研究会「原子核物理でつむぐrプロセス」
YITP, Kyoto U., May 22-24, 2019

Towards systematic and consistent nuclear data inputs for astrophysical *r*-process with Bayesian approaches

Haozhao LIANG（梁豪兆）

RIKEN Nishina Center, Japan
Graduate School of Science, the University of Tokyo, Japan

May 23, 2019

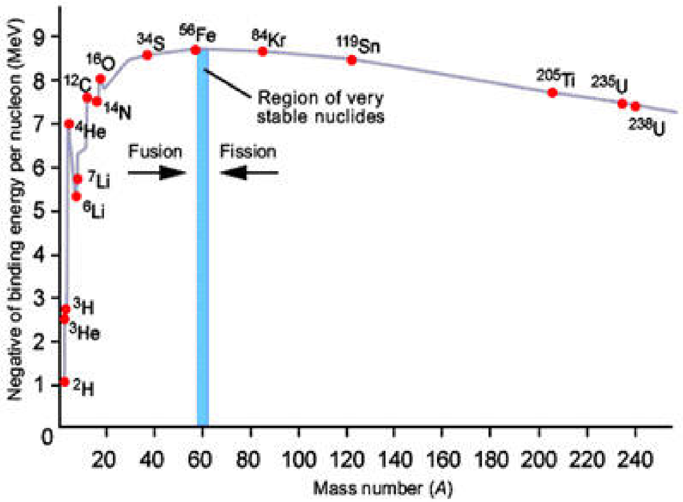
In collaboration with
Zhongming Niu (*Anhui U., China*) + ……

Contents

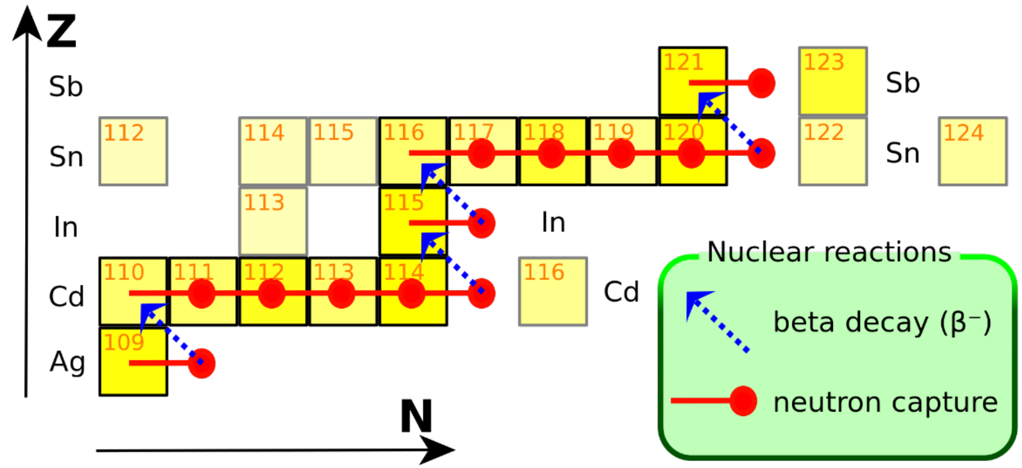
- **Introduction**
- **Nuclear inputs with Bayesian approaches**
  - A toy model
  - Nuclear Masses
  - Nuclear β-decay half-lives
- **Summary and Perspectives**

Nucleosynthesis

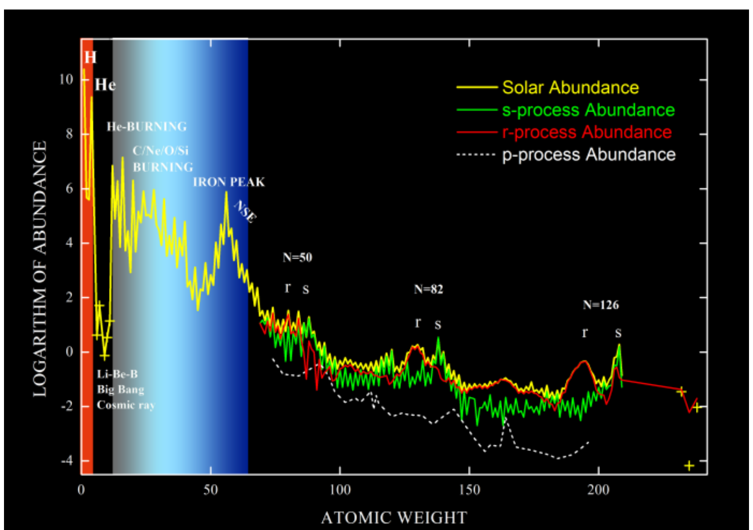
➢ Nuclear binding energy



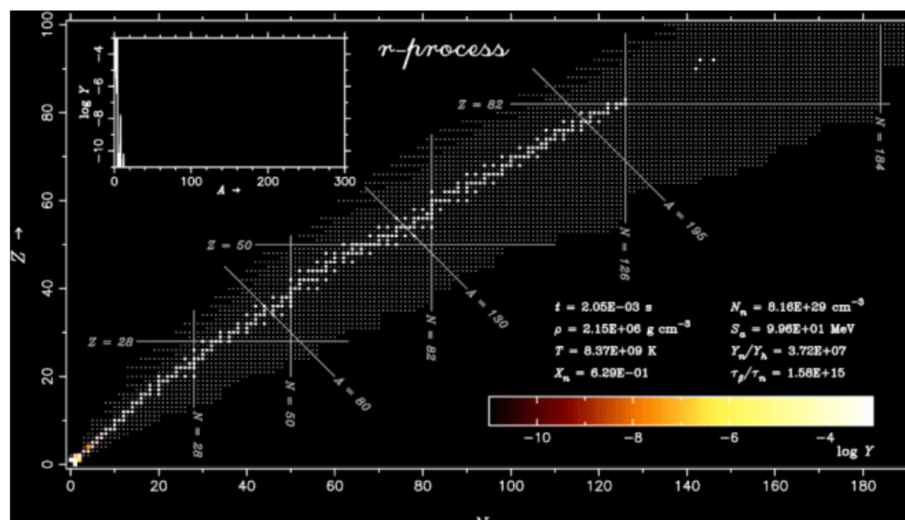
➢ Slow neutron-capture process (s-process)



➢ Abundance

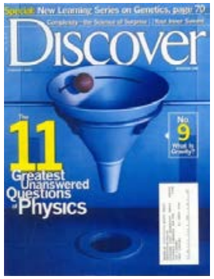


➢ Rapid neutron-capture process (r-process)

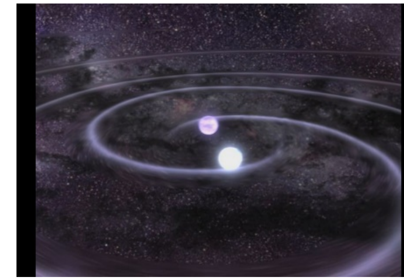


r-process nucleosynthesis and nuclear inputs

The 11 greatest unanswered questions of physics



Question 3
How were the heavy elements from iron to uranium made?



★ Nuclear data inputs for *r*-process

Quantity		Effects
S_n	neutron separation energy	path
$T_{1/2}$	β-decay half-lives	abundance pattern, time scale
P_n	β-delayed n-emission branchings	final abundance pattern
Y_i	fission (products and branchings)	endpoint, degree of fission cycling
		abundance pattern (?)
G	partition functions	path (very weakly)
$N_A \langle \sigma \nu \rangle$	neutron capture rates	conditions for waiting point approximation
		final abundance pattern during freezeout (?)

