TALENT COURSE 11
LEARNING FROM DATA: BAYESIAN METHODS AND MACHINE LEARNING

Lecture 25: Experimental design

Daniel Phillips
Ohio University
TU Darmstadt
ExtreMe Matter Institute

TALENT Course 11 is possible thanks to funding from the STFC

Recycling old examples

Recycling old examples

- Suppose you want to determine H_0 to a given accuracy σ_T. How do you do that?

$$\sigma = \sqrt{\frac{H_0(1 - H_0)}{N}}$$

Recycling old examples

- Suppose you want to determine H_0 to a given accuracy σ_T. How do you do that?

$$\sigma = \sqrt{\frac{H_0(1-H_0)}{N}}$$

- σ depends on H_0, although only weakly

Recycling old examples

- Suppose you want to determine H_0 to a given accuracy σ_T. How do you do that?

$$\sigma = \sqrt{\frac{H_0(1 - H_0)}{N}}$$

- σ depends on H_0, although only weakly

- Signal plus background: design choices matter in terms of whether you extract B better or A+B better

Recycling old examples

- Suppose you want to determine H_0 to a given accuracy σ_T. How do you do that?

$$\sigma = \sqrt{\frac{H_0(1 - H_0)}{N}}$$

- σ depends on H_0, although only weakly

- Signal plus background: design choices matter in terms of whether you extract B better or A+B better

- Binning choices can also be interesting, although not discussed here

A (sort-of) new example: convolution

Sivia, Section 7.2

A (sort-of) new example: convolution

Sivia, Section 7.2

- “Ideal data” related to PSF R and “spectrum of interest” f by

$$F_k = T\int f(x)R(x_k - x)dx + B(x_k)$$

A (sort-of) new example: convolution

Sivia, Section 7.2

- "Ideal data" related to PSF R and "spectrum of interest" f by

$$F_k = T\int f(x)R(x_k - x)dx + B(x_k)$$

- Here we take $B(x) = B$ fixed; $R(x) \propto \exp\left(-\frac{x^2}{2w^2}\right)$

A (sort-of) new example: convolution

Sivia, Section 7.2

- "Ideal data" related to PSF R and "spectrum of interest" f by

$$F_k = T \int f(x) R(x_k - x)\, dx + B(x_k)$$

- Here we take $B(x) = B$ fixed; $R(x) \propto \exp\left(-\frac{x^2}{2w^2}\right)$

- Observed data related to F_k by Poisson statistics:

$$\mathrm{pr}(\{D_k\} \mid \{F_k\}, I) = \prod_{k=1}^{N} \frac{F_k^{D_k} e^{-D_k}}{D_k!}$$

A (sort-of) new example: convolution

Sivia, Section 7.2

- "Ideal data" related to PSF R and "spectrum of interest" f by

$$F_k = T\int f(x)R(x_k - x)dx + B(x_k)$$

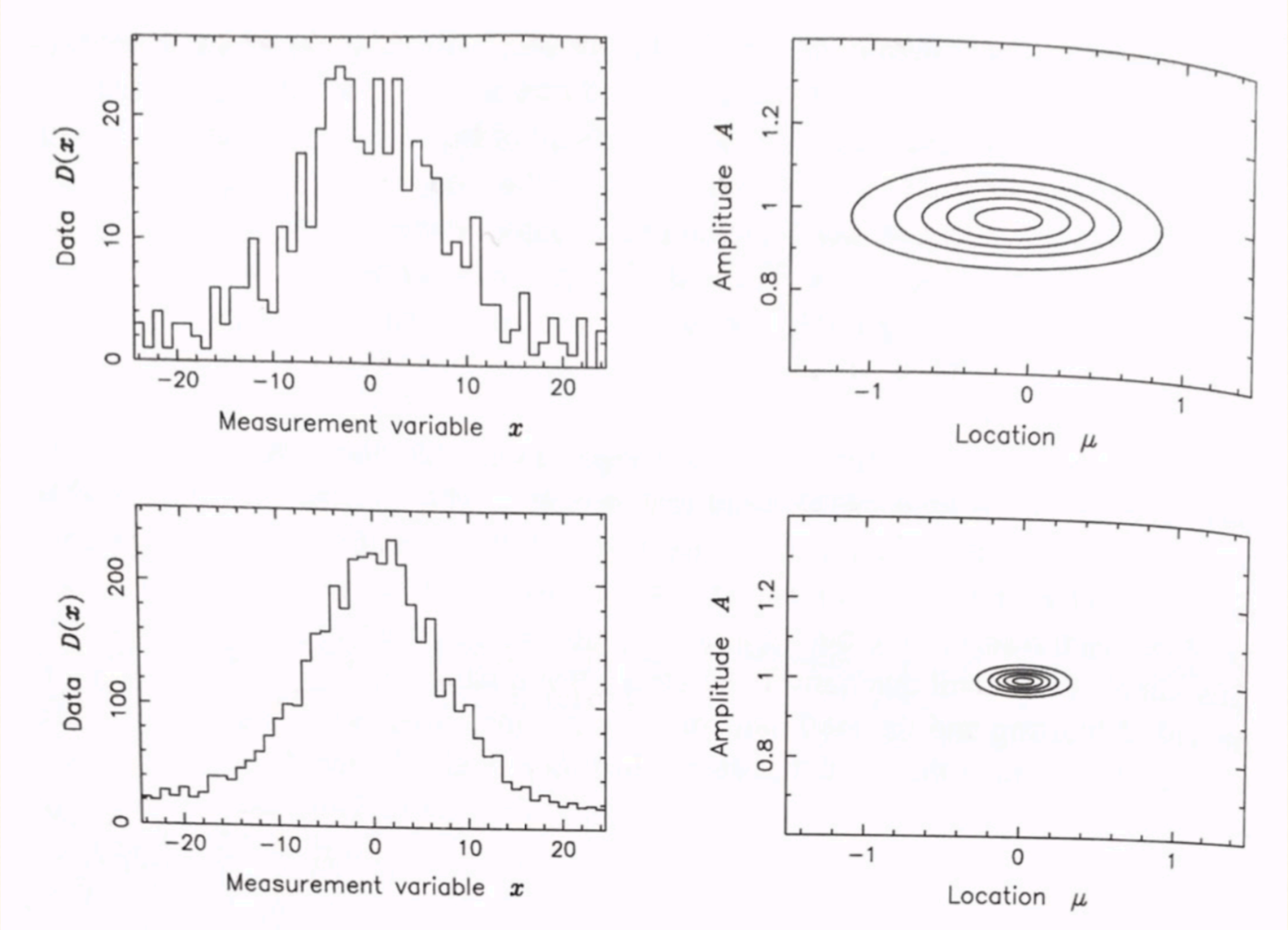
- Here we take $B(x) = B$ fixed; $R(x) \propto \exp\left(-\frac{x^2}{2w^2}\right)$

- Observed data related to F_k by Poisson statistics:

$$\mathrm{pr}(\{D_k\}\,|\,\{F_k\}, I) = \prod_{k=1}^{N} \frac{F_k^{D_k} e^{-D_k}}{D_k!}$$

- Let's do $f(x) = A\delta(x-\mu)$; goal is to choose T and w so as to obtain best measure/estimate A and/or μ

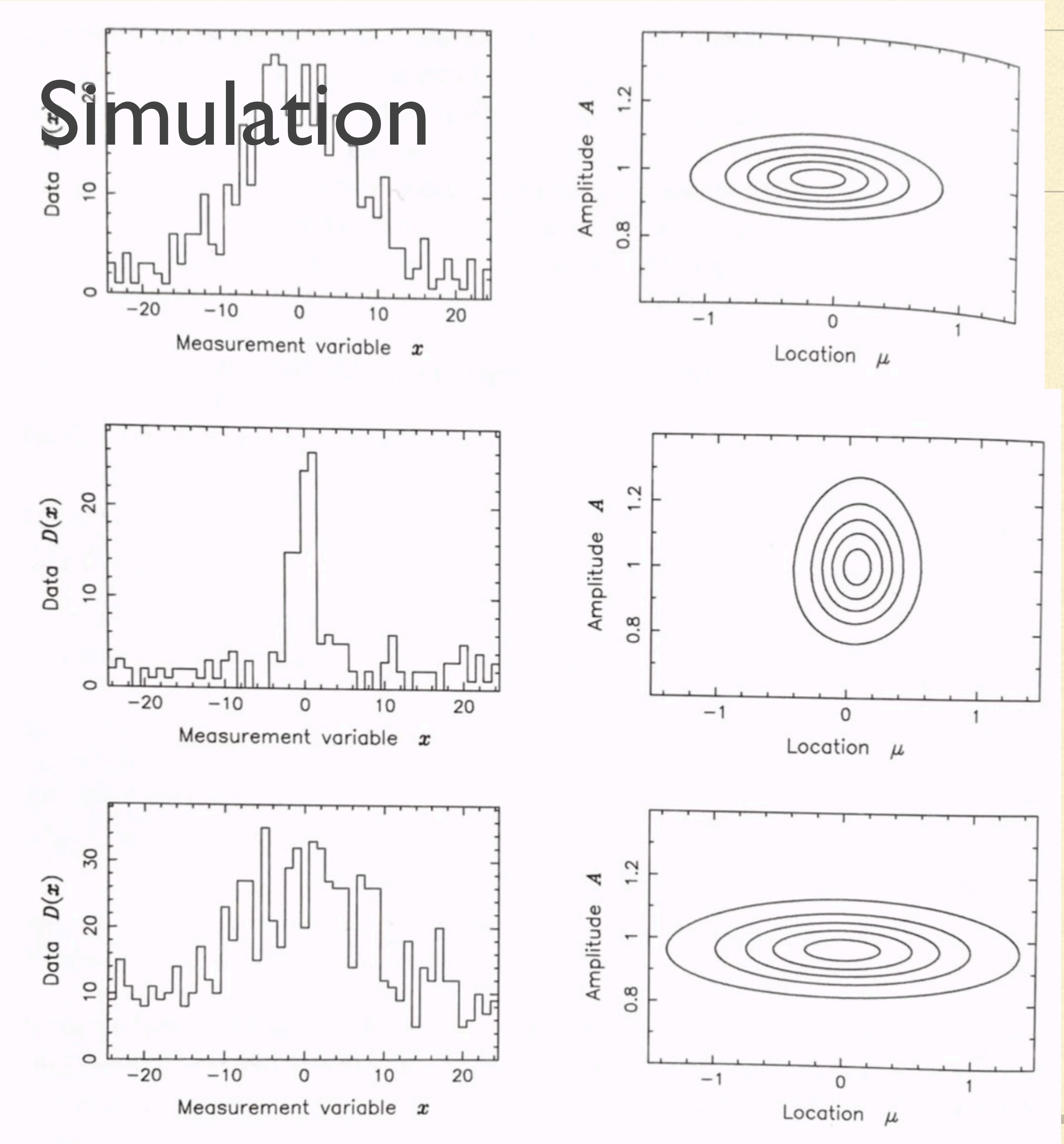
Simulation



T=20, w=10, B=T/10

T=200, w=10, B=T/10

Simulation



T=20, w=10, B=T/10

T=20, w=4, B=T/10

T=20, w=4, B=T/2

Analytic insight

Analytic insight

- Choose $A=1$; $\mu=0$

Analytic insight

- Choose A=1; μ=0
- $\hat{D}_k = T \exp(-x_k^2/(2w^2)) + B$

Analytic insight

- Choose A=1; μ=0
- $\hat{D}_k = T \exp(-x_k^2/(2w^2)) + B$
- Model $D_k = \hat{D}_k + \epsilon_k; \langle \epsilon_k \rangle = 0; \langle \epsilon_k^2 \rangle = \sigma_k^2 = \hat{D}_k$

Analytic insight

- Choose A=1; μ=0
- $\hat{D}_k = T\exp(-x_k^2/(2w^2)) + B$
- Model $D_k = \hat{D}_k + \epsilon_k; \langle \epsilon_k \rangle = 0; \langle \epsilon_k^2 \rangle = \sigma_k^2 = \hat{D}_k$
- log likelihood=$\mathrm{constant} - \chi^2/2$

