Reverend Bayes weighs in on nuclear physics

G.F. Bertsch

QCD thermodynamics, Aug. 24-26
Schloss Waldthausen

1. What is Bayes?
2. Characterizing hot matter
    a) EOS
    b) chemical freeze-out
3. Cold nuclear matter (with Derek Bingham)

What is Bayesian statistics?

$$P_{\text{new}}(\text{parameters}) = P(\text{parameters}|\text{observables})$$
$$= \mathcal{N} P(\text{observables}|\text{parameters}) P_{\text{old}}(\text{parameters})$$

One often assumes that $P_{\text{old}} = \text{constant}$

Evaluating the conditional probability $P(\text{observables}|\text{parameters})$

a) Likelihood function for the observables

$$\exp\left(-\sum_i^N (x_i^{exp} - x_i^p)^2 / 2\sigma_i^2\right)$$

b) parameters --> observables

i) Markov Chain Monte Carlo (MCMC)

ii) Gaussian process emulator

Equation of State of hot matter

Scott Pratt, et al., PRL 114 202301 (2015)
PRC 89 034917 (2014)

Observations: particle spectrum, elliptic flow, HBT correlations
Au+Au at 100 GeV/n; Pb+PB at 1.38 TeV/n

Model:

Physics	variable parameters	fixed parameters
Initial state $T_{ij}(\vec{r}, \tau_0)$	5+5	$\tau_0 = 0.8$ fm
hydrodynamic expansion	2 EOS 2 viscosity	
chemical freeze-out	0	$T_c = 165$ MeV
hadronic expansion	0	PDG + resonance scattering + σ_0

The first 1000 runs were chosen semirandomly throughout the 14-dimensional parameter space according to latin hypercube sampling. The thirty observables were then reduced to 14 principal components, which captured over 99.9% of the variance. Identically to what was done in Ref. [14], these principal components were interpolated from the 1000 runs using a Gaussian process emulator during a Markov chain Monte Carlo (MCMC) exploration of the parameter space. This yielded a posterior sampling of the parameter space, i.e., a sampling that was weighted by the likelihood to reproduce the measured observables.