Nuclear inputs for r-process

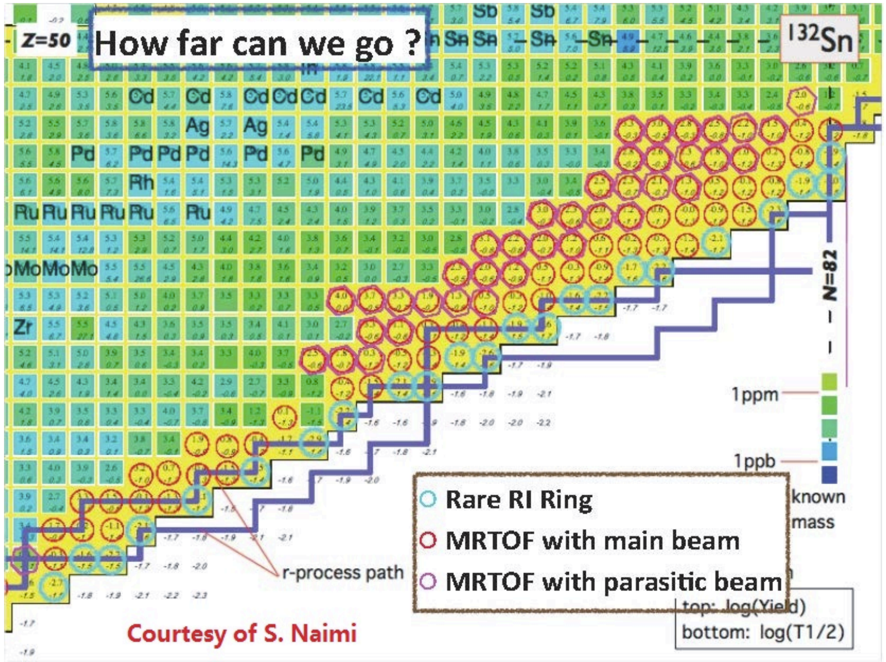
Key exp. @ RIKEN

masses

***β*-decay half-lives**

***β*-delayed n-emssions**

……



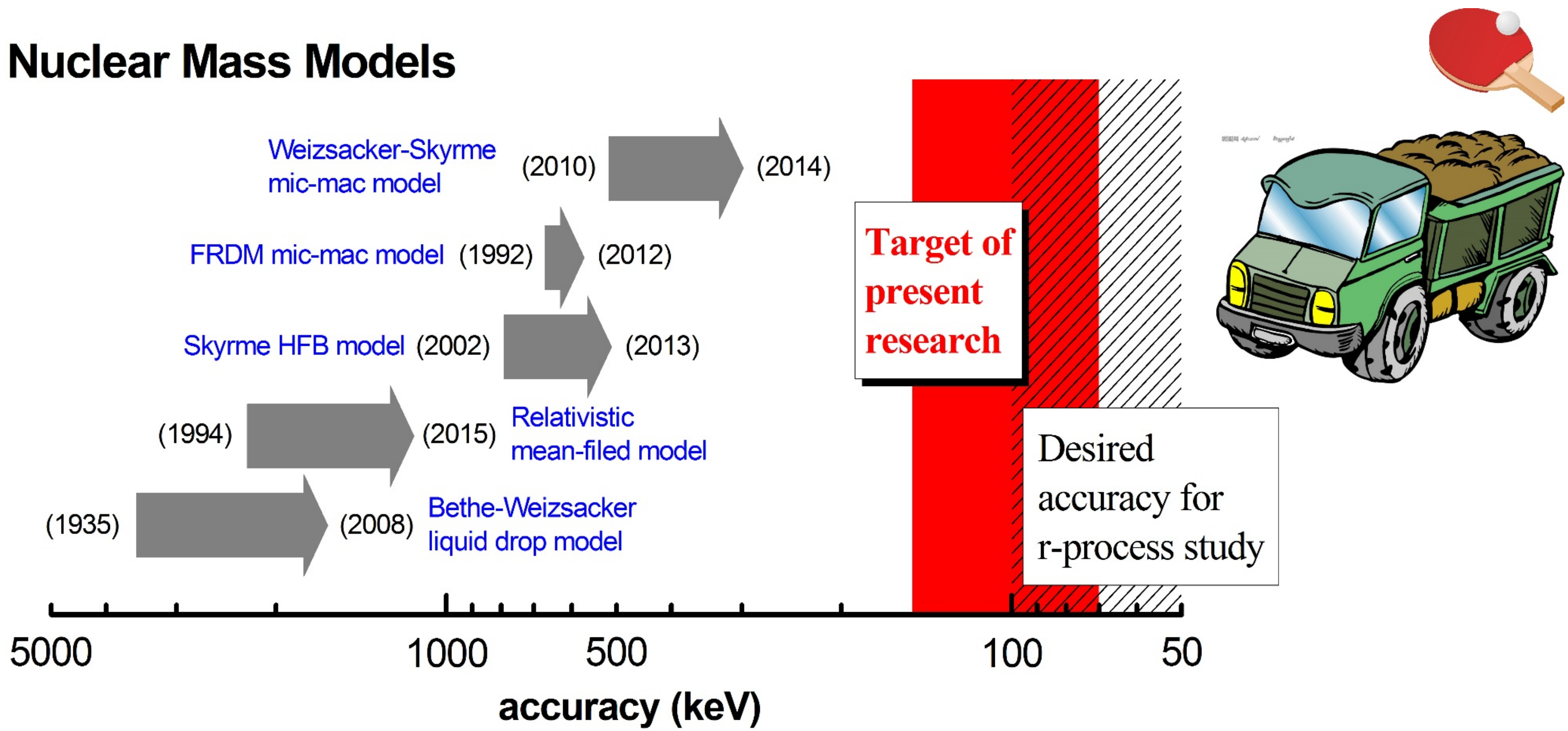
- To provide and organize **all these inputs** in a systematic and consistent way
  - e.g., changes in mass ➔ changes in half-lives, capture rates …

    (not hybrid databases !)
  - more exp. data ➔ more reliable extrapolation / smaller uncertainties

    (higher accuracy ?)

Nuclear mass models

- Theoretically, the development of nuclear mass model can be traced back to the early age of nuclear physics, known as Bethe-Weizsacker liquid drop model in 1935.
- To take into account the nuclear shell effects: the microscopic models and the microscopic-macroscopic (mic-mac) models.



Theories + Bayesian approaches (I)

Nuclear mass predictions based on Bayesian neural network approach with pairing and shell effects

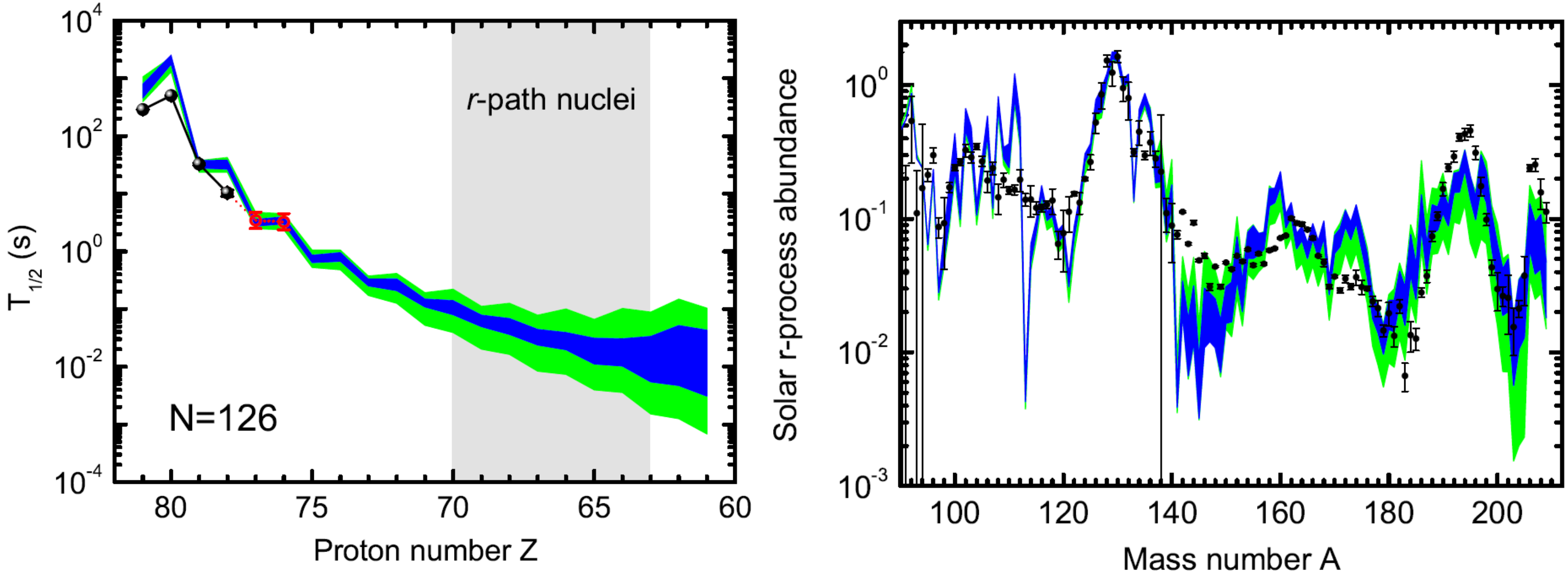
Z.M. Niu (牛中明)[a,b], H.Z. Liang (梁豪兆)[b,c,d,*]

Physics Letters B 778 (2018) 48–53

Predictions of nuclear β-decay half-lives with machine learning and their impacts on r process