Analytic insight

- Choose A=1; μ=0
- $\hat{D}_k = T \exp(-x_k^2/(2w^2)) + B$
- Model $D_k = \hat{D}_k + \epsilon_k; \langle \epsilon_k \rangle = 0; \langle \epsilon_k^2 \rangle = \sigma_k^2 = \hat{D}_k$
- log likelihood=$\text{constant} - \chi^2/2$
- Compute second derivatives, approximate sums by integrals and find if B≪T

Analytic insight

- Choose A=1; μ=0
- $\hat{D}_k = T\exp(-x_k^2/(2w^2)) + B$
- Model $D_k = \hat{D}_k + \epsilon_k; \langle \epsilon_k \rangle = 0; \langle \epsilon_k^2 \rangle = \sigma_k^2 = \hat{D}_k$
- log likelihood=$\mathrm{constant} - \chi^2/2$
- Compute second derivatives, approximate sums by integrals and find if B≪T

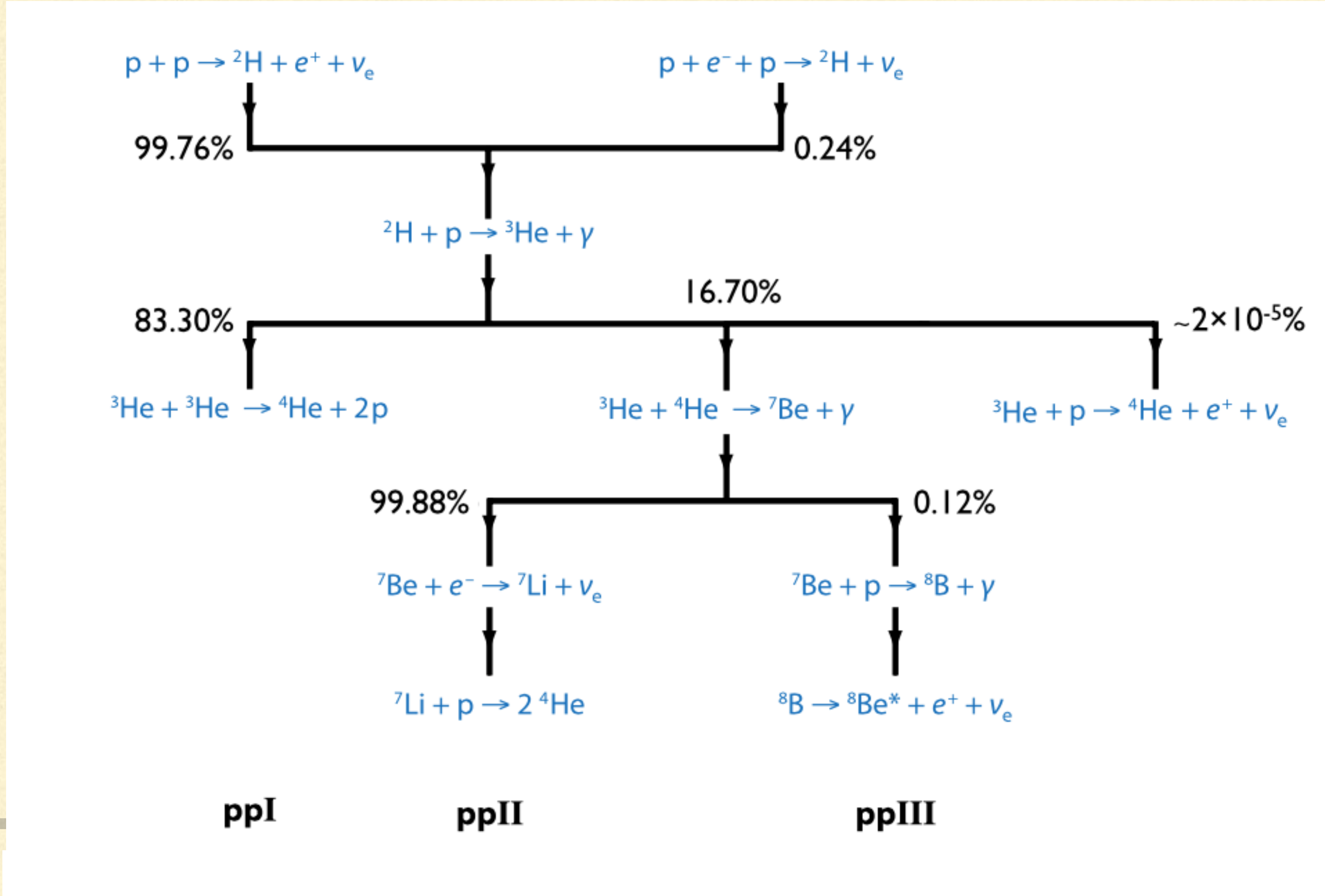
$$\begin{pmatrix} \sigma_\mu^2 & \rho\sigma_\mu\sigma_A \\ \rho\sigma_\mu\sigma_A & \sigma_A^2 \end{pmatrix} \propto \frac{1}{T}\begin{pmatrix} w & 0 \\ 0 & 1/w \end{pmatrix}$$

Experimental planning: fixed model

- Decide what you want to measure
- Understand resources (beam time, number of detectors, etc.)
- Simulate for different arrangement of resources
- Choose arrangement that gives you most precise (or should it be most accurate?) result for the quantity of interest
- May require soul searching
- Or choice of utility function

Why is $^7Be + p \rightarrow {}^8B + \gamma$ important?

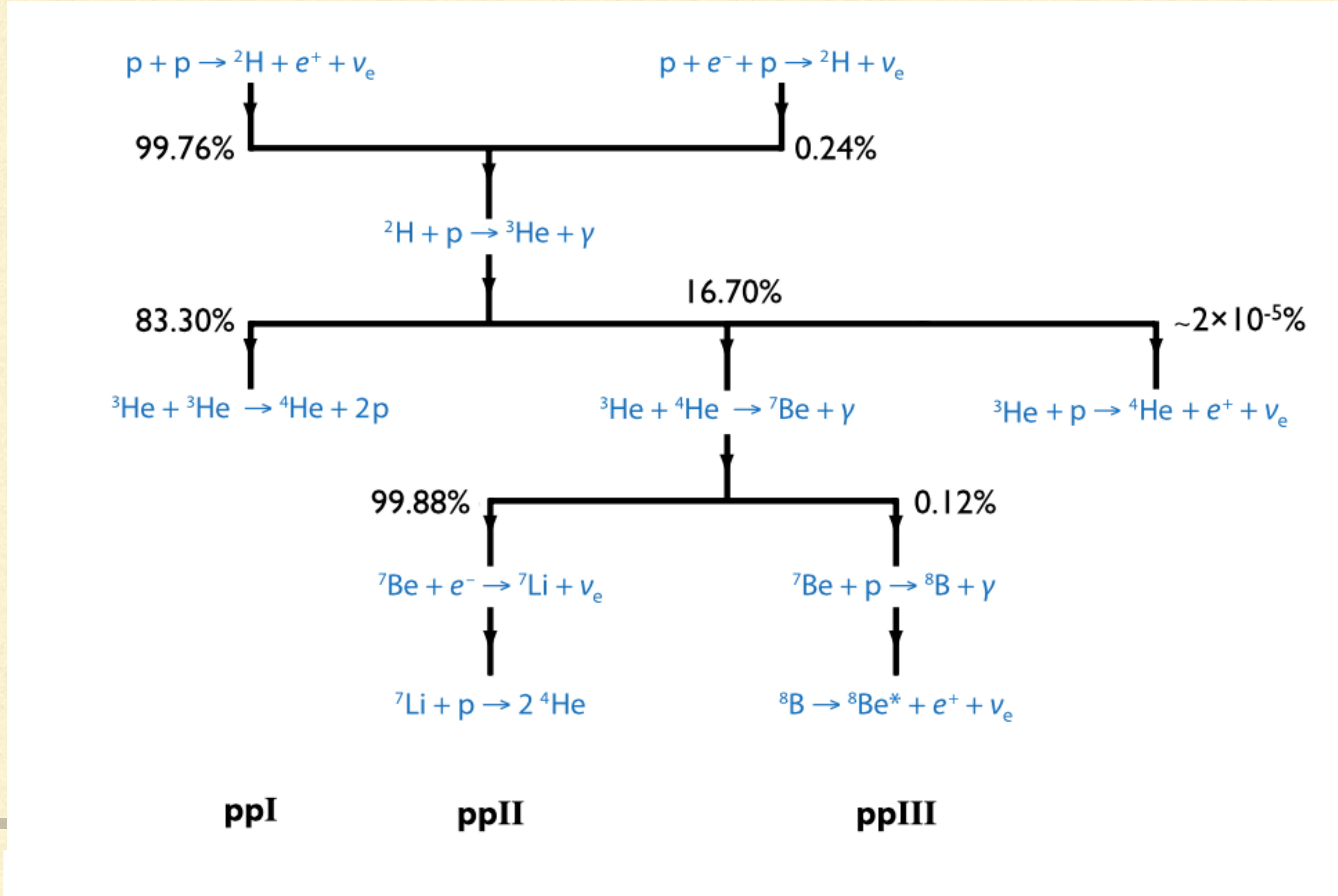
Adelberger et al., Rev. Mod. Phys. 83, 195 (2011)



Why is $^7Be + p \rightarrow {}^8B + \gamma$ important?

Adelberger et al., Rev. Mod. Phys. 83, 195 (2011)

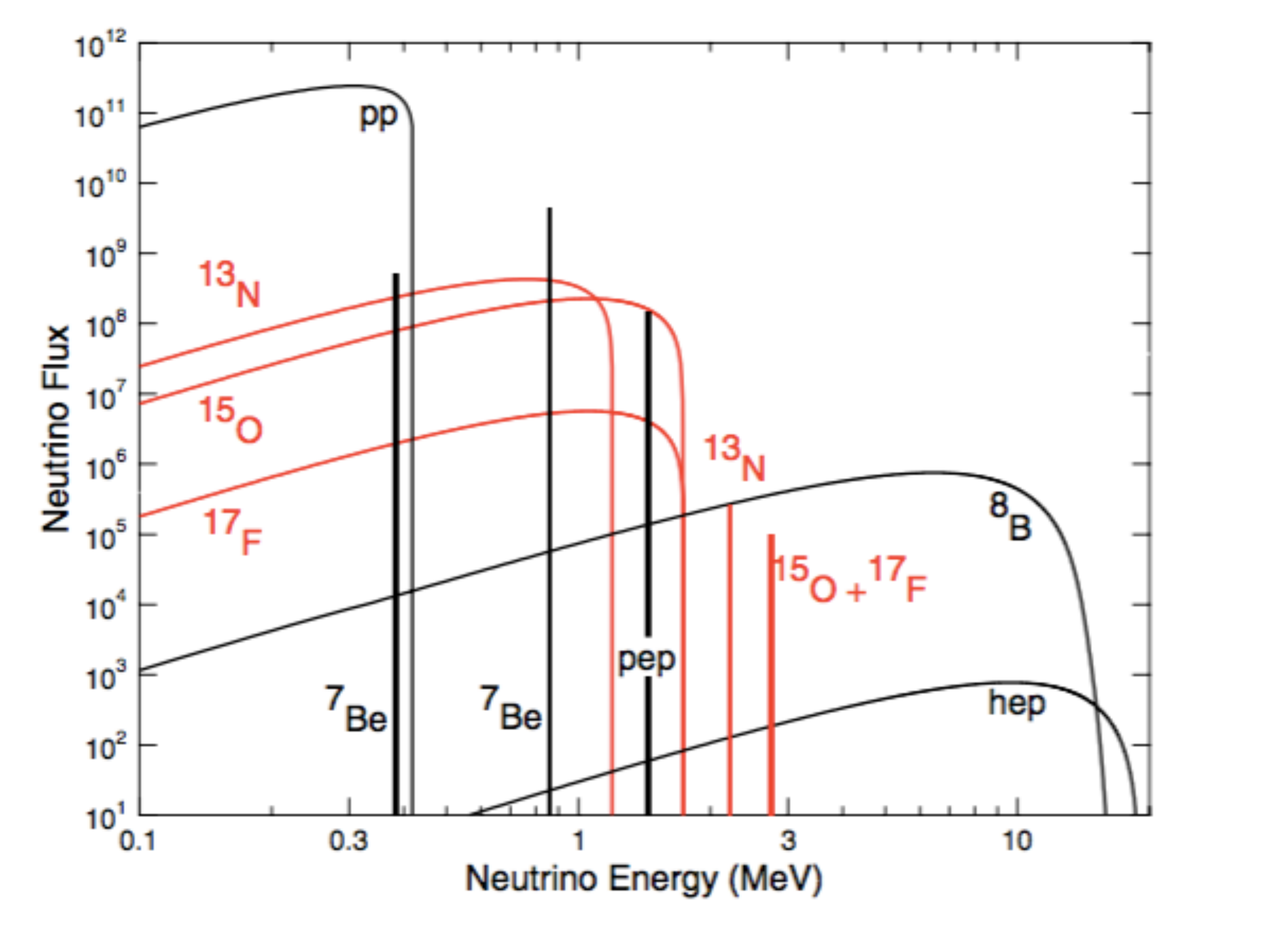
- Part of pp chain (ppIII)



Why is $^{7}Be + p \rightarrow {}^{8}B + \gamma$ important?

