2020绵阳原子核结构理论研讨会
绵阳, 2020.01.11-14

Predictions of nuclear masses and half-lives with Bayesian neural network approach

牛中明 (Zhongming Niu)

School of Physics and Materials Science, Anhui University, China

11 January 2020


内蒙古自治区
吉林
海参崴
Владивосток
辽宁
北京
天津
河北
朝鲜
平壤
평양
日本海
甘肃
宁夏回族自治区
山西
陕西
山东
黄海
韩国
釜山
부산
中国
西安
大阪
広島
福岡
河南
江苏
安徽大学
上海
安徽
湖北
成都
杭州
东海
四川
重庆
浙江
贵州
湖南
江西
福建
台北市
台湾
广西壮族自治区
广东
广州
高雄
香港
河内
Google

安徽大學 粒子与核物理研究组

RESEARCH GROUP OF PARTICLE AND NUCLEAR PHYSICS

- ✓ 国家“双一流”建设高校
- ✓ 国家“211工程”建设高校
- ✓ 安徽省人民政府与教育部共建高校
- ✓ 安徽省属重点综合性大学

安徽大學 粒子与核物理研究组
RESEARCH GROUP OF PARTICLE AND NUCLEAR PHYSICS

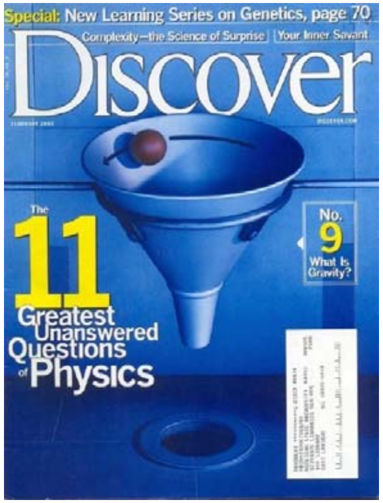
Outline

1. Introduction
2. Bayesian neural network approach
3. Results and discussion
   - ⋆ Toy model
   - ⋆ Nuclear masses
   - ⋆ Nuclear beta-decay half-lives
4. Summary and perspectives

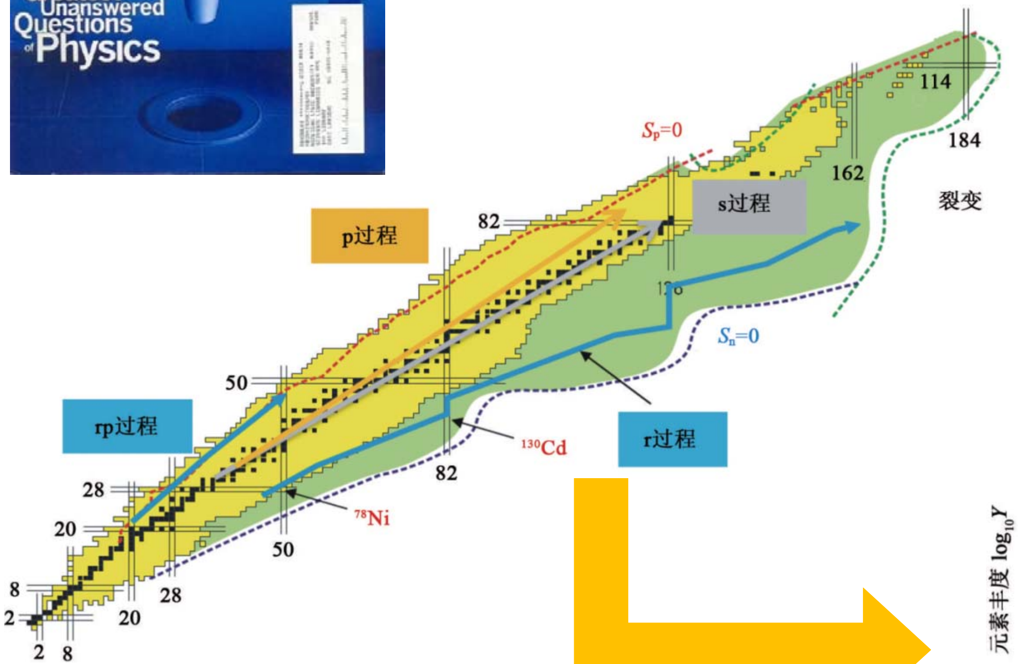
Outline

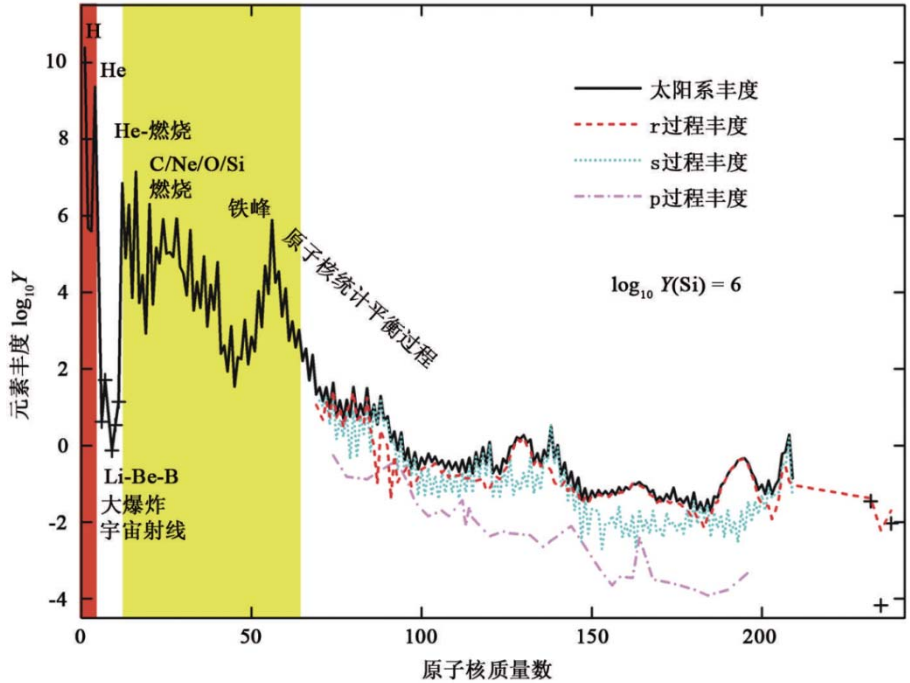
1. Introduction
2. Bayesian neural network approach
3. Results and discussion
   - ★ Toy model
   - ★ Nuclear masses
   - ★ Nuclear beta-decay half-lives
4. Summary and perspectives

Origin of heavy elements



Question 3:
How were the heavy elements from iron to uranium made? E. Haseltin, Discover 23(2), 37 (2002)





 Introduction

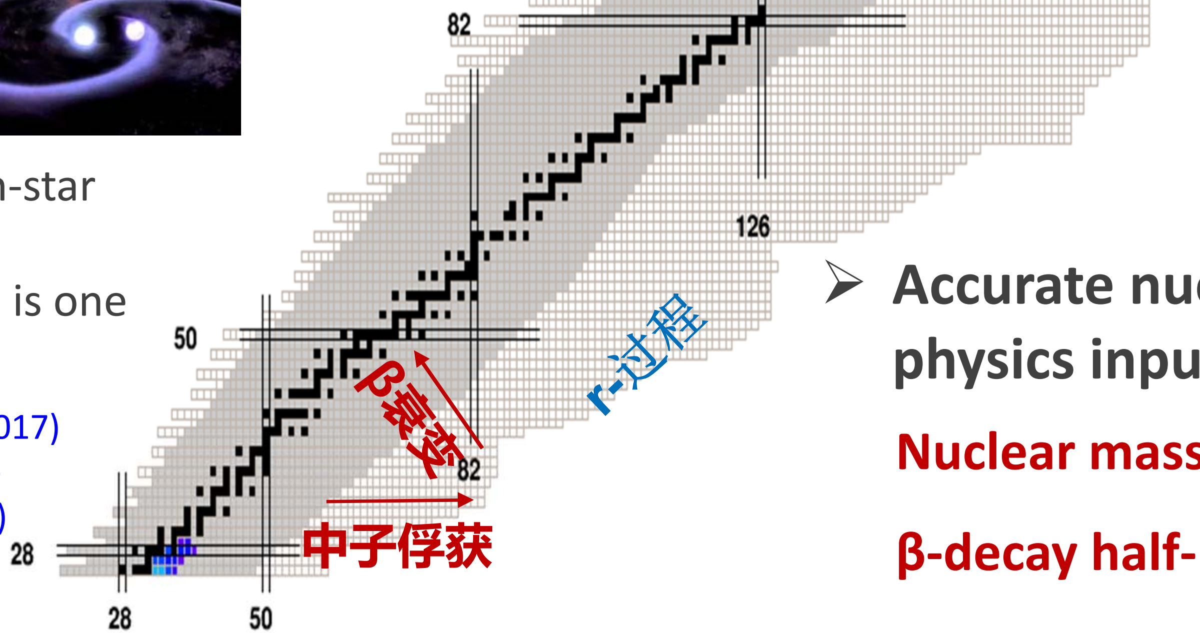
r-process

- ***r*-process site？**

Supernova　Neutron-star merge

GW170817 neutron-star merge：
neutron-star merge is one of *r*-process site

Nature **551**,64; 67; 80 (2017)
Science **358**, 1559 (2017)
ApJL **848**, L17; L19 (2017)



- **Accurate nuclear physics inputs？**

Nuclear masses

β-decay half-lives

Neutron-capture rates

…

Majority of neutron-rich nuclei related to the r-process are still out of the reach of experimental capabilities, so theoretical predictions have to be used.

