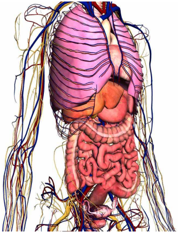
Selected Examples of Computational Activity of Wrocław Neutrino Group

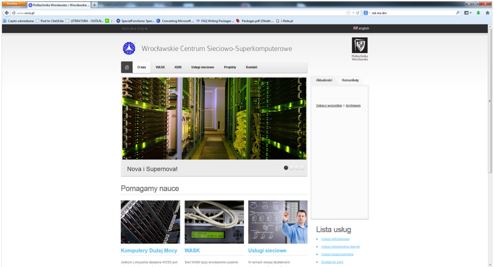
Krzysztof M. Graczyk

Instytut Fizyki Teoretycznej
Uniwersytet Wrocławski

third of February 2014, Wrocław



- Neutrino Interactions, theoretical models, Monte Carlo simulations, working with the data (Graczyk, Golan, Juszczak, Sobczyk, Zmuda)
- Heavy Ion Collisions (Prorok)
- Data Analysis
- Neural Networks (Graczyk)
- E-M Nucleon Form Factors, Proton Structure (Graczyk)
- T2K Experiment (Zmuda, Golan, Sobczyk)



Użytkownik	czas pracy procesora: (h)	(dni)	lat	liczba zadań
Tomasz Golan	**54094**	2254	6.2	**23269**
Cezary Juszczak	22529	939	2.5	764
Krzysztof Graczyk	6878	286	0.8	284
Jan Sobczyk	1			1
Dariusz Prorok	0			1
Jakób Żmuda	0			0

- root, w szczególnosci TMinuit
- gnuplot
- Mathematica (rachunki teorio-polowe), REDUCE
- NuWro (tworzony w ramach naszego zakładu)
- WNet (Bayesian Neural Network)
- GNU Scientific Library, LoopTool, FeynCalc, etc.

- Monte Carlo simulations vs. MiniBooNE experiment: statistical extraction of the physical parameters (Golan, Graczyk, Juszczak, Sobczyk, Phys.Rev. C88 (2013) 024612).
- NuWro development: implementation of the 2p2h, intranuclear cascade, and RPA (Sobczyk, Juszczak, Golan, Zmuda, Graczyk)
- Single Pion Production (Zmuda, Graczyk, Sobczyk, *in preparation*)
- Two-Photon Exchange (TPE) Effect: Analytical and Numerical calculations (Graczyk, Phys.Rev. C88 (2013) 065205)
- (TPE) by Neural Network, *paper in preparation*

Cztery żywioły – Neutrina ważnym elementem układanki

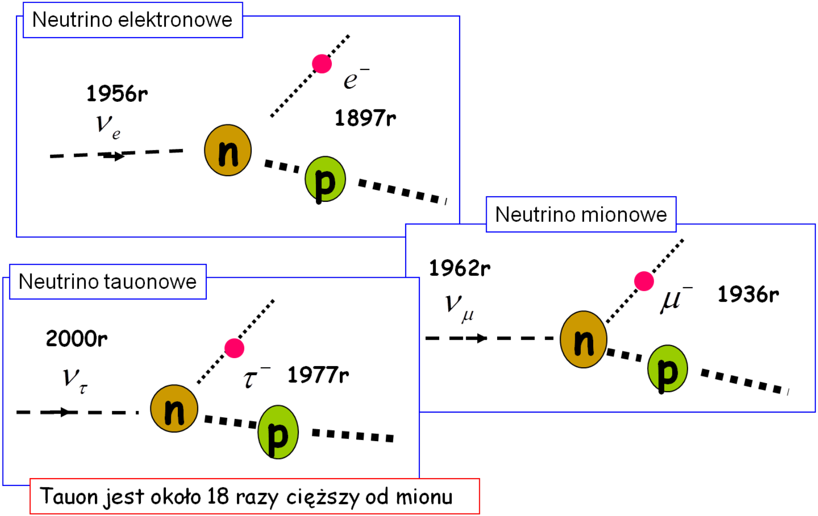
- Grawitacja - grawiton: G.
- Elektryczność i Magnetyzm - foton: γ.
- Oddziaływania Silne - gluon: g.
- Oddziaływania Słabe - bozony: W^+, W^-, Z^0.

	KWARKI		LEPTONY	
·Pierwsza rodzina	u	d	e^-	ν_e
·Druga rodzina	c	s	μ^-	ν_μ
·Trzecia rodzina	t	b	τ^-	ν_τ

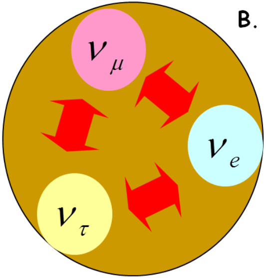
oraz odpowiednio antykwarki i antyneutrina

Trzy rodzaje neutrin!!!

Trzy rodzaje neutrin

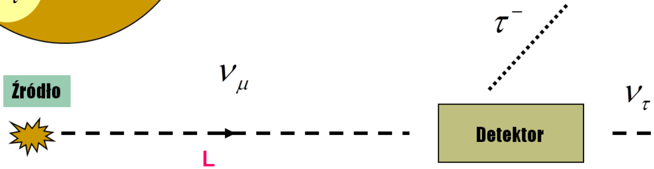


Skąd wiemy, że są masywne … Oscylacje



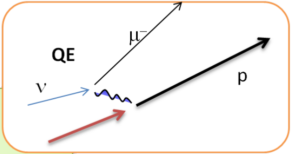
B. Pontecorvo 1958

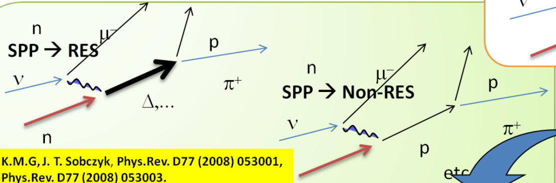
$$P(\nu_\mu \to \nu_\tau) = \sin^2 2\theta \sin^2\left(\frac{1.27\Delta m^2 L}{E}\right)$$



Oscylacje zachodzą wtedy i tylko wtedy gdy neutrina mają masę!!!

