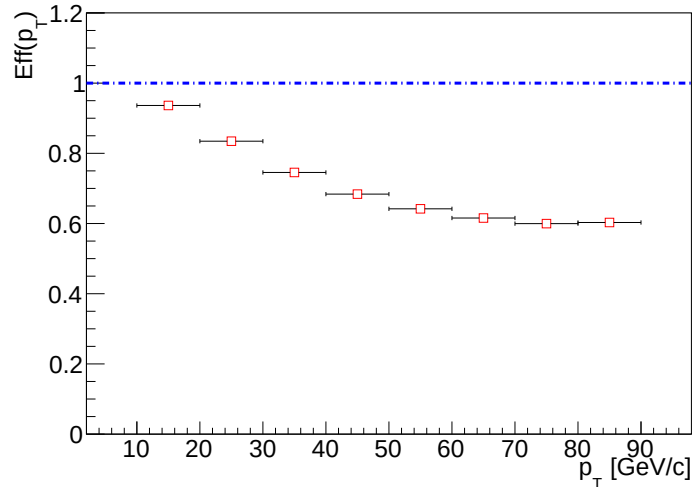
Predikovaný jet bude považován za nalezený vzhledem k pravdivému jetu, pokud bude jeho $\text{IoU} > 0,25$ a příčná hybnost určená Mask R-CNN $p_T^{\text{mrcnn}} > 0,75p_T^{\text{anti-}k_t}$. Výsledky efektivit pro p+p srážky pro konstantní p_T spektrum lze vidět na Obr. 6.7, pro realistické p_T spektrum na Obr. 6.8 a pro Au+Au srážky na Obr. 6.9.



Obrázek 6.7: Efektivita detekce jetů v p+p srážkách pomocí modelu Mask R-CNN trénovaného na konstantním p_T spektru. Červené body ukazují výsledky modelu při testování na množině dat s konstantním p_T , zelené s realistickým p_T . Lze vidět, že model trénovaný na této trénovací množině dosahuje přibližně stejných efektivit napříč celým intervalem p_T , což je zapříčiněno rovnoměrným zastoupením jetů s různými hybnostmi. Některé chyby měření jsou menší než velikosti naměřených bodů.



Obrázek 6.8: Efektivita detekce jetů v p+p srážkách pomocí modelu Mask R-CNN trénovaného na realistickém p_T spektru. Červené body ukazují výsledky modelu při testování na množině dat s plochým p_T , zelené s realistickým p_T . Na rozdíl od modelu, který byl trénován na množině dat s konstantním p_T , dosahuje tento model horších výsledků pro jety s vysokým p_T . Toto je způsobeno nerovnoměrným zastoupením jetů v dané množině dat, kde dominují jety s nízkým p_T . Některé chyby měření jsou menší než velikosti naměřených bodů.



Obrázek 6.9: Efektivita detekce jetů v Au+Au srážkách pomocí modelu Mask R-CNN. Analogicky jako v předchozím případě je i výkon tohoto modelu ovlivněn nerovnoměrným zastoupením jetů v trénovací množině. Chyby měření jsou menší než velikosti naměřených bodů.

Závěr

Tato práce se soustředila na téma aplikace strojového učení na jetovou fyziku. V první kapitole jsme popsali základní vlastnosti částic a interakcí a soustředili jsme se zejména na základy silné interakce. Ve druhé kapitole jsme popsali fyziku ultrarelativistických jádro-jaderných srážek a zmínili jsme základní měření, která potvrzují některé vlastnosti kvark-gluonového plazmatu. Problematiku rekonstrukce jetů, jejich dekontaminaci vůči měkké složce a dekonvoluci fluktuujícího pozadí jsme popsali ve třetí kapitole. Čtvrtá kapitola se pak zabývala popisem jednotlivých subdetektorů na experimentu STAR a krátce se zmínila i o urychlovači RHIC jako o celku. Pátá kapitola pak sloužila jako stručný úvod do problematiky strojového učení a aplikací strojového učení na jetovou fyziku.

V šesté kapitole byly uvedeny výsledky týkající se aplikace modelu Mask R-CNN na detekci jetů v $\eta - \varphi$ rovině. Výsledné efektivit v závislosti na příčných hybnostech jetů v p+p srážkách lze vidět na Obr. 6.7 a 6.8, pro Au+Au srážky lze efektivitu vidět na Obr. 6.9. Námi natrénované modely se naučily identifikovat jety v $\eta - \varphi$ rovině bez rekonstrukce srážky a následného klastrování pomocí jetového algoritmu. základní metoda, ke které jsme vztahovali výkon našeho modelu, nám posloužil algoritmus anti- k_t . Z námi získaných efektivit lze vidět, že se v p+p srážkách výkon modelu Mask R-CNN blíží tomu, co získáme pomocí klasického klastrování algoritmem anti- k_t . V Au+Au srážkách je tato shoda horší, jelikož proces generace dat, který jsme zvolili, nedostatečně odstraňoval pozadí, které je v jádro-jaderných srážkách velké. Problematické je taktéž to, že z hlediska experimentální fyziky je jet definovaný pomocí jetového algoritmu, tedy základní metoda nám z definice dává ten nejlepší výsledek. Naši metodu je tedy dobré chápat jako aproximaci pro náročný proces rekonstrukce srážky a následného klastrování jetů, ne jako rigorózní alternativu. Toto může mít praktické využití jako rychlý trigger pro srážky, ve kterých jsou přítomné jety s jistými charakteristikami.

