

České vysoké učení technické v Praze
Fakulta jaderná a fyzikálně inženýrská

Výzkumný úkol

České vysoké učení technické v Praze
Fakulta jaderná a fyzikálně inženýrská

Řešení evolučních rovnic QCD a jejich využití při
analýze experimentálních dat

Jiří Hejbal

Katedra fyziky

Akademický rok: 2001/2002

Školitel: Prof. Jiří Chýla, CSc., FzÚ AV ČR

V Praze 17. října 2003

Obsah

1	Úvod	2
2	Mellinova transformace	4
3	Metody inverzní Mellinovy transformace	6
3.1	Integrovaní po parabolické křivce	6
3.2	Integrace po přímce	7
4	Numerický výpočet evoluční rovnice	9
5	Popis programu užitého při výpočtu evoluční rovnice	13
6	Ověření správnosti inverzní Mellinovy transformace	15
7	Krokové metody minimalizace v mnoha proměnných	18
7.1	Síťové hledání a náhodné hledání	18
7.2	Jednparametrická variace	18
7.3	Rosenbrockova metoda	19
7.4	Metoda simplexu	19
7.5	Gradientní metody	20
7.5.1	Výpočet derivací	20
7.5.2	Nejstrmější spád	21
7.5.3	Newtonova metoda	21
7.5.4	Konjugované směry	22
7.5.5	Konjugované gradienty	23
8	Výsledky a závěry	24

1 Úvod

Jedním z nejběžnějších prostředků ke zkoumání vnitřní struktury subatomárních částic jsou experimenty založené na hluboko nepružných rozptylech. Svazek vysokoenergetických bodových částic, nejčastěji elektronů nebo mionů je rozptylován na terči se zkoumanými částicemi. Ze zjištěných energetických a úhlových rozdělení rozptýlených leptonů se pak rekonstruuje struktura zkoumané částice. V 60. letech právě takové experimenty poprvé ukázaly na vnitřní strukturu hadronů. To byl počátek kvark-partonového modelu. V něm považujeme hadron složený z elementárních konstituentů - partonů. Ty v současné době ztotožňujeme s kvarky a gluony.

K popisu partonové struktury se užívají tzv. distribuční funkce. Jejich význam je následující. Uvažujeme-li partony i -tého druhu, kterým přísluší distribuční funkce $f_i(x)$, potom výraz $f_i(x)dx$ představuje počet partonů i -tého druhu, jejichž frakce hybnosti je mezi $x + dx$.

Partonový model předpovídá závislost pouze na proměnné x . Existují však experimenty, které dokazují, že distribuční funkce závisí na další veličině. Tou je kvadrát přenesené čtyřhybnosti Q^2 . Tuto skutečnost je možné popsat v rámci kvantové chromodynamiky. Nově představená závislost na Q^2 nás přivádí k zavedení tzv. "oblečené" nesingletní distribuční funkce. Ta vyhovuje Altareli-Parisiho evoluční rovnici:

$$\begin{aligned} \frac{dq_{NS}(x, Q^2)}{d \ln Q^2} &= \frac{\alpha_s(Q^2)}{2\pi} \int_x^1 \frac{dy}{y} P_{qq}^{(0)}\left(\frac{x}{y}\right) q_{NS}(y, Q^2) = \\ \frac{\alpha_s}{2\pi} \int dz \int dy P_{qq}^{(0)}(z) q_{NS}(y) \delta(x - yz) &\equiv \frac{\alpha_s}{2\pi} P_{qq}^{(0)} \otimes q_{NS}, \end{aligned} \quad (1)$$

kde α_s , představuje vazbovou konstantu QCD. Funkce $a_s \equiv \frac{\alpha_s}{\pi}$ je dána rovnicí:

$$\frac{\partial a_s(Q^2)}{\partial \ln(Q^2)} = -\beta_0 a_s^2(Q^2) - \beta_1 a_s^3(Q^2). \quad (2)$$

Řešení této rovnice můžeme psát jako:

$$\frac{1}{a_s(Q^2)} + \frac{\beta_1}{\beta_0} \ln \left[\frac{\beta_1 a_s(Q^2)}{b_0 + \beta_1 a_s(Q^2)} \right] = \beta_0 \ln \left(\frac{Q^2}{\Lambda^2} \right). \quad (3)$$

kde Λ představuje škálovací parametr, který může být pro tuto chvíli definován jako škála, při které se nuluje levá strana rovnice (3). Tato rovnice se obvykle řeší numericky. Zavedeme-li označení $L = \ln \frac{Q^2}{\Lambda^2}$, pak lze řešení rozvinout do mocninné řady v $1/L$ ve tvaru:

$$a_s = \frac{1}{\beta_0 L} \left(1 - \frac{\beta_1 \ln L}{\beta_0^2 L} \right) + \dots \quad (4)$$

$P_{qq}^{(0)}$ v rovnici (1) představuje jednu z tzv. větvičích funkcí. Je definována:

$$P_{qq}^{(0)}(x) = P_{\bar{q}\bar{q}}^{(0)}(x) = \frac{4}{3} \left[\frac{1+x^2}{1-x} \right]_+ \quad (5)$$

V poslední rovnici vystupuje "+" distribuce, která je definována:

$$[f(x)]_+ \equiv \lim_{\beta \rightarrow 0} \left(f(x) \Theta(1-x-\beta) - \delta(1-x-\beta) \int_0^{1-\beta} f(y) dy \right) \quad (6)$$

kde $\Theta(x)$ a $\delta(x)$ jsou skoková funkce a Diracova δ -funkce. Její působení na testovací funkci je následující:

$$\int_0^1 [f(x)]_+ g(x) dx \equiv \int_0^1 f(x) (g(x) - g(1)) dx. \quad (7)$$

Pro úplnost ještě uvedeme distribuční funkce kvarků, antikvarků a gluonů. Takové rovnice jsou vázané:

$$\frac{dq_i(x, M)}{d \ln M} = \frac{\alpha_s(M)}{\pi} \left[\int_x^1 \frac{dy}{y} P_{qq}^{(0)}\left(\frac{x}{y}\right) q_i(y, M) + \int_x^1 \frac{dy}{y} P_{qG}^{(0)}\left(\frac{x}{y}\right) G(y, M) \right] \quad (8)$$

$$\frac{d\bar{q}_i(x, M)}{d \ln M} = \frac{\alpha_s(M)}{\pi} \left[\int_x^1 \frac{dy}{y} P_{q\bar{q}}^{(0)}\left(\frac{x}{y}\right) \bar{q}_i(y, M) + \int_x^1 \frac{dy}{y} P_{qG}^{(0)}\left(\frac{x}{y}\right) G(y, M) \right] \quad (9)$$

$$\frac{dG(x, M)}{d \ln M} = \frac{\alpha_s(M)}{\pi} \left[\sum_{i=1}^{n_f} \int_x^1 \frac{dy}{y} P_{Gq}^{(0)}\left(\frac{x}{y}\right) (q_i(y, M) + \bar{q}_i(y, M)) + \int_x^1 \frac{dy}{y} P_{GG}^{(0)}\left(\frac{x}{y}\right) (G(y, M)) \right] \quad (10)$$

Evoluční rovnice popisují závislost distribučních funkcí na Q^2 . To znamená, že jejich řešení vyžaduje znalost závislosti distribuční funkce na x při nějaké libovolně zvolené hodnotě Q^2 . Tuto počáteční podmínku nelze zatím v rámci poruchové QCD spočítat, musíme ji určit z experimentálních dat.

V této práci se budeme zabývat jednou z metod řešení evoluční rovnice (1) ve vedoucím řádu založené na inverzní Mellinově transformaci. V jazyce C++ byl vypracován program umožňující tuto rovnici při libovolné počáteční podmínce řešit. V závěru bude představena technika určení takové počáteční podmínky, aby řešení evoluční rovnice (1) s touto podmínkou odpovídala experimentální skutečnosti. Použití této techniky nakonec vyústí ve fit experimentálních dat ne-singletní strukturní funkce protonu.

2 Mellinova transformace

V následující kapitole bude představena Mellinova transformace jako nástroj pro řešení evolučních rovnic kvantové chromodynamiky.

Pro uvedení do tohoto problému budeme uvažovat evoluční rovnici pro distribuční funkce $f(x, Q^2)$ ve tvaru

$$Q^2 \frac{\partial f(x, Q^2)}{\partial Q^2} = P(x, Q^2) \otimes f(x, Q^2), \quad (11)$$

kde x značí zlomek hybnosti protonu nesený partonem, $P(x, Q^2)$ je Altarelli-Parisiho jádro a \otimes značí konvoluci:

$$A(x) \otimes B(x) = \int_0^1 dy \int_0^1 dz \delta(x - yz) A(y) B(z). \quad (12)$$

Rovnici (11) můžeme faktorizovat použitím Mellinových momentů:

$$A(n) = \int_0^1 dx x^{n-1} A(x) \quad (13)$$

získáme tak rovnici

$$Q^2 \frac{\partial f(n, Q^2)}{\partial Q^2} = P(n, Q^2) \cdot f(n, Q^2). \quad (14)$$

Při řešení rovnice (14) existuje několik možných postupů. Všechny spočívají ve vhodných úpravách této rovnice. Předvedeme některé z nich.

Nejprve přejdeme od proměnné Q^2 k proměnné $a_s = \alpha_s(Q^2)/\pi$. Důvodem jsou známé rozvoje (přesněji vyjádření v NLO) funkcí vystupujících v rovnici (14). Máme tedy:

$$\frac{\partial f(n, Q^2)}{\partial a_s} = \frac{P(n, a_s)}{\beta(a_s)} f(n, Q^2), \quad (15)$$

kde

$$\beta(a_s) \equiv Q^2 \frac{\partial a_s}{\partial Q^2} \quad (16)$$

Předpokladejme znalost evolučního jádra do druhého řádu poruchové QCD:

$$P(n, a_s) = a_s P^{(0)}(n) + a_s^2 P^{(1)}(n) + O(a_s^3) \quad (17)$$

V těchto řádech je funkce β dána takto:

$$\beta(a_s) = -\beta_0 a_s^2 - \beta_1 a_s^3 + O(a_s^4). \quad (18)$$

Použitím rovnic (17) a (18) v rovnici (15) a rozvojem pravé strany v $1/a_s$ máme:

$$\frac{\partial f(n, Q^2)}{\partial a_s} = -\frac{1}{\beta_0 a_s} \left[P^{(0)}(n) + a_s \left(P^{(1)}(n) - \frac{\beta_1}{\beta_0} P^{(0)}(n) \right) \right] f(n, Q^2). \quad (19)$$

Možné je také pracovat s rovnicí (15) po přímém dosazením (17) a (18). Tehdy dostáváme:

$$\frac{\partial f(n, Q^2)}{\partial a_s} = -\frac{(P^{(0)}(n) + a_s P^{(1)}(n))}{a_s(\beta_0 + \beta_1 a_s)} f(n, Q^2) \quad (20)$$

Rovnice (19) a (20) jsou formálně řešena pomocí takzvaných evolučních operátorů $E(n)$:

$$f(n, Q^2) = E(n, a_s(Q^2), a_s(Q_0^2)) f(n, Q_0^2) \quad (21)$$

Zmíníme ještě jeden tvar rovnice se kterým se dá v rámci této metody úspěšně pracovat. Jednoduše budeme uvažovat vyjádření v proměnné Q^2 , docílíme toho tím, že do rovnice (14) dosadíme s ohledem na (4) vyjádření (17). V tomto případě pak pro momenty máme:

$$Q^2 \frac{\partial f(n, Q^2)}{\partial Q^2} = \left(\frac{1}{\beta_0 L} \left(1 - \frac{\beta_1 \ln L}{\beta_0^2 L} \right) P^{(0)}(n) + \left(1 - \frac{\beta_1 \ln L}{\beta_0^2 L} \right)^2 P^{(1)}(n) \right) f(n, Q^2). \quad (22)$$

I v tomto případě bude tato rovnice formálně řešena evolučním operátorem, který nyní závisí na Q^2 a Q_0^2 :

$$f(n, Q^2) = \hat{E}(n, Q^2, Q_0^2) f(n, Q_0^2) \quad (23)$$

Rovnice pro momenty distribučních funkcí (19), (20) a (22) jsou řešitelné analyticky. Nyní uvedeme konkrétní tvary evolučních operátorů.

Evoluční operátor, který řeší rovnici (19) je pro nesingletní případ dán do druhého řádu v poruchové QCD:

$$E_\eta(n, a_s(Q^2), a_s(Q_0^2)) = \left(\frac{a_s(Q^2)}{a_s(Q_0^2)} \right)^{\frac{\gamma_0^{NS}(n)}{2\beta_0}} \left[1 + \frac{a_s(Q^2) - a_s(Q_0^2)}{2\beta_0} \left(\gamma_1^{NS}(n, \eta) - \frac{\beta_1}{\beta_0} \gamma_0^{NS}(n) \right) \right]. \quad (24)$$

Zde η nabývá dvou hodnot a to podle typu uvažované nesingletní distribuční funkce (existují dva typy nesingletních distribučních funkcí). Ve výrazu vystupují tzv. anomální dimenze $\gamma_0^{NS}(n)$ a $\gamma_1^{NS}(n)$, které souvisí s evolučním jádrem: $P_i(n) = -\frac{1}{2} \gamma_i^{NS}(n)$.

Evoluční operátor, který řeší rovnici (20) má v nesingletním případě tvar:

$$E(n, \eta, a_s(Q^2), a_s(Q_0^2)) = \left(\frac{a_s(Q^2)}{a_s(Q_0^2)} \right)^{\frac{\gamma_0(n)}{2\beta_0}} \left(\frac{\beta_0 + \beta_1 a_s(Q^2)}{\beta_0 + \beta_1 a_s(Q_0^2)} \right)^{\frac{\gamma(n, \eta)}{2\beta_1}}. \quad (25)$$

Nakonec pro evoluční operátor řešící rovnici (22), máme v nesingletním případě:

$$\begin{aligned} \hat{E}(n, \eta, Q^2, Q_0^2) = & \left(\frac{L_0}{L} \right)^{\frac{\gamma_0(n)}{2\beta_0}} \exp \left[-\frac{1}{2\beta_0^2} \left(\frac{1}{L_0} - \frac{1}{L} \right) \left(\gamma_1(n, \eta) - \frac{\beta_1}{\beta_0} \gamma_0(n) \right) + \right. \\ & \frac{\beta_1}{2\beta_0^3} \left(\frac{\ln L_0}{L_0} - \frac{\ln L}{L} \right) \gamma_0(n) + \frac{\beta_1}{4\beta_0^4} \left(\frac{1 + 2 \ln L_0}{L_0^2} - \frac{1 + 2 \ln L}{L^2} \right) \gamma_1(z, \eta) \\ & \left. - \frac{\beta_1^2}{54\beta_0^6} \left(\frac{2 + 6 \ln L_0 + 9 \ln^2 L_0}{L_0^3} - \frac{2 + 6 \ln L + 9 \ln^2 L}{L^3} \right) \gamma_1(n, \eta) \right] \end{aligned} \quad (26)$$

kde $L_i = \ln Q_i^2 / \Lambda^2$.

Nyní se budeme soustředit na otázku zpětné transformace získaných momentů distribučních funkcí. Budeme tak chtít přejít zpět k proměnným x a Q^2 . Zpětná Mellinova transformace je dána výrazem:

$$f(x, Q^2) = \frac{1}{2\pi i} \int_C dn x^{-n} f(n, Q^2) \quad (27)$$

zde C obíhá všechny singularity komplexního integrandu.

Distribuční partonové funkce jsou při zvolené počáteční škále Q_0^2 parametrizované

$$xf(x, Q_0^2) = \sum_i A_i x^{\alpha_i} (1-x)^{\beta_i} \quad (28)$$

po Mellinově transformaci nabydou tvar:

$$f(n, Q_0) = \sum_i A_i B(n + \alpha_i - 1, 1 + \beta_i), \quad (29)$$

kde $B(x, y)$ je Eulerova beta funkce. Vynecháme-li nyní argumenty Q^2 a Q_0^2 máme pro řešení (11):

$$f(x, Q^2) = \text{Re} \frac{1}{\pi i} \int_{C_s} dz E(n) R(n) \quad (30)$$

kde

$$R(n) = x^{-n} \sum_i A_i B(z + \alpha_i - 1, 1 + \beta_i), \quad (31)$$

3 Metody inverzní Mellinovy transformace

V předcházející kapitole byl naznačen postup při řešení evoluční rovnice tvaru (11). Rovnici jsme pomocí Mellinovy transformace převedli do N -prostoru, ve kterém ji lze řešit analyticky. Pro zpětnou transformaci do x -prostoru platí (27). V následující části předvedeme dva způsoby postupu při této zpětné transformaci. Oba budou spočívat ve speciální volbě integrační křivky v integrálu (27). Poznamenejme ještě, že na volbě této křivky, pokud obíhá všechny singularity v integrandu, výpočet nezávisí. Situace se však změní, pokud budeme zmíněný integrál počítat numericky. Tehdy už výsledek na volbě integračního oboru závisí.

