Axial nucleon form factor from a Bayesian analysis of the neutrino scattering data
ISNET-6

Krzysztof M. Graczyk

work with Luis Alvarez-Ruso and Eduardo Saul-Sala

Institute for Theoretical Physics

University of Wrocław
Poland

October 11, 2018

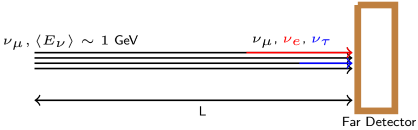
Uniwersytet Wrocławski

Motivation

- Neutrinos ν_e, ν_μ, ν_τ,..., fundamental particles
- weakly interacting particles, neutral
- neutrino oscillations:
  - neutrino oscillation parameters, mass differences and mixing angles
  - CP-violation phase
  - mass hierarchy
- ...

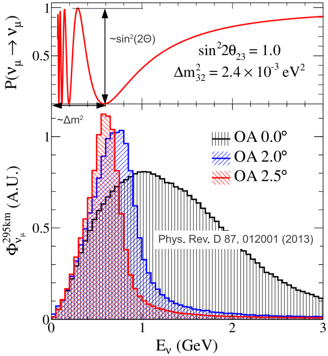
A need of better knowledge of ν-nuclei, ν-nucleon scattering cross sections (in particular for $E_\nu \sim 1$ GeV)

ν-A Interactions



$$P(\nu_\mu \to \nu_\tau) = \sin^2\theta_{23}\left(\frac{\Delta m^2_{32} L}{4E_\nu}\right)$$

- Neutrino energy reconstructed from the analysis of QuasiElastic (QE) scattering events!

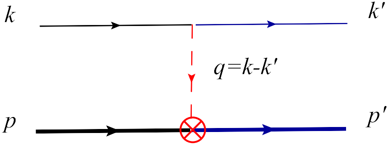


- accelerator source of neutrinos $\to E_\nu \sim 1$ GeV but neutrino energy not monochromatic: E_ν known with some precision!

Axial Nucleon Form Factor

- Consider **Charged Current Quasielastic** (CCQE)

$$\nu(k) + n(p) \to \mu^-(k') + p(p')$$



$$i\mathcal{M}_{Born} \approx i\frac{G_F \cos\theta_C}{\sqrt{2}} \cdot \overline{u}(k')\gamma^\mu(1-\gamma_5)u(k) \times \overline{u}(p')\Gamma^\mu(q)u(p)$$

- q^μ - energy-momentum transfer
- Γ^μ has a V^μ-ector $-$ A^μ-xial structure

- Vector part expressed in terms electromagnetic nucleon form factors
- *$Axial$*:

$$A^{\mu} = \left(\gamma_{\mu}\gamma_5 + \frac{q^{\mu}\gamma_5}{2M}\frac{4M^2}{m_{\pi}^2 + Q^2}\right) F_A(Q^2)$$

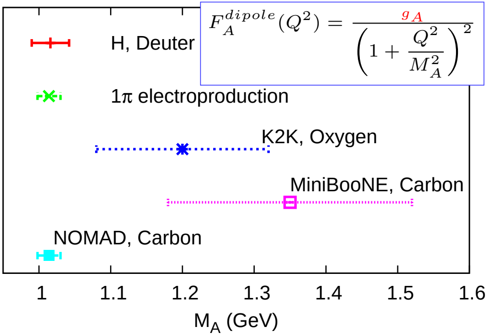
  where

$$Q^2 = -q^{\mu}q_{\mu} \equiv -q^2 > 0$$

- Neutron β-decay asymmetry $F_A(0) = g_A = 1.2723 \pm 0.0023$. C. Patrignani et al., PDG, Chin. Phys. C40, 100001 (2016)

- F_A an input for the ν-nucleus scattering models
- ν-nucleon and ν-nucleus cross sections dominated by the axial contribution

Dipole Axial Nucleon Form Factor

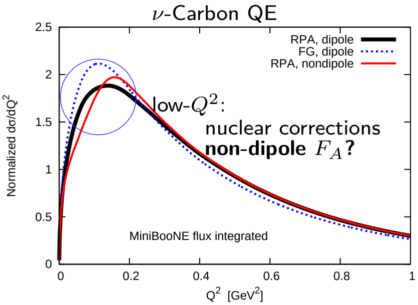


$$F_A^{dipole}(Q^2) = \frac{g_A}{\left(1 + \frac{Q^2}{M_A^2}\right)^2}$$

The most reliable information about F_A: bubble chamber data i.e. $\nu - d$ scattering (ANL, BNL, and FNAL, CERN data)

Recent measurements: ν-Carbon, ν-Oxygen, ... scattering: extraction of F_A is nuclear model dependent!