Nuclear mass models

Nuclear mass models

★ Macroscopic mass models (simple, but large deviations and bad ability of extrapolation): BW, BW2 [Weizsäcker1935ZP, Bethe1937RMP, Kirson2008NPA]

★ Macro-microscopic mass models (high accuracy, but belong to hybrid model): FRDM, KTUY, WS4 [Moller1995ADNDT, Koura2005PTP, Wang2014PLB]

★ Microscopic mass models, mainly the density functional theory (complicated, but have a better ability of extrapolation): Skyrme HFB (from HF-BCS, HFB-1 to HFB-32), RMF [Goriely2016PRC, Geng2005PTP]

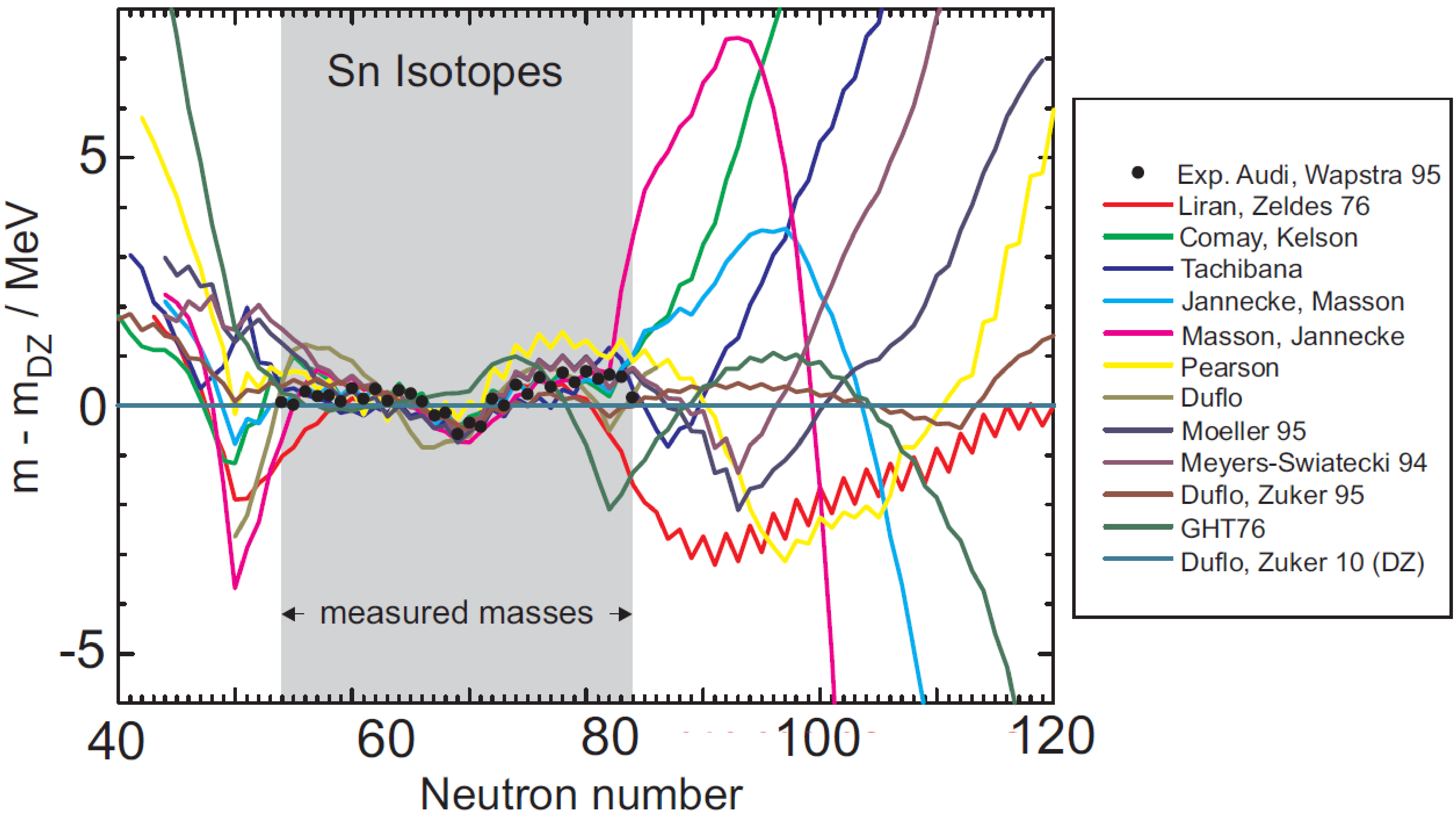
Model	BW	BW2	FRDM	KTUY	WS4	HFB-27	RMF
σ_{rms} (keV)	3296	1657	654	701	298	512	2217

data taken from AME12 [Wang2012CPC]

✓ This accuracy is still not enough to the studies of exotic nuclear structures, astrophysics nucleosynthesis, and weak-interaction processes.

✓ Different mass models giving comparable accuracy can extrapolate quite differently out to the neutron drip line.

Nuclear mass models



From Sun2008PhDThesis

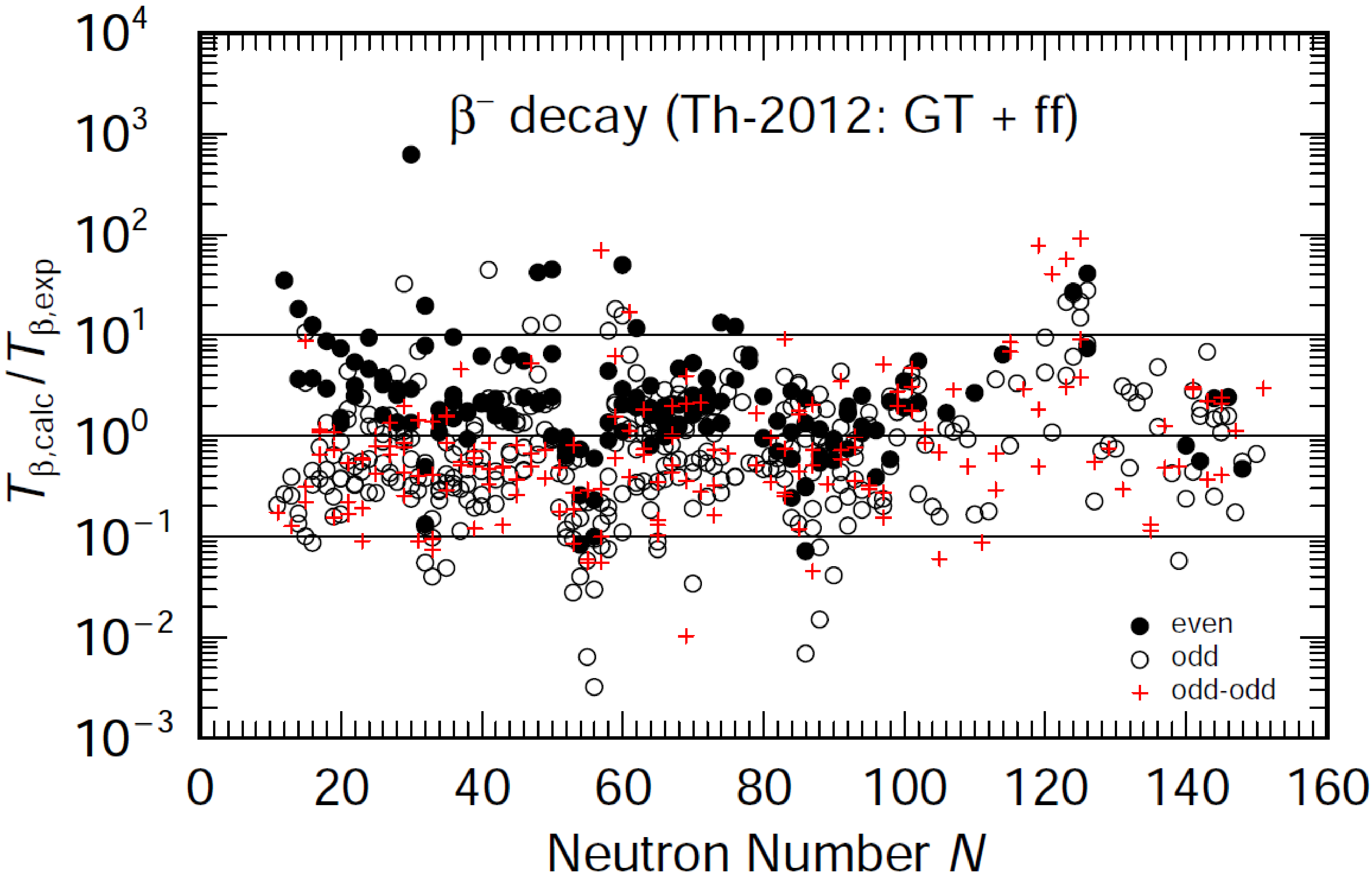
Nuclear models for β-decay half-lives

Nuclear models for β-decay half-lives

- ★ Phenomenological formula Zhang2006PRC, Zhang2007JPG
- ★ Gross theory Takahashi1973ADNDT, Takahashi1990PTP, Nakata1997NPA
- ★ Shell model Pinedo1999PRL, Caurier2002PRC, Langanke2003RMP, Zhi2013PRC
- ★ Quasiparticle random phase approximation (QRPA)

Nilsson BCS+QRPA:	Staudt1990ADNDT, Hirsch1993ADNDT, Nabi1999ADNDT
FRDM+QRPA:	Möller1997ADNDT, Möller2003PRC
Woods-Saxon+QRPA:	Ni2012JPG
SHF BCS+QRPA:	Sarriguren2005, 2010, 2011PRC
DF(Fayans)+CQRPA:	Borzov1996ZPA, Borzov2003,2005PRC, Borzov2008NPA
ETFSI(Skyrme)+CQRPA:	Borzov1997NPA, Borzov2000PRC
SHF(BCS)+(Q)RPA:	Bai2010PRL, Minato2013PRL
SHFB+QRPA:	Engel1999PRC
RHB+QRPA:	Nikšić2005PRC, Marketin2007,2016PRC, Niu2013PRC(R)
RHFB+QRPA:	Niu2013PLB

Nuclear models for β-decay half-lives



From Moller2018ADNDT

Introduction

- Some techniques have been developed to further improve the mass predictions of these nuclear models, such as the RBF approach [Wang2010PRC, Niu2013PRC, Niu2016PRC] and the image reconstruction technique [Morales2010PRC].
- Neural network approach have also been used to predict masses using the Bayesian version (BNN) [Utama2016PRC] or conventional version [Gazula1992NPA, Zhang2017JPG].
- Advantages of Bayesian approach:

  ⋆ avoids the over-fitting automatically by using hyper prior for the noise variance
  ⋆ quantifies the uncertainties in predictions

Apply the BNN approach to

✓ improve the mass predictions: 1) including hyper prior for the noise variance; 2) two extra variables related to pairing correlations and shell effects apart from Z and A.