ν-nucleon



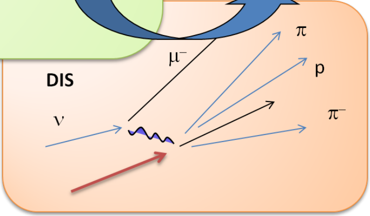


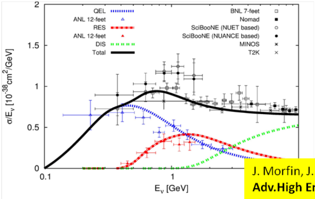
K.M.G, J. T. Sobczyk, Phys.Rev. D77 (2008) 053001,
Phys.Rev. D77 (2008) 053003.
K.M.G et al. Phys.Rev. D80 (2009) 093001.
K.M.G, AIP Conf.Proc. 1405 (2011) 134-139

- Hadronic and mesonic degrees of freedom
- Chiral filed theory motivited description
- Quark model motiveted approach
- Isobar model

KMG, C. Juszczak, J.T.Sobczyk,
Nucl.Phys. A781 (2007) 22
KMG, AIP Conf.Proc. 1222 (2010) 238

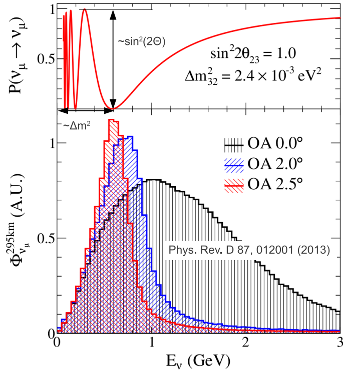
Quark-Hadron Duality





J. Morfin, J. Nieves, J. T. Sobczyk,
Adv.High Energy Phys. 2012 (2012) 934597

ν @ IFT @

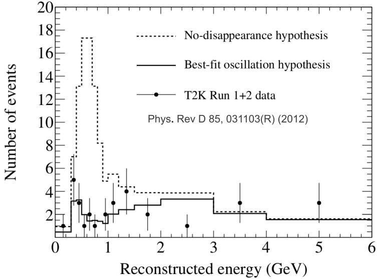


- Off-axis detector, spectrum between ∼500 - ∼1200 MeV.

- Only indirect detection, "reconstructed energy":

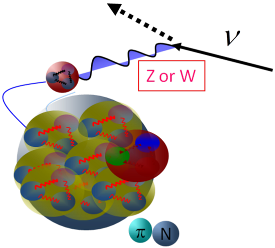
$$E_{rek.} = \frac{E_\mu M - \frac{1}{2} m_\mu^2}{M - E_\mu + p_\mu \cos(\Theta_\mu)}$$

E_μ-muon energy, $\cos(\Theta_\mu)$-muon angle. Assumption: nucleon at rest



Predictions of # events with/without oscillation

Based on Monte Carlo. What you put is what you get → dependency on theoretical lepton-nucleus interaction modeling.



- •One body currents from neutrino-nucleon scattering: an input of nuclear models
- •Impuls approximation
- •Fermi Gas model, a ground state of the nucleon
- •Spectarl function approach
- •A bare nuclear model dressed by: e.g. RPA 1n-1h excitations, calculatins within relativistic HadroDynamics .

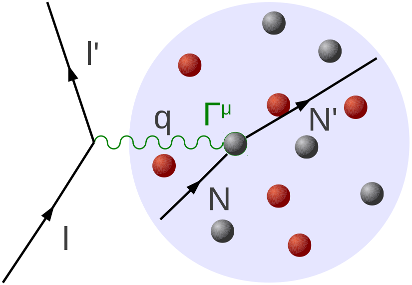
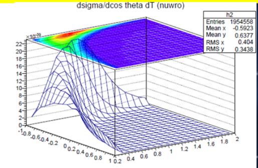


fig. J. Zmuda
`http://www.ift.uni.wroc.pl/~tgolan/`

NuWro (Neutrino-Wrocław) Monte Carlo Generator

- **Authors:**
- prof. Jan Sobczyk
- dr Cezary Juszczak
- dr Jarosław Nowak
- mgr Tomasz Golan
- dr Krzysztof Graczyk
- dr Jakub Żmuda
- Maciej Tabiszewski

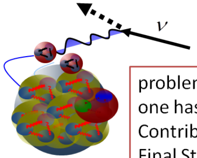
Selected papers:
- **C. Juszczak, T. Golan, J.T. Sobczyk, Phys. Rev. C86 (2012) 015505**
- **C. Juszczak , J. T. Sobczyk, J. Żmuda, Phys. Rev. C82 (2010) 045502**
- **C. Juszczak, Acta Phys.Polon. B40 (2009) 2507**
- **J. Nowak, J. T. Sobczyk, Acta Phys.Polon. B37 (2006) 1955**
- **C. Juszczak, J. A. Nowak, Nucl. Phys. Proc. Suppl. 159 (2006) 211**
- **J. T. Sobczyk, J. A. Nowak, K.M. Graczyk, Nucl.Phys.Proc.Suppl. 139 (2005) 266**



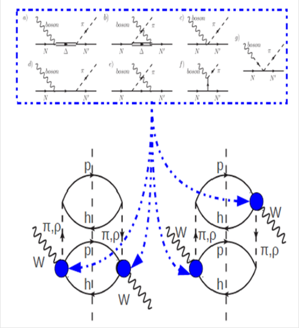
- ❑ All major neutrino-nucleus interaction channels (QEL, DIS, RES and COH)
- ❑ Covers ν energies from MeV to TeV
- ❑ density profiles and binding energies for most of nuclei
- ❑ Local Fermi Gas and Spectral Function models of nucleus.
- ❑ Intra-Nuclear cascade with pion-nucleon and nucleon-nucleon scattering.
- ❑ scattering of **complex neutrino beams** on real **detector** geometries
- ❑ the detector geometry is read from a data file (NuWro can be used by many different experiments)
- ❑ The object oriented data analysis –compatible with root CERN frametool.

- NuWro is the firts MC generator to
  - include such dynamical effects like spectral function and Meson Exchange Current.
  - has an online interface http:/nuwro.ift.uni.wroc.pl
  - It is a probably the fastest event generation, compared to other codes.

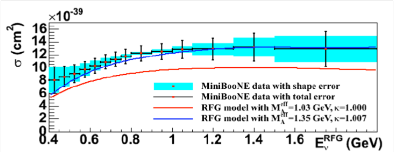
ν @ IFT @



problem of the axial mass
one has to take into account MEC
Contribution (beyond one body current)
Final State Interaction
Pion production and absorbtion etc.