Z. M. Niu,[1,2] H. Z. Liang,[3,4,*] B. H. Sun,[5] W. H. Long,[6] and Y. F. Niu[6,7]

arXiv:1810.03156



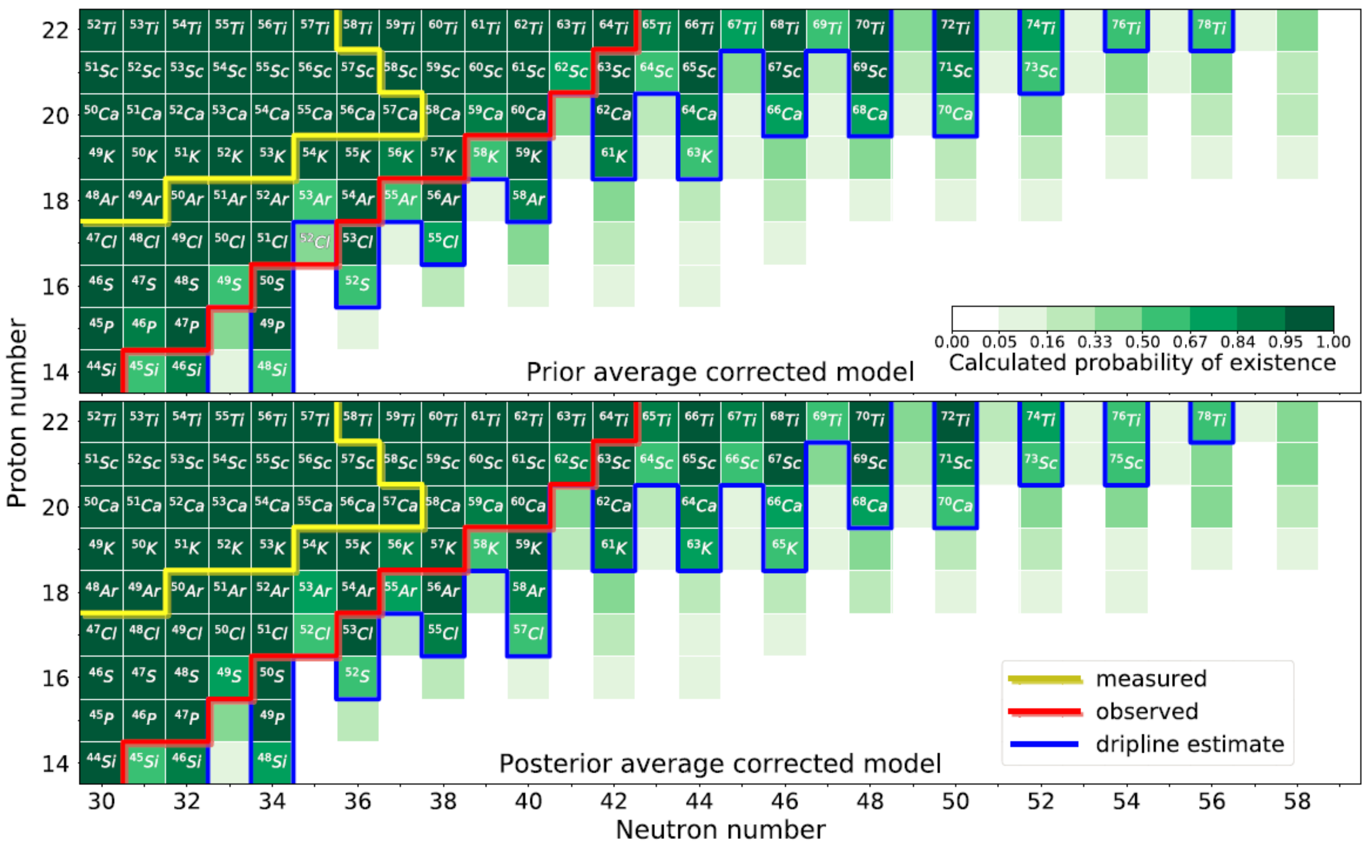
Another (ultimate) goal: to structure energy density functionals for DFT

Theories + Bayesian approaches (II)

PHYSICAL REVIEW LETTERS **122,** 062502 (2019)

Neutron Drip Line in the Ca Region from Bayesian Model Averaging

Léo Neufcourt,[1,2] Yuchen Cao (曹宇晨),[3] Witold Nazarewicz,[4] Erik Olsen,[2] and Frederi Viens[1]



Contents

- **Introduction**
- **Nuclear inputs with Bayesian approaches**
  - A toy model
  - Nuclear Masses
  - Nuclear β-decay half-lives
- **Summary and Perspectives**

Bayesian and frequentist views

- Differences between Bayesians and frequentists Bishop2006Springer

Frequentists:

- ✓ Data are a repeatable random sample --- there is a frequency
- ✓ Underlying parameters remain constant during this repeatable process
- ✓ **Parameters are unknown but fixed**

Bayesians:

- ✓ Data are observed from the realized sample
- ✓ **Parameters are unknown and described probabilistically**
- ✓ Data are fixed

➢ **Example: tossing a coin of unknown properties; probability ω of the coin landing heads**

- ✓ Choose some criterion, such as maximum likelihood
- ✓ Find the optimal estimator according to this criterion, such as the frequency of heads in past tosses

$$\omega = \frac{N^{\mathbf{head}}}{N^{\mathbf{total}}}$$

- ✓ Express this unknown properties using a probability distribution over possible values based on our intuitive believes
- ✓ Update this distribution using the **Bayes' theorem** as the outcome of each toss becomes known

$$p(\omega \mid D) = \frac{p(D \mid \omega)\, p(\omega)}{p(D)}$$

Bayesian approach in regression problem

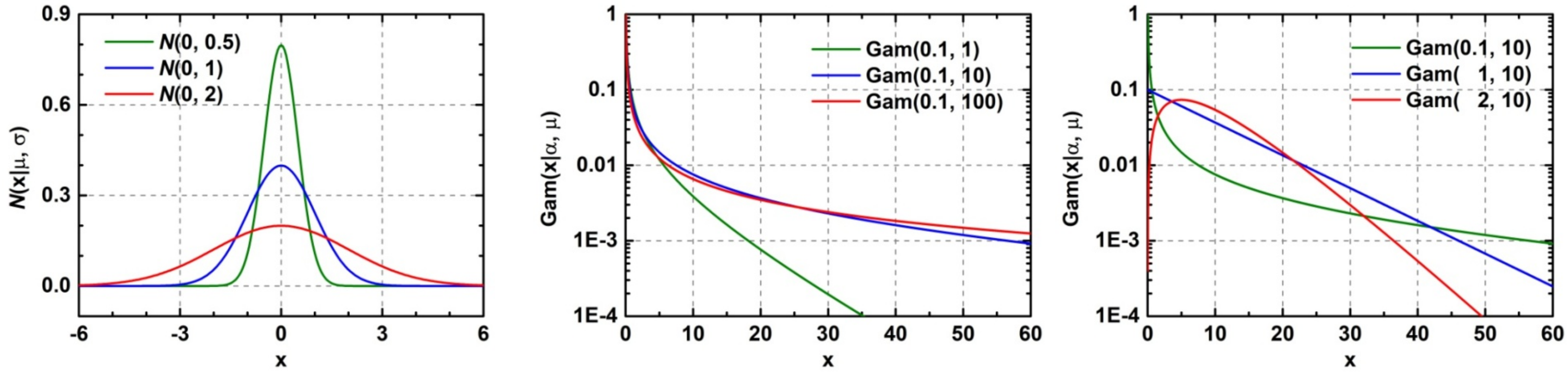
- Posterior distributions of parameters are Neal1996Springer

$$p(\omega|D)=\frac{p(D|\omega)p(\omega)}{p(D)}\propto p(D|\omega)p(\omega),\ \ D=\{(x_1,t_1),\ \ (x_2,t_2),\ \ \ldots,\ \ (x_N,t_N)\}$$

- prior distribution $p(\omega)$:

$$p(\omega)=N(\omega|0,\sigma_\omega),\ p(\tau_\omega=1/\sigma_\omega^2)=\mathrm{Gam}(\tau_\omega|\alpha_\omega,\mu_\omega)$$

$$p(\tau_n=1/\sigma_n^2)=\mathrm{Gam}(\tau_n|\alpha_n,\mu_n)$$



$$N(x|\mu,\sigma)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]\qquad \mathrm{Gam}(x|\alpha,\mu)=\frac{(\alpha/\mu)^\alpha}{\Gamma(\alpha)}x^{\alpha-1}\exp\left(-\frac{\alpha x}{\mu}\right)$$

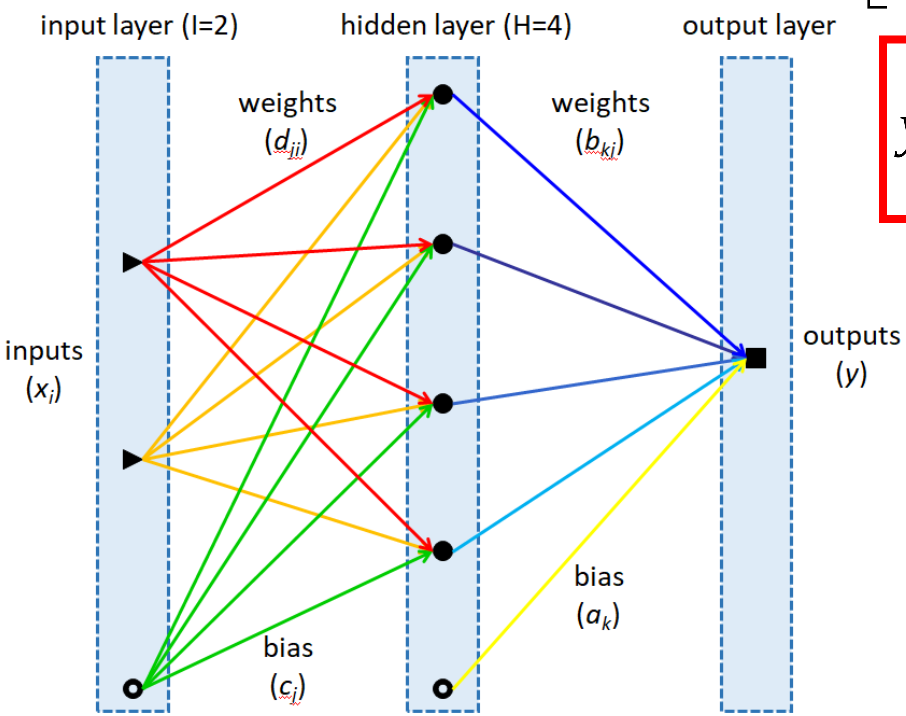
Bayesian approach in regression problem

- Posterior distributions of parameters are   Neal1996Springer

$$p(\omega|D)=\frac{p(D|\omega)p(\omega)}{p(D)}\propto p(D|\omega)p(\omega),\ \ D=\{(x_1,t_1),\ \ (x_2,t_2),\ \ \ldots,\ \ (x_N,t_N)\}$$

