

第九届华大QCD讲习班：深度学习与粒子物理/核物理

武汉, 2021.10.11-15

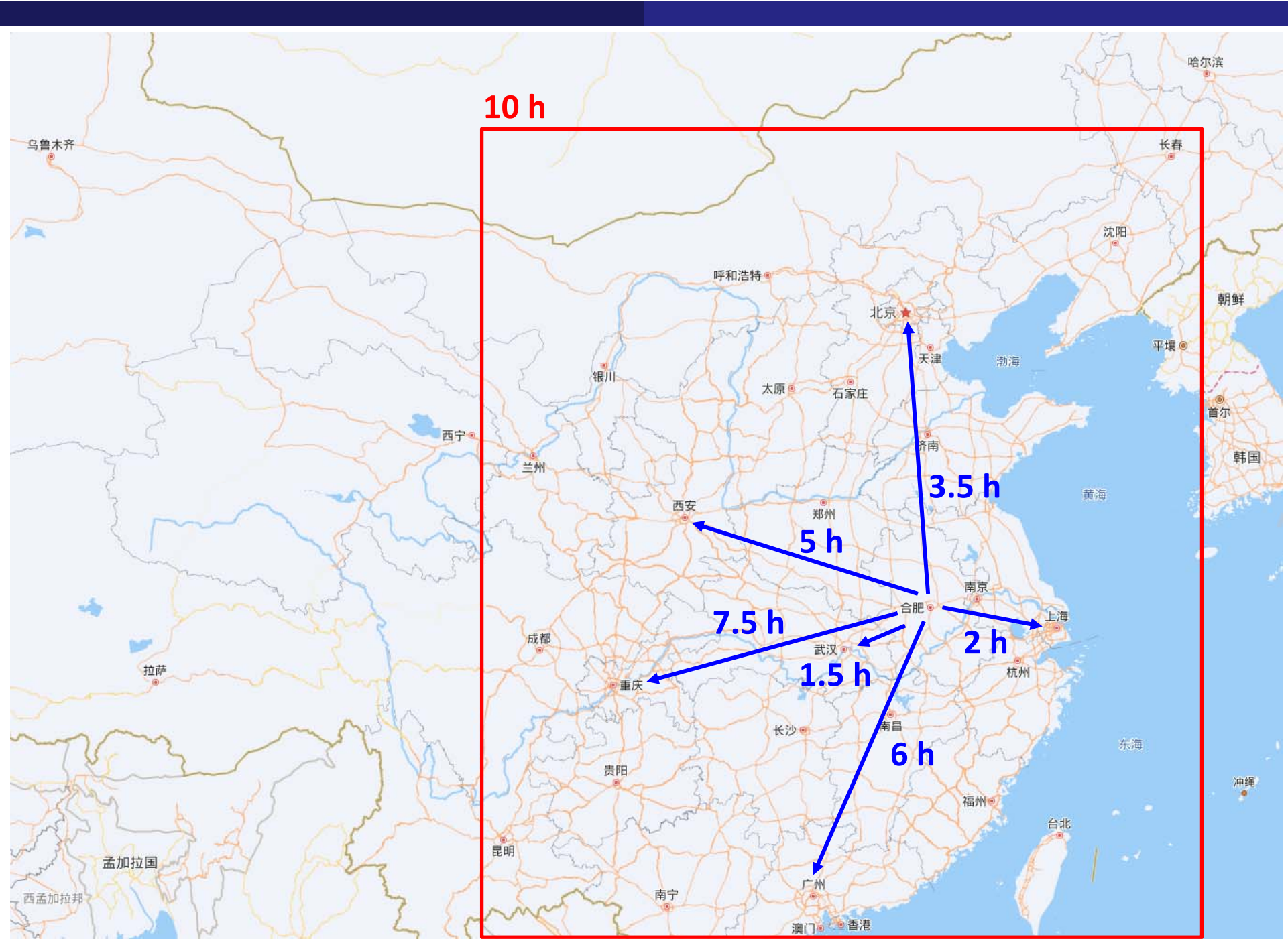
贝叶斯神经网络与核物理：质量与 β 衰变寿命

牛中明

安徽大学 物理与光电工程学院

2021年10月13日





10 h



● 安徽大学

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

● 物理学科

- ✓ 安徽大学开校学科之一（1928年）
- ✓ 首批国家级一流本科专业建设点
- ✓ 本、硕、博全链条人才培养体系
- ✓ 安徽省重点学科和省博士后岗位



- 安徽大学核与粒子物理组始于1979年2月-3月在广西南宁举办的核结构理论讲习班。我校徐辅新、李先胤两位教师参加。



- ✓ 1979年，举办了“拉卡代数和核结构”讲习班。
- ✓ 1988年，举办了第二届全国核结构大会。
- ✓ 1999年，举办了首届全国核反应会议。
- ✓ 2002年，举办了第九届全国核结构大会。
- ✓ 2012年，举办了“原子核物理中的协变密度泛函理论”讲习班。
- ✓ 2014年，举办了“核天体反应截面间接测量方法”研讨会。
- ✓ 2018年，举办了“原子核与强子物理前沿进展”研讨会。



郭建友：教授，博导，中国科技大学博士（导师：阮图南、徐辅新）；2002-2004，北京大学博士后（合作导师：孟杰）。物理系系主任、安徽大学粒子与核物理研究组负责人，校学术委员会委员。

李 刚：教授，中国科技大学博士（导师：马文淦）；2013-2014，Stanford University访问学者（合作导师：Stanley J. Brodsky）。物理与材料科学学院副院长。



牛中明：教授，博导，北京大学博士（导师：孟杰）；2016-2017，RIKEN博士后（合作导师：Tetsuo Hatsuda）。校团委兼职副书记、院学术委员会副主任。

刘 泉：教授，安徽大学博士（导师：郭建友）；2015-2016，Michigan State University访问学者（合作导师：Witold Nazarewicz）。





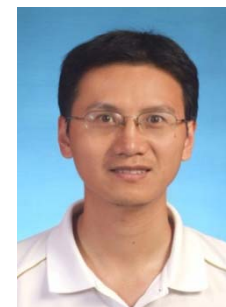
张 宇：特任研究员，中国科技大学博士（导师：马文淦）；南京大学博士后（合作导师：刘作伟）。

宋 昂：副教授，中国科技大学博士（导师：马文淦）。



衡太骅：副教授，中国科技大学博士（导师：井思聪）；2014-2015, Iowa State University访问学者（合作导师：James P. Vary）。

文群刚：讲师，中国原子能科学院博士（导师：张焕乔、周书华）；2009-2012中国科技大学博士后（合作导师：赵政国）。



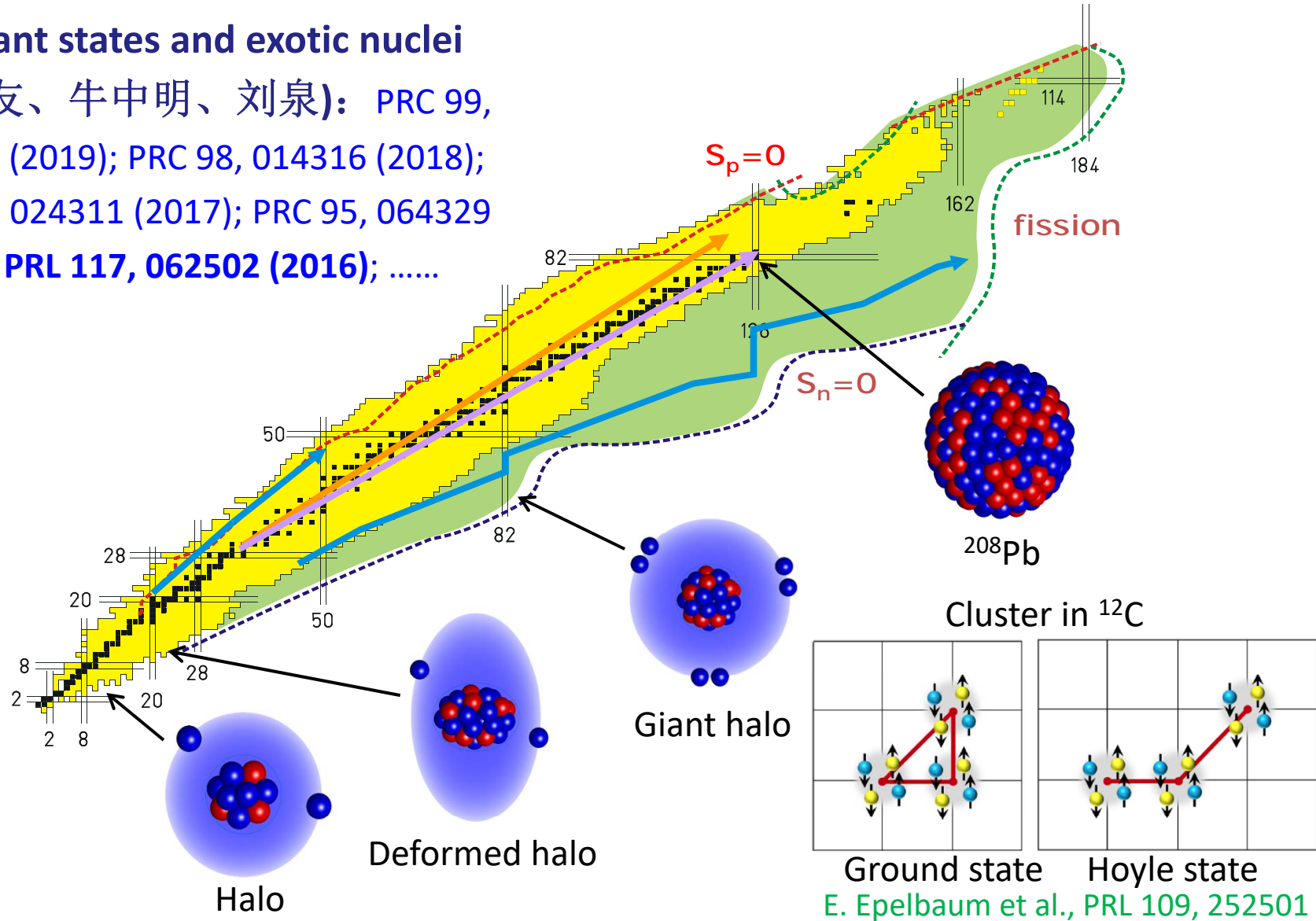
陈寿万：讲师，中国科技大学博士（导师：王群），2016-2018, Washington University (St. Louis)访问学者（合作导师：Willem H. Dickhoff）。





✓ Resonant states and exotic nuclei

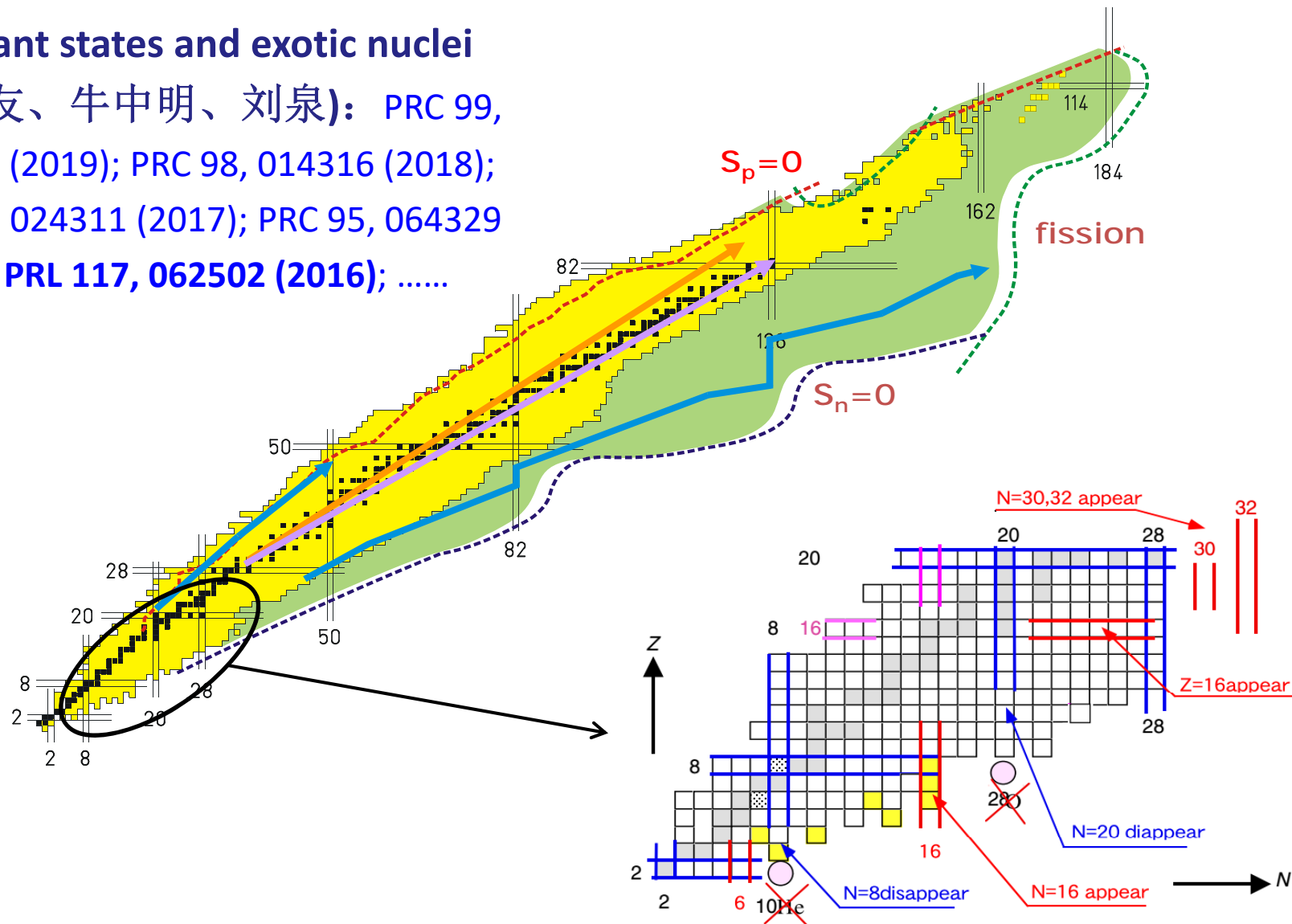
(郭建友、牛中明、刘泉): PRC 99, 014309 (2019); PRC 98, 014316 (2018); PRC 95, 024311 (2017); PRC 95, 064329 (2017); PRL 117, 062502 (2016);





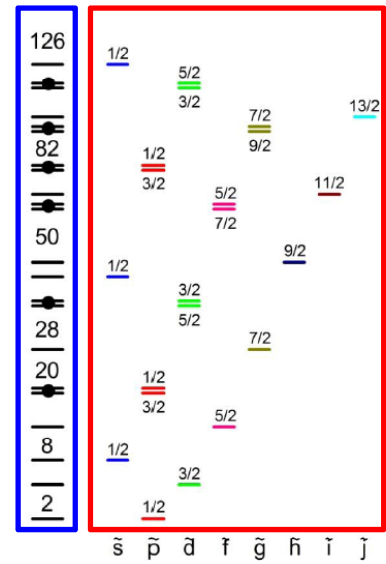
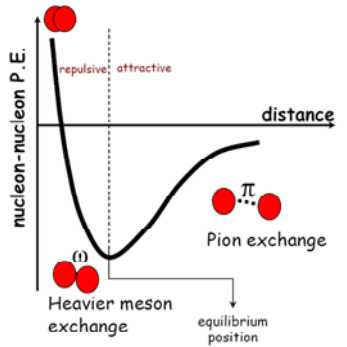
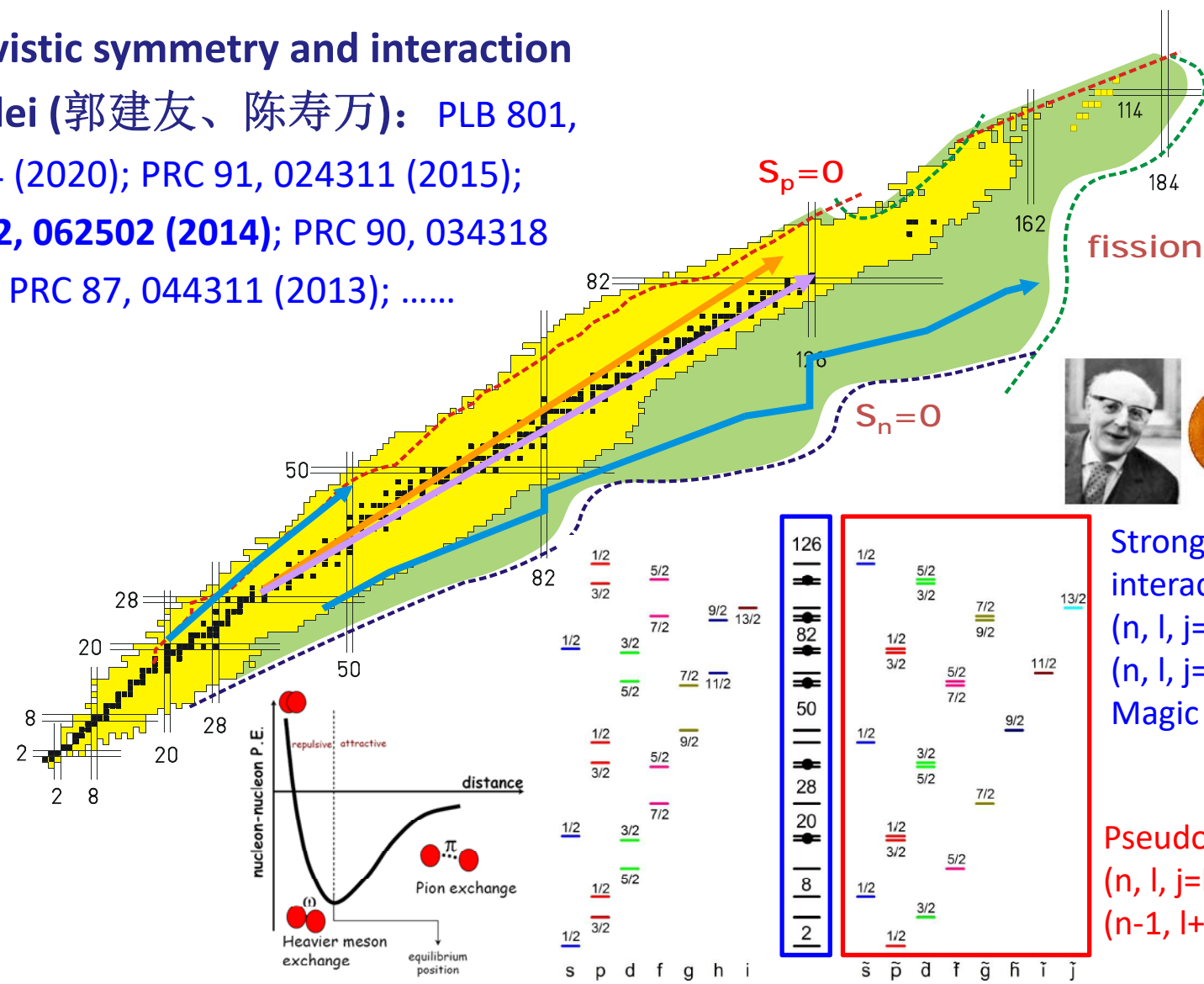
✓ Resonant states and exotic nuclei

(郭建友、牛中明、刘泉): PRC 99, 014309 (2019); PRC 98, 014316 (2018);
 PRC 95, 024311 (2017); PRC 95, 064329 (2017); PRL 117, 062502 (2016);





✓ Relativistic symmetry and interaction in nuclei (郭建友、陈寿万): PLB 801, 135174 (2020); PRC 91, 024311 (2015); PRL 112, 062502 (2014); PRC 90, 034318 (2014); PRC 87, 044311 (2013);



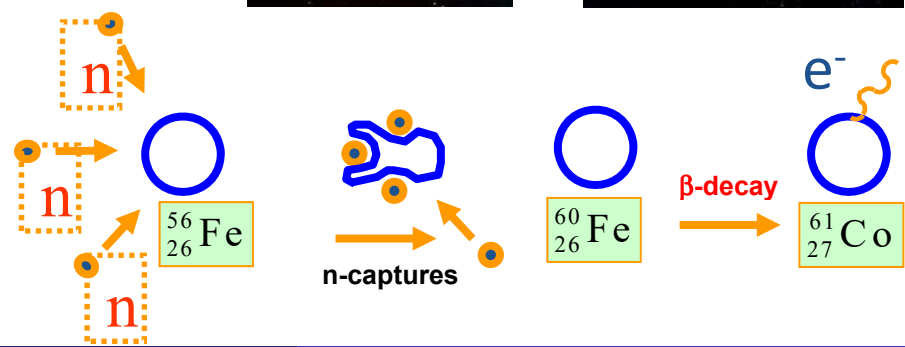
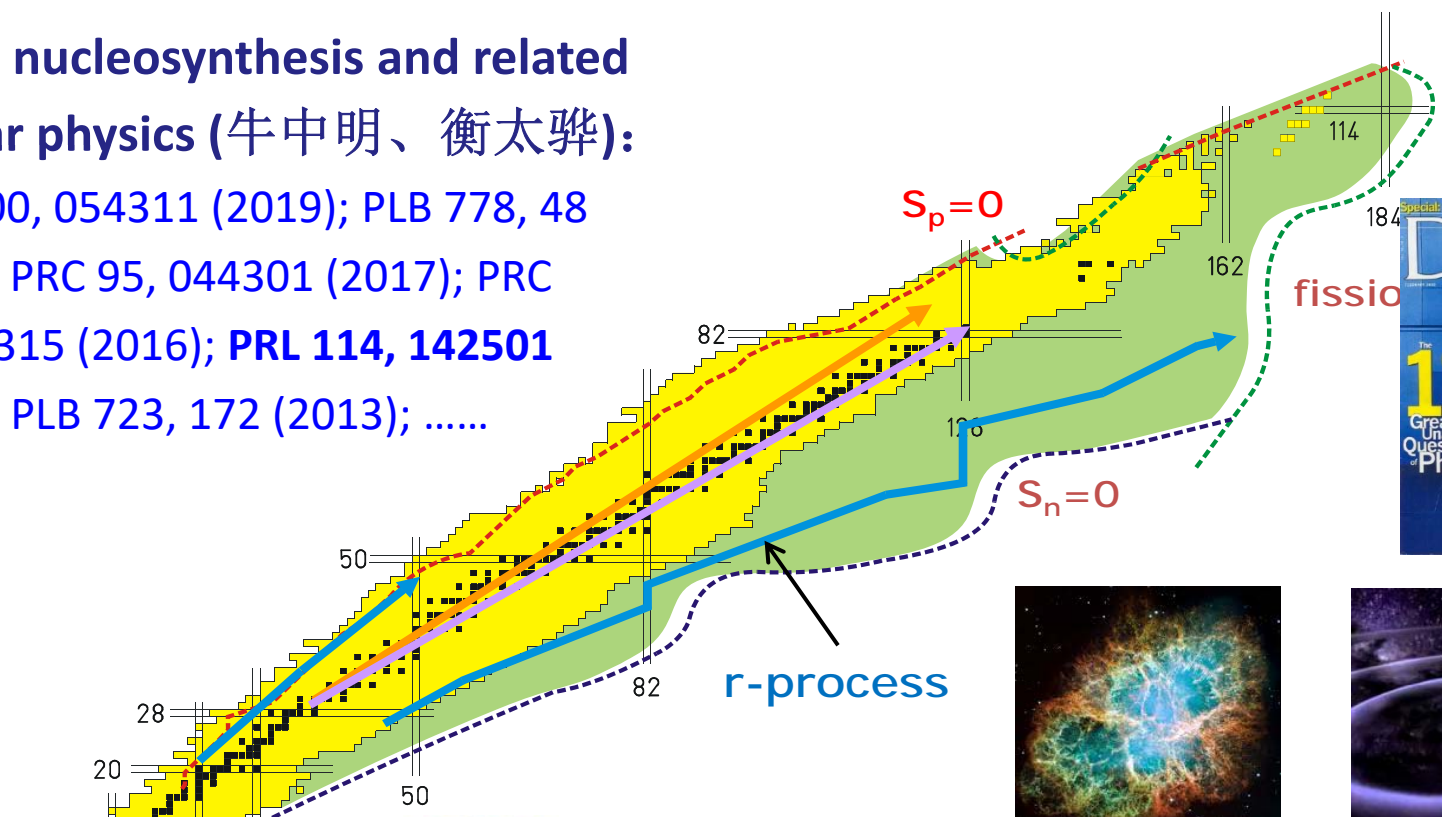
Strong spin-orbit interaction:
 $(n, l, j=l+1/2)$
 $(n, l, j=l-1/2)$
Magic numbers