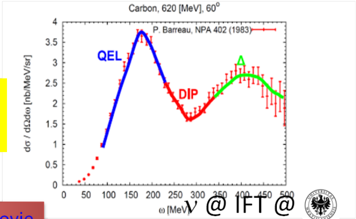


From J. Żmuda



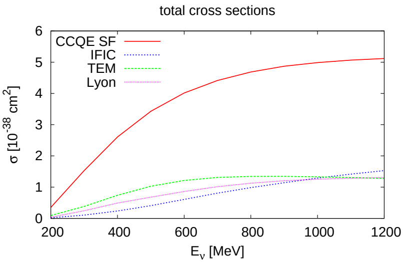
Seminar of T. Katori

J.T.Sobczyk, J.Żmuda, **arXiv:1210.6149**
J.T.Sobczyk, Phys. Rev. C86, 015504 (2012)
C. Juszczak, J.T.Sobczyk, J.Żmuda, **Phys.Rev. C82 (2010) 045502**
K.M. Graczyk, C. Juszczak, J.T.Sobczyk, T. Golan, will be available next week



FSI: propagation of nucleons throw nuclei, → Metropolis algorithm → NuWr

Tomek Golan Movie



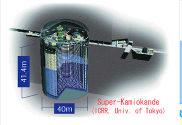
- MEC- important contribution to scattering signal.
- Theoretical uncertainties-large.

Neutrino CC double-differential cross sections

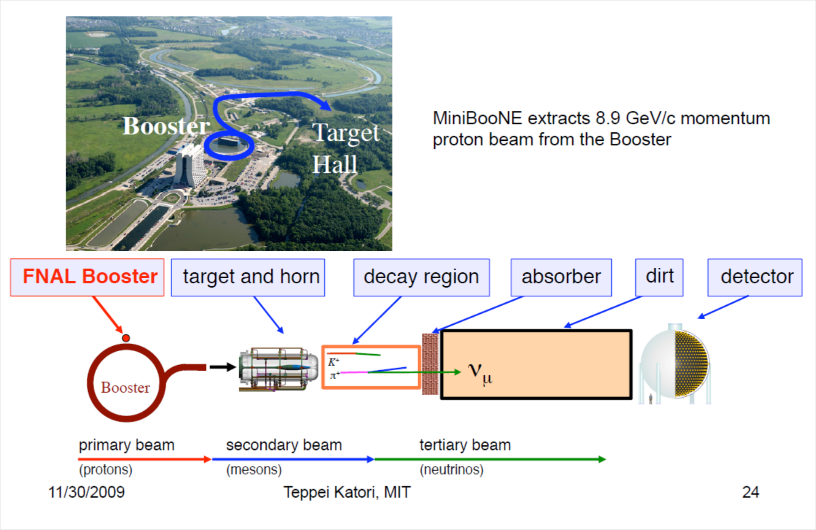
- Nieves MEC implementation in NEUT (Peter Sinclair et al. from T2K): event weights as double-differential cross section $d\sigma/dT_\mu d\cos\Theta$. Separate tables for each target, each neutrino energy E (about 90 values up to 30 GeV) each flavour and neutrino/antineutrino.
- Irregular binning, varies with flavour/nucleus/antineutrino etc. $\rightarrow$ about 6 MB needed for (anti)muon/ (anti)electron neutrino tables when repeatable values omitted.
- Quite accurate, bur relatively slow. Problem: experimental analysis needs $\sim 10^7$ MC event samples with all information about particles.

- Thanks to courtesy of Juan Nieves and Manuel Vicente-Vacas: fortran code producing hadronic tensor elements for MEC on a grid in q^0 and $|\vec{q}|$.
- Grid size : 10 MeV step up to 1.2 GeV in $|\vec{q}|$ (approximate q_{cut} for Nieves MEC model), same spacing in q^0 with $q_0 \leq |\vec{q}|$ (triangular physical region).
- Each target grid run takes 4-5 days on core-i7 3.4 GHz with OpenMP (8 cores).
- Two targets: carbon and oxygen (oter nuclei up to A=7: cross section/nucleon from carbon grid, heavier nuclei:from oxygen grid).
- Nucleon pair decay according to Jan Sobczyk Phys.Rev. C86 (2012) 015504.
- Similar results to old implementation.
- One data set: all flavors accessible as well as antineutrinos.
- Improvement both in data table size (6 MB$\longrightarrow$800 kB) and speed (order-of-magnitude factor thanks to uniform binning).
- T2K uses NuWro for reference and cross-checks!

Neutrinowy eksperyment tzw. długiej bazy - drugiej generacji T2K

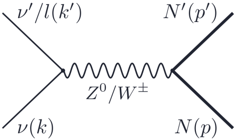


Fizycy z Warszawy, Krakowa, Katowic i Wrocławia (prof. Jan Sobczyk, mgr Tomasz Golan, mgr Jakub Żmuda)



T. Katori, Wroclaw, Neutrino Seminar 2009 prąd leptonowy:

$$j_\mu = \bar{u}(k')\gamma^\mu(1-\gamma_5)u(k)$$



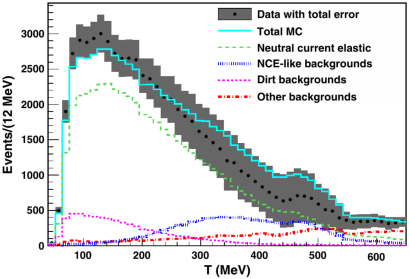
prąd hadronowy:

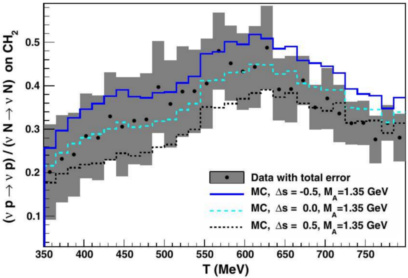
$$h_\mu = \bar{u}(p')\Gamma^\mu u(p)$$

- Γ^μ wyrażne są przez czynniki postaci: wektorowe i aksjalne.
- aksjalny czynnik postaci:

$$G_A^{p/n}(Q^2) = \pm(g_A + g_A^s)\left(1 + \frac{Q^2}{M_A}\right)^{-2}$$

- Masa aksjalna (M_A) wyznaczana z danych neutrinowych.