3.1 Integrovaní po parabolické křivce

V následujícím popíšeme první z variant. Zvolíme integrační obor tak, že v něm bude možné integrand velmi dobře aproximovat sadou ortogonálních polynomů. Volí se parabolická křivka, která bude plně určená partonovou distribuční funkcí v počáteční škále Q_0^2 . Získaný integrál je pak možné spočítat například pomocí Gauss-Laguerrovy metody pro výpočet určitého integrálu. Při této metodě je k dosažení vysoké přesnosti (0,02%) potřeba zhruba čtyř až pěti výpočtů integrandu.

Podívejme se blíže na tuto metodu v nesingletním případě. Křivka podél, které probíhá integrace je dána:

$$z(u) = z_0 + ic_2 \sqrt{u} + \frac{1}{2} c_2^2 c_3 u, \quad u = 0, \dots, \infty \quad (32)$$

kde z_0 jsou minima funkce $F(n)$ na reálné ose. Parametry c_2 a c_3 mají tvar:

$$c_2 = \sqrt{\frac{2F'(z_0)}{F''(z_0)}}, \quad c_3 = \frac{F'''(z_0)}{3F''(z_0)}. \quad (33)$$

$F'(n)$, $F''(n)$ a $F'''(n)$ jsou první tři derivace $F(n)$ podle z . S použitou parametrizací (32) jsou tyto derivace dány:

$$F' = \sum_i G'_i F_i \quad (34)$$

$$F'' = \sum_i (G''_i + G_i'^2) F_i \quad (35)$$

$$F''' = \sum_i (G_i'''' + 3G_i'' G_i' + G_i'^3) F_i \quad (36)$$

$$(37)$$

kde $F_i = x^{-z} A_i B(z + \alpha_i - 1, 1 + \beta_i)$ a

$$G_i = \ln(x^{-z} B(z + \alpha_i - 1, 1 + \beta_i)), \quad (38)$$

$$G_i' = -\ln x + \Psi(z + \alpha_i - 1) - \Psi(z + \alpha_i + \beta_i), \quad (39)$$

$$G_i'' = \Psi'(z + \alpha_i - 1) - \Psi'(z + \alpha_i + \beta_i), \quad (40)$$

$$G_i''' = \Psi'(z + \alpha_i - 1) - \Psi''(z + \alpha_i + \beta_i), \quad (41)$$

Zde Ψ , Ψ' a Ψ'' jsou polygamní funkce. S touto parametrizací máme pro integrál (30):

$$I = \frac{c_2}{2\pi} \int_0^\infty \frac{du}{\sqrt{u}} e^{-u} \operatorname{Re} [e^u (1 - ic_2 c_3 \sqrt{u}) E(z(u)) F(z(u))]. \quad (42)$$

Funkce $u^{-1/2} e^{-u}$ jsou vážené funkce Laguerrových polynomů $L_k^{-1/2}(x)$ a integrál můžeme aproximovat pomocí Gauss-Laguerrova vzorce:

$$I \approx \frac{c_2}{2\pi} \sum_{j=1}^k w_j \operatorname{Re} [e^{u_j} (1 - ic_2 c_3 \sqrt{u_j}) E(z(u_j)) F(z(u_j))]. \quad (43)$$

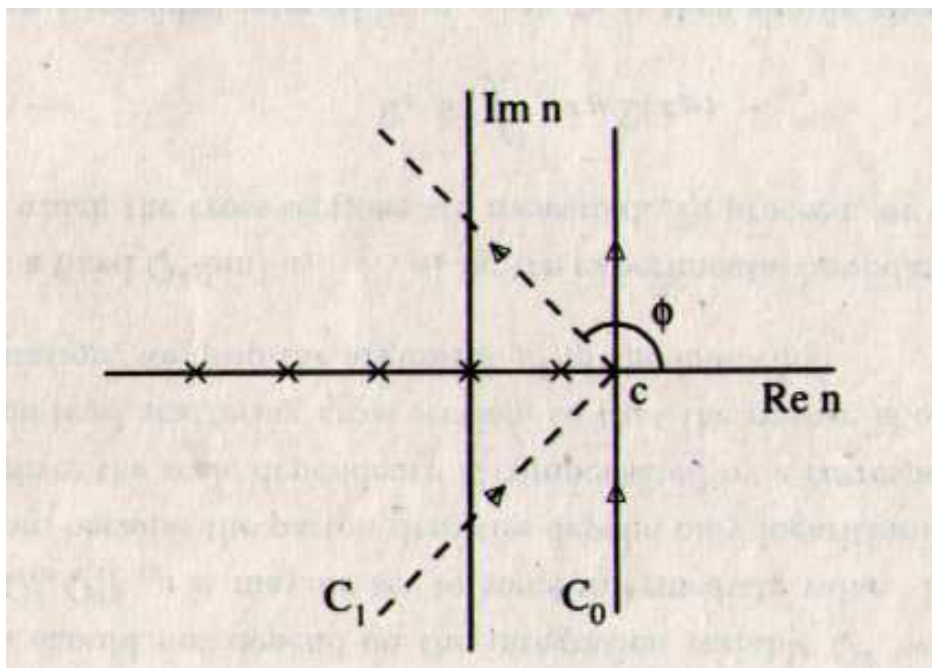
kde u_j jsou kořeny $L_k^{-1/2}(x)$ a váhy w_j jsou dány:

$$w_j = \frac{\Gamma(n + \frac{1}{2})}{n!(n+1)^2} \frac{u_j}{(L_{k+1}^{-1/2}(u_j))^2}. \quad (44)$$

3.2 Integrace po přímce

Nyní přistoupíme ke druhé možnosti numerické realizace inverzní Mellinovy transformace. Představíme ji v následující části. V kapitole 4. bude popsán postup při řešení zcela konkrétní rovnice. Ten bude spočívat na zpětné transformaci popsané právě v této kapitole.

Pokud funkce $f(x)$ rychle ubývá pro $x > 1$ a pokud je po částech spojitá pro $x > 0$, potom můžeme mít inverzní Mellinova transformace tvar:



Obrázek 1: Dvě možné volby integrační křivky při inverzní Mellinově transformaci. Důležité je, že obě leží vpravo od oblasti singularit, které jsou v obrázku vyznačeny křížky.

$$f(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} dn x^{-n} f(n), \quad (45)$$

kde reálné číslo c je zvoleno tak, aby byl integrál $\int_0^1 x^{c-1} f(x)$ absolutně konvergentní. To vyžaduje zvolit c tak, aby leželo napravo od singularity n_{max} , vzdálené na reálné ose nejvíc vpravo. Křivka užitá v integrálu (45) je zobrazena na Obrázku 1. a je označena jako C_0 . Ve stejném obrázku je ilustrativně naznačena ještě další křivka, podél které je možné integraci vést.

Je užitečné přepsat rovnici (45) jako integraci přes reálnou proměnou. Budeme se soustředit na funkce splňující $f^*(n) = f(n^*)$, kde $*$ značí komplexní sdružení. Potom je snadné ukázat, že funkce $f(x)$ splňuje, pro integrační obor určený souřadnicí c a úhlem ϕ , následující rovnost:

$$f(x) = \frac{1}{\pi} \int_0^\infty dz \operatorname{Im} \left[\exp(i\phi) x^{-c-z} \exp(i\phi) F(n = c + z \exp(i\phi)) \right], \quad (46)$$

Tento integrál nezávisí ani na c ani ϕ . Nicméně pro numerický výpočet může být vhodná volba těchto parametrů velmi užitečná. Při konkrétních požadavcích na přesnost výpočtu je tato vhodná dokonce nezbytná. Navíc, pokud je $f(n)$ analytická známá funkce, je výhodné volit $\phi > \pi/2$, speciálně pokud tato funkce neubývá příliš rychle pro $|n| \rightarrow \infty$. Potom faktor $\exp(z \log \frac{1}{x} \cos \phi)$ exponenciálně tlumí integrand (který pro malá x značně osciluje), což při numerické implementaci rovnice (46) umožňuje volit menší horní mez z_{max} .

Nicméně, obecně musí být momenty partonových účinných průřezů vypočteny numericky. Tehdy je výhodné volit integrační křivku C_0 , tedy $\phi = \pi/2$ a rovnice (46) vede na:

$$F(x) = \frac{1}{\pi} \int_0^\infty dz \operatorname{Re} \left[x^{-c-iz} F_{n=c+iz} \right]. \quad (47)$$

Při výpočtu integrálu (45) tedy volíme integraci po přímce kolmé na reálnou osu a procházející na této ose bodem c .

4 Numerický výpočet evoluční rovnice

V následující části se pokusíme aplikovat metodu inverzní Mellinovy transformace na konkrétní příklad. Uvažujme reakci $\nu_\mu(\bar{\nu}_\mu) + N \rightarrow \mu^- + X$. Ta odpovídá rozptylu mionového neutrina (anti-neutrina) na nukleonu N . Diferenciální účinný průřez této reakce je dán následující formulí:

$$\frac{d\sigma^{(\nu/\bar{\nu})p}}{dxdy} = \frac{G_F^2 S}{2\pi} \left[F_2^{(\nu/\bar{\nu})p}(x, Q^2) \left(\frac{1 + (1-y)^2}{2} \right) \pm x F_3^{(\nu/\bar{\nu})p}(x, Q^2) \left(\frac{1 + (1-y)^2}{2} \right) \right] \quad (48)$$

kde G_F je Fermiho vazbová konstanta. Srovnáním této formule s diferenciálními účinnými průřezy tohoto procesu v partonového modelu dostáváme pro strukturní funkce F_3 :

$$xF_3^{\nu p} = 2x[s(x) + d(x) - \bar{u}(x) - \bar{c}(x)] \quad (49)$$

$$xF_3^{\bar{\nu} p} = 2x[u(x) + c(x) - \bar{s}(x) - \bar{d}(x)] \quad (50)$$

Uvažujme následující výraz:

$$\frac{1}{2} (F_3^{\nu p} + F_3^{\bar{\nu} p}) = (u + d + s + c - \bar{u} - \bar{d} - \bar{s} - \bar{c}) \doteq \frac{1}{2} (F_3^{\nu p} + F_3^{\nu n}) \equiv F_3^{\nu N} \quad (51)$$

Ten lze určit experimentálně neboť je kombinací diferenciálních účinných průřezů:

$$F_3^{\nu N} \sim \frac{d\sigma^{\nu N}}{dxdy} - \frac{d\sigma^{\bar{\nu} N}}{dxdy} \quad (52)$$

Je patrné, že výraz $F_3^{\nu N}$ je kombinací valenčních distribučních funkcí. Z linearity evoluční rovnice pro nesingletní distribuční funkce vyplývá, že i funkce $F_3(x, Q^2)$ je jejím řešením. Tedy:

$$\frac{dF_3(x, Q^2)}{d \ln Q^2} = \frac{\alpha_s(Q^2)}{2\pi} \int_x^1 \frac{dy}{y} P_{qq}^{(0)} \left(\frac{x}{y} \right) F_3(y, Q^2), \quad (53)$$

V následující části naznačíme základní kroky při řešení takové rovnice. Tento postup bude analogický s algoritmem řešení popsaném v kapitolách 2. a 3.2. Ještě předtím ale stručně připomeneme základní schéma postupu při řešení evolučních rovnic.

K tomu abychom mohli řešit rovnici typu (1) je v první řadě zapotřebí znalost hraniční podmínky. Ta zahrnuje informace o závislosti partonové distribuční funkce na x při nějaké počáteční škále Q_0^2 . Výběr této počáteční škály je poměrně delikátní záležitostí, neboť hledané řešení rovnice na výběru Q_0^2 pochopitelně závisí. Obvykle se Q_0^2 volí v rozsahu několika GeV . Poruchová teorie tyto počáteční podmínky neumožňuje vypočítat, musí být získány z experimentálních dat.

Standardní analýza experimentálních dat se provádí v následujících krocích: provede se volba počáteční škály Q_0^2 , zvolí se vhodná parametrizace počáteční podmínky tj. funkce $f(x, Q_0^2)$. Příkladem může být:

$$f(x, Q_0^2) = Ax^\alpha(1-x)^\beta(1+\gamma x) \quad (54)$$

kde A, α, β, γ jsou volné parametry. Pro zvolenou sadu těchto parametrů spolu s hodnotou hlavního parametru QCD Λ , vstupujícím do $\alpha_s(Q/\Lambda)$ řešíme rovnice (1). Takto získáme $f(x, Q^2)$ pro všechna x a Q^2 . Tato řešení se pak porovnávají s experimentálními daty a na základě těchto srovnání se hodnoty uvedených parametrů dále zpřesňují. Výsledkem je sada parametrů A, α, β, γ určující hraniční podmínku tak, že řešením evolučních rovnic s touto podmínkou jsou hledané distribuční funkce.

Obraťme nyní pozornost na řešení evoluční rovnice (53). Z důvodu zachování značení přeznačíme řešení takové rovnice na $f(x, Q^2)$. Uvažujme tedy evoluční rovnici:

$$\frac{df(x, Q^2)}{d \ln Q^2} = \frac{\alpha_s(Q^2)}{\pi} P_{qq}^{(0)}(Q^2) \otimes f(x, Q^2) \quad (55)$$

kde

$$P_{qq}^{(0)}(x) = \frac{4}{3} \left[\frac{1+x^2}{1-x} \right]_+ \quad (56)$$

Po použití Mellinových momentů máme:

$$\frac{dq(n, Q^2)}{d \ln Q^2} = \frac{\alpha_s(Q^2)}{\pi} P_{qq}^{(0)}(Q^2) \cdot q_{NS} \quad (57)$$

Tuto rovnici již umíme vyřešit analyticky. Takové řešení, včetně hraničních podmínek hledejme ve tvaru:

$$f(n, Q^2) = E(n, Q^2, Q_0^2) f(n, Q_0^2) \quad (58)$$

V rámci našeho numerického výpočtu se omezíme na nejnižší řády známých rozvoju všech veličin. V takovém případě pro evoluční operátor $E(n, Q^2, Q_0^2)$ máme (viz.(24)):

$$E(n, a_s(Q^2), a_s(Q_0^2)) = \left(\frac{a_s(Q^2)}{a_s(Q_0^2)} \right)^{-\frac{P^{(0)}(n)}{4\beta_0}}. \quad (59)$$

V evolučním operátoru vystupuje moment větvicí funkce $P_{qq}^{(0)}$. Ten je pochopitelně možné počítat numericky z definice, nicméně je možné pro něj nalézt explicitní vyjádření. To poněkud zjednodušuje algoritmus výpočtu a zároveň zatěžuje výsledek menší chybou. Odbočíme tedy na chvíli a pokusíme se nalézt momenty pro úplnost všech větvicích funkcí.