Resulting EOS

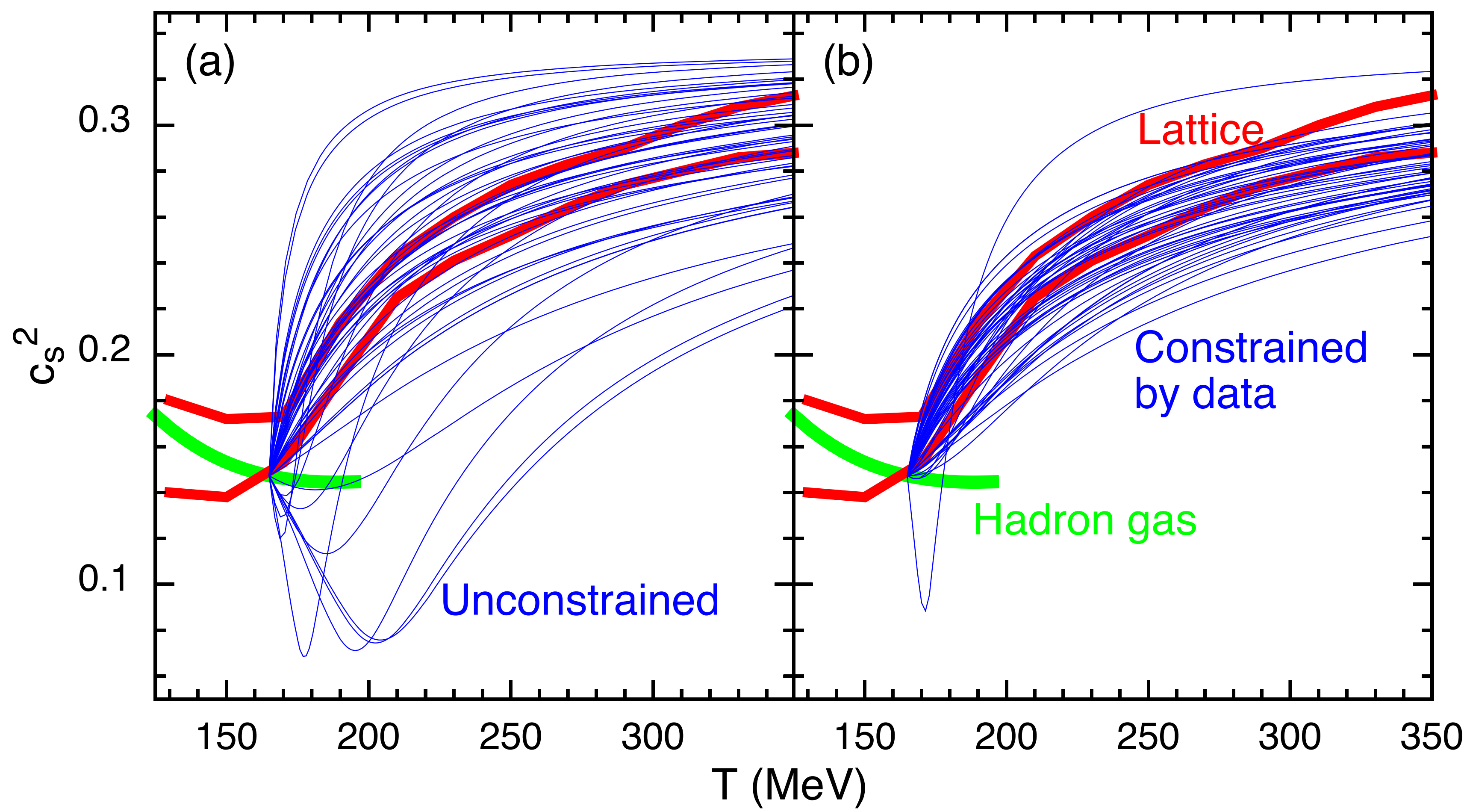


FIG. 5 (color online). (a) Fifty equations of state were generated by randomly choosing X' and R in Eq. (2) from the prior distribution and weighted by the posterior likelihood (b). The two

S. Pratt, et al., PRL 114 202301 (2015).

Chemical freeze-out J.E. Bernhard,...,U. Heinz, arXiv:1605.0394

Observations: flow, $\left.\frac{dN}{dy}\right|_{\pi^{\pm},K^{\pm},p}$ in Pb+PB at 1.38 TeV/n

Model: 9 variable parameters including T_c

$$T_c = 148 \pm 2 \ \text{MeV}$$

How should theorists deal with uncertainties from imperfections in their models?

Editorial, Phys. Rev. A 040001 (2011):

"It is all too often the case that the numerical results are presented without uncertainty estimates. Authors sometimes say that it is difficult to arrive at error estimates. Should this be considered an adequate reason for omitting them? [...] There is a broad class of papers where estimates of theoretical uncertainties can and should be made. Papers presenting the results of theoretical calculations are expected to include uncertainty estimates for the calculations whenever practicable."

Consider a model $x_i = M(p,v_i)$ v_i: input variables
p: parameter set
x_i: model output

Let's assume that the model is an approximation to a theory that is exact. Then we can write

$$x_i^{exp} = M(p_t, v_i) + U(v_i)$$

p_t: true parameter set

$U(v_i)$ is the unknown correction that makes the theory exact.

We can define an an ensemble of likely U's based on the performance of the model by itself.

Step 1. Optimize the model to get a best parameter set p_0.
Step 2. Determine the statistical properties of the residuals

$$r_i = x_i^{exp} - x_x^{p_0}$$

Step 3. An ensemble of likely U's is defined to satisfy

$$\langle U_m^2 \rangle = \langle r_i^2 \rangle$$

If the $U_m(v_i)$ are distributed as independent and Gaussian, we arrive at a chi-squared prescription for the parameter uncertainties.

Usual chi-squared

$$P(p) = \exp\left(-\sum_i^N (x_i^{exp} - x_i^p)^2 / 2\sigma_i^2 \right)$$

Likelihood function

$$\sigma_i^2 = (\sigma_i^{exp})^2 + (\sigma_i^{num})^2 + (\sigma_i^{the})^2$$

$(\sigma_i^{the})^2$ from residuals of best fit

Parameter uncertainty is found at $\chi^2(p_0 + \Delta p) = \chi^2(p_0) + 1$

J. Phys. G **41** 074001 (2014).

Phys.Rev. Lett. **114** 122501 (2015)