- Przy założeniu $g_A^s = 0$ najlepsze dopasowanie znaleziono dla:

$$M_A = 1.39 \pm 0.11 \text{ GeV}$$

- Niezgodnosć z wczeniejszymi pomiarami: $M_A \sim 1$ GeV.

- Przy założeniu $M_A = 1.35$ GeV najlepsze dopasowanie znaleziono dla:

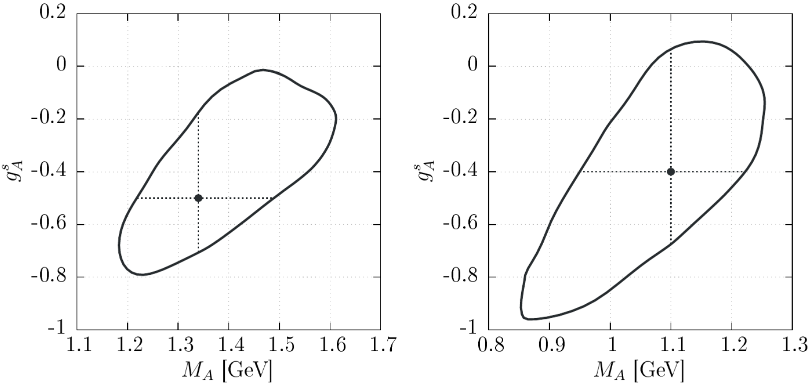
$$g_A^s = 0.08 \pm 0.26$$

- $\nu p \to \nu p$: zdarzenia z protonem powyżej progu Czerenkowa.

- Powtórzenie analizy z wykorzystaniem NuWro.
- Zbadanie wpływu wkładu od oddziaływań przez prądy wymiany mezonów na wyniki.
- Wykonanie symultanicznej ekstrakcji obu parametrów (M_A, g_A^s).

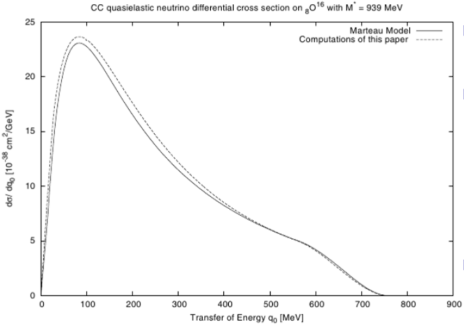
- Wykonanie dwuwymiarowego fitu wymaga ogromnej mocy obliczeniowej.
- Wstępna siatka:
  - $M_A \in (0.6, 2.5)$ GeV co 0.05 GeV $\rightarrow$ 39 pozycji
  - $g_A^s \in -1, 1)$ co $0.1 \rightarrow$ 21 pozycji
  - Razem: 819 symulacji
- Pojedyncza symulacja zajmuje ok. 6h. Obliczenie całej siatki na pojedynczym procesorze zajłoby 4914h $\approx$ 205 dni $\approx$ pół roku.
- WCSS $\rightarrow$ ok. 100 symulacji jednoczesnie (bo kolejka), co skróciło czas obliczeniowy do 2-3 dni!
- Wstępna siatka została zagęszczona w okolicach minimów χ^2, co wydłużyłoby oczekiwanie na wyniki do ok. roku.

Kontur 1σ został, wyznaczony w programie Mathematica.



(a) Without $np - nh$ (b) With $np - nh$

NuWro RPA implementation



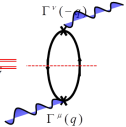
K.M. Graczyk, JTS, Eur. Phys. J C31 (2003) 177

For QE Marteau is basically the same as Martini model.

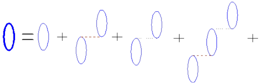
- a part of K.M. Graczyk PhD thesis
- based on a theoretical paper K.M. Graczyk, JTS, *The algebraic solution of RPA for the CC quasielastic neutrino nucleus scattering*, Eur. Phys. J C31 (2003) 177
- implementation (C. Juszczak): a CCQE event is generated; its weight is multiplied by

$$\frac{\frac{d^2\sigma^{RPA}}{d^2T_\mu d\cos\theta}}{\frac{d\sigma^{LFG}}{dT_\mu d\cos\theta}}$$

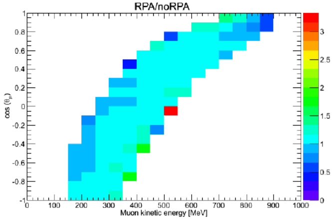
$\mathrm{Im}\ \Pi^{\mu\nu}_{RFG} =$



$$i\Pi^{\mu\nu}(q) = \int \frac{d^4p}{(2\pi)^4} Tr\left(G(p+q)\Gamma^{\mu}G(p)\Gamma^{\nu}\right)$$



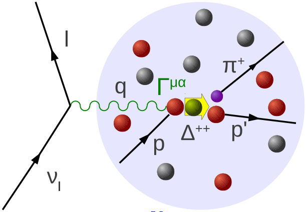
NuWro RPA implementation – example



Carbon,
$E_\nu = 1$ GeV.
Ratio: $\frac{FG+RPA}{FG}$.

Reduction of cross section wrt Fermi gas model at low Q^2 is clearly seen.

fig. by Jan Sobczyk

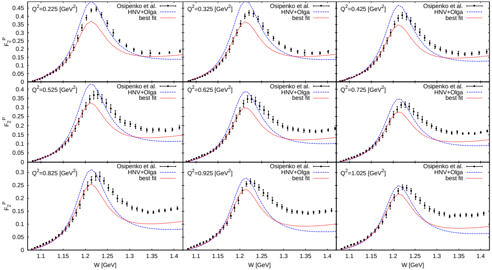


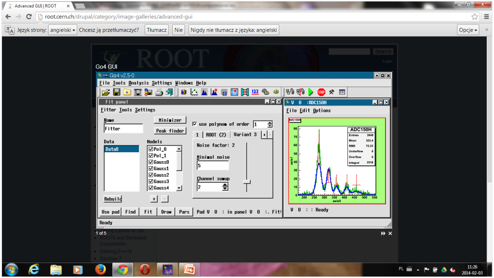
- At T2K mean energy $\nu_\mu p \to \mu^- p \pi^+$ channel: intermediate $\Delta(1232)$ baryon.
- Other channels: large nonresonant background.
- $E_{rec.}$ biased by pion.