Pseudospin symmetry
 $(n, l, j=l+1/2)$
 $(n-1, l+2, j=l+3/2)$



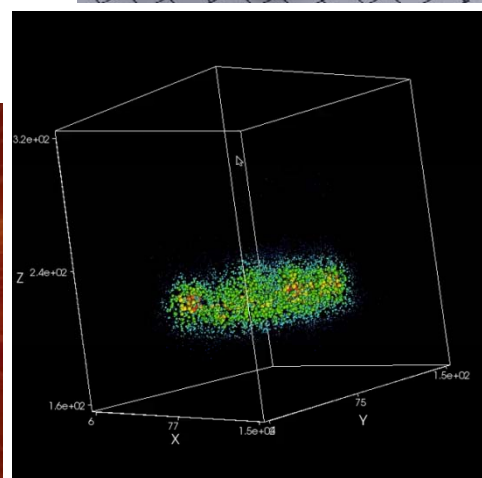
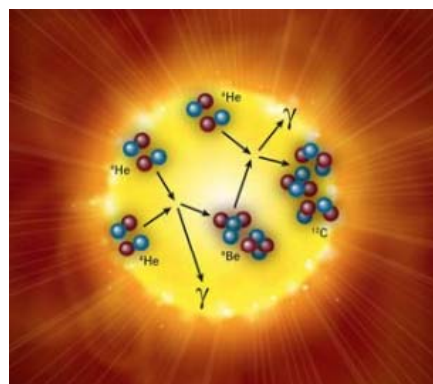
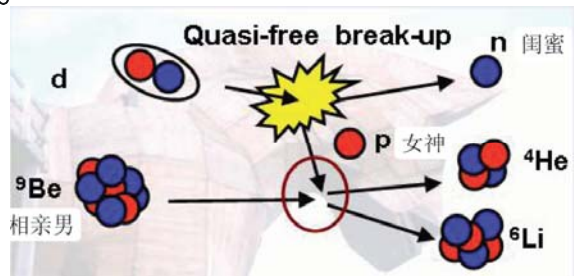
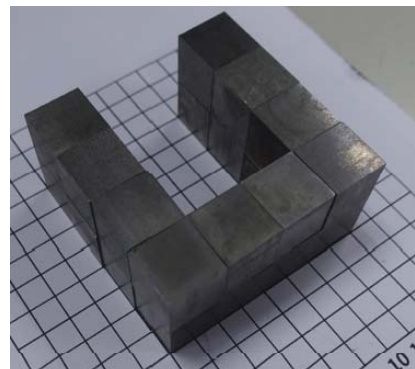
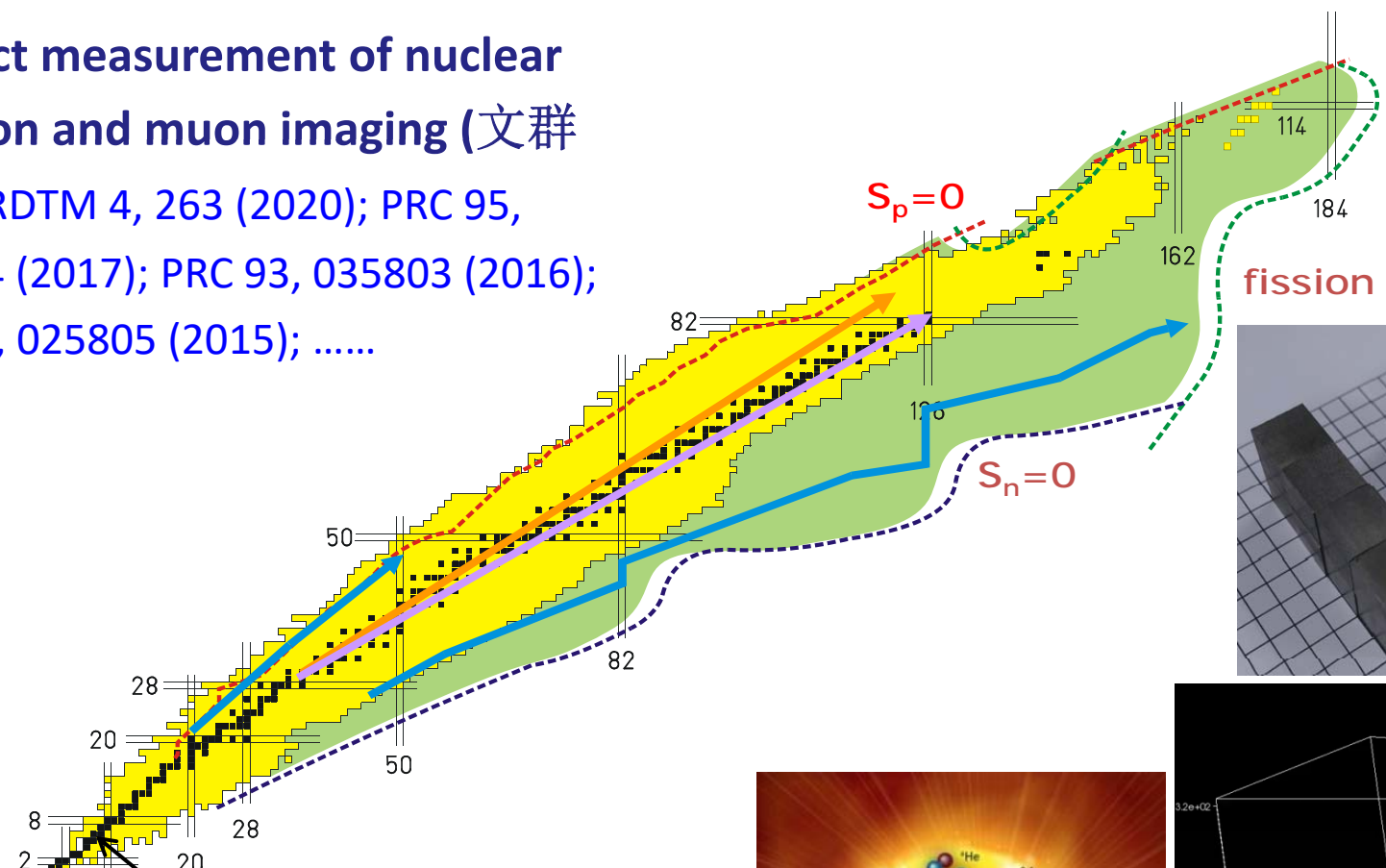
✓ Stellar nucleosynthesis and related nuclear physics (牛中明、衡太骅):

PRC, 100, 054311 (2019); PLB 778, 48 (2018); PRC 95, 044301 (2017); PRC 94,054315 (2016); **PRL 114, 142501 (2015)**; PLB 723, 172 (2013);





✓ Indirect measurement of nuclear reaction and muon imaging (文群刚): RDTM 4, 263 (2020); PRC 95, 035804 (2017); PRC 93, 035803 (2016); PRC 92, 025805 (2015);



安徽大学2021年引进人才待遇及支持条件参考标准

岗位设置及引才对象	岗位薪酬	安家费
<p>I类岗位：杰出人才 在国内外享有崇高声誉和重大影响，具有高深的学术造诣，在人才培养、科学研究、工程技术、成果转化、社会服务等方面做出系统的重大的、创造性的成就和贡献。</p>	面议	面议
<p>II类岗位：领军人才 在人才培养、科学研究、工程技术、成果转化、社会服务等方面做出重大贡献和成就，得到领域广泛认可，具有赶超或引领国际先进水平的能力，能够带领团队协同攻关。</p>	年薪90-120万元	350万元
<p>III类岗位：拔尖人才 在人才培养、科学研究、工程技术、成果转化、社会服务等方面取得突出成果，具有协助所在学科赶超或保持国际先进水平的能力。</p>	年薪50-70万元	200万元
<p>IV类岗位：青年拔尖人才 在人才培养、科学研究、工程技术、成果转化、社会服务等方面取得优秀成果，具有推动所在学科发展的能力。</p>	紧缺专业年薪45万元 一般专业年薪35万元	紧缺专业110万元 一般专业90万元
<p>V岗位：学术骨干人才 近5年的业绩成果一般应达到我校相应学科教学科研型教授平均水平；或具有海外工作经历且达到我校科研为主型教授平均水平。</p>	紧缺专业年薪30万元 一般专业年薪25万元	紧缺专业65万元 一般专业60万元

Outline

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

Outline

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

Nuclear physics and machine learning

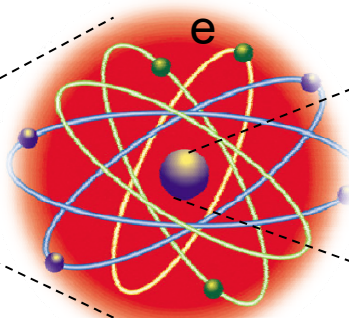
足球



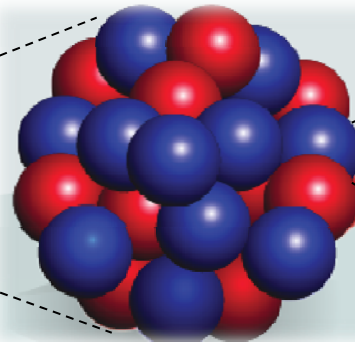
分子



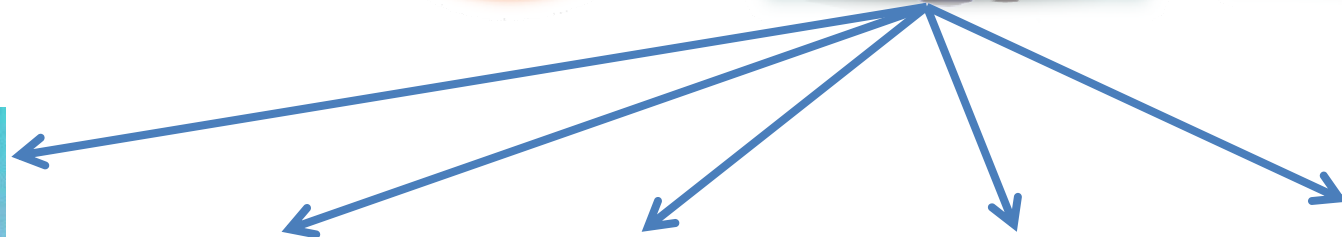
原子



原子核



核子



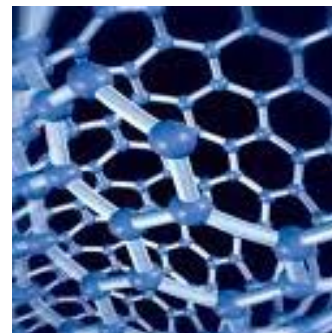
国防建设



能源动力



医学健康



材料科学

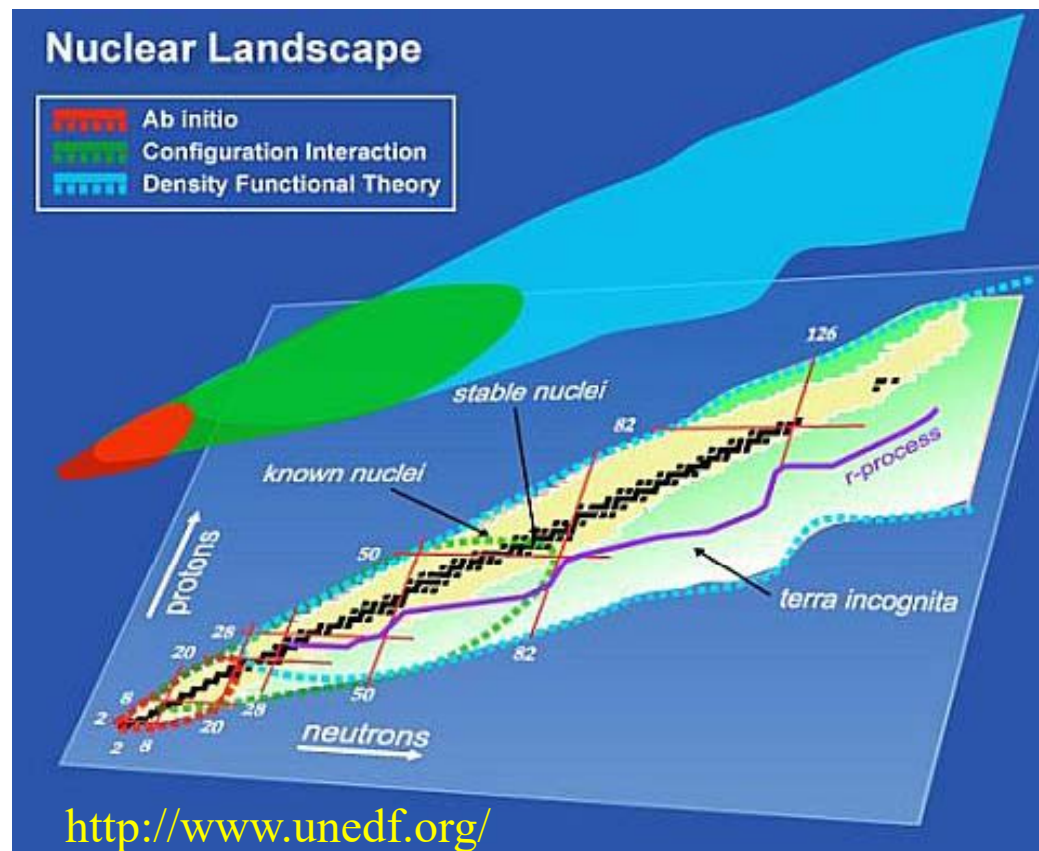
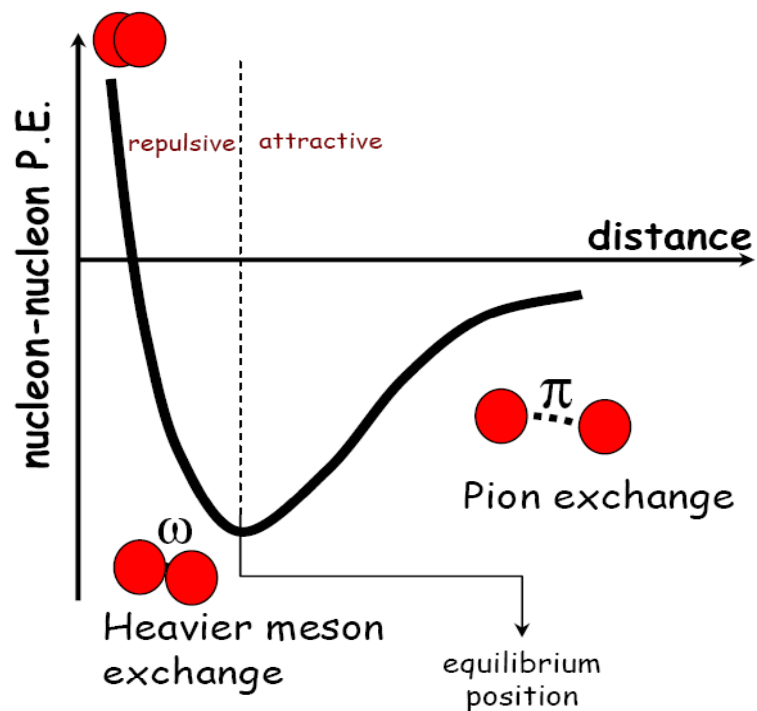
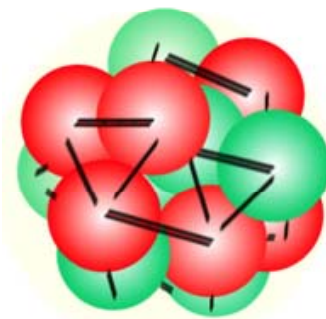


天体物理&宇宙学

Nuclear physics

● 原子核是由质子和中子组成的**强相互作用**束缚的**量子多体系统**。

- ★ 核力：现实核力、有效核力
- ★ 核多体方法：从头计算、壳模型、密度泛函理论



Nuclear physics

核数据:

- 质量: ~2500
AME2020: [Wang2021CPC](#)
- 电荷半径: ~1000
[Angeli2013ADNDT](#)
- 激发能量: ~400, 900, 800, ...
 0_2^+ , 2_1^+ , 4_1^+ , ... <http://www.nndc.bnl.gov/>
- 衰变寿命: ~3000
NUBASE2020: [Kondev2021CPC](#)
- 核反应实验数据库 **EXFOR**: ~25000
<https://www-nds.iaea.org/exfor/>

世界五大通用核数据库

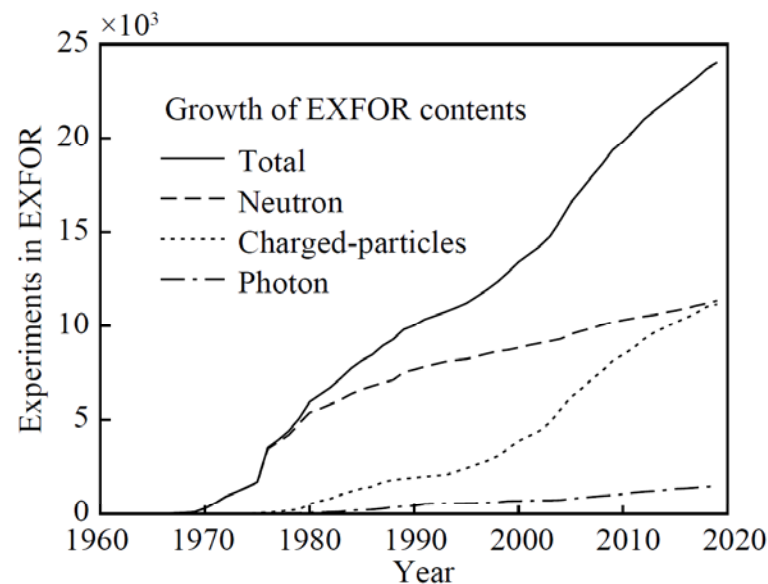
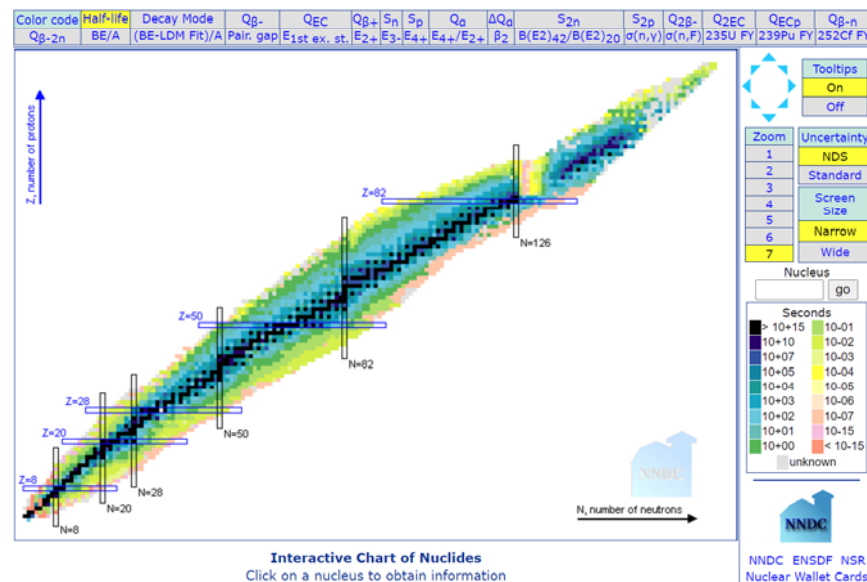
中国: CENDL: <http://www.nuclear.csdb.cn/>

美国: ENDF: <http://www.nndc.bnl.gov/>

日本: JENDL: <http://www.nndc.jaea.go.jp/>

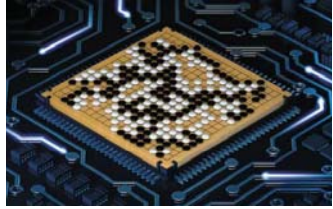
欧洲: JEFF: <http://www.oecd-nea.org/>

俄国: BROND: <http://www-nds.iaea.org/>



Wang2021NPR

Machine learning



game of go



computer vision



translation



autonomous car



voice recognition

Machine learning in nuclear physics

- **NN**: Neural Network
- **BNN**: Bayesian Neural Network
- **CNN**: Convolutional Neural Network
- **MDN**: Mixture Density Network
- **BGP**: Bayesian Gaussian Processes
- **CGP**: Constrained Gaussian Processes
- **DT**: Decision Tree
- **NBP**: Naive Bayesian Probability Classifier
- **SVM**: Support Vector Machines
- **RBF**: Radial Basis Function
- **KRR**: Kernel Ridge Regression
- **CLEAN**: CLEAN Image Reconstruction
- ...

Nuclear physics and machine learning

● Nuclear structure:

➤ Masses:

NN: [Gazula1992NPA](#), [Athanasopoulos2004NPA](#), [Bayram2014ANE](#), [Zhang2017JPG](#),
[Yuksel2021IJMPE](#); DT: [Gao2021arXiv](#); NBP: [Liu2021PRC](#); CLEAN: [Morales2010PRC](#)
 BNN: [Utama2016PRC](#), [Niu2018PLB](#), [Niu2019PRC](#), [Rodriguez2019EPL](#)
 RBF: [Wang2011PRC](#), [Niu2013.2016PRC](#); KRR: [Wu2020PRC](#), [Wu2021PLB](#)
 SVM: [Clark2006IJMPB](#); BNN (Q_α): [Rodriguez2019JPG](#)
 BNN and BGP ($S_{1n/2n}$): [Neufcourt2018,2020PRC](#), [Neufcourt2019PRL](#)

➤ Nuclear spins and parities:

NN: [Gernoth1993PLB](#); SVM: [Clark2006IJMPB](#)

➤ Charge radii:

BNN: [Utama2016JPG](#); CNN: [Cao2020APS](#); NN: [Wu2020PRC](#); NBP: [Ma2020PRC](#)

➤ Excited states:

NN (energies of 2_1^+): [Akkoyuna2020arXiv](#)

NN (excitation spectrum): [Lasseri2020PRL](#)

NN (giant dipole resonance key parameters): [Bai2021PLB](#), [Wang2021PRC](#)

➤ Extrapolation problems in ab initio method:

NN: [Negoita2019PRC](#), [Jiang2019PRC](#); CGP: [Yoshida2020PRC](#)

Nuclear physics and machine learning

● Nuclear decays and reactions:

➤ α -decay half-lives:

DT, NN: [Saxena2021JPG](#), [Ma2021CPC](#)

➤ β -decay half-lives:

BNN: [Niu2019PRC](#); SVM, NN: [Costiris2008arXiv](#); NN: [Costiris2009PRC](#)

➤ Fission yields:

BNN: [Wang2019PRL](#), [Qiao2021PRC](#); MDN: [Lovell2019EPJWC](#)

➤ Cross-sections in proton induced spallation reactions:

BNN: [Ma2020CPC](#)

➤ Neutron-nucleus scattering data:

BNN: [Liang2021Thesis](#)

➤ Fusion reaction cross-sections:

NN: [Akkoyun2020NIMB](#)

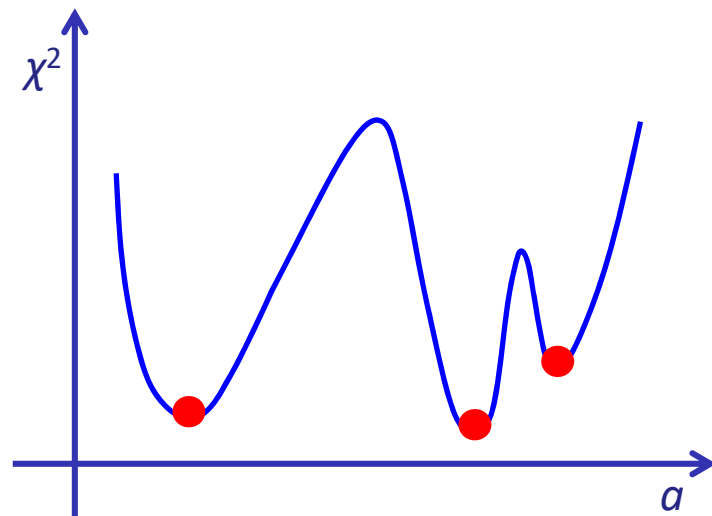
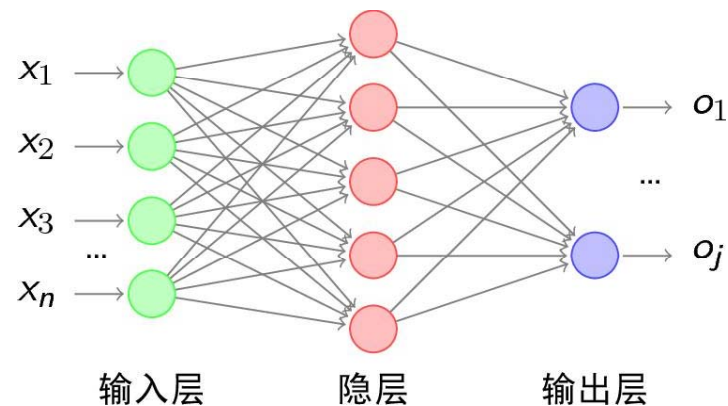
Introduction

● 贝叶斯神经网络:

- ★ 强大的数据拟合能力
- ★ 能够推广并且预测未知数据
- ★ 学习规则简单，便于计算机实现

- ★ 利用先验自动避免过拟合
- ★ 量化预言不确定性
- ★ 自动识别不相关输入

■ 黑盒算法、数据饥饿、计算耗时



Introduction

Physics Letters B 778 (2018) 48–53



ELSEVIER

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Physics Letters B

www.elsevier.com/locate/physletb



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,*}

PHYSICAL REVIEW C **99**, 064307 (2019)

^a School of Physics and

^b Interdisciplinary The

^c RIKEN Nishina Cente

^d Department of Physi

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

Z. M. Niu (牛中明)^{1,2}, H. Z. Liang (梁豪兆)^{3,4,*}, B. H. Sun (孙保华)⁵, W. H. Long (龙文辉)⁶ and Y. F. Niu (牛一斐)^{6,7}

¹School of Physics and Materials Science, Anhui University, Hefei 230601, China

²Institute of Physical Science and Information Technology, Anhui University, Hefei 230601, China

³RIKEN Nishina Center, Wako 351-0198, Japan

⁴Department of Physics, Graduate School of Science, The University of Tokyo, Tokyo 113-0033, Japan

⁵School of Physics and Nuclear Energy Engineering, Beihang University, Beijing 100191, China

⁶School of Nuclear Science and Technology, Lanzhou University, Lanzhou 730000, China

⁷ELI-NP, “Horia Hulubei” National Institute for Physics and Nuclear Engineering, RO-077125, Bucharest-Magurele, Romania

Outline

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

Example: toss coin

- Example: toss 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
- ✓ 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

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

$$p(\omega | D) = \frac{p(D | \omega)p(\omega)}{p(D)}$$

Maximum likelihood method

● 最大似然法

设进行了 N 次试验，正面朝上的次数为 k ，若正面朝上的概率为 ω ，则似然函数为

$$p(N, k | \omega) = C_N^k \omega^k (1 - \omega)^{N-k}$$

$$\frac{dp(N, k | \omega)}{d\omega} = 0 \Rightarrow \omega = \frac{k}{N}$$

例如： $N = 10, k = 6$ ：

✓ $\omega = 0.5$ 时：

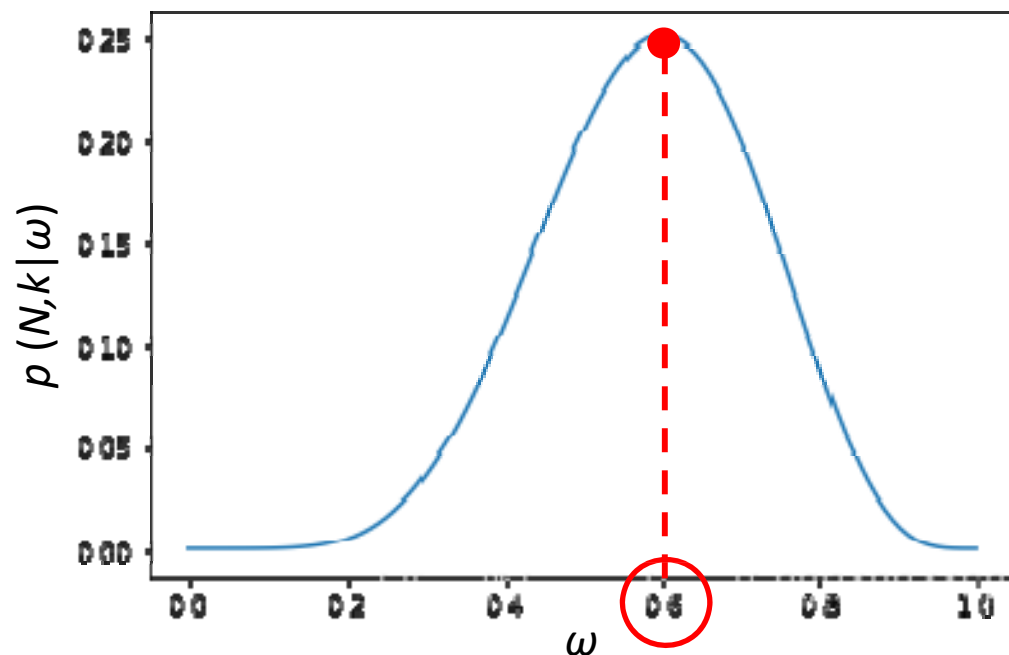
$$p = C_{10}^6 0.5^6 (1 - 0.5)^4 \approx 0.21$$

