贝叶斯推断在核物理中的应用

张振
中山大学，中法核工程与技术学院

第九届华大QCD讲习班，武汉，2021.10.13

一校、三城、五校园

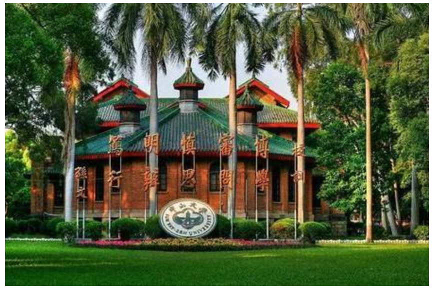
中山大学（珠海校区）
中法核工程与技术学院

中山大学

一校、三城（广州、珠海、深圳）、五校园

珠海校园

广州南校园



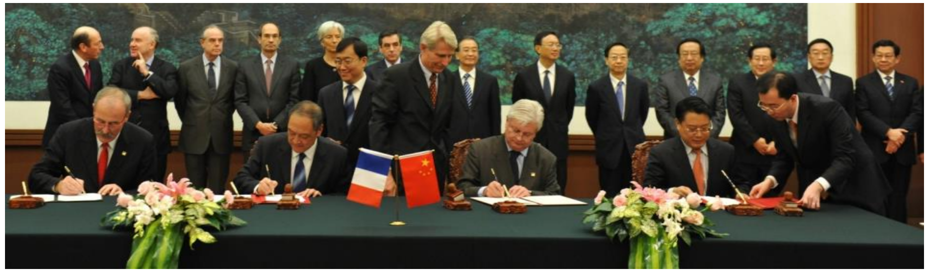
广州北校园

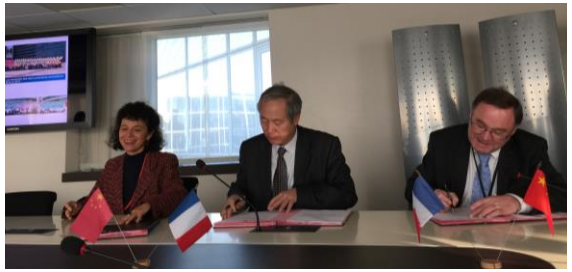
广州东校园

深圳校园

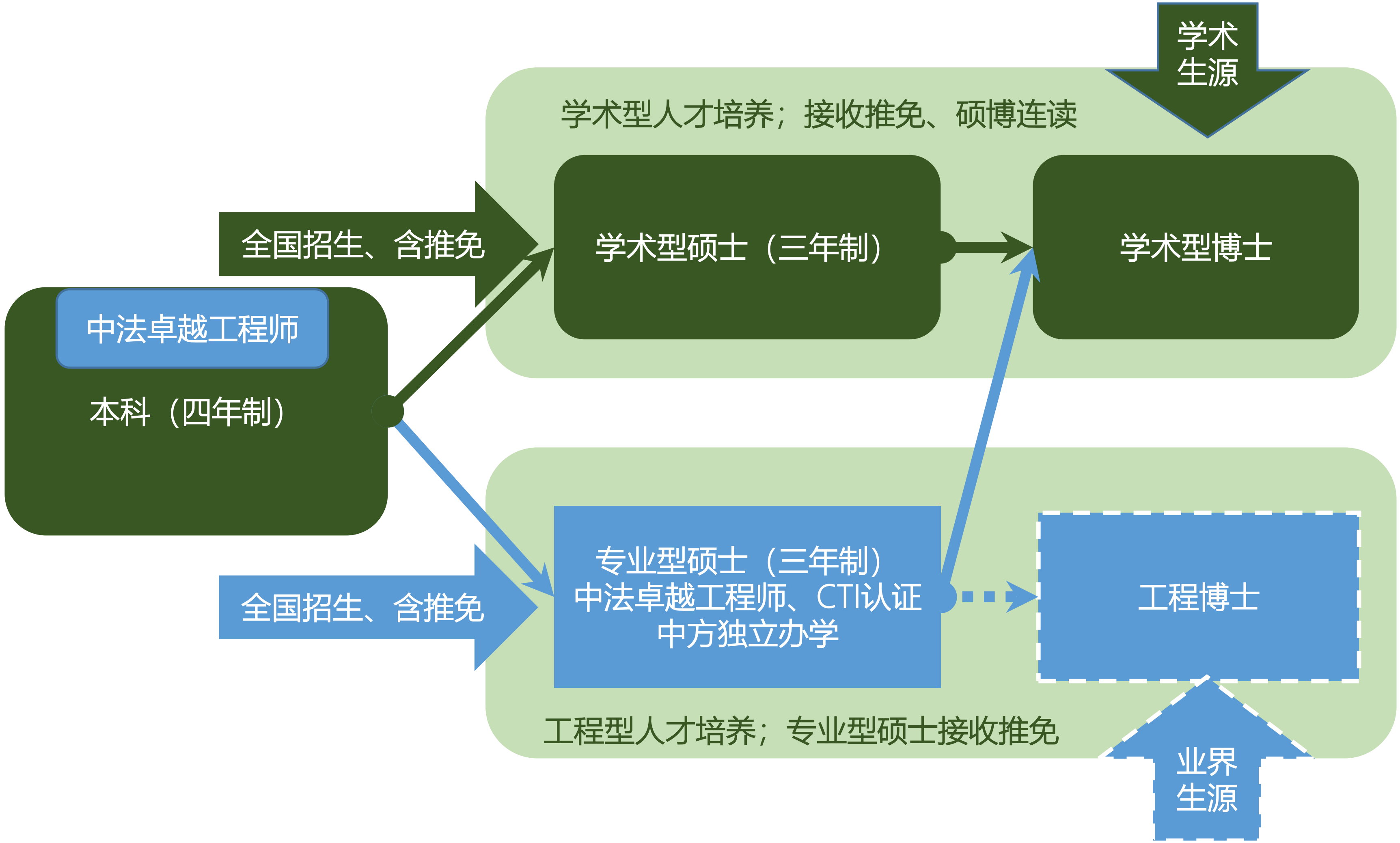
学院简介：学院简史

- 2009年：中山大学与法国格勒诺布尔理工大学为首的民用核能工程师教学联盟（FINUCI）合作成立中山大学中法核工程与技术学院
- 2010年：第一届学生入学
- 2011年：核科学与技术一级学科硕士点
- 2012年：核能与核技术工程二级学科专业硕士点
- 2016年：核工程力学二级学科博士点；第一届CTI认证硕士生毕业
- 2018年：第八届广东省教育教学成果奖一等奖
- 2019年：入选首批国家级一流本科专业建设点（双万计划）
- 2020年：核科学与技术一级学科博士点通过学校自主审议，报送教育部





学院简介：办学模式



学术生源
学术型人才培养；接收推免、硕博连读
全国招生、含推免
学术型硕士（三年制）
学术型博士
中法卓越工程师
本科（四年制）
全国招生、含推免
专业型硕士（三年制）
中法卓越工程师、CTI认证
中方独立办学
工程博士
工程型人才培养；专业型硕士接收推免
业界生源

科研方向

核燃料循环与材料

- 核工程材料与力学
- 核化学与放射化学

核能科学与工程

- 核仿真与安全
- 反应堆热工水力
- 先进核能系统

核物理

中低能核物理、天体物理、核数据

核科学与技术学科

粒子物理核物理交叉

- 核环境辐射监测与应急
- 辐射防护与安全

辐射防护与环境保护

- 核探测与探测器
- 核技术及应用

核技术及应用

粒子物理、新型探测器、中微子应用

粒子物理

基础研究：粒子物理与原子核物理

中微子物理

- 中微子质量
- 反应堆中微子能谱

（王为、魏月环、张玉美）

核结构与反应

- 核结构理论和实验
- 低能核反应理论
- 中低能重离子核反应理论

（袁岑溪、 苏军、祝龙、梅波、张振、滑伟、郭琛琛）

核数据

- 裂变截面和产额
- 重要材料的中子辐照损伤截面

（陈胜利、梅波、袁岑溪、苏军）

相关交叉研究

- 粒子（核）天体物理
- 粒子探测技术
- 核电子学

（王为、方晓、张振、魏月环、梅波、黄土琛）

Outline

- Motivation
- Bayesian inference approach
  - Basic concept
  - Markov Chain Monte Carlo (MCMC)
  - Gaussian process (GP)
  - Principal component analysis (PCA)
- Applications in nuclear physics
- Summary

Inverse problems in nuclear physics

Lattice QCD

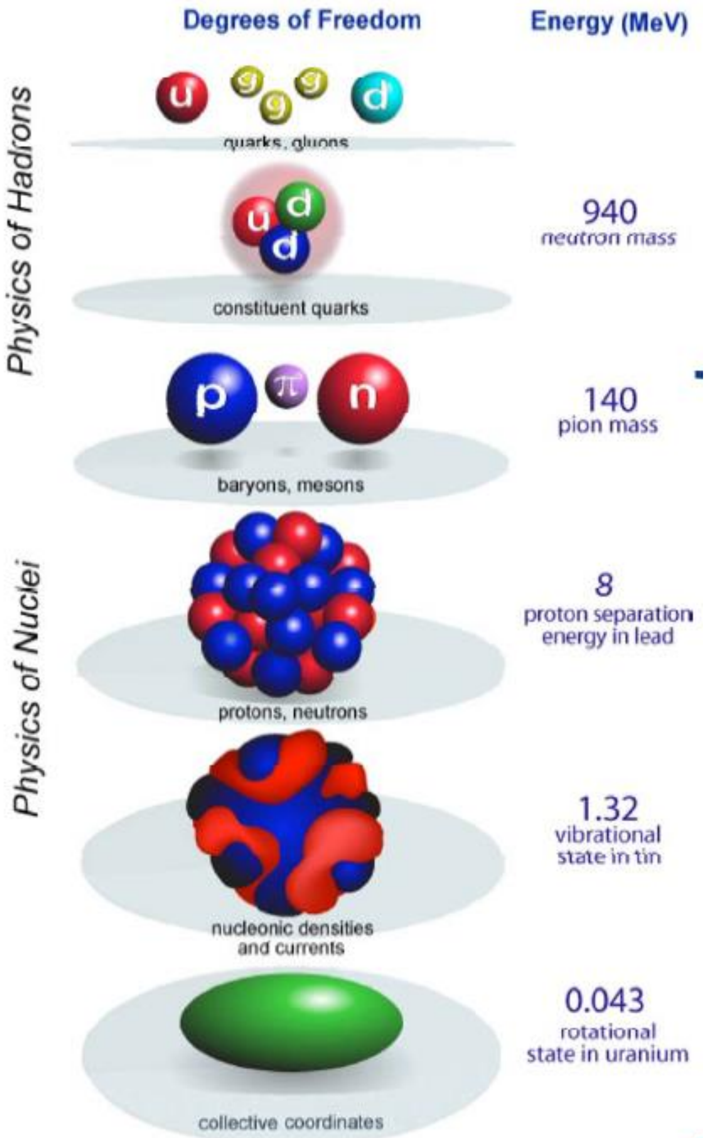
Quark model

Ab initio

Configuration interaction

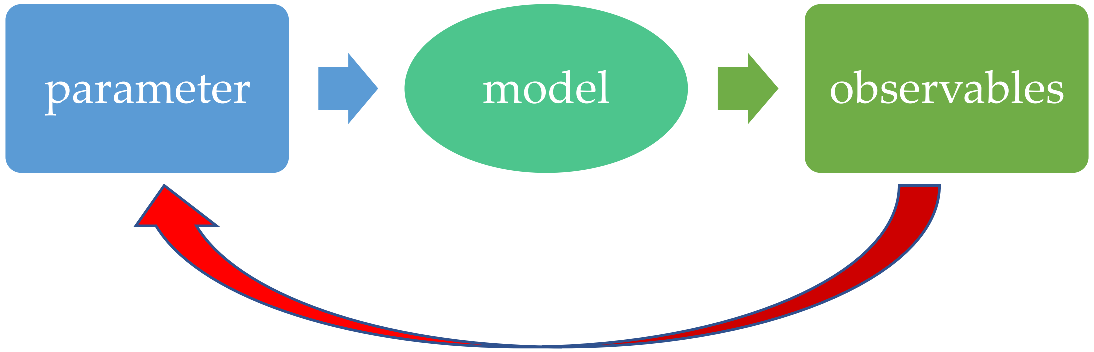
Energy density functional theory

collective models



J. Phys. G 43, 044002 (2016)

- Model everywhere



Inverse problem

- Error propagation?