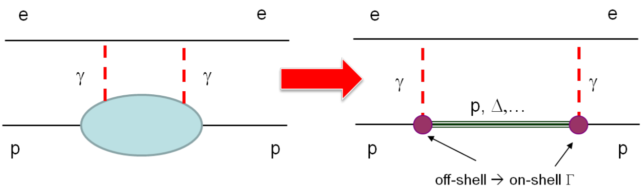
$$\Gamma^{\mu\alpha} = \left[\frac{C_3^V}{M}(g^{\alpha\mu}\not{q} - q^\alpha\gamma^\mu) + \frac{C_4^V}{M^2}(g^{\alpha\mu}q\cdot p_\Delta - q^\alpha p_\Delta^\mu) + \frac{C_5^V}{M^2}(g^{\alpha\mu}q\cdot p - q^\alpha p^\mu) + g^{\alpha\mu}C_6^V\right]\gamma^5 +$$
$$+ \left[\frac{C_3^A}{M}(g^{\alpha\mu}\not{q} - q^\alpha\gamma^\mu) + \frac{C_4^A}{M^2}(g^{\alpha\mu}q\cdot p_\Delta - q^\alpha p_\Delta^\mu) + C_5^A g^{\alpha\mu} + \frac{C_6^A}{M^2}q^\alpha q^\mu\right]$$

- Vector part: rather well-known from photo- and electroproduction data.
- Axial part: dominated by $C_5^A(Q^2) = \frac{C_5^A(0)}{(1+Q^2/M_{A\Delta}^2)^2}$.
  $C_5^A(0) \approx 1.2 \propto f^*$. $M_{A\Delta}$: fits to ANL/BNL data.

- Together with Krzystof and Jan we consider single pion production models containind $\Delta(1232)$ resonance and background.
- Problem: default HNV model does not reproduce electromagnetic or weak data in a satisfactory way.
- My part: providing code which calculates cross sections or proton structure functions for different form factor parametrizations.
- C++ code handles Dirac and Lorentz structures present in nucleon$\longrightarrow$nucleon+pion transition current.
- Fits performed using KG's fitting code. Recentlty- electromagnetic part with F_2^p data from Osipenko et al."The Proton structure function F_2 with CLAS",arXiv:hep-ex/0309052".
- Very fast algorithms using Minuit/MIGRAD.
- Data inclusive: cut in invariant mass to 2π threshold.
- Still work in progress: additional constraints on helicity amplitudes are needed, but at the same time model-dependent!



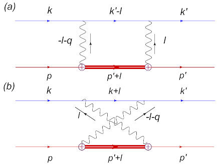




- Quantum Field like approach:
  MT(60), Blunden et al., PRL**91** (2003) 142304, PRC**72** (2005) 034612,... ($\gamma\gamma$), Zhou et al.,PRC **81** (2010) 035208 ($Z^0\gamma$).
  - take the Nucleon, $\Delta(1232)$ (P_{33}),..., to model hadronic intermediate state;
  - off-shell electro-weak $\rightarrow$ on-shell nucleon vertices;
- Dominant contribution from elastic state, *Kondratyuk et al.PRL***95** *(2005) 172503*
- Agreement (in low and intermediate Q^2 range) with **dispersion calculations** by Borisyuk and Kobushkin, PRC **78** (2008) 025208;

$$W_{TPE} \equiv \sum_{spin} 2\mathrm{Re}\left\{(i\mathcal{M}_{1\gamma})^* i\mathcal{M}_{2\gamma}\right\}, \quad \delta_{2\gamma} = \frac{W_{TPE}}{\sum_{spin} \left|i\mathcal{M}_{1\gamma}\right|^2}$$

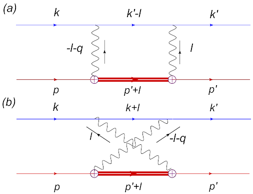
$$W_{TPE} = \frac{1}{2}\frac{e^2}{Q^2}\mathrm{Im}\left\{w_N^{\parallel} + w_N^{\times} + w_\Delta^{\parallel} + w_\Delta^{\times}\right\}.$$



One should subtract Mo-Tsai TPE correction from the data!

$$\Delta_{2\gamma} = \delta_{2\gamma}(full) - \delta_{2\gamma}(MT),$$

- **ELASTIC TPE**: nucleon intermediate state
- **INELASTIC TPE**: $P_{33}(1232), \ldots$ intermediate states



N or $\Delta(1232)$ resonance as intermediate state.

$$j^{\mu\nu}_{\parallel} = \overline{u}(k')\gamma^{\mu}(\hat{k'} - \hat{l} + m)\gamma^{\nu}u(k)$$
$$j^{\mu\nu}_{\times} = \overline{u}(k')\gamma^{\mu}(\hat{k} + \hat{l} + m)\gamma^{\nu}u(k)$$
$$j^{\mu} = \overline{u}(k')\gamma^{\mu}\gamma^{\nu}u(k)$$
$$h_{\mu} = \overline{u}(p')\Gamma_{\mu}(q+l)u(p)$$

$$L^{\alpha\mu\nu}_{\parallel,\times} \equiv \sum_{spin} j^{\alpha *} j^{\mu\nu}_{\parallel,\times}$$
$$\mathcal{H}^{N,\Delta}_{\alpha\mu\nu} \equiv \sum_{spin} h^{*}_{\alpha} h^{N,\Delta}_{\mu\nu}$$

$$h^{N}_{\mu\nu} = \overline{u}(p')\Gamma_{\nu}(-l)(\hat{p'} + \hat{l} + M_p)\Gamma_{\mu}(q+l)u(p)$$
$$h^{\Delta}_{\mu\nu} = \overline{u}(p')\Gamma^{\Delta,in}_{\mu\xi}(-l, p'+l)\left[\hat{p'} + \hat{l} + M_{\Delta}\right] \times \Lambda^{\xi\eta}(p'+l)\Gamma^{\Delta,out}_{\eta\nu}(q+l, p'+l)u(p).$$

