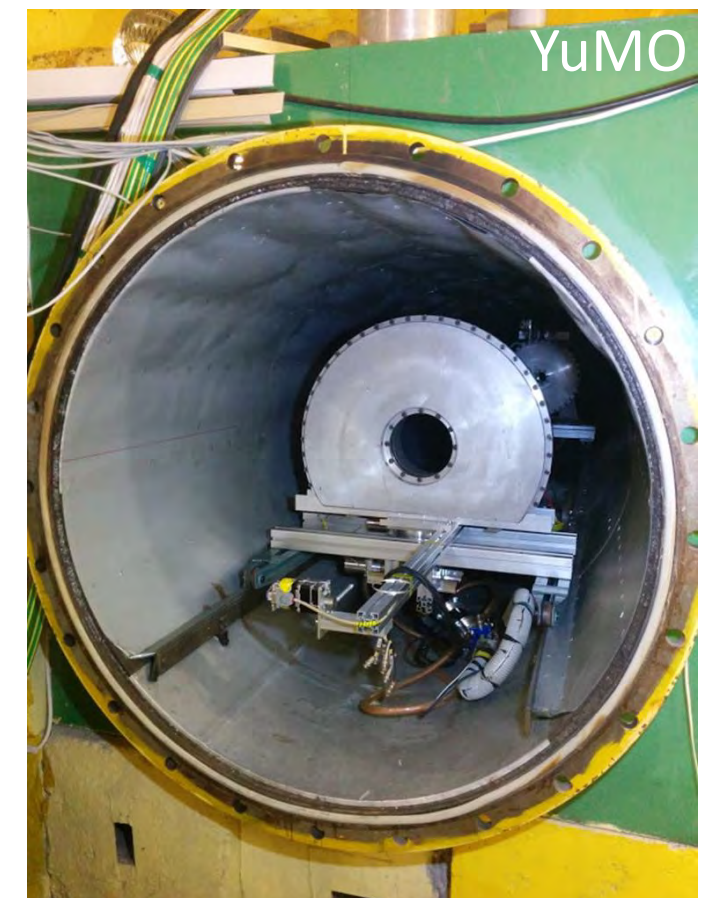




A WEB APPLICATION FOR FITTING EXPERIMENTAL DATA USING JINR CLOUD INFRASTRUCTURE AND ROOT PACKAGE TOOLS

K. Lukyanov, A. Soloviev, T. Solovjeva, E. Zemlyanaya
MLIT, JINR, Dubna, Russia

INTRODUCTION



FITTER

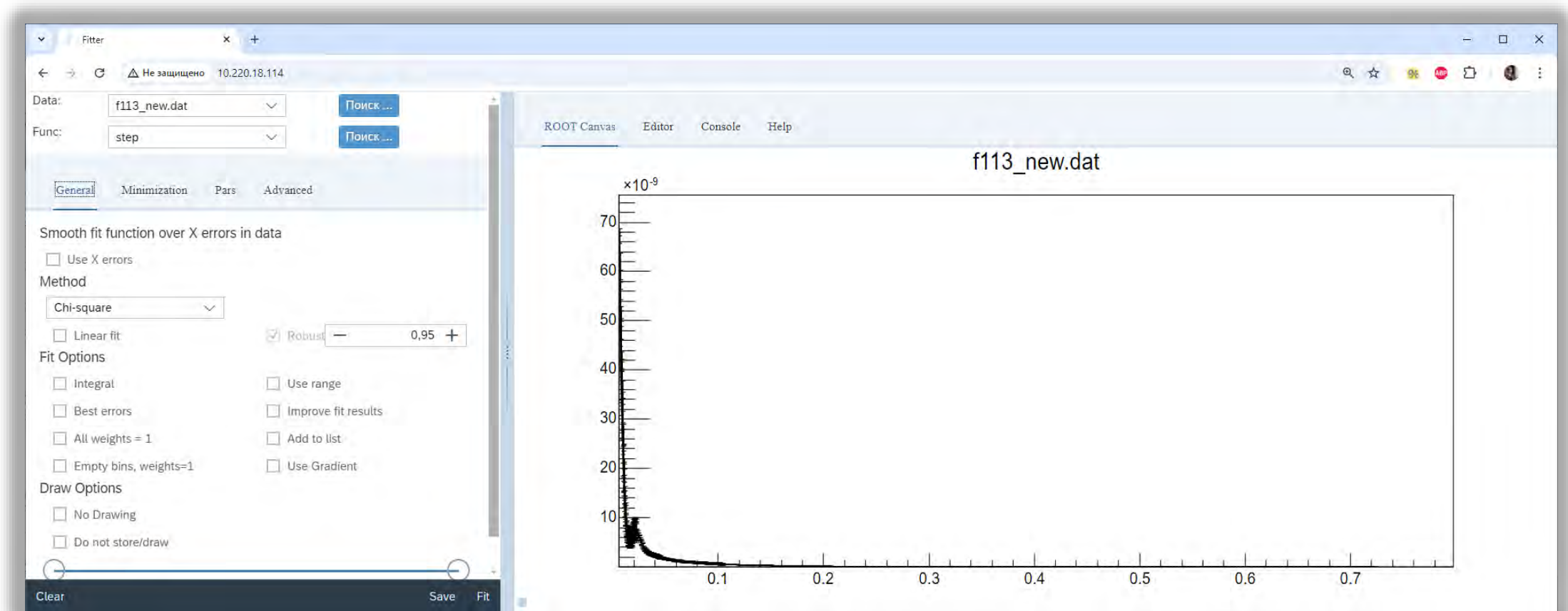
To find the parameters of a theoretical model, it is necessary to minimize the functional, which is a measure of the deviation between the theoretical curve and experimental data. Usually, this is done using the chi-square criterion

$$\chi^2 = \frac{1}{N - N_{params}} \sum_{i=1}^N \left(\frac{f(x_i) - y_i}{\Delta y_i} \right)^2$$

Processing of data obtained on a small-angle neutron scattering spectrometer makes it possible to analyze the structure of the particles being studied. The shape of the particle is approximated by simple geometric bodies - ellipsoids, cylinders, prisms. The analytical equations for these bodies are part of the FITTER.