Adelberger et al., Rev. Mod. Phys. 83, 195 (2011)

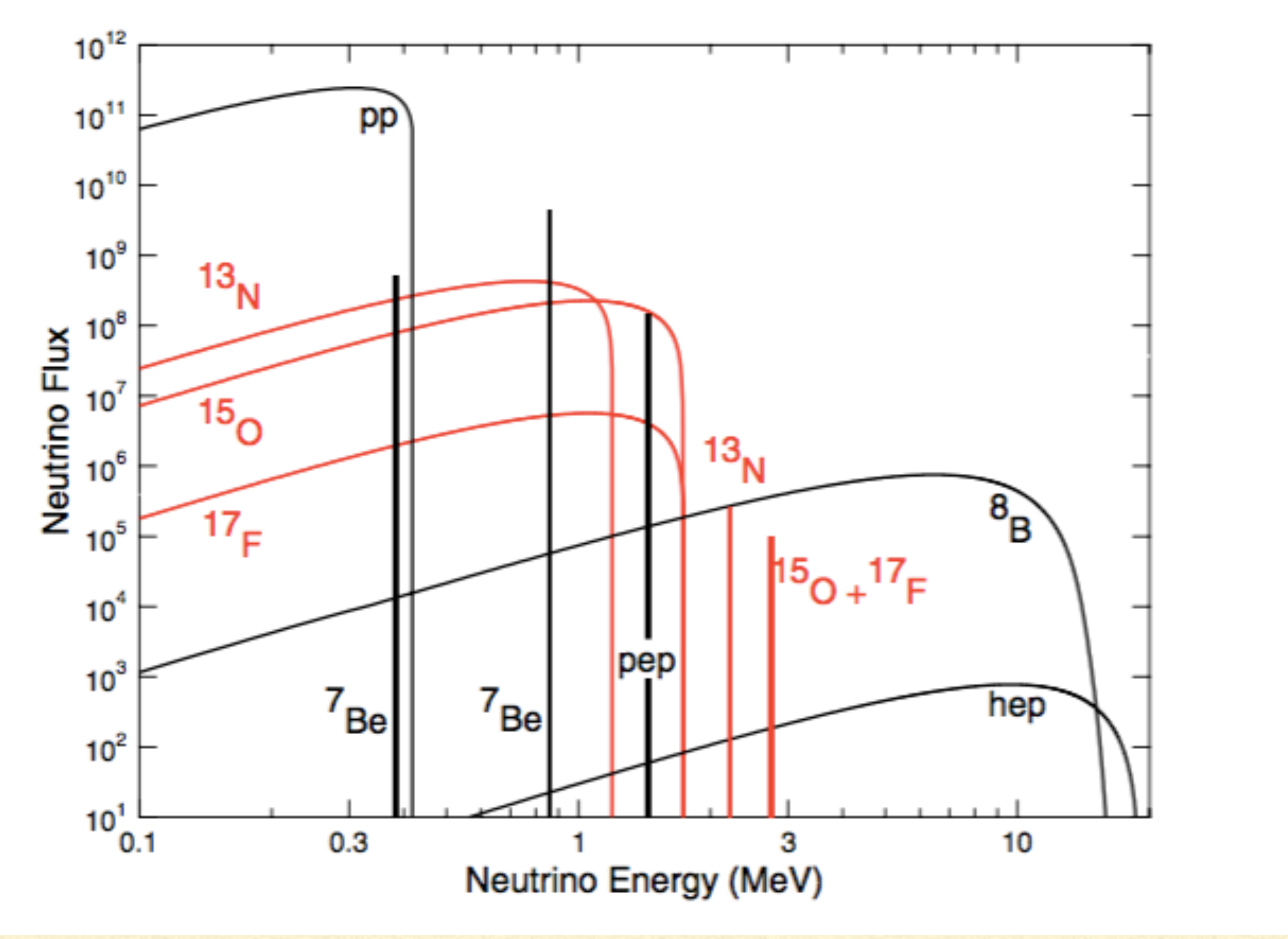
- Part of pp chain (ppIII)
- Key for predicting flux of solar neutrinos, especially high-energy (^{8}B) neutrinos



Why is $^7Be + p \rightarrow {}^8B + \gamma$ important?

Adelberger et al., Rev. Mod. Phys. 83, 195 (2011)

- Part of pp chain (ppIII)
- Key for predicting flux of solar neutrinos, especially high-energy (^8B) neutrinos
- Accurate knowledge of $^7Be(p,\gamma)$ needed for inferences from solar-neutrino flux regarding solar composition →solar-system formation history

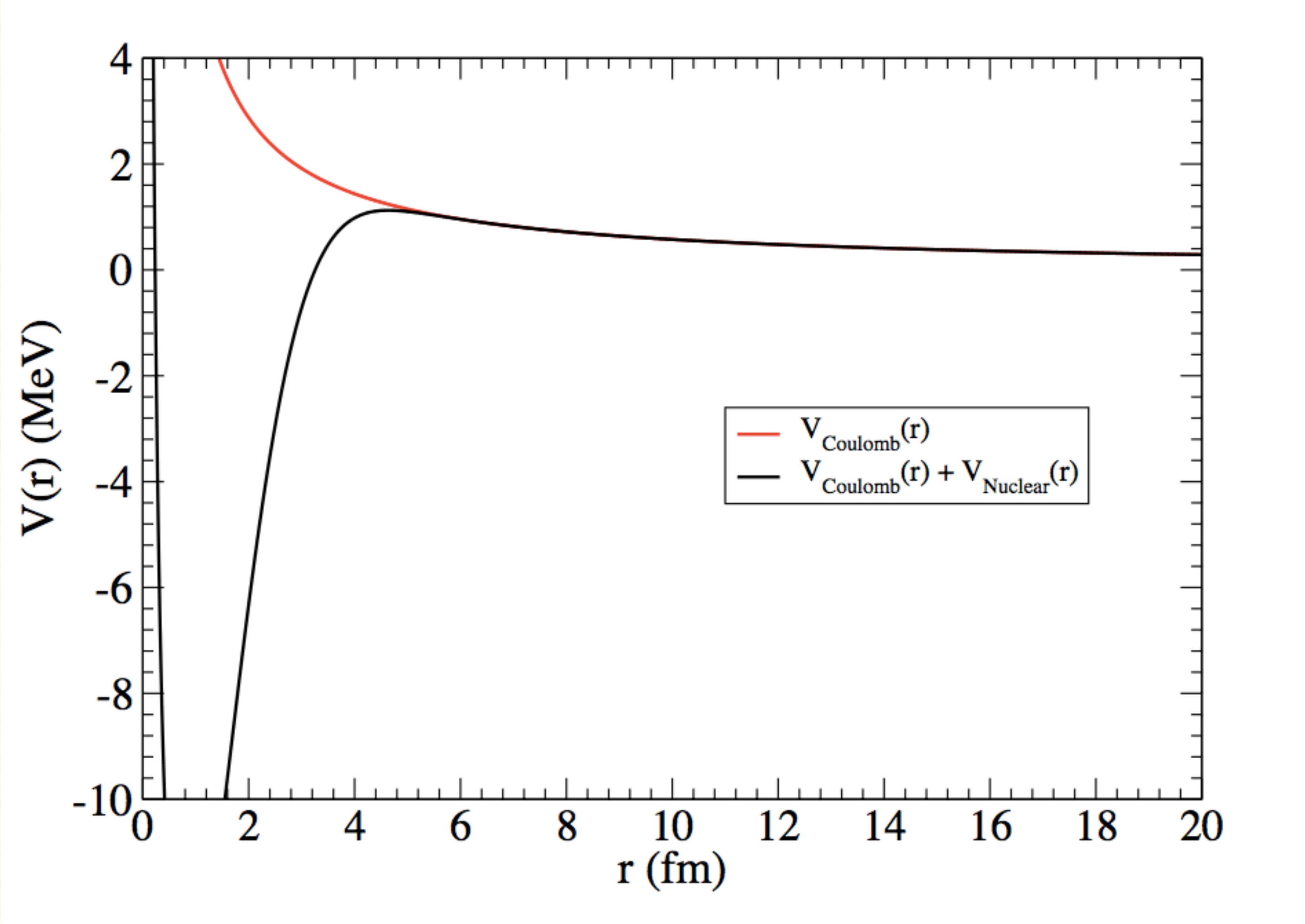


This is an extrapolation problem

Thermonuclear reaction rate $\propto \langle v\sigma \rangle \propto \int_0^\infty dE \, \exp\left(-\frac{E}{k_B T}\right) E\,\sigma(E)$

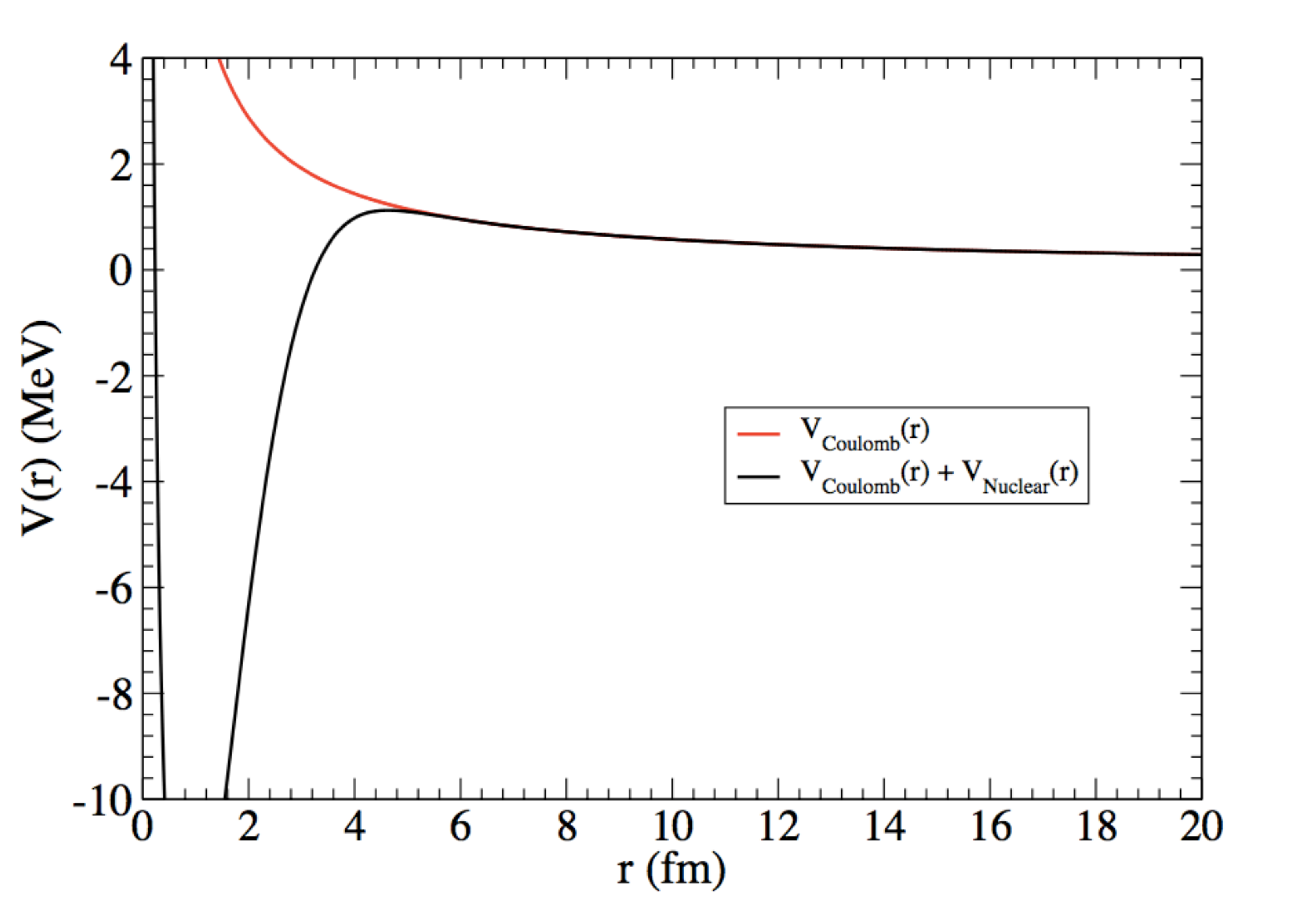
This is an extrapolation problem

Thermonuclear reaction rate $\propto \langle v\sigma \rangle \propto \int_0^\infty dE \, \exp\left(-\frac{E}{k_B T}\right) E \, \sigma(E)$