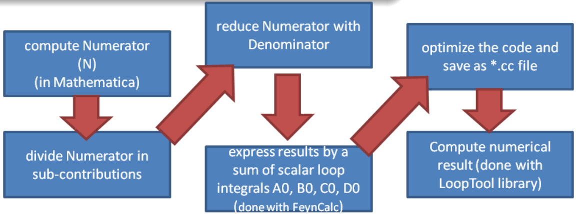
$$w^{\parallel}_{N,\Delta} = e^4 \int \frac{d^4 l}{(2\pi)^4} \frac{L^{\alpha\mu\nu}_{\parallel}\mathcal{H}^{N,\Delta}_{\alpha\mu\nu}}{[(q+l)^2 + i\epsilon][l^2 + i\epsilon][(k'-l)^2 - m^2 + i\epsilon][(p'+l)^2 - M^2_{p,\Delta} + i\Gamma_{\Delta}M_{\Delta}]}$$
$$w^{\times}_{N,\Delta} = e^4 \int \frac{d^4 l}{(2\pi)^4} \frac{L^{\alpha\mu\nu}_{\times}\mathcal{H}^{N,\Delta}_{\alpha\mu\nu}}{[(q+l)^2 + i\epsilon][l^2 + i\epsilon][(k+l)^2 - m^2 + i\epsilon][(p'+l)^2 - M^2_{p,\Delta} + i\Gamma_{\Delta}M_{\Delta}]}.$$

$$\int \frac{d^4 l}{(2\pi)^4} \frac{N(l^2, p' \cdot l, q \cdot l, k' \cdot l)}{[(q+l)^2 - m_b^2]^n [l^2 - m_c^2]^m [(q+l)^2 + i\epsilon][l^2 + i\epsilon][(k'-l)^2 - m^2 + i\epsilon][(p'+l)^2 - M_{N,\Delta}^2 + i\epsilon]}$$

$n, m = 1, 2.$

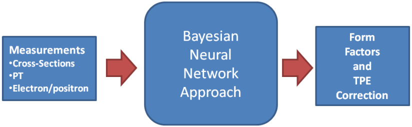
- For $\Box_N$ diagrams $N < l^5$
- For $\Box_\Delta$ diagrams $N < l^7$

$$F_i(t)/C_i^V(t) = \sum_{k=1} \sum_{j=1} \frac{f_i^{k,j}/c_i^{k,j}}{(t - M_{i,k,j}^2)^k},$$



C:/Users/kgraczyk/Documents/Tex/Publications/Preprints/arXiv1306.5991/mathematica/electron_proton_D/stabilne/electron_proton_Delta_5.3.nb

- Wyliczenie pełnej amplitudy dla P_{33} programem Mathematica (12 dni), więcej czasu potrzeba dla obliczenia ekwiwalentu dla CCQE.
- Kompilacja programu (C++) po optymalizacji: około 32h, wymagane conajmniej 18 GB RAM bądź 16 GB RAM i swap.
- Problemy z użyciem biblioteki FeynCalc i LoopTool
- Należało napisać własną implementację algorytmu rozkładającego amplitudy...
- problemy ze stabilnoscią wyników
- pochodne liczone numerycznie
- Amplituda zapisana jest w 44 plikach $*$.cc dla dango przykładu modelu fizycznego
- Fit modelu TPE (elastyczne plus nieelastyczne wkłady) do danych elastycznych ep ponad dwa tygodnie pracy.



- Theoretical problems with estimation of the form factors and TPE? $\leftrightarrow$ what about measurements?
- Represent G_{Ep}, G_{Mp} as well as $\Delta\tilde{C}_{2\gamma}$ by statistical model, which is based on the experimental measurements;
- Introduce the objective method for distinguishing between models

Neural Networks in Physisc

- HEP (Detector Physics), Particle reconstruction and identification, *Denby, Computer Physics Communications 49 (1988), 429; Kurek, Rondio, Sulej, Zaremba, Meas. Sci. Technol. 18 (2007) 2486;....*
- Nucleon Structure Functions: Ball et al. Nucl.Phys. B874 (2013) 36; Askanazi et al. arXiv:1309.7085; Kumericki, et al. JHEP 1107 (2011) 073

Physical quantities $\Leftarrow$ model independent $\Leftarrow$ Measurements

- Similar idea: NNPDF group, parton distribution functions parametrized by neural networks
  First paper: JHEP 0205 (2002) 062,
- Consider as many as possible neural network parametrizations, and classify them with a help of Bayesian algorithm. Then a physical observable/desired quantity is given by an average over all models.

$$\langle \mathcal{O}(\mathcal{N})\rangle = \int_{\mathcal{N}\in\mathcal{F}} \mathcal{D}\mathcal{N}\,\mathcal{O}(\mathcal{N})\mathcal{P}(\mathcal{N})$$

- **Searching for the model which will have ability to give good predictions outside current data domain (GOOD PREDICTIVE POWER)**

- feed-forward network with one hidden layer to model G_{Ep};
- * Cybenko theorem, Math. Control Signals System (1989) 2, 303;

$$G_{Ep}(Q^2, \vec{w}) : \mathcal{R} \ni Q^2 \to G_{Ep} \in \mathcal{R}$$

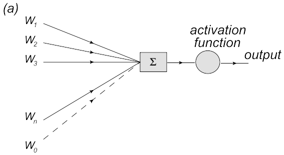
- Sigmoid function as an activation function – limited support.

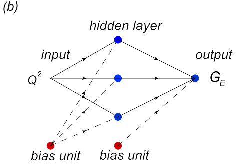
$$f_{act.}\left(\sum_{i=0}^{n} w_i f^i_{act.}(previous\ layer)\right),$$

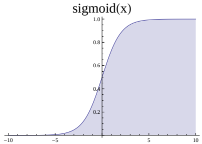
- N experimental points: $\mathcal{D}$: $\{(x_1, t_1, \Delta t_1), \ldots, (x_N, t_N, \Delta t_N)\}$
- Error function

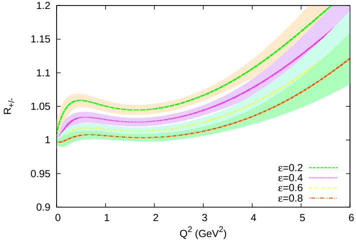
$$S_{ex}(\vec{w}, \mathcal{D}) = \underbrace{\sum_{i=1}^{N} \left(\frac{y(x_i, \vec{w}) - t_i}{\Delta t_i}\right)^2}_{\chi^2(\vec{w}, \mathcal{D})} + ?!$$

- "Teaching a network": searching for w_i's and α parameters.
- Overfitting problem – bias-variance trade off problem.