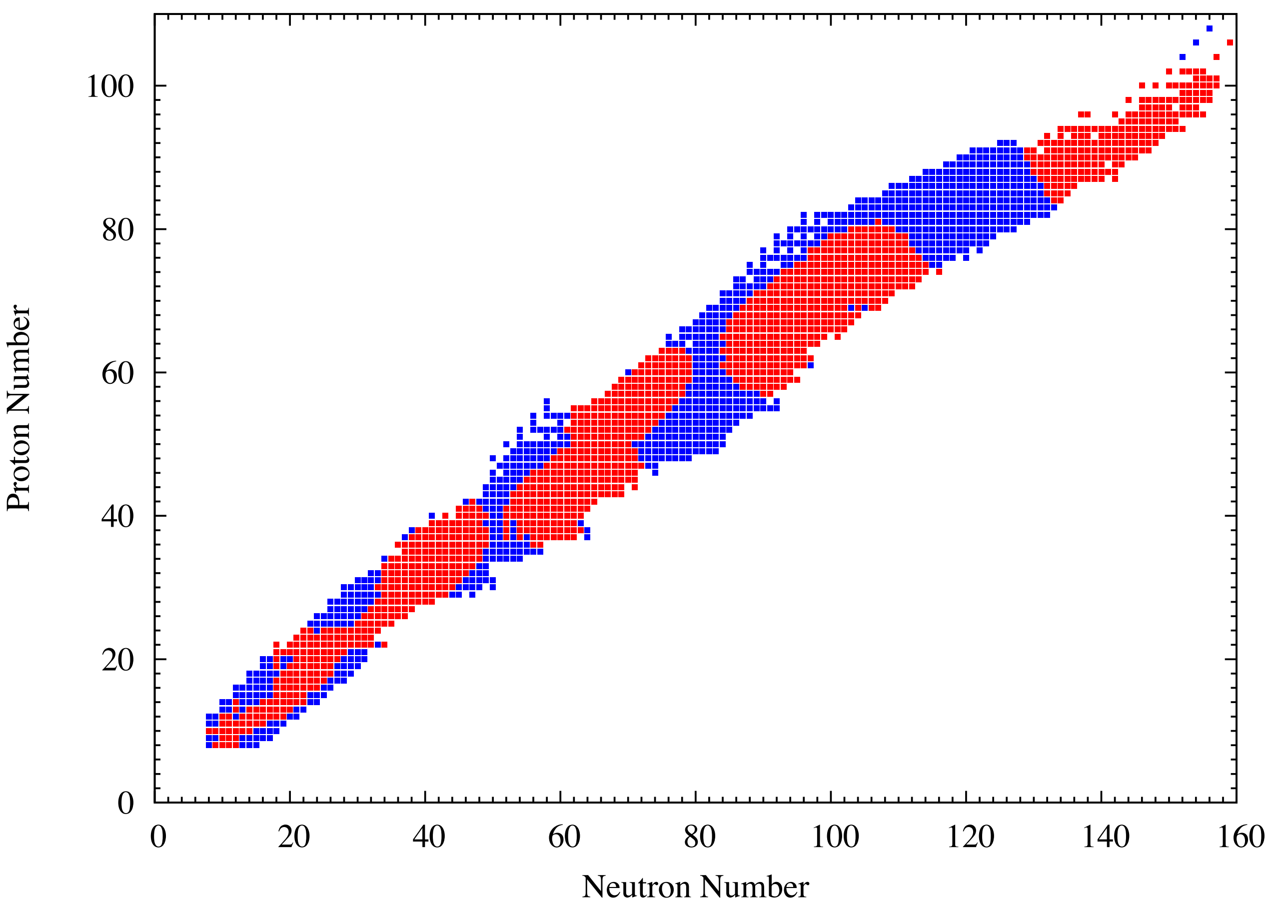
The chi-squared error estimation fails completely for models of nuclear binding energies.