Uncertainty and correlation

The Editors, Phys. Rev. A 83, 040001 (2011)

> *It is all too often the case that the numerical results are presented without uncertainty estimates. Authors sometimes say that it is difficult to arrive at error estimates. Should this be considered an adequate reason for omitting them? [...] There is a broad class of papers where estimates of theoretical uncertainties can and should be made. Papers presenting the results of theoretical calculations are expected to include uncertainty estimates for the calculations whenever practicable, and especially under the following circumstances: (1) if the authors claim high accuracy, or improvements on the accuracy of previous work; (2) if the primary motivation for the paper is to make comparisons with present or future high precision experimental measurements; (3) if the primary motivation is to provide interpolations or extrapolations of known experimental measurements.*

Referee's comment on *Zhang et al. PLB 777,73 (2018)*

> Finally, given that the authors have calibrated a new set of extended Skyrme interactions, I puzzle why in this day and age their predictions are not accompanied by theoretical uncertainties and correlation plots.

Information content of new measurements

PHYSICAL REVIEW C **81**, 051303(R) (2010)

Information content of a new observable: The case of the nuclear neutron skin

P.-G. Reinhard[1] and W. Nazarewicz[2,3,4,5]

1. Considering the current theoretical knowledge, what novel information does new measurement bring in?
2. How can new data reduce uncertainties of current theoretical models?

Why Bayesian analysis?

Bayesian analysis for

- ✓ Parameter calibration
- ✓ Uncertainty quantification for parameters and predictions
- ✓ Correlation analysis
- ✓ Information content of new measurements

……

Frequentist approach

Optimize parameters by minimizing the object function

$$\chi^2(\mathbf{p}) = \sum_{\mathcal{O}} \left(\frac{\mathcal{O}^{(\mathrm{th})}(\mathbf{p}) - \mathcal{O}^{(\mathrm{exp})}}{\Delta\mathcal{O}} \right)^2$$

Around the optimal parameter set p_0:

$$\chi^2(\boldsymbol{p}) \approx \chi^2(\boldsymbol{p}_0) + \sum_{i,j}^{N_p} (p_i - p_{0i}) \mathcal{M}_{ij} (p_j - p_{0j}), \text{ with } \mathcal{M}_{ij} = \frac{1}{2} \frac{\partial^2 \chi^2}{\partial p_i \partial p_j} \bigg|_{\boldsymbol{p}_0}.$$

Multivariate Gaussian distribution:

$$\mathcal{N} \exp\left[-\frac{1}{2} (\boldsymbol{p} - \boldsymbol{p}_0)^T \mathcal{C}^{-1} (\boldsymbol{p} - \boldsymbol{p}_0) \right]$$

Covariance matrix

$$\mathcal{C} = \frac{\chi^2(\boldsymbol{p}_0)}{N_d - N_p} \mathcal{M}^{-1}$$

Variance of A

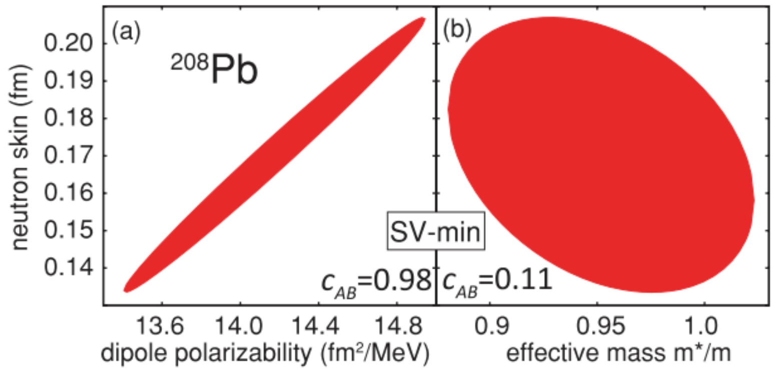
$$\sigma_A^2 = \sum_{ij} \left(\frac{\partial A}{\partial p_i} \right)_0 \mathcal{C}_{ij} \left(\frac{\partial A}{\partial p_j} \right)_0$$

Covariance of A and B

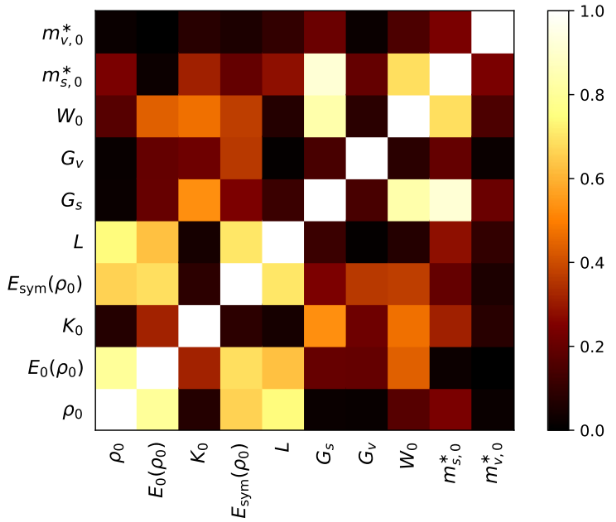
$$Cov(A, B) = \sum_{ij} \left(\frac{\partial A}{\partial p_i} \right)_0 \mathcal{C}_{ij} \left(\frac{\partial B}{\partial p_j} \right)_0$$

Correlation

$$C_{AB} = \frac{Cov(A,B)}{\sigma_A \sigma_B}$$



Reinhard and Nazarewicz, PRC81, 051303(R) (2010)
Dobaczewski et al, JPG41, 074001 (2014)



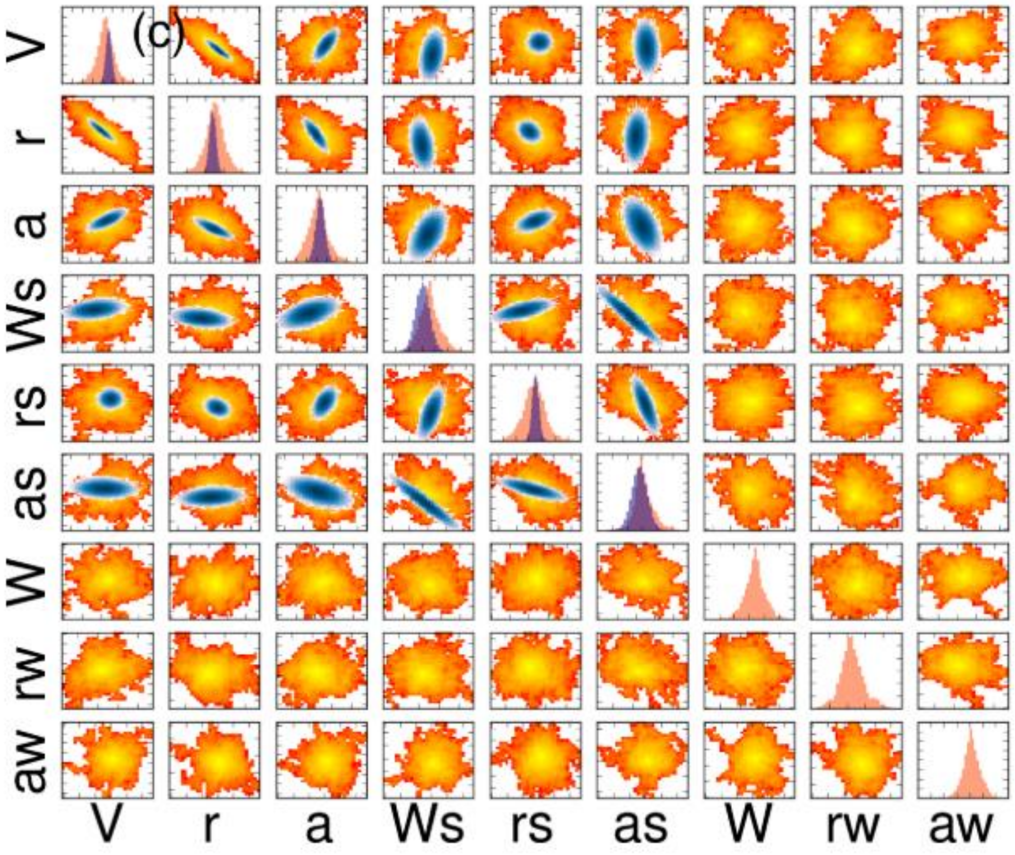
Zhang et al. PLB 777,73 (2018)

Frequentist vs Bayesian

King et al., PRL 122, 232502 (2019).

Frequentist
- Local minimum or global minimum?
- Errors and correlations are determined by $\partial^2\chi^2/\partial p_i \partial p_j$ at minimum point. (reliable for a well-constrained model).

Bayesian
- More flexible in exploring parameter space.
- More reliable uncertainties and correlations.
- Prior knowledge can be easily taken into account.