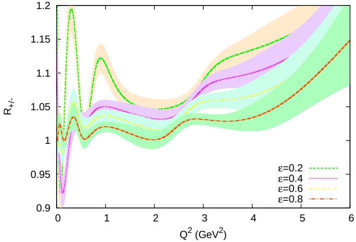








optimal model (ev=-547.968 E=545.01 H/G=7/4)



overfitted model (ev=-646.099 E=381.972 H/G=15/6)

Figs. obtained from updated (not yet published) analysis of TPE effect by improved version of WNet

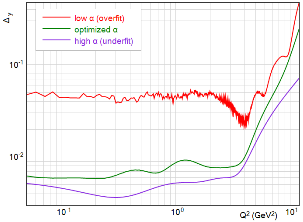


Fig. done by R. Sulej, with NetMaker (Neutron Form Factor)

- Overfitted models reproduce the statistical errors and overestimate the uncertainties!
- Let's introduce the penalty term,

$$S_{ex}(\vec{w},\mathcal{D}) \to S_{ex}(\vec{w},\mathcal{D}) + \alpha \underbrace{(w_1^2 + \ldots + w_W^2)}_{E_W(\vec{w})}.$$

- Cross-validation (applied by **NNPDF**),..., and many others approaches,...
- Bayesian approach!?

The Bayesian framework for the feed forward neural networks was developed to provide consistent and objective methods, which allow to

- establish optimal structure of the network (in practise number of the hidden units, layers);
- find optimal values of the weights and the α parameters;
- establish optimal values of the learning algorithm parameters;
- compute the neural network output uncertainty, and uncertainties for the weight and α parameters.
- classify and compare models quantitatively.

* MacKay, Neural Computation 4 (3), (1992) 415; Neural Computation 4 (3), (1992) 448;
Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press 2008

Occam's razor

Method in natural way prefers simpler than complex models

WNet

- C++ library developed for TPE analysis;
- can be applied for any analysis;
- several learning algorithms;
- linux application (run from the command line level)
- Authors: K.M.G

* (in the preliminary stage cooperation with Rober Sulej (NetMaker) and Piotr Płoński)

The BNN approach requires minimal input from the user. The idea of the approach is to replace the user's common sense by the mathematical objective procedures. Obviously some user's input is necessary.

- $\mathcal{S} \in$ neural networks with different number of units.
- Let's restrict the set S to the neural networks with only one hidden layer (the choice supported by Cybenko theorem).
- Each network $\mathcal{N}_\beta \in \mathcal{S}$ ($\beta = 1, 2, \ldots$) approximates physical quantities (Form Factors) based on the data $\mathcal{D}$.
- Let's introduce the conditional probability $P(\mathcal{N}_\beta|\mathcal{D})$.

$$\langle \mathcal{O}(\mathcal{N}) \rangle_{\mathcal{D}} = \int_{\mathcal{S} \in \mathcal{N}} \mathcal{D}\mathcal{N} \mathcal{O}(\mathcal{N}) P(\mathcal{N}_\beta|\mathcal{D})$$

- The evidence $P(\mathcal{D}|\mathcal{N}_\beta)$,

$$P(\mathcal{N}_\beta|\mathcal{D}) = \frac{P(\mathcal{D}|\mathcal{N}_\beta) P(\mathcal{N}_\beta)}{P(\mathcal{D})}.$$

$P(\mathcal{D})$ – normalization factor, does not depend on $\mathcal{N}_\beta$.

- $P(\mathcal{N}_\beta)$ – prior probability, at the beginning of any analysis there is no reason to prefer a particular model (network),

$$P(\mathcal{N}_1) = P(\mathcal{N}_2) = P(\mathcal{N}_3) = \ldots$$

- the evidence differs from $P(\mathcal{N}_\beta|\mathcal{D})$ by only a constant normalization factor.
- Let's introduce the physical initial assumptions,

$$\mathcal{P}(\mathcal{D}|\mathcal{N}_\beta) \to \mathcal{P}\left(\mathcal{D}|\{\mathcal{I}_{Phys.}\}, \mathcal{N}_\beta\right)$$

Hierarchical Approach

Pattern

$$P(\vec{w}|\mathcal{D}, \alpha, \mathcal{N}) = \frac{\mathcal{P}(\mathcal{D}|\vec{w}, \alpha, \mathcal{N})\mathcal{P}(\vec{w}|\alpha, \mathcal{N})}{\mathcal{P}(\mathcal{D}|\alpha, \mathcal{N})}$$

$$\mathcal{P}(\alpha|\mathcal{D}, \mathcal{N}) = \frac{\mathcal{P}(\mathcal{D}|\alpha, \mathcal{N})\mathcal{P}(\alpha|\mathcal{N})}{\mathcal{P}(\mathcal{D}|\mathcal{N})}$$

$$\mathcal{P}(\mathcal{N}|\mathcal{D}) = \frac{\mathcal{P}(\mathcal{D}|\mathcal{N})\mathcal{P}(\mathcal{N})}{\mathcal{P}(\mathcal{D})}$$

- $* \to *$

- 27 cross section data sets, 50 years of measurements;

* **Systematic normalization error has to be taken into account!**

$$\sigma^R_{1\gamma+2\gamma}(Q^2,\epsilon) = \sigma^R_{1\gamma}(Q^2,\epsilon) + \Delta C_{2\gamma}(Q^2,\epsilon)$$
$$\sigma^R_{1\gamma}(Q^2,\epsilon) = \tau G_M^2(Q^2) + \epsilon G_E^2(Q^2)$$

- 14 PT, form factor ratio $\mu_p G_E/G_M$ data sets, about 14 years of measurements;

$$\mathcal{R}_{1\gamma}(Q^2) = \mu_p \frac{G_E(Q^2)}{G_M(Q^2)}$$

* **I assumed that PT data does not fill TPE correction!**

- **3** data sets for $d\sigma(e^+p)/d\sigma(e^-p)$.

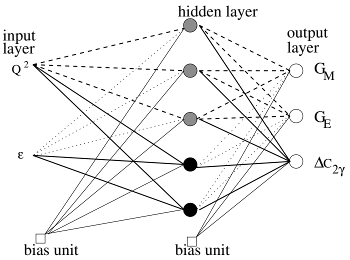
$$\mathcal{R}_\pm(Q^2,\epsilon) = 1 - \frac{2\Delta C_{2\gamma}(Q^2,\epsilon)}{\sigma_{1\gamma+2\gamma,R}(Q^2,\epsilon)}.$$

Three unknown functions and three types of the data sets!