✓ $\omega = 0.6$ 时：

$$p = C_{10}^6 0.6^6 (1 - 0.6)^4 \approx 0.25$$

✓ others:

见左图



Bayesian method

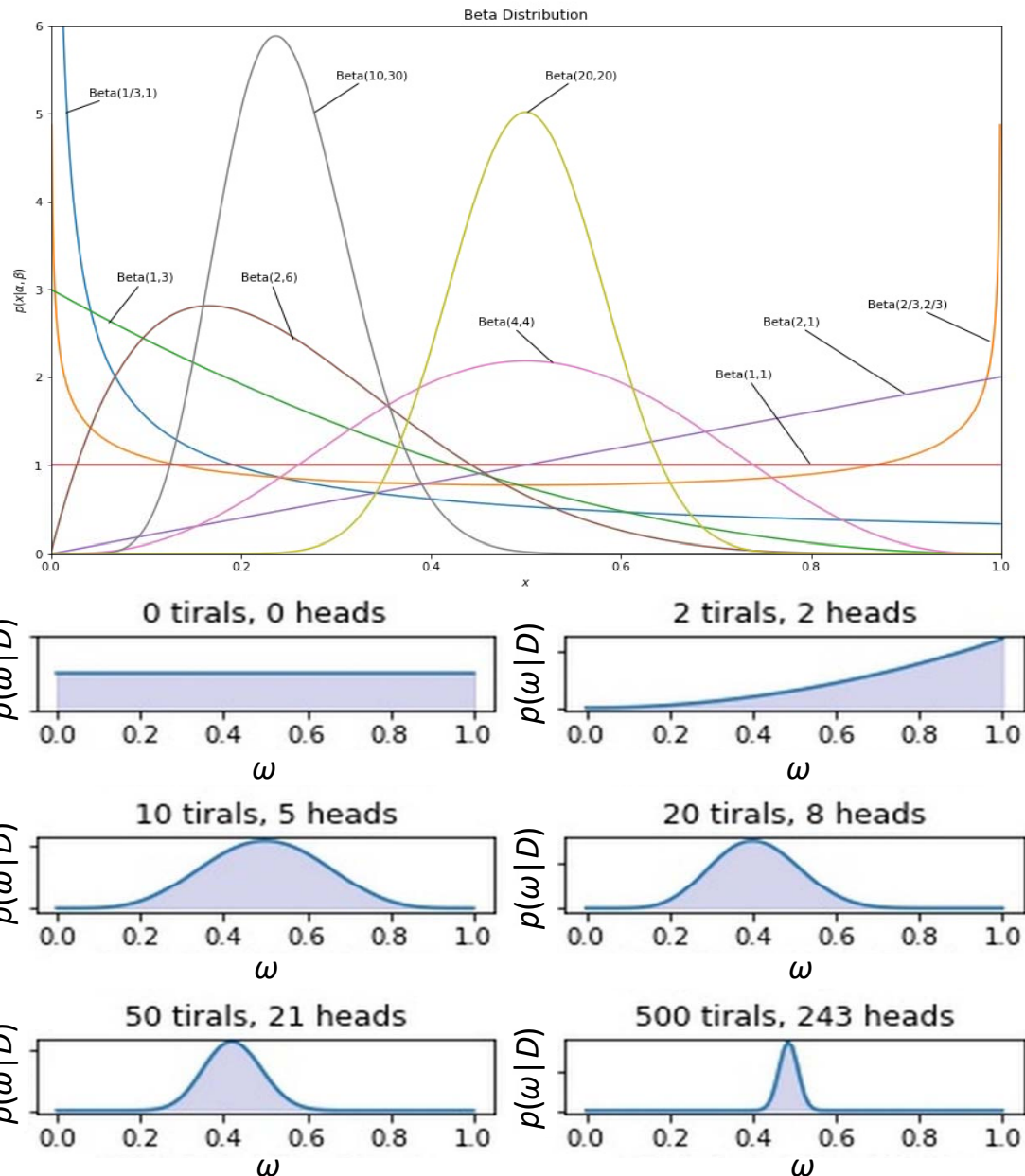
● 贝叶斯方法

$$p(\omega | D) = \frac{p(D | \omega) p(\omega)}{p(D)}$$

$$p(\omega) \sim \text{Beta}(\alpha, \beta)$$

如: $\text{Beta}(1,1) \equiv \text{Uniform}(0,1)$

$$\begin{aligned} p(D | \omega) &= P(N, k | \omega) \\ &= C_N^k \omega^k (1 - \omega)^{N-k} \end{aligned}$$



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**

Influence of prior

Differences between Bayesians and frequentists



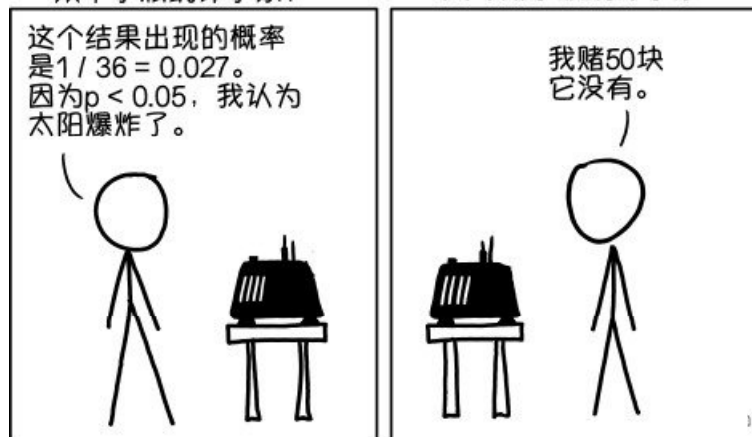
机器说假话的概率：

$$p = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36} \approx 0.028 < 5\%$$

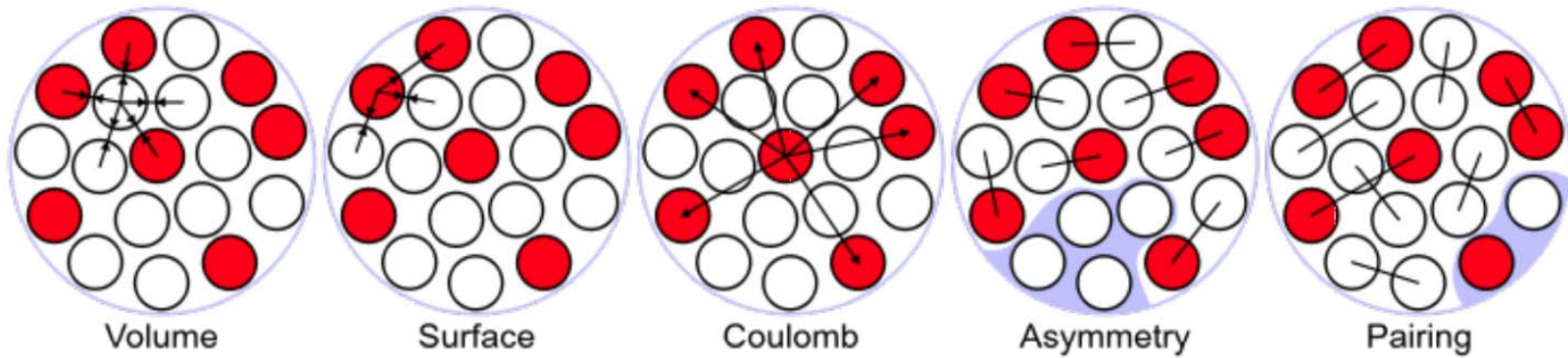
- 频率学派：因 $p < 5\%$ ，位于可置信区间内，所以机器可信，即认为太阳爆炸了
- 贝叶斯学派：
 - ✓ 因为太阳不容易爆炸，假设太阳爆炸的先验概率为0.0000000001
 - ✓ 即使 $p < 5\%$ ，但后验概率仍然很小，所以认为太阳没有爆炸。

频率学派统计学家：

贝叶斯学派统计学家：



Example: LDM



$$B(Z, A) = a_v A - a_{surf} A^{2/3} - a_c \frac{Z^2}{A^{1/3}} - a_{sym} \frac{(N - Z)^2}{A} + a_p \frac{(-1)^Z + (-1)^N}{2\sqrt{A}}$$

➤ Minimize χ^2 or maximize likelihood

$$\chi^2 = \sum_{Z, N \geq 8} \left[\frac{B_{exp}^{Z, N} - B_{LDM}^{Z, N}}{\delta_m} \right]^2$$

likelihood $\propto \exp(-\chi^2 / 2)$



实验数据取自AME2012
CPC 36, 1603 (2012)

a_v	15.5868
a_{surf}	17.0871
a_c	0.7066
a_{sym}	23.1537
a_p	12.2047

Example: LDM

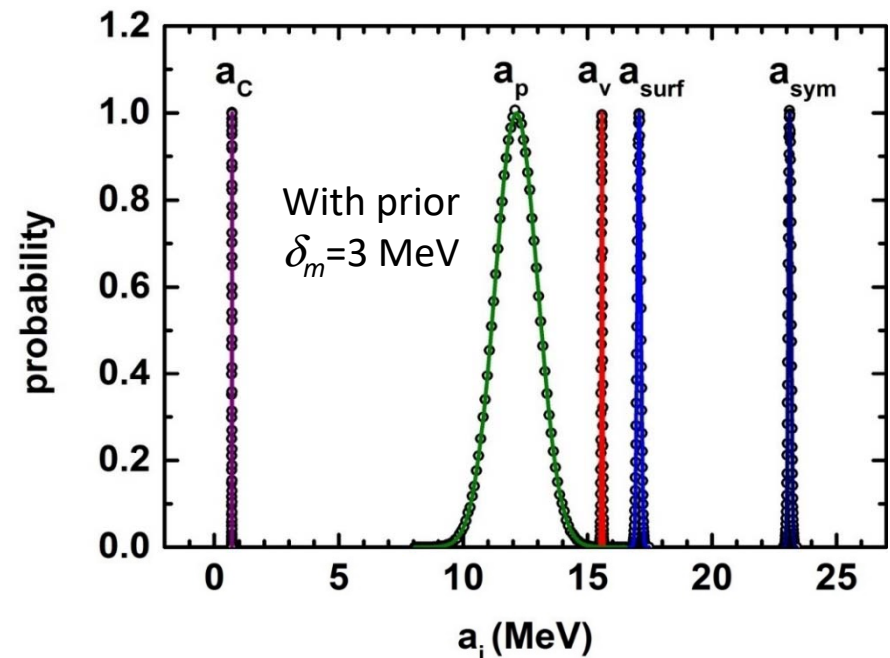
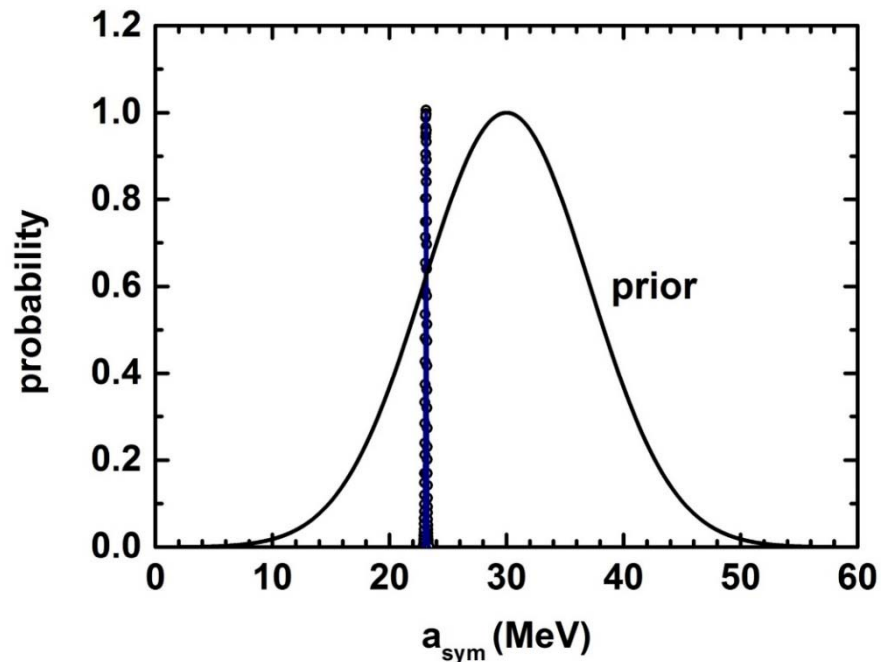
● Bayes' theorem: $p(\omega | D) = p(D | \omega) p(\omega) / p(D)$

$$B(Z, A) = a_v A - a_{surf} A^{2/3} - a_c \frac{Z^2}{A^{1/3}} - a_{sym} \frac{(N-Z)^2}{A} + a_p \frac{(-1)^Z + (-1)^N}{2\sqrt{A}}$$

$$p(\omega) \rightarrow p(a_i) \propto \exp\left[-(a_i - a_{i0})^2 / 2\sigma_{a_{i0}}^2\right] \quad a_{i0} = 16, 20, 0.7, 30, 10; \sigma_{a_{i0}} = 5, 10, 0.5, 10, 5$$

$$p(D | \omega) \rightarrow p(m_{exp} | a_i) \propto \exp(-\chi^2 / 2), \quad \chi^2 = \sum_{Z, N \geq 8} \left[\frac{B_{exp}^{Z, N} - B_{LDM}^{Z, N}}{\delta_m} \right]^2$$

实验数据取自AME2012
CPC 36, 1603 (2012)



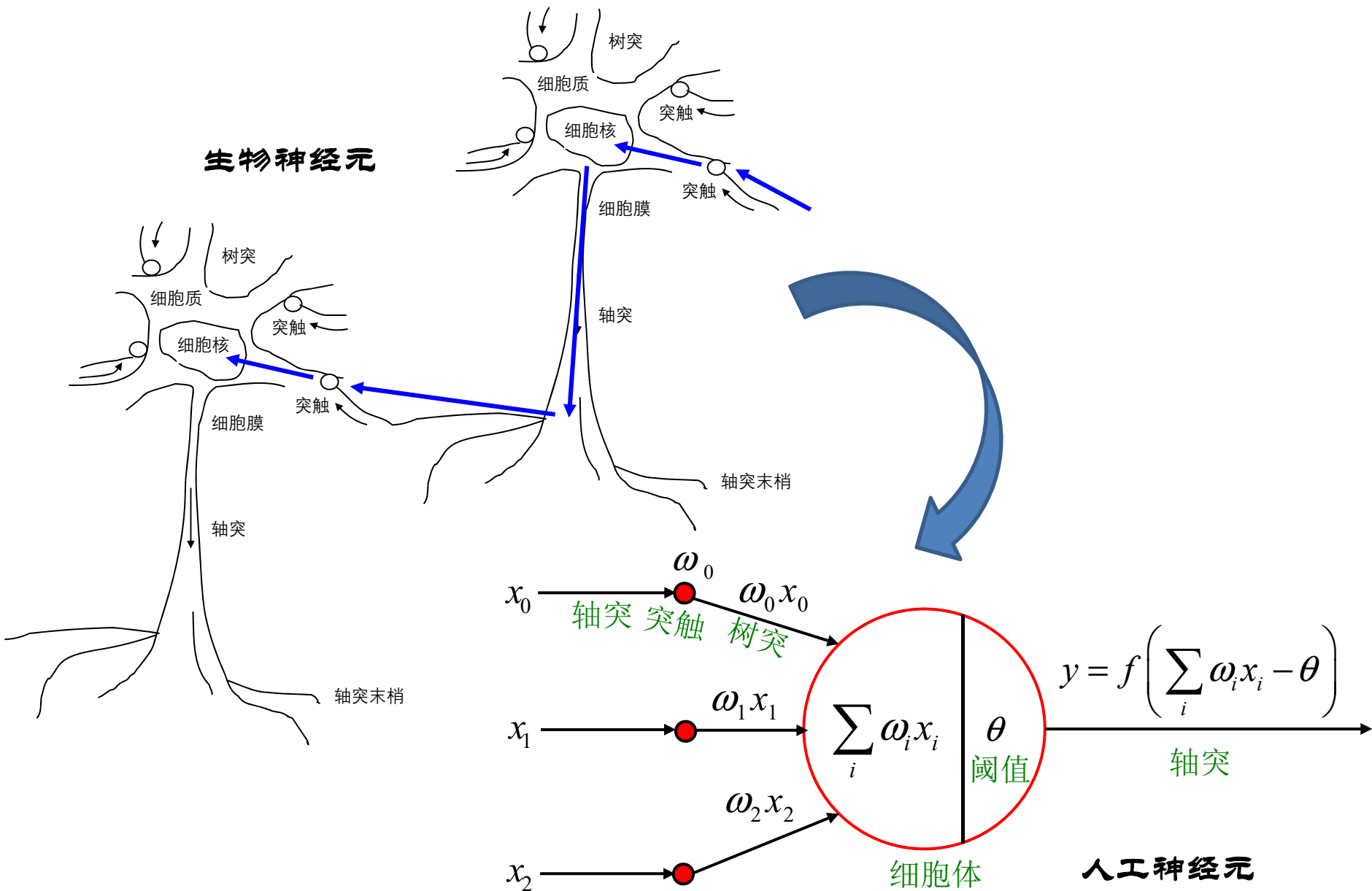
Example: LDM

	With prior ($\delta_m=3$ MeV)		Without prior ($\delta_m=3$ MeV)		With prior ($\delta_m=10$ MeV)		least square method	
a_v	15.5786	± 0.0238	15.5785	± 0.0238	15.5794	± 0.0795	15.5868	± 0.0244
a_{surf}	17.0706	± 0.0742	17.0705	± 0.0742	17.0734	± 0.2474	17.0871	± 0.0758
a_c	0.7054	± 0.0017	0.7054	± 0.0017	0.7054	± 0.0055	0.7066	± 0.0017
a_{sym}	23.125	± 0.0595	23.1249	± 0.0595	23.1275	± 0.1985	23.1537	± 0.0607
a_p	12.1387	± 0.8565	12.2029	± 0.8692	11.6717	± 2.4851	12.2047	± 0.8853

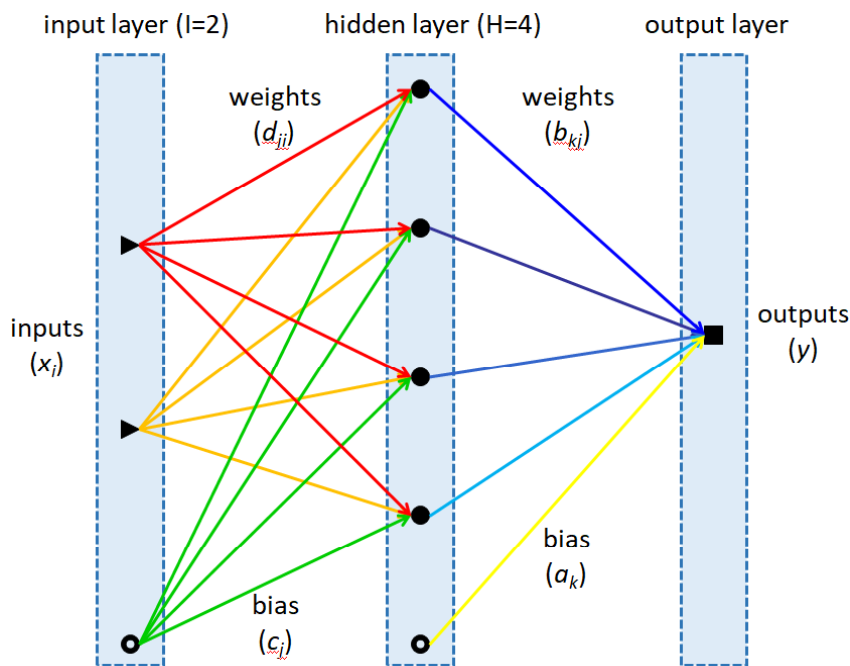
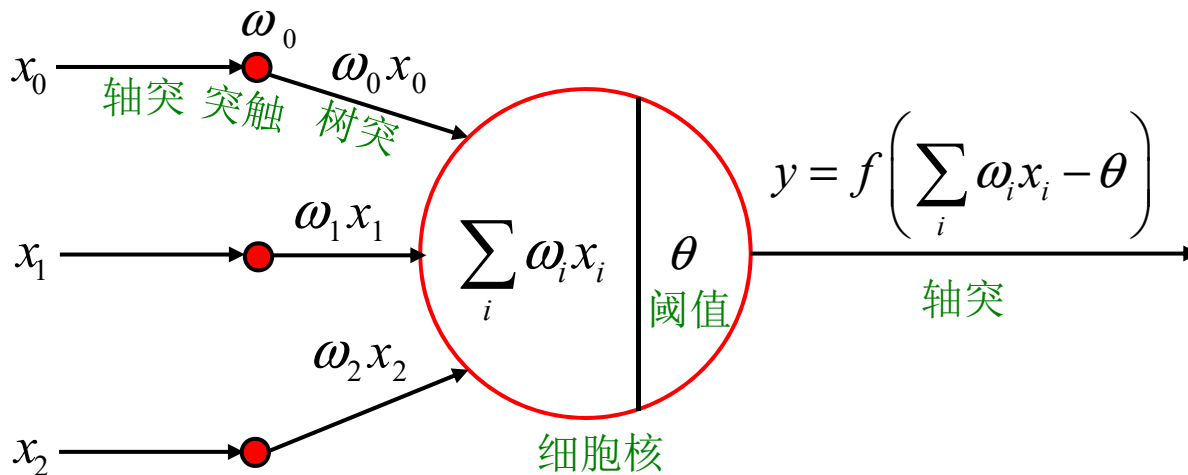
- ✓ When the number of the data is large enough, the influence of prior distribution is almost negligible and the deduced results are similar to those from least square method.
- ✓ If δ_m are very large, the values of a_i are still similar, while their uncertainties would increase.

Neural network

生物神经元

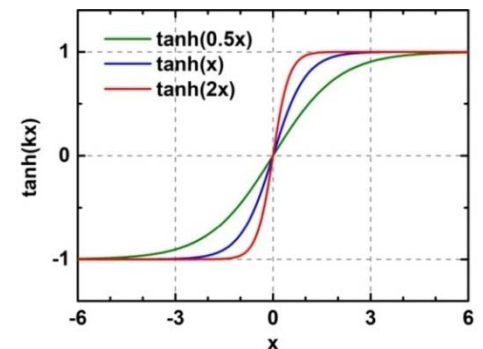


Neural network



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

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



Bayesian approach in regression problem

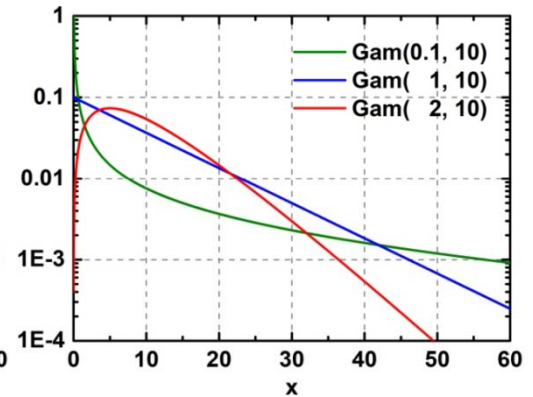
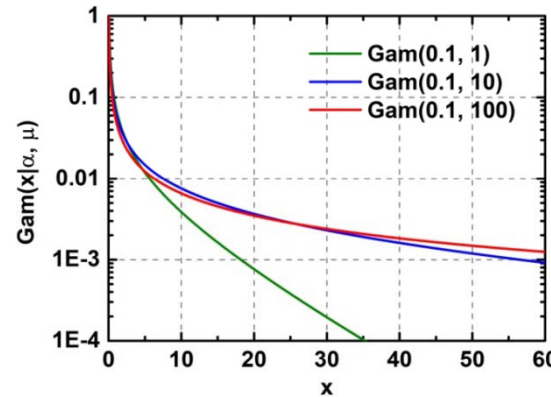
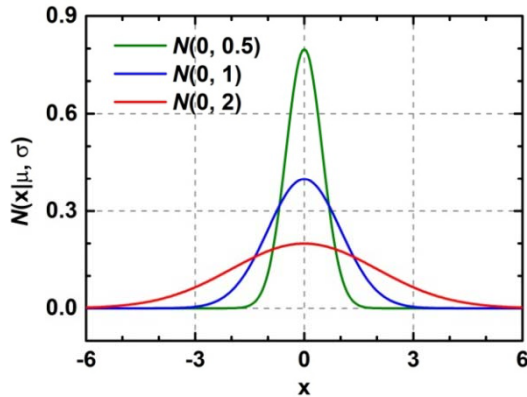
- Posterior distributions of parameters are [Neal1996Springer](#)

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

- ✓ prior distribution $p(\omega)$:

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

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



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

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

Bayesian approach in regression problem

- Posterior distributions of parameters are [Neal1996Springer](#)

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

✓ sampling with Markov chain Monte Carlo (MCMC) method

- Make predictions

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

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

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.