WEB-BASED VERSION OF FITTER

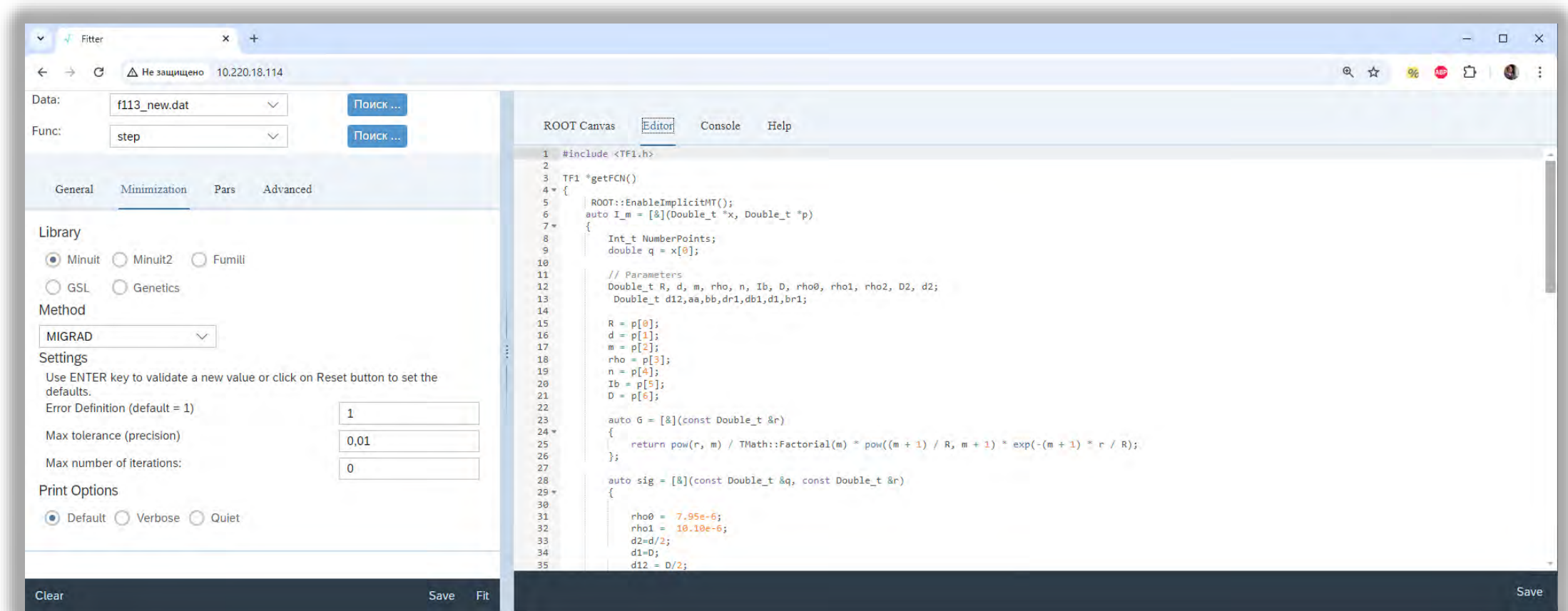
<http://fitter.jinr.ru/>



Soloviev A.G., Solovjeva T.M., Lukyanov K.V. Deployment of a web application for fitting experimental data in the JINR cloud infrastructure // PEPAN, vol.55, issue 3, pp.627-631, 2024.

WEB-BASED VERSION OF FITTER

<http://fitter.jinr.ru/>



Soloviev A.G., Solovjeva T.M., Lukyanov K.V. Deployment of a web application for fitting experimental data in the JINR cloud infrastructure // PEPAN, vol.55, issue 3, pp.627-631, 2024.

WEB-BASED VERSION OF FITTER

<http://fitter.jinr.ru/>

The screenshot shows the Fitter web application interface. The left panel contains input fields for 'Data' (f113_new.dat) and 'Func' (step), along with tabs for 'General', 'Minimization', 'Pars', and 'Advanced'. The 'Advanced' tab is active, showing 'Number of Points' (40), 'Parameter 1' (R), 'Parameter 2' (d), 'Confidence Level' (0.683), and a 'Fill Colour' selector. The right panel shows the 'Console' output, which includes the following text:

```

step
FCN=609.53 FROM MIGRAD STATUS=CONVERGED 556 CALLS 557 TOTAL
EDM=7.47518e-08 STRATEGY= 1 ERROR MATRIX UNCERTAINTY 2.1 per cent
EXT PARAMETER STEP FIRST
NO. NAME VALUE ERROR SIZE DERIVATIVE
1 R 2.04487e+02 2.74779e+00 -1.19515e-03 -7.52879e-05
2 d -5.11556e+01 9.98987e-01 -4.99653e-04 -7.28284e-04
3 m 1.75510e+01 1.65200e+00 -2.11334e-04 1.80507e-04
4 rho 9.71474e-06 1.48309e-08 3.09544e-12 -1.54780e+05
5 n 2.75983e-09 1.09501e-10 -3.45041e-14 -2.54062e+06
6 lb 1.86774e-11 2.53199e-13 3.56813e-16 -1.13917e+08
7 D 1.99368e+01 3.08499e-01 -4.75301e-06 -3.72398e-03

