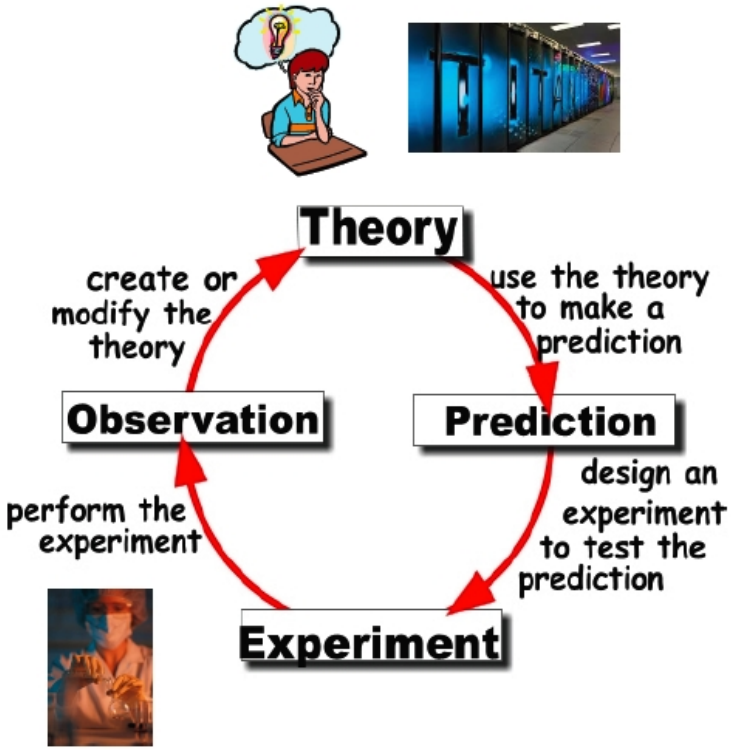
Bootstrap: a nuclear physics point of view

A. Pastore

Department of Physics, University of York, Heslington, York, YO10 5DD, UK

ISNET-5 York

Introduction



The scientific method

- A problem is identified
- Relevant data are gathered
- An hypothesis is formulated from the data
- The hypothesis is empirically tested