Dále rozebereme některé nedostatky tohoto přístupu a způsoby, kterými je lze odstranit. Grafy efektivit ukazují jasnou závislost mezi tvarem p_T spektra jetů a tvarem efektivit. Z teoretického hlediska je správné modely strojového učení trénovat na rozděleních, která jsou co nejbližší reálně pozorovaným. Ovšem jak vidíme na Obr. 6.8, toto vede ke zhoršení výkonu modelu při vysokých p_T v důsledku toho, že trénovací množina má malý počet jetů s vysokým p_T . Tento problém by šel vyřešit tak, že by se trénovací množina rozdělila na jety po 10 GeV/c a při trénování by se každý jet váhoval relativní četností vzhledem k jetům s nejvyšším p_T .

Dále je potřeba provést modifikaci procesu pro generaci dat. Seleční kritérium

$p_{T,raw}^{jet} - \rho A > 10 \text{ GeV}/c$ na přijímané jety v Au+Au srážkách slouží pouze jako první přiblížení, jelikož takové jety, ačkoliv už nejsou tvořeny jenom náhodnými fluktuacemi pozadí, mají stále velkou příměs měkké složky. Jako vhodným kandidátem na dekontaminaci může sloužit metoda **constituent-subtractor** [80], která se ukázala být vhodným nástrojem pro redukci pozadí v jádro-jaderných srážkách. Je také potřeba zanést detektorové efekty do trénovacích dat, hlavně simulace šumu v detektoru, a taktéž rozmazávání p_T rozdělení jednotlivých částic.

Další výzkum v této oblasti se může týkat natrénování modelu na realistických datech či implementaci této metody pro výběr fyzikálně zajímavých srážek na úrovni triggeru experimentu STAR.

Literatura

- [1] Wikipedia, the free encyclopedia, Standard model of elementary particles, https://upload.wikimedia.org/wikipedia/commons/0/00/Standard_Model_of_Elementary_Particles.svg [Online; 17.3.2018].
- [2] S. Dulat *et al.*, Phys. Rev. **D93**, 033006 (2016), 1506.07443.
- [3] K. Fukushima and C. Sasaki, Prog. Part. Nucl. Phys. **72**, 99 (2013), hep-ph/1301.6377.
- [4] P. Rosnet, Quark-Gluon Plasma: from accelerator experiments to early Universe, in *11th Rencontres du Vietnam: Cosmology: 50 years after CMB discovery Quy Nhon, Vietnam, August 16-22, 2015*, 2015, hep-hp/1510.04200.
- [5] K. Tuchin, Adv. High Energy Phys. **2013**, 490495 (2013), hep-ph/1301.0099.
- [6] M. L. Miller, K. Reygers, S. J. Sanders, and P. Steinberg, Ann. Rev. Nucl. Part. Sci. **57**, 205 (2007), nucl-ex/0701025.
- [7] STAR Collaboration, L. Adamczyk *et al.*, Phys. Rev. Lett. **121**, 032301 (2018), 1707.01988.
- [8] STAR Collaboration, J. Adam *et al.*, Phys. Rev. **C99**, 034908 (2019), 1812.10224.
- [9] ALICE Collaboration, J. Adam *et al.*, JHEP **03**, 081 (2016), 1509.06888.
- [10] STAR Collaboration, J. Adams *et al.*, Phys. Rev. Lett. **91**, 072304 (2003), nucl-ex/0306024.
- [11] STAR Collaboration, J. Adams *et al.*, Phys. Rev. Lett. **97**, 162301 (2006), nucl-ex/0604018.
- [12] ATLAS Collaboration, G. Aad *et al.*, Phys. Rev. Lett. **114**, 072302 (2015), 1411.2357.
- [13] STAR Collaboration, L. Adamczyk *et al.*, Phys. Rev. Lett. **119**, 062301 (2017), 1609.03878.
- [14] CMS Collaboration, S. Chatrchyan *et al.*, Phys. Rev. **C84**, 024906 (2011), 1102.1957.
- [15] M. Cacciari, G. P. Salam, and G. Soyez, JHEP **04**, 063 (2008), 0802.1189.