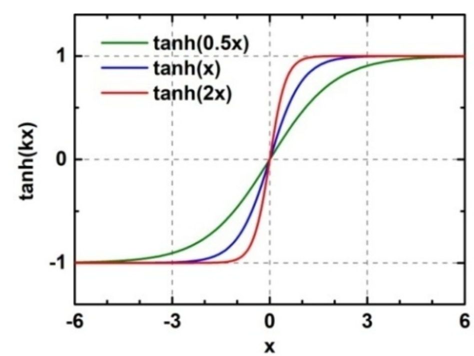
- likelihood function $p(D|\omega)$

$$p(x,t|\omega)=\exp(-\chi^2/2),\ \ \chi^2=\sum_{n=1}^{N}\left[\frac{t_n-y(x_n,\omega)}{\sigma_n}\right]^2$$



$$y(x,\omega)=a+\sum_{j=1}^{H}b_j\tanh\left(c_j+\sum_{i=1}^{I}d_{ji}x_i\right)$$

$$\tanh x=\frac{\sinh x}{\cosh x}=\frac{e^x-e^{-x}}{e^x-e^{-x}}$$



Bayesian approach in regression problem

- Posterior distributions of parameters are Neal1996Springer

$$p(\omega|D)=\frac{p(D|\omega)p(\omega)}{p(D)}\propto p(D|\omega)p(\omega),\quad D=\{(x_1,t_1),\ (x_2,t_2),\ \ldots,\ (x_N,t_N)\}$$

sampling with **Markov chain Monte Carlo (MCMC) method**

- Make predictions

$$\langle y_n\rangle=\int y(x_n,\omega)p(\omega|x,t)d\omega=\frac{1}{K}\sum_{k=1}^{K}y(x_n,\omega_k)$$

$$\Delta y_n=\sqrt{\langle y_n^2\rangle-\langle y_n\rangle^2}$$

- The BNN approach can give the **joint probability distribution of all parameters**, from which we can get the **correlations among parameters**, so the number of independent parameters may be much less the number of BNN parameters.

A toy model

True : $y = 0.3 + 0.4x + 0.5\sin(2x)$

Data : $y = 0.3 + 0.4x + 0.5\sin(2x) + 0.2 \times \text{randn}$

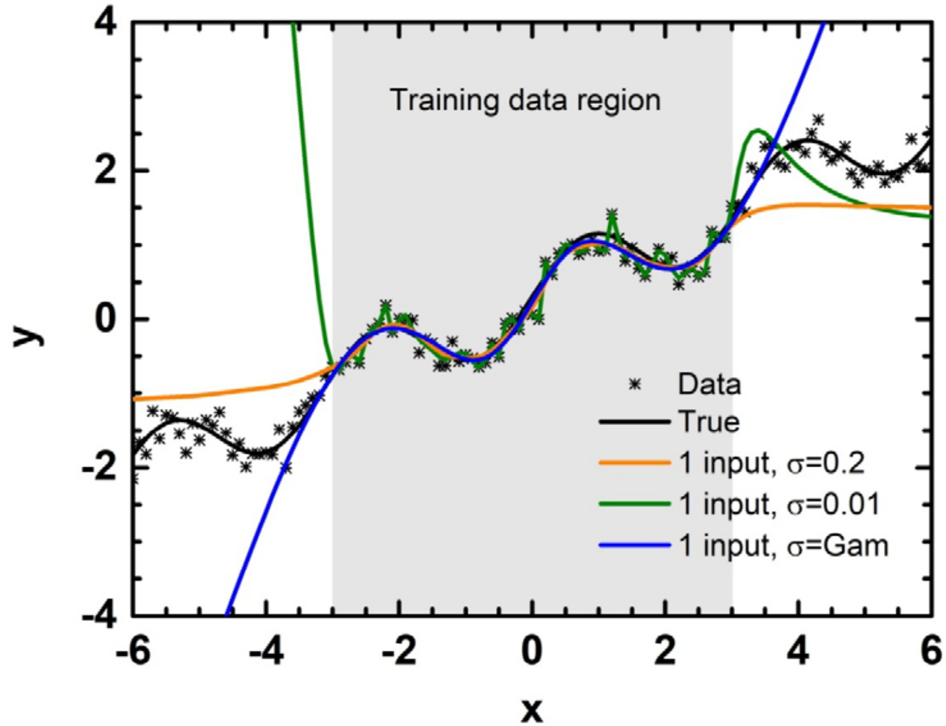
- Number of training data: N=61, x∈[-3, 3]

1 input : $y = f(x)$

2 inputs : $y = f[x_1 = x, x_2 = \sin(2x)]$

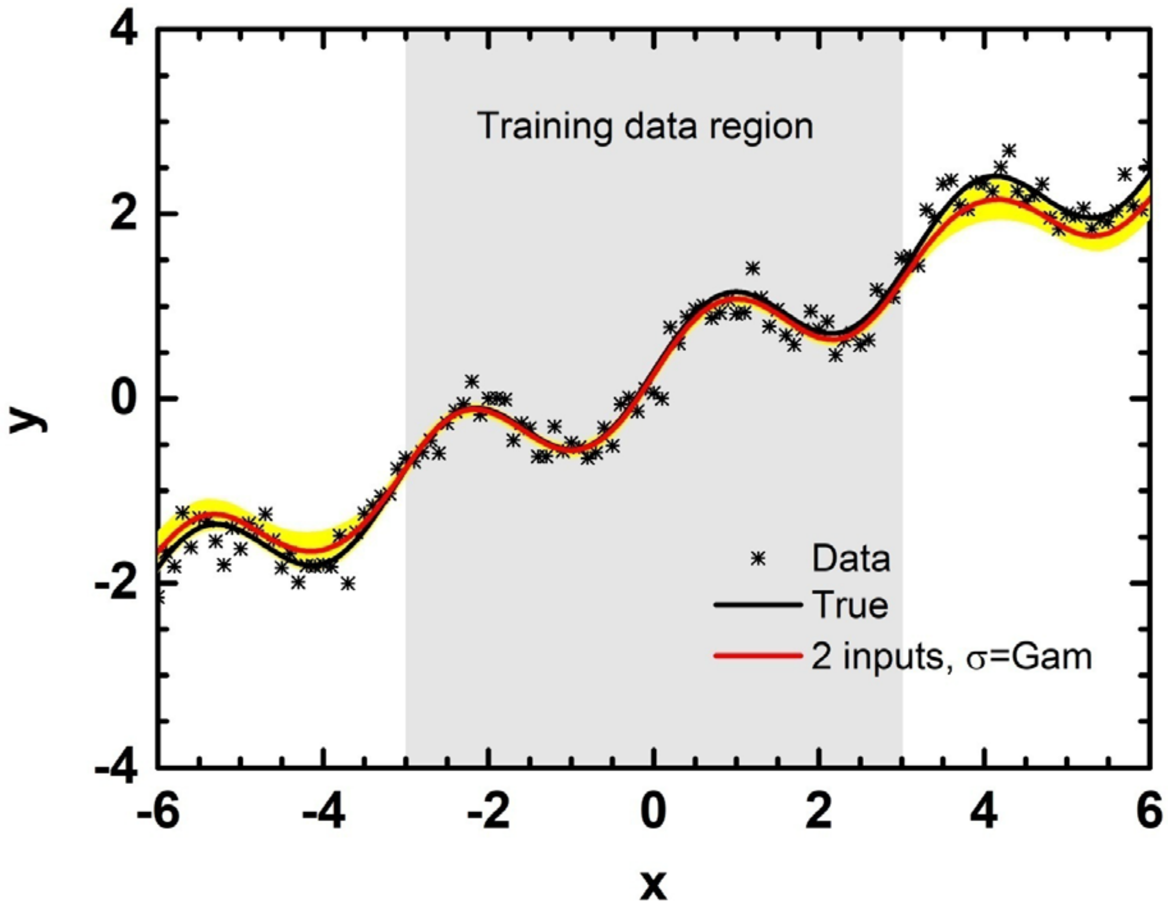
- Number of hidden unit:
  H=20 for $f(x)$; H=15 for $f(x_1, x_2)$
- Number of parameters: 61

$$\text{Likelihood function} : p(x, y \mid \omega) = \exp(-\chi^2/2),\ \chi^2 = \sum_{i=1}^{N}\left(\frac{y_i - f(x_i, \omega)}{\sigma}\right)^2$$



- BNN can **avoid overfitting** if a Gamma distribution is taken as the noise prior.
- Direct BNN fitting with x as the only input variable *can only extrapolate around a few steps* from known region.

A toy model



- **Including reasonable variable** is very effective for the extrapolation of neural network.
- **Uncertainties of predictions** are also reasonable.