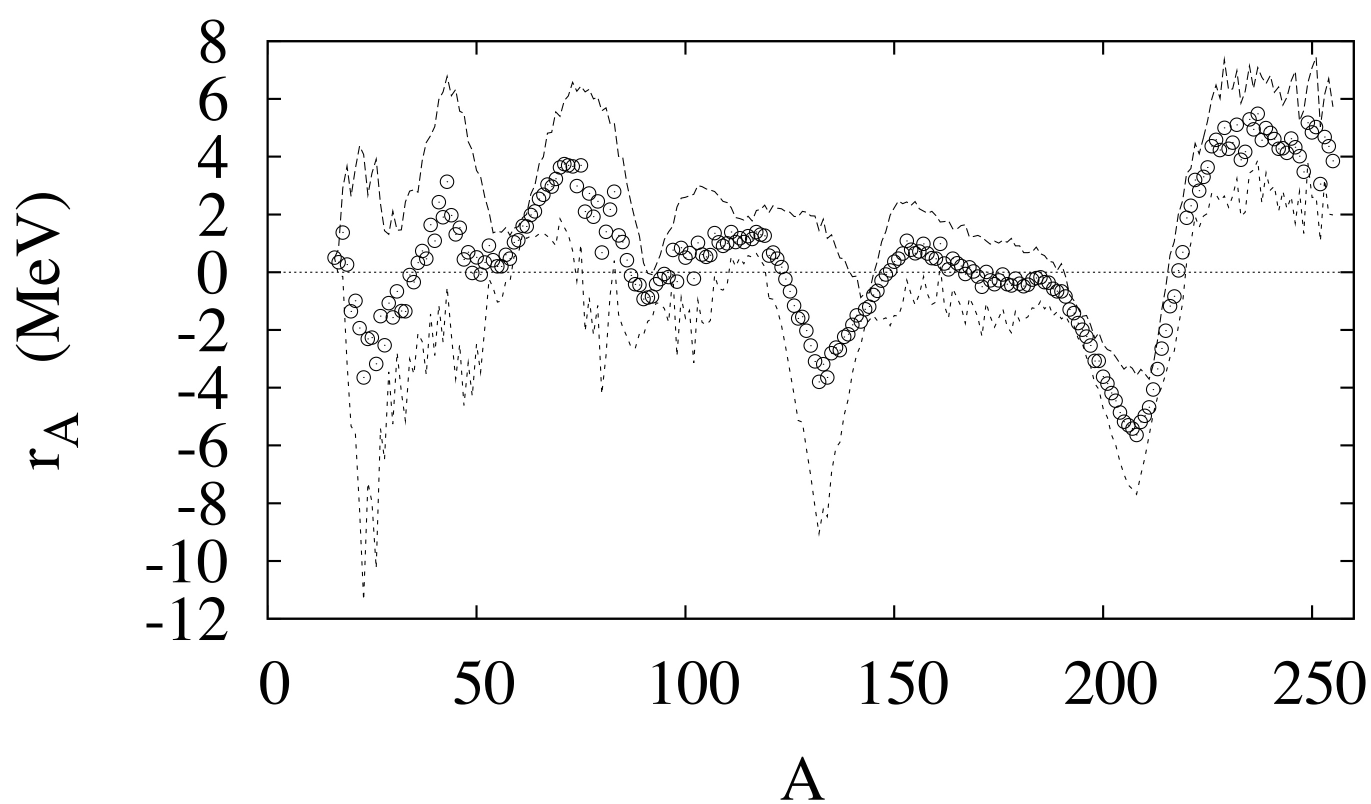
The liquid drop model as an example:

$$B(Z,A) = a_v A - a_s A^{2/3} - a_c \frac{Z^2}{A^{1/3}} - a_a \frac{(N-Z)^2}{A} - \delta \frac{\mathrm{mod}(Z,2) + \mathrm{mod}(N,2) - 1}{A^{1/2}}$$