- [16] J. Rusnak, *Jet Reconstruction in Au+Au Collisions at RHIC*, PhD thesis, Czech Tech. Univ., 2016.
- [17] STAR Collaboration, J. Rusnak, PoS **EPS-HEP2013**, 191 (2013).
- [18] STAR Collaboration, Drupal, <https://drupal.star.bnl.gov/STAR/book/export/html/7278> [Online; 2.6.2019].
- [19] LBL News Center, <https://newscenter.lbl.gov/wp-content/uploads/sites/2/HFT-figure2.jpg> [Online; 2.6.2019].
- [20] K. He, G. Gkioxari, P. Dollár, and R. B. Girshick, Mask R-CNN, in *ICCV*, pp. 2980–2988, IEEE Computer Society, 2017.
- [21] S. Ren, K. He, R. B. Girshick, and J. Sun, IEEE Trans. Pattern Anal. Mach. Intell. **39**, 1137 (2017).
- [22] J. Cogan, M. Kagan, E. Strauss, and A. Schwartzman, JHEP **02**, 118 (2015), 1407.5675.
- [23] M. Paganini, L. de Oliveira, and B. Nachman, Phys. Rev. **D97**, 014021 (2018), 1712.10321.
- [24] Particle Data Group, M. Tanabashi *et al.*, Phys. Rev. **D98**, 030001 (2018).
- [25] J. C. Collins, D. E. Soper, and G. F. Sterman, Adv. Ser. Direct. High Energy Phys. **5**, 1 (1989), hep-ph/0409313.
- [26] V. N. Gribov and L. N. Lipatov, Sov. J. Nucl. Phys. **15**, 438 (1972), [Yad. Fiz.15,781(1972)].
- [27] G. Altarelli and G. Parisi, Nucl. Phys. **B126**, 298 (1977).
- [28] Y. L. Dokshitzer, Sov. Phys. JETP **46**, 641 (1977), [Zh. Eksp. Teor. Fiz.73,1216(1977)].
- [29] K. Yagi, T. Hatsuda, and Y. Miake, Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. **23**, 1 (2005).
- [30] R. Vogt, *Ultrarelativistic heavy-ion collisions* (Elsevier, Amsterdam, 2007).
- [31] Y. L. Dokshitzer and D. E. Kharzeev, Phys. Lett. **B519**, 199 (2001), hep-ph/0106202.
- [32] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, Nucl. Phys. **B406**, 187 (1993).
- [33] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, JHEP **08**, 001 (1997), hep-ph/9707323.
- [34] M. Wobisch, *Measurement and QCD analysis of jet cross-sections in deep inelastic positron proton collisions at $\sqrt{s} = 300$ GeV*, PhD thesis, Aachen, Tech. Hochsch., 2000.

- [35] G. D’Agostini, Improved iterative Bayesian unfolding, in *Alliance Workshop on Unfolding and Data Correction Hamburg, Germany, May 27-28, 2010*, 2010, 1010.0632.
- [36] A. Hocker and V. Kartvelishvili, Nucl. Instrum. Meth. **A372**, 469 (1996), hep-ph/9509307.
- [37] BRAHMS Collaboration, I. Arsene *et al.*, Nucl. Phys. **A757**, 1 (2005), nucl-ex/0410020.
- [38] B. B. Back *et al.*, Nucl. Phys. **A757**, 28 (2005), nucl-ex/0410022.
- [39] PHENIX Collaboration, K. Adcox *et al.*, Nucl. Phys. **A757**, 184 (2005), nucl-ex/0410003.
- [40] STAR Collaboration, J. Adams *et al.*, Nucl. Phys. **A757**, 102 (2005), nucl-ex/0501009.
- [41] M. Anderson *et al.*, Nucl. Instrum. Meth. **A499**, 659 (2003), nucl-ex/0301015.
- [42] STAR Collaboration, M. Beddo *et al.*, Nucl. Instrum. Meth. **A499**, 725 (2003).
- [43] C. Chasman *et al.*, LBNL-PUB-5509-2008 (2008).
- [44] G. Contin *et al.*, Nucl. Instrum. Meth. **A907**, 60 (2018), 1710.02176.
- [45] L. Arnold *et al.*, Nucl. Instrum. Meth. **A499**, 652 (2003), physics/0211083.
- [46] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)* (Springer-Verlag, Berlin, Heidelberg, 2006).
- [47] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT Press, 2016), <http://www.deeplearningbook.org>.
- [48] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, New York, NY, USA, 2004).
- [49] Y. Nesterov, *Introductory Lectures on Convex Optimization: A Basic Course*, 1 ed. (Springer Publishing Company, Incorporated, 2014).
- [50] L. Bottou and O. Bousquet, The tradeoffs of large scale learning, in *Proceedings of the 20th International Conference on Neural Information Processing Systems*, NIPS’07, pp. 161–168, USA, 2007, Curran Associates Inc.
- [51] L. Bottou, On-line learning in neural networks, chap. On-line Learning and Stochastic Approximations, pp. 9–42, Cambridge University Press, New York, NY, USA, 1998.
- [52] Y. LeCun, L. Bottou, G. B. Orr, and K.-R. Müller, Efficient backprop, in *Neural Networks: Tricks of the Trade, This Book is an Outgrowth of a 1996 NIPS Workshop*, pp. 9–50, London, UK, UK, 1998, Springer-Verlag.