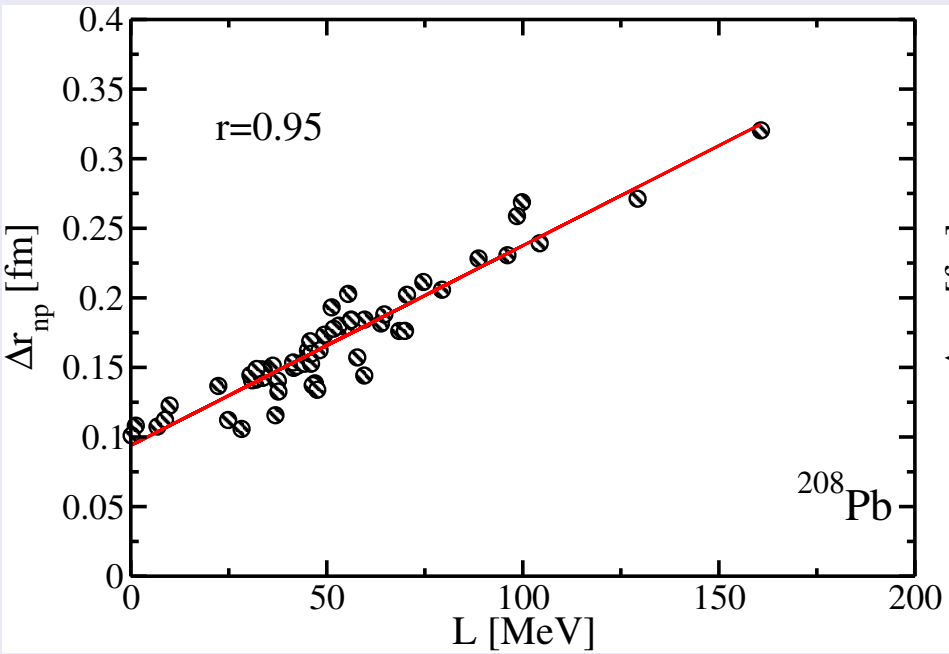
Journal of Physics G

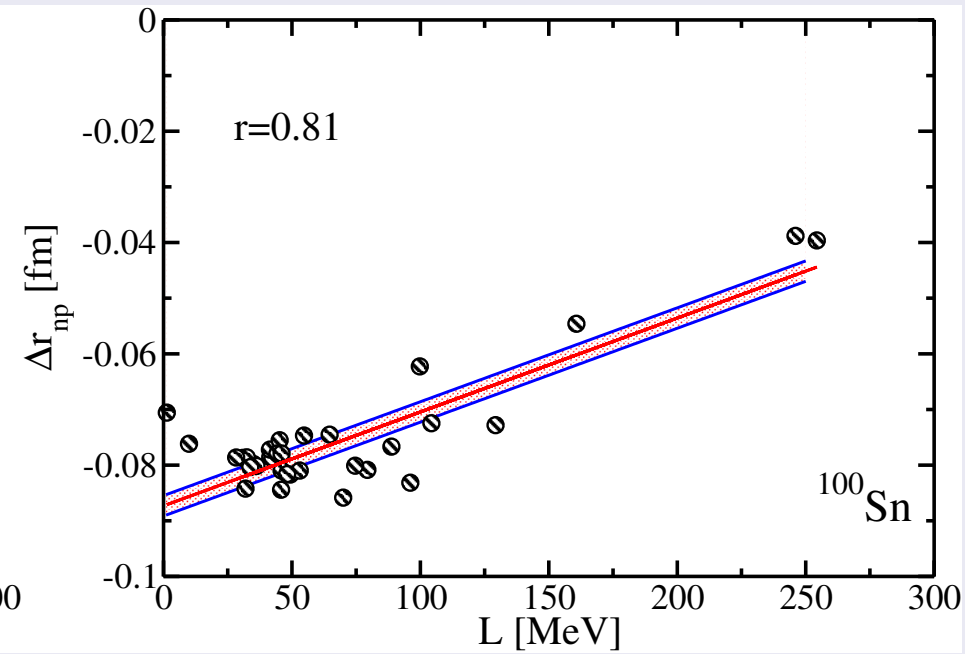
Enhancing the interaction between nuclear experiment and theory through information and statistics (ISNET) (2015)

What is Bootstrap?

- Non-Parametric Bootstrap is a statistical method introduced by Efron in 1979 B. Efron and R. Tibshirani, An introduction to the bootstrap (CRC press, 1994).
- Based on resampling (with replacement) of data
- No specific assumptions on the *parent* distribution
- Possible extension: parametric Bootsrap

Analysis of correlations





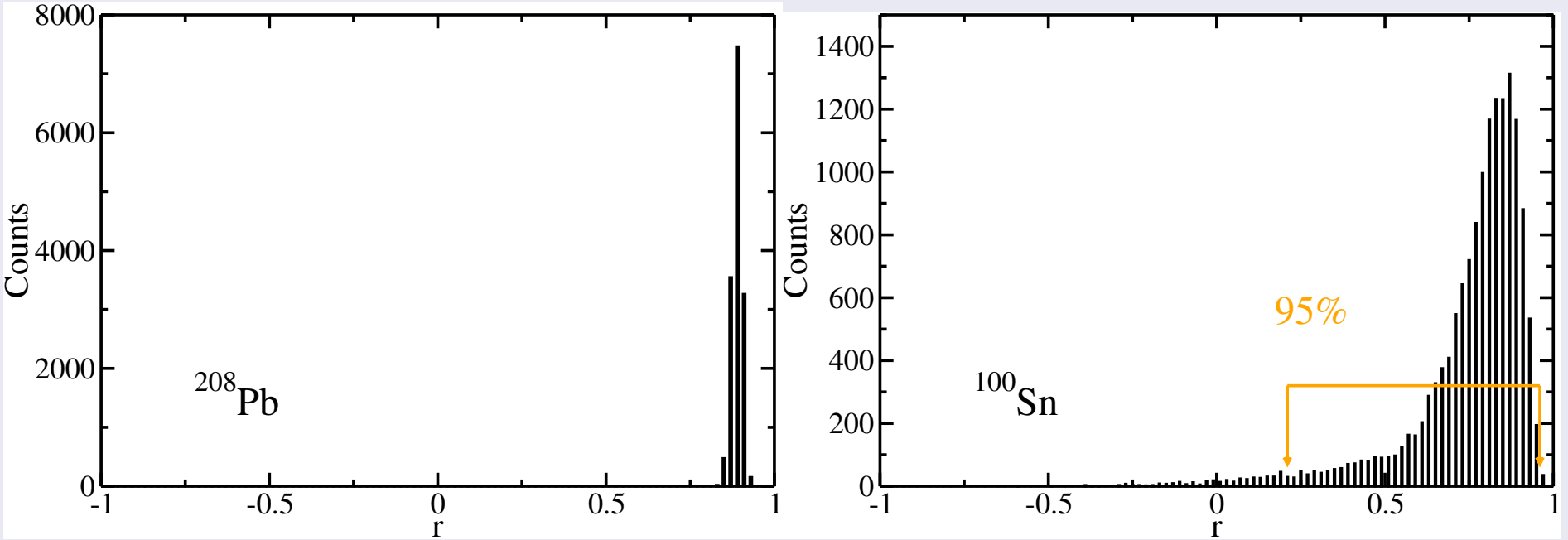
Bootstrap @ work!