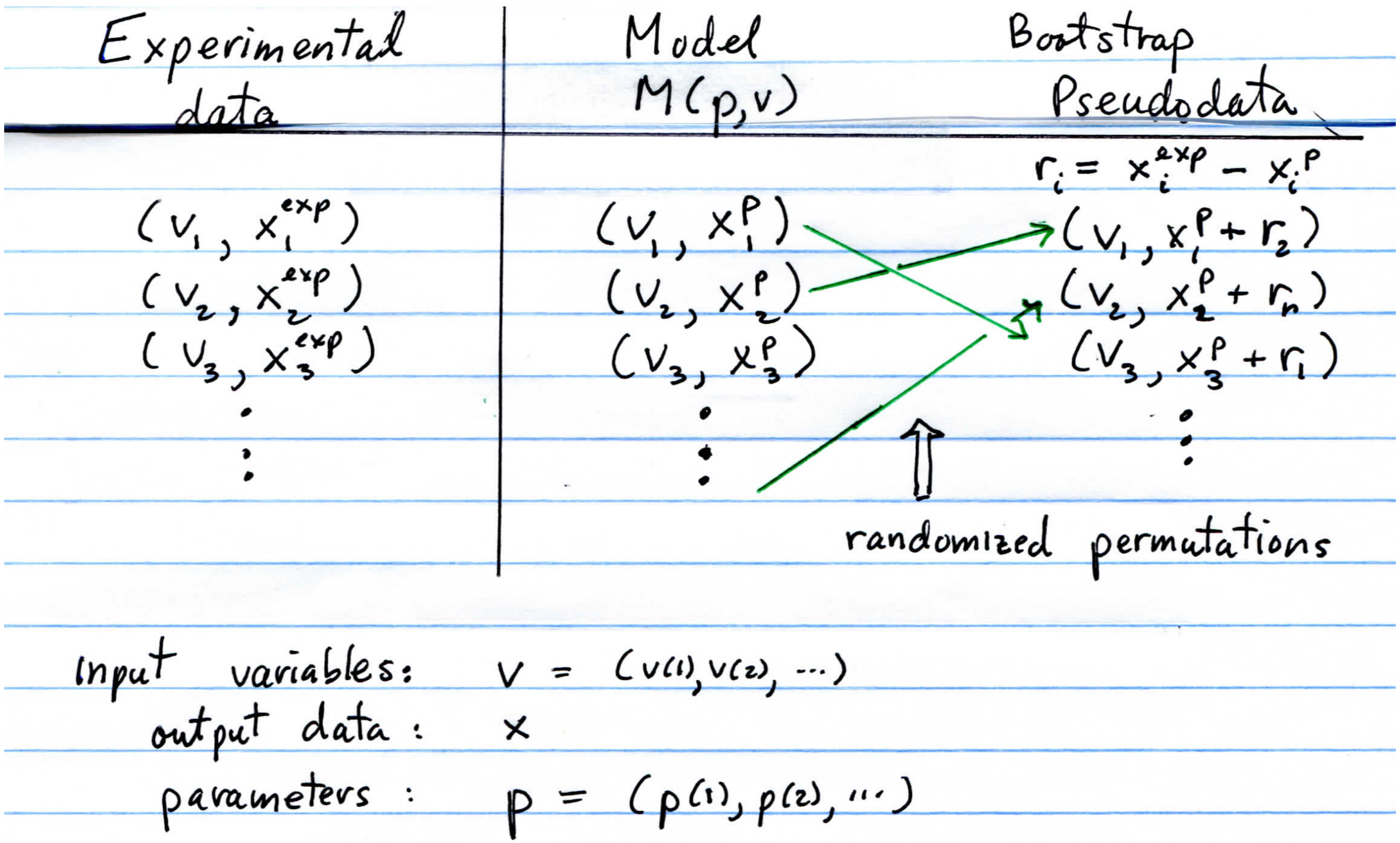
Least-squares + chi-square gives $a_v = 15.59 \pm 0.03 \ \ \mathrm{MeV}$

The problem is that the r_i's are correlated.