Blue: Frequentist approach
Orange: Bayesian based on MCMC

Bayesian Inference approach

- Take probabilistic view of parameters.
- Update knowledge on parameters upon observed data.
- Tools:

  MADAI http://madai.phy.duke.edu (PCA+GP+MCMC)

  Pymc3 https://docs.pymc.io/

  …

Bayes' theorem

- Conditional probability

$$P(A \mid B) : \text{probability of event A occurring, given event B}$$

- Product rule:

$$P(AB) = P(A \mid B)P(B) = P(B \mid A)P(A)$$

- Bayes' theorem

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Bayesian inference

- **Likelihood**: probability of observed data given parameter $p(\mathcal{D}\,|\,\boldsymbol{\theta})$
- **Prior**: knowledge about the parameters before seeing data $p(\boldsymbol{\theta})$

  Usually flat
- **Posterior**: probability of parameters given observed data $p(\boldsymbol{\theta}\,|\,\mathcal{D})$

  e.g., $\exp\left[-\sum_i \frac{(o_i - o_i^{\mathrm{exp}})^2}{2\sigma_i^2}\right], \quad \sigma_i = \sqrt{\sigma_{i,\mathrm{th}}^2 + \sigma_{i,\mathrm{exp}}^2}$
- Apply Bayes' rule:

$$p(\boldsymbol{\theta}\,|\,\mathcal{D}) = \frac{p(\mathcal{D}\,|\,\boldsymbol{\theta})\,p(\boldsymbol{\theta})}{p(\mathcal{D})} \propto p(\mathcal{D}\,|\,\boldsymbol{\theta})\,p(\boldsymbol{\theta}),$$

  with $p(D)$ being evidence or marginal distribution of the data.
- Make predictions:

$$p(O\,|\,\mathcal{D}) = \int p(O\,|\,\boldsymbol{\theta})\,p(\boldsymbol{\theta}\,|\,\mathcal{D})\,d\boldsymbol{x}$$

Bayesian inference-uncertainty quantification

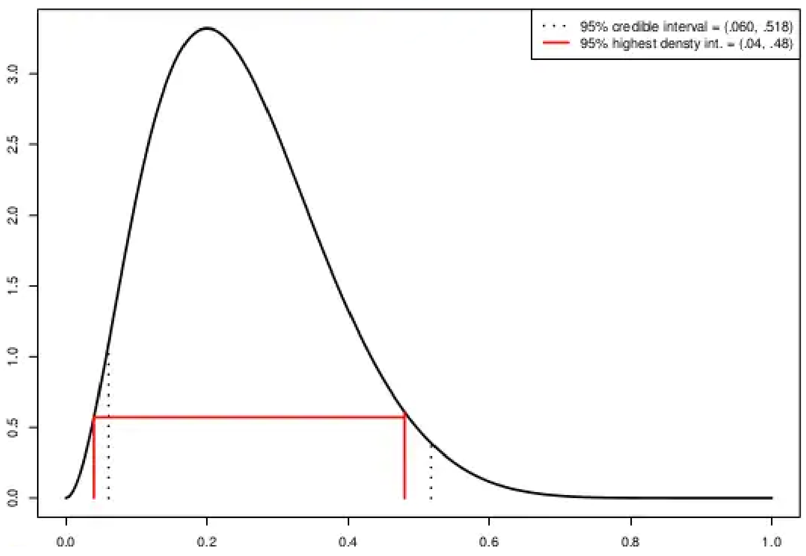
- **边缘分布**（Marginal distribution）

$$p(\theta_i\,|\,\mathcal{D}) = \frac{\int p\,(\boldsymbol{\theta}\,|\,\mathcal{D})\,\mathrm{d}\prod_{j\neq i}\theta_j}{\int p\,(\boldsymbol{\theta}\,|\,\mathcal{D})\,\mathrm{d}\prod_{j}\theta_j}$$

- **置信区间**（credible interval) [a,b]:

$$P(a \le \theta_i \le b) = \int_a^b p(\theta_i\,|\,\mathcal{D})\,dx_i = 1-\alpha$$

通常可选取 $P(\theta_i < a) = P(\theta_i > b) = \alpha/2$



- **平均值**（mean value):

$$\langle\theta_i\rangle = \int \theta_i\, p(\theta_i\,|\,\mathcal{D})\,d\boldsymbol{\theta}$$

- **中位数** (median value) U:

$$P(\theta_i < U) = 0.5$$

- **Maximum a posterior**

Bayesian inference-correlation analysis

- **联合分布**（joint distribution）

$$p[(\theta_i,\theta)\mid\mathcal{D}]=\frac{\int p\,(\boldsymbol{\theta}\mid\mathcal{D})\mathrm{d}\prod_{k\neq i,j}\theta_k}{\int p\,(\boldsymbol{\theta}\mid\mathcal{D})\mathrm{d}\prod_{k}\theta_k}$$

- **协方差**（covariance）：

$$Cov(\theta_i,\theta_j)=\int(\theta_i-\langle\theta_i\rangle)\,(\theta_j-\langle\theta_j\rangle)\,p[(\theta_i,\theta_j)\mid\mathcal{D}]\,d\theta_i d\theta_j$$

- **相关系数**（correlation coefficient)：

$$R=\frac{Cov(\theta_i,\theta_j)}{\sigma_i\sigma_j}$$

with $\sigma_i=\sqrt{\langle(\Theta_\mathrm{i}-\langle\Theta_i\rangle)^2\rangle}$ being the standard deviation of Θ_i

Technique difficulties

$$p(\boldsymbol{\theta} \mid \mathcal{D}) = \frac{p(\mathcal{D} \mid \boldsymbol{\theta})\, p(\boldsymbol{\theta})}{\int p(\mathcal{D} \mid \boldsymbol{\theta})\, p(\boldsymbol{\theta})\, d\boldsymbol{\theta}}$$

Difficulties:

- Non-analytical
- High-dimensional (many parameters)
- Slow model

Technique difficulties

$$p(\boldsymbol{\theta} \mid \mathcal{D}) = \frac{p(\mathcal{D} \mid \boldsymbol{\theta})\, p(\boldsymbol{\theta})}{\int p(\mathcal{D} \mid \boldsymbol{\theta})\, p(\boldsymbol{\theta})\, d\boldsymbol{\theta}}$$

Difficulties:

- Non-analytical
- High-dimensional (many parameters)

- Slow model

Solutions:

- Variational inference
- **Markov Chain Monte Carlo (MCMC)**

- **Emulator:**
  Gaussian process (GP)
  + Principal component analysis (PCA)
  (Dimensionality Reduction for multi-observables)

Metropolis-Hastings Algorithm

- **Metropolis-Hastings algorithm** (Metropolis et al. 1953, Hastings 1970):

  Construct a Markov Chain for θ by employing an auxiliary distribution that is easy to sample from.

- **Steps:**
  1. Choose a starting value $\theta^{(\mathrm{i})}$ ($i = 0$);
  2. Produce a candidate value θ^* according to a **proposal distribution** $q(\theta^*|\theta^{(i)})$;
  3. Calculate the **acceptance probability**

$$a(\theta^{(i)},\theta^*) = \min\left\{1, \frac{p(\theta^* \mid \mathcal{D})q(\theta^{(i)} \mid \theta^*)}{p(\theta^{(i)} \mid \mathcal{D})q(\theta^* \mid \theta^{(i)})}\right\}$$

Ratio independent of the normalization factor

  4. Generate $u \sim U(0,1)$,

     If u < a, let $\theta^{(i+1)} = \theta^*$, else $\theta^{(i+1)} = \theta^i$.
  5. i=i+1, go to the 2nd step.

- The Markov Chain converges to the stationary distribution $p(\boldsymbol{\theta} \mid \mathcal{D})$

Random-walk Metropolis-Hastings

- Take $q(\theta^*|\theta^{(i)})$ to be **symmetric**, i.e.,

$$q(\theta^* \mid \theta^{(i)}) = q(\theta^{(i)} \mid \theta^*) = q(|\,\theta^* - \theta^{(i)}\,|).$$

  Popular choices are (multivariate) Gaussians or t-distributions.

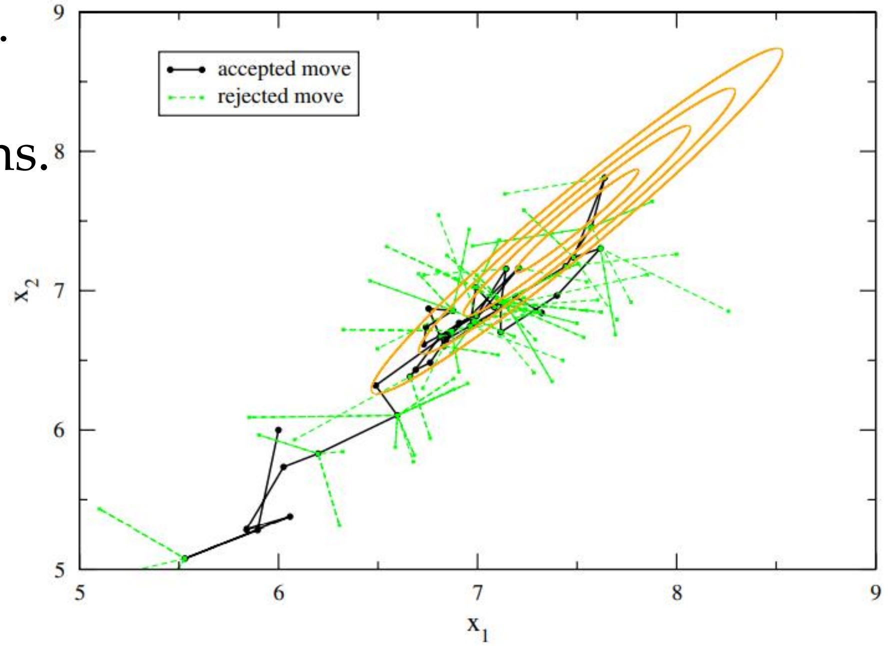
- Draw $\epsilon \sim q$, $\theta^* = \theta^{(i)} + \epsilon$
- **Acceptance probability**

$$a(\theta^{(i)}, \theta^*) = \min\left\{1, \frac{p(\theta^* \mid \mathcal{D})}{p(\theta^{(i)} \mid \mathcal{D})}\right\}$$

- **Scale parameter** controls the acceptance rates.

- **Q: convergency? Autocorrelation?**

  Interactive MCMC Sampling Visualizer by Chi Feng
  https://chi-feng.github.io/mcmc-demo



U. von Toussaint, Rev. Mod. Phys. 83. 943 (2011)

Gaussian process

The Gaussian Processes Web Site : http://www.gaussianprocess.org/

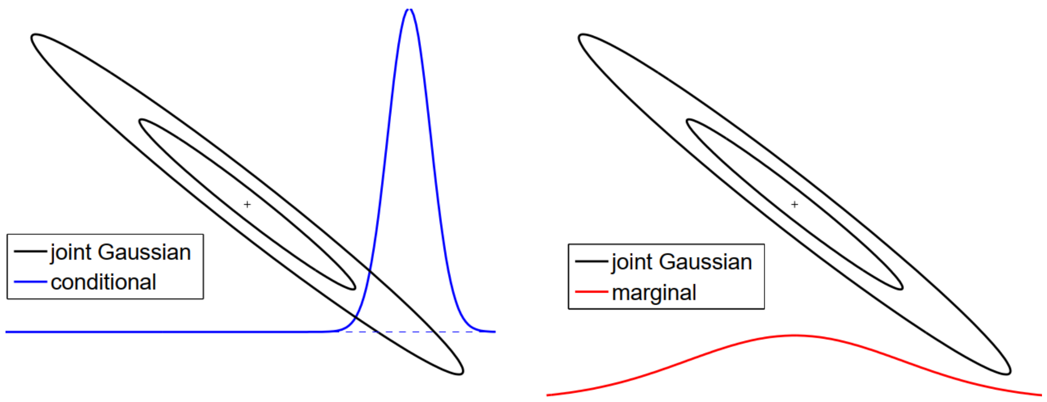
- **Definition**: *A Gaussian process is a collection of random variables, any finite number of which have a joint Gaussian distribution.*

  Rasmussen and Williams, *Gaussian Processes for Machine Learning*, MIT Press, 2006

- **Multivariate Gaussian distribution**

$$p\left(\boldsymbol{x};\boldsymbol{\mu},\Sigma\right)=\frac{1}{\left(2\pi\right)^{d/2}|\Sigma|^{1/2}}\exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)$$

where μ is the mean vector and Σ is the covariance matrix.