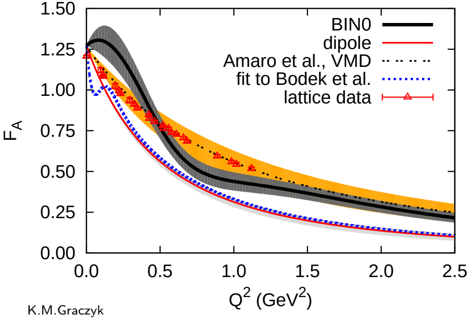
Uniwersytet Wrocławski



Non-dipole dependence of F_A?

Electric and magnetic form factors of the nucleon have non-dipole Q^2-shape!

recent studies: Bodek, et al. EPJC53, 349 (duality constrained); Adamuscin et al. PRC78 035201 (monopole); Amaro et al. PRD93, 053002 (VMD, two monopoles); Meyer et al., PRD93, 113015 (z-expansion)



Goal

- analyze the bubble chamber ν-deuteron scattering data trying to find deviations of F_A from dipole shape
- obtain predictions of F_A with uncertainties

- F_A obtained from the data
- consider the class of flexible functions $\rightarrow$ feed-forward neural networks

$$F_A(Q^2) = F_A^{\text{dipole}}(Q^2) \times \mathcal{N}_M(Q^2; \{w_i\})$$

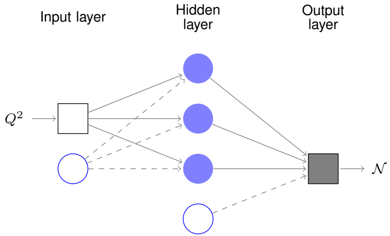
- choose the "best" fit using Bayesian algorithm

* (used to extract G_E, G_M and 2γ correction from elastic ep scattering data)

Main assumptions:

(I) $F_A(Q^2)$ is assumed to be a continuous function of Q^2 in its validity domain

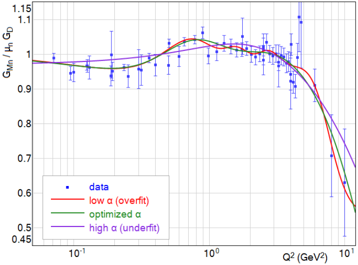
(II) As $F_A(Q^2)$ is bounded, there must be a $C > 1$: $F_A(Q^2)/F_A^{\text{dipole}}(Q^2) < C$ in the whole Q^2 interval of $(0, 3)$ GeV^2

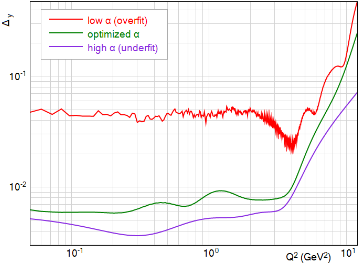


Bias-variance trade off

- too complex fit reproduces the statistical noise and overfits the data $\rightarrow$ poor generalization
- too simple fit underfits the data

* unrealistic uncertainties





Bayesian statistics embodies Occam's razor

Bayesian Framework for neural networks: MacKay, Neural Computation 4 (3), (1992) 415; Neural Computation 4 (3), (1992) 448

- consider function $\mathcal{N}$ (depending on W weights $\{w_i\}$)

$$P(\{w_i\}|\mathcal{D},\mathcal{N}) = \frac{\overbrace{P(\mathcal{D}|\{w_i\},\mathcal{N})}^{likelihood}\overbrace{P(\{w_i\}|\mathcal{N})}^{prior}}{\underbrace{P(\mathcal{D}|\mathcal{N})}_{evidence}}$$

- Evidence for model $\mathcal{N}$

$$P(\mathcal{D}|\mathcal{N}) = \int \prod_k dw_k \, P(\mathcal{D}|\{w_i\},\mathcal{N})P(\{w_i\}|\mathcal{N})$$

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

  $P(\mathcal{D})$ - normalization
- Assume that $P(\mathcal{N}_1) = P(\mathcal{N}_2) = ...$ then $P(\mathcal{D}|\mathcal{N})$ ranks models...