The bootstrap method to simulate the ensembles



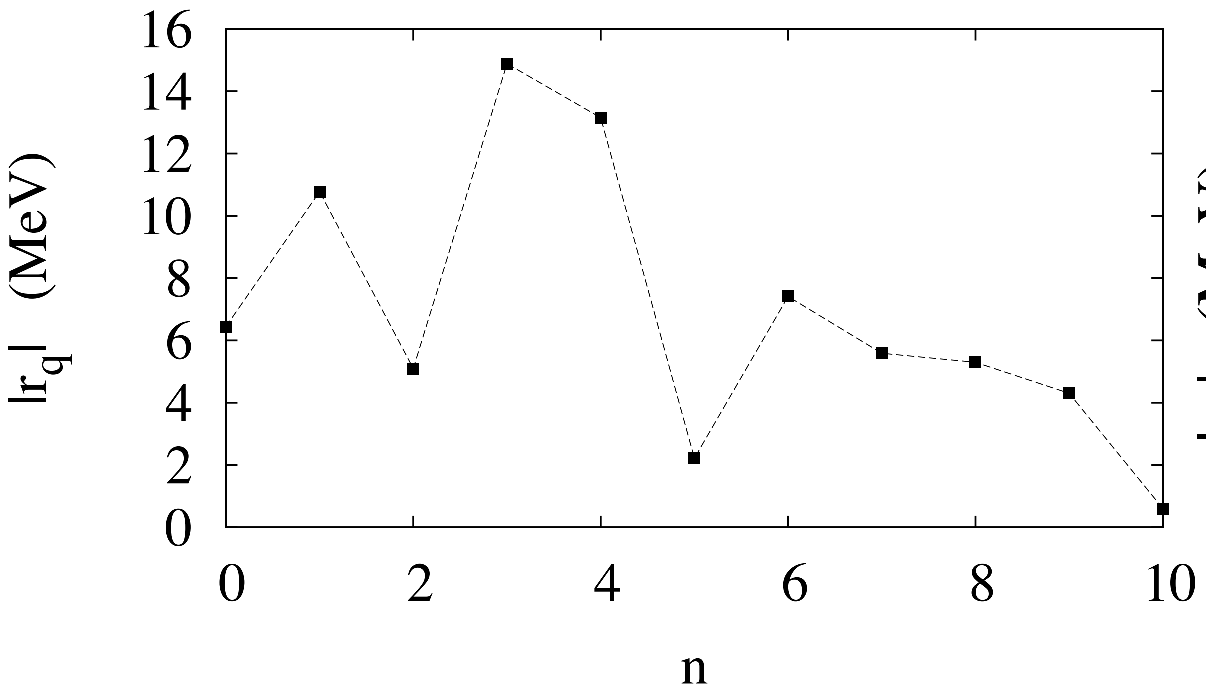
A bootstrap with correlations

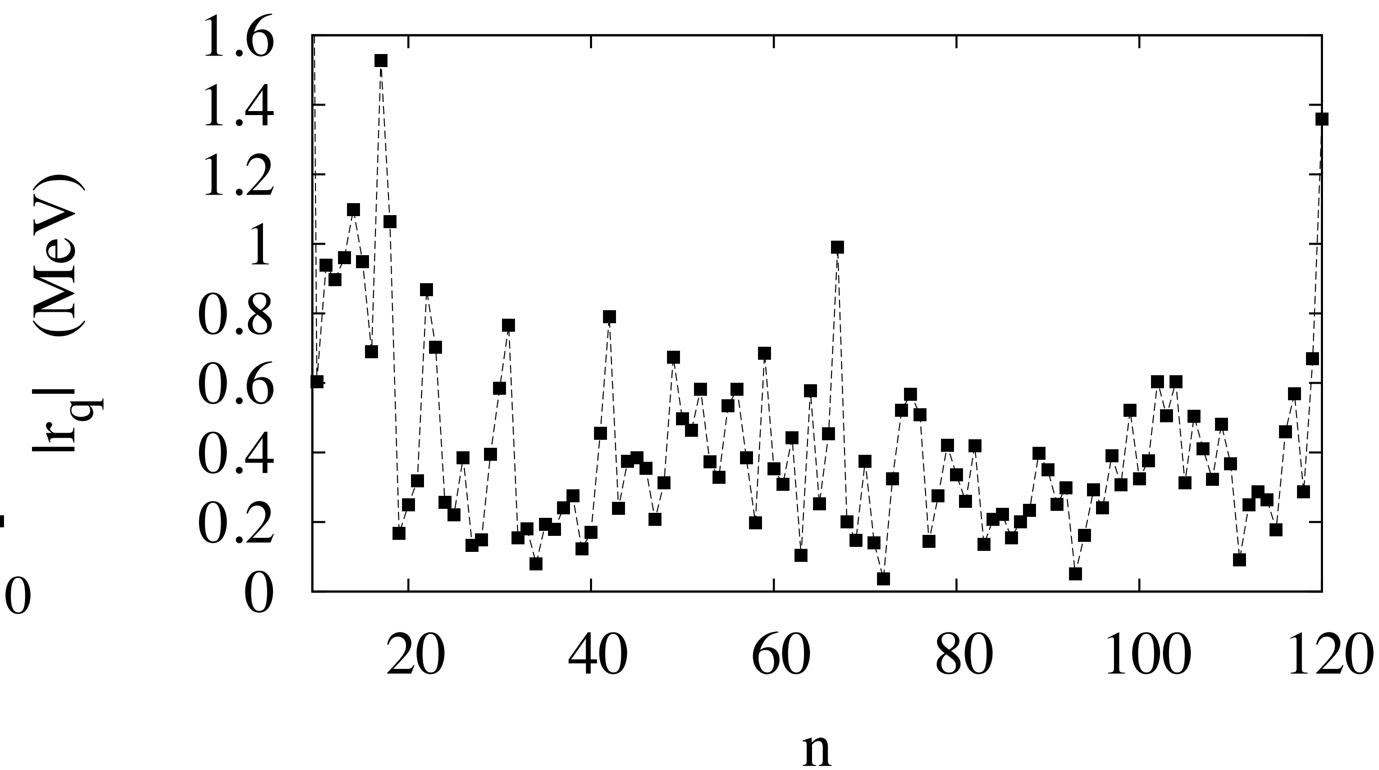
Build a likelihood function for U by the frequency-domain bootstrap.