Conditional and marginal distributions of a multivariate Gaussian are gaussian

Gaussian process emulator

- **GP** can serve as a fast surrogate of a slow model.

  Given training data from model calculations, a GP can guess model prediction y_* at x_*.

  $X = (\vec{x}_1, \cdots, \vec{x}_m),\ \boldsymbol{y} = (y_1, \cdots, y_m)$

- Considering $\boldsymbol{y}$ and y_* as **random variables**, seek for $p(y_* | \boldsymbol{y})$.
  - y_* and $\boldsymbol{y}$ should be correlated and the correlation depends on $|x_* - x_i|$.

- **Prior:**

Covariance matrix

$$\begin{bmatrix} \boldsymbol{y} \\ y_* \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma^0(X,X) & \Sigma^0(X,\vec{x}_*) \\ \Sigma^0(\vec{x}_*,X) & \Sigma^0(\vec{x}_*,\vec{x}_*) \end{bmatrix} \right)$$

- Conditional distribution:

$$y_* \mid \boldsymbol{y} \sim \mathcal{N}(\mu_*, \sigma_*^2)$$

$$\mu_* \mid \boldsymbol{y} = \Sigma^0(x_*,X)\Sigma^0(X,X)^{-1}\boldsymbol{y}$$

$$\sigma_*^2 = \sigma(x_*,x_*) - \Sigma^0(x_*,X)\Sigma^0(X,X)^{-1}\Sigma^0(X,x_*)$$

Gaussian process emulator

A typical covariance function:

$$\Sigma^0(X,X)=\begin{pmatrix}\sigma(\vec{x}_1,\vec{x}_1) & \cdots & \sigma(\vec{x}_1,\vec{x}_m)\\ \vdots & \ddots & \vdots \\ \sigma(\vec{x}_m,\vec{x}_1) & \cdots & \sigma(\vec{x}_m,\vec{x}_m)\end{pmatrix},\quad \sigma(\vec{x}_i,\vec{x}_j)=\sigma_{\mathrm{GP}}^2\exp\left(-\sum_k\frac{(x_{i,k}-x_{j,k})^2}{2l_k^2}\right)+\sigma_n^2\delta_{ij}$$

Decrease with distance

Hyper-parameters:

- Noise σ_n = uncertainty of a stochastic model
- Amplitude σ_{GP} only affects σ_* but not affect μ_* if σ_n=0
- Length scale l_k : usually μ_* is relatively insensitive to l_k

- Reproduce the training points without noise
- Make predictions with statistical uncertainties
- Values of *hyper-parameters*?

Maximize marginal likelihood

Bernhard et al., PRC 91, 054910 (2015)

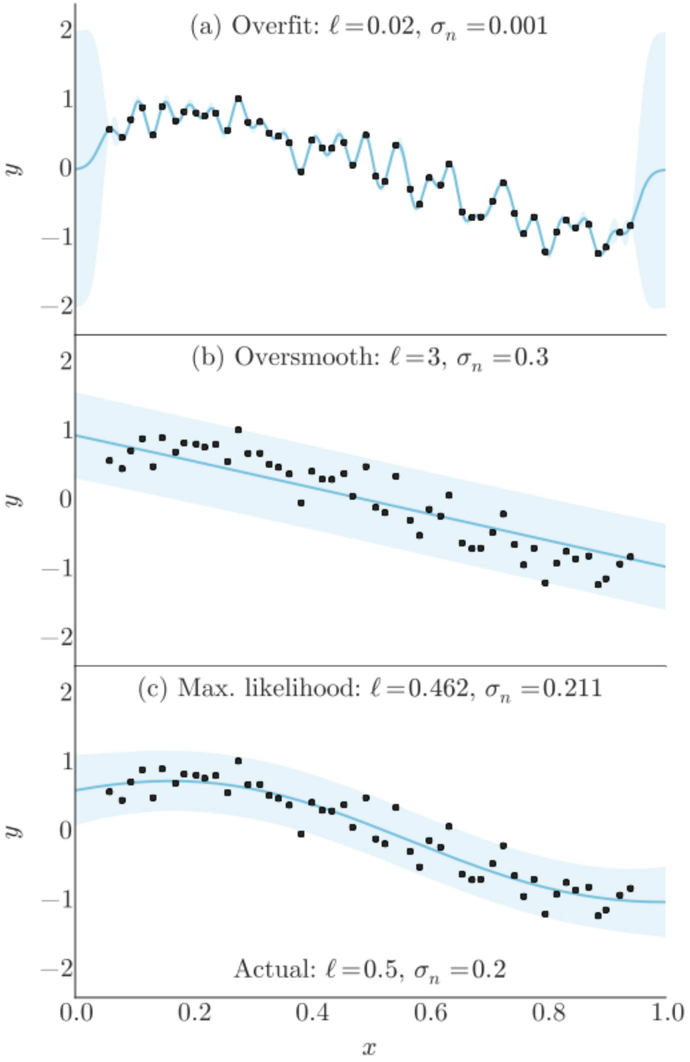
- Log marginal likelihood :

$$\log P(\mathbf{y} \mid X, \theta) = -\frac{1}{2}\mathbf{y}^{\top}(\Sigma^0)^{-1}\mathbf{y} - \frac{1}{2}\log|\Sigma^0| - \frac{m}{2}\log 2\pi$$

- Hyper-parameters can be determined by maximizing the marginal likelihood
- Example:

$$\sigma(x, x') = \exp\left(-\frac{|x - x'|^2}{2\ell^2}\right) + \sigma_n^2 \delta_{xx'}$$

$$\mathcal{GP}(x_*, (X, \boldsymbol{y})) \sim \mathcal{M}(x_*)$$



- GP for a single output. Multiple output case?

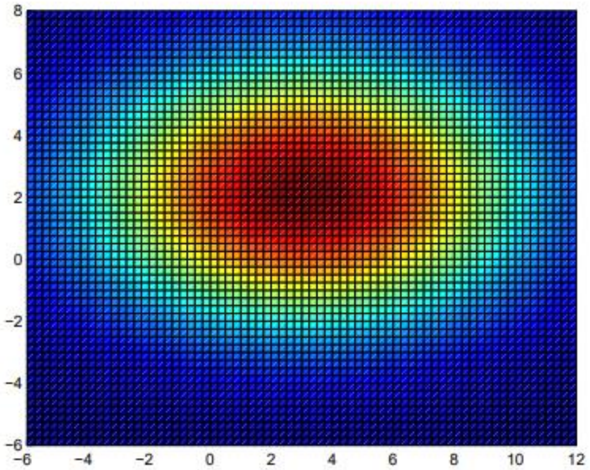
Principal component analysis

The central idea of principal component analysis is to reduce the dimensionality of a data set in which there are a large number of **interrelated variables**, while retaining as much as possible of the variation present in the data set. This reduction is achieved by transforming to a new set of variables, the principal components, which are uncorrelated, and which are ordered so that the first few retain most of the variation present in all of the original variables.

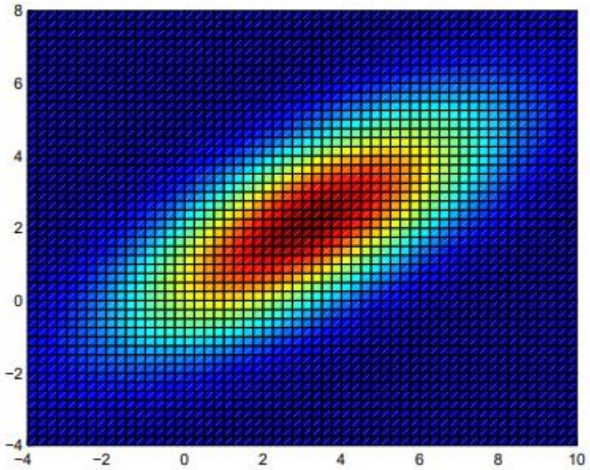
I.T. Jolliffe, *Principal Component Analysis*, Second Edition

uncorrelated

$$\mathcal{N}\left(\mu=\begin{bmatrix}3\\2\end{bmatrix},\Sigma=\begin{bmatrix}25 & 0\\0 & 9\end{bmatrix}\right)$$



correlated



$$\mathcal{N}\left(\mu=\begin{bmatrix}3\\2\end{bmatrix},\Sigma=\begin{bmatrix}10 & 5\\5 & 5\end{bmatrix}\right)$$

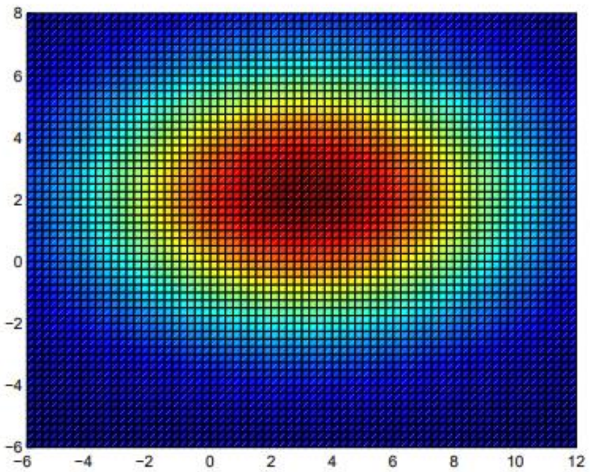
Principal component analysis

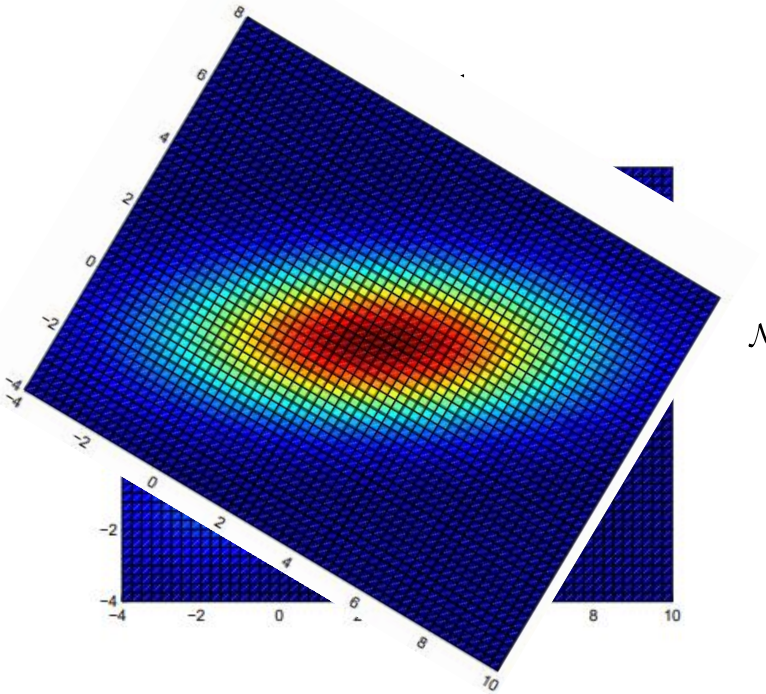
The central idea of principal component analysis is to reduce the dimensionality of a data set in which there are a large number of **interrelated variables**, while retaining as much as possible of the variation present in the data set. This reduction is achieved by transforming to a new set of variables, the principal components, which are uncorrelated, and which are ordered so that the first few retain most of the variation present in all of the original variables.