```

Below the console output, the minimizer is identified as Minuit / Migrad, and the following statistics are displayed:

```

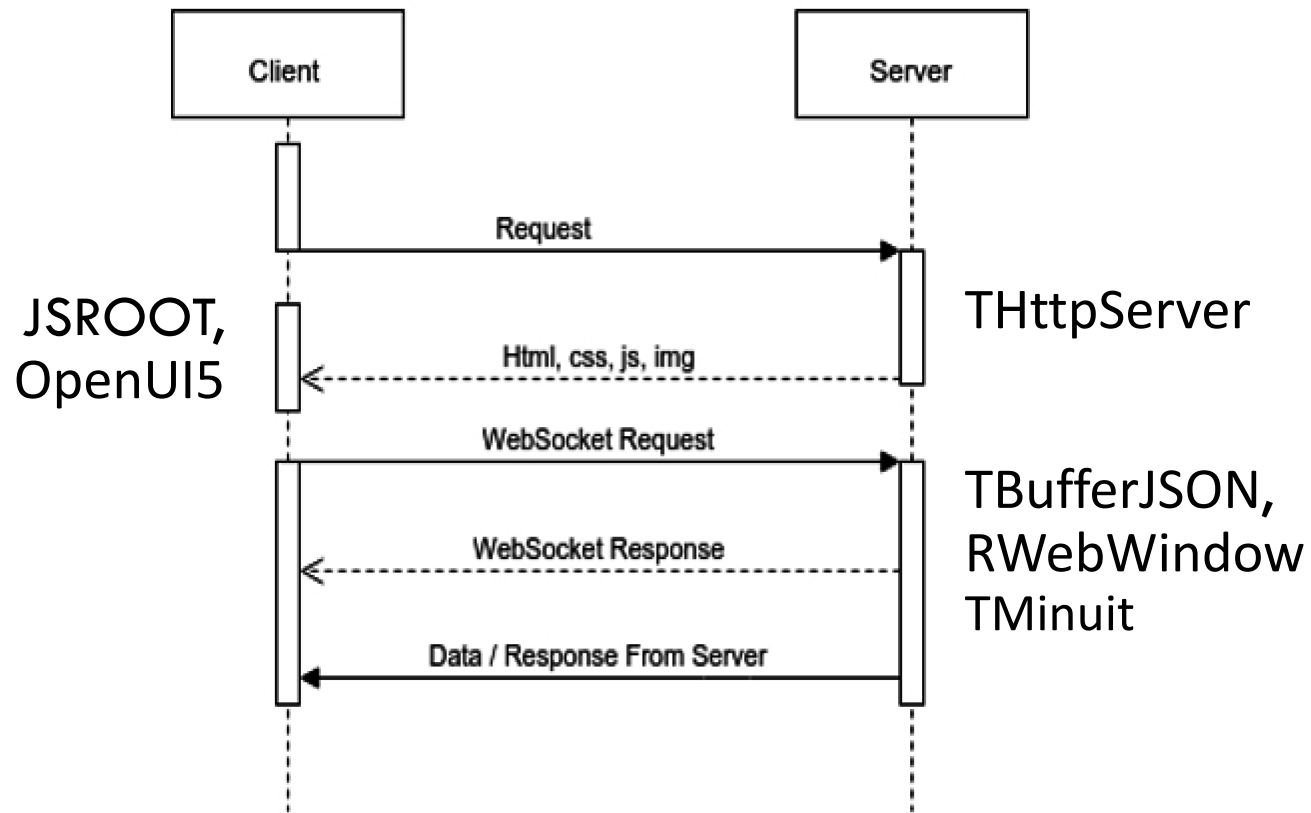
*****
Minimizer is Minuit / Migrad
Chi2      = 609.53
NDf       = 923
Edm       = 7.47518e-08
NCalls    = 557
R         = 204.487 +/- 2.74779
d         = -51.1556 +/- 0.998987
m         = 17.551 +/- 1.652
rho       = 9.71474e-06 +/- 1.48309e-08
n         = 2.75983e-09 +/- 1.09501e-10
lb        = 1.86774e-11 +/- 2.53199e-13

```

Soloviev A.G., Solovjeva T.M., Lukyanov K.V. Deployment of a web application for fitting experimental data in the JINR cloud infrastructure // PEPAN, vol.55, issue 3, pp.627-631, 2024.