- We build BB new data sets with n points (as original) by resampling

$$\left(\begin{array}{c} 2n-1 \\ n \end{array} \right) = \frac{(2n-1)!}{n!(n-1)!}$$

- We calculate for each data-set $\hat{r}$
- We build the distribution

Errors on $\hat{r}$: confidence interval 68% quantile.



Bootsrap fit

- Possible alternative ot χ^2
- Bootsrap of residuals $y = f(\vec{a}, x) + \varepsilon$ (4000 samples)
- Creation of *new* data sets $y^* = y + \varepsilon^*$
- Minimisation of a distance $(y^* - f(\vec{a}, x))^2$ or $|y^* - f(\vec{a}, x)|$

Example: Liquid-Drop model

We want to fit a linear model (as a start point!)

$$f(N, Z; av, as, ac, aa) = av - as(N+Z)^{-1/3} - ac\frac{Z^2 - Z}{(N+Z)^{4/3}} - aa\frac{(N-Z)^2}{(N+Z)^2}$$

We use 2236 *experimental* points. No error on the data (as a start)

Direct χ^2 fit

MATHEMATICA results

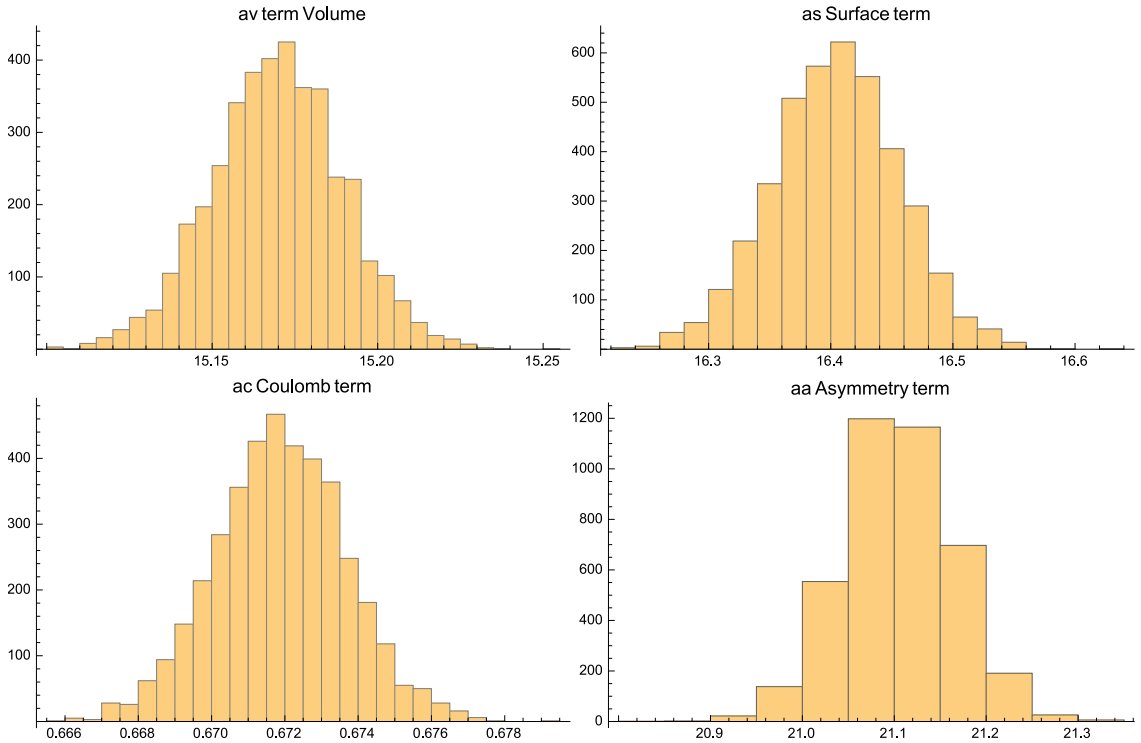
	Estimate	Standard Error	t- Statistic	P- Value	
av	15.1697	0.0197484	768.15	1.35875104398	$\times 10^{-2806}$
as	16.4042	0.0525803	311.985	4.85895485094	$\times 10^{-1904}$
ac	0.671887	0.00180058	373.151	8.83790086779	$\times 10^{-2082}$
aa	21.1048	0.0608933	346.586	2.61381401095	$\times 10^{-2008}$