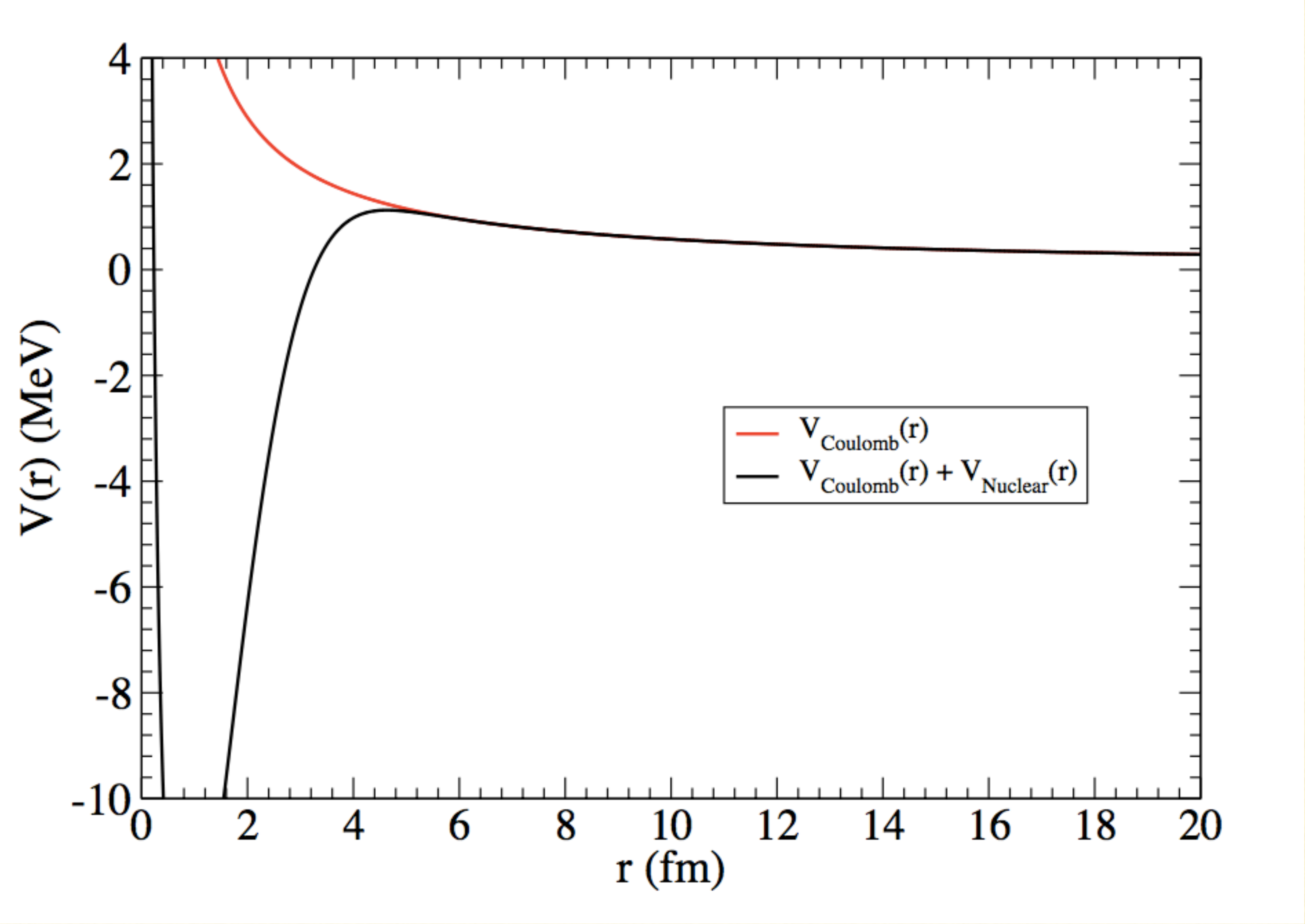
This is an extrapolation problem

Thermonuclear reaction rate $\propto \langle v\sigma \rangle \propto \int_0^\infty dE \, \exp\left(-\frac{E}{k_B T}\right) E \, \sigma(E)$



This is an extrapolation problem

Thermonuclear reaction rate $\propto \langle v\sigma \rangle \propto \int_0^\infty dE \, \exp\left(-\frac{E}{k_B T}\right) E \, \sigma(E)$



This is an extrapolation problem

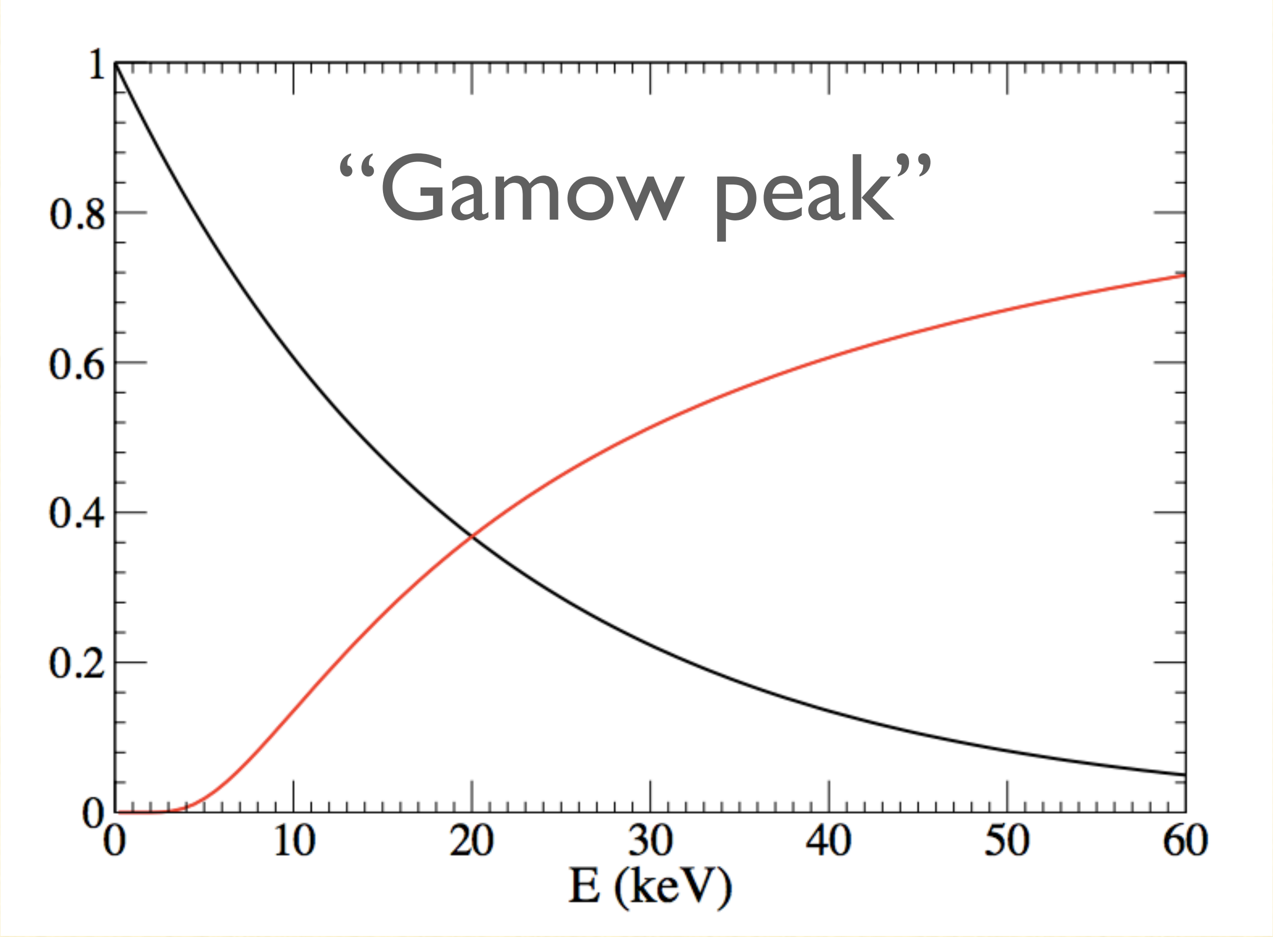
Thermonuclear reaction rate $\propto \langle v\sigma \rangle \propto \int_0^\infty dE \, \exp\left(-\frac{E}{k_B T}\right) E \, \sigma(E)$

$$\sigma(E) = \frac{S(E)}{E} \exp\left(-\pi Z_1 Z_2 \alpha_{\text{em}} \sqrt{\frac{m_R}{2E}}\right)$$

This is an extrapolation problem

Thermonuclear reaction rate $\propto \langle v\sigma \rangle \propto \int_0^\infty dE \, \exp\left(-\frac{E}{k_B T}\right) E\,\sigma(E)$

$$\sigma(E) = \frac{S(E)}{E} \exp\left(-\pi Z_1 Z_2 \alpha_{\mathrm{em}} \sqrt{\frac{m_R}{2E}}\right)$$