✓ improve the predictions of nuclear beta-decay half-lives.

Outline

1. Introduction
2. Bayesian neural network approach
3. Results and discussion
   - ★ Toy model
   - ★ Nuclear masses
   - ★ Nuclear beta-decay half-lives
4. Summary and perspectives

Bayesian and frequentist (traditional) 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

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

✓ 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

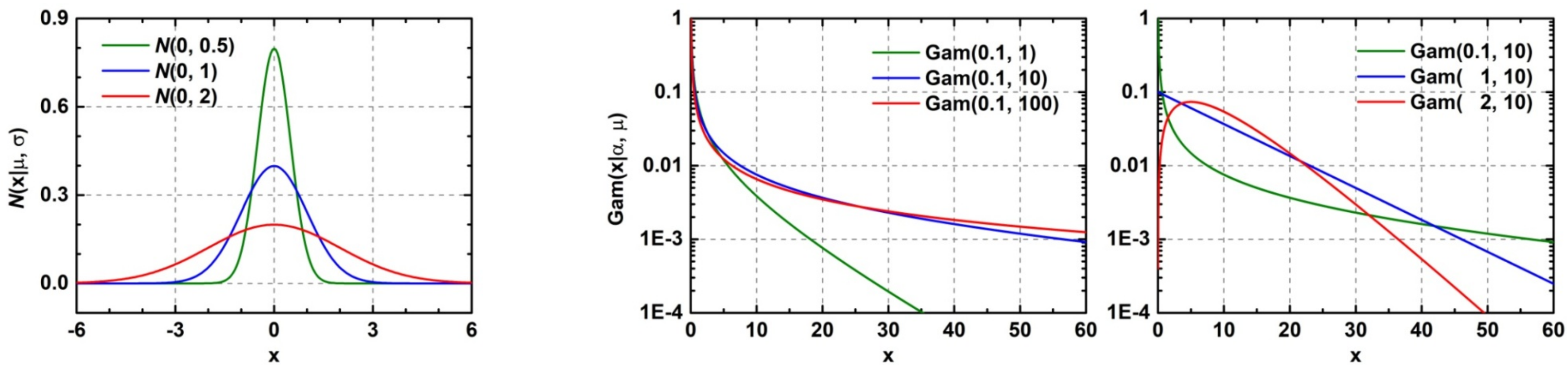
- Posterior distributions of parameters are Neal1996Springer

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

✓ 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 neural network approach

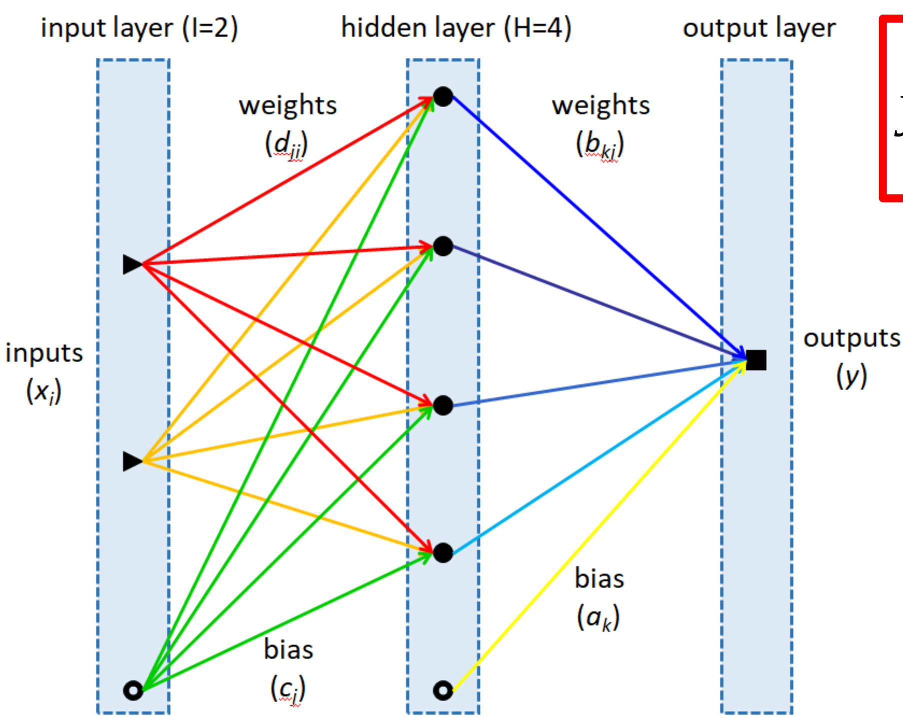
Bayesian approach in regression problem

- Posterior distributions of parameters are Neal1996Springer

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

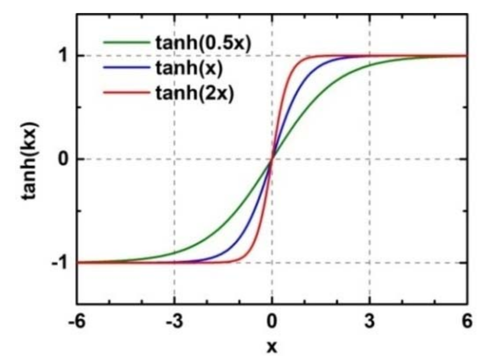
✓ likelihood function $p(D|\omega)$

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



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

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



Bayesian approach in regression problem

- Posterior distributions of parameters are Neal1996Springer

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

✓ sampling with Markov chain Monte Carlo (MCMC) method

- Make predictions

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

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

Remark:

- 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.

Outline

1. Introduction
2. Bayesian neural network approach
3. Results and discussion
   - ⋆ Toy model
   - ⋆ Nuclear masses
   - ⋆ Nuclear beta-decay half-lives
4. Summary and perspectives

Toy model

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

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

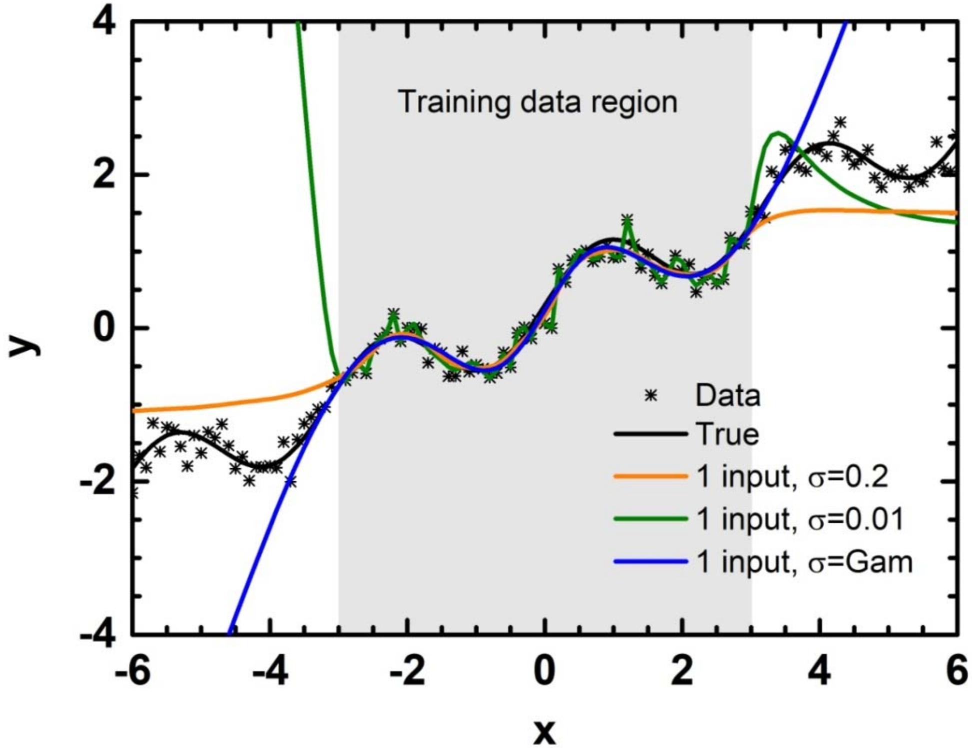
- Number of training data: N=61, x∈[-3, 3]