随机变量的抽样

● 单位均匀分布的随机数

最简单且最基本的连续型随机变量的分布是单位均匀分布，即[0, 1]上的均匀分布，其分布密度函数为：

$$f(x) = \begin{cases} 1, & 0 \leq x \leq 1 \\ 0, & \text{其他} \end{cases}$$

- ✓ 随机数表：随机数表由等概率出现的0, ..., 9十个数字组成，依次取出表中 n 个相邻的数字合并在一起作为随机数。
- ✓ 物理方法：利用某些随机物理现象，如计算机固有噪声，通过某些特殊设备，可以在计算机上直接产生随机数。
- ✓ 伪随机数：同余法

$$x_{n+1} = (ax_n + c) \bmod m$$

$$\xi_{n+1} = x_{n+1} / (m - 1)$$

a 、 c 和 m 分别为乘子、增量和模， x_0 为初值。

随机变量的抽样

● 直接抽样法（反函数法）

设随机变量 X 的分布密度函数为 $f(x)$

$$0 \leq F(x) = \int_{-\infty}^x f(t) dt \leq 1$$

$f(x)$ 必须归一！

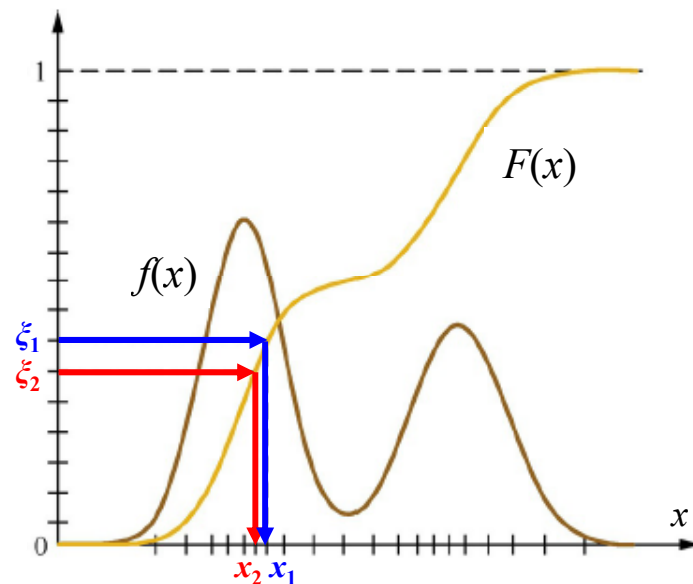
产生 $[0, 1]$ 均匀分布的随机数 ξ ，令 $\xi = F(\eta)$ 或 $\eta = F^{-1}(\xi)$

$$1d\xi = dF(\eta) = F'(\eta)d\eta = f(\eta)d\eta$$

即按反函数计算出的 η 就是服从分布密度函数 $f(x)$ 的随机变量。

➤ 抽样步骤：

- (1) 计算 $f(x)$ 的分布函数 $F(x)$
- (2) 产生 $[0, 1]$ 均匀分布随机数 ξ
- (3) 计算 $F^{-1}(\xi)$ ，令 $\eta = F^{-1}(\xi)$
- (4) 重复(2)和(3)



随机变量的抽样

➤ 例：对指数分布的直接抽样

分布密度函数为：

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0, \lambda > 0 \\ 0, & x \leq 0 \end{cases}$$

积分得到分布函数：

$$F(\eta) = \int_{-\infty}^{\eta} f(x) dx = \int_0^{\eta} \lambda e^{-\lambda x} dx = 1 - e^{-\lambda \eta}$$

令 $\xi = F(\eta) = 1 - e^{-\lambda \eta}$ ，则指数分布随机变量的抽样为：

$$\eta = -\frac{1}{\lambda} \ln(1 - \xi) = -\frac{1}{\lambda} \ln \xi$$

(1-ξ)和ξ同样服从[0, 1]均匀分布

随机变量的抽样

● 变换抽样法

若对 $f(x)$ 的抽样比较复杂，而对 $g(y)$ 的抽样已知且简单。设法找到变换关系： $y=y(x)$ ，使得

$$f(x)dx = g(y)dy$$

按变换 $x=x(y)$ 计算出的 x 就是服从分布密度函数 $f(x)$ 的随机变量。

➤ 抽样步骤：

- (1) 找到变换关系 $y=y(x)$
- (2) 对 $g(y)$ 进行抽样得到 y
- (3) 计算 $x=x(y)$
- (4) 重复(2)和(3)

若 $g(y)$ 为 $[0, 1]$ 的均匀分布，则

$$\begin{aligned} f(x)dx &= dy \\ \Rightarrow dy / dx &= y' = f(x) \\ \Rightarrow y(x) &= \int_{-\infty}^x f(x)dx = F(x) \\ \Rightarrow x &= F^{-1}(y) \end{aligned}$$

即直接抽样法。

计算定积分的蒙特卡罗方法

● 用算术平均代替数学期望

$$J = \int_a^b h(x) dx = \int_a^b \frac{h(x)}{f(x)} f(x) dx$$

$$\text{令 } g(x) = \frac{h(x)}{f(x)} \Rightarrow J = \int_a^b g(x) f(x) dx$$

按照概率密度函数 $f(x)$ 进行随机抽样 x_i ，用数学期望代替积分值

$$J = \int_a^b g(x) f(x) dx = E(g) = \frac{1}{n} \sum_{i=1}^n g(x_i) = \overline{g(x)}$$

✓ 若 $f(x)$ 选为 $[a, b]$ 均匀分布的概率密度函数

$$f(x) = \frac{1}{b-a} \Rightarrow g(x) = \frac{h(x)}{f(x)} = (b-a)h(x)$$

$$\longrightarrow J = \frac{1}{n} \sum_{i=1}^n g(x_i) = \frac{1}{n} \sum_{i=1}^n (b-a)h(x_i) = \frac{b-a}{n} \sum_{i=1}^n h(x_i)$$

随机变量的抽样

● 重要抽样法

$$J = \int_a^b g(x) f(x) dx = \frac{1}{n} \sum_{i=1}^n g(x_i)$$

$$\sigma^2 = \int_a^b [g(x) - J]^2 f(x) dx = \frac{1}{n} \sum_{i=1}^n [g(x_i) - J]^2$$

当 $g(x)$ 在定义域内有显著起伏变化时，方差 σ^2 较大，进而蒙卡积分 J 的误差 σ/\sqrt{N} 较大。

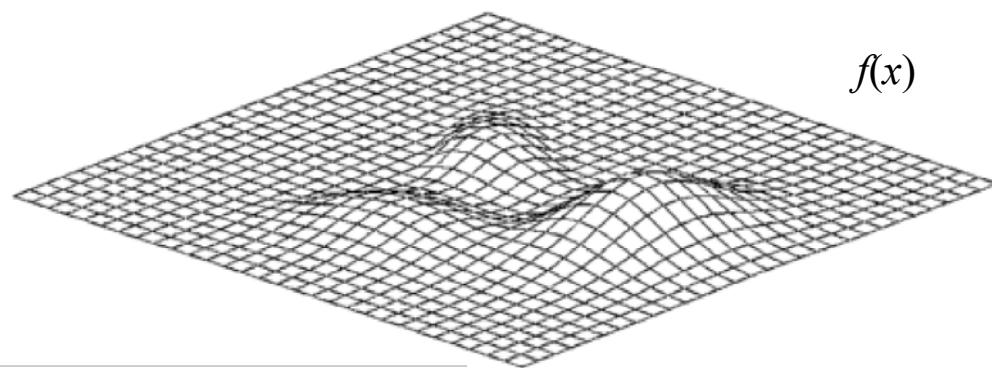
✓ 适当选取偏倚分布密度函数 $f^*(x)$ ，使得 $g^*(x)$ 在定义域内变化比较平坦或者 $g(x)f(x)$ 大的地方 $f^*(x)$ 值也较大。然后产生分布密度函数为 $f^*(x)$ 的随机变量 x_i ，则有

$$J = \int_a^b g(x) f(x) dx = \int_a^b \frac{g(x) f(x)}{f^*(x)} \underbrace{f^*(x)}_{\text{偏倚分布密度函数}} dx = \int_a^b g^*(x) f^*(x) dx = \frac{1}{n} \sum_{i=1}^n g^*(x_i)$$

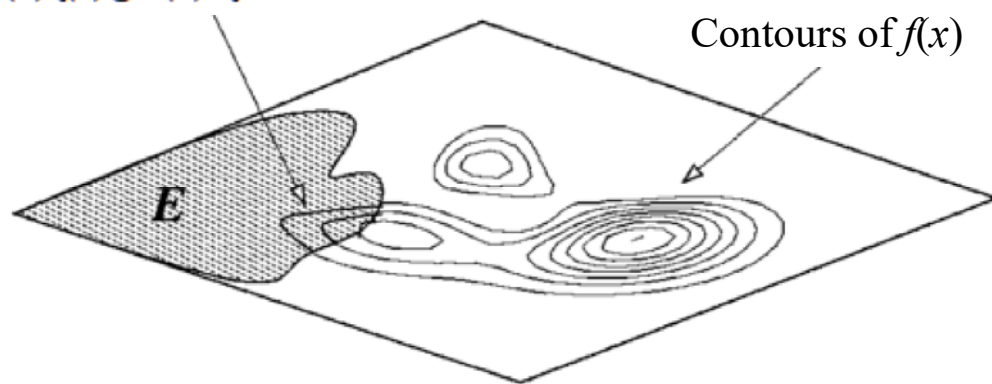
偏倚分布密度函数

随机变量的抽样

$$J = \int_a^b g(x) f(x) dx = \int_a^b \frac{g(x) f(x)}{f^*(x)} f^*(x) dx = \int_a^b g^*(x) f^*(x) dx = \frac{1}{n} \sum_{i=1}^n g^*(x_i)$$



Contours of $f(x)$



$$\text{例 } J = \int_0^{\infty} e^{-x/\lambda} f(x) dx$$

取 $f^*(x)$ 为指数分布 $e^{-x/\lambda} / \lambda$

$$J = \int_0^{\infty} \frac{e^{-x/\lambda} f(x)}{f^*(x)} f^*(x) dx$$

$$= \lambda \int_0^{\infty} f(x) f^*(x) dx$$

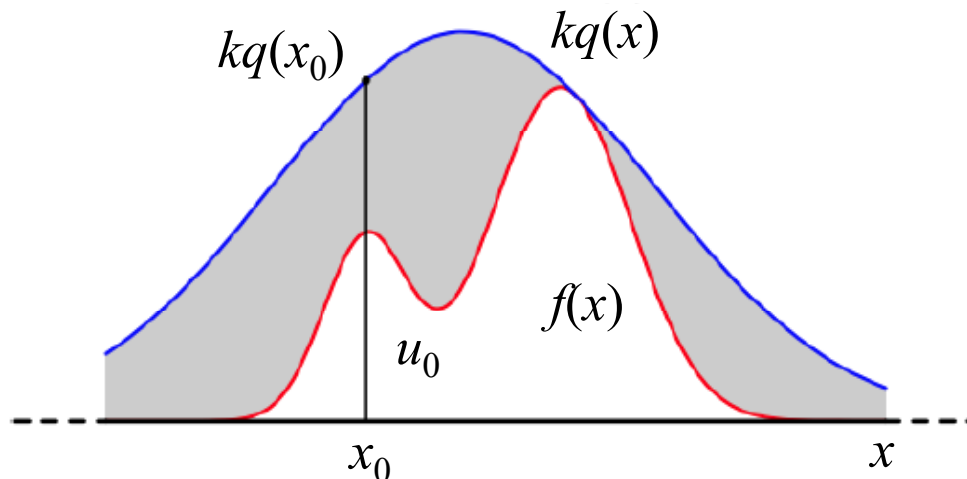
按照 $f^*(x)$ 抽样 x_i

$$J = \frac{\lambda}{n} \sum_{i=1}^n f(x_i)$$

随机变量的抽样

● 拒绝抽样法

若 $f(x)$ 太复杂无法直接抽样，可以设定一个可抽样的分布 $q(x)$ ，如高斯分布，然后按照一定规则拒绝某些样本，实现对 $f(x)$ 的抽样



➤ 抽样步骤：

- (1) 确定 $q(x)$ 和常量 k ，使得 $f(x)$ 总在 $kq(x)$ 的下方
- (2) x 轴的方向：从 $q(x)$ 分布抽样取得 x_0
- (3) y 轴的方向：从 $[0, kq(x_0)]$ 均匀分布中抽样得到 u_0
若 $u_0 > f(x_0)$ ，拒绝抽样 x_0 ；否则接受 x_0
- (4) 重复(2)和(3)

Markov chain

- **马尔科夫链**: 设 $\{X_n, n=1, 2, \dots\}$ 是一个随机序列, 状态空间 E 为有限或可列集, 对于任意的正整数 n , 若 $j, i_k \in E (k=1, \dots, n-1, n)$, 有

$$P\{X_{n+1} = j | X_n = i_n, X_{n-1} = i_{n-1}, \dots, X_1 = i_1\} = P\{X_{n+1} = j | X_n = i_n\}$$

则称 $\{X_n, n=1, 2, \dots\}$ 为马尔科夫链, $P\{X_{n+1}=j | X_n=i_n\}$ 称为转移概率。

- **齐次马氏链**: 设 $\{X_n, n=1, 2, \dots\}$ 是一个马氏链, 若 $P\{X_{n+1}=j | X_n=i_n\}$ 与 n 无关, 即

$$P\{X_{n+1} = j | X_n = i_n\} = p_{ij}$$

则称 $\{X_n, n=1, 2, \dots\}$ 为时齐的马氏链 (也称时齐马氏链), p_{ij} 为系统由状态 i 转移到状态 j 的转移概率。

- **转移矩阵**: 由转移概率 p_{ij} 组成的矩阵 P 称为转移矩阵, 其具有性质

$$(1) \text{ 对一切 } i, j \in E, 0 \leq p_{ij} \leq 1$$

$$(2) \text{ 对一切 } i \in E, \sum_{j \in E} p_{ij} = 1$$

Markov chain

● 齐次马氏链的遍历性：设齐次马氏链 $\{X_n, n=1, 2, \dots\}$ 的状态空间为 $E=\{1, 2, \dots\}$ ，若对于一切状态 $i, j \in E$ ，存在不依赖于 i 的常数 π_j ，为其转移概率的极限，即

$$\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \pi_j, \quad i, j \in E$$

$$P^{(n)} = P^n = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1j} & \dots \\ p_{21} & p_{22} & \dots & p_{2j} & \dots \\ \dots & \dots & \dots & \dots & \dots \\ p_{i1} & p_{i2} & \dots & p_{ij} & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix} \xrightarrow{n \rightarrow \infty} \begin{pmatrix} \pi_1 & \pi_2 & \dots & \pi_j & \dots \\ \pi_1 & \pi_2 & \dots & \pi_j & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \pi_1 & \pi_2 & \dots & \pi_j & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

则称此齐次马氏链具有遍历性。若 $\pi_j, j \in E$ 满足

$$\pi_j \geq 0, \quad \sum_j \pi_j = 1$$

则称其为转移概率的极限分布。备注：有限状态的遍历马氏链必存在极限分布，无限（即可列）状态的遍历马氏链不一定存在极限分布，只有其极限概率构成概率分布时才存在极限分布。

Markov chain

● 齐次马氏链的平稳分布：设 $\{X_n, n=1, 2, \dots\}$ 为一齐次马氏链，若存在实数集合 $\{\pi_j, j \in E\}$ ，满足

$$\begin{cases} \pi_j \geq 0, j \in E \\ \sum_{j \in E} \pi_j = 1 \end{cases}$$

则称 $\{\pi_j, j \in E\}$ 为概率分布。如果此概率分布满足

$$\pi_j = \sum_{i \in E} \pi_i p_{ij}, j \in E$$

则 $\{X_n, n \geq 0\}$ 是一平稳齐次马氏链，并称 $\{\pi_j, j \in E\}$ 为该过程的一个平稳分布。备注：有限马氏链转移概率的极限分布一定是平稳分布，因此有限状态遍历马氏链的极限分布就是平稳分布；但无限马氏链转移概率的极限分布不一定是平稳分布。

Markov chain

例：设齐次马氏链的状态空间为 $E=\{1, 2\}$ ，其转移概率矩阵为

$$P = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

可见 $P^n=P$ ，但 P^n 某一系列的所有元素并不是同一值，所以此马氏链不是遍历的。设有 π_1, π_2 ，满足 $\pi_1+\pi_2=1$ 且 $0<\pi_1, \pi_2<1$ ，有

$$(\pi_1 \quad \pi_2) = (\pi_1 \quad \pi_2) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

可见，平稳分布是存在的，而且有无穷多个。

备注：虽然遍历的马氏链具有平稳性，但是平稳的马氏链不一定具有遍历性，即不遍历的马氏链也可以具有平稳性。

Markov chain

● **遍历定理1**: 对于有限状态的齐次马氏链 $\{X_n, n=1, 2, \dots\}$, 设状态空间为 $E=\{1, 2, \dots, k\}$, 若存在正整数 m , 对任意状态 $i, j \in E$, 其 m 步的转移概率均大于0, 即

$$p_{ij}^{(m)} > 0$$

则此马氏链具有遍历性, 且 $\{\pi_j\}=\{\pi_1, \pi_2, \dots, \pi_k\}$ 是方程组

$$\pi_j = \sum_i \pi_i p_{ij}$$

满足条件 $\sum_j \pi_j = 1, \pi_j > 0$ 的唯一解。

备注: 利用该定理, 可以判断有限状态齐次马氏链的遍历性以及求出稳态概率 π_j , 此时稳态概率即为平稳分布。

Markov chain

例：设齐次马氏链的状态空间为 $E=\{1, 2, 3\}$ ，其转移概率矩阵为

$$P = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix}$$

试问此链是否具有遍历性？若有，试求其稳态概率。

$$\text{解： } P^{(2)} = P^2 = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix} \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix} = \begin{pmatrix} 1/2 & 1/4 & 1/4 \\ 1/4 & 1/2 & 1/4 \\ 1/4 & 1/4 & 1/2 \end{pmatrix}$$

可见所有的二步转移概率均大于0，由上面定理可知，此链具有遍历性。再由转移概率与稳态概率满足的方程组得

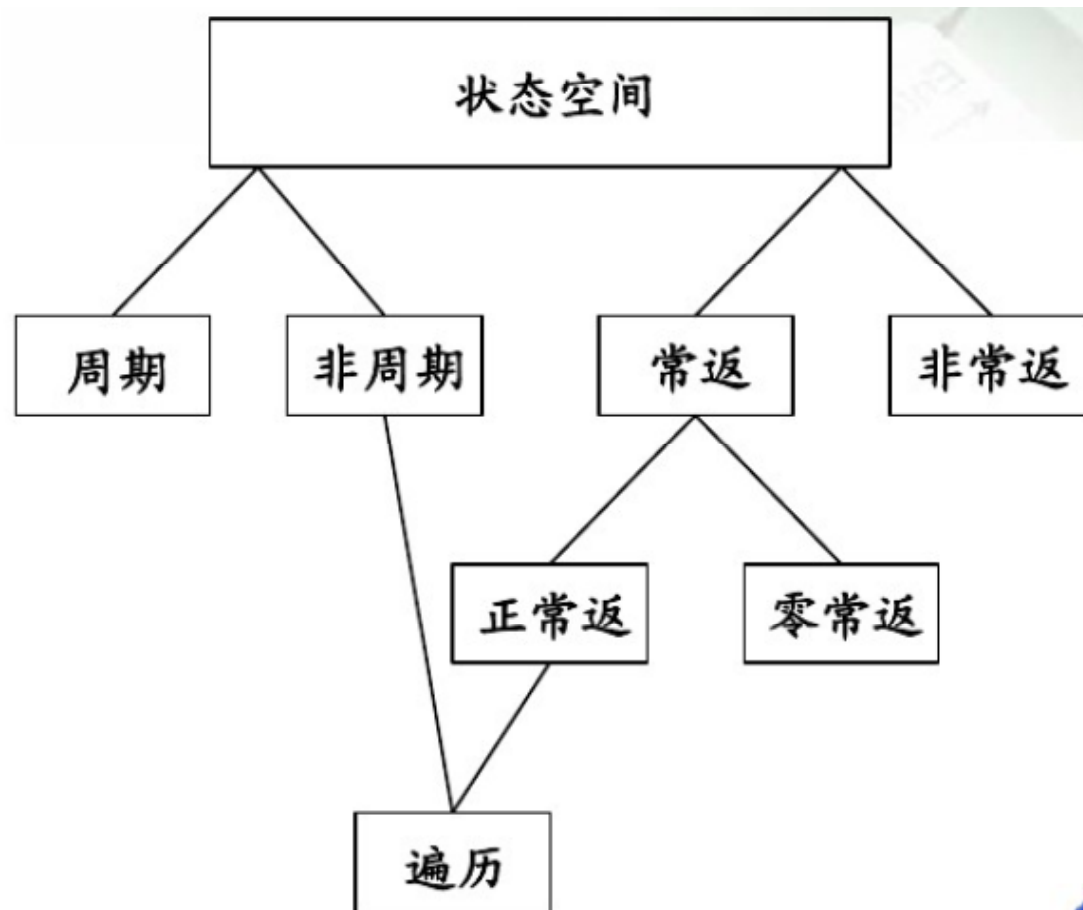
$$\begin{cases} \pi_1 = \pi_1 \frac{1}{2} + \pi_2 \frac{1}{2} + \pi_3 0 \\ \pi_2 = \pi_1 \frac{1}{2} + \pi_2 0 + \pi_3 \frac{1}{2} \\ \pi_3 = \pi_1 0 + \pi_2 \frac{1}{2} + \pi_3 \frac{1}{2} \end{cases} \quad \text{且} \quad \begin{cases} \pi_1, \pi_2, \pi_3 > 0 \\ \pi_1 + \pi_2 + \pi_3 = 1 \end{cases} \quad \text{解之可得稳态概率 } \pi_1 = \pi_2 = \pi_3 = \frac{1}{3}$$

Markov chain

遍历定理2: 对于不可约且非周期的可列状态的齐次马氏链 $\{X_n, n=1, 2, \dots\}$, 设状态空间为 $E=\{1, 2, \dots\}$, 其存在平稳分布的充要条件是, 这个链所有状态都是正常返的, 并且此时极限分布

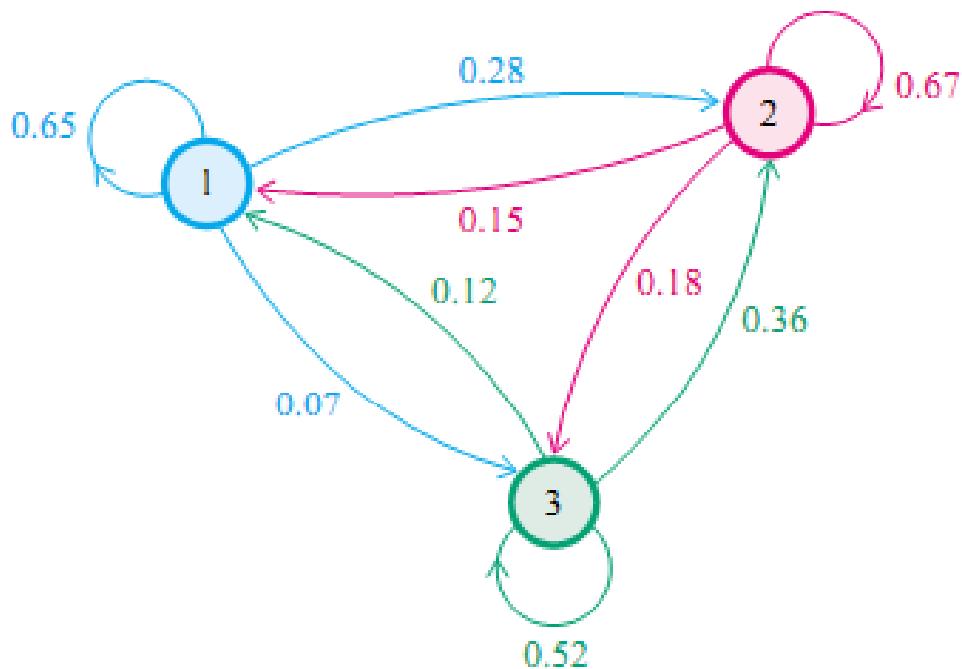
$$\pi_j = \lim_{n \rightarrow \infty} p_{ij}^{(n)} > 0$$

是唯一的平稳分布。注意: 对于有限状态的齐次马氏链 $\{X_n, n=1, 2, \dots\}$, 设状态空间为 $E=\{1, 2, \dots, k\}$, 若其是不可约的, 则其所有的状态都是正常返的, 因此, 其存在平稳分布的充要条件仅为其是不可约且非周期的马氏链。



Markov chain

		子代		
State		1	2	3
父代	1	0.65	0.28	0.07
	2	0.15	0.67	0.18
	3	0.12	0.36	0.52



转移概率矩阵

$$Q = \begin{bmatrix} 0.65 & 0.28 & 0.07 \\ 0.15 & 0.67 & 0.18 \\ 0.12 & 0.36 & 0.52 \end{bmatrix}$$

假设当代人处在下层、中层、上层的比例是

$$p_0 = [p_0(1), p_0(2), p_0(3)]$$

那么他们的子代、孙子代、...、第 n 代的分布比例将是

$$p_1 = p_0 Q$$

$$p_2 = p_1 Q = p_0 Q^2$$

...