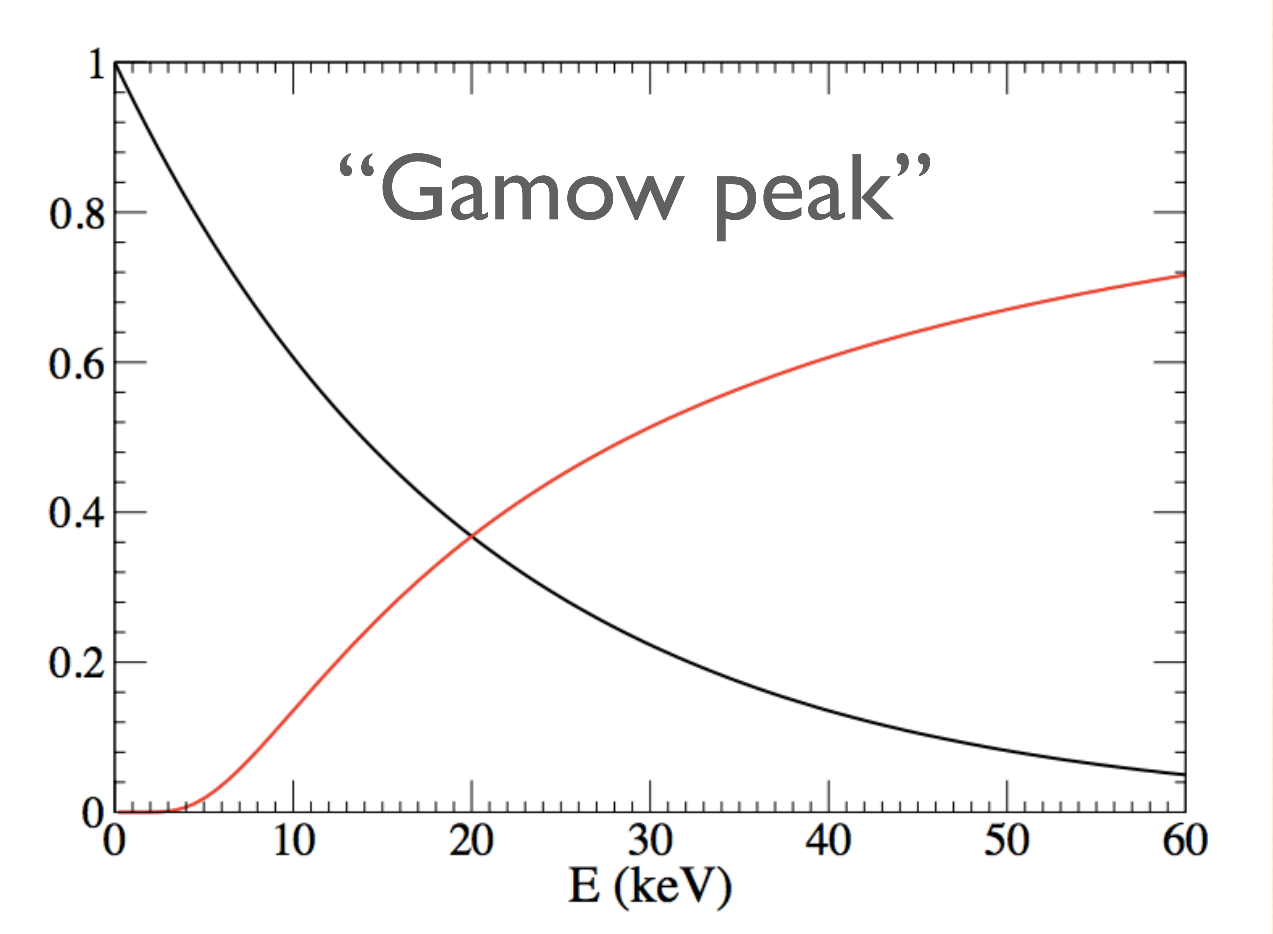


This is an extrapolation problem

Thermonuclear reaction rate $\propto \langle v\sigma \rangle \propto \int_0^\infty dE \, \exp\left(-\frac{E}{k_B T}\right) E\,\sigma(E)$

$$\sigma(E) = \frac{S(E)}{E} \exp\left(-\pi Z_1 Z_2 \alpha_{\rm em} \sqrt{\frac{m_R}{2E}}\right)$$

- E1 capture: ^{7}Be + p→^{8}B + γ
- Energies of relevance 20 keV



What matters where?

$$\mathcal{M}(E) \propto \int dr A_1 \exp(-\gamma_1 r) \left(1 + \frac{1}{\gamma_1 r}\right) r u_E(r); \quad \gamma_1 = 1/(13 \text{ fm})$$

Dominated by ^{7}Be-p separations ~10s of fm

What matters where?

$$\mathcal{M}(E) \propto \int dr\, A_1 \exp(-\gamma_1 r)\left(1 + \frac{1}{\gamma_1 r}\right) r u_E(r); \quad \gamma_1 = 1/(13\ \mathrm{fm})$$

ANC

Dominated by ^{7}Be-p separations ~10s of fm

What matters where?

$$\mathcal{M}(E) \propto \int dr\, A_1 \exp(-\gamma_1 r)\left(1 + \frac{1}{\gamma_1 r}\right) r u_E(r); \quad \gamma_1 = 1/(13 \text{ fm})$$

ANC

Dominated by ^{7}Be-p separations ~10s of fm

Numbers that matter: $k_C = Q_c Q_n \alpha_{EM} M_R = 24$ MeV; $p = \sqrt{2 m_R E}$;
$\gamma_1 = \sqrt{2 m_R B}$; a: parameterizes strength of p-^{7}Be strong scattering

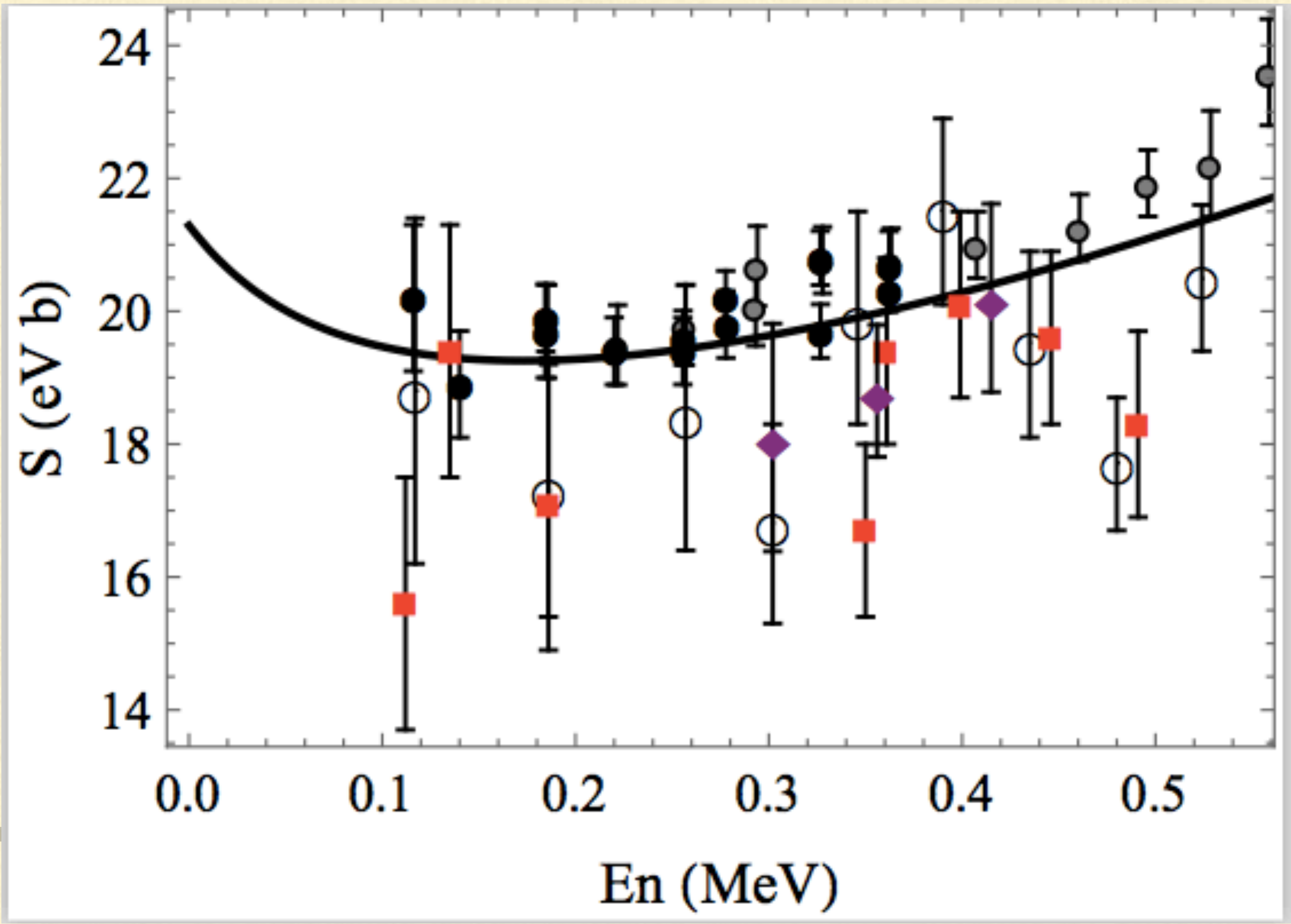
What matters where?

$$\mathcal{M}(E) \propto \int dr\, A_1 \exp(-\gamma_1 r)\left(1 + \frac{1}{\gamma_1 r}\right) r u_E(r); \quad \gamma_1 = 1/(13\ \mathrm{fm})$$

ANC

Dominated by ^{7}Be-p separations ~10s of fm

Numbers that matter: $k_C = Q_c Q_n \alpha_{EM} M_R = 24$ MeV; $p = \sqrt{2\, m_R\, E}$;
$\gamma_1 = \sqrt{2\, m_R\, B}$; a: parameterizes strength of p-^{7}Be strong scattering



What matters where?

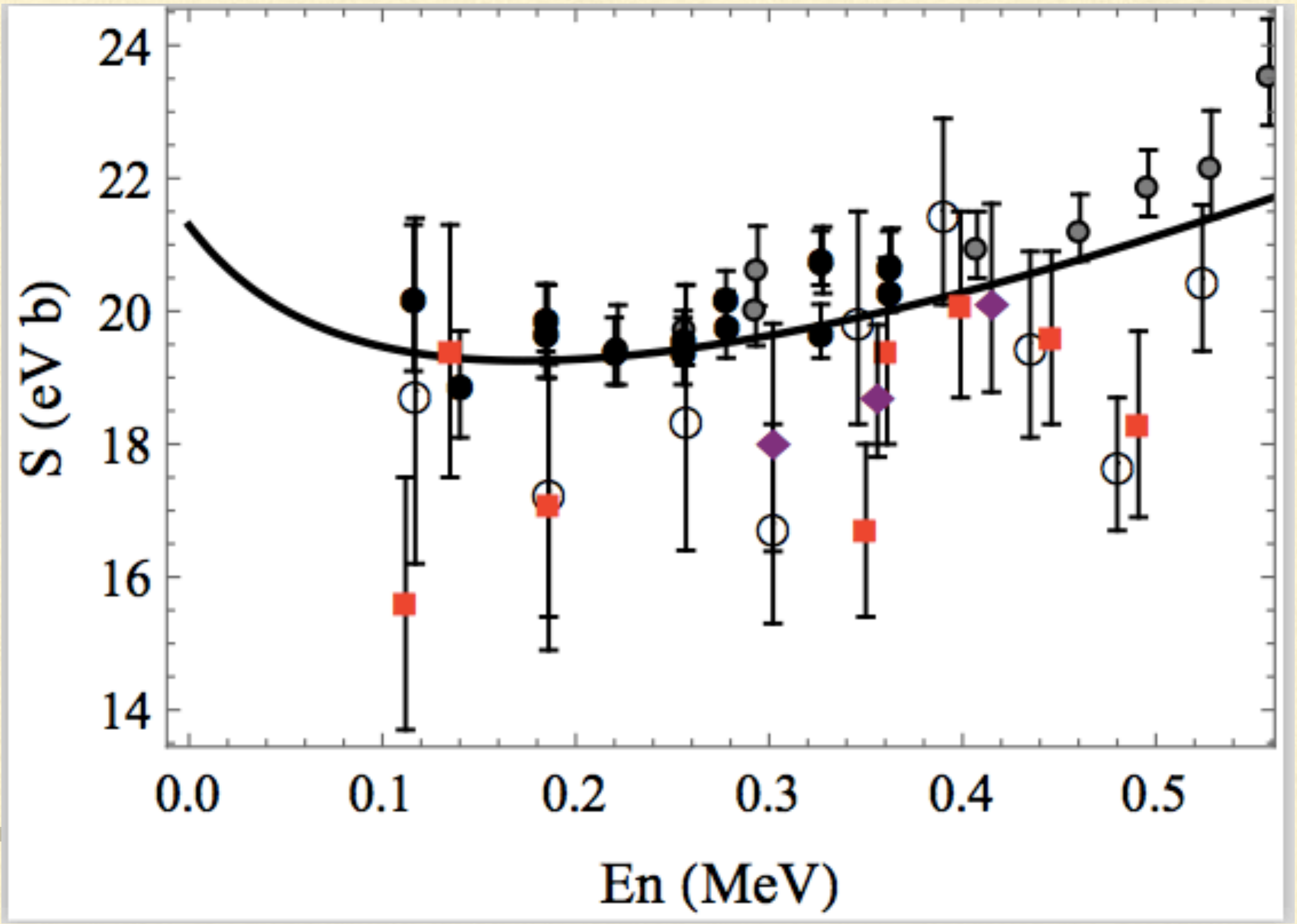
$$\mathcal{M}(E) \propto \int dr\, A_1 \exp(-\gamma_1 r)\left(1+\frac{1}{\gamma_1 r}\right) r u_E(r); \quad \gamma_1 = 1/(13\ \mathrm{fm})$$

ANC

Dominated by ^{7}Be-p separations ~10s of fm

Numbers that matter: $k_C = Q_c Q_n \alpha_{EM} M_R = 24$ MeV; $p = \sqrt{2\, m_R\, E}$; $\gamma_1 = \sqrt{2\, m_R\, B}$; a: parameterizes strength of p-^{7}Be strong scattering