J-P. Kreiss and S.N. Lahiri, Handbook of Statistics **30** 3 (2012)

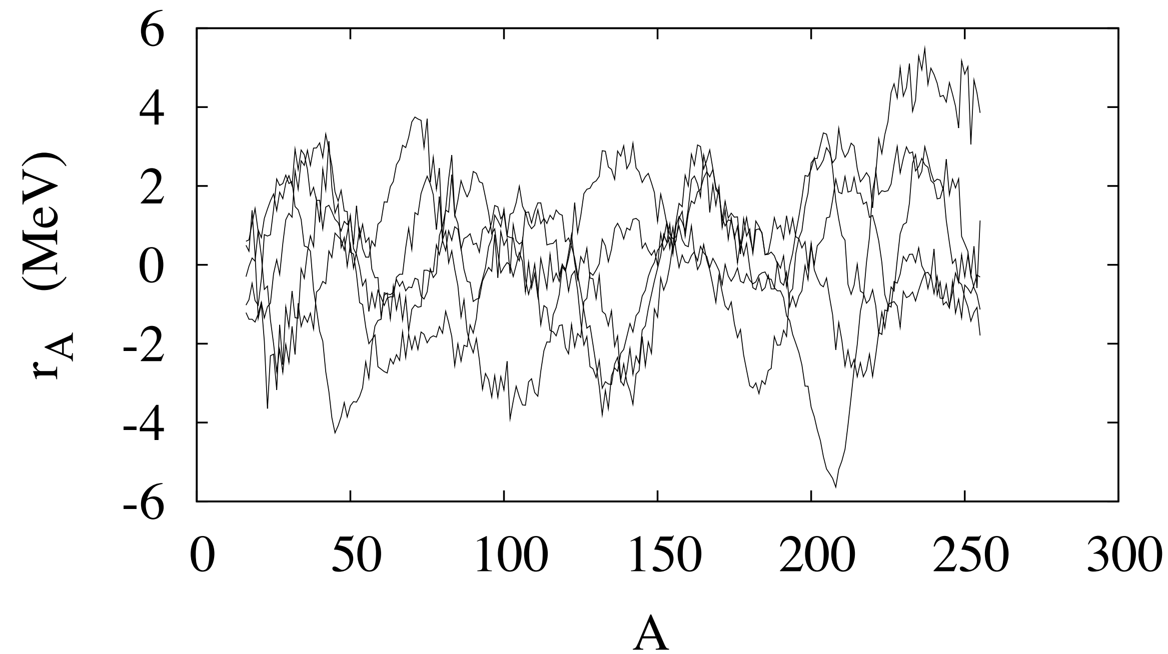
Step 1. Fourier-transform the residuals with respect to A