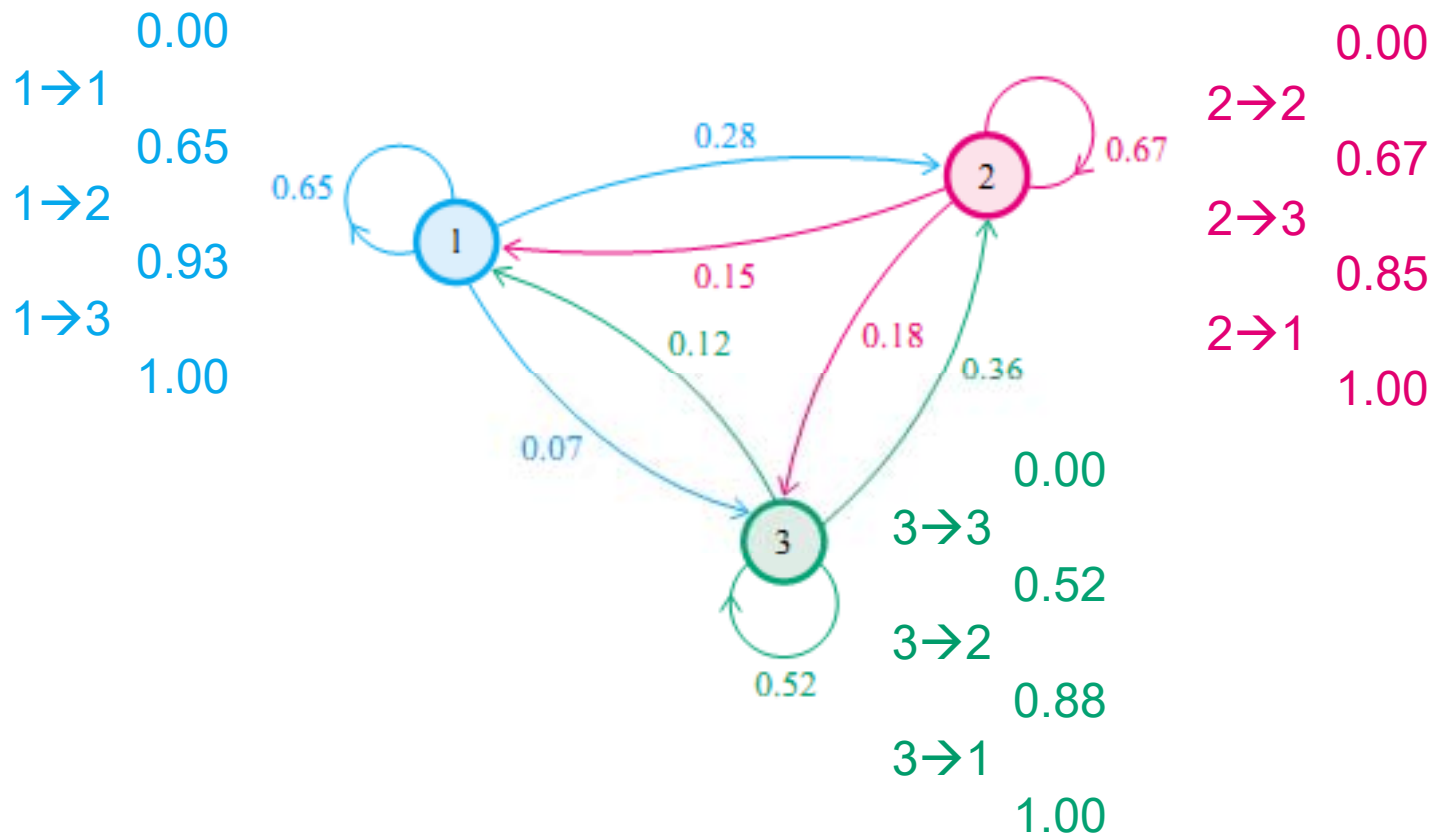
$$p_n = p_{n-1} Q = p_0 Q^n$$

Markov chain

第 n 代人	下层	中层	上层	第 n 代人	下层	中层	上层
0	0.210	0.680	0.110	0	0.75	0.15	0.1
1	0.252	0.554	0.194	1	0.522	0.347	0.132
2	0.270	0.512	0.218	2	0.407	0.426	0.167
3	0.278	0.497	0.225	3	0.349	0.459	0.192
4	0.282	0.490	0.226	4	0.318	0.475	0.207
5	0.285	0.489	0.225	5	0.303	0.482	0.215
6	0.286	0.489	0.225	6	0.295	0.485	0.220
7	0.286	0.489	0.225	7	0.291	0.487	0.222
8	0.289	0.488	0.225	8	0.289	0.488	0.225
9	0.286	0.489	0.225	9	0.286	0.489	0.225
10	0.286	0.489	0.225	10	0.286	0.489	0.225
...

即使初始状态不同，但当 n 足够大时， Q^n 矩阵的每一行都稳定地收敛到 $[0.286, 0.489, 0.225]$ 这个概率分布。

Markov chain



1—(0.61)—1—(0.37)—1—(0.61)—1—(0.68)—2—(0.34)—2—(0.61)—2—
 (0.88)—1—(0.97)—3—(0.66)—2—(0.55)—2—(0.87)—1—...

抽样: 1, 1, 1, 1, 2, 2, 2, 1, 3, 2, 2, 1, (1, 2, 3出现的概率分别为0.286, 0.489, 0.225)

Markov Chain Monte Carlo

- 马尔科夫链蒙特卡罗方法（MCMC）的基本思想：
 - ✓ 给定待采样的概率分布 $p(x)$
 - ✓ 构造马尔可夫链的转移矩阵 Q ，使该马尔可夫链的平稳分布为 $p(x)$
 - ✓ 随机选取初始状态 x_0
 - ✓ 进行马尔科夫过程转移，得到一系列状态值： $\{x_0, x_1, \dots, x_n, x_{n+1}, \dots\}$
 - ✓ 检查马尔科夫过程收敛的步数，若为第 n 步，那么 $\{x_n, x_{n+1}, \dots\}$ 即为分布 $p(x)$ 的样本

Markov Chain Monte Carlo

- 细致平稳条件：如果非周期马氏链的转移矩阵 Q 和分布 $p(x)$ 满足：

$$p(i)q(i, j) = p(j)q(j, i)$$

则 $p(x)$ 就是马氏链的平稳分布， Q 矩阵元 $q(i, j)$ 表示从状态 i 到 j 的转移概率，上式被称为细致平稳条件（detailed balance condition）。

- ✓ **Metropolis-Hastings**方法：设有转移矩阵为 Q 的马氏链，一般来说

$$p(i)q(i, j) \neq p(j)q(j, i)$$

即细致平稳条件不成立，所以 $p(x)$ 不可能是这个马氏链的平稳分布。为了满足细致平衡条件，引入接受率 $\alpha(i, j)$

$$\begin{cases} \alpha(i, j) = p(j)q(j, i) \\ \alpha(j, i) = p(i)q(i, j) \end{cases} \Rightarrow p(i) \underbrace{q(i, j)\alpha(i, j)}_{q'(i, j)} = p(j) \underbrace{q(j, i)\alpha(j, i)}_{q'(j, i)}$$

于是，转移矩阵 Q' 的马氏链恰好满足细致平稳条件，即马氏链 Q' 的平稳分布就是 $p(x)$ 。

Metropolis et al., J. Chem. Phys. 21, 1087 (1953)
Hastings, Biometrika. 57, 97 (1970)

Markov Chain Monte Carlo

? 若接受率 $\alpha(i, j)$ 过小，采样过程由于拒绝率过高，容易原地踏步，从而使得马氏链遍历所有空间耗时太长，采样效率过低。如： $\alpha(i, j)=0.1, \alpha(j, i)=0.2$ ，则有

$$p(i)q(i, j) \times 0.1 = p(j)q(j, i) \times 0.2$$

上式两边扩大5倍

$$p(i)q(i, j) \times 0.5 = p(j)q(j, i) \times 1$$

此时仍满足细致平稳条件，但提高了接受率。因此，为提高接受率，可以将 $\alpha(i, j)$ 和 $\alpha(j, i)$ 同比例放大，使其中较大的一个放大至1，即令

$$\alpha(i, j) = \min \left\{ \frac{p(j)q(j, i)}{p(i)q(i, j)}, 1 \right\} \text{ 或 } \min \left\{ \frac{p(j)q(i | j)}{p(i)q(j | i)}, 1 \right\}$$

不再要求 $p(x)$ 为归一化的概率分布

Markov Chain Monte Carlo

➤ Metropolis-Hastings抽样步骤:

- ✓ 选择提议分布 q
- ✓ 初始化马尔科夫链初始状态为 $X_0 = x_0$
- ✓ 对 $t=0, 1, 2, \dots$, 循环以下过程进行抽样
 - 第 t 个时刻马氏链状态为 $X_t=x_t$, 采样 $x^* \sim q(x|x_t)$
 - 从 $[0, 1]$ 均匀分布采样 u
 - 如果 $u < \alpha(x_t, x^*) = \min \left\{ \frac{p(x^*)q(x_t | x^*)}{p(x_t)q(x^* | x_t)}, 1 \right\}$, 则接受转移, 即 $X_{t+1}=x^*$; 否则不接受转移, 即 $X_{t+1}=x_t$

备注: 如提议分布 q 取高斯分布, 常将 x_t 作为高斯分布的均值, 其标准差 σ 就是采样步长。 σ 过小, α 将较大, 但由于步长太小采样较慢; σ 过大, 虽然步长较大, 但由于 α 将过小采样效率也较低。通常来说 α 在0.1-0.4左右采样效率较高(0.234), 常根据 α 来选择合理的采样步长。

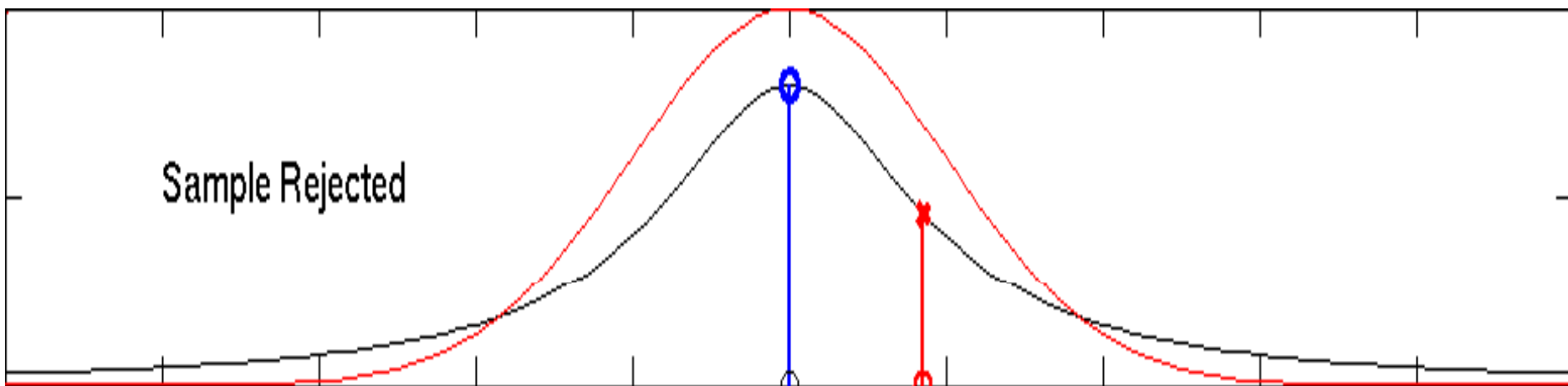
Markov Chain Monte Carlo

采样过程:

$$x_0 \rightarrow \begin{cases} x_* \in N(x_0, \sigma) \\ \alpha(x_0, x_*) \\ u \in U(0, 1) \end{cases} \rightarrow x_1 = \begin{cases} x_0, & u > \alpha(x_0, x_*) \\ x_*, & u \leq \alpha(x_0, x_*) \end{cases}$$

$$x_1 \rightarrow \begin{cases} x_* \in N(x_1, \sigma) \\ \alpha(x_1, x_*) \\ u \in U(0, 1) \end{cases} \rightarrow x_2 = \begin{cases} x_1, & u > \alpha(x_1, x_*) \\ x_*, & u \leq \alpha(x_1, x_*) \end{cases}$$

.....



Outline

- 1 Introduction
- 2 Bayesian neural network approach
- 3 Results and discussion**
 - ★ Toy model
 - ★ Nuclear masses
 - ★ Nuclear β -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: $N=61, x \in [-3, 3]$

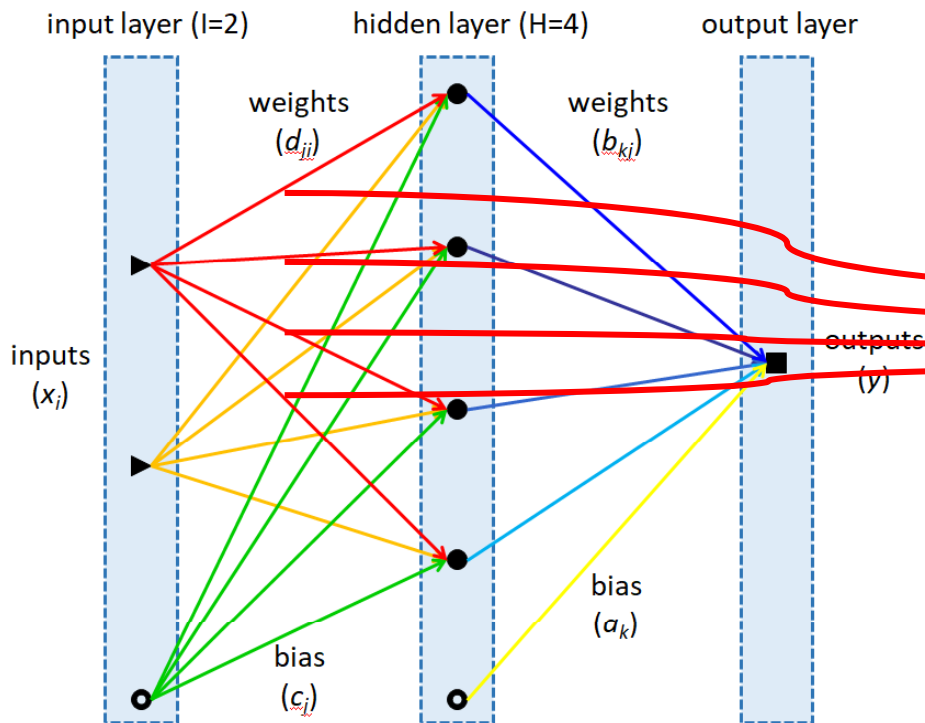
1 input : $y = f(x)$

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

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

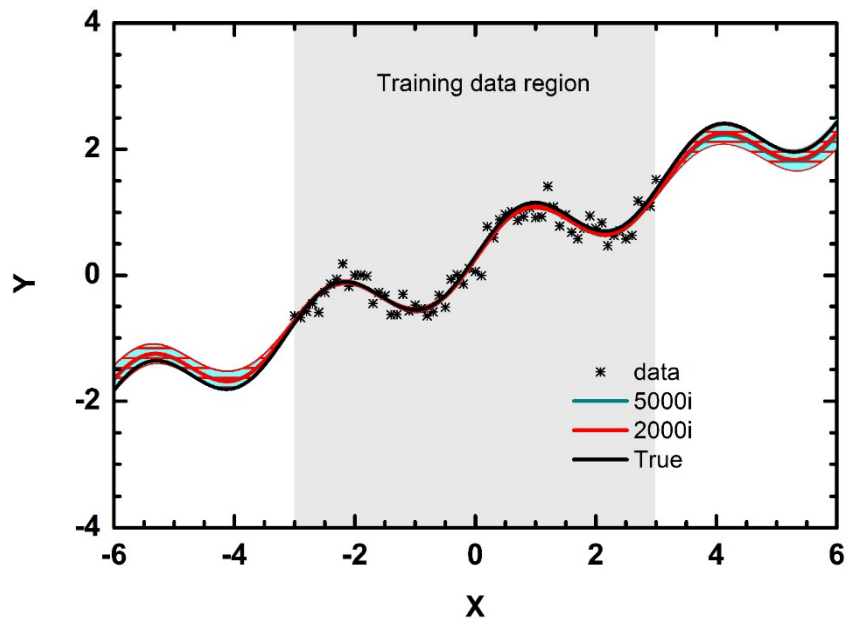
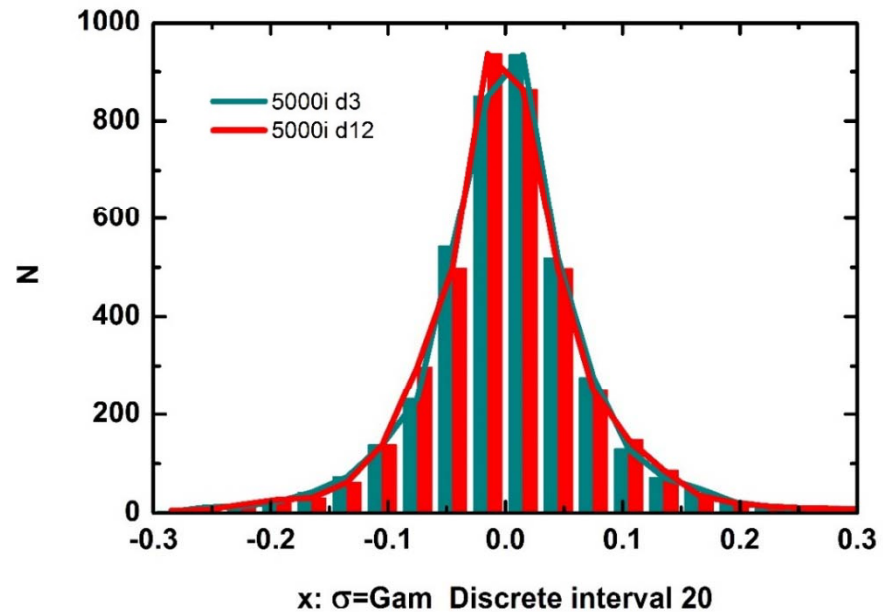
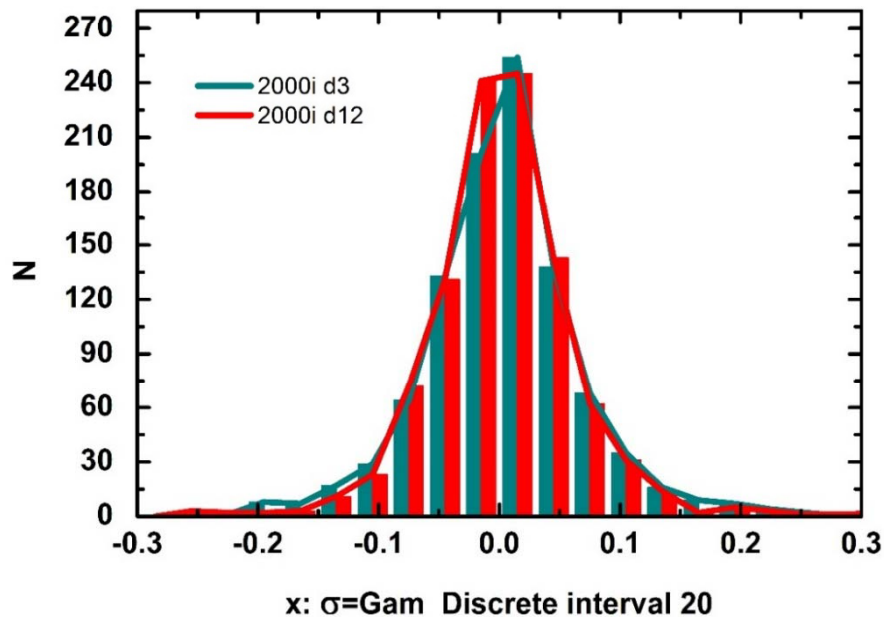
➤ Number of parameters: 61

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



$d_{11}, d_{21}, d_{31},$
 d_{41}, \dots should
have the same
probability
distribution!!!

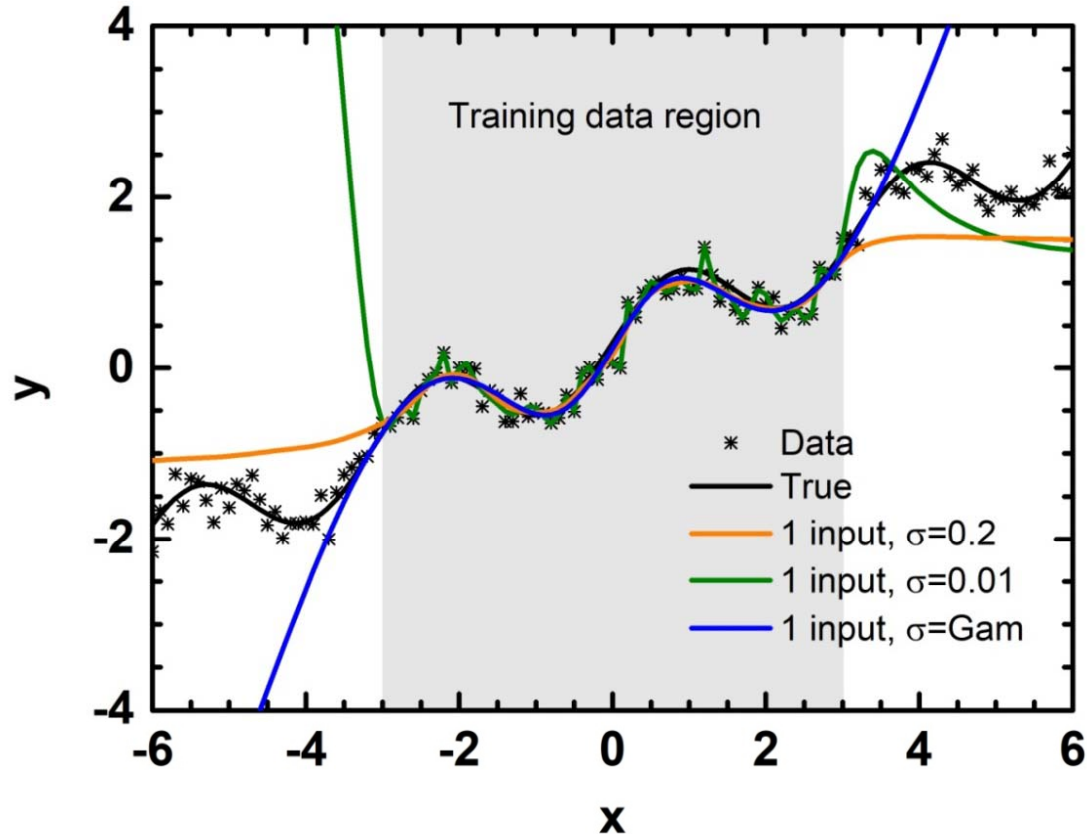
Toy model



✓ 2000步迭代之后，不同 d 参数分布几乎一致。

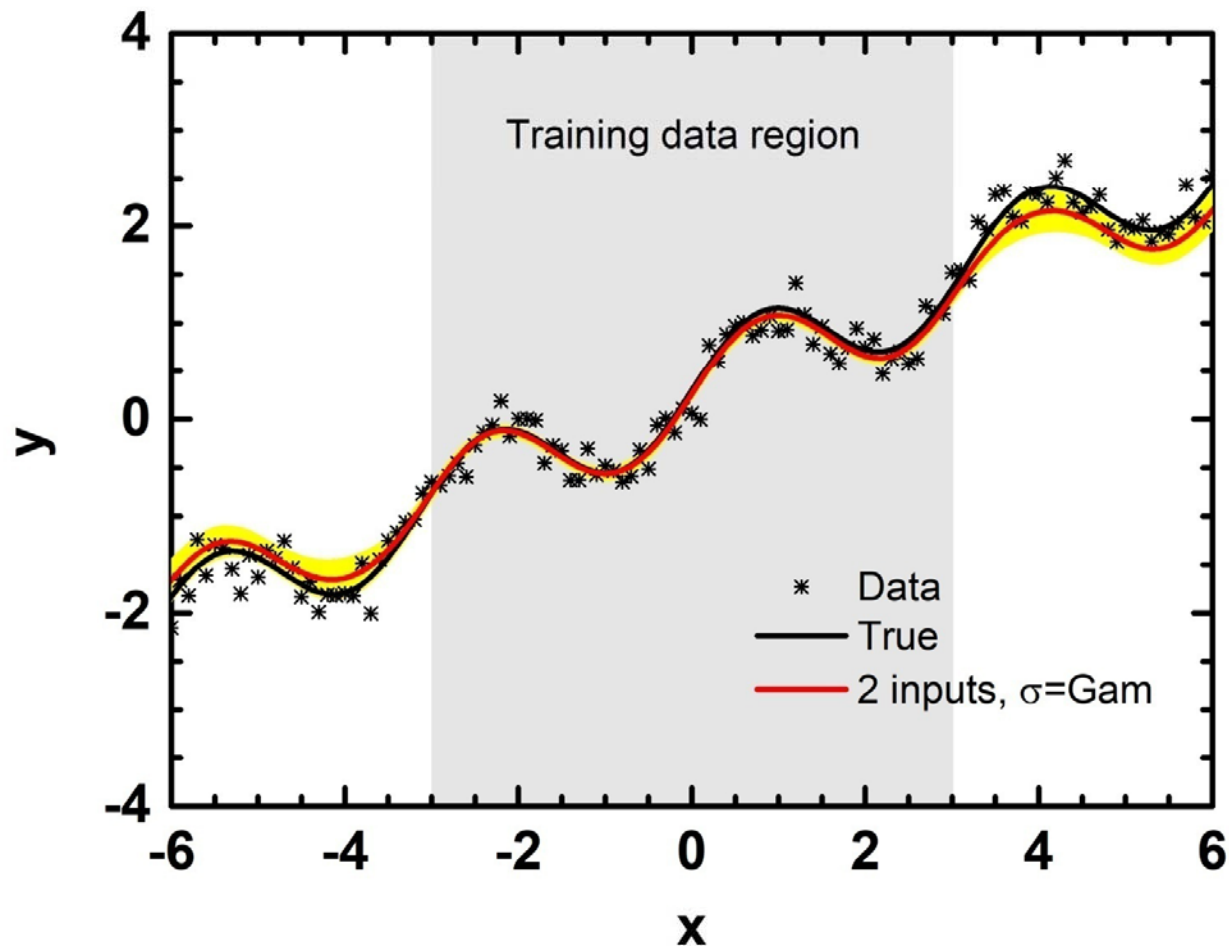
✓ 2000步与5000步采样所得分布及预言结果几乎一致。

Toy model



- ✓ 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.