Připomeňme si větvicí funkce QCD:

$$P_{qq}^{(0)}(x) = P_{\bar{q}\bar{q}}^{(0)}(x) = \frac{4}{3} \left[\frac{1+x^2}{1-x} \right]_+ \quad (60)$$

$$P_{Gq}^{(0)}(x) = P_{G\bar{q}}^{(0)}(x) = \frac{4}{3} \left[\frac{1+(1-x)^2}{x} \right] \quad (61)$$

$$P_{qG}^{(0)}(x) = P_{\bar{q}G}^{(0)}(x) = \left[\frac{x^2 + (1-x)^2}{2} \right] \quad (62)$$

$$P_{GG}^{(0)}(x) = 6 \left\{ \left[\frac{x}{1-x} \right]_+ + \frac{1-x}{x} + x(1-x) + \left(\frac{33-2n_f}{36} - 1 \right) \delta(1-x) \right\} \quad (63)$$

Tyto větvičí funkce jsou stejné pro všechny kvarky q_i, \bar{q}_i . Pro vyjádření této skutečnosti bylo použito symbolů q, \bar{q}

Ve větvičích funkcích (60) a (63) vystupuje " + " distribuce, která je definována:

$$[f(x)]_+ \equiv \lim_{\beta \rightarrow 0} \left(f(x) \Theta(1-x-\beta) - \delta(1-x-\beta) \int_0^{1-\beta} f(y) dy \right) \quad (64)$$

kde $\Theta(x)$ a $\delta(x)$ jsou skoková funkce a Diracova δ -funkce. Její působení na testovací funkci je následující:

$$\int_0^1 [f(x)]_+ g(x) dx \equiv \int_0^1 f(x) (g(x) - g(1)) dx. \quad (65)$$

Nyní již není složité za pomoci formule (65) a definice Mellinových momentů (13) spočítat momenty obecné mocniny x^r :

$$F(n) = \int_0^1 x^{n-1} x^r dx = \left[\frac{x^{n+r}}{n+r} \right]_0^1 = \frac{1}{n+r}, \quad \text{Re } n > -r. \quad (66)$$

a libovolné " + " distribuce tvaru $[x^r/(1-x)]_+$:

$$F(n) = \int_0^1 x^{n-1} \left[\frac{x^r}{(1-x)} \right]_+ dx = \int_0^1 \frac{x^{n+r-1} - x^r}{1-x} dx. \quad (67)$$

S využitím vztahu:

$$\int_0^1 \frac{1-t^{z-1}}{1-t} dt = \psi(z) + \gamma_E, \quad \text{Re } z > 0, \quad (68)$$

kde γ_E je Eulerova konstanta,

můžeme po jednoduché úpravě pro momenty v (67) psát:

$$F(n) = \int_0^1 \left(\frac{1-x^r}{1-x} - \frac{1-x^{n+r-1}}{1-x} \right) dx = \psi(r+1) - \psi(n+r), \quad \text{Re } n > -r \quad (69)$$

S použitím těchto výrazů pak pro momenty větvičích funkcí máme:

$$P_{qq}^{(0)}(n) = \frac{4}{3} (\psi(1) - \psi(n) + \psi(3) - \psi(n+2)), \quad (70)$$

$$P_{Gq}^{(0)}(x) = \frac{4}{3} \left(\frac{2}{n-1} - \frac{2}{n} + \frac{1}{n+1} \right), \quad (71)$$

$$P_{qG}^{(0)}(x) = \frac{1}{n+2} - \frac{1}{n+1} + \frac{1}{2n}, \quad (72)$$

$$P_{GG}^{(0)}(x) = 6 \left[\psi(2) - \psi(n+1) + \frac{1}{n-1} - \frac{1}{n} + \frac{1}{n+1} - \frac{1}{n+2} + \frac{33-2n_f}{36} - 1 \right]. \quad (73)$$

Vraťme se nyní zpět k popisu algoritmu inverzního transformování. Nyní již můžeme zapsat momenty hledané distribuční funkce:

$$f(n, Q^2) = x^{-n} \left(\frac{a_s(Q^2)}{a_s(Q_0^2)} \right)^{-\frac{\psi(1)-\psi(n)+\psi(3)-\psi(n+2)}{3\beta_0}} AB(n + \alpha - 1, 1 + \beta) \quad (74)$$

Vše je tedy připraveno pro zpětnou transformaci. Jejím provedením pak pro hledané distribuční funkce nalezneme:

$$f(x, Q^2) = \frac{1}{\pi} \int_0^\infty dz \operatorname{Re} \left[x^{-c-iz} \left(\frac{a_s(Q^2)}{a_s(Q_0^2)} \right)^{-\frac{\psi(1)-\psi(n)+\psi(3)-\psi(n+2)}{3\beta_0}} AB(n + \alpha - 1, 1 + \beta) \right] \quad (75)$$

Do tohoto integrálu vstupuje celkem pět parametrů: A, α, β, Q_0^2 a c . První čtyři parametrizují hraniční podmínku (54) (v našich výpočtech je $\gamma = 0$). Jejich určení je pro řešení evoluční rovnice klíčové. Poruchová teorie je neumožňuje spočítat, zjištěny jsou na základě srovnání s experimentálními daty. Pokud jde o parametr c , na něm výpočet integrálu (75) nezávisí, pokud je zvolen tak, že leží napravo od singularity umístěné na reálné ose nejvíc vpravo. Singulárními body integrandu v (75) jsou reálná čísla $\{\dots, -3, -2, -1, 0\}$. Zodpovědný je za ně moment větvičí $P_{qq}^{(0)}(x)$. ψ funkce v něm obsažená totiž v těchto bodech není definována. V případě naší rovnice má tedy tento parametr význam vzdálenosti integrační křivky od nejkrajnější, vpravo umístěné singularity. Obecně je to bod ve kterém integrační křivka v komplexní rovině protíná reálnou osu.

Na poloze integrační křivky za výše zmíněných předpokladů výpočet integrálu (45) nezávisí. To je ovšem pravda jen do té doby, do kdy výpočet tohoto integrálu neprovádíme numericky. Při numerickém výpočtu už výsledek na volbě c závisí. Přivede nás k tomu zkoumání vlivu parametru c na integrovanou funkci. Průběhy integrandu pro různá c jsou znázorněna Obrázku 2. Parametrizace okrajové podmínky byla zvolena následující: $A = 8.6$, $\alpha = 0.85$, $\beta = 3.7$, $Q_0^2 = 5 \text{ GeV}^2$, $\Lambda = 0.377 \text{ MeV}$. Integrací bychom obdrželi distribuční funkci v bodě $x = 0.3$ a $Q^2 = 50 \text{ GeV}^2$. Je patrné, že se vzrůstajícím c integrovaná funkce značně osciluje a to na stále větší oblasti definičního oboru. Je jasné, že pro velká c je potřeba volit úměrně velký integrační obor, což u numerického výpočtu naráží na problém časové náročnosti. Navíc integrací takto oscilujících funkcí značně narůstá chyba výpočtu. Abychom tuto chybu co nejvíce eliminovali, budeme volit c co nejbližší singularity. Tehdy je průběh integrované nejméně komplikovaný – dochází zde nejméně ke vzájemnému rušení kladných a záporných příspěvků k počítanému integrálu. Pro numerický výpočet je tedy tato volba nejspokojnější.

Při numerickém výpočtu je nutné nahradit nekonečno v integrálu nějakou konečnou hodnotou. Musíme ovšem počítat s tím, že se tím dopouštíme určité chyby. Při volbě horní meze je možné využít zjištěných průběhů integrandu. Volíme takovou horní mez, abychom integrovaly jen přes relevantní část integrandu. Integrovaná funkce totiž poměrně rychle ubývá a nevhodnost zatížení výpočtu integrací v oblasti, kde jsou příspěvky k integrálu velmi malé je nasnadě. Příslušná volba horní meze musí samozřejmě korespondovat s volbou c .

5 Popis programu užitého při výpočtu evoluční rovnice

V následujícím textu bude nastíněna činnost programu, který umožňuje numericky řešit evoluční rovnice QCD. Program byl napsán v jazyce C++, zdrojový kód je připojen v příloze. Jeho činnost spočívá ve dvou základních dovednostech. První z nich je výpočet distribuční funkce pro libovolné x a Q^2 , druhou je pak fitování obecného řešení evoluční rovnice na naměřené hodnoty distribuční funkce, dodané experimentem. Obě tyto činnosti nyní popíšeme trochu podrobněji.

- Princip a činnost části programu, umožňující výpočet distribuční funkce je velmi jednoduchý. Program si žádá sadu parametrů A, α, β, Q_0 , polohu integrační křivky c , parametr Λ a pochopitelně hodnotu x a Q . Vnitřními parametry jsou potom horní mez integrálu zpětné transformace a velikost absolutní a relativní chyby výpočtu. Tyto hodnoty se během základních početních operací většinou nemění. Po dodání všech těchto parametrů se spustí algoritmus, jehož základní kroky byly popsány v kapitole 4. Výsledkem je hodnota distribuční funkce v žádaných bodech x a Q . Při analýze výsledku je obvyklé zpracovat grafy s naznačenými průběhy jedné proměnné při fixování druhé. Tento program umožňuje dodat dostatečný počet příslušných funkčních hodnot pro tvorbu těchto grafů v obou variantách. (tj. průběh distribuční funkce v proměnné x při pevném Q^2 a naopak). Všechny výsledky jsou ukládány do souborů a uchovány tak pro jejich případně zpracování dalšími aplikacemi.

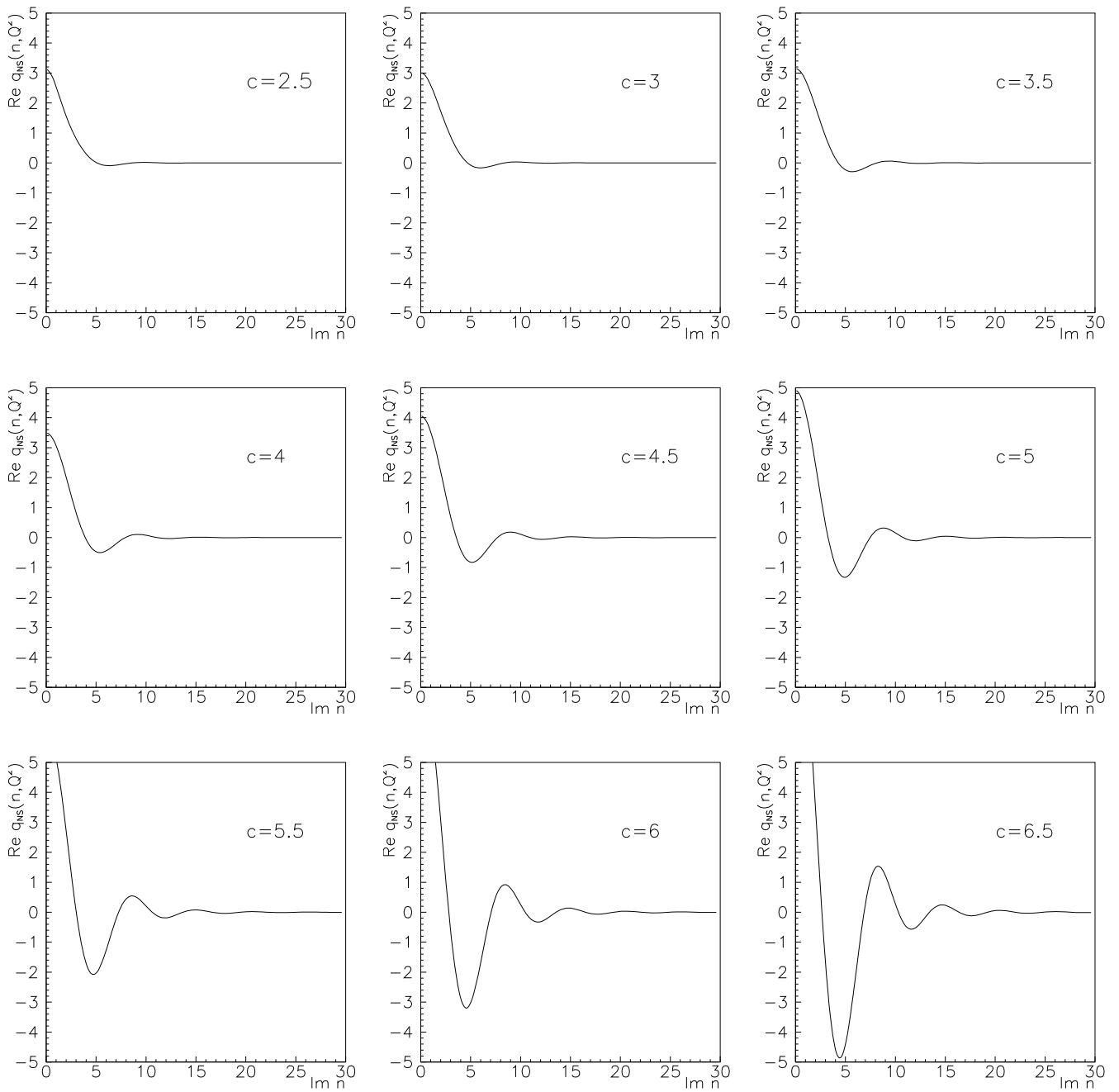
Součástí vypracovaného programu je i několik testů ověřujících správnost numerického výpočtu inverzní Mellinovy transformace. Jejich popis fungování spolu s výsledky budou popsány ve zvláštní kapitole 6.

- Druhá významná součást programu má na starosti fitování obecného řešení evoluční rovnice na konkrétní experimentální data. To se děje systematickým vyhledáváním a postupným zpřesňováním parametrů A, α, β a Λ , které parametrizují okrajovou podmínku evoluční rovnice a parametru Λ . Snahou je nalézt takové parametry, aby řešení rovnice s hraniční podmínkou určenou těmito parametry byla těmi řešeními, která v příslušných x a Q^2 odpovídají experimentálním hodnotám. Do vyhledávacích manévřů stále zahrnujeme i volný parametr Λ . Princip vyhledávání těchto parametrů je postaven na minimalizaci funkce

$$\chi^2(A, \alpha, \beta, \Lambda) = \sum_{k=1}^K \left(\frac{e_k - xf(A, \alpha, \beta, \Lambda)}{\sigma_k} \right)^2, \quad (76)$$

kde e_k a σ_k jsou naměřené hodnoty a jejich chyby, $xf(A, \alpha, \beta, \Lambda)$ jsou napočítané hodnoty distribuční funkce s příslušnými parametry. Hodnoty x a Q^2 použité při výpočtu distribuční funkce si v každém sčítanci odpovídají s hodnotami, za jakých byly naměřeny hodnoty e_k . Počet měření K bývá obvykle mnohem větší, než je počet parametrů n vstupujících do výpočtu hodnot $xf(x, Q^2)$.

Je tedy potřeba zvolit počáteční sadu parametrů A, α, β, Q_0 a Λ a pro ni napočítat χ^2 . Program potom funkci χ^2 iterativně minimalizuje. V každém kroku dochází ke zpřesnění všech parametrů. Program umožňuje vývoj funkce χ^2 i příslušné hodnoty parametrů v každém kroku sledovat.



Obrázek 2: Integrandy inverzní Mellinovy transformace pro různé polohy integrační křivky.

6 Ověření správnosti inverzní Mellinovy transformace

Po sepsání výše zmíněného programu bylo vhodné ověřit, že skutečně počítá to co počítat má. Existuje několik postupů, kterými lze správnost výpočtů kontrolovat nebo přímo ověřovat.

Pokud jde o přímé testování správnosti inverzního integrování, provedeme srovnání nějaké zvolené funkce s výsledkem inverzní Mellinovy transformace momentů této funkce. Očekávat budeme stejné hodnoty. Za testovací funkci byla zvolena parametrizovaná hraniční podmínka (54) s parametry zvolenými tak, aby práce s touto funkcí simulovala výpočty prováděné při řešení naší původní evoluční rovnice. Výhodnost takové volby ve vztahu k našemu řešenému problému se ozřejmí, pokud si uvědomíme, že volbou $Q = Q_0$ ve výrazu (75) dostáváme:

$$f(x, Q_0^2) = \frac{1}{\pi} \int_0^\infty dz \operatorname{Re} \left[x^{-c-iz} AB(n + \alpha - 1, 1 + \beta) \right]. \quad (77)$$

Inverzní transformování momentů hraniční podmínky evoluční rovnice s vhodnými parametry je tedy speciálním případem obecné inverzní transformace (75) vedoucí k řešení naší evoluční rovnice. Pro náš test byla zvolena funkce:

$$g(x) = 8.6x^{8.5}(1-x)^{3.94} \quad (78)$$

Při výpočtu srovnávací funkce bylo použito parametrů $A = 8.6$, $\alpha = 8.5$, $\beta = 3.94$, $c = 2.5$, $\Lambda = 0.377 \text{ MeV}$ a $Q = Q_0$. Výsledky jsou zaznamenány v Tabulce 1. V prvním sloupci jsou ve vybraných bodech hodnoty funkce $g(x)$, ve druhém jsou pak v odpovídajících bodech invertované momenty funkce $g(x)$. Odhadované chyby těchto výsledků jsou k nalezení ve třetím sloupci. V posledním sloupci jsou pak skutečné chyby, kterých jsme se při inverzní transformaci dopustili tj. |Výsledek 1-Výsledek 2|.

Uvedeme ještě jeden způsob možnosti ověření, tentokrát jen dílčího výpočtu při inverzní Mellinově transformaci. Týká se napočítaných momentů (74). Ještě než dojde k jejich zpětné transformaci, je vhodné ověřit, že mají odpovídající vlastnosti.