Assumption

- The PT data is less sensitive to the TPE correction than the cross section measurements, hence the TPE contribution to $\mathcal{R}_{1\gamma}$ ratio can be neglected.
- Inconsistency between PT and cross-section data should allow to find BNN a missing correction, interpreted as TPE...

$$\mathcal{N}_{g,t}: \mathbb{R}^2 \mapsto \mathbb{R}^3, \quad \mathcal{N}_{g,t}(Q^2,\epsilon;\vec{w}) = \begin{pmatrix} G_M \\ G_E \\ \Delta C_{2\gamma} \end{pmatrix}.$$



ANN: 2-(3-2)-3, $\mathcal{N}_{3,2}$, 3 months of working of 40 CPU now, 5 days of working with wcss!

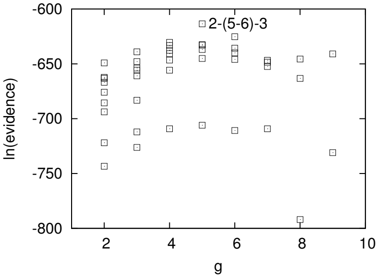
Evidence

The logarithm of the evidence (only model independent terms are written) reads

$$\ln \mathcal{P}\left(\mathcal{D}|\mathcal{A}_{g,t}\right) \approx -S_{ex}(\mathcal{D},\vec{w}_{MP}) + \frac{W}{2}\ln\alpha_{MP} - \alpha_{MP}E_w(\vec{w}_{MP}) - \frac{\ln|A|}{2} - \frac{1}{2}\ln\frac{\gamma}{2} + \underbrace{(g+t)\ln(2) + \ln(g!) + \ln(t!)}_{symmetry\ factor},$$

W – number of weigh parameters; $|A|$ is determinant of the hessian matrix: $A_{ij} = \nabla_i \nabla_j\ S_{ex}|_{\vec{w}=\vec{w}_{MP}} + \alpha_{MP}$.

- Misfit the approximated data usually of low-value
- Occam factor penalizes the complex models.
- Symmetry contribution to this quantity is given be the symmetry factor.

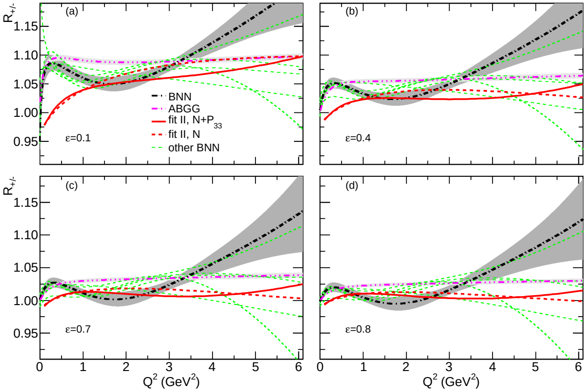


Network 2-(5-6)-3 is the most suitable to describe the data. K.M.G. Phys.Rev. C84 (2011) 034314

- About 1000 training processes sucesfull for every discussed architecture;
- 45 different architectures;
- $4 \leq g + t = M \leq 12$;

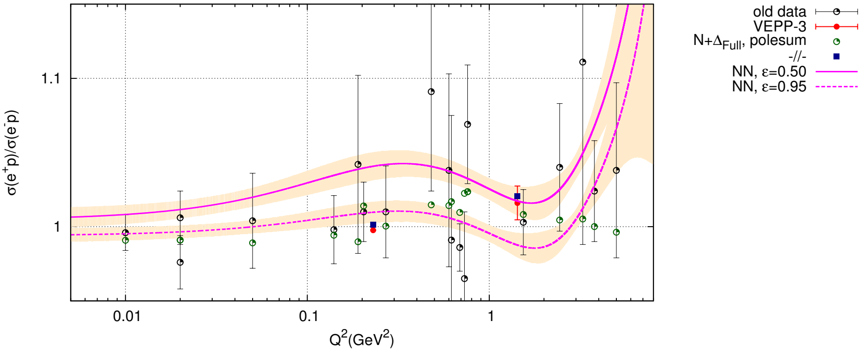
$$\left\langle \mathcal{O}(G_E, G_M, \Delta\tilde{C}_{2\gamma})(Q^2,\epsilon)\right\rangle = \sum_{m=1}^{\infty}\sum_{g=1,t=1}^{g+t=m} \mathcal{O}(G_E^{\mathcal{N}_{g,t}}, G_M^{\mathcal{N}_{g,t}}, \Delta\tilde{C}_{2\gamma}^{\mathcal{N}_{g,t}})\mathcal{P}(\mathcal{N}_{g,t}) \quad (1)$$

$$\approx \mathcal{O}(G_E^{\mathcal{N}_{5,6}}, G_M^{\mathcal{N}_{5,6}}, \Delta\tilde{C}_{2\gamma}^{\mathcal{N}_{5,6}}) \quad (2)$$



$$R_{+/-} = \frac{d\sigma(e^+p)}{d\sigma(e^-p)} \approx 1 - 2\Delta_{2\gamma}$$

- **Rejected models** accseptable due to conventional non-Bayesian point of view...
- **ABGG**: conventional analysis with $\chi^2/NDF < 1$, as well as with fits of the Form Factors consistent with other phenomenological discussions...
- But **ABGG** does not agree with theoretical predictions...



- BNN and hadronic model predictions agree well with measurements.

$$R_{+/-} = \frac{d\sigma(e^+p)}{d\sigma(e^-p)} \approx 1 - 2\Delta_{2\gamma}$$

- $\nu_\mu \to \nu_e$: θ_{13} measurement and then CP violation parameter, in present and future neutrino experiments (e.g. T2K, NOνA).
-
  - $E_\nu \sim 1$ GeV.
  - Charged Current Quasi-Elastic (CCQE) Scattering: a dominant process
  - estimate of the systematic differences between $\nu_e N$ and $\nu_\mu N$ CCQE cross sections important for data analysis:
    Day, McFarland, PRD**86** (2012) 053003, also a talk by M. Day, **NuFact12**.
  - Are the Radiative Corrections (RC's) a potential source of difference between cross sections for ν_e and ν_μ scattering?
  - * $m_e \ll m_\mu$: **electron radiates more than muon!**
  - **What about the problem of axial mass effect?**