Contents

- **Introduction**
- **Nuclear inputs with Bayesian approaches**
  - A toy model
  - Nuclear Masses
  - Nuclear β-decay half-lives
- **Summary and Perspectives**

Numerical details

Likelihood function $p(D|\omega)$

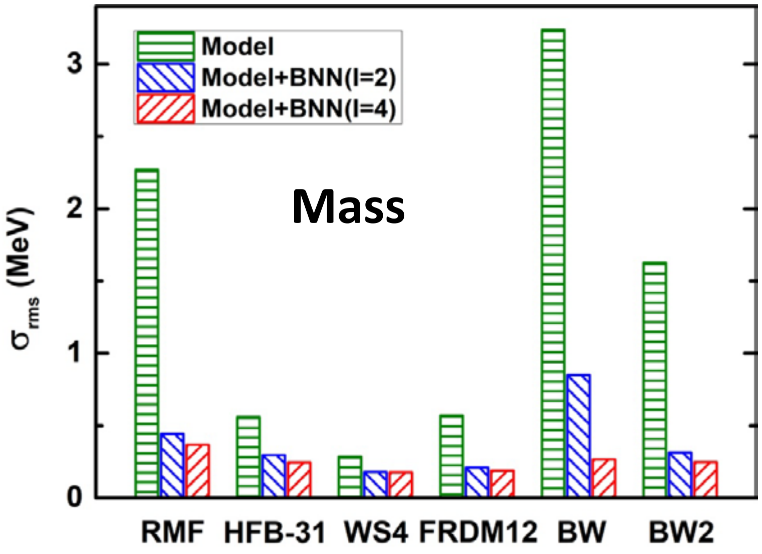
$$p(D\,|\,\omega)=\exp(-\chi^2/2),\ \chi^2=\sum_{n=1}^{N}\left[\frac{t_n-y_n(x,\omega)}{\sigma_n}\right]^2$$

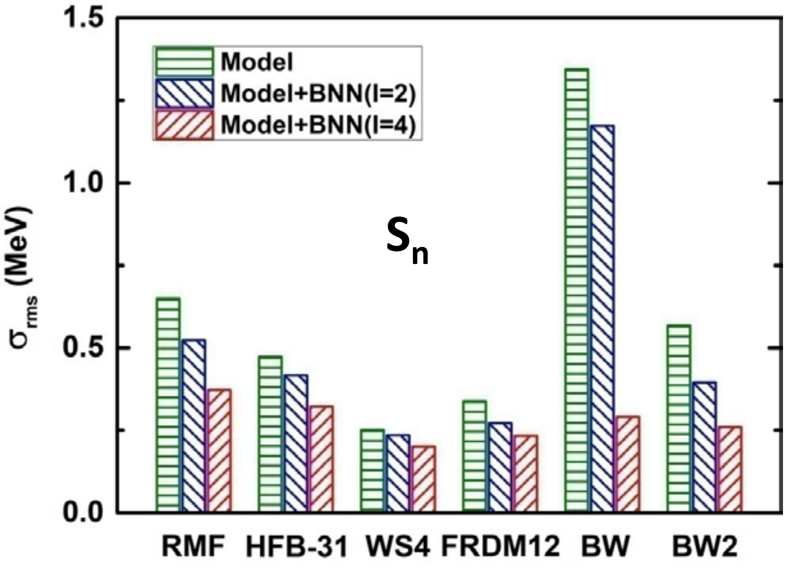
$$t_n=M_n^{\mathrm{exp}}-M_n^{\mathrm{th}},\ y(x,\omega)=a+\sum_{j=1}^{H}b_j\tanh\left(c_j+\sum_{i=1}^{I}d_{ji}x_i\right)\Rightarrow M_n'^{\mathrm{th}}=M_n^{\mathrm{th}}+y(x,\omega)$$

- ★ Inputs:
  - ✓ 2 inputs (I=2): Z, A
  - ✓ 4 inputs (I=4): Z, A, δ, P; $\delta=[(-1)^Z+(-1)^N]/2$, $P=\nu_n\nu_p/(\nu_p+\nu_n)$

    $$\nu_p=\min(|Z-Z_0|),\ \nu_n=\min(|N-N_0|)$$
- ★ Hidden units:
  - ✓ 2 inputs (I=2): H=42
  - ✓ 4 inputs (I=4): H=28
- ★ Number of parameters: 169
- ★ **Data:** Huang et al., CPC 41 030002; Wang et al., CPC 41 030003.
  - ✓ **Entire set**: 2272 nuclei in AME2016 (Z, N>=8 and σ^{exp}<=100 keV)
  - ✓ **Learning set**: 1800 data randomly selected from entire set
  - ✓ **Validation set**: the remaining 472 data in entire set

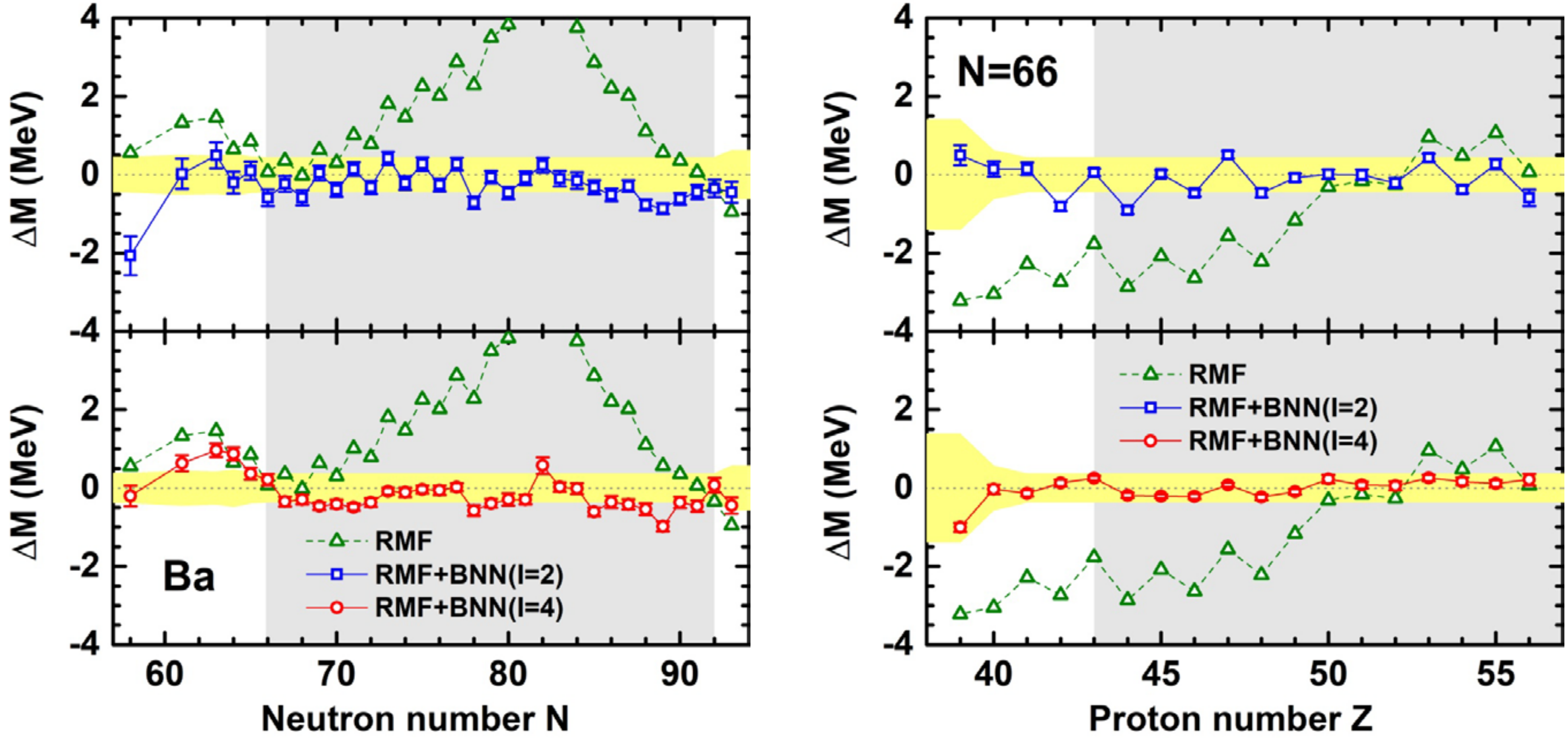
Rms deviations of mass and S_n





- The predictions of nuclear mass and neutron-separation energy are **significantly improved** with the BNN approach.
- After the improvement using the BNN approach with four inputs, the rms deviations are generally around **200 keV**.
- The BNN **with four inputs** is more powerful than the BNN with two inputs, especially for the neutron separation energy.

Mass extrapolation



- The **smooth deviations** can be **improved significantly**, while the **odd-even staggering** can only remarkably reduced with BNN-I4 approach.
- The BNN corrections are still reasonable if the extrapolation is *not far away from the training region*.

Contents

- **Introduction**
- **Nuclear inputs with Bayesian approaches**
  - A toy model
  - Nuclear Masses
  - Nuclear β-decay half-lives
- **Summary and Perspectives**

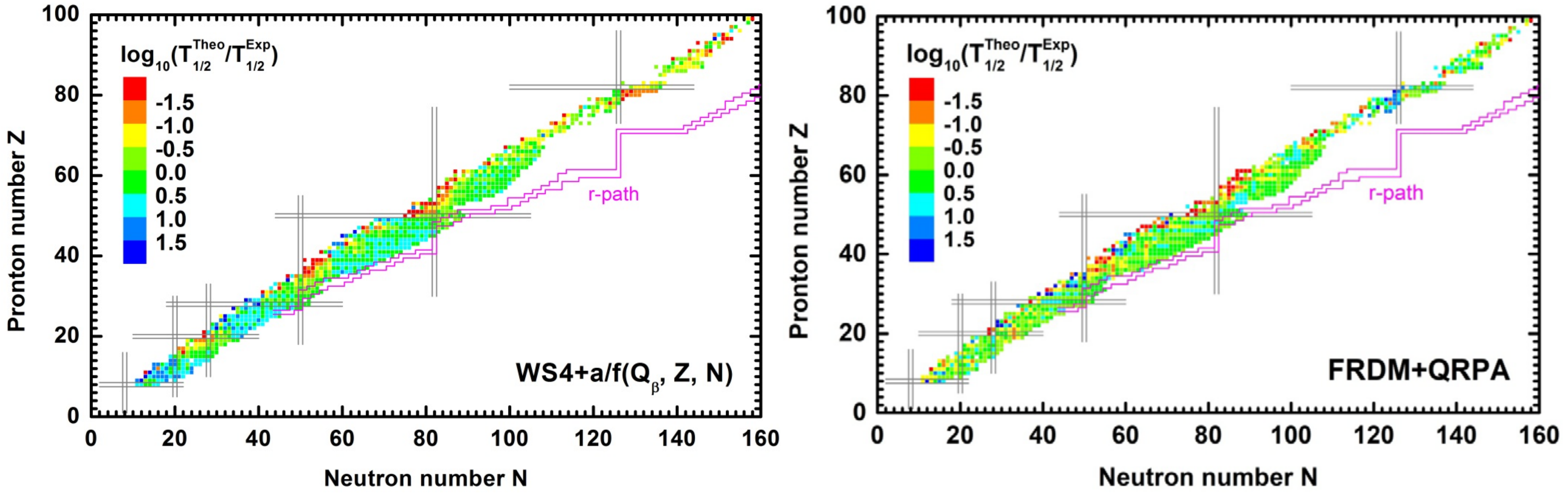
Nuclear β-decay half-lives

- The nuclear β-decay half-life in the allowed Gamow-Teller approximation reads as follows:

$$T_{1/2} = \frac{\ln 2}{\lambda_\beta} = \frac{D}{g_A^2 \sum_m B_{\rm GT}(E_m) f(Z,A,E_m)} \rightarrow T_{1/2} = a / f(Z,A,E_m{=}Q_\beta - c(\delta-1)/\sqrt{A})$$

where $D = \frac{\hbar^7 2\pi^3 \ln 2}{g^2 m_e^5 c^4} = 6163.4$ s, $g_A = 1$, $B_{\rm GT}(E_m)$ is the transition probability, and E_m is the maximum value of β-decay energy. The phase volume is

$$f(Z,A,E_m) = \frac{1}{m_e^5} \int_{m_e}^{E_m} p_e E_e (E_m - E_e)^2 F(Z,A,E_m) dE_e,$$



Numerical details

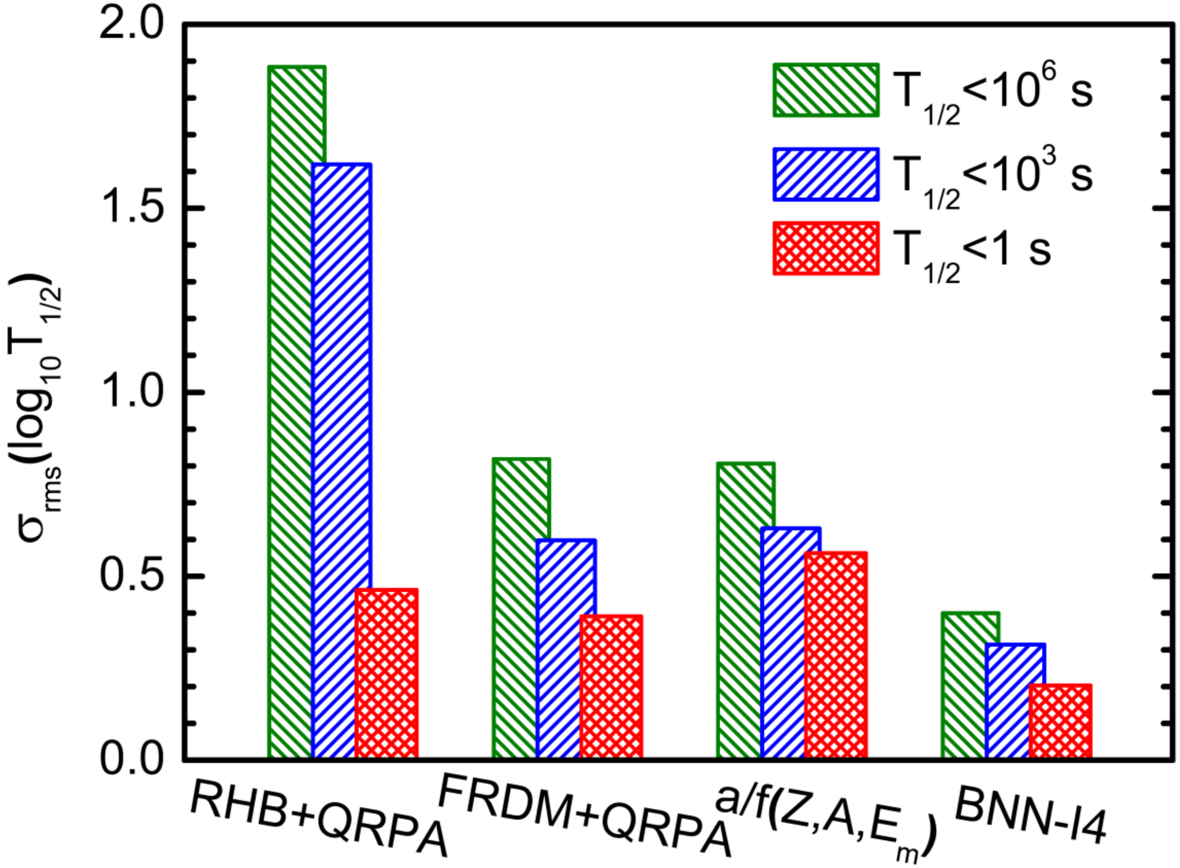
Likelihood function $p(D|\omega)$

$$p(D\,|\,\omega)=\exp(-\chi^2/2),\ \chi^2=\sum_{n=1}^{N}\left[\frac{t_n-y_n(x,\omega)}{\sigma_n}\right]^2$$

$$t_n=\log(T_n^{\exp}/T_n^{\rm th}),\ y(x,\omega)=a+\sum_{j=1}^{H}b_j\tanh\left(c_j+\sum_{i=1}^{I}d_{ji}x_i\right)\Rightarrow \log(T_n'^{\rm th})=\log(T_n^{\rm th})+y(x,\omega)$$

- Inputs:
  - ✓ 2 inputs (BNN-I2): *Z*, *N*
  - ✓ 4 inputs (BNN-I4): *Z*, *N*, $\delta=[(-1)^Z+(-1)^N]/2$, Q_β
- Hidden units:
  - ✓ 2 inputs (BNN-I2): *H*=30
  - ✓ 4 inputs (BNN-I4): *H*=20
- Number of parameters: 121
- **Data**: Audi et al., CPC 41, 030001 (2017)
  - ✓ **Entire set**: 1009 nuclei in NUBASE2016 (*Z*, *N*>=8 and β^--decay fraction=100%)
  - ✓ **Learning set**: 900 data randomly selected from entire set
  - ✓ **Validation set**: the remaining 109 data in entire set

Half-lives with BNN approaches

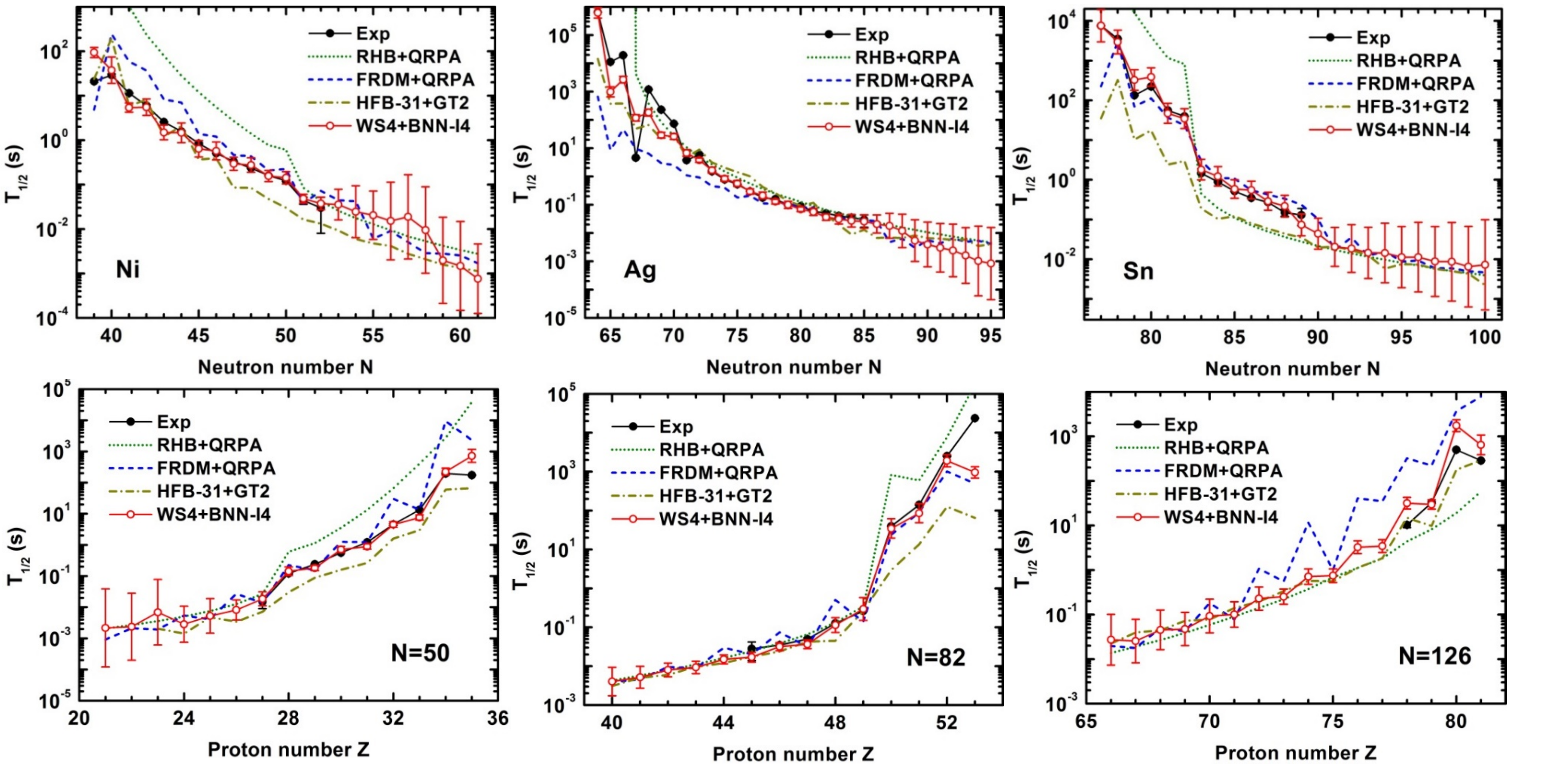


$$\sigma_{rms} = \sqrt{\frac{\sum_{i=1}^{n}\left[\log_{10}\left(T_{1/2}^{\rm Exp}/T_{1/2}^{\rm Theo}\right)\right]^2}{n}}$$