- Let consider the N independent measurements y_i with uncertainty σ_i $\mathcal{D} = \{(y_1(x_1), \sigma_1), ...\}$
- likelihood (of the Gaussian form)

$$2 \ln \mathcal{P}(\mathcal{D}|\{w_i\}, \mathcal{N}) \sim -\chi^2_{ex}(\mathcal{D}, \{w_i\}, \quad \chi^2_{ex}(\mathcal{D}, \mathcal{N}) = \sum_{i=1}^{N} \frac{(y_i(x_i) - f(\mathcal{N}))^2}{\sigma_i^2}$$

- prior

$$\mathcal{P}(\{w_i\}|\alpha, \mathcal{N}) \quad \sim \quad \exp\left(-\frac{\alpha}{2} E_w\right), \quad E_w = \sum_{k=1}^{W} w_k^2$$

  α regularizer (hyperparameter), W - number of weights (neural network parameters)
- Posterior

$$2 \ln P(\{w_i\}|\mathcal{D}, \mathcal{N}, \alpha) \sim -(\chi^2_{ex}(\mathcal{D}, \{w_i\}) + \alpha E_w) = -\mathcal{E}(\mathcal{D}, \{w_i\})$$

- $\mathcal{E}(\mathcal{D}, \{w_j\})$ has a minimum at $\{w_j\}_{MP}$
- α: overfit, underfit $\to$ optimal model

- Analytic calculations (in Gaussian approximation)

$$
\begin{aligned}
\mathcal{P}\left(\{w_i\}|\,\mathcal{D},\alpha,\mathcal{N}\right) &= \frac{\mathcal{P}\left(\mathcal{D}|\,\{w_i\},\alpha,\mathcal{N}\right)\mathcal{P}\left(\{w_i\}|\,\alpha,\mathcal{N}\right)}{\mathcal{P}\left(\mathcal{D}|\,\alpha,\mathcal{N}\right)} \rightarrow \\
\mathcal{P}\left(\alpha|\,\mathcal{D},\mathcal{N}\right) &= \frac{\mathcal{P}\left(\mathcal{D}|\,\alpha,\mathcal{N}\right)\mathcal{P}\left(\alpha|\,\mathcal{N}\right)}{\mathcal{P}\left(\mathcal{D}|\,\mathcal{N}\right)} \rightarrow \\
\mathcal{P}\left(\mathcal{N}|\,\mathcal{D}\right) &= \frac{\mathcal{P}\left(\mathcal{D}|\,\mathcal{N}\right)\mathcal{P}(\mathcal{N})}{\mathcal{P}(\mathcal{D})}
\end{aligned}
$$

$$2\alpha_{MP}E_w = \sum_{i=1}^{W}\frac{\lambda_i}{\lambda_i+\alpha_{MP}} \equiv \gamma(\alpha_{MP})$$

λ_i - eigenvalues of $H_{kj} = \nabla_k\nabla_j\chi_{ex}$

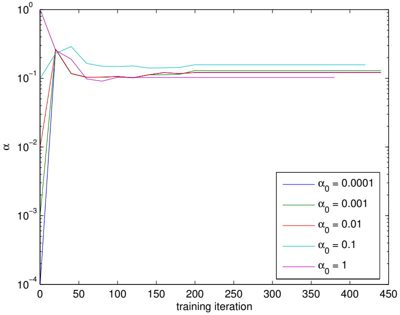
$\rightarrow$ α_{MP}: α is iterated according to (12)
$\alpha_{\mathbf{k+1}} = \gamma(\alpha_{\mathbf{k}})/\mathbf{2E_w}$