Z definice větvičí funkce $P_{qq}^{(0)}(x)$ snadno zjistíme, že:

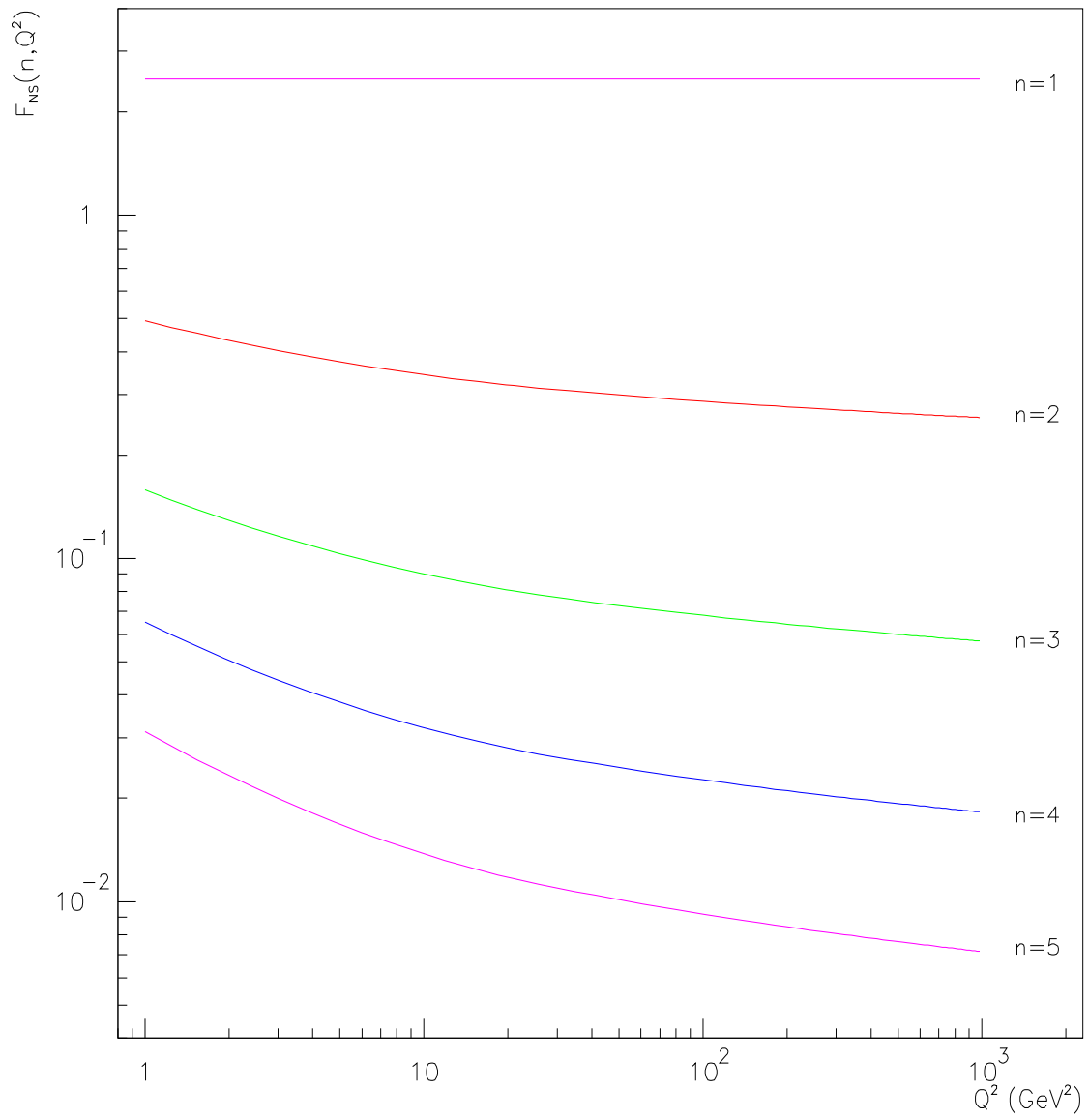
$$P_{qq}^{(0)}(1) = \int_0^1 P_{qq}^{(0)}(x) dx = 0. \quad (79)$$

První moment větvičí funkce $P_{qq}^{(0)}(x)$ je tedy nulový, tudíž první moment distribuční funkce by měl být konstantní.

Pro přirozené momenty této větvičí funkce větší než jedna snadno nalezneme:

$$P_{qq}^{(0)}(n) = \frac{4}{3} \left(1 + \frac{1}{2} - \sum_{i=1}^{n-1} \frac{2}{i} - \frac{1}{n} - \frac{1}{n+1} \right). \quad (80)$$

Tento výraz je vždy záporný, tedy exponent výrazu $a_s(Q^2)/a_s(Q_0^2)$ v rovnici (74) je kladný. Vzhledem k tomu, že je funkce $a_s(Q^2)/a_s(Q_0^2)$ v proměnné Q^2 klesající (viz (4)), měly by být (pro všechna přirozená $n > 1$) v proměnné Q^2 klesající i momenty distribučních funkcí.



Obrázek 3: Prvních pět přirozených momentů distribuční funkce. Podrobný popis je v textu.

Přesně tyto vlastnosti mají mnou zjištěné výsledky. Graficky jsou shrnuty v Obrázku 3. Zde je vyneseno prvních pět přirozených momentů distribuční funkce. Při výpočtu byly zvoleny následující parametry: $A = 8.6$, $\alpha = 0.85$, $\beta = 3.94$, $Q_0^2 = 5 GeV^2$, $\Lambda = 0.337 MeV$.

7 Krokové metody minimalizace v mnoha proměnných

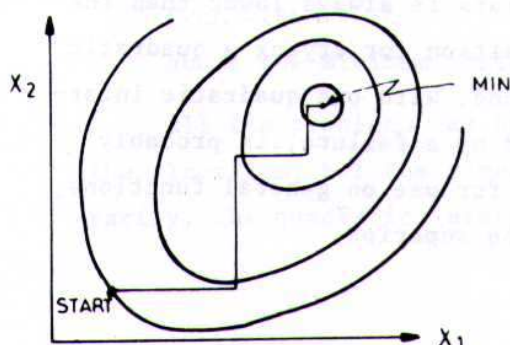
Metody minimalizace ve více dimenzích jsou většinou zobecněnými variantami jednodimenzionálních případů. Nežřídka však tato zobecnění sebou přinášejí problémy, které nás nutí volit jiné cesty.

7.1 Síťové hledání a náhodné hledání

Příkladem enormního zvýšení složitosti při přechodu do prostoru vyšších dimenzí je použití metody uzlů pro funkce více proměnných. Pro lokalizaci minima s přesností 1% v případě jedné proměnné je potřeba 100 výpočtů funkce. V případě funkce 10 proměnných je už vyžadováno 10^{20} napočtených bodů. U minimalizací funkcí platí obecně, že metoda, která se osvědčila v jednodimenzionálním případě už nemusí být rozšířením do více rozměrů přijatelná. Dokonce ani pokud tyto metody vylepšíme o zpřesnění vyhledávacích oblastí v jednotlivých proměnných. Tehdy se jejich efektivita zvyšuje jen velmi pomalu a proto je vhodnější obrátit pozornost na jiné, vhodnější techniky.

7.2 Jednopermetrická variace

Jelikož námi hledaný extrém považujeme za stacionární bod v n proměnných x_i , bude nulovat všech n prvních derivací $\partial F/\partial x_i$. Bylo by možné pokusit se při jeho hledání nulovat každou derivaci zvlášť.

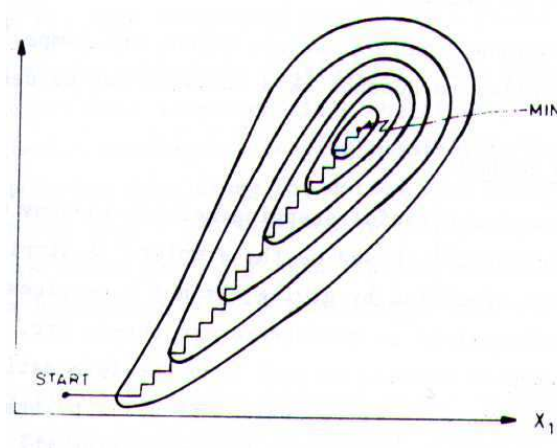


Obrázek 4: Schématické znázornění postupu při minimalizaci metodou jednopermetrické variace v případě funkce dvou proměnných.

Takový postup představuje poměrně starou metodu jednopermetrické variace, kdy se pokoušíme hledat minimum vzhledem k jedné proměnné a to za pomoci některé jednodimenzionální techniky. Po skončení minimalizace vzhledem k jedné proměnné pochopitelně nenacházíme minimum vzhledem k jiným proměnným a proto je potřeba celou proceduru stále opakovat. Takový algoritmus nicméně obvykle konverguje. Ilustrovat to můžeme na příkladu funkce dvou proměnných. Její schéma je znázorněné na Obrázku 4. Uzavřené křivky značí oblasti se stejnou funkční hodnotou, rovné úseky značí postup při minimalizaci v jednotlivých krocích. V prvním kroku hledáme minimum vzhledem k proměnné x_1 , ve druhém vzhledem k proměnné x_2 , ve třetím opět vzhledem k x_1 atd. V tomto případě metoda konverguje po pouhých čtyřech jednopermetrických minimalizacích.

Uvažujme nyní funkci reprezentovanou schématem na Obrázku 5. V tomto případě naše metoda postupuje velmi pomalu. Je to způsobeno "rovnými" úseky spojující body se stejnými funkčními hodnotami. Takové chování při procesu minimalizace může být považováno za neúnosně pomalé.

V následující části popíšeme několik úspěšných vylepšení zaměřených na eliminaci právě takového chování.



Obrázek 5: Schematické znázornění méně příznivé varianty při minimalizaci metodou jednoparametrické variace.

7.3 Rosenbrockova metoda

Rosenbrockův algoritmus začíná v podstatě úplně stejně jako jednoparametrická minimalizace. Poté co proběhne úplný cyklus se všemi proměnnými, zavede se nový výchozí souřadný systém a to tak, že jedna z os bude určena vektorem z počátečního bodu minimalizace do koncového bodu jednoho cyklu. Směr tohoto vektoru představuje pro pozdější výpočet vhodnou volbu. V případě výše uvedeného "úzkého údolí" probíhají jednotlivé kroky minimalizace víceméně podél stěn a úspěšně se tak vyhýbá "cik-cak" chování. V dalším cyklu je jednoparametrická minimalizace provedena se sadou proměnných nového souřadného systému.

Rosenberckova metoda je vcelku účinná a bývá stabilní. Se vzrůstajícím počtem proměnných, jeho výkonnost klesá. Navíc ve srovnání s metodami popsányi později konverguje pomaleji.

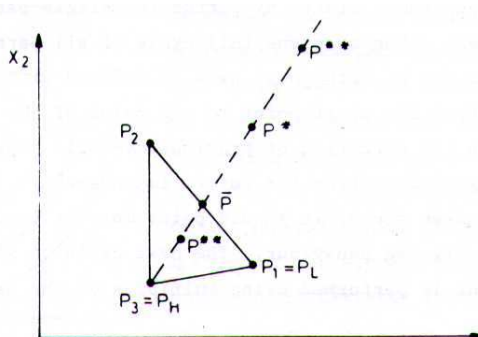
7.4 Metoda simplexu

Jedna z nejúspěšnějších metod užívaná v případě mnoha proměnných pochází z rukou Neldera a Meada a je založena na útvaru simplexu. Simplex je n -dimensionální útvar určený svými $n + 1$ vrcholy. Je to kupříkladu trojúhelník ve dvou dimenzích, čtyřstěn ve třech dimenzích apod. Simplex dal název této metodě proto, že s funkcí v každém kroku pracujeme v jejích $n + 1$ hodnotách. Pro lepší představu se blíže podíváme na užití této metody ve dvou dimenzích. Zvolme tři simplexní body (třeba i náhodně) a vypočítejme hodnotu funkce v těchto bodech. Označme P_H bod s největší funkční hodnotou (pro náš případ minimalizace bod nejméně přijatelný) a P_L bod naopak s nejnižší hodnotou. Budiž \bar{P} těžištěm všech bodů simplexu kromě bodu P_H , tedy:

$$\bar{P} = \frac{1}{n} \left(\sum_{i=1}^{n+1} P_i - P_H \right) \quad (81)$$

Z původního simplexu vytvoříme nový záměnou bodu P_H nějakým vhodnějším bodem. Pokusíme se tento bod nalézt zobrazením bodu P_H ve středové symetrii podle bodu \bar{P} , tedy $P^* = \bar{P} + (\bar{P} - P_H)$. Pokud $F(P^*) < F(P_L)$, volíme bod $P^{**} = \bar{P} + 2(\bar{P} - P_H)$. Pokud $F(P^*) > F(P_H)$, volíme bod $P^{**} = \bar{P} - 1/2(\bar{P} - P_H)$. Nejvhodnější bod potom v simplexu pro další krok nahrazuje bod P_H . Pokud takový nenalezneme, vytvoříme kolem bodu P_L nový simplex.

Tuto metodu je možné při hledání nového bodu na přímce $P_H\bar{P}$ upravovat volbou číselného faktoru vystupujícího ve výrazu pro P^{**} . Důležité je, aby tento nový bod nebyl příliš blízko \bar{P} . V tom případě by náš simplex kolapsoval do přímky (nebo obecně do nadplochy v n rozměrném prostoru), což by znemožnilo další iterace.



Obrázek 6: Znázornění postupu při minimalizaci metodou simplexu. Podrobný popis je v textu.

Simplexní algoritmus je vytvořen tak, aby postupoval v co největších krocích. Je tedy poněkud méně citlivý v případě mělkých minim. Výhodou je však nutnost jen několika výpočtů funkce, obvykle jednoho nebo dvou v každé iteraci. V každém kroku se postupuje v nejvhodnějším směru - od nejvyšší hodnoty do průměru nejnižších hodnot.

Vhodným kritériem konvergence pro simplexní metodu je rozdíl $F(P_H) - F(P_L)$. Iterace jsou ukončeny pokud je tento rozdíl menší než nějaká předvolená hodnota. Konečným krokem výpočtu je potom výpočet funkce v \bar{P} .

7.5 Gradientní metody

7.5.1 Výpočet derivací

Za gradientní metody budeme považovat takové, které užívají informace z velmi malé oblasti proměnných k určení relativně vzdálených bodů. Nemusí to nutně znamenat, že sledují směr gradientu, ale gradient nebo vyšší derivace jsou při výpočtu nějakým způsobem užity. Mnoho výkonných algoritmů níže uvedených vyžadují znalost derivace funkce, minimalizační program tedy musí být schopen na základě konečných rozdílů odhadnout derivaci minimalizované funkce.

První derivace může být odhadnuta z:

$$\frac{\partial F}{\partial x}|_{x_0} \approx \frac{F(x_0 + d) - F(x_0)}{d} \quad (82)$$

kde d je malé. Chyba dána nejnižším rozvojem Taylorova rozvoje pak bude dána:

$$\delta \approx \frac{d}{2} \frac{\partial^2 F}{\partial x^2}|_{x_0}. \quad (83)$$

Pochopitelně je výhodné uvažovat d co nejmenší, ale zároveň tak velké, aby chyba při výpočtu F nebyla větší než právě představené δ . Protože druhé derivace nejsou známy, není možné nalézt optimální d . Je tedy nutné zapojit intuici a šikovně ho odhadnout.

Mnohem bezpečnější bude zvolit body symetricky na každé straně x_0 :

$$\frac{\partial F}{\partial x}|_{x_0} \approx \frac{F(x_0 + d) - F(x_0 - d)}{2d}, \quad (84)$$

v tomto případě je chyba δ ve druhém řádu nulová a člen nejnižšího řádu rozvoje je úměrný třetí derivaci. Nevýhodou této metody je, že vyžaduje $2n$ volání funkce pro odhad prvních n derivací narozdíl od asymetrické metody, kdy jich je ke stejnému účelu zapotřebí jen $n + 1$ (nebo jen n , pokud již známe $F(x_0)$). Výhodou je naopak fakt, že druhou derivaci dostáváme jako mezivýsledek výpočtu první derivace:

$$\frac{\partial F}{\partial x}|_{x_0} \approx \frac{F(x_0 + d) + F(x_0 - d) - 2F(x_0)}{d^2}, \quad (85)$$

7.5.2 Nejstrmější spád

Pokud známe první derivace minimalizované funkce, je přirozené vydat se při hledání minima ve směru záporně vzatého gradientu, neboť to je směr ve kterém funkce klesá nejrychleji. Tento postup použil již před 130 lety Cauchy.

Tato metoda sestává ze série jednodimenzionálních minimalizací každá ve směru nejstrmějšího poklesu funkce v bodě, kde se hledání počíná. Směr gradientu pochopitelně nezůstává konstantní, je tedy nutné počítat s mnoha iteracemi.

7.5.3 Newtonova metoda

Uvažujme obecnou kvadratickou funkci. Tu můžeme minimalizovat jen pokud budeme mít k dispozici hodnoty funkce, její první i druhou derivaci. Kvadratickou funkci napíšeme jako:

$$F(\underline{x}) = F(\underline{x}_0) + g^T(\underline{x} - \underline{x}_0) + \frac{1}{2}(\underline{x} - \underline{x}_0)^T G(\underline{x} - \underline{x}_0), \quad (86)$$

kde gradient g je vypočten v bodě \underline{x}_0 a matice G druhé derivace je konstanta. Potom je minimum dáno:

$$\underline{x}_m = \underline{x}_0 - G^{-1}g = \underline{x}_0 - Vg, \quad (87)$$

kde jsme inverzní matici druhé derivace označili V . To je v podstatě vícedimenzionální ekvivalent kvadratické interpolace. K výhodám patří fakt, že jednotlivé iterační kroky nemohou být libovolně dlouhé, jsou totiž přesně určeny touto metodou. Výhodou je i skutečnost, že směr při hledání není vždy pevně daný gradientem, ale bývá korelován smíšenými druhými derivacemi.