$10^{0.2} \sim 1.6$

★ Logarithmic rms deviations with respect to the known β-decay half-lives from NUBASE2016

Predictions of nuclear half-lives

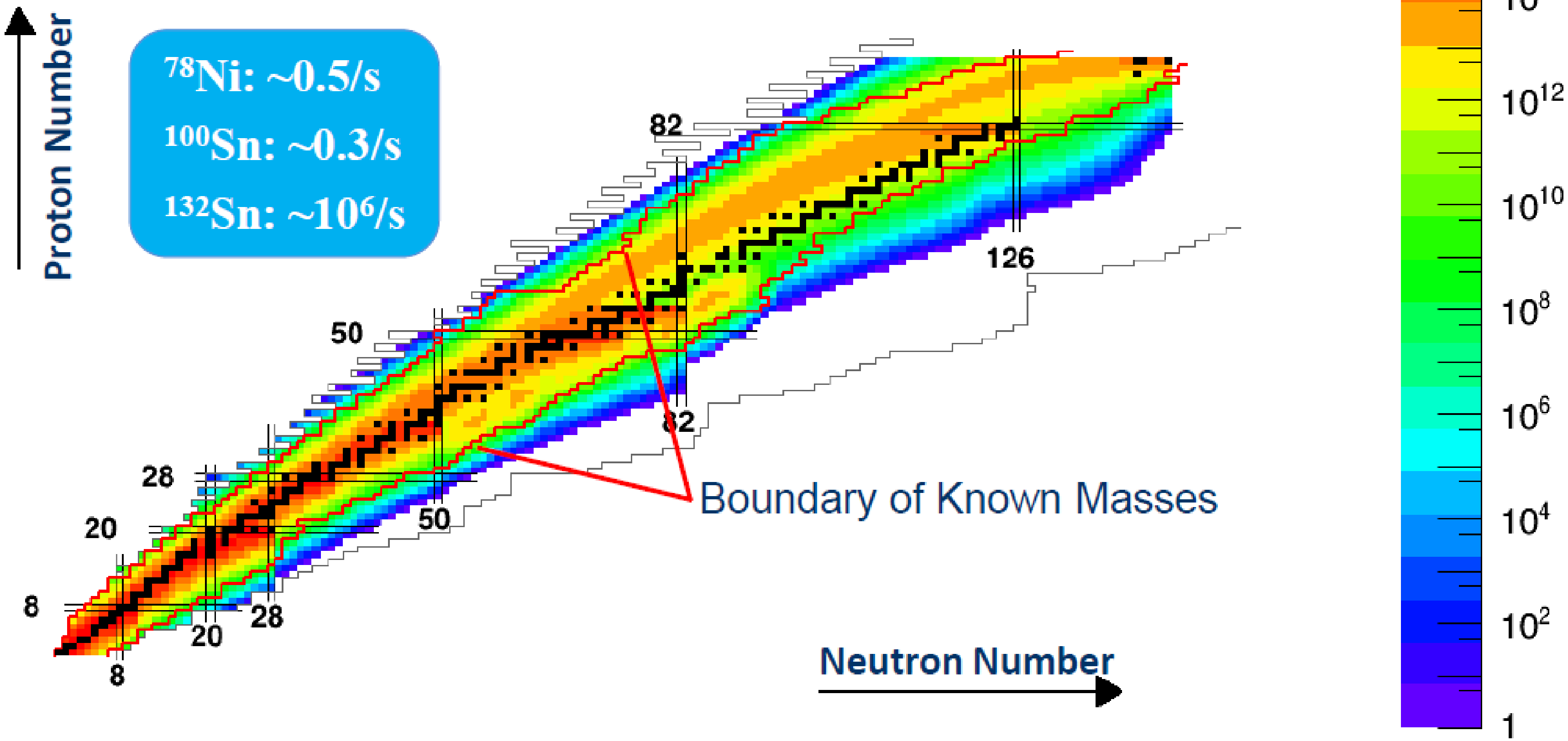


- The results of **WS4+BNN-I4** approach are **in a good agreement** with the experimental data, in particular, for those with **$T_{1/2}$ < 1 s**.
- Extrapolating to unknown region, **uncertainties** of WS+BNN-I4 **increase remarkably**.
- Results of other models generally agree with WS4+BNN-I4 **within uncertainties**.

Capability of Producing Nuclides

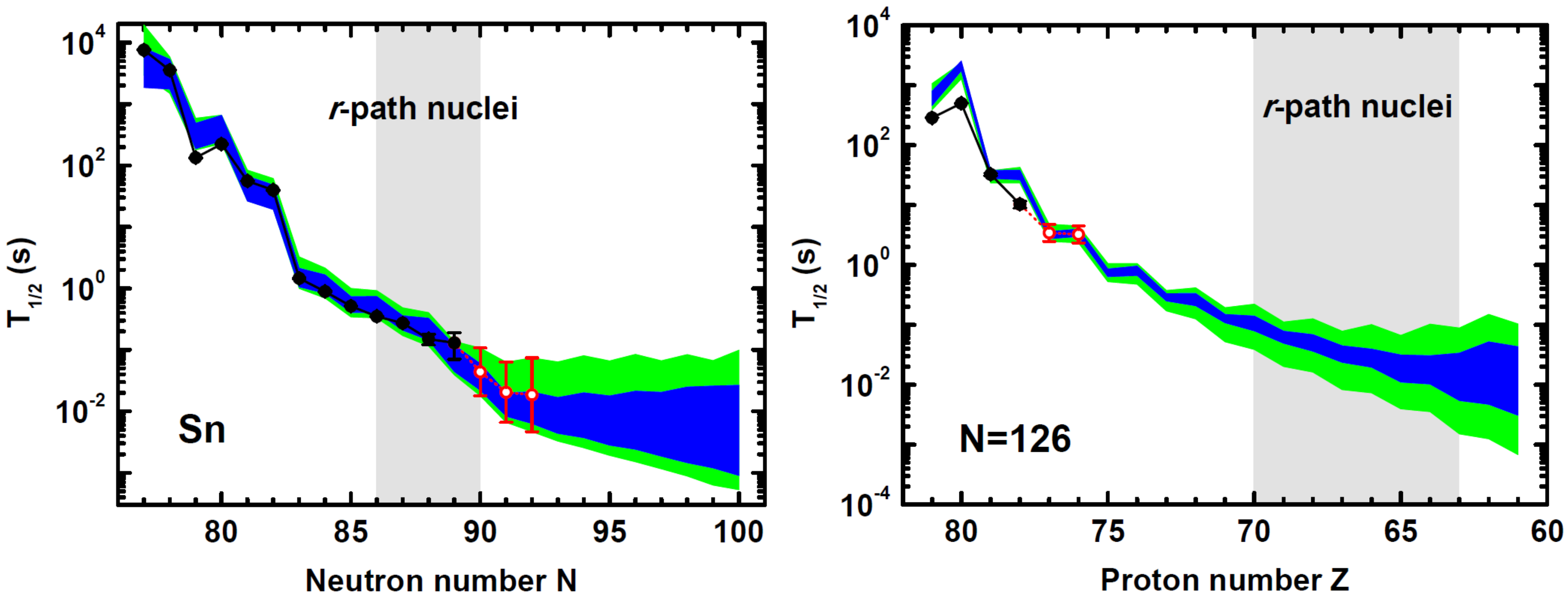
Nuclides Available (Production Yield) at HIAF

One of the world's most powerful facilities to explore the nuclear chart



Prolific sources of nuclides far away from the stability line will be provided using projectile fragmentation, in-flight fission, multi-nucleon transfer, and fusion reactions. The limits shown are the production rate of one nuclide per day, which enable the "discovery experiments"