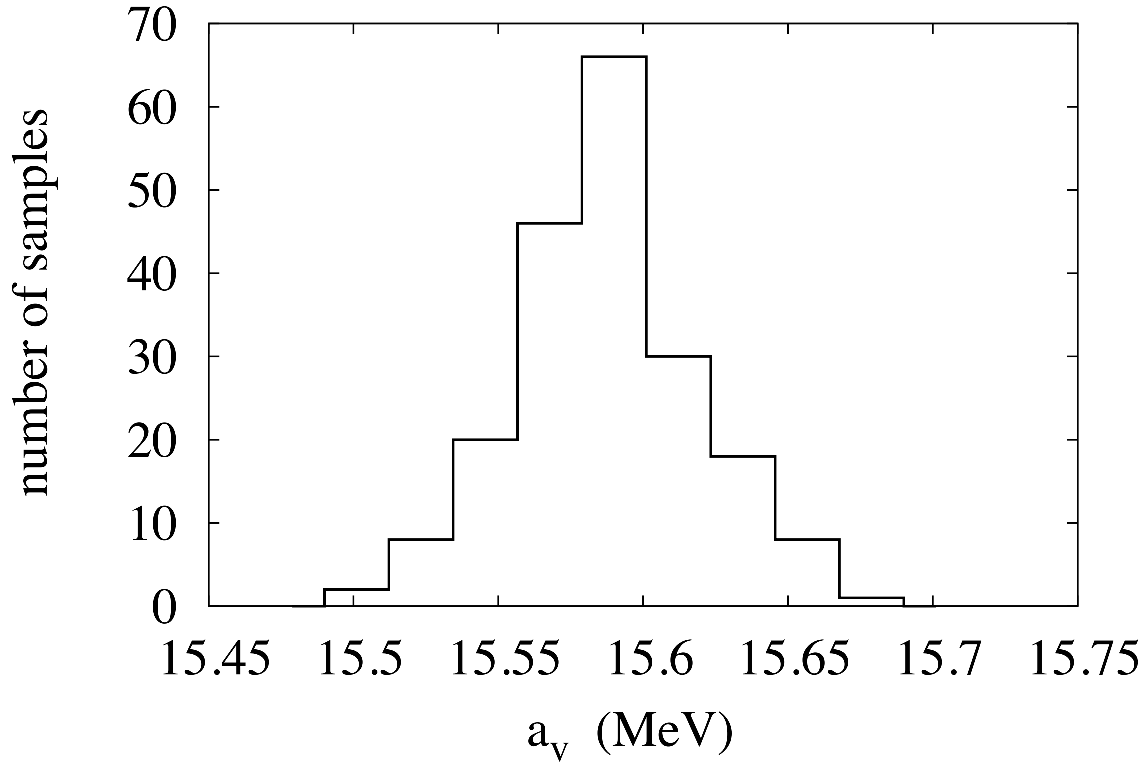
$$r_q(m) = N^{-1/2} \sum_n \exp(2i\pi n_i m/N) r_i$$





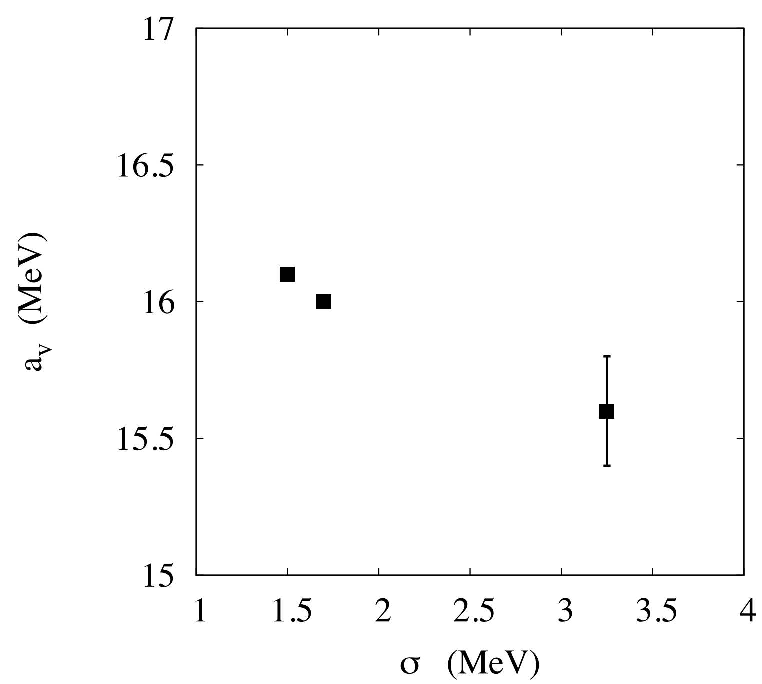
Step 2. The ensemble to sample has the same $|r_q|$ set but with arbitrary phases.





$$a_v = 15.6 \pm 0.2 \ \text{MeV}$$

Is the derived error robust?
Better models (i.e. DFT) should fall in the error band.



kk

Theory	r.m.s. residual	$a_v(c)$	$a_s(c)$
SLy4 [8]	1.7	−16.06	32.0
SkP-based [11]	1.7	−16.11	31.1
BSk4-based [12]	1.7	−16.03	29.6
Skxce-based [13]	1.5	−16.10	31.0
LD	3.1	−15.6	23.3