K záporům patří skutečnost, že metoda diverguje, pokud je matice G (nebo V) pozitivně definitní. Přesto je to velmi účinný nástroj a mnoho užitečných algoritmů je na Newtonově metodě založena.

7.5.4 Konjugované směry

Vektory \underline{d}_i a \underline{d}_j nazveme konjugované vzhledem k pozitivně definitní symetrické matici A , pokud:

$$\underline{d}_i^T A \underline{d}_j = 0 \quad \text{pro } i \neq j. \quad (88)$$

Pokud by A byla jednotková matice, byly by tyto konjugované vektory ortogonální. Konjugaci tedy můžeme považovat za zobecnění ortogonalitu. Množina n konjugovaných vektorů generuje n -dimensionální prostor, tedy libovolný vektor toho prostoru můžeme zapsat jako lineární kombinaci n konjugovaných vektorů. Sadu konjugovaných vektorů můžeme zkonstruovat pomocí Gram-Smidtova ortogonalizačního procesu. Zvolme si libovolný vektor \underline{d}_1 . Potom je snadné ověřit, že vektor

$$\underline{d}_2 = A \underline{d}_1 - \frac{\underline{d}_1^T A A \underline{d}_1}{\underline{d}_1^T A \underline{d}_1} \underline{d}_1 \quad (89)$$

je konjugovaný k vektoru \underline{d}_1 . Tento proces můžeme zopakovat a získat tak vektor \underline{d}_3 , který bude konjugovaný k \underline{d}_1 i \underline{d}_2 . Takto bychom mohli pokračovat. Získali bychom n konjugovaných vektorů $\{\underline{d}_1, \underline{d}_2, \dots, \underline{d}_n\}$.

Takové vektory pro nás budou z hlediska minimalizace zajímavé, pokud budou konjugované vzhledem k matici druhé derivace G . V takovém případě se dá ukázat, že posloupnost lineárních minimalizací v každém z n konjugovaných směrů minimalizuje obecnou kvadratickou funkci n proměnných. Tuto skutečnost můžeme snadno předvést. Uvažujme kvadratickou funkci:

$$F(\underline{x}) = F(\underline{0}) + \underline{g}^T \underline{x} + \frac{1}{2} \underline{x}^T G \underline{x} \quad (90)$$

a n vektorů konjugovaných vzhledem ke G :

$$\underline{d}_i^T G \underline{d}_j = 0, \quad i \neq j. \quad (91)$$

Vektory \underline{x} a \underline{g} můžeme vyjádřit jako lineární kombinace

$$\underline{x} = \sum_i y_i \underline{d}_i \quad (92)$$

$$\underline{g} = \sum_i c_i \underline{d}_i. \quad (93)$$

Naši obecnou kvadratickou funkci tedy můžeme nyní psát jako:

$$F(\underline{x}) = F(\underline{0}) + \left(\sum_i c_i \underline{d}_i^T \right) \left(\sum_j y_j \underline{d}_j \right) + \frac{1}{2} \left(\sum_i y_i \underline{d}_i^T \right) G \left(\sum_j y_j \underline{d}_j \right) \quad (94)$$

Pokud nyní poslední člen přeskupíme do dvojnásobné sumy, vypadnou kvůli konjugaci podmínky členy, pro které $i \neq j$. Poslední výraz tedy můžeme zjednodušit:

$$F(\underline{x}) = F(\underline{0}) + \sum_i \sum_j c_i \underline{d}_i^T \underline{d}_j y_j + \frac{1}{2} \sum_j y_j^2 \underline{d}_j^T G \underline{d}_j \quad (95)$$

$$= \sum_j (b_j y_j + b'_j y_j^2) \quad (96)$$

kde

$$b_j = \sum_i c_i \underline{d}_i^T \underline{d}_j \quad (97)$$

a

$$b'_j = \sum \underline{d}_j^T G \underline{d}_j \quad (98)$$

jsou konstanty. Vyjádřením kvadratické funkce v proměnné y místo x jsme ji rozdělili na sumu nezávislých kvadratických funkcí jedné proměnné. Minimalizace vzhledem k y_i pak bude nezávislá na minimalizaci v ostatních konjugovaných směrech.

Lze namítnout, že pro konstrukci konjugovaných vektorů je zapotřebí znalost matice druhé derivace G a tehdy lze okamžitě použít Newtonovu metodu. Užitečnost užití konjugovaných vektorů však spočívá v tom, že existují metody umožňující jejich výpočet bez úvodní znalosti celé matice G . Jelikož nám jde o výpočet všech konjugovaných vektorů, půjde zřejmě o postupy, ve kterých budeme znalost matice G suplovat ekvivalentními informacemi. Nicméně, takový postup umožňuje zpracovávat informace pro minimalizaci průběžně a to se u zvláště u rozsáhlých výpočtů osvědčuje - takové algoritmy bývají stabilnější.

7.5.5 Konjugované gradienty

Pokud jsou známé první derivace minimalizované funkce, lze užít velmi elegantní metodu konjugovaných gradientů. Předpokládejme, že je funkce vypočtena ve dvou bodech \underline{x}_0 a \underline{x}_1 . Uvažujme nyní rozdíly:

$$\underline{\Delta x} = \underline{x}_1 - \underline{x}_0 \quad (99)$$

$$\underline{\Delta g} = \underline{g}_1 - \underline{g}_0. \quad (100)$$

Pokud uvažujeme kvadratickou funkci s maticí druhých derivací G , potom platí:

$$\underline{\Delta g} = G \underline{\Delta x}. \quad (101)$$

Libovolný vektor \underline{d}_1 ortogonální k $\underline{\Delta g}$ je konjugovaný k $\underline{\Delta x}$, neboť:

$$\underline{d}_1^T \underline{\Delta g} = \underline{d}_1^T G \underline{\Delta x} = 0. \quad (102)$$

Máme tedy návod na výpočet konjugovaných vektorů bez znalosti G založené na změně gradientu v nějakém zvoleného směru.

Za počáteční směr volíme $d_0 = -g_0$, tedy směr nejstrmějšího spádu v počátečním bodě x_0 . Označme minimum nalezené v tomto směru jako x_1 a gradient v tomto směru jako g_1 . Další

vyhledávací směr bude dán vektorem od kterého požadujeme, aby byl konjugovaný k \underline{d}_0 . Musí být lineární kombinací vektorů, které máme v danou chvíli k dispozici, konkrétně:

$$\underline{d}_1 = -\underline{g}_1 + b\underline{d}_0. \quad (103)$$

Podmínka konjugace dává:

$$\underline{d}_1^T G \underline{d}_0 = \underline{d}_1^T G (\underline{x}_1 - \underline{x}_0) = 0 \quad (104)$$

nebo

$$(-\underline{g}_1^T + b\underline{d}_0^T) G \underline{d}_0 = (-\underline{g}_1^T - b\underline{g}_0^T) (\underline{g}_1 - \underline{g}_0) = 0. \quad (105)$$

Protože \underline{x}_1 je minimum nelezené ve směru $\underline{d}_0 = -\underline{g}_0$, bude směr \underline{g}_0 ortogonální ke gradientu funkce v bodě \underline{x}_1 , tedy $\underline{g}_1^T \underline{g}_0 = 0$. Potom pro koeficient b dostáváme:

$$b = \frac{\underline{g}_1^T \underline{g}_1}{\underline{g}_0^T \underline{g}_0}. \quad (106)$$

Tedy nový směr bude dán konjugovaným vektorem:

$$\underline{d}_1 = -\underline{g}_1 + \frac{\underline{g}_1^T \underline{g}_1}{\underline{g}_0^T \underline{g}_0} \underline{d}_0. \quad (107)$$

Tímto způsobem můžeme vygenerovat n navzájem konjugovných vektorů, přičemž i -tý bude mít tvar:

$$\underline{d}_{i+1} = -\underline{g}_{i+1} + \frac{\underline{g}_{i+1}^T \underline{g}_{i+1}}{\underline{g}_i^T \underline{g}_i} \underline{d}_i. \quad (108)$$

8 Výsledky a závěry

V následující kapitole shrneme výsledky, které byly získány pomocí vypracovaného programu. Nejprve předvedeme některá řešení evoluční rovnice (55), pote výsledky fitování těchto řešení na experimentální data.

Program umožňuje pro zadanou počáteční podmínku numericky napočítat hodnotu distribuční funkce pro libovolné x a Q^2 . Výpočtem dostatečného počtu takových hodnot lze získat představu o průběhu funkcí řešících příslušnou evoluční rovnici. Důležitý pro nás bude především charakter vývoje distribuční funkce (jako funkce v x) v Q^2 . Nyní si představíme konkrétní výsledky. Na Obrázku 7. je znázorněna závislost distribučních funkcí řešících rovnici (55) na x pro různé hodnoty Q^2 , na Obrázku 8. jsou potom vyznačeny závislosti na Q^2 pro různá x . Zvolený rozsah x i Q^2 je pod možnostmi současných urychlovačů, nicméně mnou zjištěné výsledky se snaží mít pouze ilustrativní charakter a pro ten účel jsou dostačující.

Důležité je, že je z těchto výsledků patrná jedna zajímavá vlastnost distribučních funkcí. Pro malá x jsou to klesající funkce v Q^2 , kdežto pro velká x jsou v Q^2 funkcemi klesajícími. Fyzikální vysvětlení této skutečnosti je následující. S rostoucím předaným kvadrátem čtyřhybnosti Q^2 se zvětšuje rozlišovací schopnost experimentu. To znamená, že to, co se nám při menším Q^2 jeví jako jeden parton nesoucí např. 10 GeV^2 se nám při větším předaném čtyřimpulsu jeví jako shluk

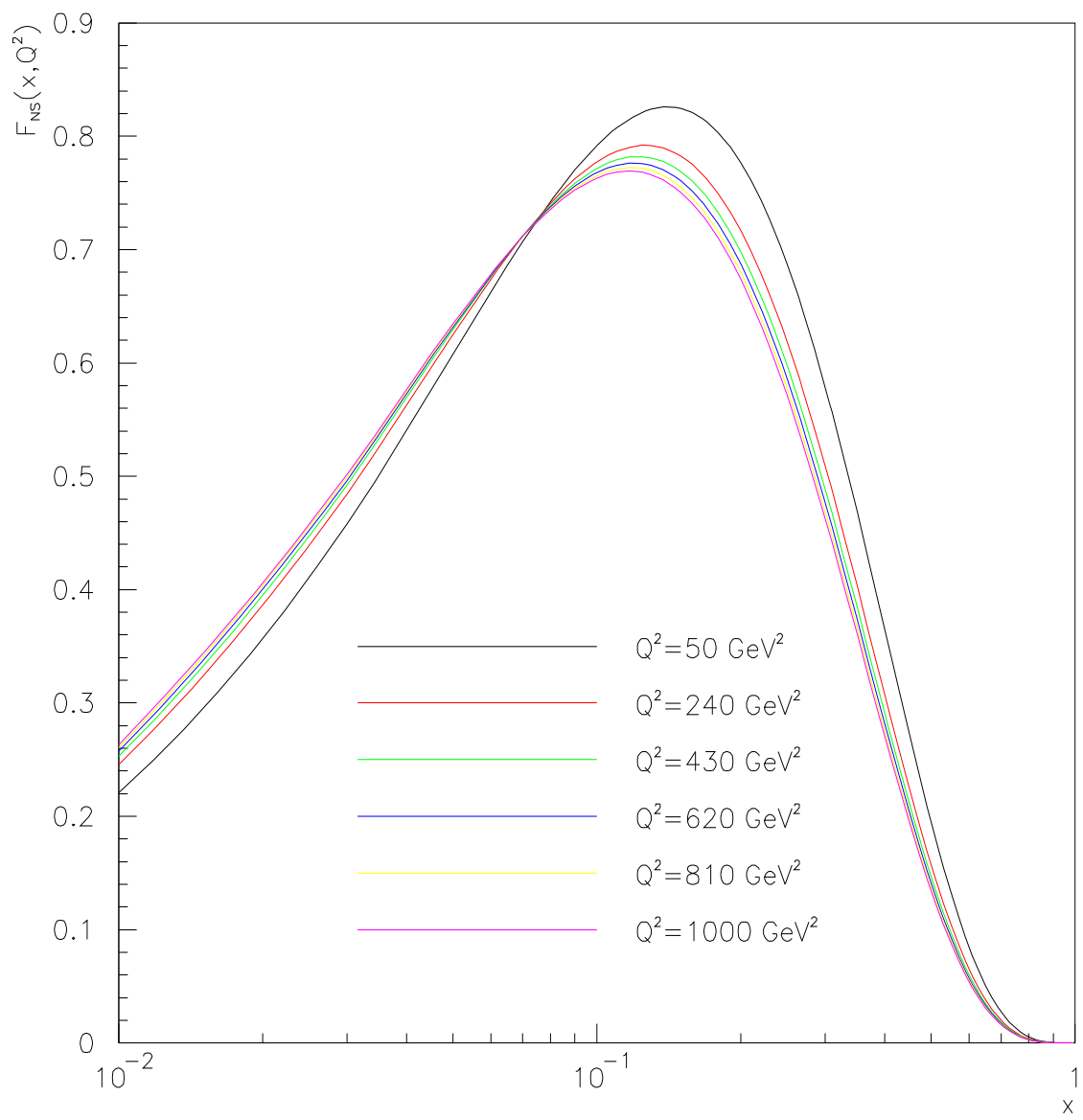
například dvou partonů, z nichž každý nese energii 5GeV^2 . Je tedy zřejmé, že výskyt partonů nesoucích menší podíly celkové čtyřhybnosti nukleonu s rostoucím Q^2 roste, kdežto výskyt partonů nesoucích větší podíly celkové čtyřhybnosti s rostoucím Q^2 klesá.

Nyní obrátíme pozornost na výsledky fitů. Experimentální data byla získána z adresy [7]. Pro minimalizaci funkce $\chi^2(A, \alpha, \beta, \Lambda)$ byly použity dvě metody, Nelderova a Meadova simplexní metoda (více kapitola ?.) a metoda konjugovaných gradientů (více kapitola 2.). Postup vývoje minimalizace obou metod je zaznamenán v Příloze 1. a v Příloze 2. Jsou v nich zaznamenány jednotlivé iterační kroky. V každém z nich dochází k nalezení nových hodnot parametrů, které snižují hodnotu funkce χ^2 . V obou metodách se ukázalo velmi užitečné předdefinovat parametr A ve výrazu (54) na A^2 . Minimalizovaná funkce se stává citlivější na změnu tohoto parametru. V případě metody konjugovaných gradientů navíc tento parametr vstupuje do všech složek gradientu (původně byla jedna složka gradientu na tomto parametru nezávislá). To všechno přispělo k výraznému zlepšení výsledků minimalizací v případě jedné i druhé metody. Oběma metodami bylo dosaženo takřka stejné minimální hodnoty $\chi = 75$. Takový výsledek můžeme s ohledem na zvolenou parametrizaci považovat za uspokojivý. Odlišnost je jen v hodnotách parametrů, v nichž je toto minimum nabyto.