- Evidence approximation $\rightarrow$

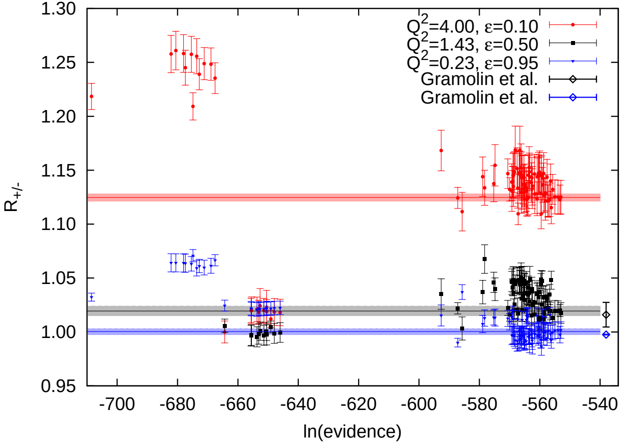
$$\ln\mathcal{P}\left(\mathcal{D}|\,\mathcal{N}\right) \approx -\mathcal{E}(\mathcal{D},\{w_j\}_{MP})/2 - \frac{1}{2}\ln|A|$$

$$+\frac{W}{2}\ln\alpha_{MP} - \frac{1}{2}\ln\frac{\gamma_{MP}}{2} + combinatorial$$

where $A = \nabla_i\nabla_j S(\mathcal{D},\{w_j\})$

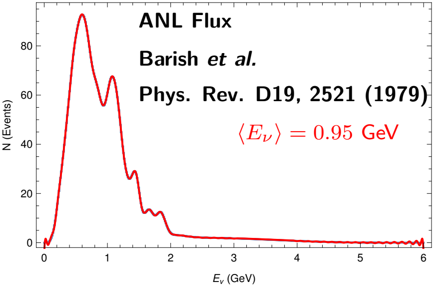


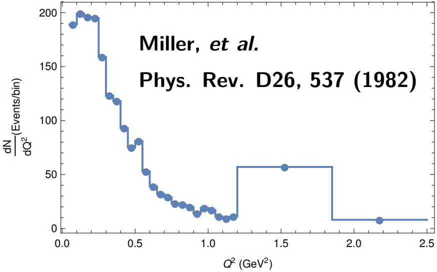
elastic ep scattering data and 2γ exchange correction



Bubble Chamber Data: $\nu_\mu + d \to \mu^- + p + p$

- **ANL** data: Argonne National Laboratory data (today's talk)
- * Distribution of events: $Q^2 \in (0.05, 2.5)$ GeV^2,
- → **BNL**: Brookhaven National Laboratory, $\langle E_\nu \rangle \sim 1.56$ GeV, $Q^2 \in (0.03, 2.8)$ GeV^2
- → **FNAL** data: $\langle E_\nu \rangle \sim 22.6$ GeV, $Q^2 \in (0., 3)$ GeV^2





Analysis: arXiv:1805.00905

- The least square-function:

$$\chi^2_{ex} = \chi^2_{\mathrm{ANL}} + \chi^2_{g_A}$$

where

$$\chi^2_{\mathrm{ANL}} = \sum_{i=k}^{n_{\mathrm{ANL}}} \frac{\left(N_i - pN_i^{th}\right)^2}{N_i} + \left(\frac{1-p}{\Delta p}\right)^2$$

see D'Agostini, NIMA346, 306, Roe, arXiv:1506.09077, Baker and Cousins, NIM221 437

- Δp is a systematic uncertainty for $\#N$ of events
-

$$N_i^{th} = \int_0^\infty dE_\nu \frac{\frac{d\sigma}{dQ^2}(E_\nu, F_A)}{\sigma(E_\nu, F_A)} \frac{dN}{dE_\nu}, \quad \sigma(E_\nu, F_A) = \int_{min}^{max} \frac{d\sigma}{dQ^2}(E_\nu, F_A) dQ^2$$

see Meyer et al. (2017)

- $F_A(Q^2 = 0)$,

$$\chi^2_{g_A} = \left(\frac{F_A(0) - g_A}{\Delta g_A}\right)^2$$

g_A and Δg_A from PDG

Model of $\nu_\mu + d \to \mu^- + p + p$

$$\frac{d\sigma}{dQ^2}(\nu_\mu n \to \mu^- p) = \frac{G_F^2 M}{8\pi E_\nu^2} \times \left[A(Q^2, E_\nu)F_A^2 + B(Q^2, E_\nu)F_A + C(Q^2, E_\nu)\right]$$