ARCHITECTURE

<http://fitter.jinr.ru/>

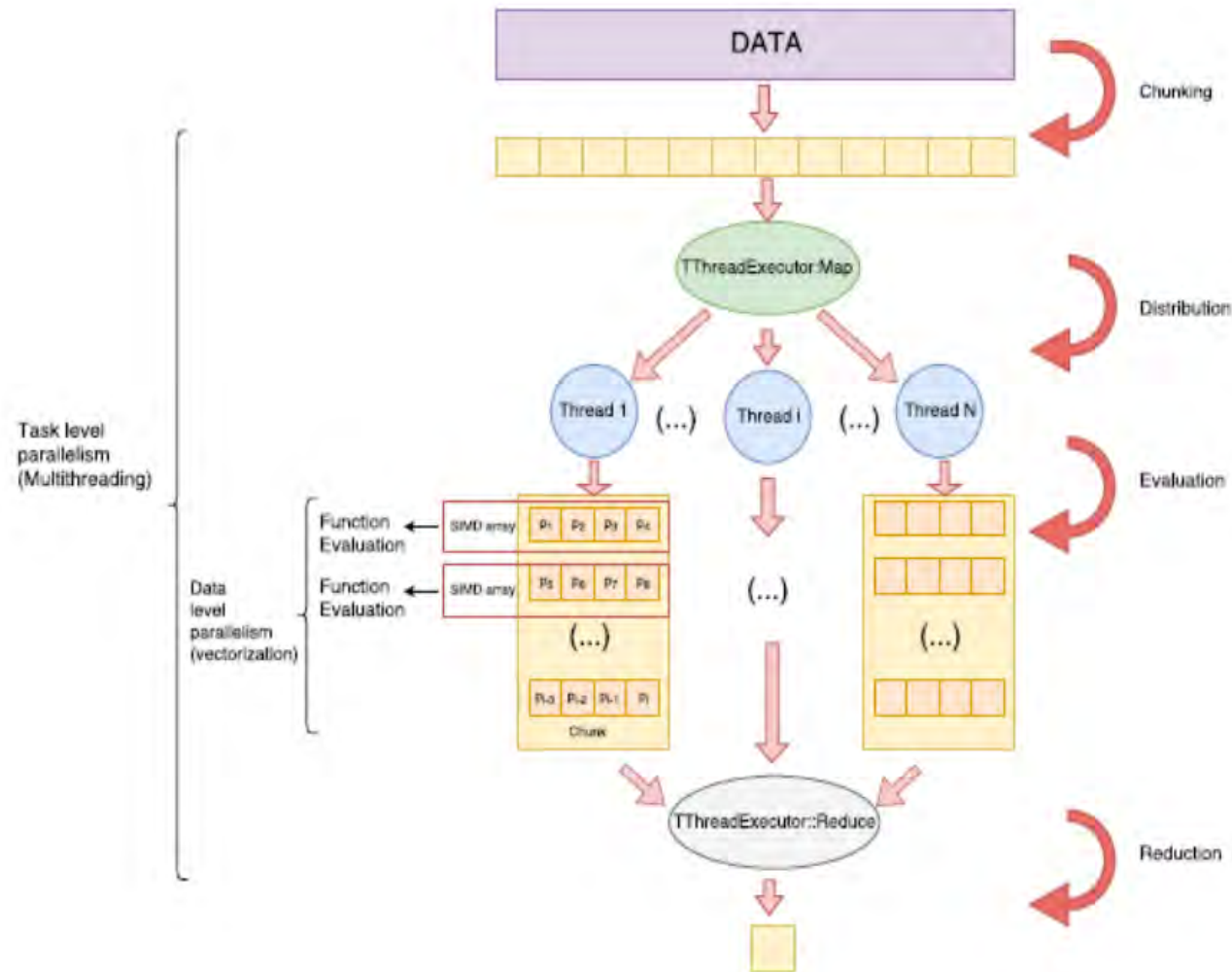


```

↑ 10:10:0:READY=
↓ 1:10:1:SHOWPANEL:localapp.view.Main
↑ 1:10:1:PANEL_READY
↓ 1:10:1:OPTIONS:{"_typename": "ROOT::Experimental::RFitPanelModel", "fTitle": "", "fDataSet": {}
↑ 1:10:0:KEEPALIVE
↑ 0:9:0:KEEPALIVE
↑ 0:8:1:OPTIONS:{"_typename": "ROOT::Experimental::RFitPanelModel", "fTitle": "", "fDataSet": [{"_type
↑ 0:7:1:DATA: 6.4707398e-03 5.7622315e-08 9.7453739e-09 6.5988727e-03 5.8733149e-08 9.9395164e-09
↓ 4:10:1:DATA:{"_typename": "TGraphErrors", "fUniqueID": 0, "fBits": 1032, "fName": "f113_new
↓ 0:9:1:OPTIONS:{"_typename": "ROOT::Experimental::RFitPanelModel", "fTitle": "", "fDataSet": {}
↑ 2:10:0:KEEPALIVE

1  0:7:1:DATA: 6.4707398e-03 5.7622315e-08 9.7453739e-09
2  6.5988727e-03 5.8733149e-08 9.9395164e-09
3  6.7270063e-03 5.8177527e-08 9.8922512e-09
4  6.8551399e-03 5.6085984e-08 9.6427664e-09
5  6.9832735e-03 5.4336522e-08 9.5603091e-09
6  7.1114071e-03 5.1121204e-08 9.2211981e-09
7  7.2395399e-03 4.9306049e-08 8.9781282e-09
8  7.3676735e-03 4.8772709e-08 8.9811363e-09
9  7.4958071e-03 4.7488745e-08 8.8890914e-09
10 7.6239407e-03 4.5697627e-08 8.6491182e-09
11 7.7520743e-03 4.4534891e-08 8.5005185e-09
12 7.8802072e-03 4.3626876e-08 8.3913524e-09
  
```

PARALLELIZATION



TASK-LEVEL PARALLELISM

- data fragmentation
- distribution and evaluation
- reduction

ACTIVATION

ROOT::EnableImplicitMT()