1 input : $y = f(x)$

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

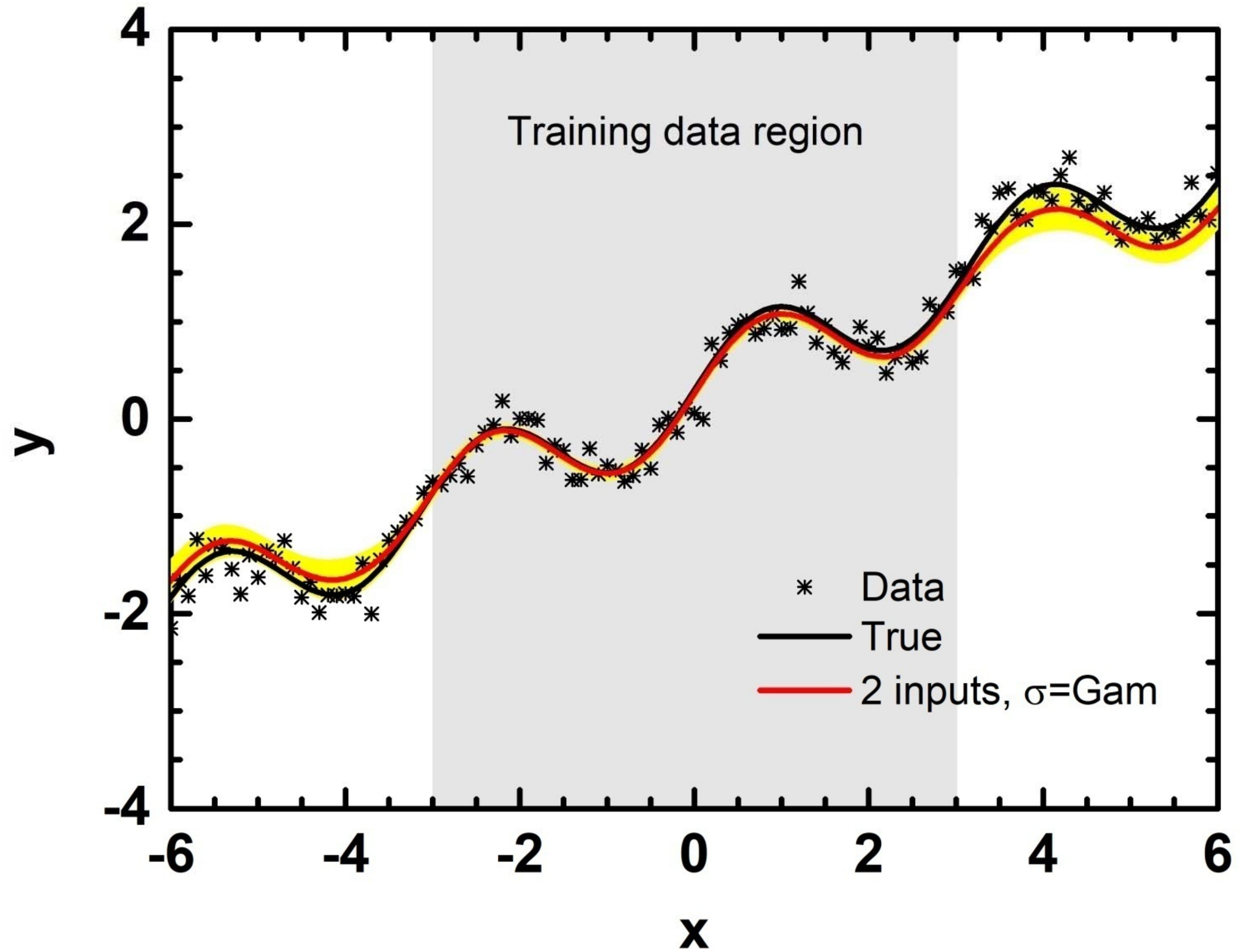
- Number of hidden unit: H=20 for $f(x)$; H=15 for $f(x_1, x_2)$
- Number of parameters: 61

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



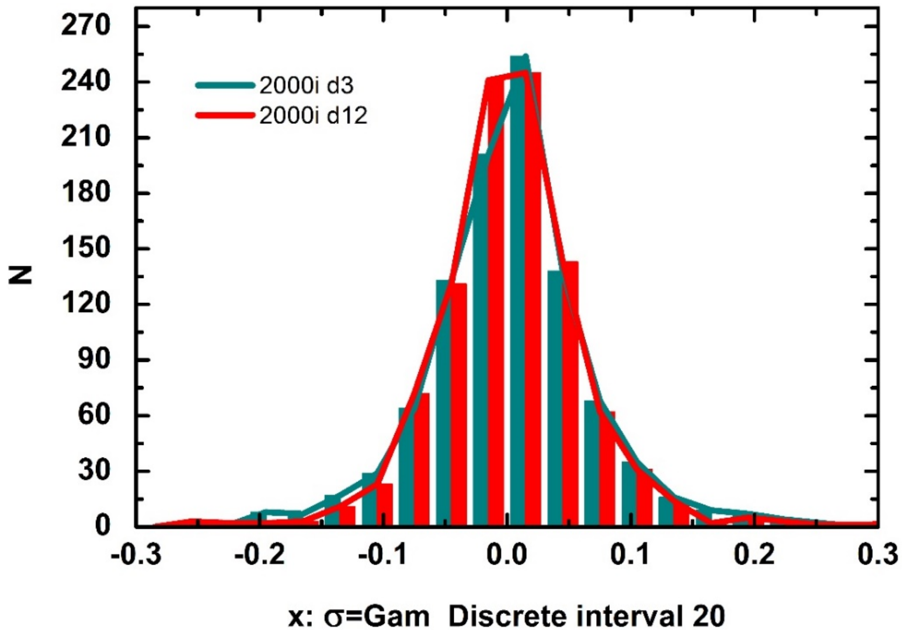
- ✓ BNN can avoid overfitting if a Gamma distribution is taken as the noise prior.
- ✓ Direct BNN fitting with *x* as the only input variable can only extrapolate around a few steps from known region, while the overfitting would make the extrapolation unacceptable.

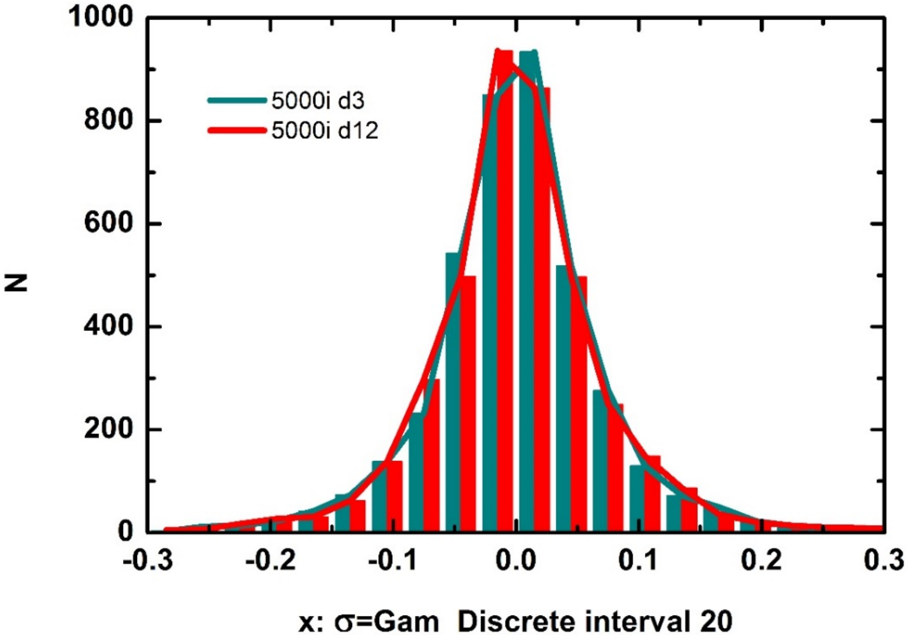
Toy model

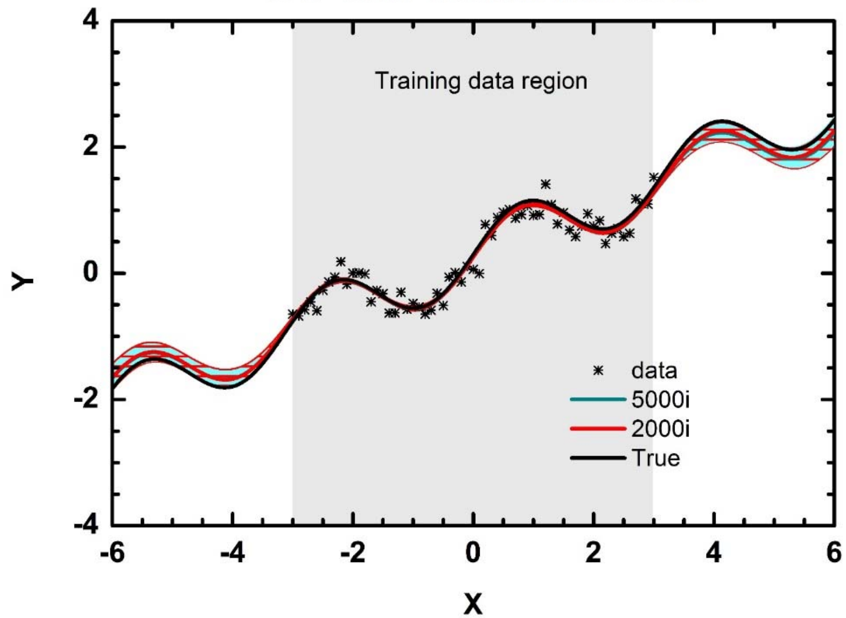


✓ Including reasonable variable is very effective for the extrapolation of neural network and the uncertainties of predictions are also reasonable.

Toy model







✓ 2000步迭代之后，不同d参数分布几乎一致。

✓ 2000步采样所得结果与5000步采样所得结果几乎一致。

Outline

1. Introduction
2. Bayesian neural network approach
3. Results and discussion
   - ⋆ Toy model
   - ⋆ Nuclear masses
   - ⋆ Nuclear beta-decay half-lives
4. Summary and perspectives

Numerical details

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

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

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

- ★ 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: W.J. Huang et al., CPC 41 030002; M. Wang et al., CPC 41 030003.
  - ✓ Entire set: 2272 nuclei in AME2016 (Z, N>=8 and $\sigma^{\exp}$<=100 keV)
  - ✓ Learning set: 1800 data randomly selected from entire set
  - ✓ Validation set: the remaining 472 data in entire set