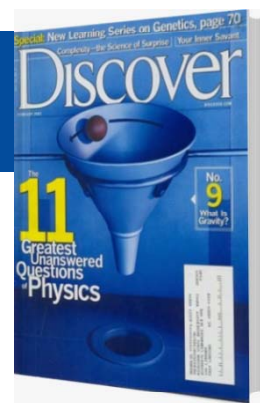
Outline

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

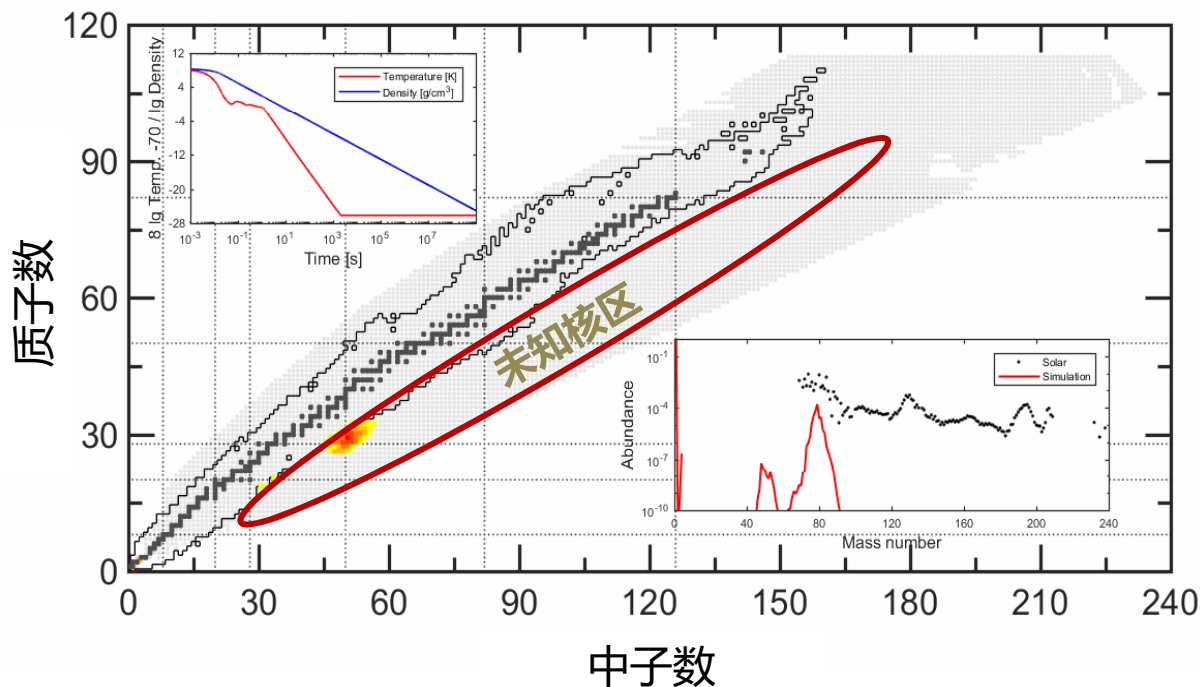
Origin of elements

比铁重的元素是如何合成的？

《发现》杂志：
“物理学十一大待解之谜”



● 主要难点：快中子俘获过程 (r-过程)



r-过程关键核物理量：

- ✓ 原子核质量决定了r-过程路径
- ✓ β 衰变寿命决定了r-过程时标



原子核质量与 β 衰变寿命的精确理论预言
是理解r-过程的关键

Influence of masses

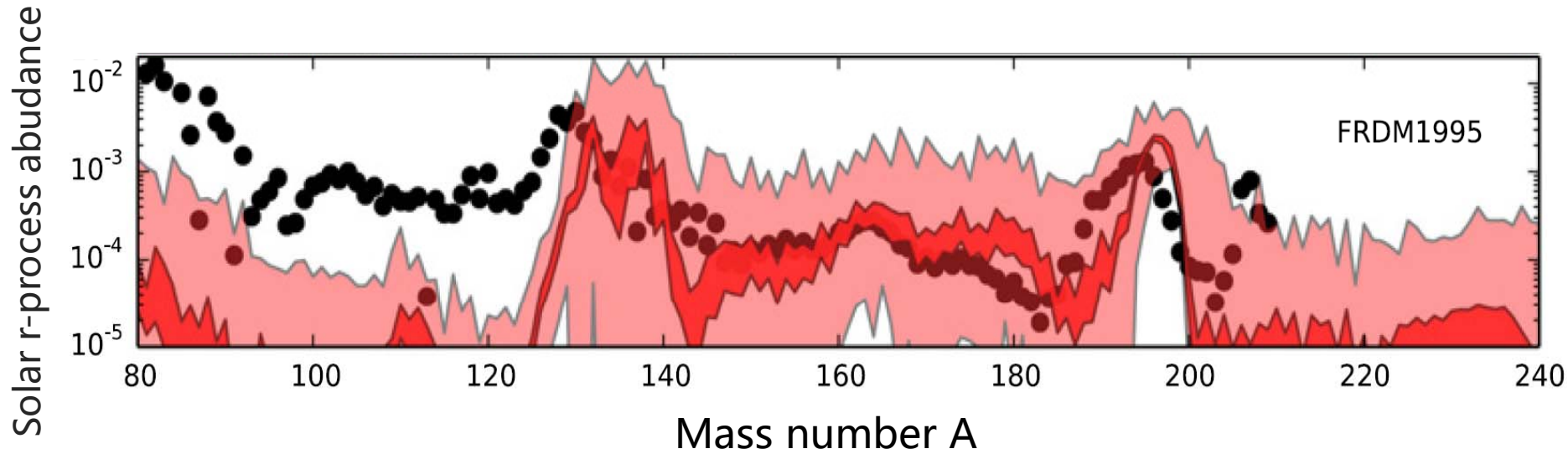


Figure: Variance in isotopic abundance patterns from FRDM1995 mass model predictions compared to the solar data (dots). [M.R. Mumpower *et al.*, PPNP 86, 86 \(2016\)](#)

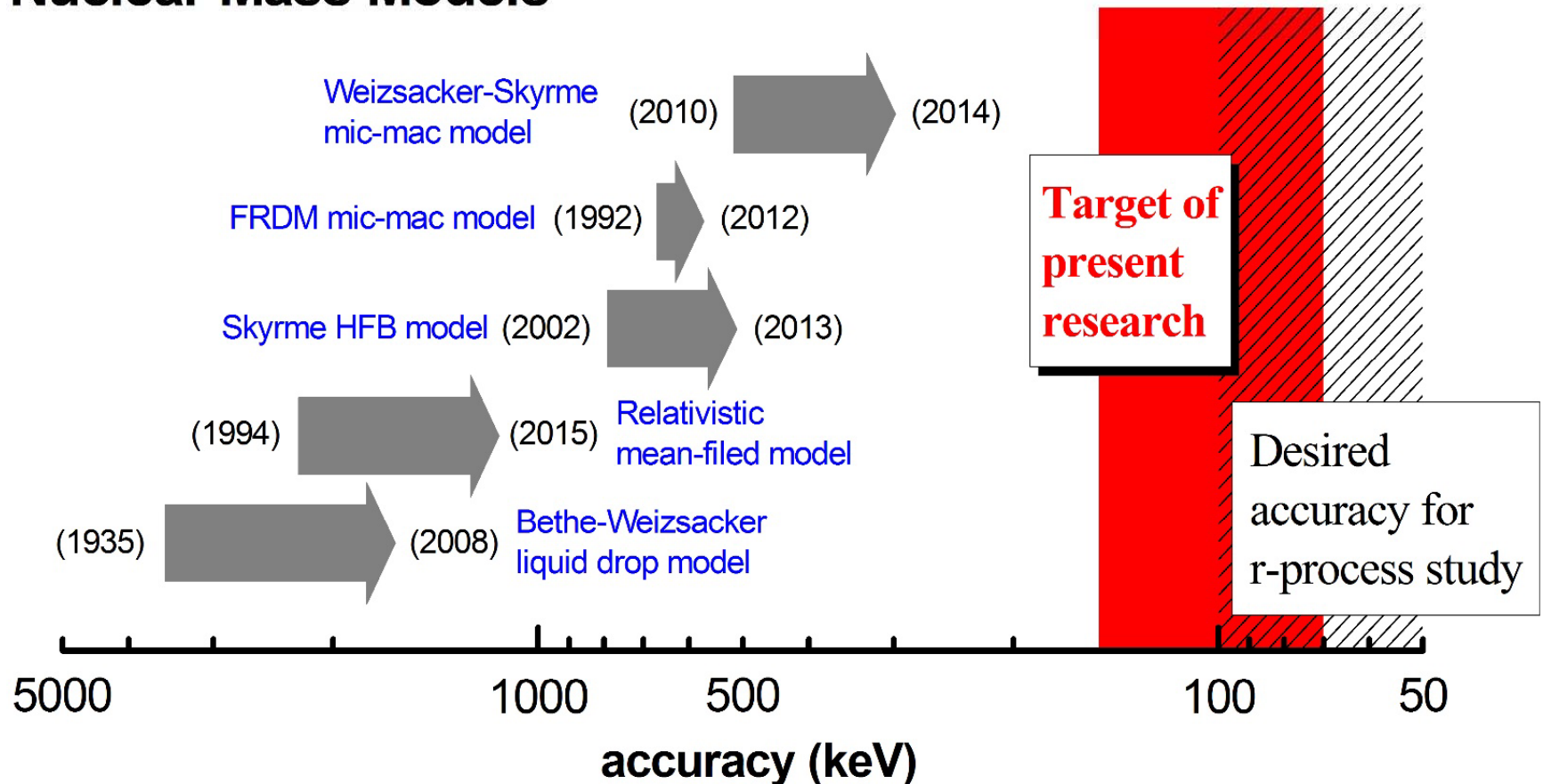
浅色和深色阴影带分别代表 0.5 MeV 和 0.1 MeV 的核质量不确定度对r-过程丰度的影响

精确描述r-过程元素丰度需要核质量预言精度至 0.1 MeV

Nuclear mass models

- ★ Macroscopic mass models: BW [[Weizsäcker1935ZP](#), [Bethe1937RMP](#), [Kirson2008NPA](#)]
- ★ Macro-microscopic mass models: FRDM, WS4 [[Moller2012PRL](#), [Wang2014PLB](#)]
- ★ Microscopic mass models: Skyrme HFB, RMF [[Goriely2016PRC](#), [Geng2005PTP](#)]

Nuclear Mass Models

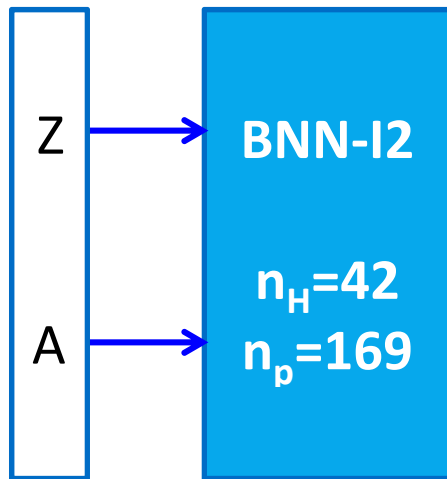


Numerical details

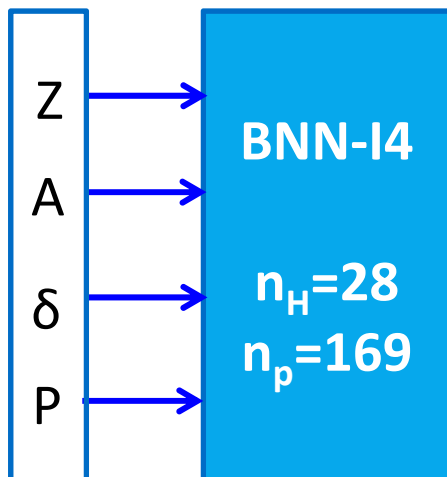
$$y_n(x, \omega) = a + \sum_{j=1}^H b_j \tanh \left(c_j + \sum_{i=1}^I d_{ji} x_{in} \right)$$

Data: 2272 nuclei from AME2016
($Z, N \geq 8$ and $\sigma^{\text{exp}} \leq 100$ keV)

M. Wang et al., CPC 41 030003



$$y_n + M_n^{\text{th}} : t_n = M_n^{\text{exp}} \Leftrightarrow y_n : t_n = M_n^{\text{exp}} - M_n^{\text{th}} \\ \Rightarrow M_n^{\text{th}} = y_n + M_n^{\text{th}}$$

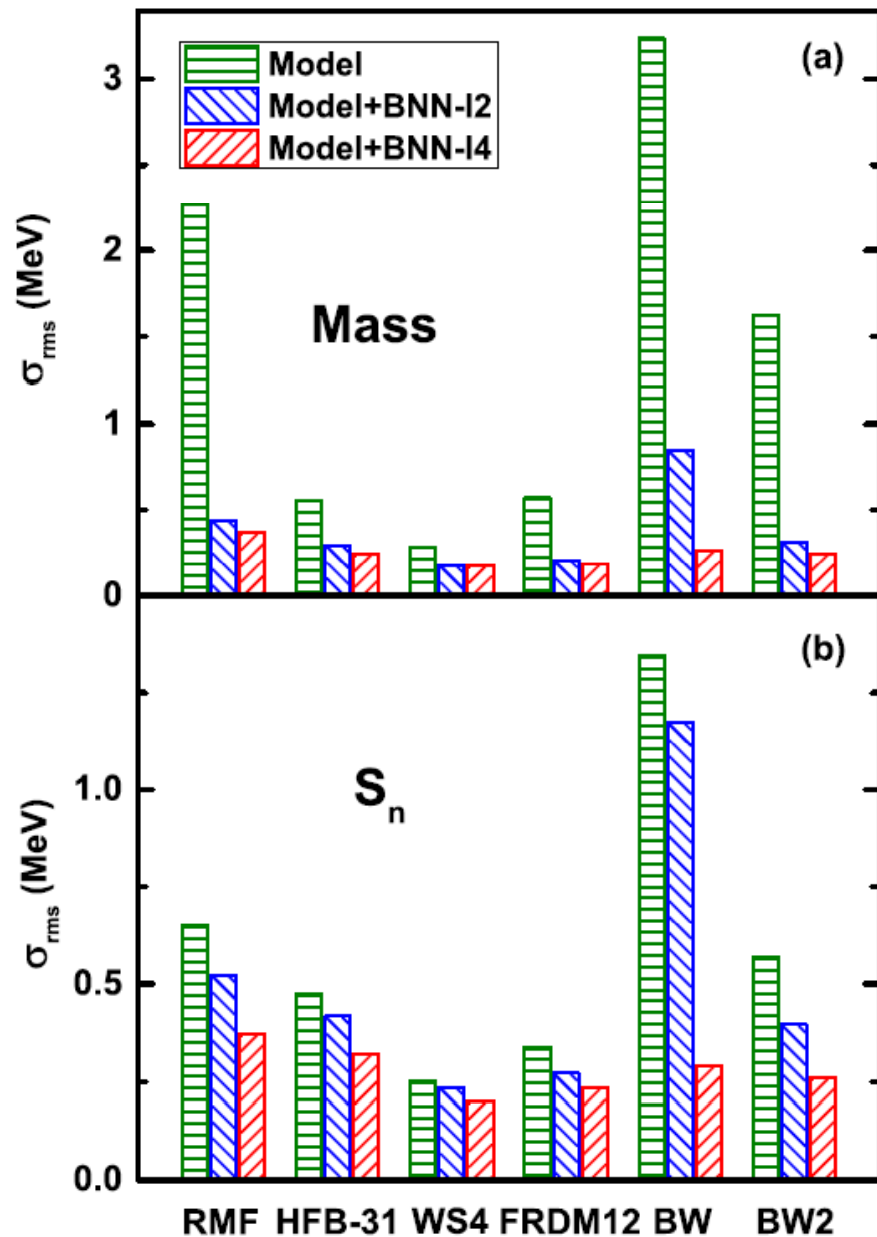


$$y_n + M_n^{\text{th}} : t_n = M_n^{\text{exp}} \Leftrightarrow y_n : t_n = M_n^{\text{exp}} - M_n^{\text{th}} \\ \Rightarrow M_n^{\text{th}} = y_n + M_n^{\text{th}}$$

$$\delta = [(-1)^Z + (-1)^N] / 2, P = v_n v_p / (v_p + v_n)$$

$$v_p = \min(|Z - Z_0|), v_n = \min(|N - N_0|)$$

Rms deviations of mass and S_n

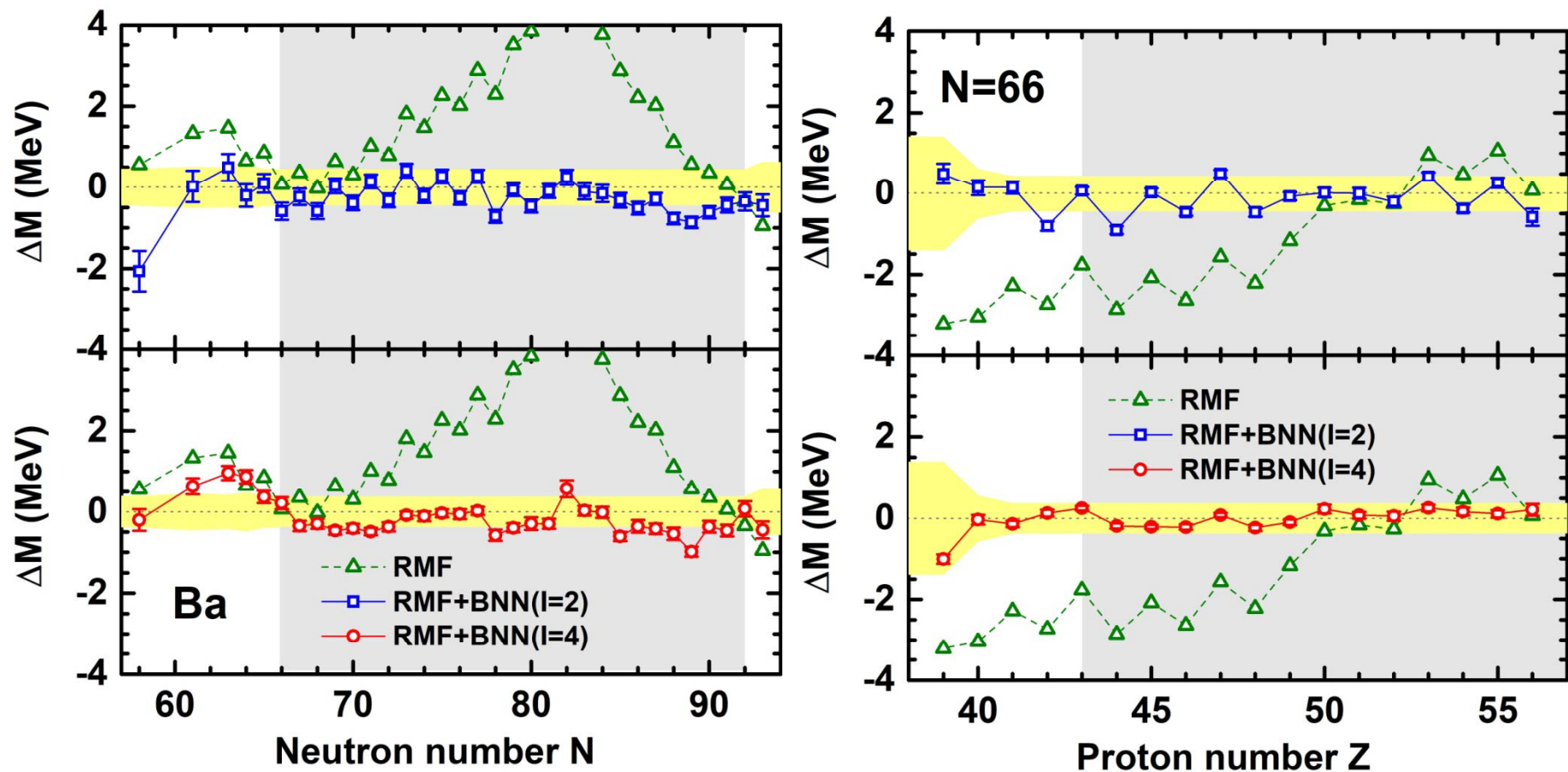


► The predictions of nuclear mass and neutron-separation energy are significantly improved with the BNN approach.

► After the improvement using the BNN approach with four inputs, the rms deviations are generally around 200 keV.

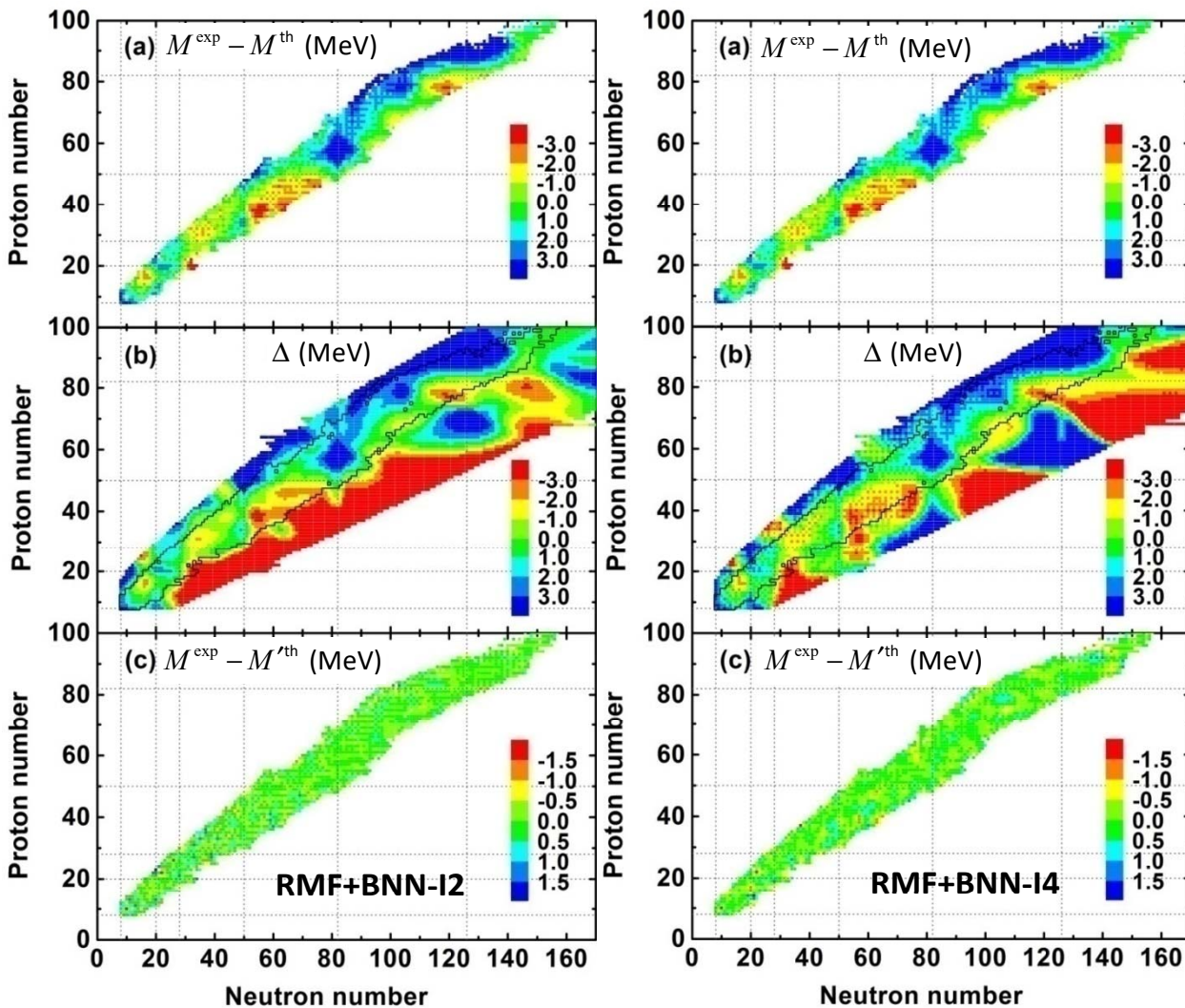
► The BNN with four inputs is more powerful than the BNN with two inputs, especially for the neutron separation energy.

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



► 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 $(50, 126)$.