I.T. Jolliffe, *Principal Component Analysis*, Second Edition

uncorrelated

$$\mathcal{N}\left(\mu=\begin{bmatrix}3\\2\end{bmatrix},\Sigma=\begin{bmatrix}25&0\\0&9\end{bmatrix}\right)$$





$$\mathcal{N}\left(\mu=\begin{bmatrix}3\\2\end{bmatrix},\Sigma=\begin{bmatrix}10&5\\5&5\end{bmatrix}\right)$$

Principal component analysis

- m correlated observables/model output o_i
- o_i are functions of model parameters, but linear combinations of o_i may be independent of parameters.

$$\tilde{o}_i = \frac{o_i - \langle o_i \rangle}{\sigma_i}$$ average over training data

- Covariance matrix:

$$M = \langle \tilde{o}_i \tilde{o}_j \rangle \ , \quad (\langle \tilde{o}_i \rangle = \langle \tilde{o}_j \rangle = 0)$$

- Eigenvalue decomposition

$$M = U \Lambda U^T$$

$$\Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m \end{bmatrix}$$

Eigenvalue

$$U = (\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_m)$$

Normalized Eigenvectors

Principal component analysis

$$M = U \Lambda U^T$$
$$UU^T = I$$

Covariance Matrix

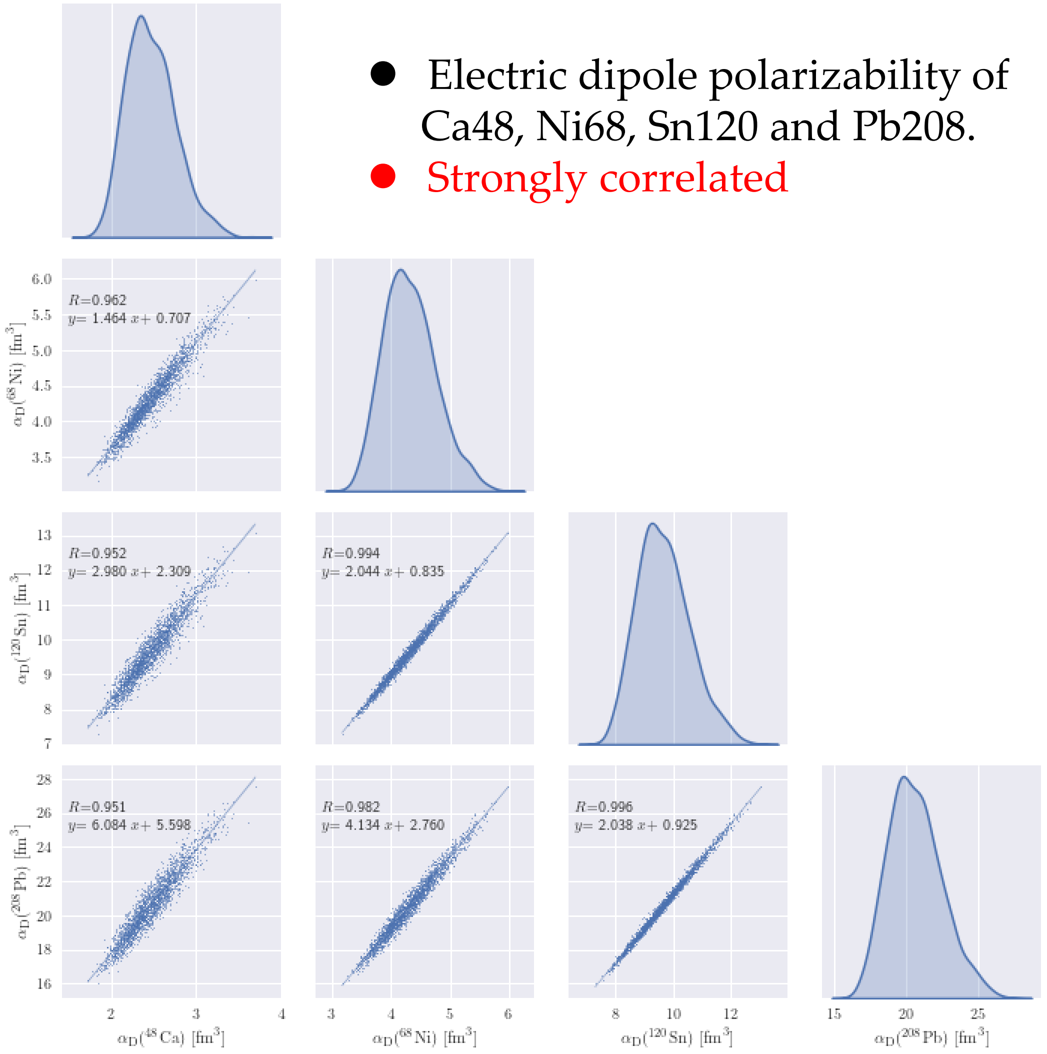
Physics output	Principal component
$\tilde{\boldsymbol{o}}$	$\boldsymbol{z} = \tilde{\boldsymbol{o}} U$
$\langle \tilde{\boldsymbol{o}} \rangle = 0$	$\langle \boldsymbol{z} \rangle = 0$
$\langle \tilde{o}_i \tilde{o}_j \rangle = M$	$\langle z_i z_j \rangle = U^T \langle \tilde{o}_i \tilde{o}_j \rangle U = \Lambda$
Correlated	Uncorrelated

$$\mathcal{L} \propto \exp\left[-\sum_i \frac{(o_i - o_i^{\exp})^2}{2\sigma_i^2}\right]$$
$$= \exp\left[-\frac{1}{2}\sum_i (\tilde{o}_i - \tilde{o}_i^{\exp})^2\right]$$
$$= \exp\left[-\frac{1}{2}\sum_i (z_i - z_i^{\exp})^2\right]$$

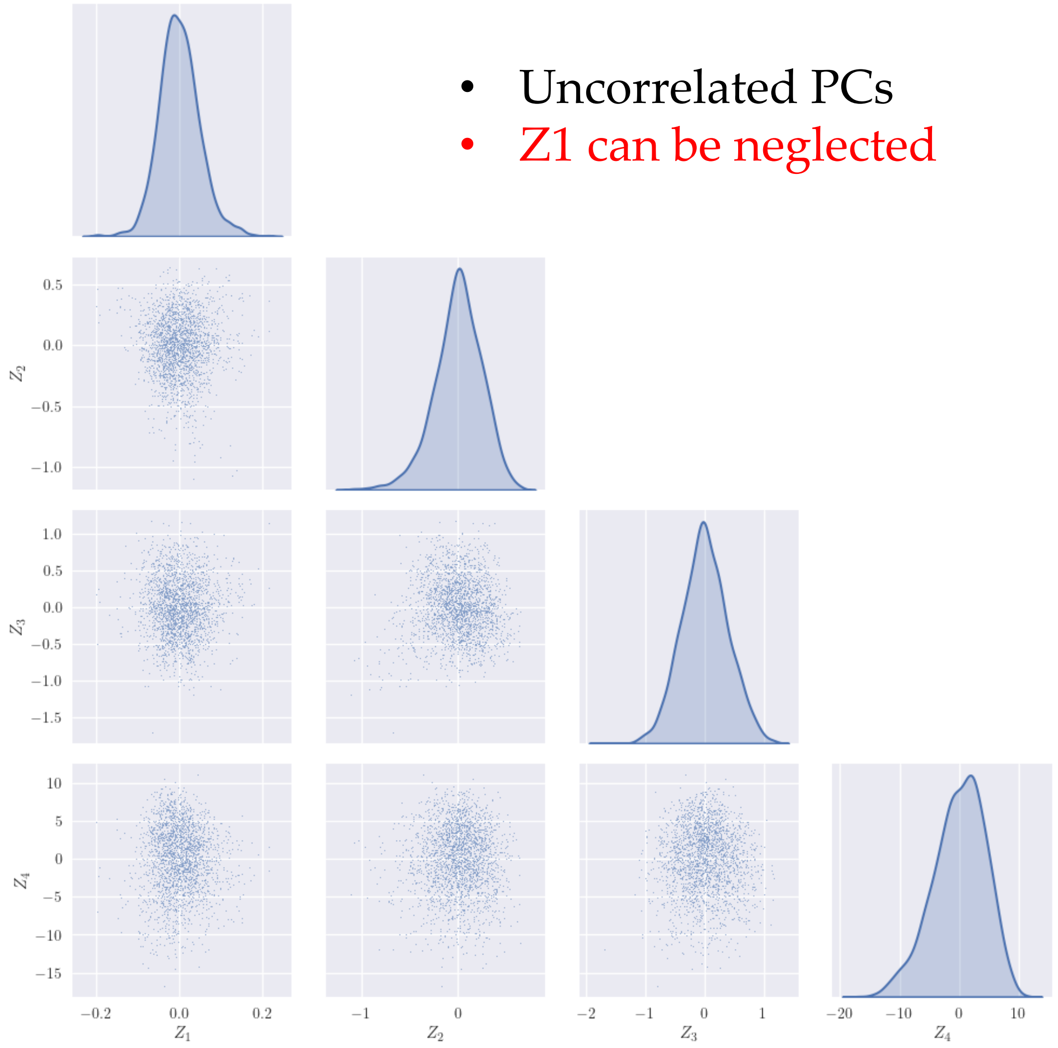
- λ_i is the variance of z_i.
- Retain the first q PCs with the largest eigenvalues.
- m correlated outputs → q uncorrelated PCs.

$$\sum_{i=1}^{q} \lambda_i \Big/ \sum_{i=1}^{m} \lambda_i \leq V$$

PCA-examples



- Electric dipole polarizability of Ca48, Ni68, Sn120 and Pb208.
- Strongly correlated