Influence of noise variance

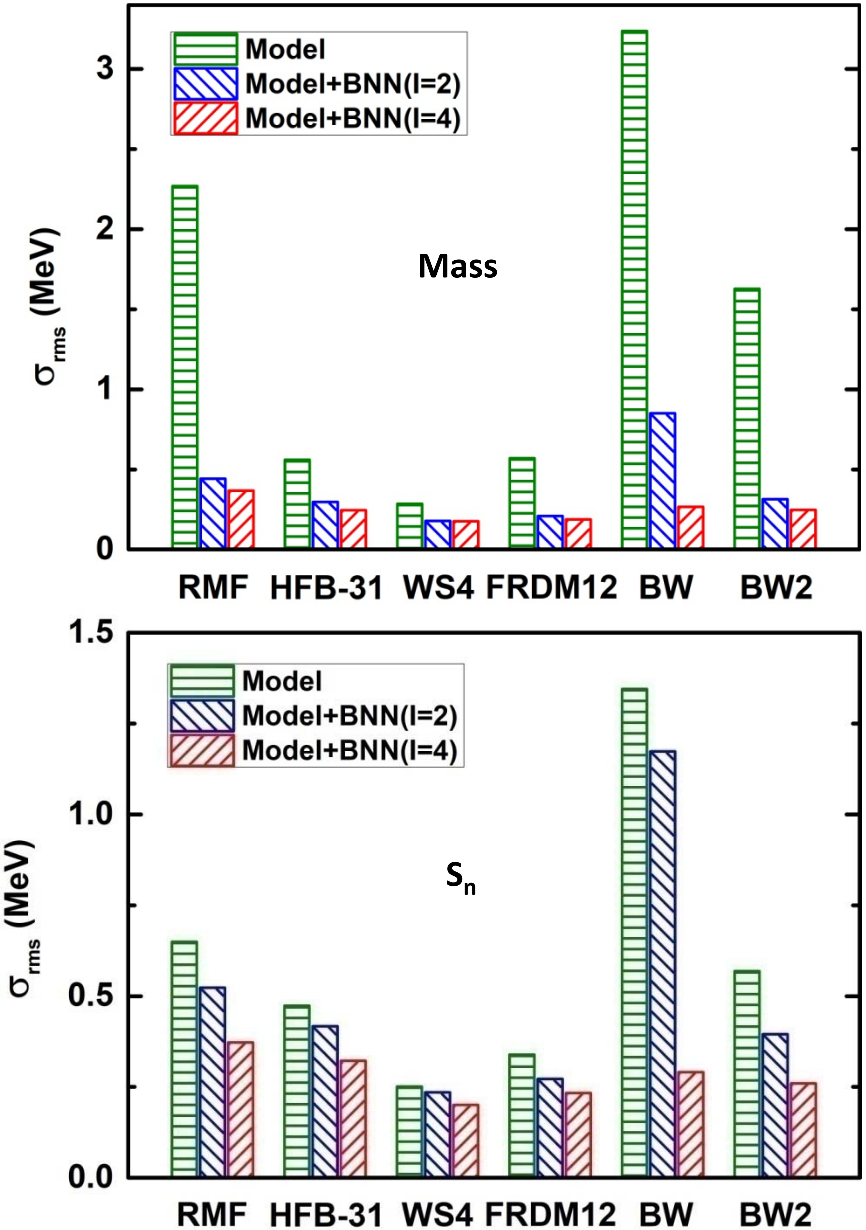
Model	Fixed	Gamma
RMF	0.732	0.443
HFB-31	0.354	0.296
WS4	0.203	0.178
FRDM2012	0.282	0.208
BW	1.035	0.850
BW2	0.497	0.313

Figure: rms deviations between experimental data and mass predictions for various models improved by BNN approach. The 2nd and 3rd columns denote the results with fixed value and gamma distribution for noise variance, respectively.

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

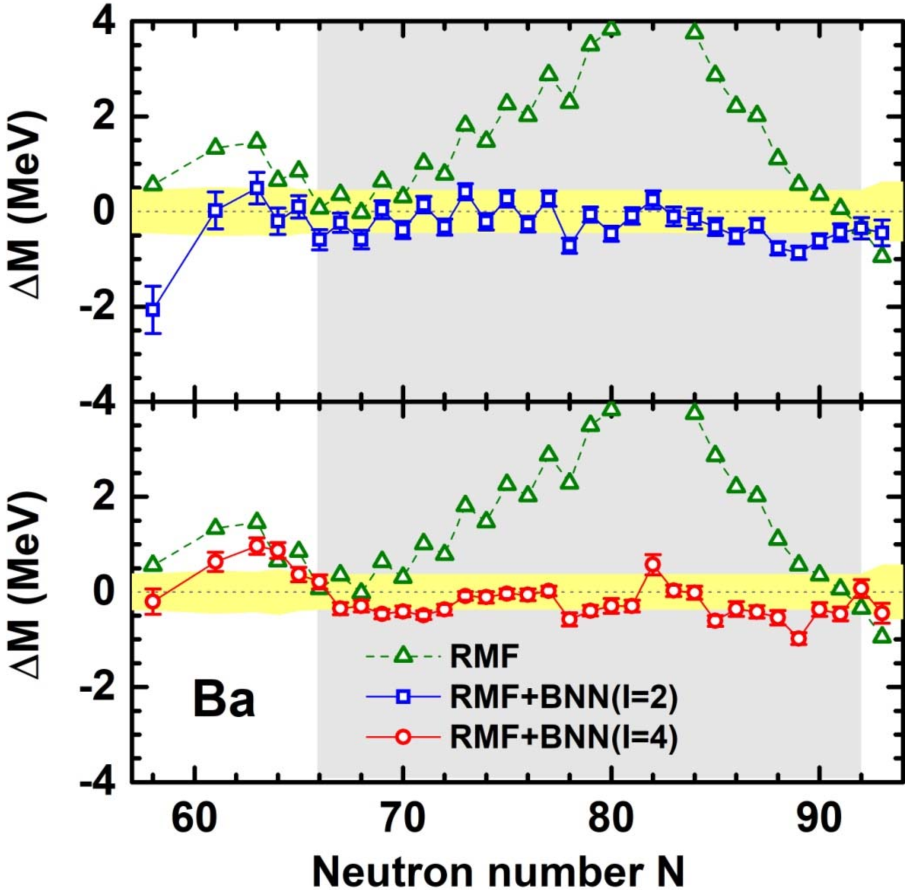
- The BNN approach can automatically find the optimal value for the noise variance, which can reduce the rms deviations by about 20%.

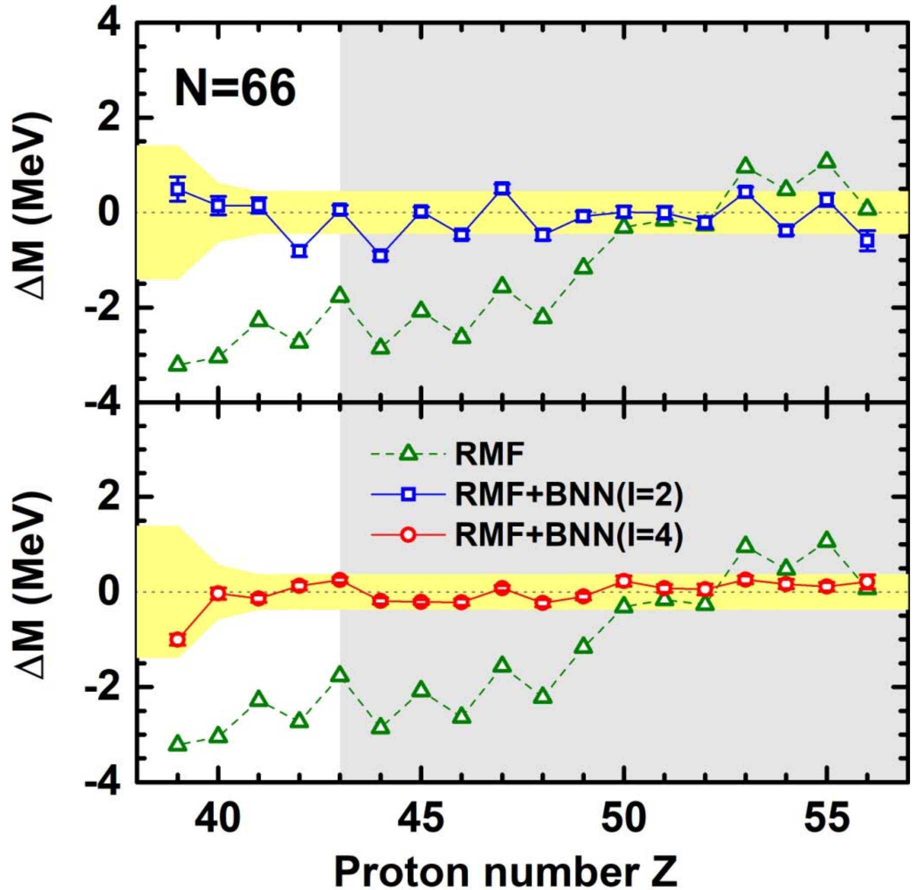
Rms deviations of mass and S_n



- 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 with both BNN approaches, 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.