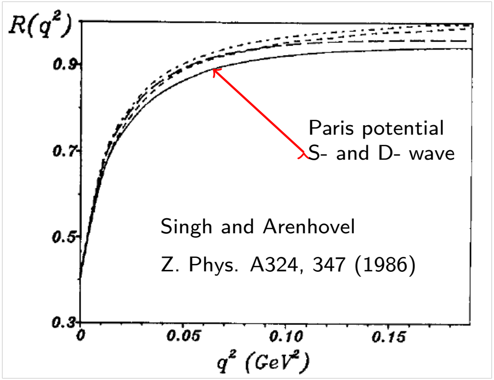
- Deuteron corrections:

$$\frac{d\sigma}{dQ^2} \to R(Q^2)\frac{d\sigma}{dQ^2}$$

where

$$R(Q^2) = \frac{\frac{d\sigma}{dQ^2}(\nu_\mu d \to \mu^- pp)}{\frac{d\sigma}{dQ^2}(\nu_\mu n \to \mu^- p)} \quad (1)$$

- correction concerns first three bins of ANL data!



Analysis

(i) consider network with $M = 1$;

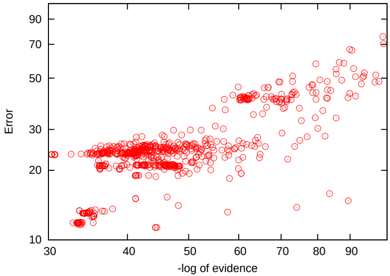
(ii) perform the network training;

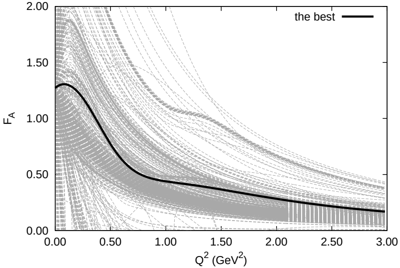
(iii) repeat steps (i)-(iii) for various $\{w_j\}_0$;

(iv) choose the best fit according to the evidence;

(v) repeat steps (i)-(iv) for $M = 2, ...;$

(vi) among the all the best fits choose the model with the highest evidence





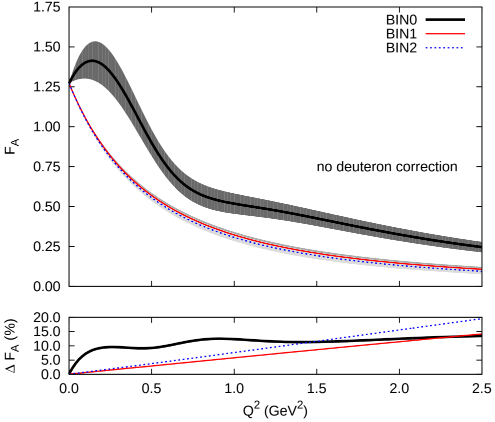
Uniwersytet Wrocławski

Analysis

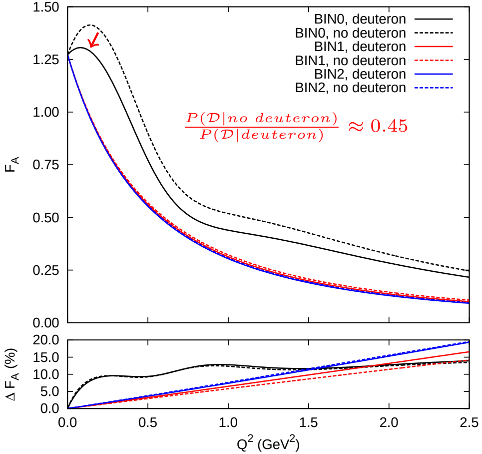
(i) BIN0: all ANL bins included

(ii) BINk: where $k = 1$ or $k = 2$: ANL bins without the first k bins

- Unexpected F_A fit for all bins (BIN0)?!
- The slope of F_A at $Q^2 = 0$ not consistent with other determinations...
- fits of BIN1 and BIN2 data consistent with original ANL analysis w dipole par. (Miller et al.)!
- similar result for BNL data, but not for FNAL data!(preliminary)