DATA-LEVEL PARALLELISM

- vectorization of the model function during evaluation

VM AT THE JINR CLOUD INFRASTRUCTURE

- **HARDWARE**

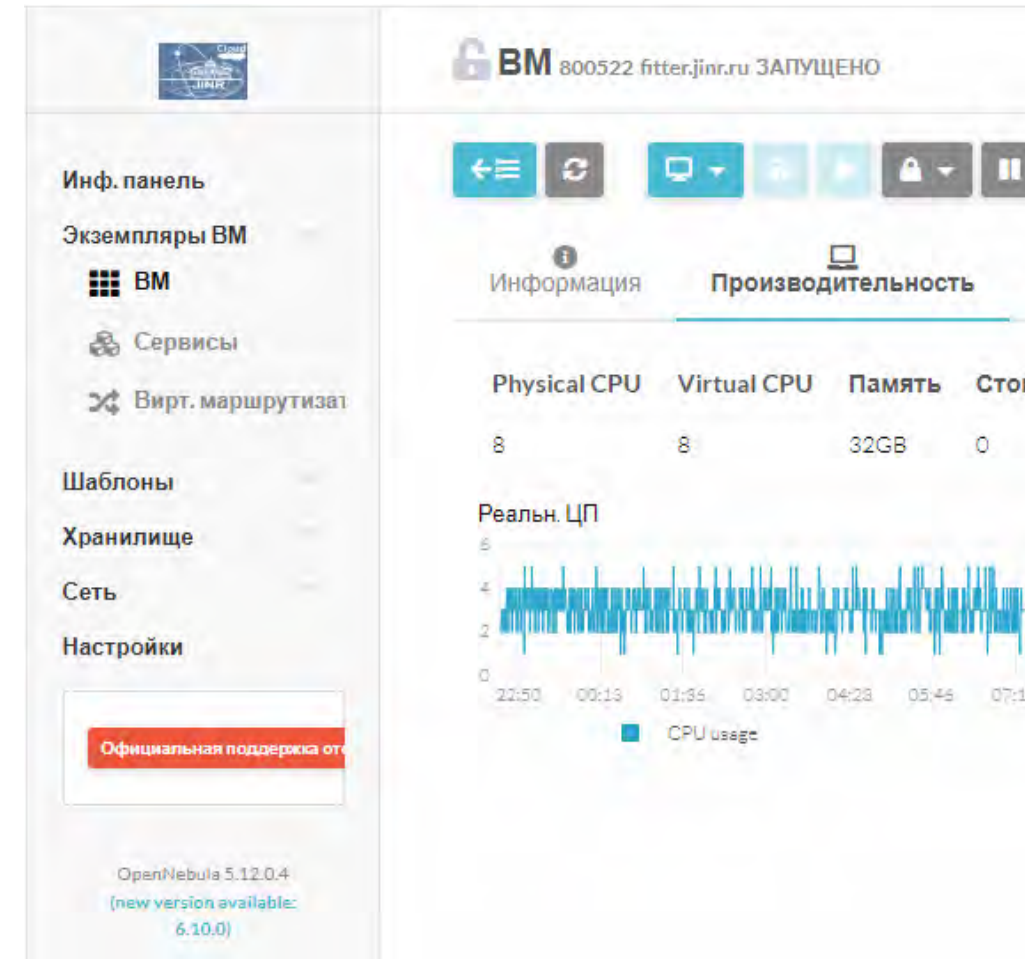
- **8 cores** (CPU Intel Xeon E5-2650 v4 @ 2,20 GHz)
- 32 Gb RAM
- 20 Gb SSD

- **SOFTWARE**

- Ubuntu 22.04.1
- ROOT 6.24.06
- FITTER_WEB from <https://git.jinr.ru/yumo/fitter-next>

- **SECURITY**

- Public key auth
- Standart JINR security tests for web-sites passed



COMPARISON OF FITTING CPU TIME

Models of particles of ferromagnetic liquid sample	fitter.jinr.ru 8 cores local	fitter.jinr.ru 8 cores web	hybrilit.jinr.ru 6 cores local	hybrilit.jinr.ru 6 cores web	computer Intel Core™ i9 10 cores local	computer Intel Core™ i9 10 cores web
BallPolydispersity	0.21 +/- 0.02	0.22 +/- 0.02	0.28+/-0.03	0.29 +/- 0.05	0.07+/-0.01	0.07 +/- 0.01
Ellipsoid	39.3 +/- 0.2	41.5 +/- 0.2	51.4 +/-0.2	54.5 +/- 0.3	19.83 +/- 0.05	19.97 +/- 0.08
Parallelepiped	374 +/- -3	475 +/- 3	490 +/- 1	527 +/- 2	188 +/- 0.5	186 +/- 0.5

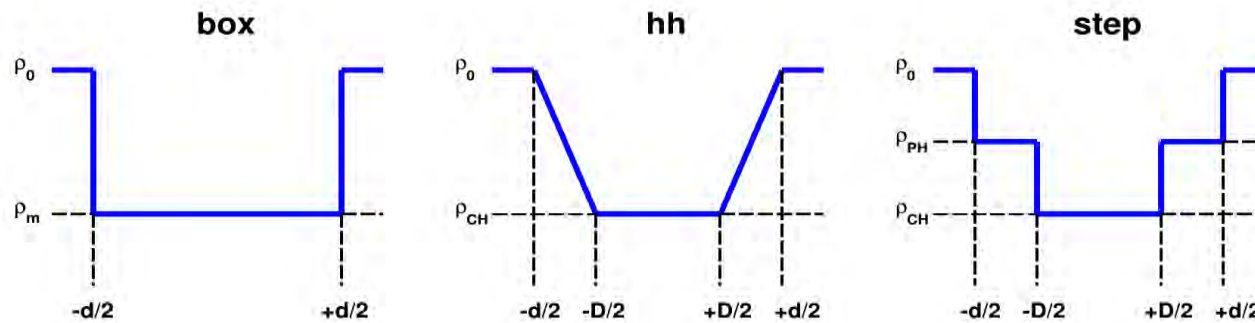
QUESTIONS

1. Is 8 cores enough? How much cores needed to obtain maximum efficiency?
2. How much memory needed in current and “optimal” configuration? *
3. How multiuser calculations affects the fitter application?
4. What is the balance solution between computing time and the amount of allocated by VM resources?