- Extrapolation is not a polynomial: non-analyticities in p/k_C, p/γ_1, and $p\,a$



What matters where?

$$\mathcal{M}(E) \propto \int dr A_1 \exp(-\gamma_1 r) \left(1 + \frac{1}{\gamma_1 r}\right) r u_E(r); \quad \gamma_1 = 1/(13 \text{ fm})$$

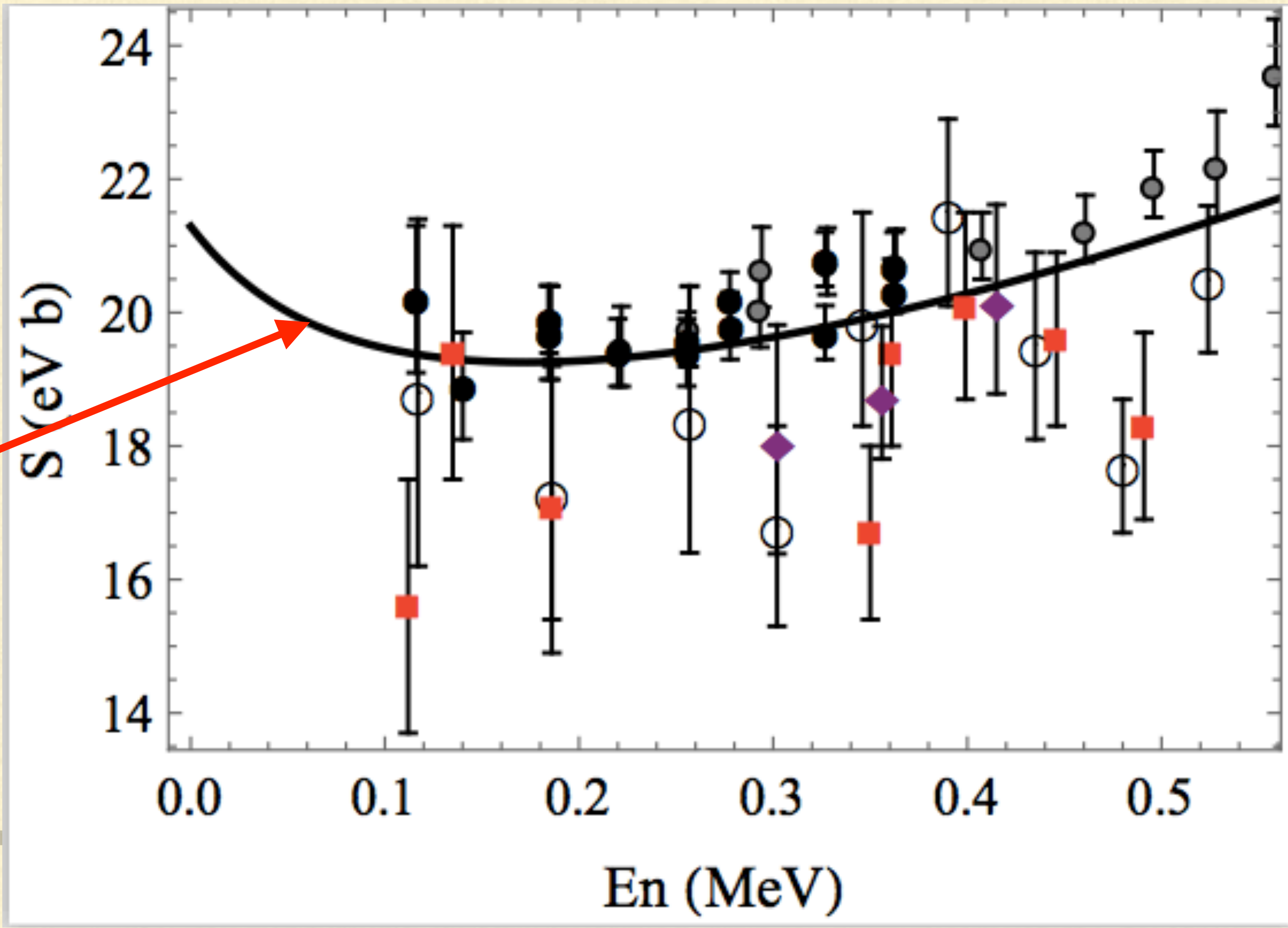
ANC

Dominated by ^{7}Be-p separations ~10s of fm

Numbers that matter: $k_C=Q_cQ_n\alpha_{EM}M_R=24$ MeV; $p=\sqrt{2\, m_R\, E}$; $\gamma_1=\sqrt{2\, m_R\, B}$; a: parameterizes strength of p-^{7}Be strong scattering

- Extrapolation is not a polynomial: non-analyticities in p/k_C, p/γ_1, and $p\,a$

Bound state (ANC & γ_1) + Coulomb



What matters where?

$$\mathcal{M}(E) \propto \int dr\, A_1 \exp(-\gamma_1 r) \left(1 + \frac{1}{\gamma_1 r}\right) r u_E(r); \quad \gamma_1 = 1/(13\ \mathrm{fm})$$

ANC

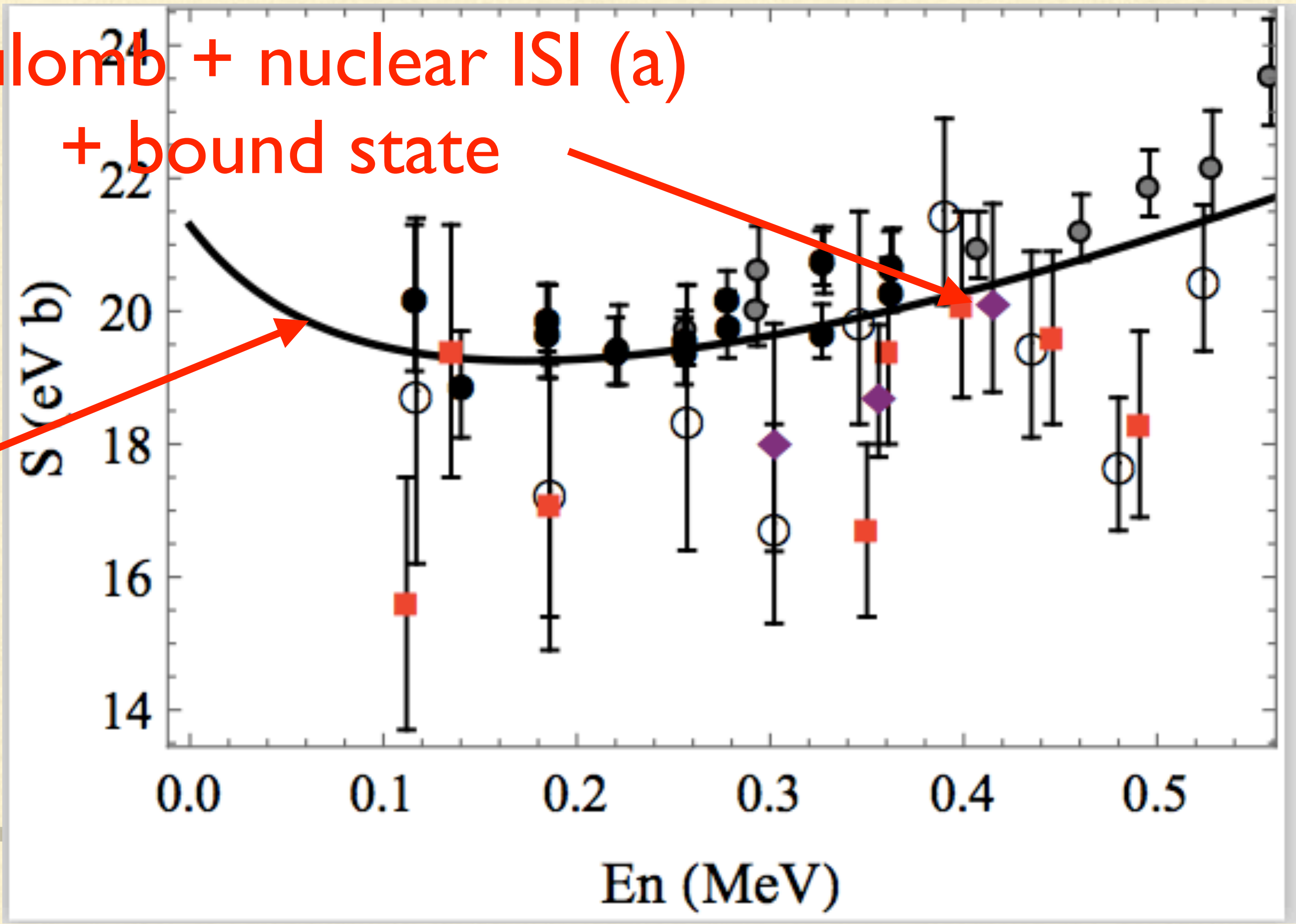
Dominated by ^{7}Be-p separations ~10s of fm

Numbers that matter: $k_C = Q_c Q_n \alpha_{EM} M_R = 24$ MeV; $p = \sqrt{2\, m_R\, E}$; $\gamma_1 = \sqrt{2\, m_R\, B}$; a: parameterizes strength of p-^{7}Be strong scattering

- Extrapolation is not a polynomial: non-analyticities in p/k_C, p/γ_1, and p a

Coulomb + nuclear ISI (a) + bound state

Bound state (ANC &γ_1) + Coulomb



What matters where?

$$\mathcal{M}(E) \propto \int dr\, A_1 \exp(-\gamma_1 r)\left(1+\frac{1}{\gamma_1 r}\right) r u_E(r); \quad \gamma_1 = 1/(13\ \mathrm{fm})$$

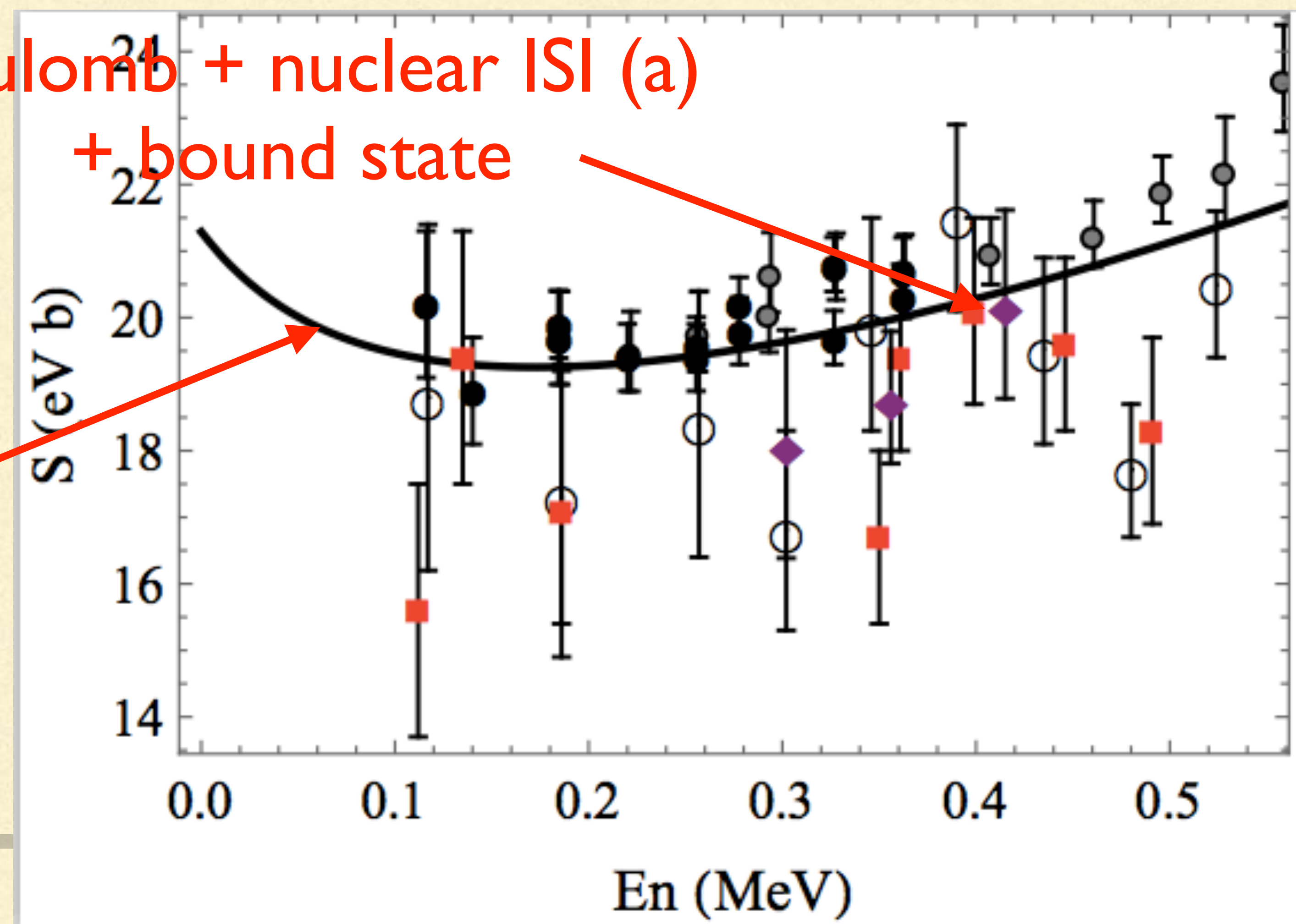
ANC

Dominated by ^{7}Be-p separations ~10s of fm

Numbers that matter: $k_C = Q_c Q_n \alpha_{EM} M_R = 24$ MeV; $p = \sqrt{2\, m_R\, E}$; $\gamma_1 = \sqrt{2\, m_R\, B}$; a: parameterizes strength of p-^{7}Be strong scattering