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 \(2018\)](#)

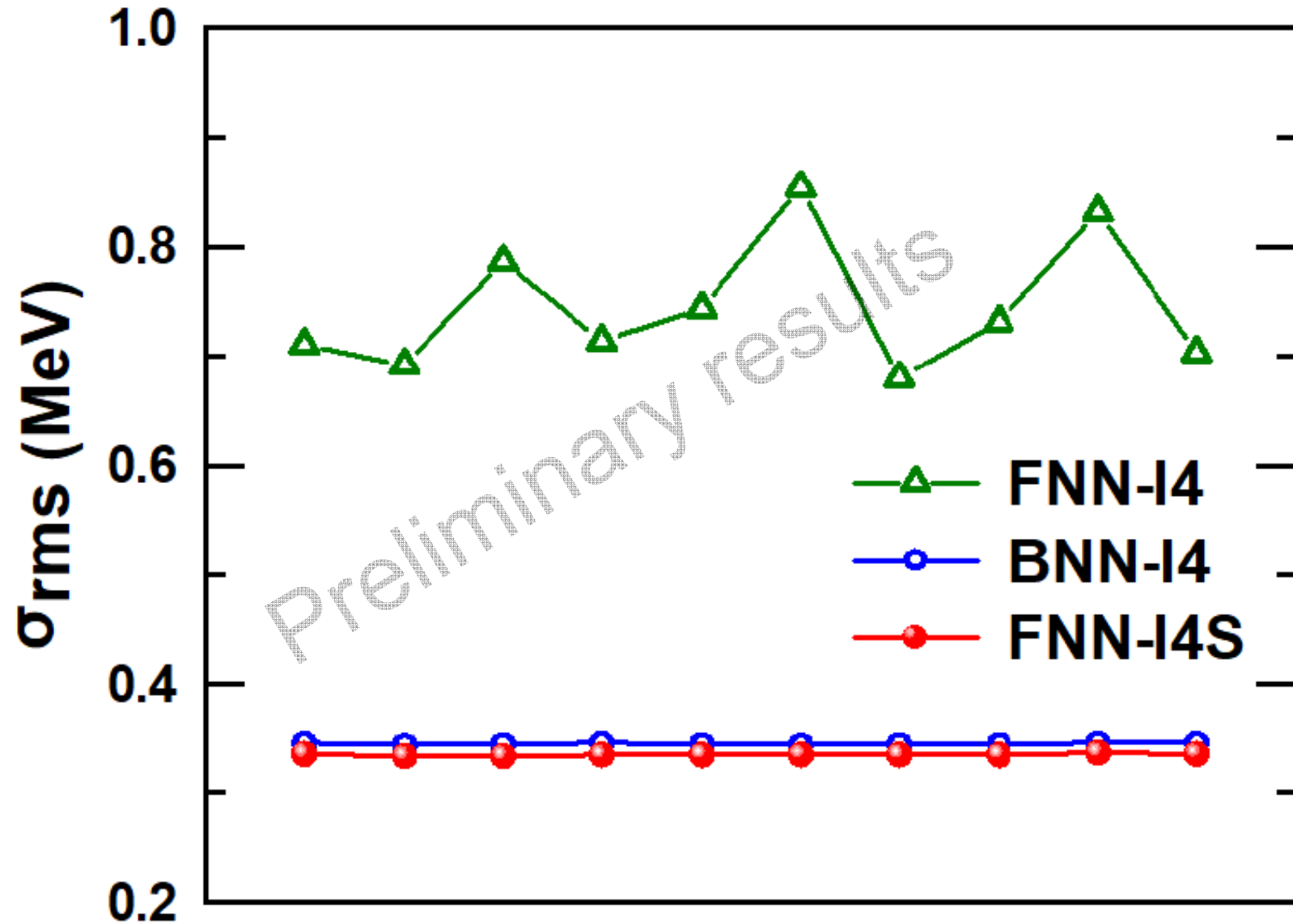
FNN VS BNN

模型	Mass rms (MeV)		
	Adam	RMSprop	SGD
BW	3.068		
FNN-I2a ($Z, N \rightarrow M_{\text{exp}}$)	39.745	33.192	41.226
FNN-I2b ($Z, N \rightarrow M_{\text{exp}} - M_{\text{BW}}$)	1.186	2.247	2.812
FNN-I4 ($Z, N, \delta, P \rightarrow M_{\text{exp}} - M_{\text{BW}}$)	0.685	1.066	2.080
BNN-I4 ($Z, N, \delta, P \rightarrow M_{\text{exp}} - M_{\text{BW}}$)	0.333		

4个输入: $n_H=28$
 2个输入: $n_H=42$
 参数均为169个

通过输入层和输出层的合理设计引入物理可以显著提升神经网络预言能力

FNN VS BNN



Outline

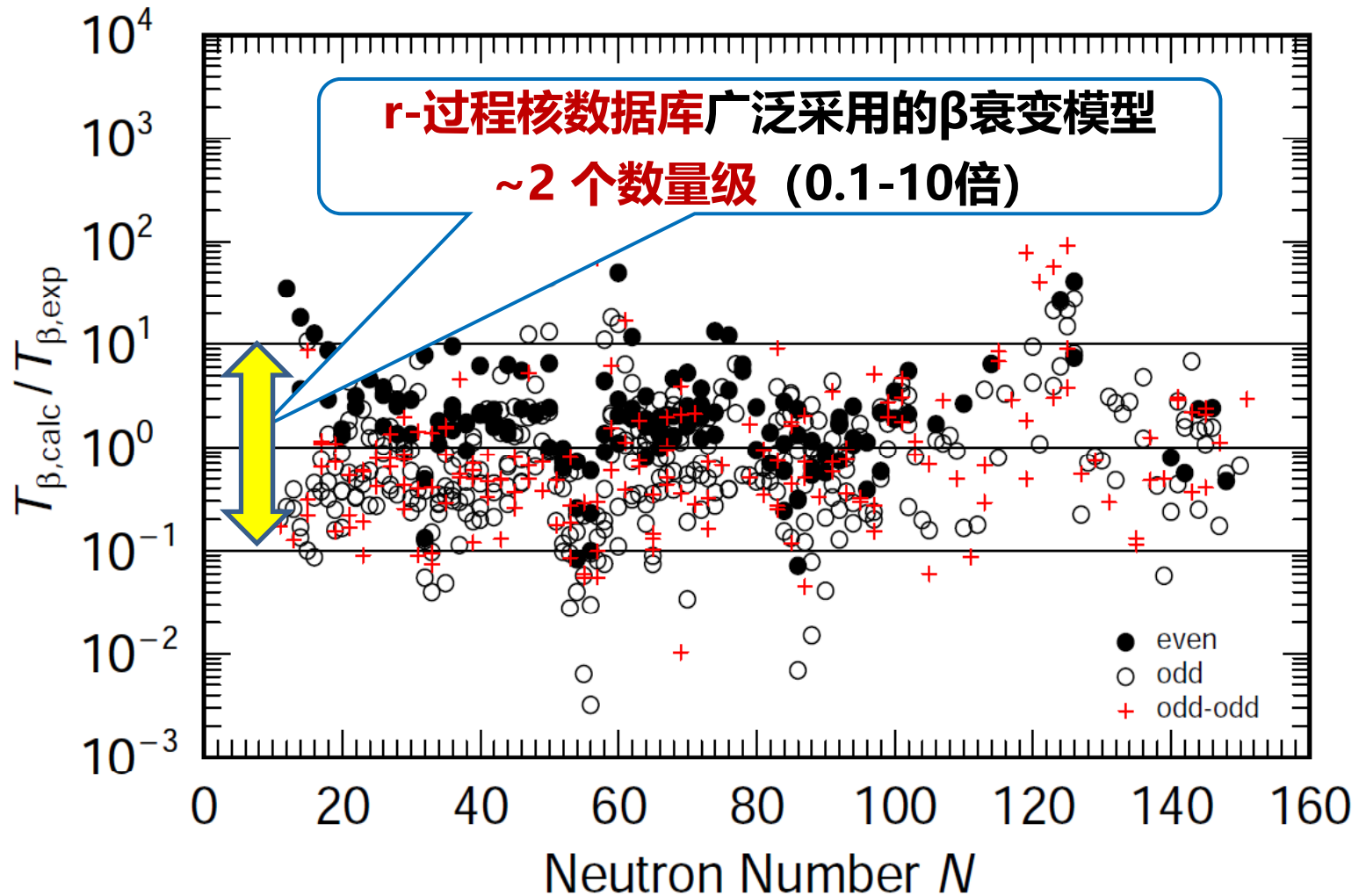
- 1 Introduction
- 2 Bayesian neural network approach
- 3 Results and discussion**
 - ★ Toy model
 - ★ Nuclear masses
 - ★ Nuclear β -decay half-lives
- 4 Summary and perspectives

Nuclear models for β -decay half-lives

Nuclear models for β -decay half-lives

- ★ Phenomenological formula Zhang2006PRC, Zhang2007JPG, Zhou2017SCPMA
- ★ Gross theory Takahashi1973ADNDT,1990PTP, Nakata1997NPA, Koura2017PRC
- ★ Shell model Pinedo1999PRL, Caurier2002PRC, Langanke2003RMP, Zhi2013PRC
- ★ Quasiparticle random phase approximation (QRPA)
 - Nilsson BCS+QRPA: Staudt1990ADNDT, Hirsch1993ADNDT, Nabi1999ADNDT
 - FRDM+QRPA:** Möller1997,2018ADNDT, 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, Niu2018PLB, FAM: Ney2020PRC
 - RHB+QRPA:** Nikšić2005PRC, Marketin2007,2016PRC, Niu2013PRC(R)
 - RHFB+QRPA: Niu2013PLB

Nuclear models for β -decay half-lives



Möller *et al.*, ADNDT 125, 1 (2019)

Influence of half-lives and capture rates

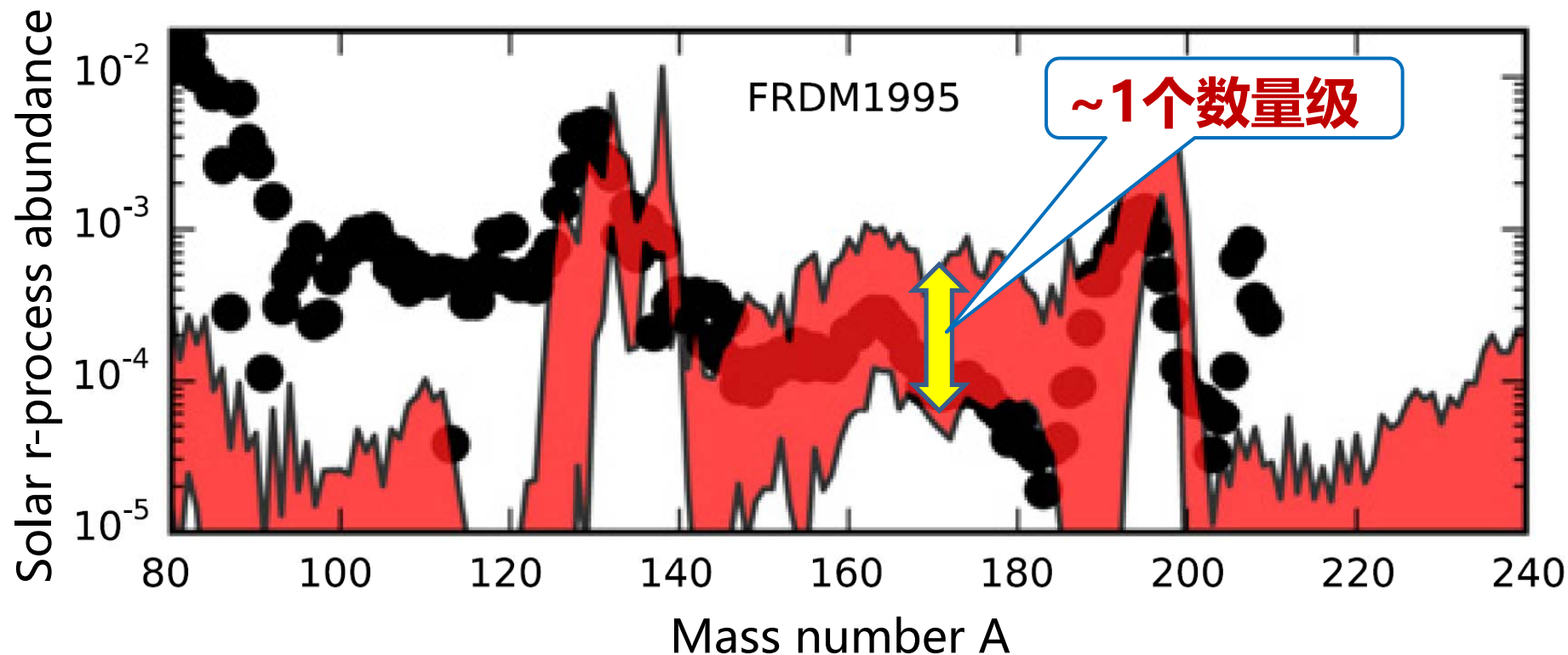


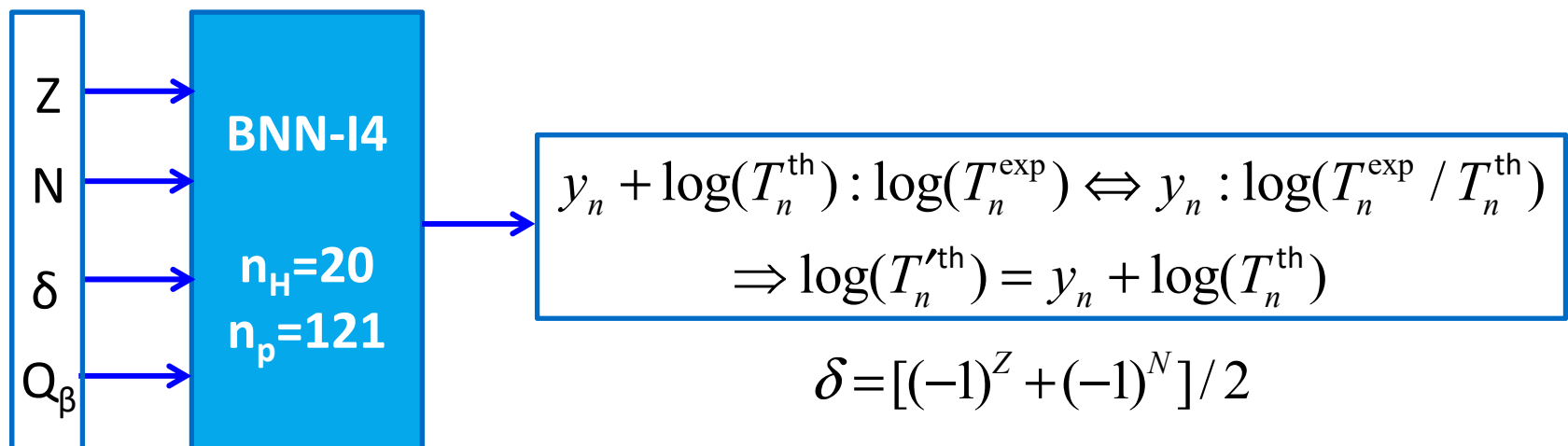
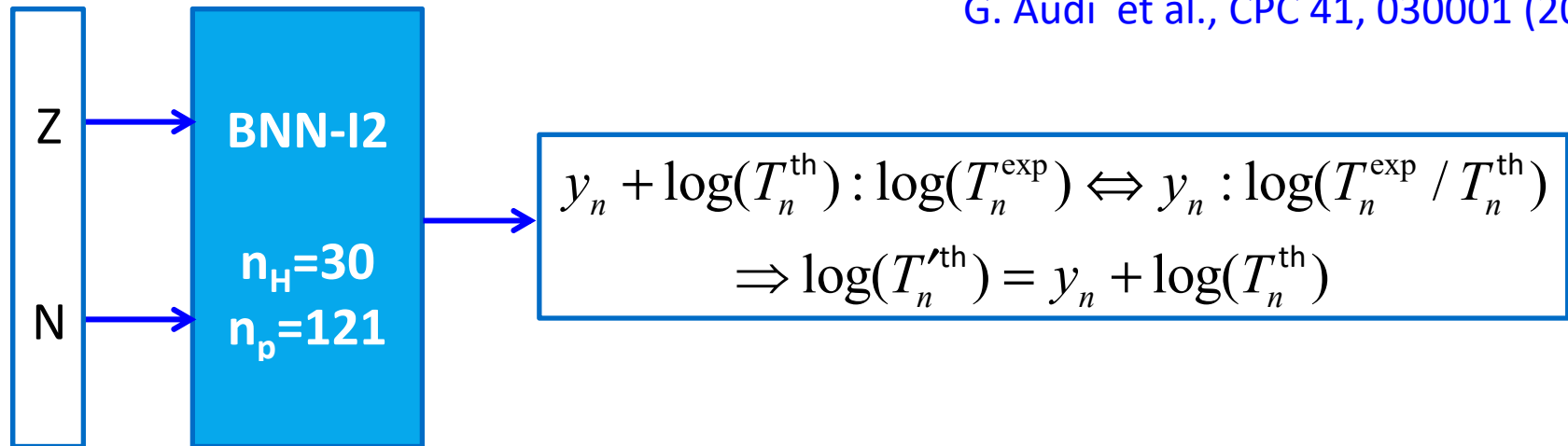
Figure: Variance in isotopic abundance patterns from FRDM1995 mass model predictions compared to the solar data (dots). [M.R. Mumpower *et al.*, PNP 86, 86 \(2016\)](#)

Numerical details

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

Data: 1009 nuclei in NUBASE2016 ($Z, N \geq 8$ and β^- -decay fraction=100%)

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



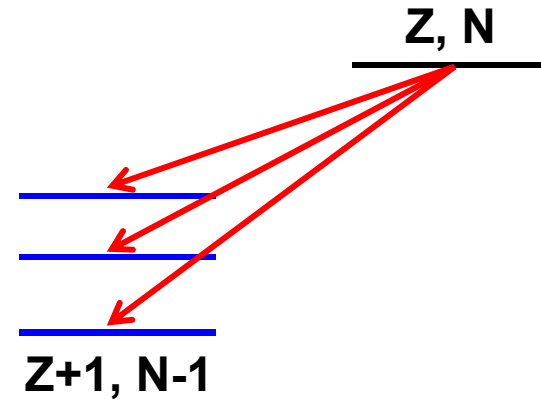
Nuclear β -decay half-lives

- The nuclear β -decay half-life in allowed GT approximation reads as follows:

$$T_{1/2} = \frac{\ln 2}{\lambda_\beta} = \frac{D}{g_A^2 \sum_m B_{GT}(E_m) f(Z, A, E_m)}$$

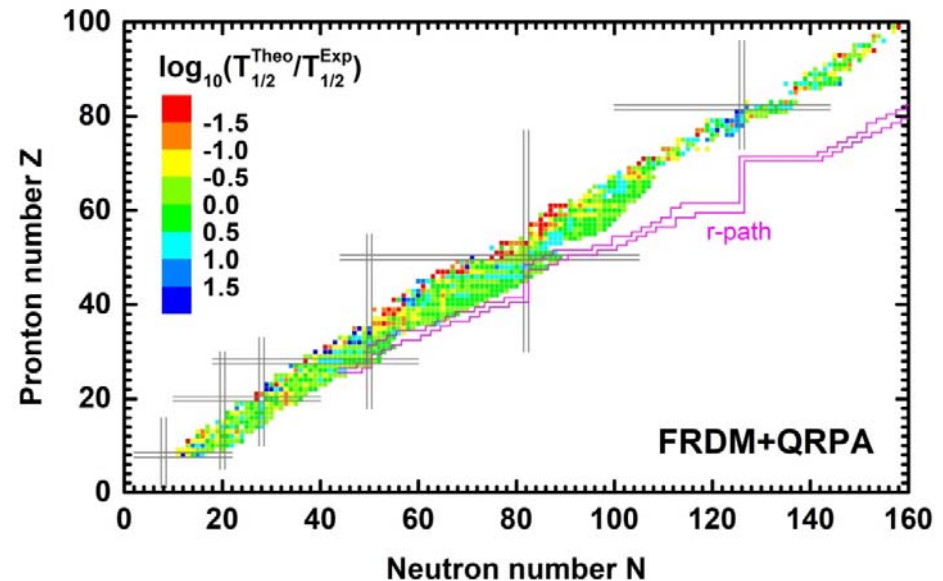
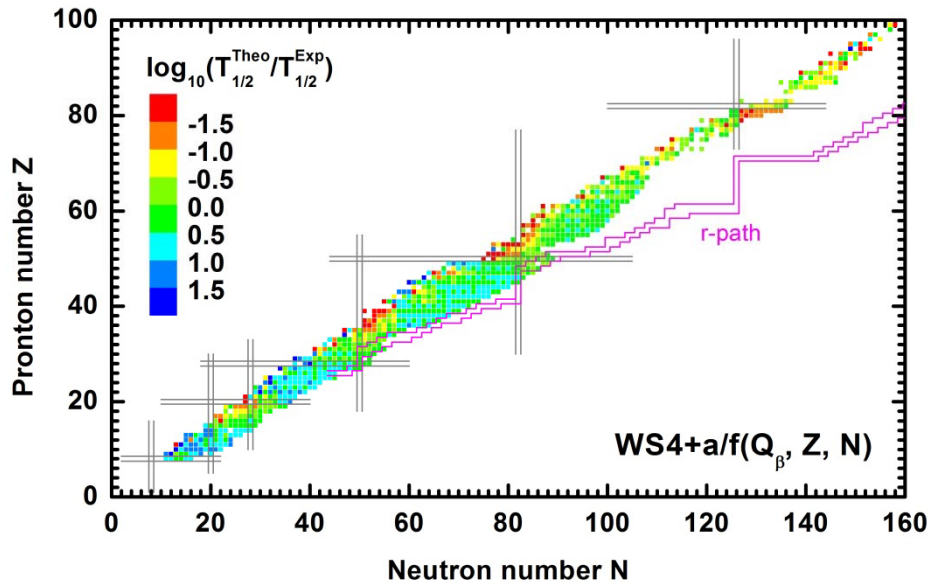
$$\rightarrow T_{1/2} = a / f(Z, A, E_m = Q_\beta - c(\delta - 1)/\sqrt{A})$$

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

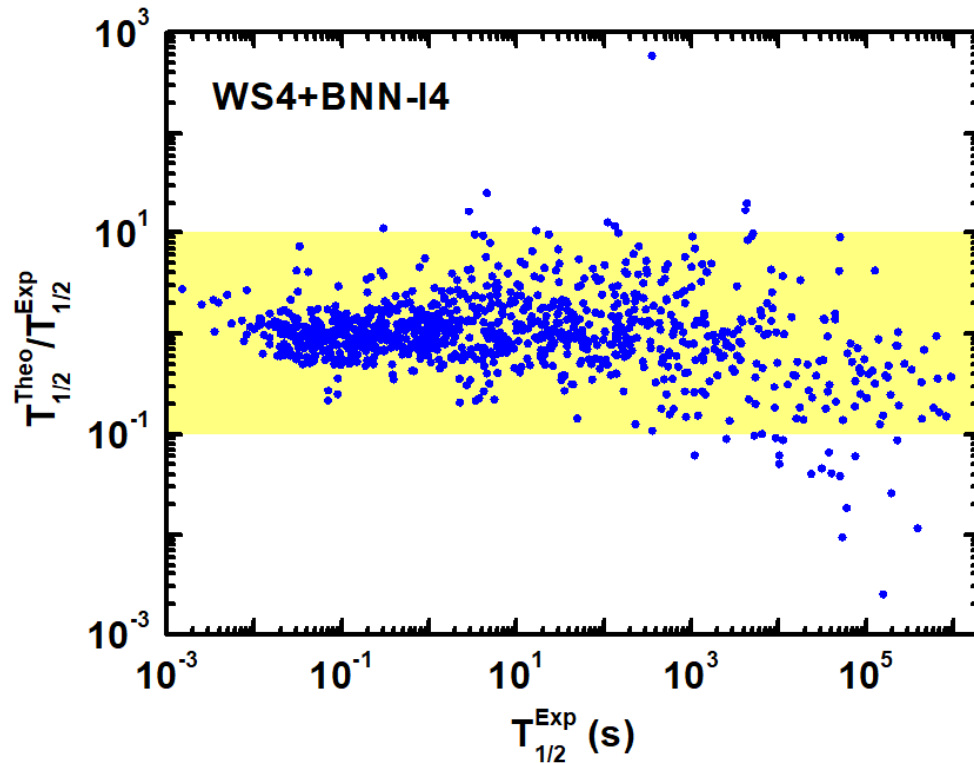


The phase volume is

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



Half-lives of BNN approaches

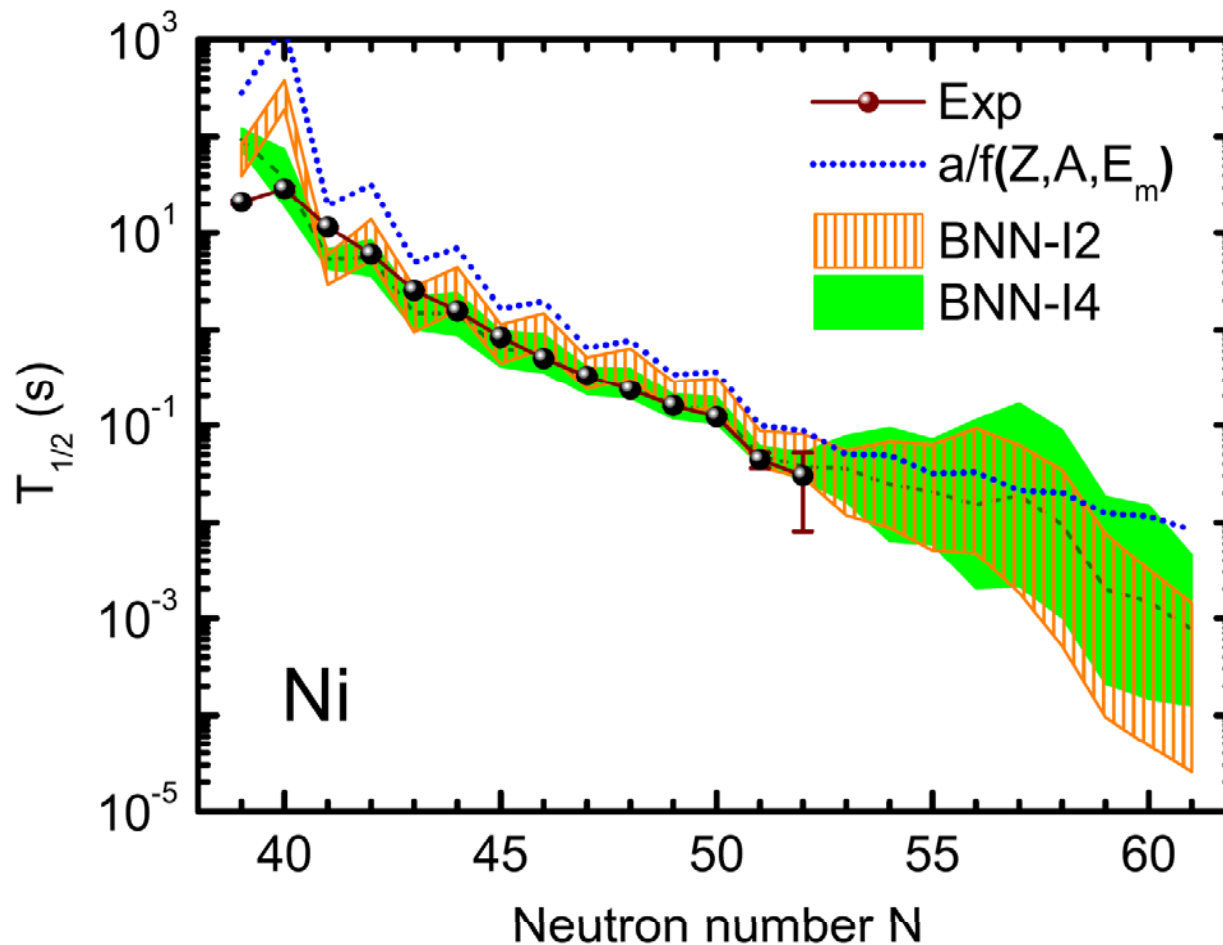


- The WS4+BNN-I4 approach usually better reproduce the half-lives of short-lived nuclei.
- The WS+BNN-I4 approach gives the best results, which can describe nuclear half-lives around $10^{0.2}=1.6$ times of experimental data for nuclei with half-lives shorter than 1 s.