Mass predictions of RMF+BNN model

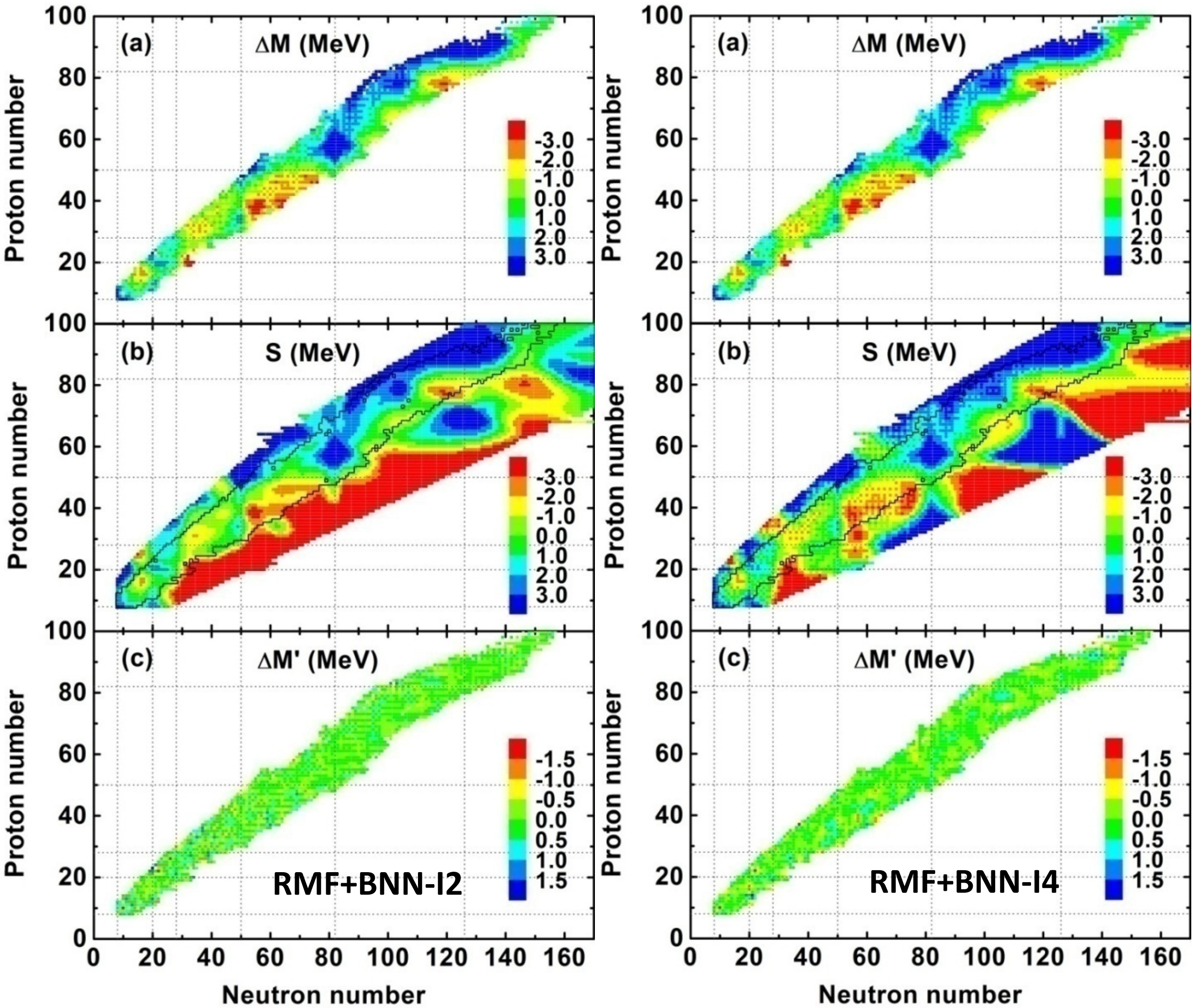


Figure: (a) Mass differences between the experimental data in AME16 and the predictions of the RMF model. (b) BNN corrections. (c) Mass differences after BNN improvement. Niu and Liang, PLB 778, 48 (2018)

- Smooth mass deviations can be easily removed by both BNN approaches, while the odd-even staggering can be well reproduced only using BNN-I4 approach.
- The extrapolation of BNN correction show more structure information for the BNN-I4 approach, especially the shell effects around (Z,N)=(28, 82) and (50, 126).

Outline

1. Introduction
2. Bayesian neural network approach
3. Results and discussion
   - ★ Toy model
   - ★ Nuclear masses
   - ★ Nuclear beta-decay half-lives
4. Summary and perspectives

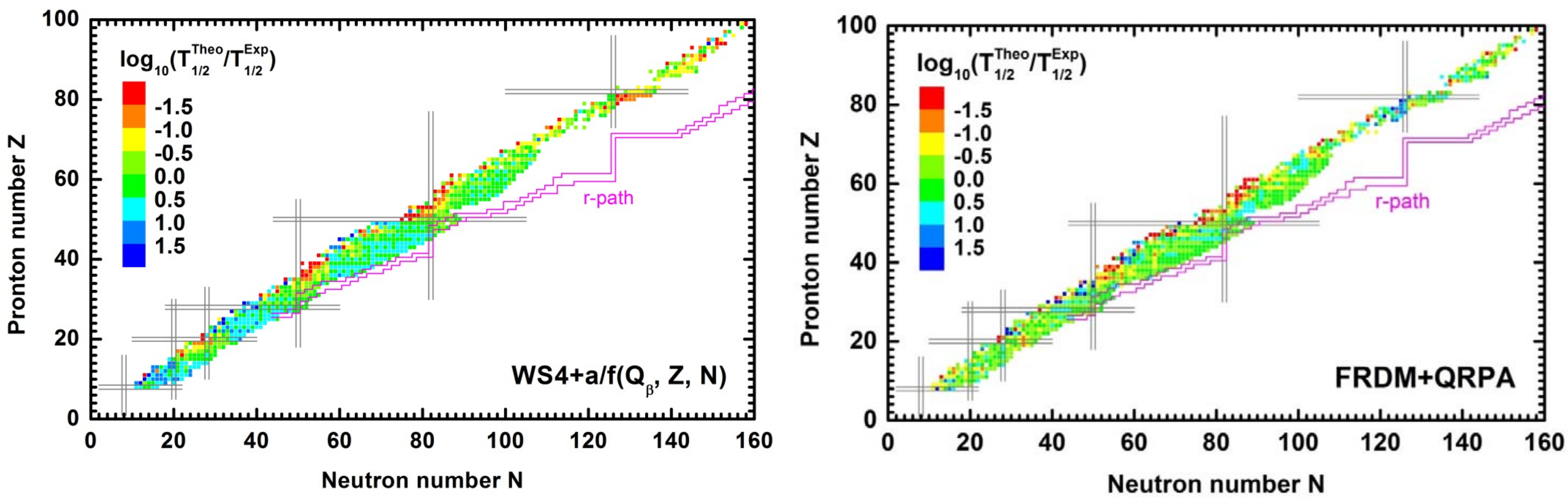
Nuclear β-decay half-lives

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

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

where $D = \frac{\hbar^7 2\pi^3 \ln 2}{g^2 m_e^5 c^4} = 6163.4$ s, $g_A = 1$, $B_{\mathrm{GT}}(E_m)$ 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 $p(D|\omega)$

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

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

★ 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

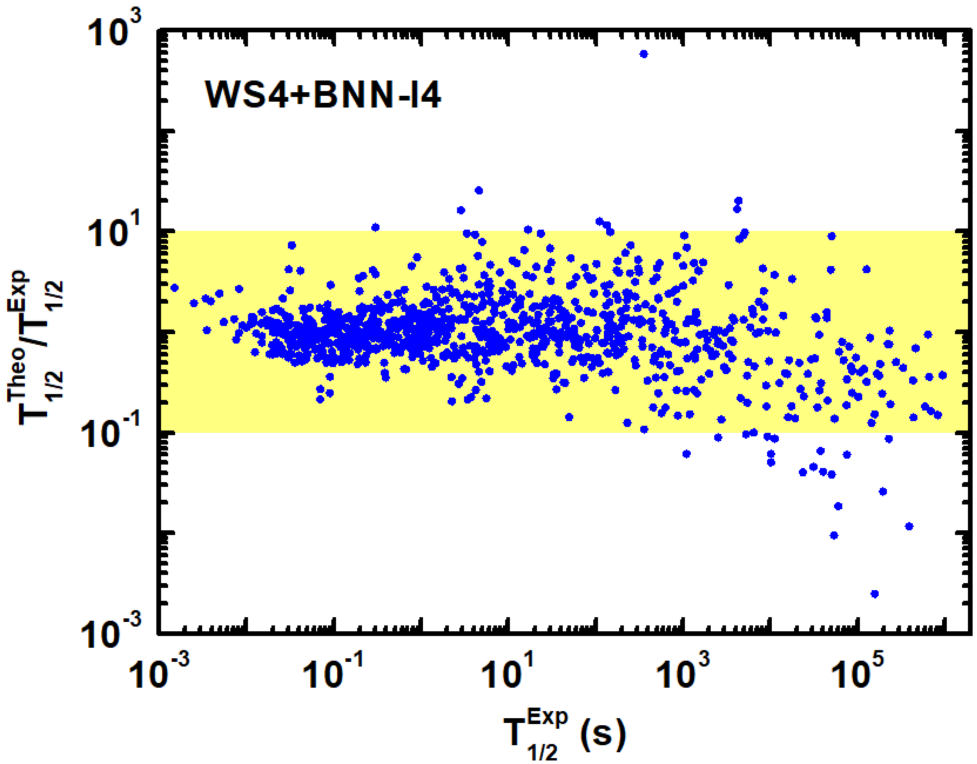
★ Data: G. Audi et al., CPC 41, 030001 (2017)