* N. Balashov, I. Kuprikov, N. Kutovskiy, A. Makhalkin, Ye. Mazhitova, I. Pelevanyuk, R. Semenov, and D. Shpotya. Changes and Challenges at the JINR and Its Member States Cloud Infrastructures. Physics of Particles and Nuclei, 2024, Vol. 55, No. 3, pp. 366–370

METHODOLOGICAL TASK

THE STRUCTURE OF VESICULAR SYSTEMS



Some of parameters:

R – radius of the vesicle

d – bilayer thickness

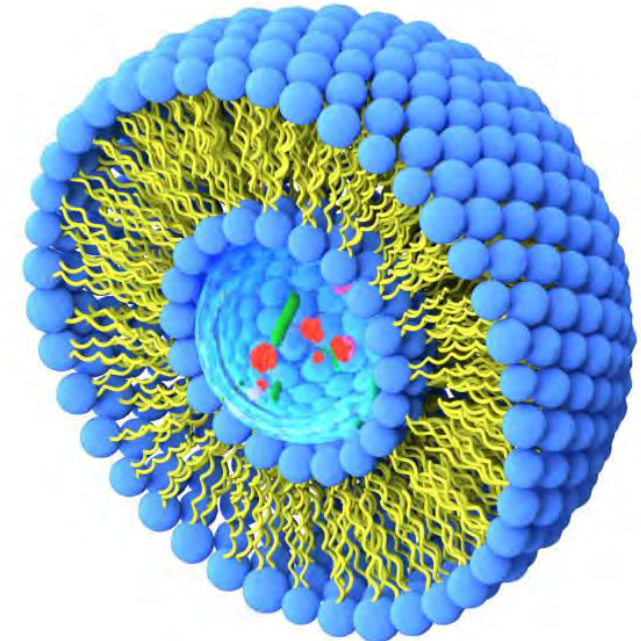
m – polydispersity coefficient

n – number of vesicles

ρ_m – membrane density

ρ_{PH} – density in the area of polar heads

D – hydrophobic thickness

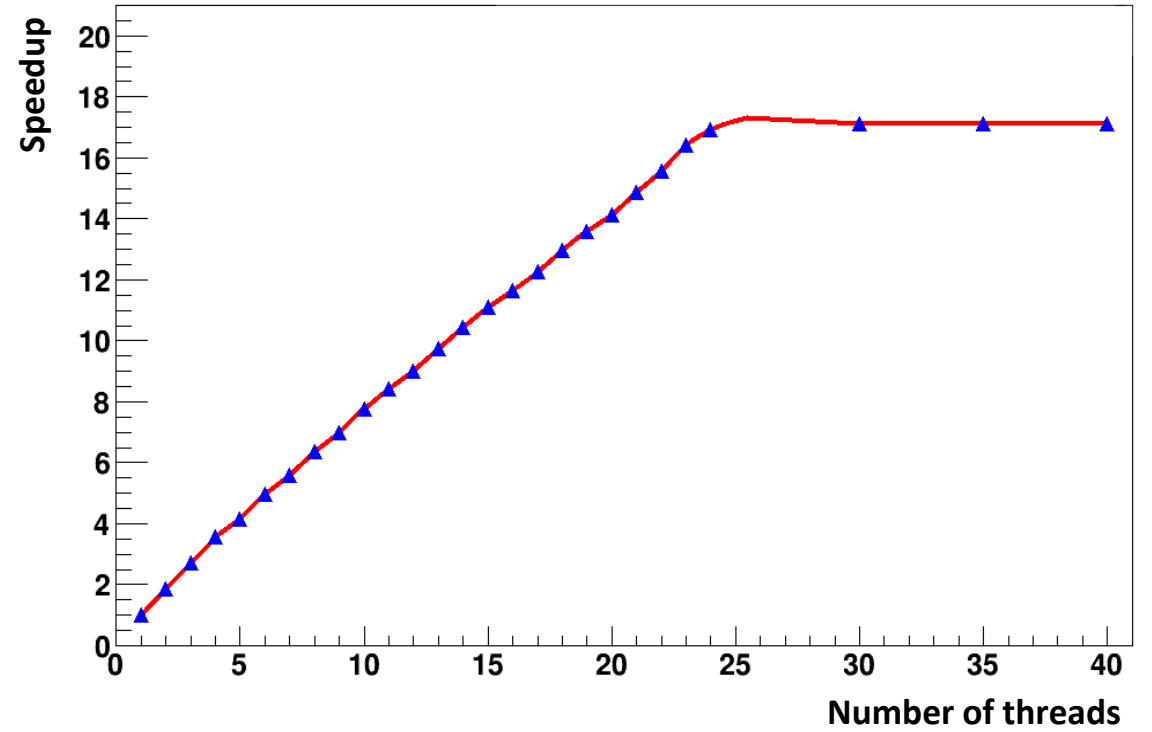
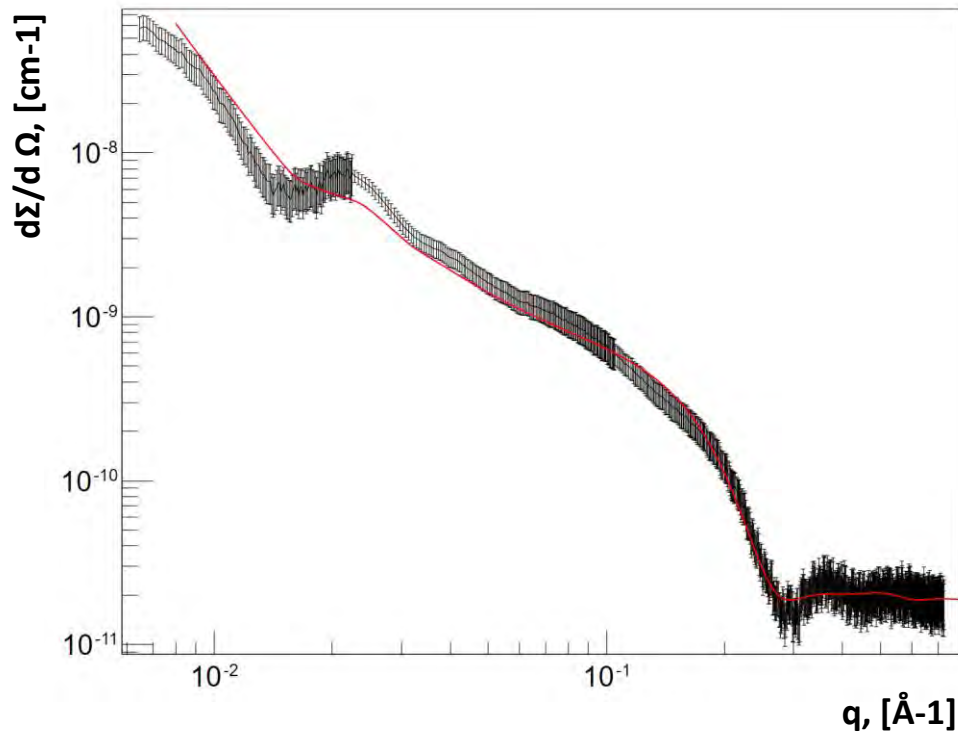