We also obtain the covariance matrix

$$\begin{pmatrix} av & as & ac & aa \\ 1 & 0.981 & 0.970 & 0.654 \\ & 1 & 0.921 & 0.590 \\ & & 1 & 0.588 \\ & & & 1 \end{pmatrix}$$

We use these results to benchmark the Bootstrap. We fix the residual ε for the *best* fit.

Bootsrap to assess errors (Gaussian here)

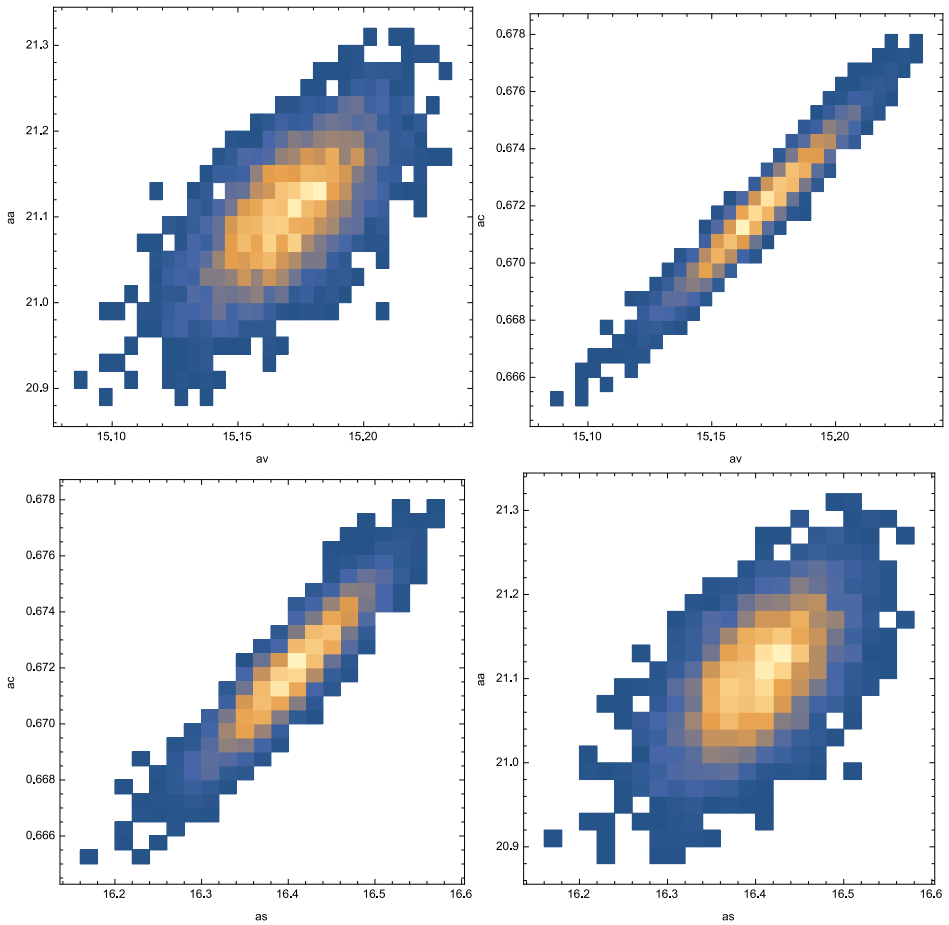


Comparing values

The mean values are $av = 15.1699;\ as = 16.404;\ ac = 0.672;\ aa = 21.104.$

The differences with MATHEMATICA: $\delta \approx 0.0003/4$

Bootsrap to assess correlations



Bootsrap distribution

We have access to the distribution of the parameters around the minimum!

Conclusions

Bootsrap is a very simple tool

- Simple to implement
- We have used at moment only on error analysis
- Need more CPU than χ^2, but more informations

Possible ideas

- Bootsrap to go beyond parabolic approximations
- Tool to avoid derivative perform numerical derivatives of χ^2
- Possible to assess non-Gaussian errors

THANK YOU!!!