	$T_{1/2} < 10^6$ 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

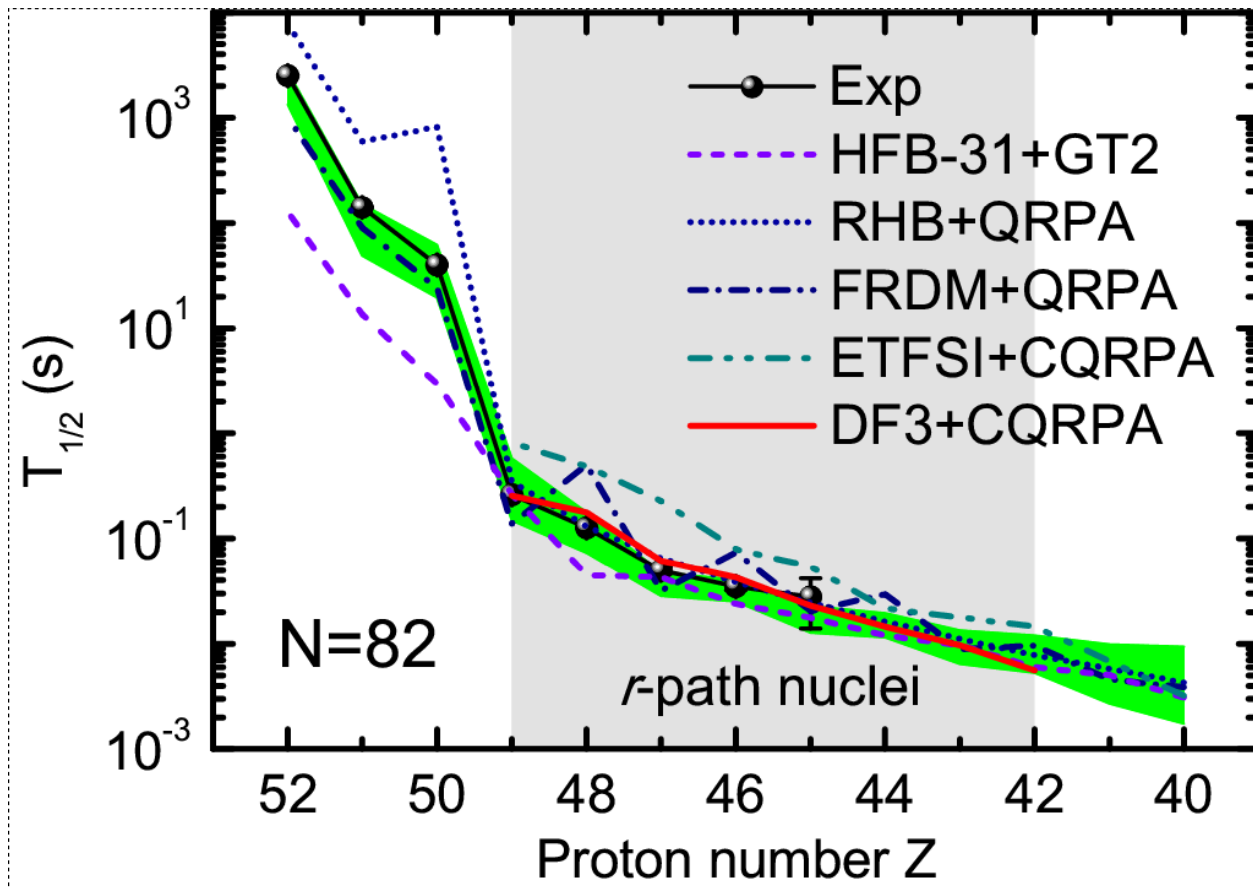
$$\sigma_{\text{rms}} = \sqrt{\frac{\sum_{i=1}^n \left[\log_{10} \left(T_{1/2}^{\text{Exp}} / T_{1/2}^{\text{Theo}} \right) \right]^2}{n}}$$

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



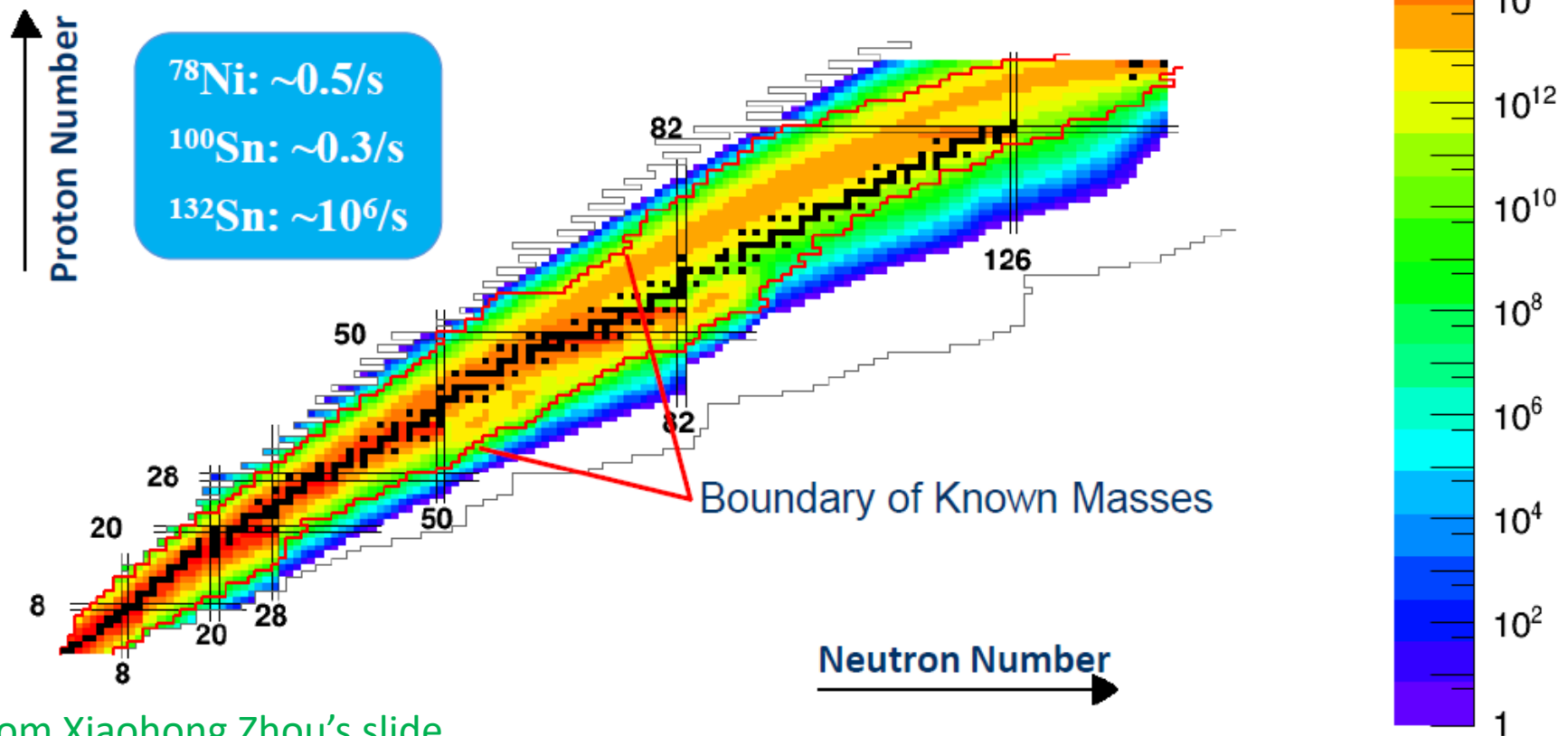
- The results of WS4+BNN-I4 approach are in good agreement with the experimental data, even completely agree with the experimental data within uncertainties for short-lived nuclei.
 - When extrapolate from known region, the results of other models generally agree with WS4+BNN-I4 predictions within uncertainties.
- Z. M. Niu et al., PRC 99, 064307 (2019)



Capability of Producing Nuclides

Nuclides Available (Production Yield) at HIAF

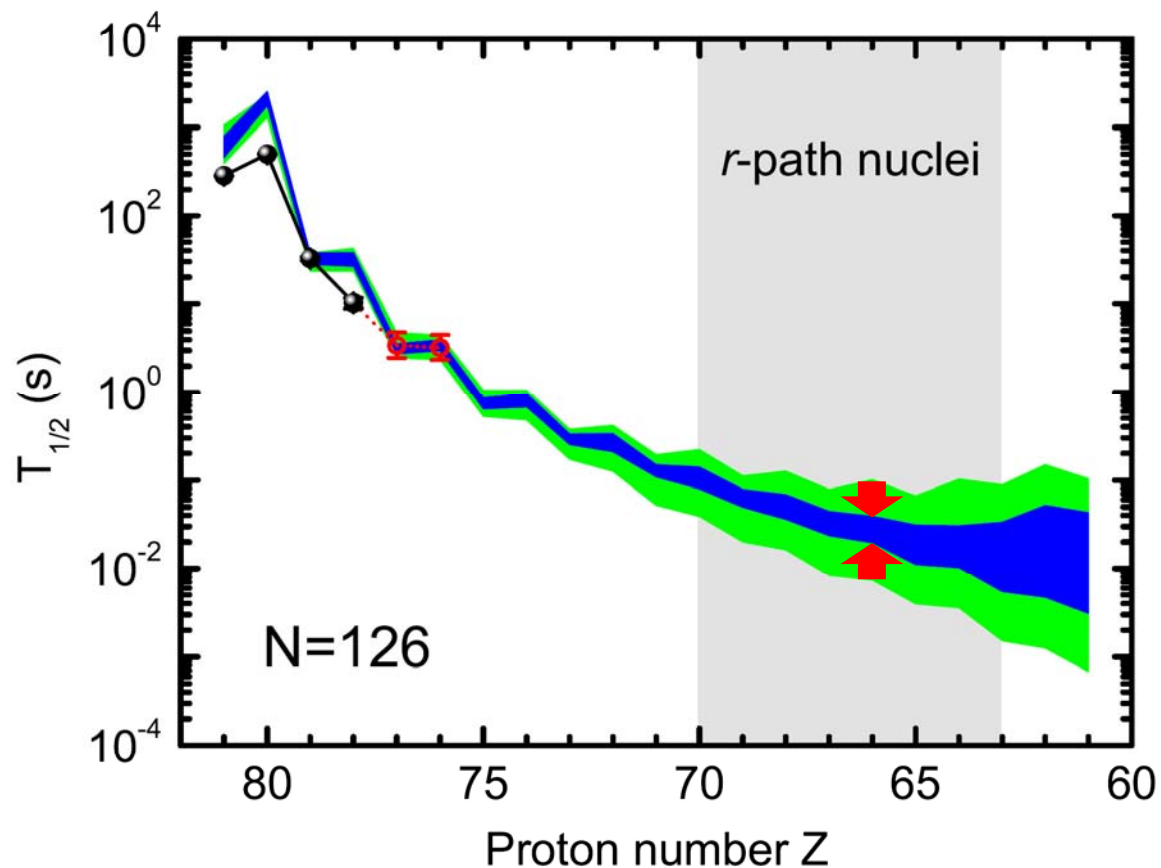
One of the world's most powerful facilities to explore the nuclear chart



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”

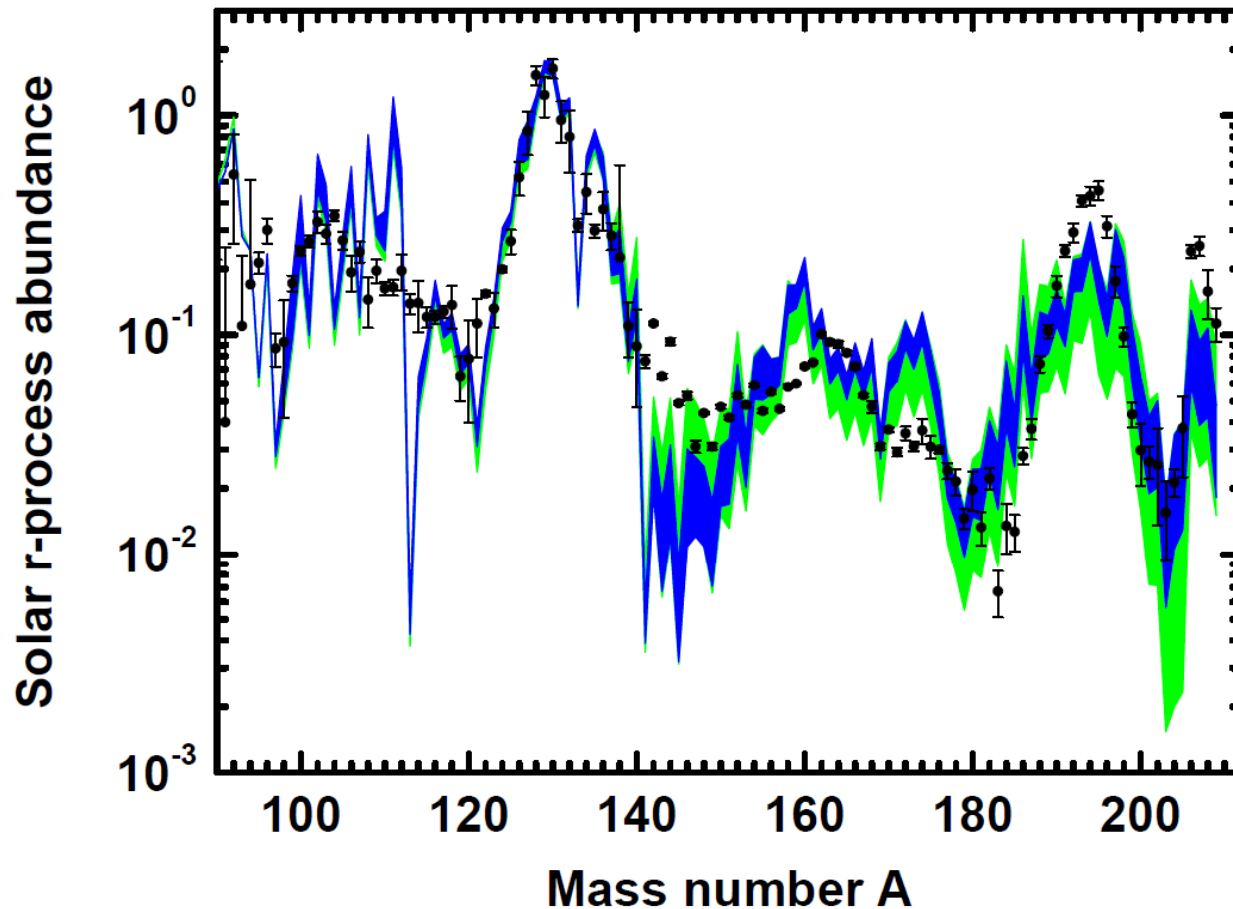
Predictions of nuclear half-lives



- 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 $A > \sim 140$, 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 β -decay half-lives
 - ★ Some preliminary results
- ④ Summary and perspectives

BMM predictions

Model	M	S_n	S_{2n}	S_p	S_{2p}	S_D	Q_β
BMM	0.084	0.078	0.105	0.083	0.111	0.096	0.099
HFB31	0.559	0.451	0.456	0.489	0.496	0.566	0.557
FRDM12	0.576	0.340	0.442	0.341	0.420	0.411	0.450
WS4	0.285	0.254	0.261	0.261	0.300	0.324	0.327

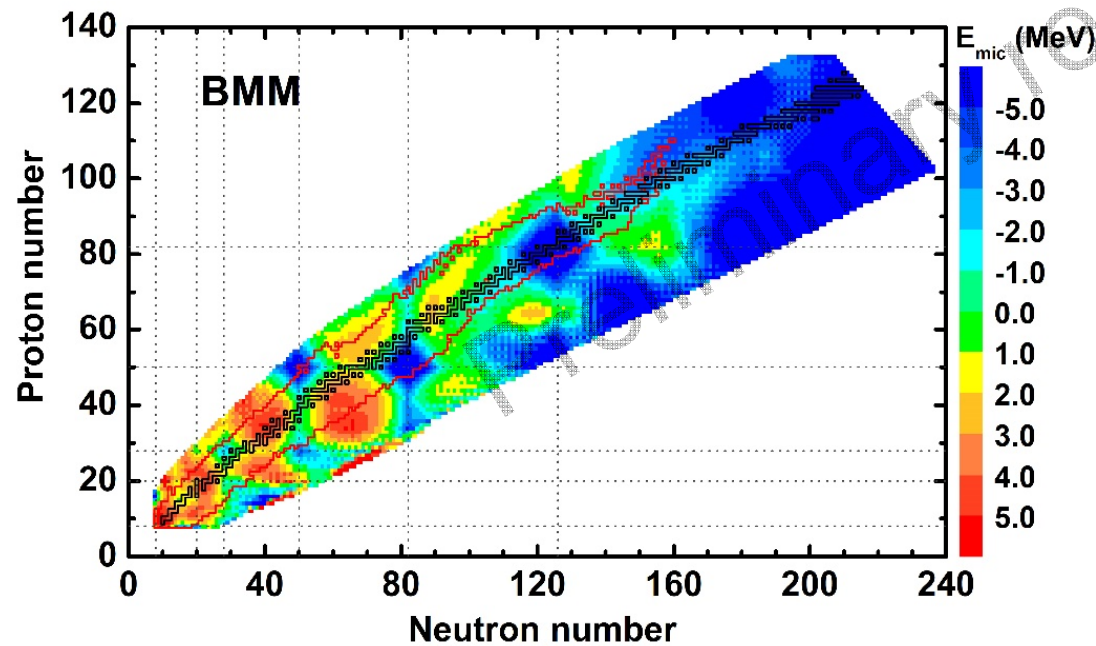


Fig: E_{mic} of BMM with the training data from AME16.

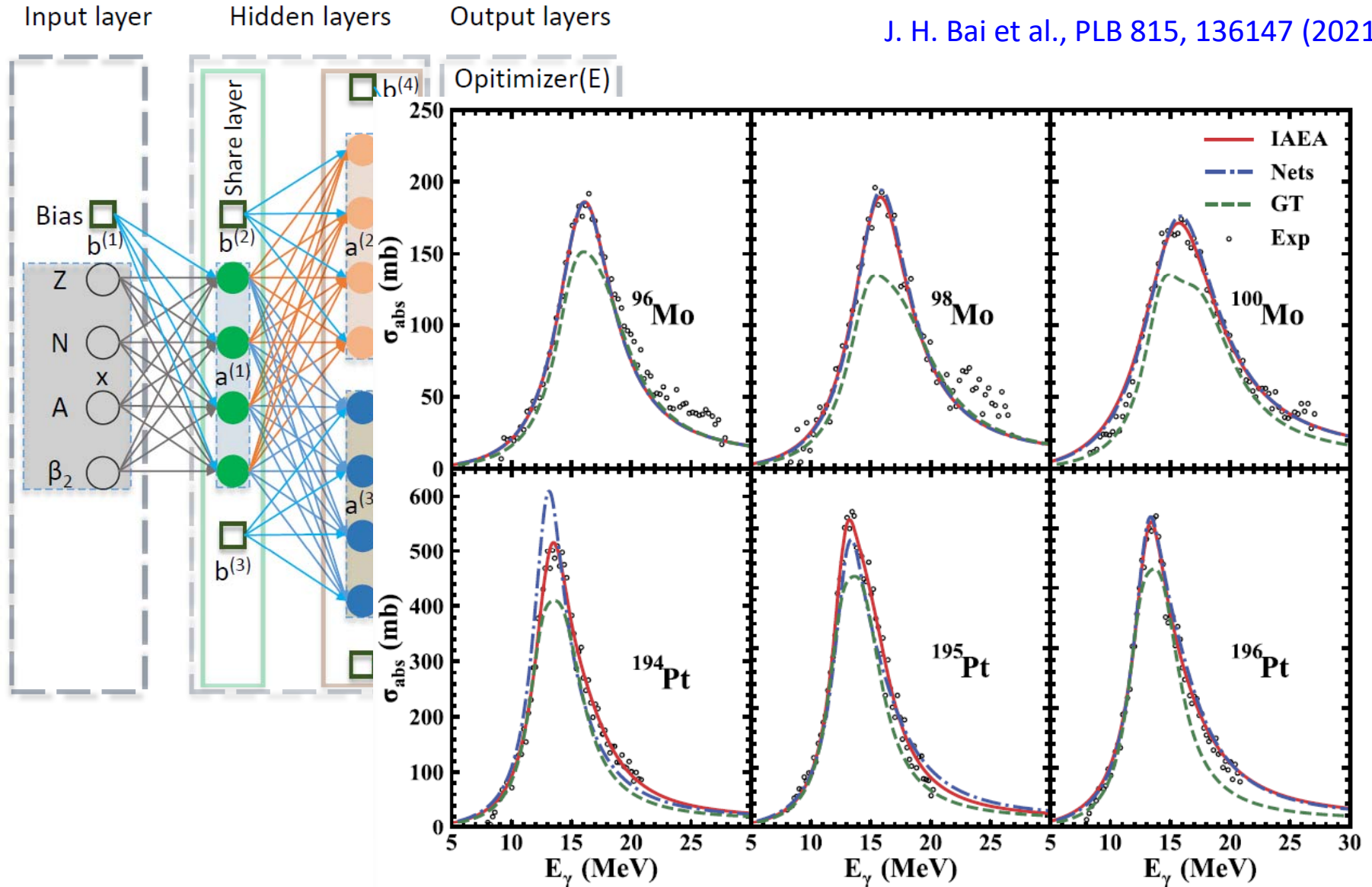
- ★ The first nuclear mass model with accuracy smaller than 100 keV is constructed.

- ★ Its accuracies to S_x and Q_x are at least about 3 times higher than other mass models.

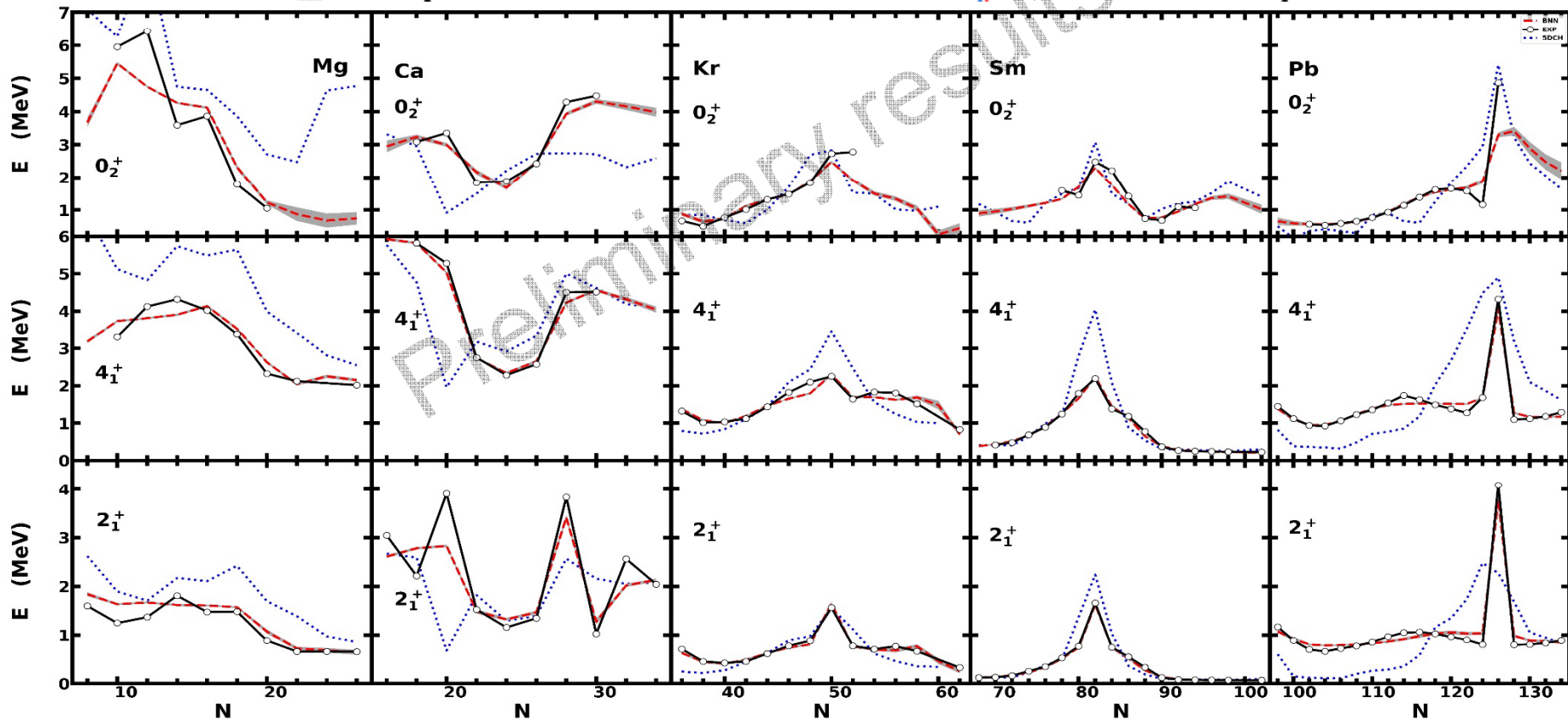
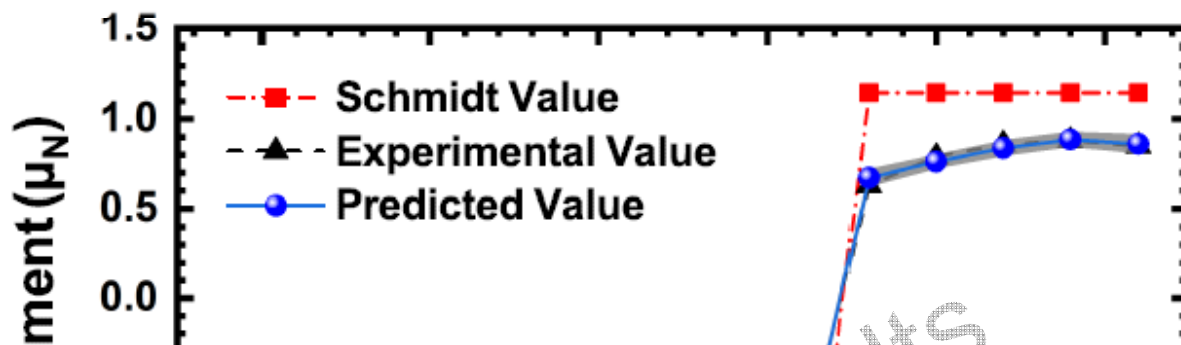
- ★ The shell structure in the known region is well reproduced.

Giant dipole resonance

J. H. Bai et al., PLB 815, 136147 (2021)



Magnetic moments and excitation energies



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 and Perspectives

- 总结：采用BNN方法研究了原子核质量和 β 衰变寿命。发现：
 - ★ BNN方法显著提高了原子核质量和 β 衰变寿命预测的准确性。
 - ★ 包含了更多的物理对于获得更好的预测性能非常重要：
 - 输入层：质量： δ 和 P ；半衰期： δ 和 Q_β
 - 输出层：实验-理论
- 展望：
 - ★ 将机器学习应用于其他核性质
 - ★ 设计适用于核性质的专业神经网络结构
 - ★ 考虑更多的物理提升机器学习性能
 - ★ 机器学习的可解释性与核理论的共同促进

Acknowledgements

合作者：

安徽大学：郭建友、王逸夫

北京航空航天大学：孙保华

吉林大学：李剑、袁子龙、田大川、尚天帅

兰州大学：龙文辉、牛一斐、孙保元、白景虎

西南大学：李志攀

东京大学：梁豪兆

Thank you!