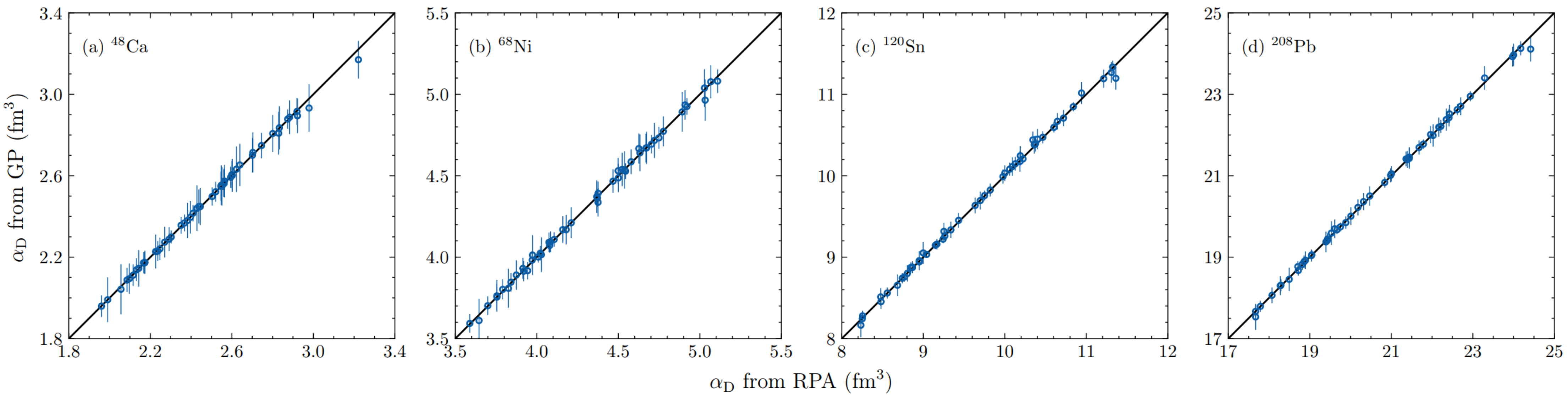


- Uncorrelated PCs
- Z1 can be neglected

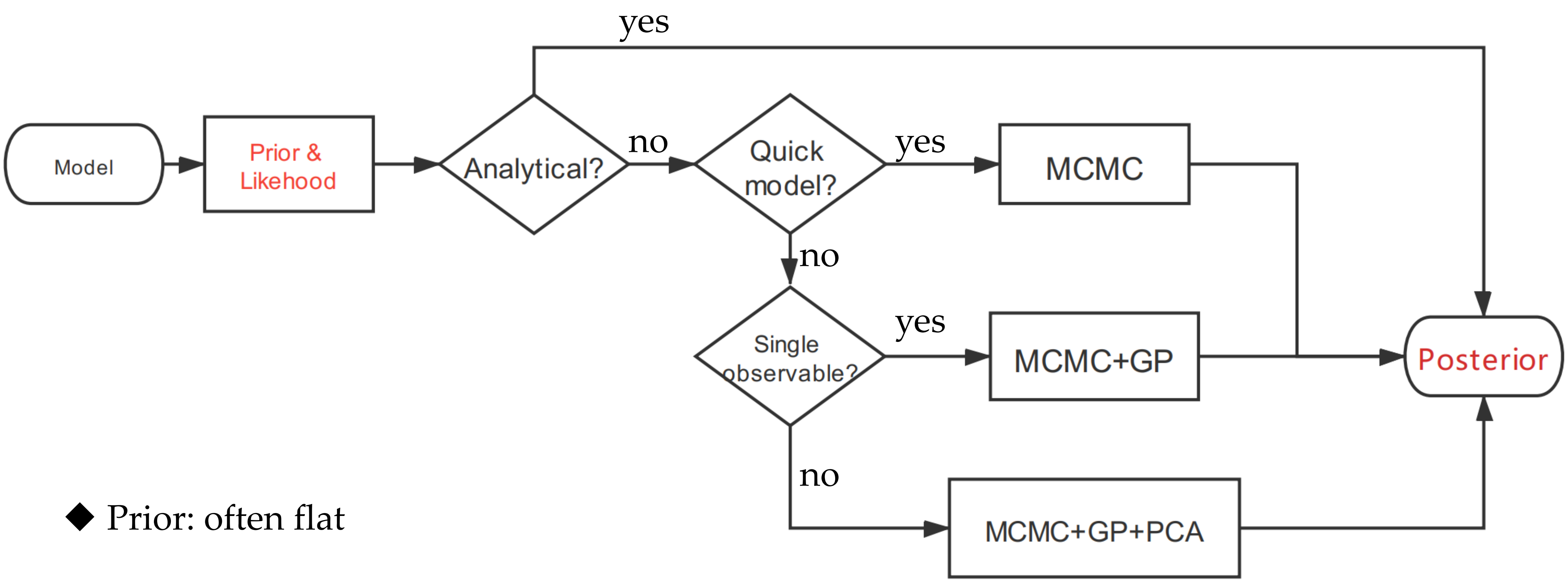
~ 0.2 ~ 1 ~ 1 ~ 10

PCA+GP

GP vs Model



◆ Tune GPs for dominant PCs instead of all model outputs.
◆ PCA+GPs => a multivariate emulator.



- Prior: often flat
- Likelihood: often Gaussian

Applications in nuclear physics

Reconstruction of impact parameter

PHYSICAL REVIEW C **104**, 034609 (2021)

Model independent reconstruction of impact parameter distributions for intermediate energy heavy ion collisions

J. D. Frankland,[1,*] D. Gruyer,[2] E. Bonnet,[3] B. Borderie,[4] R. Bougault,[2] A. Chbihi,[1] J. E. Ducret,[1] D. Durand,[2] Q. Fable,[2] M. Henri,[1] J. Lemarié,[1] N. Le Neindre,[2] I. Lombardo,[5] O. Lopez,[2] L. Manduci,[2,6] M. Pârlog,[2,7] J. Quicray,[2] G. Verde,[5,8] E. Vient,[2] and M. Vigilante[9]
(INDRA Collaboration)

X: an observable depending on b

$$P(X) = \int_0^\infty P(b)\,P(X|b)\,db.$$

Centrality: $$c_b \equiv \int_0^b P(b')\,db',$$

$$P(X) = \int_0^1 P(c_b)\,P(X|c_b)\,dc_b = \int_0^1 P(X|c_b)\,dc_b$$

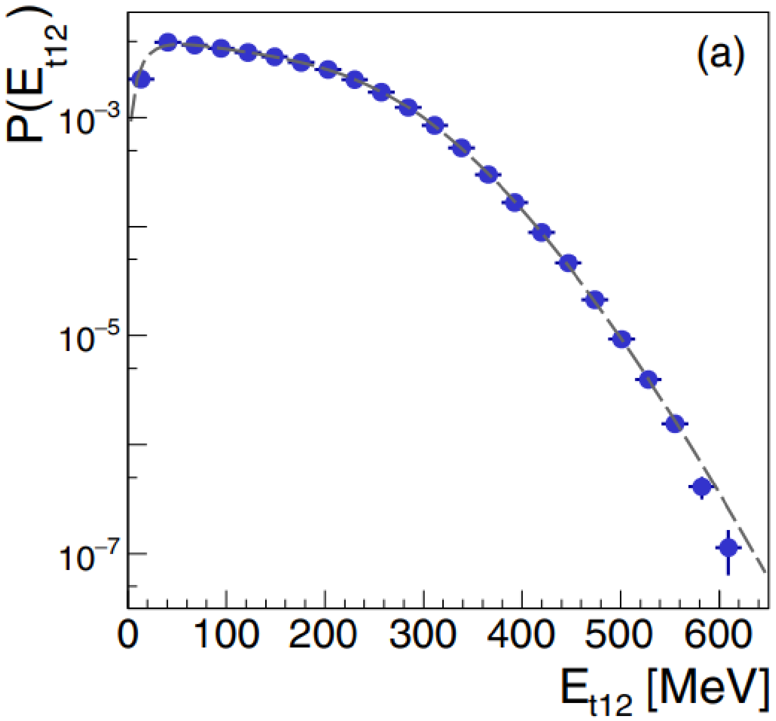
$$P(c_b, X) = P(c_b|X)P(X) = P(X|c_b)P(c_b)$$

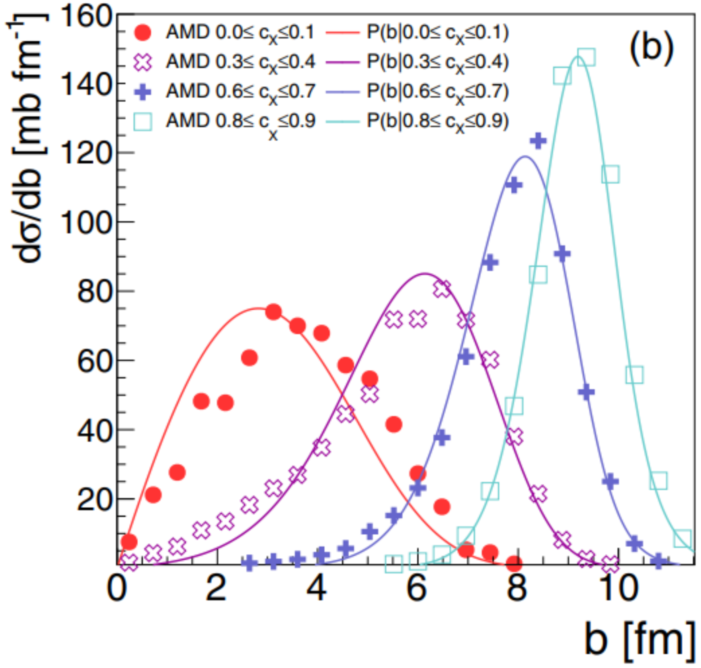
$$P(c_b|x_1 \leqslant X \leqslant x_2) = \frac{\int_{x_1}^{x_2} P(c_b, X)\,dX}{\int_{x_1}^{x_2} P(X)\,dX} = \frac{\int_{x_1}^{x_2} P(c_b|X)P(X)\,dX}{\int_{x_1}^{x_2} P(X)\,dX}$$

Reconstruction of impact parameter

$$P(X|c_b) = \frac{1}{\Gamma(k)\theta^k} X^{k-1} e^{-X/\theta},\ \ k(c_b) = k_{max}(1 - c_b^{\alpha})^{\gamma} + k_{min}$$

5 parameters determined by fitting experimental data

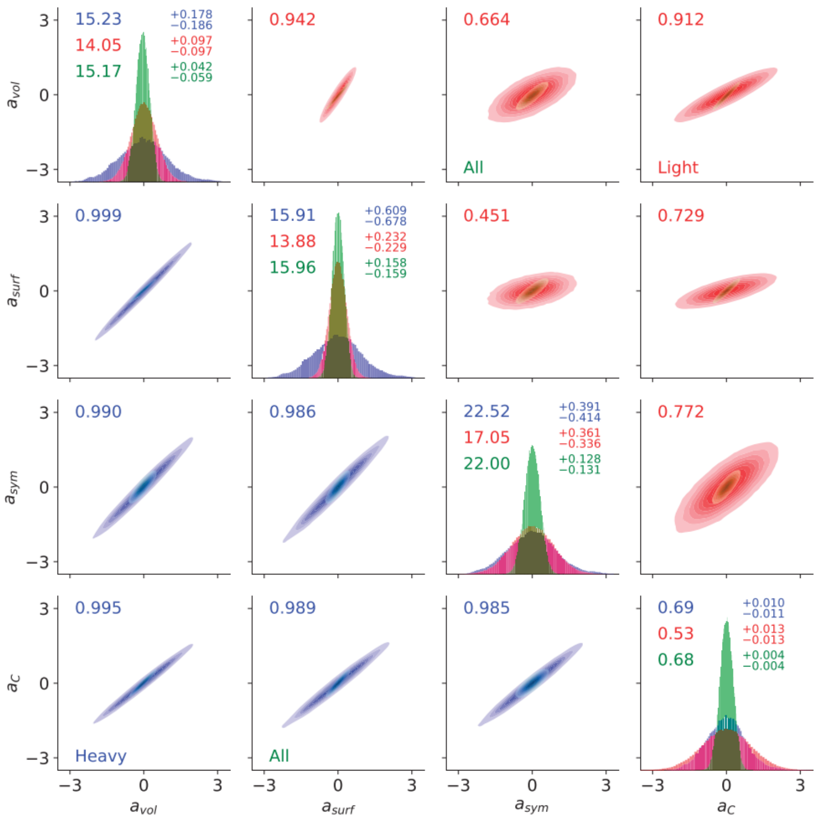




Liquid drop model

- The semi-empirical mass formula of the LDM parametrizes the binding energy of the nucleus (Z, N)as:

$$E_{\mathrm{LDM}}(N,Z) = a_{\mathrm{vol}}A - a_{\mathrm{surf}}A^{2/3} - a_{\mathrm{sym}}\frac{(N-Z)^2}{A} - a_{\mathrm{C}}\frac{Z(Z-1)}{A^{1/3}}$$



Kejzlar et al., JPG 47, 094001 (2020)

- Prior: Gaussian with larger variances
- Likelihood: exponential square
- Data:

(a) LDM(A) – LDM fitted on all 595 even–even nuclei.
(b) LDM(L) – LDM restricted to the light domain (153 nuclei).
(c) LDM(H) – LDM restricted to the heavy domain (287 nuclei).