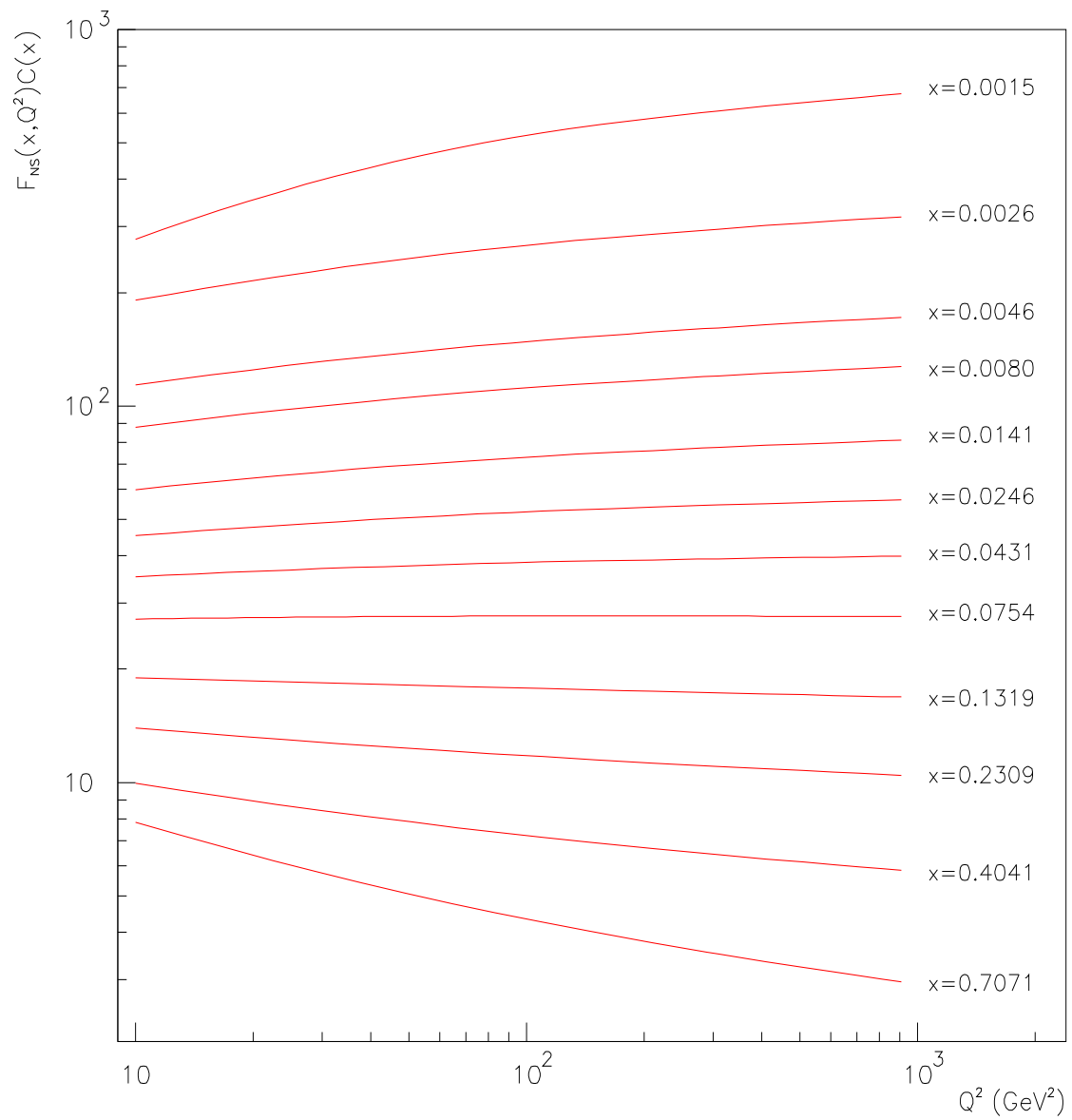
Na sérii obrázků 9.-19. je proveden fit řešení evoluční rovnice na experimentální data pro různé hodnoty Q^2 . Byly použity hodnoty parametrů zjištěné metodou konjugovaných vektorů, tedy $A = 2.414$, $\alpha = 0.707$, $\beta = 3.301$ a $\Lambda = 0.081 \text{ MeV}$; $Q^2 = 5 \text{ GeV}^2$. V těchto grafech je také zaznamenán vývoj při procesu fitování. Je znázorněn přerušovanými křivkami, které představují řešení evoluční rovnice odpovídající parametrům zvoleným na počátku minimalizace a napočítaným parametrům z 9-té iterace ($\chi^2 = 2766$).

Výsledek 1	Výsledek 2	Odhadovaná chyba	Skutečná chyba
33.9605	34.1589	0.0000	0.1985
18.4426	18.4489	0.0338	0.0063
13.4200	13.4200	0.0002	0.0000
10.6406	10.6366	0.0000	0.0040
8.7917	8.7893	0.0000	0.0024
7.4395	7.4380	0.0000	0.0014
6.3923	6.3935	0.0177	0.0012
5.5498	5.5491	0.0055	0.0007
4.8538	4.8543	0.0024	0.0006
4.2672	4.2667	0.0020	0.0006
3.7656	3.7662	0.0019	0.0005
3.3318	3.3316	0.0020	0.0002
2.9531	2.9529	0.0020	0.0002
2.6203	2.6206	0.0021	0.0003
2.3262	2.3262	0.0022	0.0000
2.0651	2.0648	0.0022	0.0003
1.8325	1.8325	0.0023	0.0000
1.6247	1.6249	0.0024	0.0002
1.4388	1.4388	0.0025	0.0000
1.2721	1.2719	0.0026	0.0002
1.1226	1.1225	0.0028	0.0001
0.9885	0.9885	0.0029	0.0001
0.8681	0.8682	0.0000	0.0002
0.7601	0.7601	0.0000	0.0001
0.6633	0.6632	0.0000	0.0001
0.5766	0.5765	0.0000	0.0001
0.4991	0.4991	0.0000	0.0001
0.4301	0.4301	0.0000	0.0001
0.3686	0.3687	0.0000	0.0001
0.3140	0.3141	0.0000	0.0001
0.2658	0.2658	0.0000	0.0000
0.2234	0.2233	0.0000	0.0001
0.1861	0.1860	0.0000	0.0001
0.1537	0.1536	0.0000	0.0001
0.1255	0.1255	0.0000	0.0000
0.1013	0.1014	0.0000	0.0001
0.0806	0.0807	0.0000	0.0001
0.0631	0.0632	0.0000	0.0001

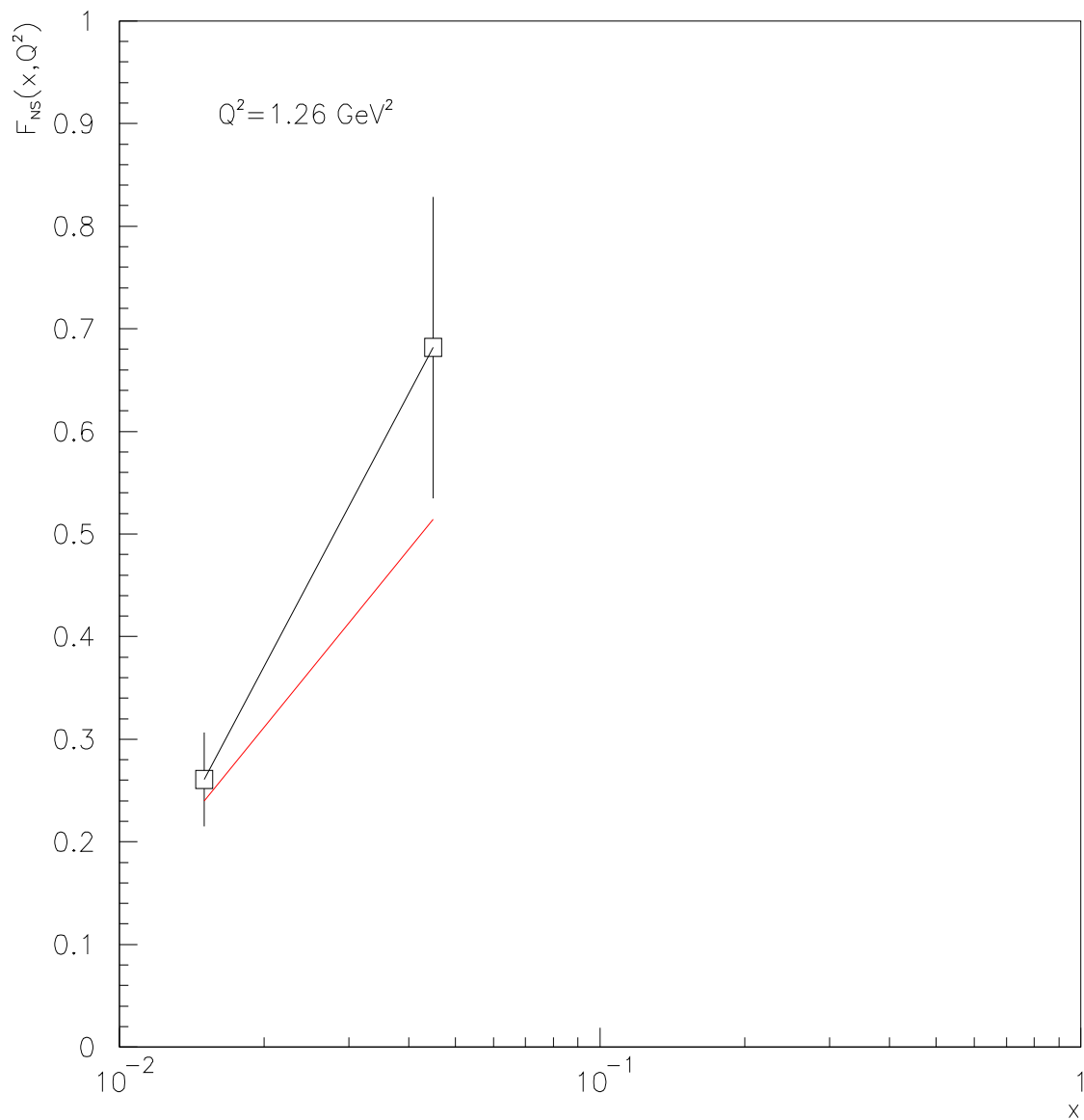
Tabulka 1: Ověření správnosti inverzní Mellinovy transformace. Jako testovací funkce byla zvolena $8.6x^{0.85}(1-x)^{3.94}$. V prvním sloupci jsou hodnoty této funkce zjištěné přímým výpočtem, ve druhém jsou potom invertované momenty této funkce. Tabulka je ještě doplněna o odhad chyb při inverzním transformování a o skutečné chyby, kterých jsme se výpočtu dopustili.



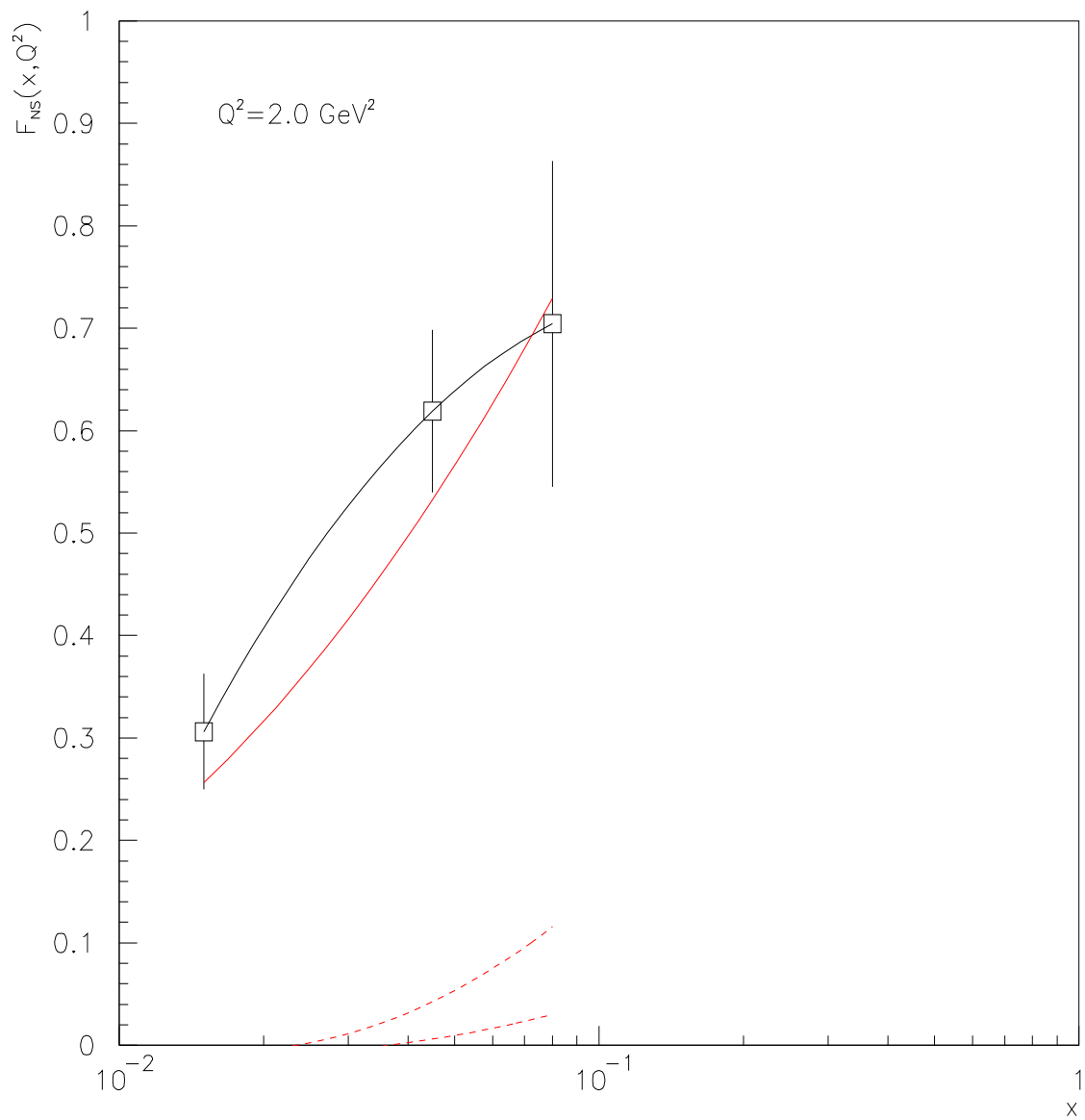
Obrázek 7: Vypočtené závislosti distribuční funkce na x pro různé hodnoty Q^2 .



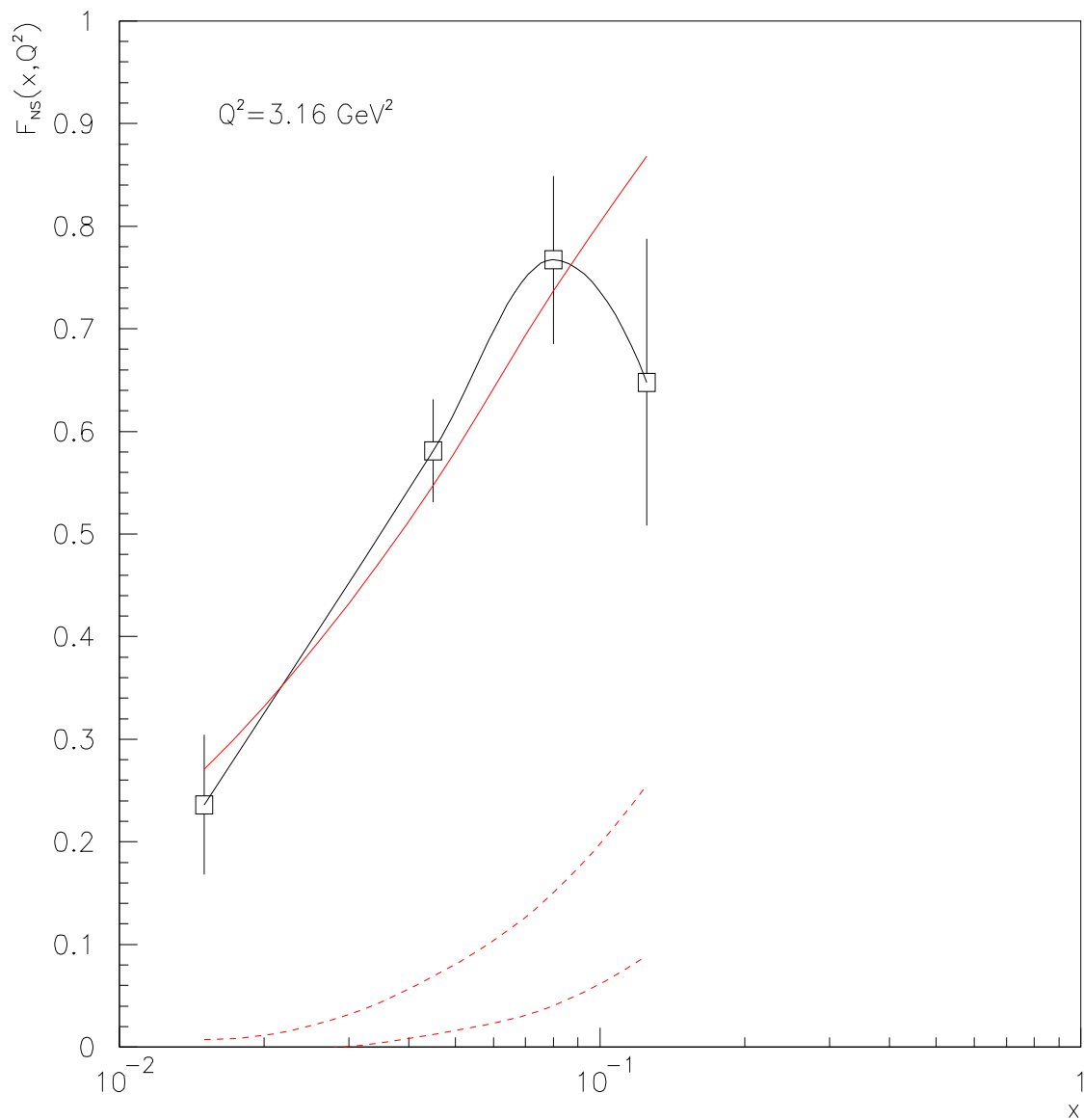
Obrázek 8: Vypočtené závislosti distribuční funkce na Q^2 pro různé hodnoty x .