✓ 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 of BNN approaches

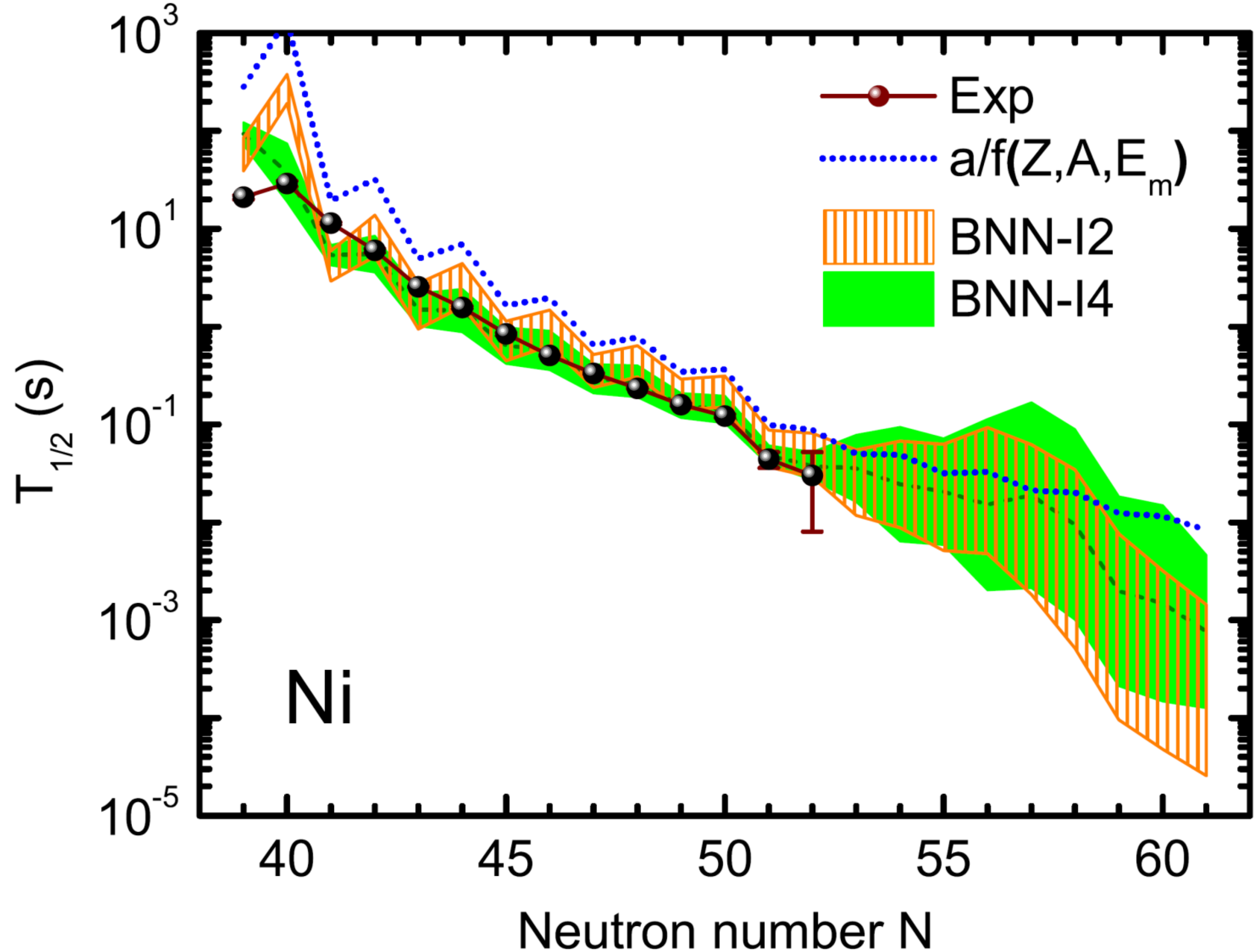


	$T_{1/2}<10^6$ s	$T_{1/2}<10^3$ s	$T_{1/2}<1$ s
WS4+a/f	0.8060	0.6302	0.5631
WS4+BNN-I2	0.4766	0.3542	0.2383
WS4+BNN-I4	**0.3999**	**0.3146**	**0.2036**
FRDM+QRPA	0.8190	0.5969	0.3906
RHB+QRPA	1.8844	1.6196	0.4631

- The WS4+BNN-I4 approach usually better reproduce the half-lives of short-lived nuclei.
- The WS+BNN-I4 approach gives the best results, which can describe nuclear half-lives around $10^{0.2}$=1.6 times of experimental data for nuclei with half-lives shorter than 1 s.

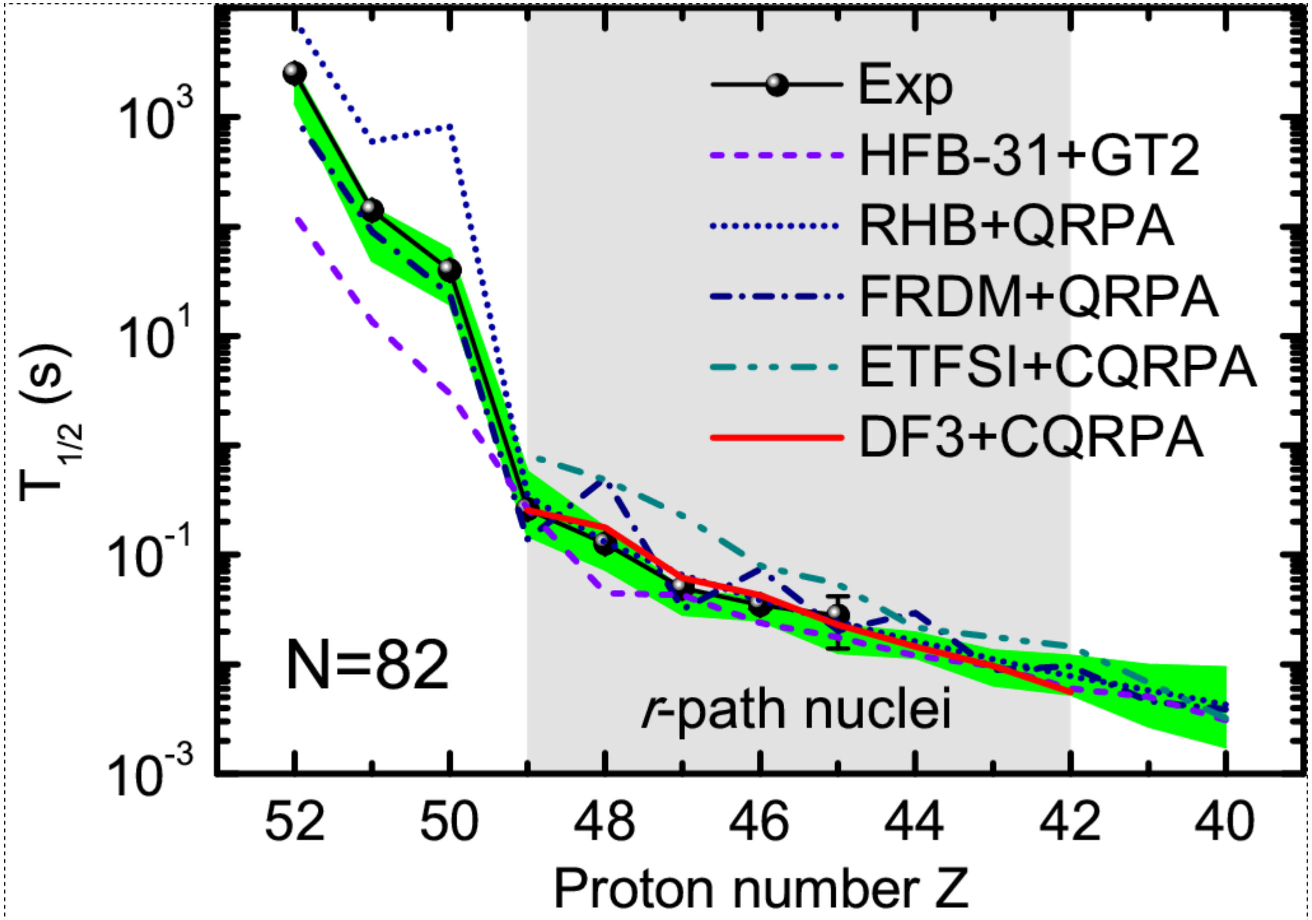
$$\sigma_{\mathrm{rms}} = \sqrt{\frac{\sum_{i=1}^{n}\left[\log_{10}\left(T_{1/2}^{\mathrm{Exp}} / T_{1/2}^{\mathrm{Theo}}\right)\right]^2}{n}}$$

Half-lives with BNN approaches



- $T_{1/2}$=$a/f(Q_\beta, Z, N)$ generally overestimates the odd-even staggering in half-lives.
- BNN-I2 approach cannot easily remove odd-even staggering in half-lives, while BNN-I4 approach well reproduce the experimental data. Z. M. Niu et al., PRC 99, 064307 (2019)

Predictions of nuclear half-lives



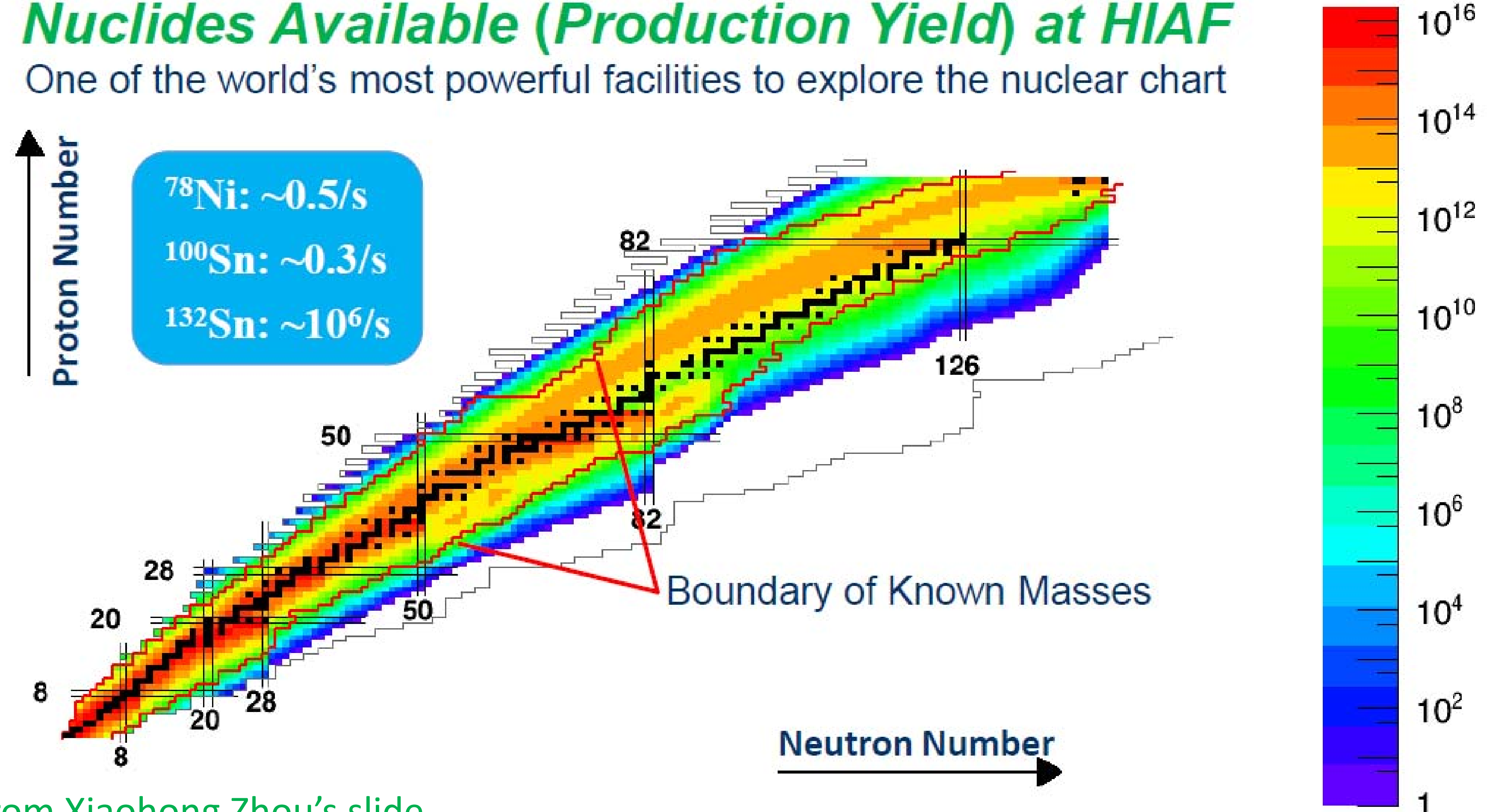
- The results of WS4+BNN-I4 approach are in good agreement with the experimental data, even completely agree with the experimental data within uncertainties for short-lived nuclei.
- When extrapolate from known region, the results of other models generally agree with WS4+BNN-I4 predictions within uncertainties.