- MCMC

Nuclear energy density functional (EDF)

Energy density functional : $E(\rho_n, \rho_p, \boldsymbol{s}_n, \boldsymbol{s}_p, \cdots)$

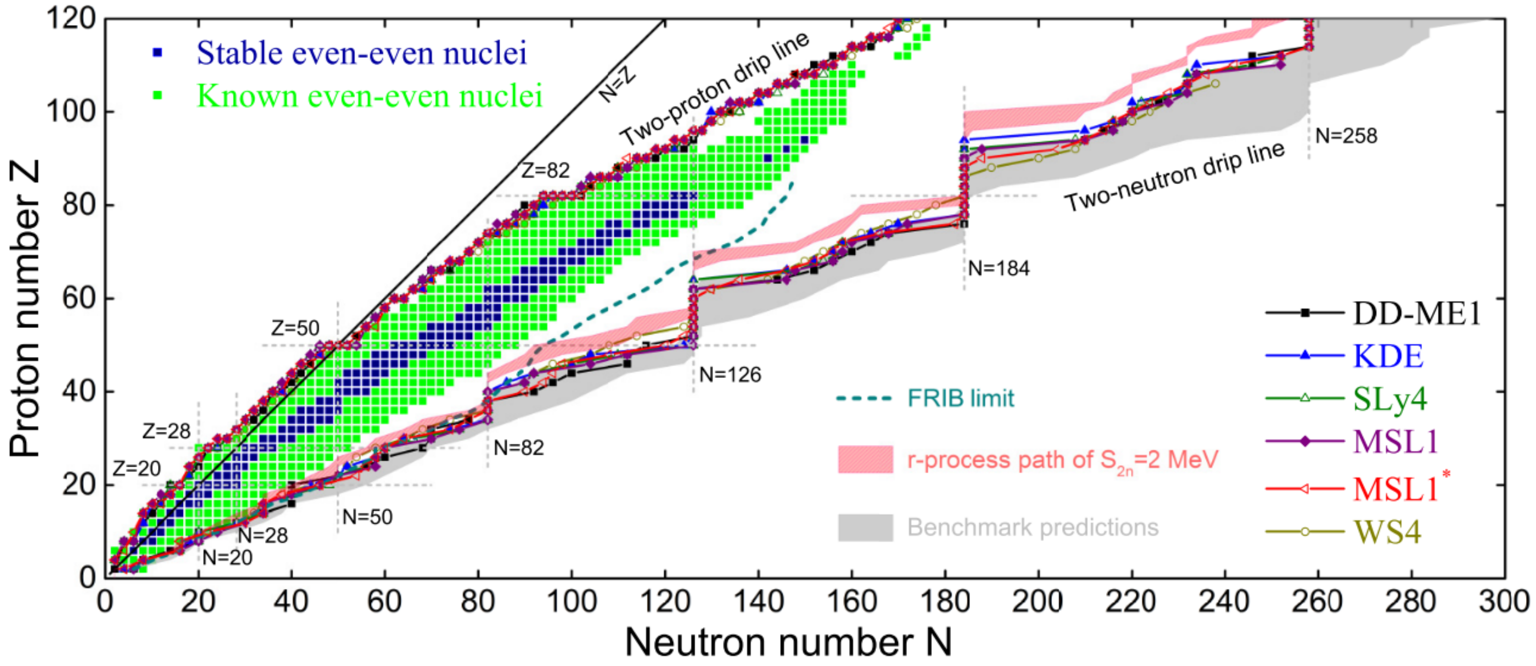
- Nonrelativistic: Skyrme, Gogny
- Relativistic: Relativistic mean-field.

…

Applications :

- Density profile
- Mass formula
- Neutron Drip line
- Fission barrier
- Neutron stars

…



Wang and Chen, PRC 92, 031303(R) (2015)

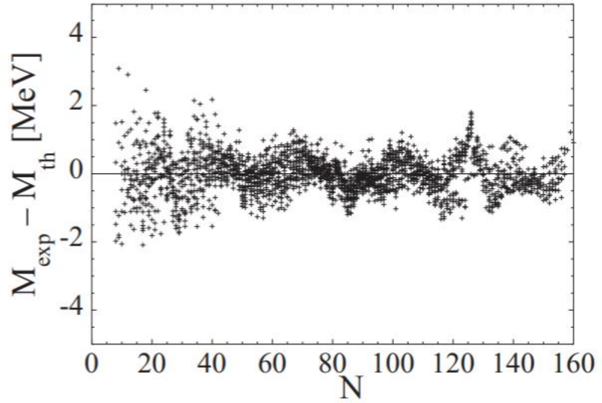


FIG. 2. Difference between measured [16] and HFB-31 masses as a function of the neutron number *N*.

Goriely et al., PRC93, 034337 (2016)

Nuclear energy density functional

PRL 114, 122501 (2015) PHYSICAL REVIEW LETTERS week ending 27 MARCH 2015

Uncertainty Quantification for Nuclear Density Functional Theory and Information Content of New Measurements

J. D. McDonnell,[1,2] N. Schunck,[2] D. Higdon,[3] J. Sarich,[4] S. M. Wild,[4] and W. Nazarewicz[5,6,7]

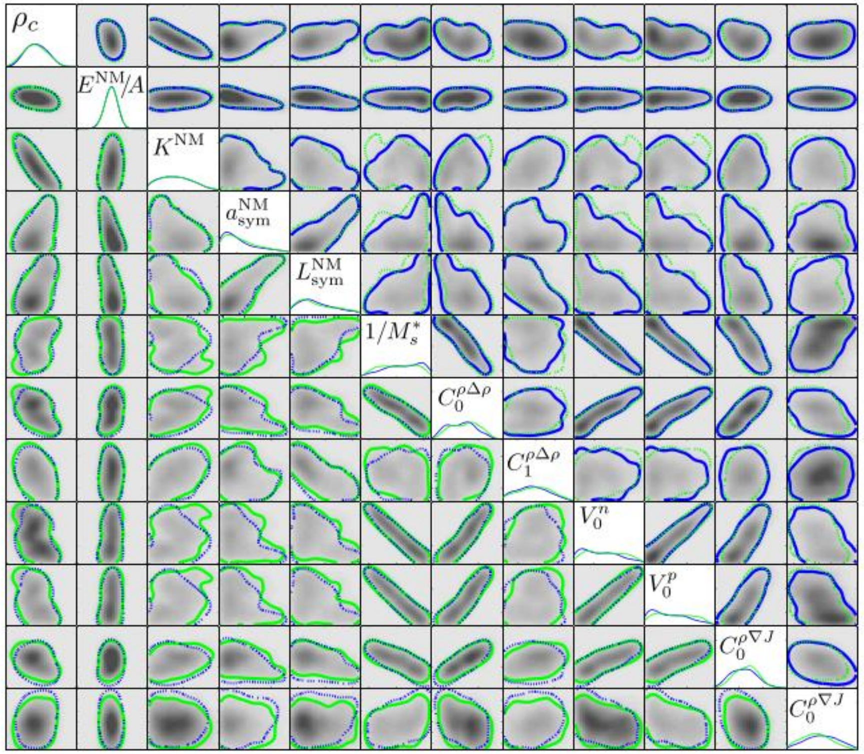
- Model: Skyrme EDF with 12 parameters.
- Data: 115 (masses, radii…)+17 new mass measurements.
- Prior: Uniform
- Computing χ^2 requires 5 min with over 800 cores

$$\chi^2(\boldsymbol{x}) = \frac{1}{n_d - n_x} \sum_{t=1}^{n_T} \sum_{j=1}^{n_t} \left(\frac{y_{tj}(\boldsymbol{x}) - d_{tj}}{\sigma_t} \right)^2.$$

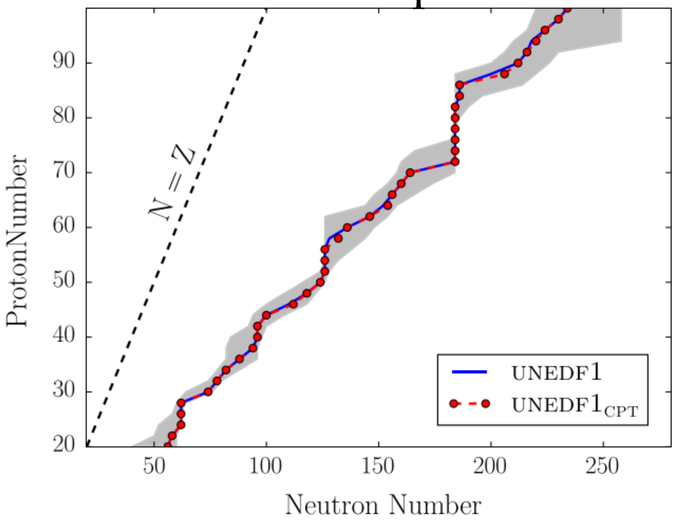
- Gaussian process based on 200 training points.
- (PCA+)GP+MCMC

Higdon et al., J. Phys. G 42, 034009 (2015).