1000 experimental points, ~50 minutes CP in single-threaded mode.

QUESTION 1: NUMBER OF CORES

HARDWARE: 24 cores, 120 Gb RAM.

SOFTWARE – the same



M. Bashashin, E. Zemlyanaya, M. Kiselev, K. Lukyanov, and K. Turapbay. SFF Analysis of Small Angle Scattering Data from Phospholipid Vesicle Systems: Parallel Implementation and Online Interface. *Physics of Particles and Nuclei Letters*, 2022, Vol. 19, No. 5, pp. 554–557.

QUESTION 2: MEMORY

	N=230	N=460	N=930
VIRT, MB	2108	2109	2114
RES, MB	1127	1129	1136
CP₁, SEC	459	1494	2980

Each process takes the same amount of memory – about 1GB RAM.

Max number of processes for single fitting procedure = number of cores (24)

Max RAM for single fitting procedure: 24cores * 1GB ~ 24GB

```

ubuntu@ip-10-220-18-114: ~
0[ 0.0%] 3[| 0.7%] 6[|||||100.0%] 9[ 0.0%] 12[ 0.0%] 15[|||||100.0%] 18[ 0.0%] 21[|||||100.0%]
1[|||||100.0%] 4[|||||100.0%] 7[ 0.0%] 10[|||||100.0%] 13[| 0.7%] 16[|||||100.0%] 19[| 0.7%] 22[ 0.0%]
2[|||||100.0%] 5[ 0.0%] 8[ 0.0%] 11[|||||100.0%] 14[ 0.0%] 17[ 0.0%] 20[ 0.0%] 23[ 0.0%]
Mem[|||||] 2.38G/62.8G Tasks: 61, 80 thr; 10 running
Swp[ ] OK/OK Load average: 8.91 5.47 2.42
Uptime: 104 days (!), 22:41:14

  PID USER   PRI  NI  VIRT   RES   SHR  S  CPU% MEM%   TIME+  Command
 317081 root    20   0 2087M 1101M 85368 R 100.  1.7  5:38.46 ./fitter-next
 317093 root    20   0 2100M 1113M 86136 R 100.  1.7  3:26.70 ./fitter-next
 317096 root    20   0 2102M 1115M 85404 R 100.  1.7  3:14.19 ./fitter-next
 317099 root    20   0 2103M 1117M 85468 R 100.  1.7  2:41.48 ./fitter-next
 317075 root    20   0 2080M 1093M 85384 R 99.9  1.7  8:04.80 ./fitter-next
 317078 root    20   0 2083M 1097M 85376 R 99.9  1.7  6:36.69 ./fitter-next
 317084 root    20   0 2092M 1104M 85420 R 99.9  1.7  5:03.72 ./fitter-next