Z. M. Niu et al., PRC 99, 064307 (2019)

Capability of Producing Nuclides

Nuclides Available (Production Yield) at HIAF

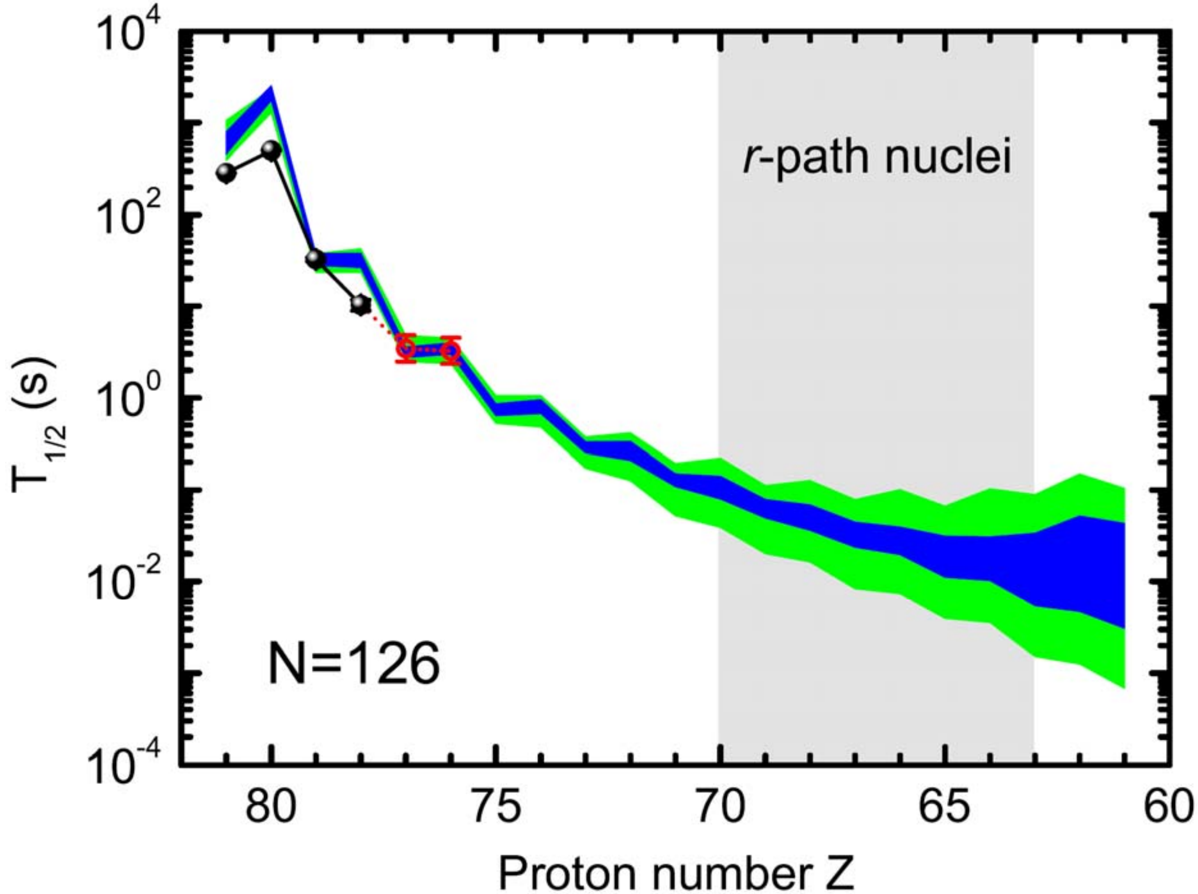
One of the world's most powerful facilities to explore the nuclear chart



From Xiaohong Zhou's slide

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"

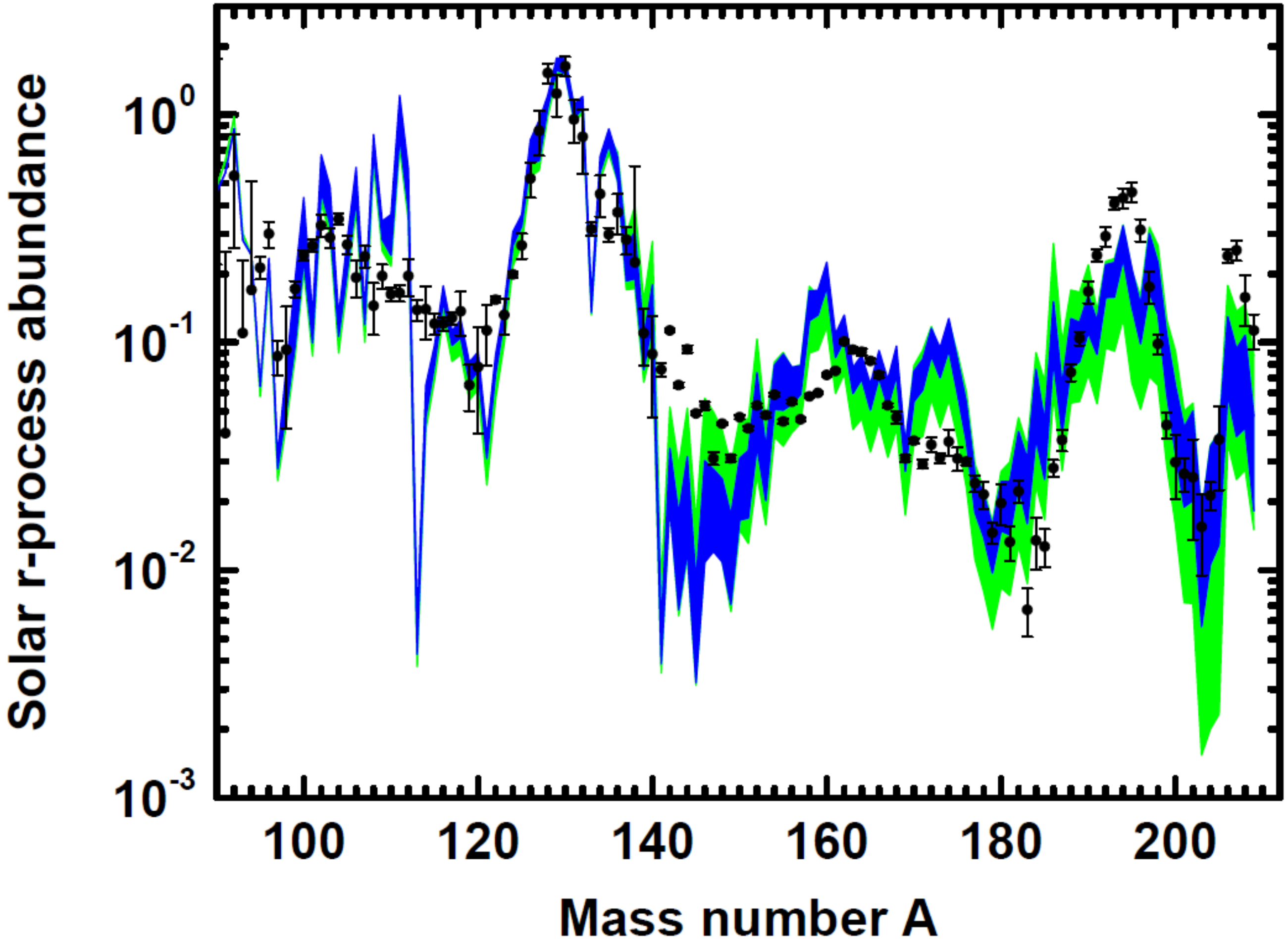
Predictions of nuclear half-lives



- If we can further measure three more β-decay half-lives for each isotopes
  - ✓ 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.

Z. M. Niu et al., PRC 99, 064307 (2019)

Predictions of r-process abundances



- Uncertainties from β-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. Z. M. Niu et al., PRC 99, 064307 (2019)

Outline

1. Introduction
2. Bayesian neural network approach
3. Results and discussion
   - ⋆ Toy model
   - ⋆ Nuclear masses
   - ⋆ Nuclear beta-decay half-lives
4. Summary and perspectives

Summary

- BNN approach was employed to predict nuclear masses and nuclear β-decay half-lives:
  - ⋆ BNN approach significantly improve the accuracies of nuclear mass and β-decay half-life predictions.
  - ⋆ It is found that the inclusion of more physical features (masses: δ and P; half-lives: δ and Q_β) is very important to achieve better predictive performance.

Acknowledgement

合作者：

北京航空航天大学（BUAA）：孙保华

兰州大学（LZU）：龙文辉、牛一斐

日本理化学研究所（RIKEN）：梁豪兆

Thank you!