- [53] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, in *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 2015.
- [54] M. D. Zeiler, CoRR **abs/1212.5701** (2012), 1212.5701.
- [55] I. Loshchilov and F. Hutter, SGDR: stochastic gradient descent with warm restarts, in *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*, 2017.
- [56] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, Gradient-based learning applied to document recognition, in *Proceedings of the IEEE* Vol. 86, pp. 2278–2324, 1998.
- [57] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, Parallel distributed processing: Explorations in the microstructure of cognition, vol. 1, chap. Learning Internal Representations by Error Propagation, pp. 318–362, MIT Press, Cambridge, MA, USA, 1986.
- [58] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, *Nature* **323**, 533 (1986).
- [59] X. Glorot and Y. Bengio, Understanding the difficulty of training deep feedforward neural networks, in *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, edited by Y. W. Teh and M. Titterton, , Proceedings of Machine Learning Research Vol. 9, pp. 249–256, Chia Laguna Resort, Sardinia, Italy, 2010, PMLR.
- [60] K. He, X. Zhang, S. Ren, and J. Sun, Delving deep into rectifiers: Surpassing human-level performance on imagenet classification, in *Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV)*, ICCV ’15, pp. 1026–1034, Washington, DC, USA, 2015, IEEE Computer Society.
- [61] A. M. Saxe, J. L. McClelland, and S. Ganguli, Exact solutions to the nonlinear dynamics of learning in deep linear neural networks, in *ICLR*, 2014.
- [62] L. de Oliveira, M. Kagan, L. Mackey, B. Nachman, and A. Schwartzman, *JHEP* **07**, 069 (2016), 1511.05190.
- [63] P. T. Komiske, E. M. Metodiev, and M. D. Schwartz, *JHEP* **01**, 110 (2017), 1612.01551.
- [64] D. Guest *et al.*, *Phys. Rev.* **D94**, 112002 (2016), 1607.08633.
- [65] GEANT4, S. Agostinelli *et al.*, *Nucl. Instrum. Meth.* **A506**, 250 (2003).
- [66] J. Allison *et al.*, *Nucl. Instrum. Meth.* **A835**, 186 (2016).
- [67] J. Allison *et al.*, *IEEE Trans. Nucl. Sci.* **53**, 270 (2006).
- [68] I. J. Goodfellow *et al.*, Generative adversarial nets, in *NIPS*, pp. 2672–2680, 2014.

- [69] T. Sjostrand, S. Mrenna, and P. Z. Skands, JHEP **05**, 026 (2006), hep-ph/0603175.
- [70] T. Sjöstrand *et al.*, Comput. Phys. Commun. **191**, 159 (2015), 1410.3012.
- [71] I. P. Lokhtin *et al.*, Comput. Phys. Commun. **180**, 779 (2009), 0809.2708.
- [72] I. P. Lokhtin and A. M. Snigirev, Eur. Phys. J. **C45**, 211 (2006), hep-ph/0506189.
- [73] N. S. Amelin *et al.*, Phys. Rev. **C74**, 064901 (2006), nucl-th/0608057.
- [74] N. S. Amelin *et al.*, Phys. Rev. **C77**, 014903 (2008), 0711.0835.
- [75] K. Tywoniuk, I. C. Arsene, L. Bravina, A. B. Kaidalov, and E. Zabrodin, Eur. Phys. J. **C49**, 193 (2007), hep-ph/0608003.
- [76] R. Brun and F. Rademakers, Nucl. Instrum. Meth. **A389**, 81 (1997).
- [77] W. Abdulla, Mask r-cnn for object detection and instance segmentation on keras and tensorflow, https://github.com/matterport/Mask_RCNN, 2017.
- [78] F. Chollet *et al.*, Keras, <https://keras.io>, 2015.
- [79] K. He, X. Zhang, S. Ren, and J. Sun, Deep residual learning for image recognition, in *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27-30, 2016*, pp. 770–778, 2016.
- [80] P. Berta, M. Spousta, D. W. Miller, and R. Leitner, JHEP **06**, 092 (2014), 1403.3108.