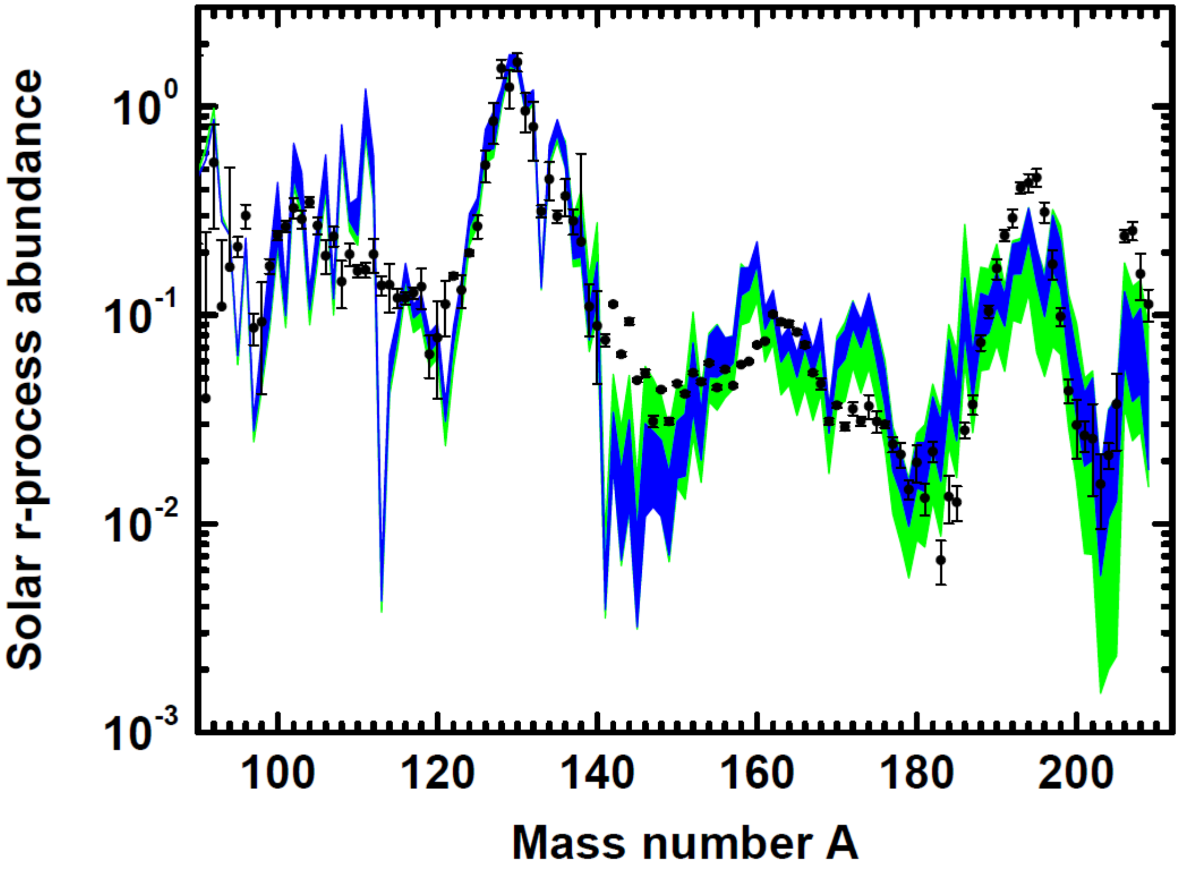
Predictions of nuclear half-lives



If we can further measure **three more β-decay half-lives** for each isotopes towards neutron-drip line

- the uncertainties of BNN predictions are similar in the training region
- they will **be decreased about 3 times** when extrapolate to the region far from known region.

Predictions of r-process abundances



- The uncertainties from nuclear beta-decay half-lives lead to large uncertainties for the r-process abundances of elements with ***A* > ~140**, which can **be remarkably reduced** if we can further measure three more β-decay half-lives.

Contents

- **Introduction**
- **Nuclear inputs with Bayesian approaches**
  - A toy model
  - Nuclear Masses
  - Nuclear β-decay half-lives
- **Summary and Perspectives**

Summary and Perspectives

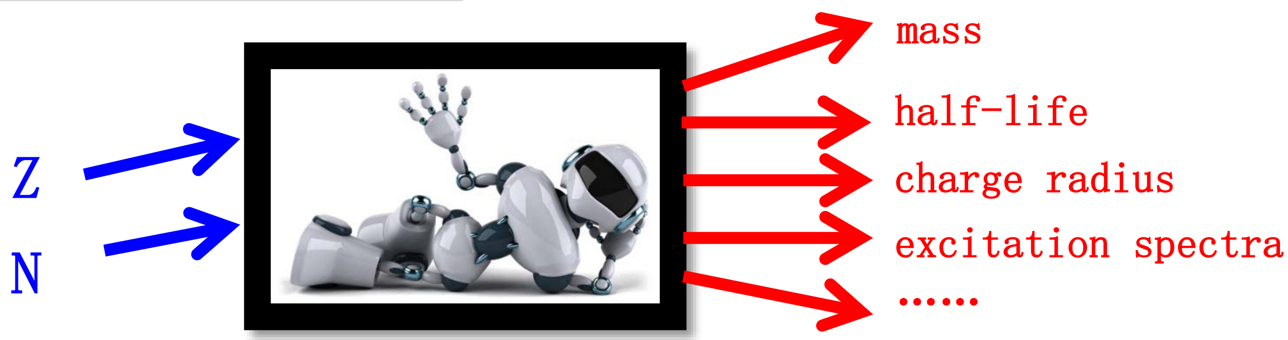
- To provide and organize **all these inputs** in a systematic and consistent way
  - e.g., changes in mass ➔ changes in half-lives, capture rates …

    (not hybrid databases !)
  - more exp. data ➔ more reliable extrapolation / smaller uncertainties

    (higher accuracy ?)

★ Nuclear data inputs for *r*-process

Quantity		Effects
S_n	neutron separation energy	path
$T_{1/2}$	β-decay half-lives	abundance pattern, time scale
P_n	β-delayed n-emission branchings	final abundance pattern
Y_i	fission (products and branchings)	endpoint, degree of fission cycling
		abundance pattern (?)
G	partition functions	path (very weakly)
$N_A \langle \sigma \nu \rangle$	neutron capture rates	conditions for waiting point approximation
		final abundance pattern during freezeout (?)



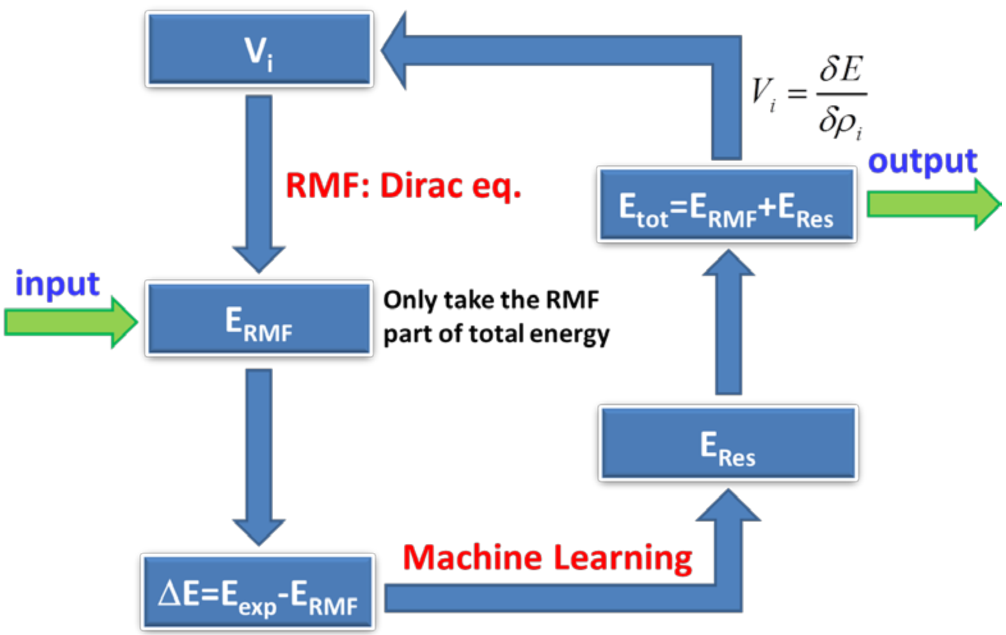
Summary and Perspectives

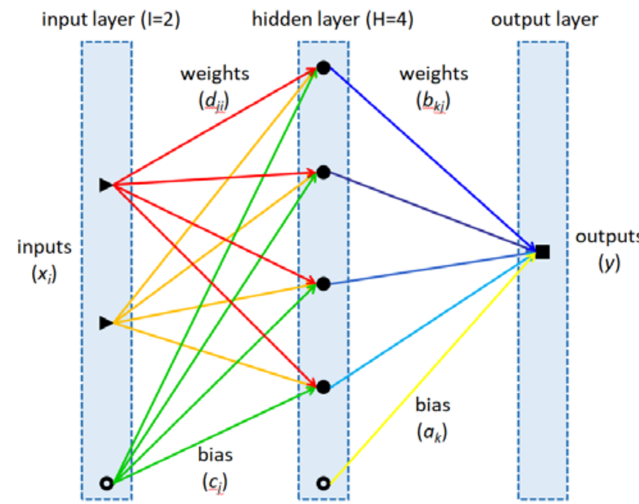
☐ To construct density functional with BNN approach?

$$E_{\text{tot}} = E_{\text{RMF}} + E_{\text{res}}$$

$$E_{\text{RMF}} = \sum_{k=1}^{A} \varepsilon_k - \int \left[\frac{1}{2}\alpha_S \rho_S^2 + \frac{1}{2}\delta_S \rho_S \Delta\rho_S + \frac{1}{2}\alpha_V \rho_V^2 + \frac{1}{2}\delta_V \rho_V \Delta\rho_V + \ldots \right] d\vec{r}$$

$$E_{\text{res}} = a + \sum_{j=1}^{H} b_j \int \tanh\left(c_j + d_{j1}\rho_S + d_{j2}\Delta\rho_S + d_{j3}\rho_V + d_{j4}\Delta\rho_V + \ldots\right) d\vec{r}$$





$$y(x,\omega) = a + \sum_{j=1}^{H} b_j \tanh\left(c_j + \sum_{i=1}^{I} d_{ji} x_i \right)$$

Thank you