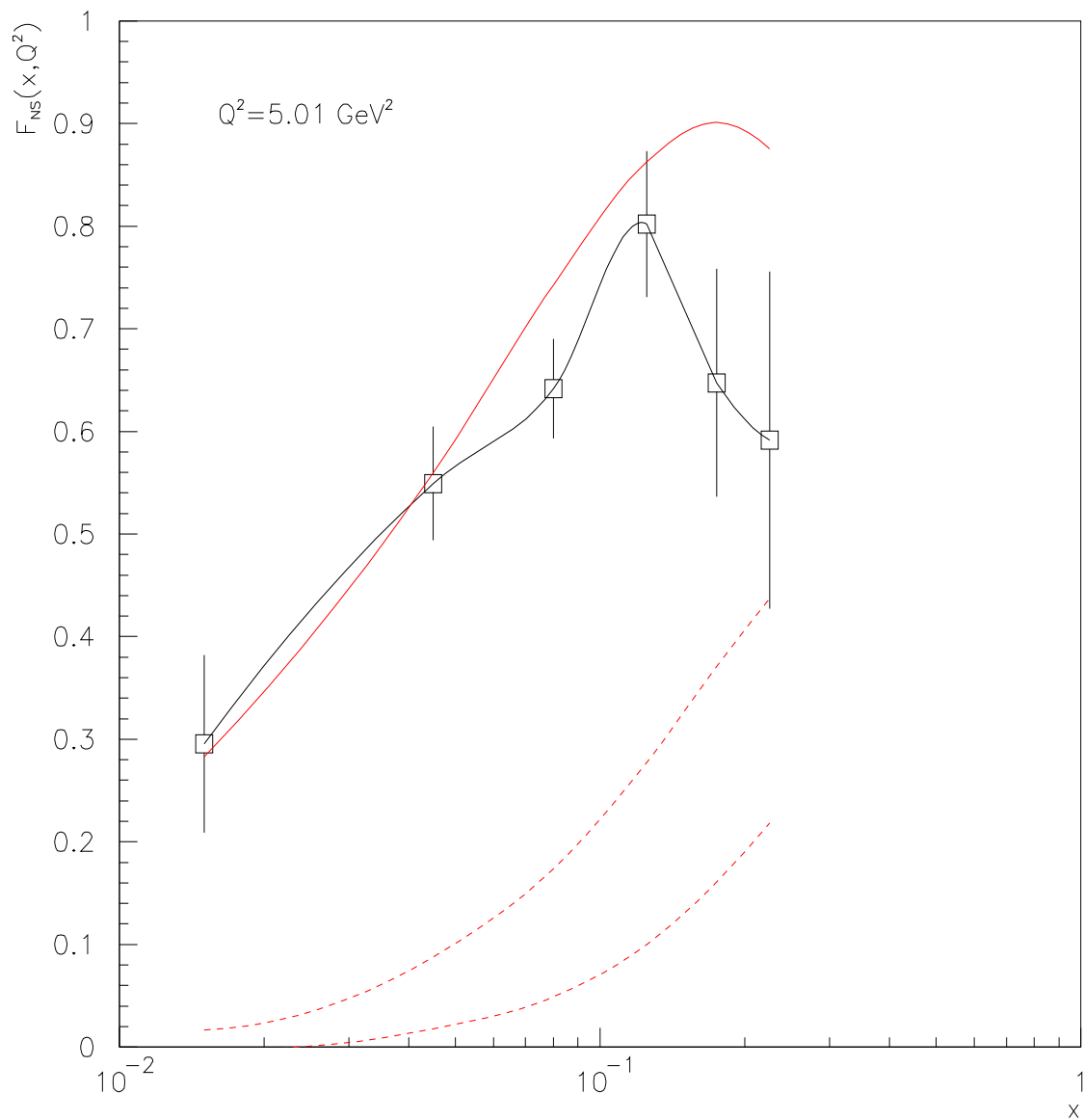
Obrázek 9: Srovnání experimentálních dat pro F_{NS} při $Q^2 = 1.26 \text{ GeV}^2$ s hodnotami vypočtenými řešením evolučních rovnic s postupně zpřesňovanými parametry A , α , β a Λ .



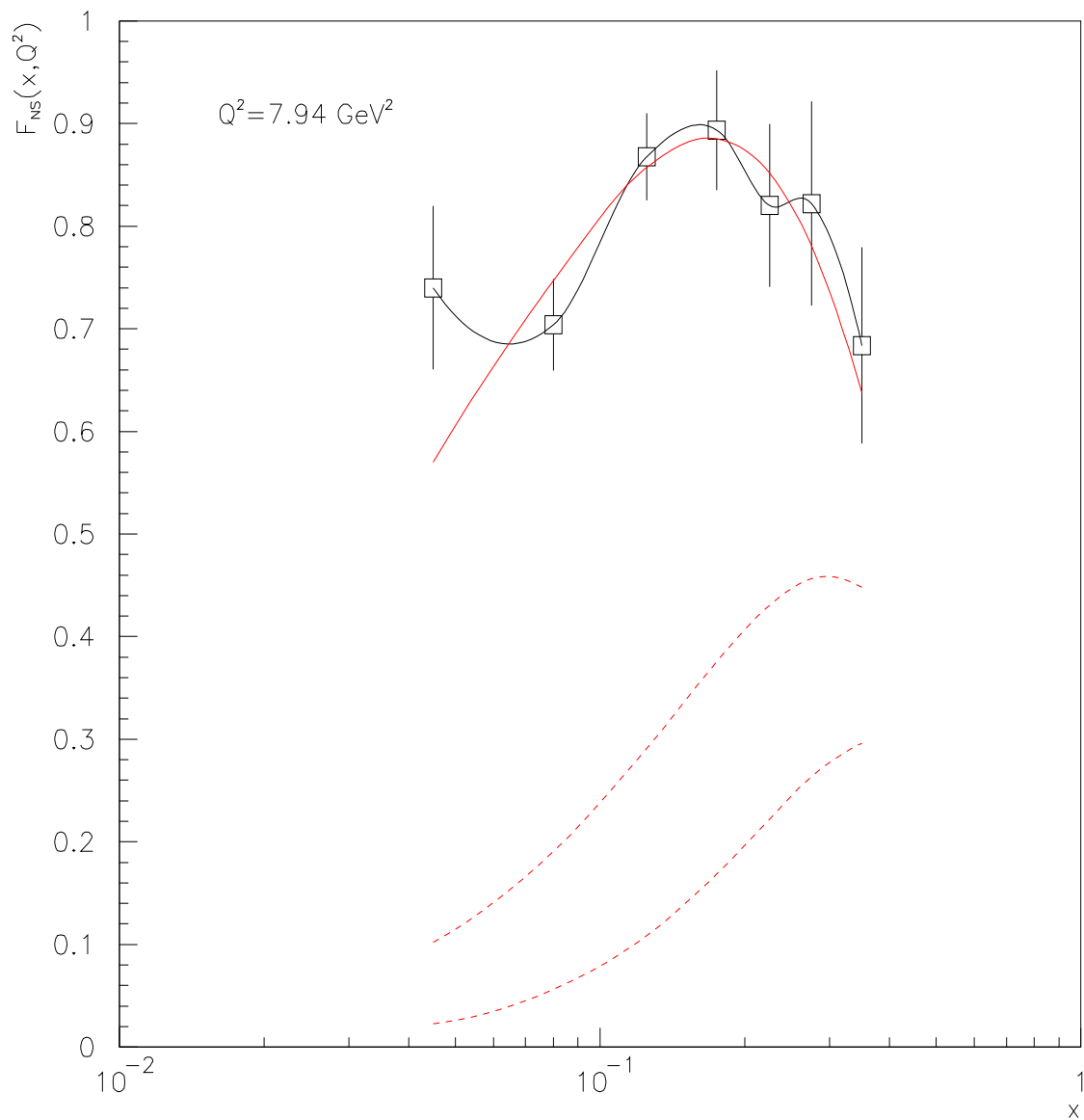
Obrázek 10: Srovnání experimentálních dat pro F_{NS} při $Q^2 = 2.0 \text{ GeV}^2$ s hodnotami vypočtenými řešením evolučních rovnic s postupně zpřesňovanými parametry A , α , β a Λ .



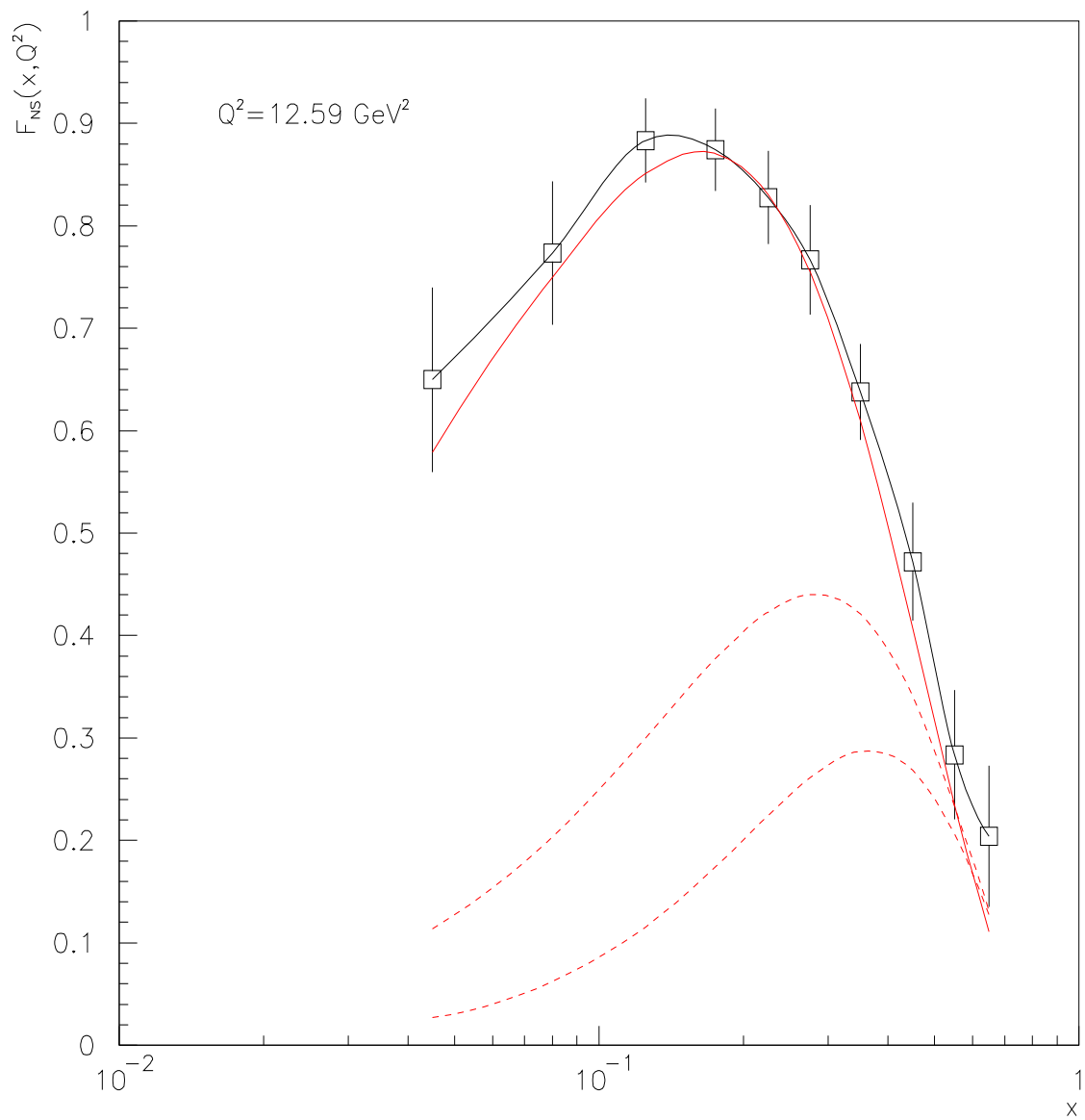
Obrázek 11: Srovnání experimentálních dat pro F_{NS} při $Q^2 = 3.16 \text{ GeV}^2$ s hodnotami vypočtenými řešením evolučních rovnic s postupně zpřesňovanými parametry A , α , β a Λ .



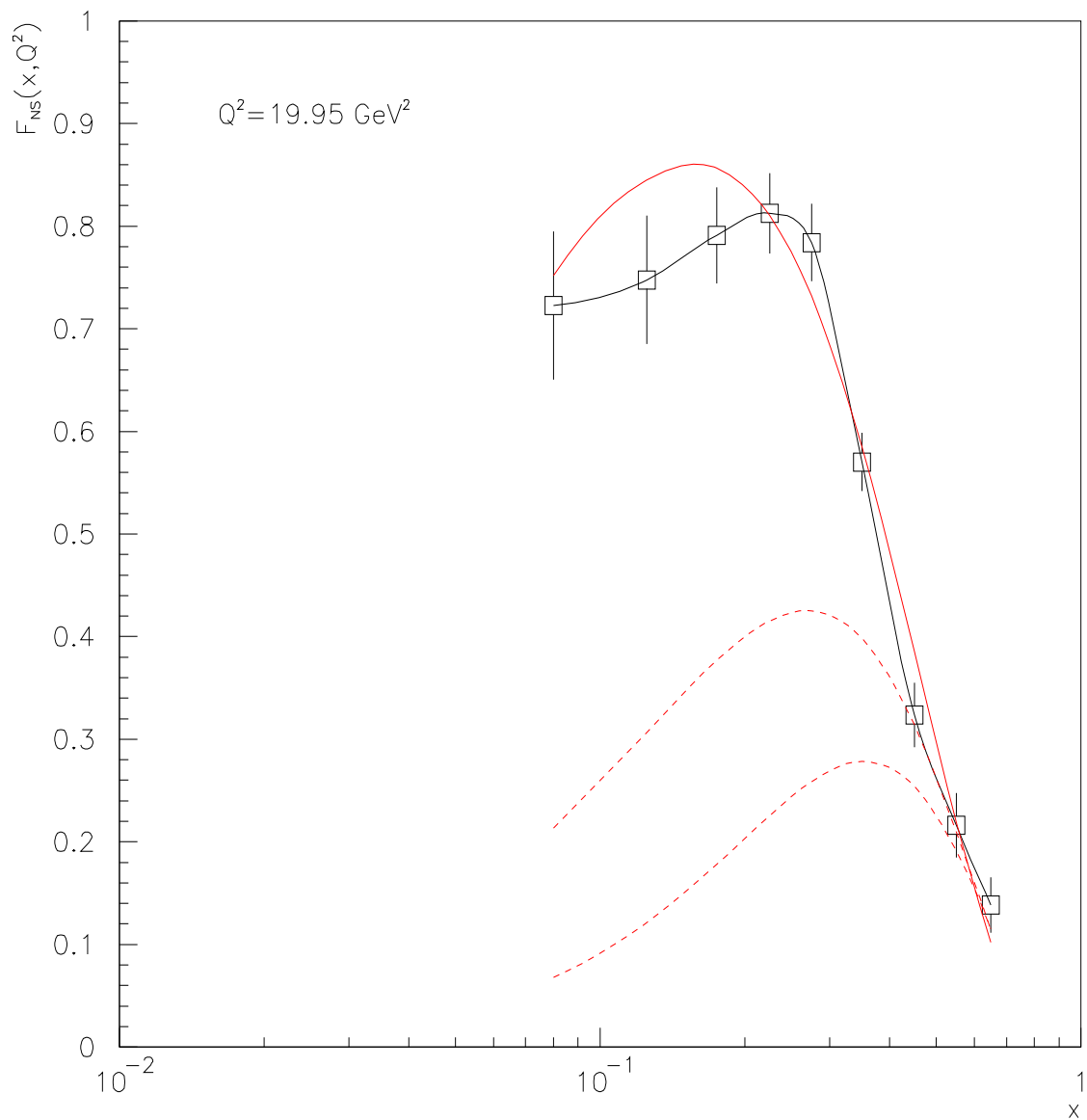
Obrázek 12: Srovnání experimentálních dat pro F_{NS} při $Q^2 = 5.01 \text{ GeV}^2$ s hodnotami vypočtenými řešením evolučních rovnic s postupně zpřesňovanými parametry A , α , β a Λ .