- Extrapolation is not a polynomial: non-analyticities in p/k_C, p/γ_1, and $p\, a$
- Sub-leading polynomial behavior in E/E_{core}

Coulomb + nuclear ISI (a) + bound state



Bound state (ANC & γ_1) + Coulomb

Data for $^{7}Be + p \rightarrow {}^{8}B + \gamma_{E1}$

- 42 data points for 100 keV < $E_{c.m.}$ < 500 keV
  - Junghans (BE1 and BE3)
  - Fillipone
  - Baby
  - Hammache (1998 and 2001)

Data for $^{7}Be + p \rightarrow {}^{8}B + \gamma_{E1}$

- 42 data points for 100 keV < $E_{c.m.}$ < 500 keV
  - Junghans (BE1 and BE3)
  - Fillipone
  - Baby
  - Hammache (1998 and 2001)
- CMEs
  - 2.7% and 2.3%
  - 11.25%
  - 5%
  - 2.2% (1998)

Data for ^{7}Be + p→^{8}B + γ_{E1}

- 42 data points for 100 keV < $E_{c.m.}$ < 500 keV
  - Junghans (BE1 and BE3)
  - Fillipone
  - Baby
  - Hammache (1998 and 2001)
- CMEs
  - 2.7% and 2.3%
  - 11.25%
  - 5%
  - 2.2% (1998)

- Subtract M1 resonance: negligible impact at 500 keV and below
- Deal with CMEs by introducing five additional parameters, ξ_i

Building the pdf

$$\mathrm{pr}(\overrightarrow{\theta}, \{\xi_j\} \,|\, D, I) \propto \mathrm{pr}(D \,|\, \overrightarrow{\theta}, \{\xi_j\}, I)\mathrm{pr}(\overrightarrow{\theta}, \{\xi_j\} \,|\, I)$$

$$\ln \mathrm{pr}(D \,|\, \overrightarrow{\theta}, \{\xi_j\}, I) = c - \frac{1}{2} \sum_{j=1}^{N_{\mathrm{expt}}} \sum_{i=1}^{N^j_{\mathrm{data}}} \frac{(d_{ji} - \xi_j S(E_{ji}; \overrightarrow{\theta}))^2}{\sigma_{ji}^2}$$

Building the pdf

- Bayes:

$$\mathrm{pr}(\overrightarrow{\theta}, \{\xi_j\} \,|\, D, I) \propto \mathrm{pr}(D \,|\, \overrightarrow{\theta}, \{\xi_j\}, I)\mathrm{pr}(\overrightarrow{\theta}, \{\xi_j\} \,|\, I)$$

- First factor: likelihood

$$\ln \mathrm{pr}(D \,|\, \overrightarrow{\theta}, \{\xi_j\}, I) = c - \frac{1}{2} \sum_{j=1}^{N_{\mathrm{expt}}} \sum_{i=1}^{N_{\mathrm{data}}^{j}} \frac{(d_{ji} - \xi_j S(E_{ji}; \overrightarrow{\theta}))^2}{\sigma_{ji}^2}$$

Building the pdf

- Bayes:

$$\mathrm{pr}(\overrightarrow{\theta}, \{\xi_j\} \,|\, D, I) \propto \mathrm{pr}(D \,|\, \overrightarrow{\theta}, \{\xi_j\}, I)\mathrm{pr}(\overrightarrow{\theta}, \{\xi_j\} \,|\, I)$$

- First factor: likelihood

$$\ln \mathrm{pr}(D \,|\, \overrightarrow{\theta}, \{\xi_j\}, I) = c - \frac{1}{2} \sum_{j=1}^{N_{\mathrm{expt}}} \sum_{i=1}^{N^j_{\mathrm{data}}} \frac{(d_{ji} - \xi_j S(E_{ji}; \overrightarrow{\theta}))^2}{\sigma_{ji}^2}$$

- Second factor: priors
  - Independent gaussian priors for ξ_j, centered at zero and with width=CME
  - Gaussian priors for $a_{S=1}$ and $a_{S=2}$, based on Angulo et al. measurement
  - Other EFT parameters, $r_{S=1}$, $r_{S=2}$, L_1, L_2, ANCs, ϵ_1, assigned flat priors, corresponding to natural ranges
  - No s-wave resonance below 600 keV

Building the pdf

- Bayes:

$$\mathrm{pr}(\vec{\theta},\{\xi_j\}\,|\,D,I) \propto \mathrm{pr}(D\,|\,\vec{\theta},\{\xi_j\},I)\,\mathrm{pr}(\vec{\theta},\{\xi_j\}\,|\,I)$$

- First factor: likelihood

$$\ln \mathrm{pr}(D\,|\,\vec{\theta},\{\xi_j\},I) = c - \frac{1}{2}\sum_{j=1}^{N_{\mathrm{expt}}}\sum_{i=1}^{N^j_{\mathrm{data}}}\frac{(d_{ji}-\xi_j S(E_{ji};\vec{\theta}))^2}{\sigma_{ji}^2}$$

- Second factor: priors
  - Independent gaussian priors for ξ_j, centered at zero and with width=CME
  - Gaussian priors for $a_{S=1}$ and $a_{S=2}$, based on Angulo et al. measurement

Extrinsic information

  - Other EFT parameters, $r_{S=1}$, $r_{S=2}$, L_1, L_2, ANCs, ϵ_1, assigned flat priors, corresponding to natural ranges
  - No s-wave resonance below 600 keV

Final result

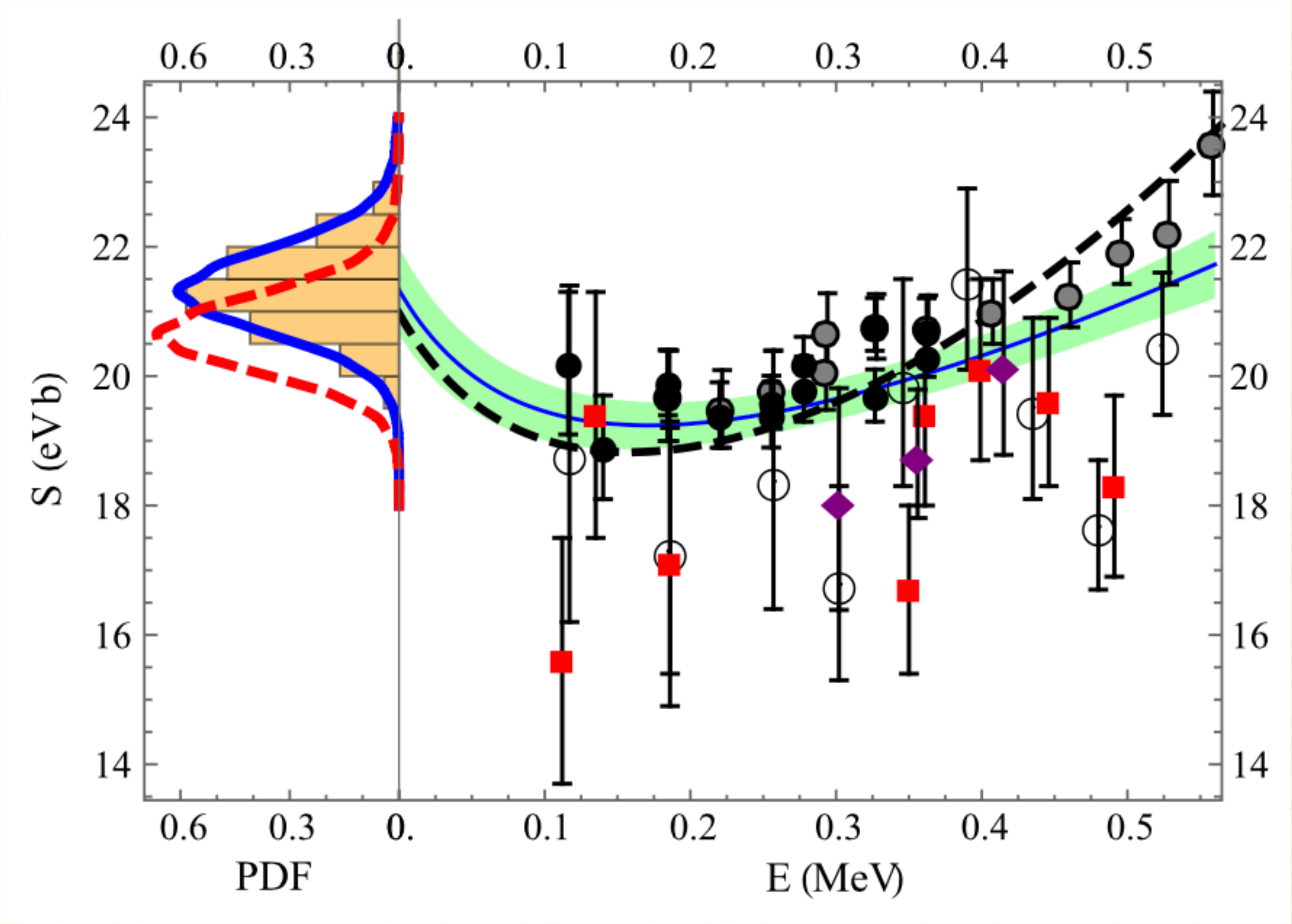
Zhang, Nollett, DP, PLB, 2015

$$\mathrm{pr}\left(\bar{F}|D;T;I\right) = \int \mathrm{pr}\left(\vec{g},\{\xi_i\}|D;T;I\right)\delta(\bar{F}-F(\vec{g}))d\xi_1\ldots d\xi_5 d\vec{g}$$