```


QUESTION 3: MULTIUSER CALCULATIONS

- THttpServer limitation: 10 simultaneous web-threads by default.
- Each web-thread can run the fitting procedure on all cores
 => 240 processes
 => 240GB RAM
 => Unreal, need limitations

```

ubuntu@ip-10-220-18-114: ~
0 [|||||91.4%] 3 [|||||86.2%] 6 [|||||92.7%] 9 [|||||98.0%]
1 [|||||92.7%] 4 [|||||95.4%] 7 [|||||100.0%] 10 [|||||99.3%]
2 [|||||100.0%] 5 [|||||95.4%] 8 [|||||100.0%] 11 [|||||100.0%]
Mem [||||| 1.89G/62.8G]
Swp [ OK/OK]

  PID USER      PRI  NI  VIRT   RES   SHR  S  CPU% MEM%   TIME+  Command
 317430 root        20   0 2141M 1165M 86096 R  620.  1.8   2:32.22  ./fitter-ne
317423 root        20   0 2149M 1165M 86096 R  586.  1.8   3:08.69  ./fitter-ne
317427 root        20   0 2145M 1165M 86096 R  583.  1.8   2:52.77  ./fitter-ne
317440 root        20   0 2141M 1165M 86096 R  484.  1.8   2:14.13  ./fitter-ne
317456 root        20   0 2141M 1165M 86096 R  84.1  1.8   0:17.79  ./fitter-ne
317453 root        20   0 2141M 1165M 86096 R  82.1  1.8   0:17.49  ./fitter-ne
317448 root        20   0 2145M 1165M 86096 R  81.4  1.8   0:20.07  ./fitter-ne
317436 root        20   0 2149M 1165M 86096 R  78.1  1.8   0:21.94  ./fitter-ne
317445 root        20   0 2145M 1165M 86096 R  77.4  1.8   0:20.18  ./fitter-ne
317452 root        20   0 2141M 1165M 86096 R  76.1  1.8   0:17.57  ./fitter-ne
317433 root        20   0 2149M 1165M 86096 R  75.5  1.8   0:22.06  ./fitter-ne
317437 root        20   0 2149M 1165M 86096 R  74.1  1.8   0:21.99  ./fitter-ne
317455 root        20   0 2141M 1165M 86096 R  72.1  1.8   0:17.04  ./fitter-ne
317438 root        20   0 2149M 1165M 86096 R  71.5  1.8   0:22.27  ./fitter-ne
317444 root        20   0 2145M 1165M 86096 R  71.5  1.8   0:20.01  ./fitter-ne
317454 root        20   0 2141M 1165M 86096 R  71.5  1.8   0:17.61  ./fitter-ne
317451 root        20   0 2141M 1165M 86096 R  70.8  1.8   0:17.43  ./fitter-ne
317447 root        20   0 2145M 1165M 86096 R  66.9  1.8   0:19.99  ./fitter-ne
317450 root        20   0 2141M 1165M 86096 R  66.9  1.8   0:17.27  ./fitter-ne
317439 root        20   0 2149M 1165M 86096 R  65.5  1.8   0:22.14  ./fitter-ne

```

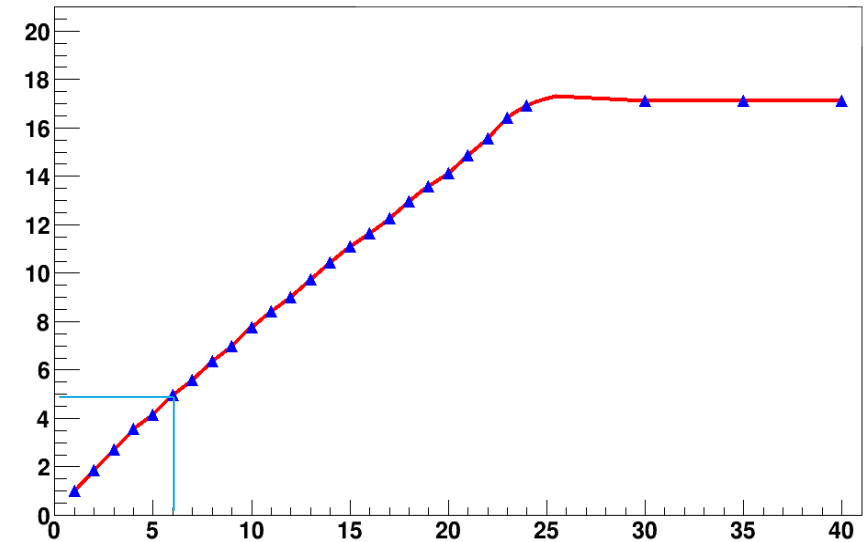
DISCUSSION

LIMITING ASSUMPTIONS:

1. Maximum 2 simultaneous fittings (current usage stats)
2. Calculation time limit of 10 minutes for 1 user

RESOURCE LIMITATIONS (PROJECT):

1. 1-core fit takes about 3000 sec CP (50min)
2. 10 minutes requires 5x speedup => 6 cores; **12 cores** for 2 users
3. 2 parallel fits = 24 server processes => **24 GB RAM (+4GB for operating system)**
4. Limit THttpServer number of **web-threads to 2 (+ 1 reserve)**



```

R      = 204.487 +/- 2.74779
d      = -51.1556 +/- 0.998981
m      = 17.551 +/- 1.65201
rho    = 9.71474e-06 +/- 1.48309e-08
n      = 2.75983e-09 +/- 1.09501e-10
lb     = 1.86774e-11 +/- 2.532e-13
D      = 19.9368 +/- 0.308499

```

Real time 0:04:28, CP time 3021.250

Real time 0:08:57, CP time 3119.410

CONCLUSION

Both methodological calculations and approbation for fitting data in real physical tasks were made using the presented web application. Some links were listed in the footnotes of the slides.

Based on the results of methodological calculations, the following can be stated:

1. A web application implements parallelization quite well, which gives it a certain "margin of safety", allowing it to use processors with a larger number of cores.
2. The minimum amount of memory per 1 process required under different operating conditions of the application is determined.
3. Estimation of the required number of resources (CPU cores and RAM memory) based on the desired time of calculation and number of users is demonstrated.