Impact of deuteron corrections

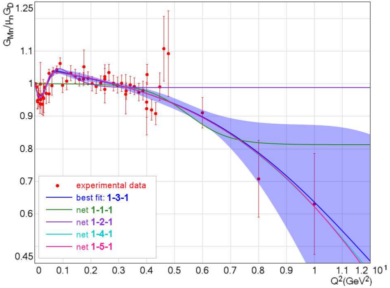


- problem with the first bin?
- deuteron correction important for the first bin

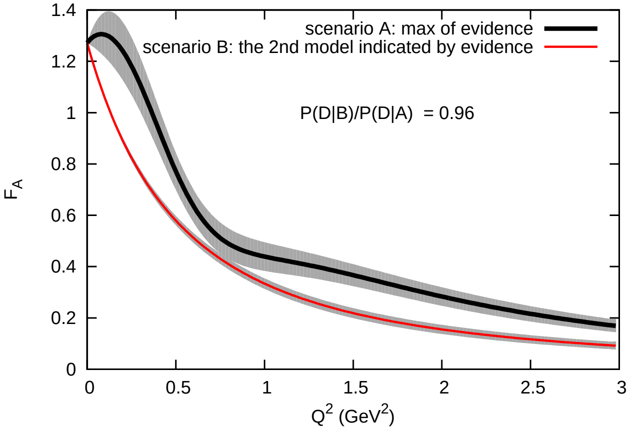
Problem with 1st bin?

Possible explanation:

- a low quality of the measurements at low-Q^2 due to low and not well understood efficiency
- an improper description of the nuclear corrections
- the actual value of the slope $dF_A(Q^2 = 0)/dQ^2$ might not be properly estimated because of the lack of very low-Q^2 data



two competitive scenarios: on the edge of dipole...

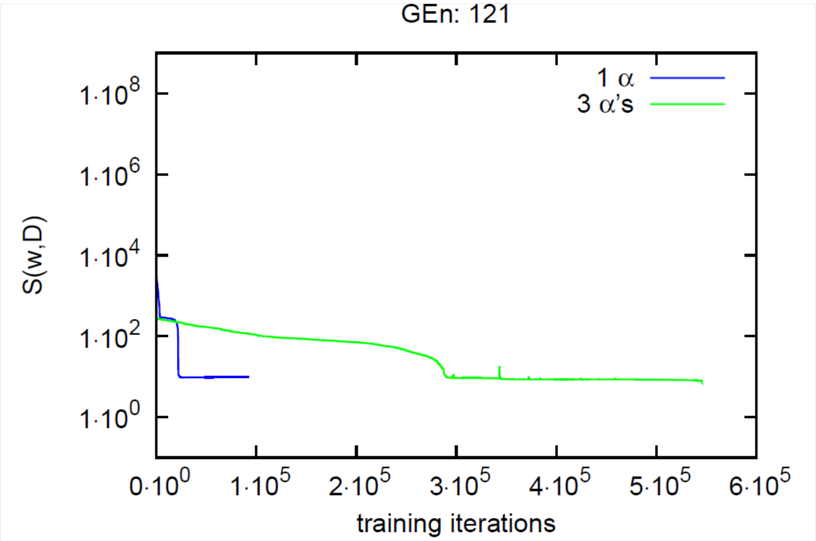


- small effects play a role...
- there is a tension between the first bin and constraint for F_A at $Q^2 = 0$
- assumption: $dF_A(Q^2 = 0)/dQ^2 < 0 \rightarrow$ the fit in agreement with dipole shape...

Conclusions

- The ANL data are not enough informative to study deviation of F_A from the dipole shape
- The fits of BIN1 and BIN2 in agreement with dipole fits
- Future: global analysis of ANL, BNL and FNAL data
- New measurements of ν-H, ν-D necessary

THIS IS THE END


GEn: 121
1 α
3 α's
S(w,D)
1·10^8
1·10^6
1·10^4
1·10^2
1·10^0
0·10^0
1·10^5
2·10^5
3·10^5
4·10^5
5·10^5
6·10^5
training iterations

least-square vs. Poisson

	Fit	N_{fit}	$M_A[MeV]$	$r_A^2[fm^2]$	bins
deuteron -2LL	52.7	1.11 $\pm$ 0.03	1032 $\pm$ 60	0.438 $\pm$ 0.051	49
deuteron χ^2	27.6	1.09 $\pm$ 0.03	1042 $\pm$ 60	0.430 $\pm$ 0.050	25

$$-2LL = 2\sum_i^N \left[N_i^{th} - N_i + N_i \ln\left(\frac{N_i}{N_i^{th}}\right)\right] \tag{2}$$