Final result

Zhang, Nollett, DP, PLB, 2015

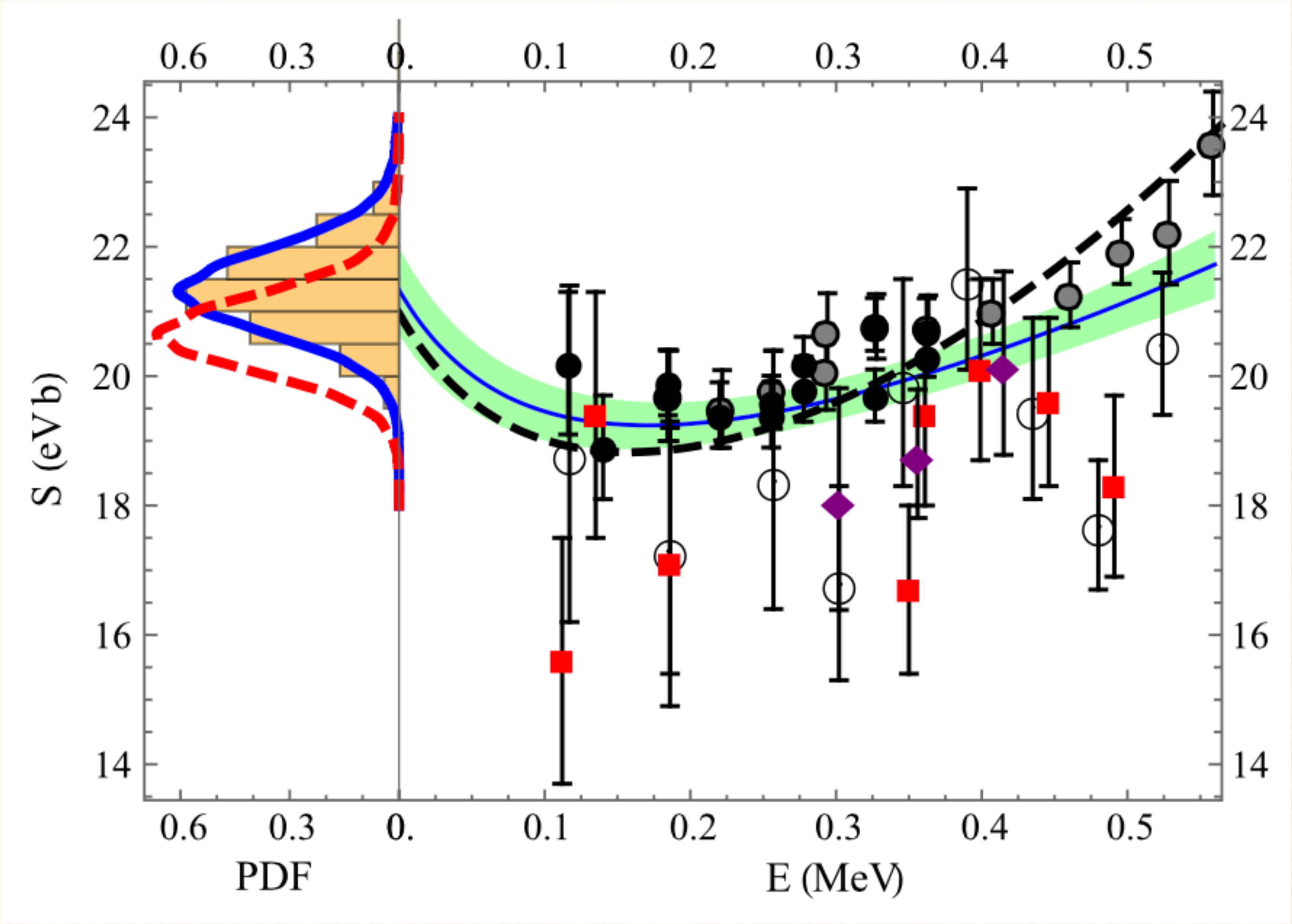
$$\mathrm{pr}\left(\bar{F}|D;T;I\right) = \int \mathrm{pr}\left(\vec{g},\{\xi_i\}|D;T;I\right)\delta(\bar{F} - F(\vec{g}))d\xi_1 \ldots d\xi_5 d\vec{g}$$



Final result

Zhang, Nollett, DP, PLB, 2015

$$\mathrm{pr}\left(\bar{F}|D;T;I\right) = \int \mathrm{pr}\left(\vec{g},\{\xi_i\}|D;T;I\right)\delta(\bar{F}-F(\vec{g}))d\xi_1\ldots d\xi_5 d\vec{g}$$

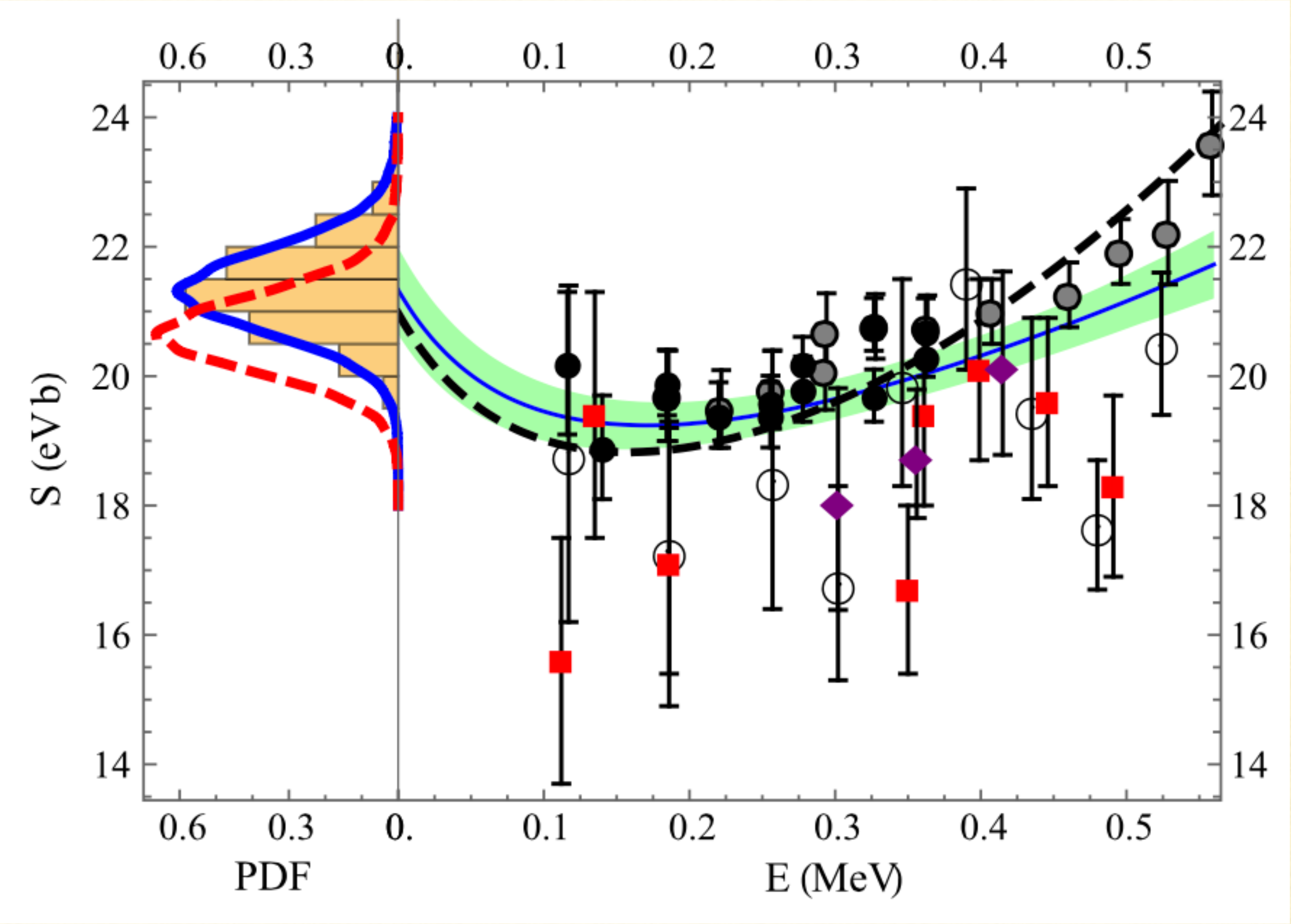


$$S(0) = 21.33^{+0.66}_{-0.69}\ \mathrm{eV\ b}$$

Final result

Zhang, Nollett, DP, PLB, 2015

$$\mathrm{pr}\left(\bar{F}|D;T;I\right) = \int \mathrm{pr}\left(\vec{g},\{\xi_i\}|D;T;I\right)\delta(\bar{F}-F(\vec{g}))d\xi_1\ldots d\xi_5 d\vec{g}$$

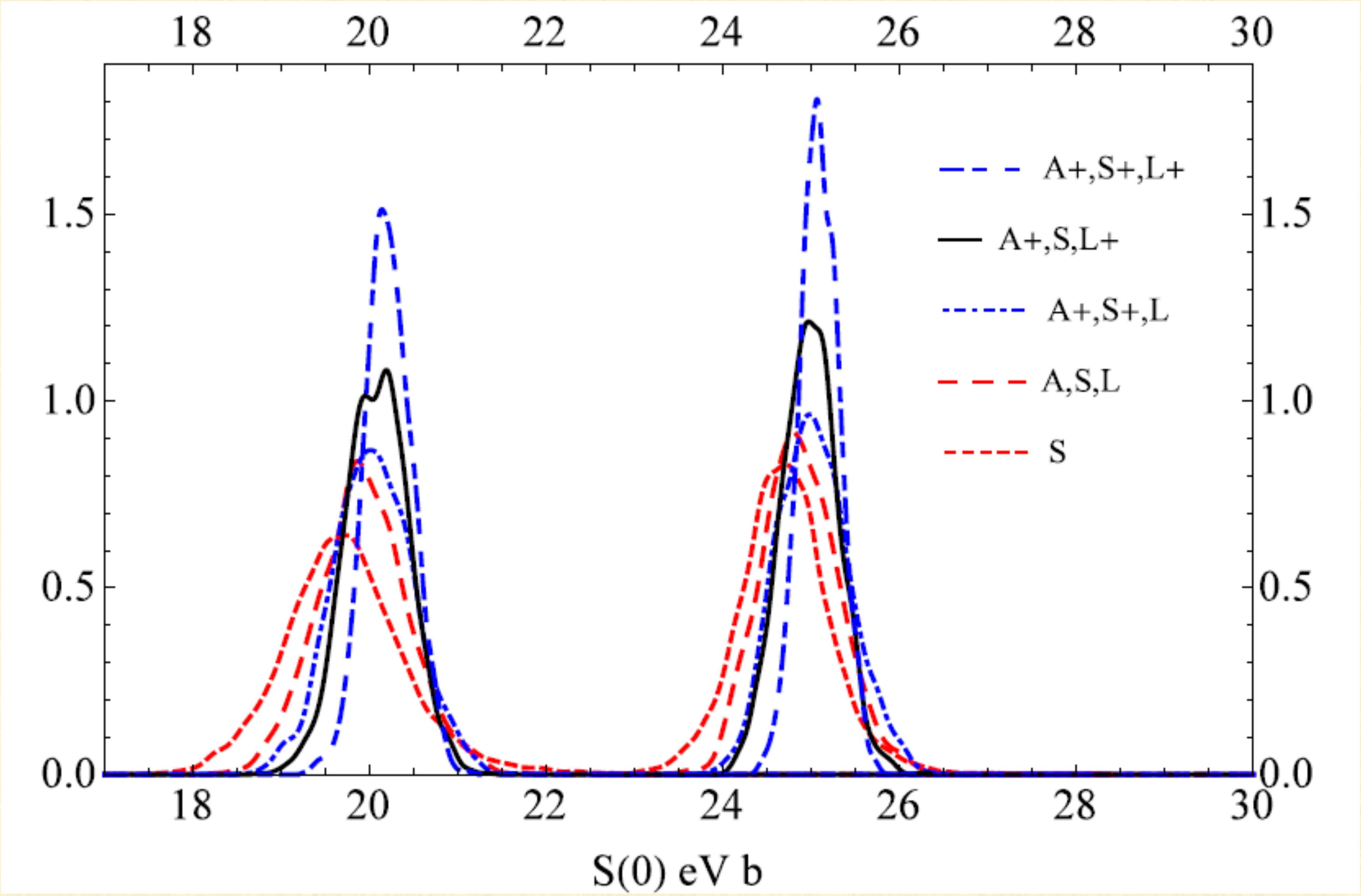


$$S(0) = 21.33^{+0.66}_{-0.69}\ \mathrm{eV\ b}$$

Uncertainty reduced by factor of two: model selection

Planning improvements

Use extrapolant to simulate impact of hypothetical future data that could inform posterior pdf for S(0)



Left-to-right:
42 data points all of similar quality to Junghans et al.

A: ANC
S: $a_{S=1}$ and $a_{S=2}$
L: short-distance

Note that 1 keV uncertainty in S_{1p} of ^{8}B may not be negligible effect

But this does not capture all the possibilities

But this does not capture all the possibilities

- Calculation now includes uncertainty in model parameters
- But should sample possible central values of A, S, L, and new data from distribution computed in model
- What priority should be given to:
  - more precise data;
  - improving A, S, and/or L?
- If one gets done (e.g. new experiment) then does that affect which thing one should go after next? Is the answer to this question dependent on the result of the experiment. (Sequential decision making.)
- Lots and lots of sampling!