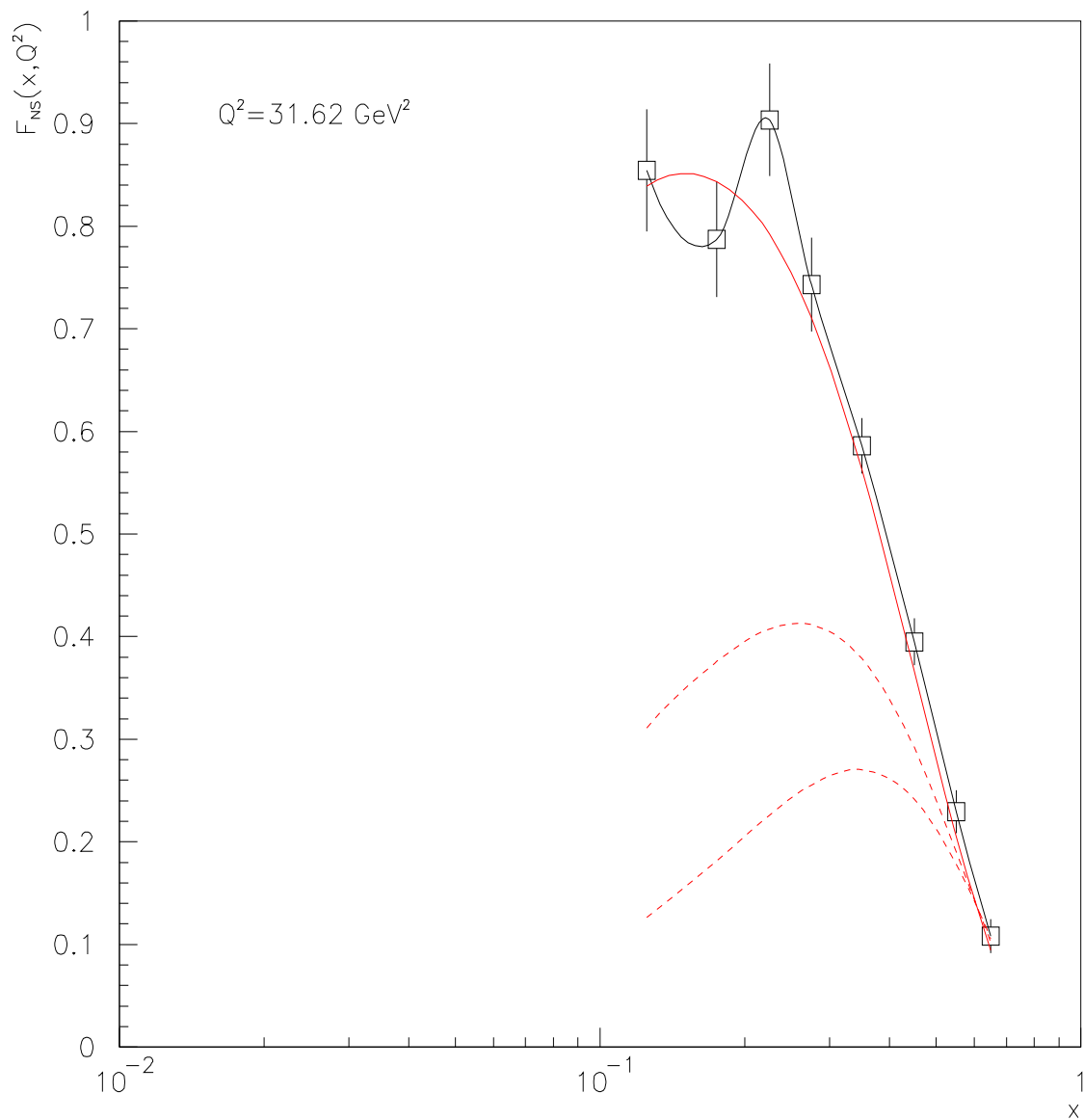
Obrázek 13: Srovnání experimentálních dat pro F_{NS} při $Q^2 = 7.94 \text{ GeV}^2$ s hodnotami vypočtenými řešením evolučních rovnic s postupně zpřesňovanými parametry A , α , β a Λ .



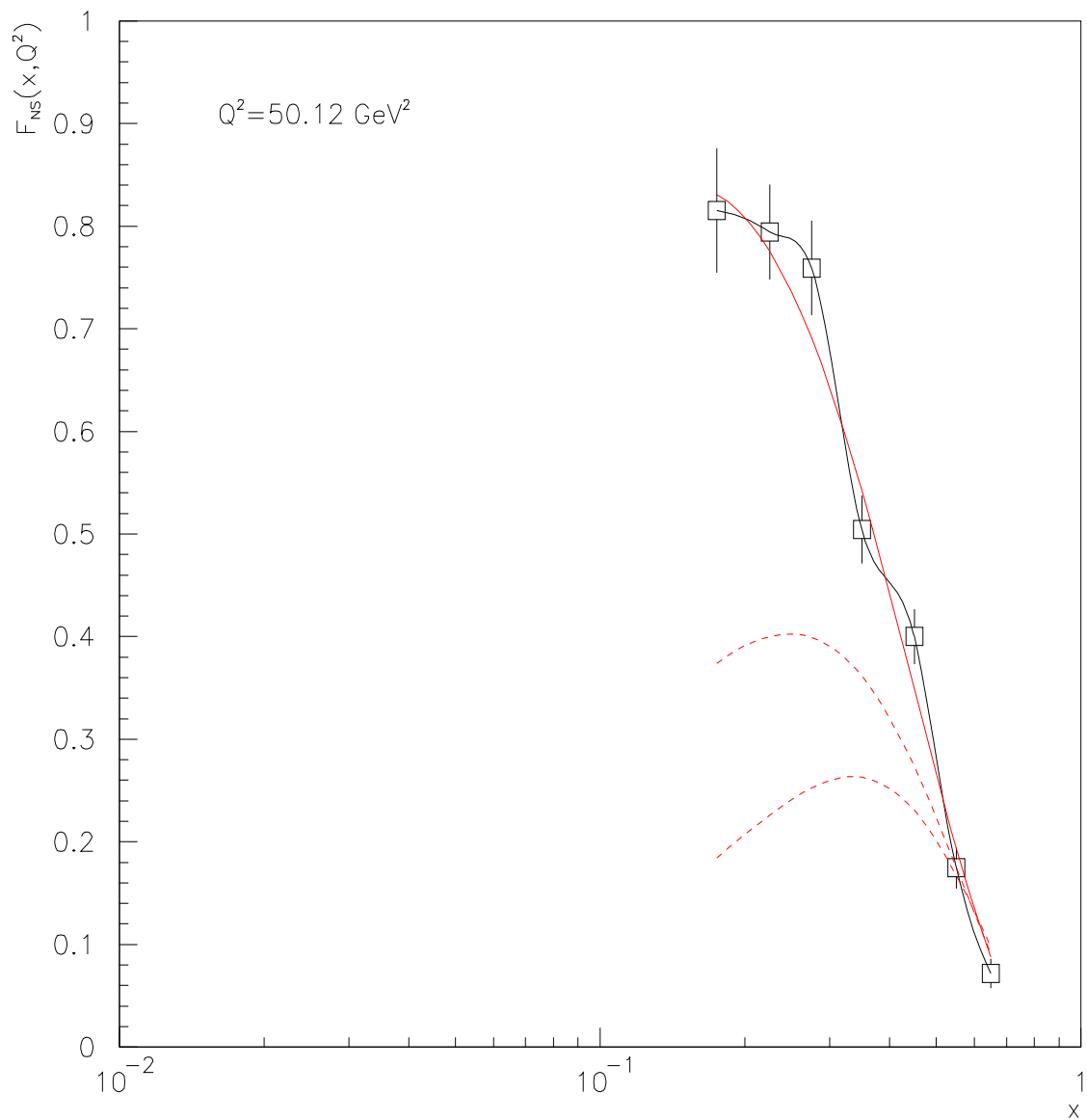
Obrázek 14: Srovnání experimentálních dat pro F_{NS} při $Q^2 = 12.59 \text{ GeV}^2$ s hodnotami vypočtenými řešením evolučních rovnic s postupně zpřesňovanými parametry A , α , β a Λ .



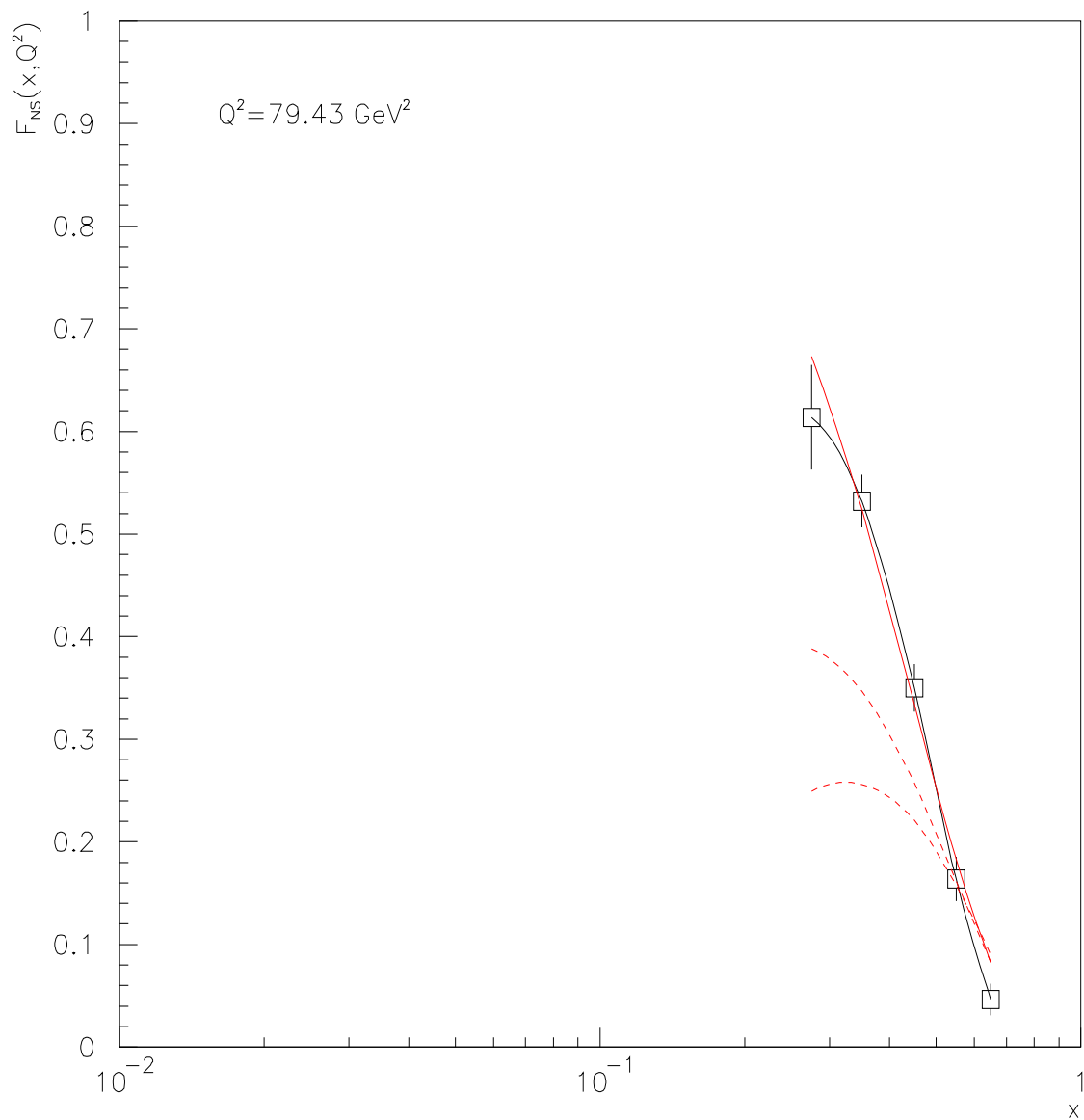
Obrázek 15: Srovnání experimentálních dat pro F_{NS} při $Q^2 = 19.95 \text{ GeV}^2$ s hodnotami vypočtenými řešením evolučních rovnic s postupně zpřesňovanými parametry A , α , β a Λ .



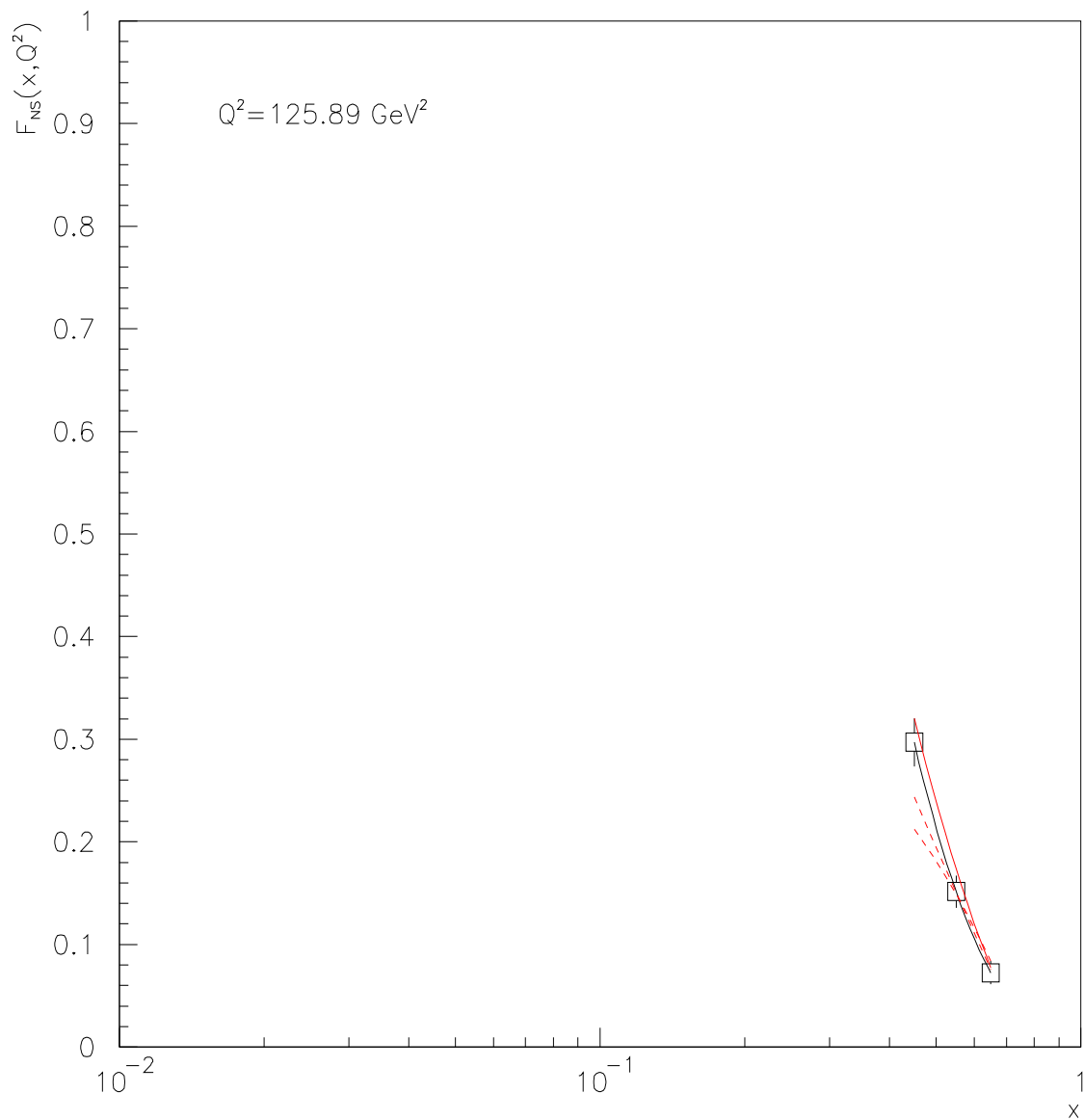
Obrázek 16: Srovnání experimentálních dat pro F_{NS} při $Q^2 = 31.62 \text{ GeV}^2$ s hodnotami vypočtenými řešením evolučních rovnic s postupně zpřesňovanými parametry A , α , β a Λ .



Obrázek 17: Srovnání experimentálních dat pro F_{NS} při $Q^2 = 50.12 \text{ GeV}^2$ s hodnotami vypočtenými řešením evolučních rovnic s postupně zpřesňovanými parametry A , α , β a Λ .



Obrázek 18: Srovnání experimentálních dat pro F_{NS} při $Q^2 = 79.43 \text{ GeV}^2$ s hodnotami vypočtenými řešením evolučních rovnic s postupně zpřesňovanými parametry A , α , β a Λ .



Obrázek 19: Srovnání experimentálních dat pro F_{NS} při $Q^2 = 125.89 \text{ GeV}^2$ s hodnotami vypočtenými řešením evolučních rovnic s postupně zpřesňovanými parametry A , α , β a Λ .

Reference

- [1] Chýla, J.: *Quarks, partons a Quantum Chromodynamics*, Fyzikální ústav AV ČR, 2001
- [2] Weinzierl, S.: *Fast evolution of parton distributions*, Univerzita Parma, 2002
- [3] James, F.: *Function minimization*, CERN, 1972
- [4] James, F.: *Minuit - reference manual*, CERN, 1992
- [5] Graudenz, D.: *The Mellin transform technique for the extraction of the gluon density*, DESY 95-107, 1995
- [6] Groom, D.E. et al.: *Particle Physics Booklet*, The European Physical Journal C15, 2000
- [7] <http://durpdg.dur.ac.uk/hepdata/online/f2/structindex.html>