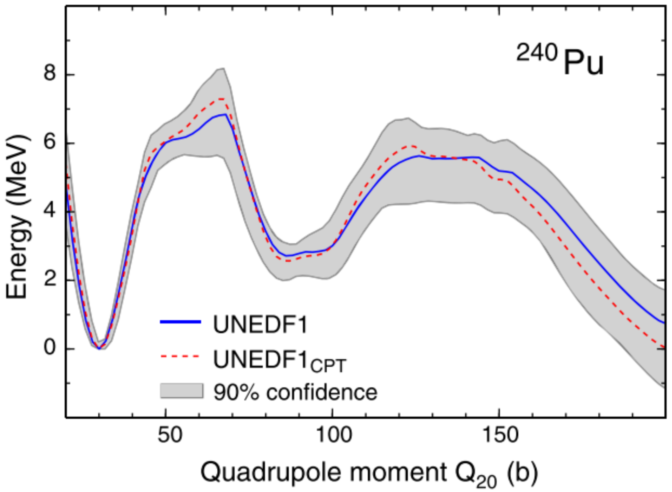
Nuclear energy density functional



Neutron drip line



Fission barrier



- Univariate and bivariate marginal estimates of the posterior distribution
- Blue: 115 data points; Green: 115 +17

Nuclear equation of state

- Nuclear EoS: binding energy per nucleon in nuclear matter

$$E = E(\rho,\delta) = E_0(\rho) + E_{\text{sym}}(\rho)\delta^2 + E_{\text{sym},4}(\rho)\delta^4 + \mathrm{O}(\delta^6),\ \ \delta = (\rho_n - \rho_p)/\rho$$

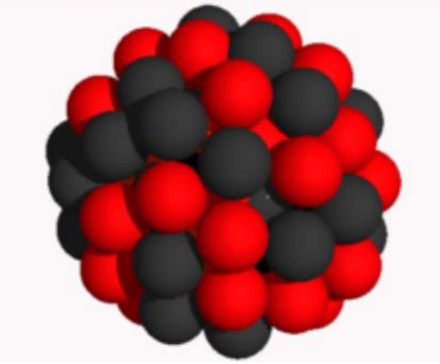
- EoS of symmetric nuclear matter.

$$E_0(\rho) = E_0(\rho_0) + \frac{1}{2!}K_0\chi^2 + \frac{1}{3!}J_0\chi^3 + \mathcal{O}(\chi^4),\ \ \chi = (\rho-\rho_0)/3\rho_0$$

- Symmetry energy

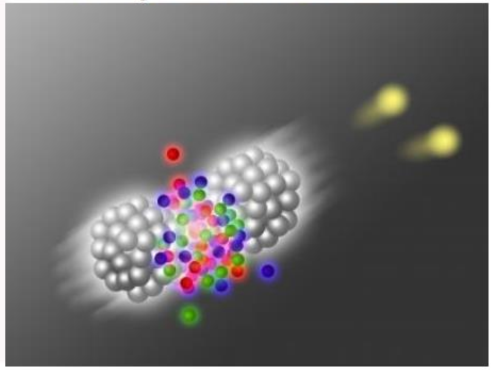
$$E_{\text{sym}}(\rho) = \frac{1}{2}\frac{\partial^2 E(\rho,\delta)}{\partial\delta^2}\bigg|_{\delta=0} = E_{\text{sym}}(\rho_0) + L\chi + \frac{1}{2!}K_{\text{sym}}\chi^2 + \mathcal{O}(\chi^3)$$

finite nulei



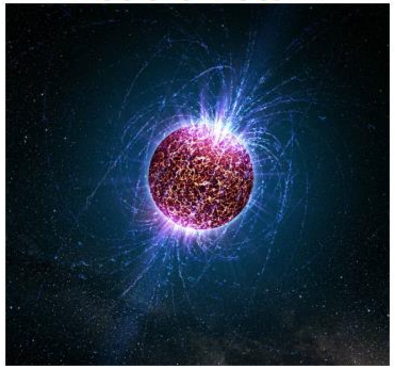
$\leq \rho_0$

heavy ion collision



a wide density range

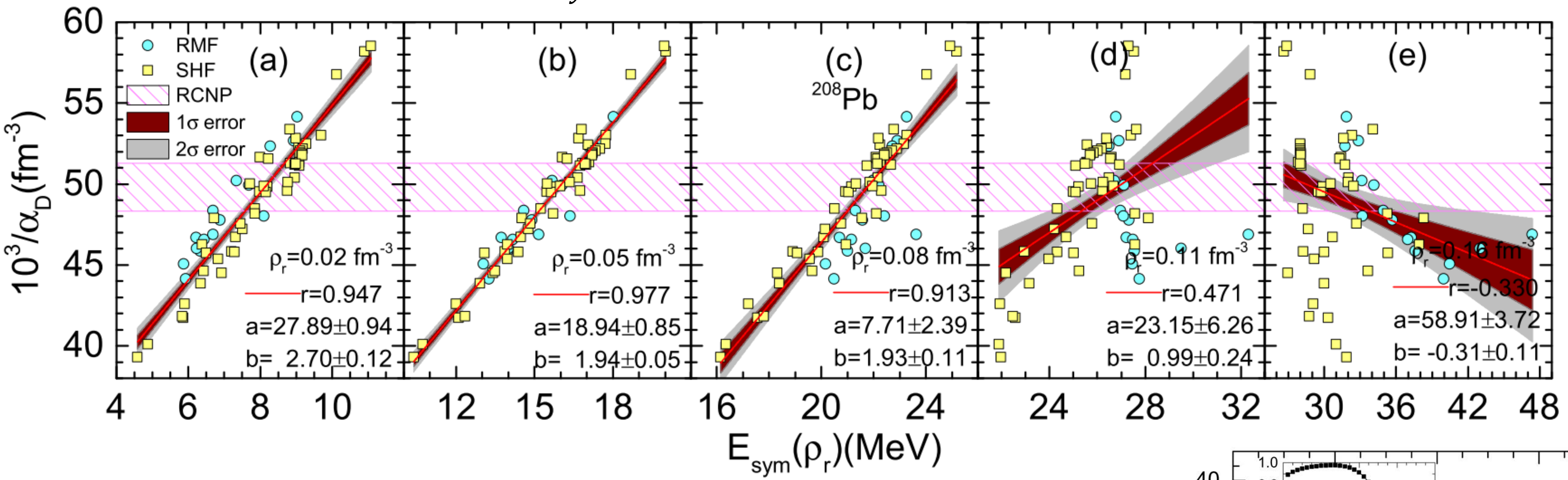
neutron star



>2~ 3 ρ_0

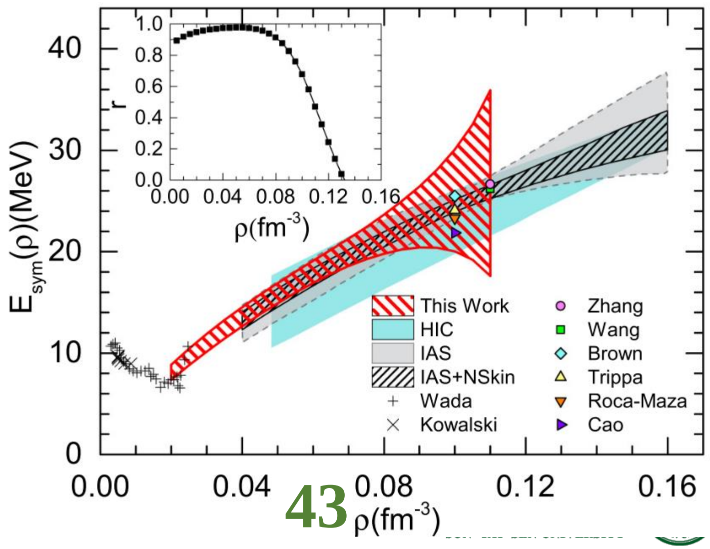
Esym from nuclear dipole polarizability

$1/\alpha_D$-$E_{sym}(\rho)$ at different densities



Zhang and Chen PRC 92, 031301(R) (2015)

- Strong linear correlation at $\rho_r = 0.05\ \mathrm{fm}^{-3}$.
- Value and uncertainty are extracted from linear fits.



Physics Letters B 810 (2020) 135820

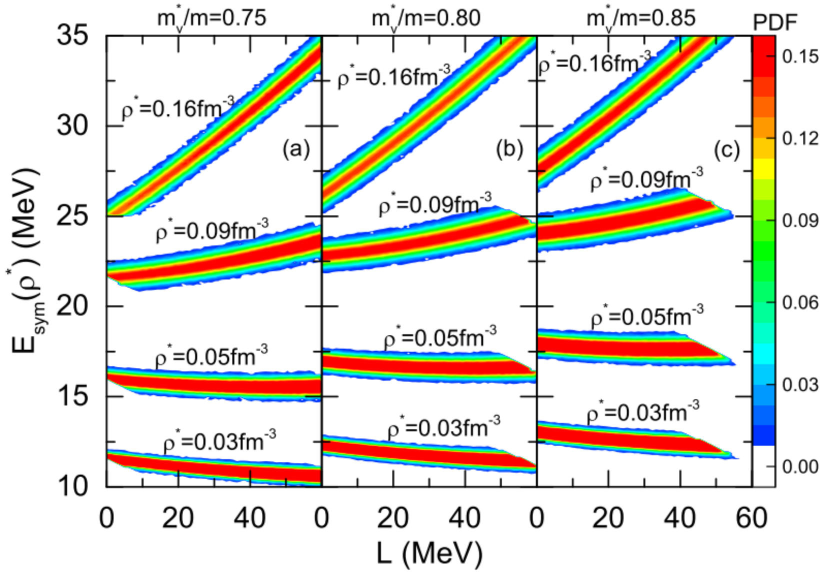
Contents lists available at ScienceDirect

Physics Letters B

www.elsevier.com/locate/physletb

Constraining isovector nuclear interactions with giant resonances within a Bayesian approach

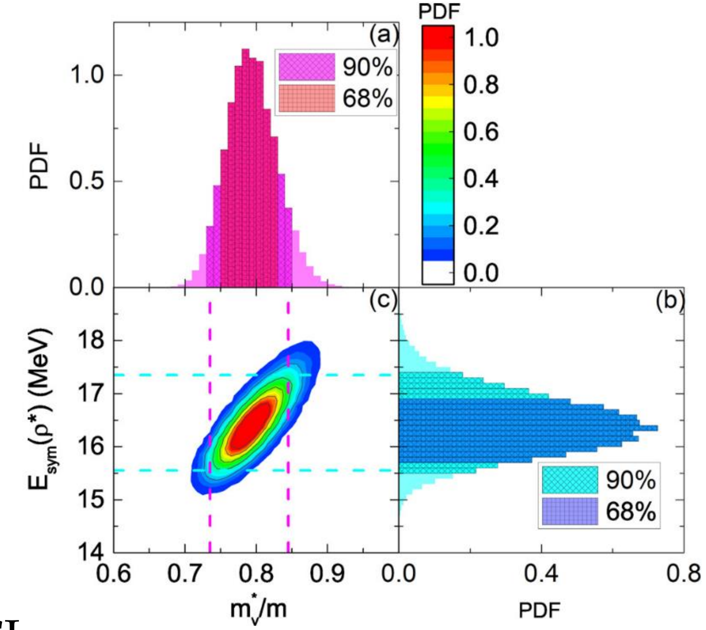
Jun Xu[a,b,*], Jia Zhou[b,c], Zhen Zhang[d], Wen-Jie Xie[e], Bao-An Li[f]



- Skyrme EDF: 3 isovector parameters
- Data: α_D and GDR energy of Pb208
- Likelihood:

$$P[D(d_{1,2})|M(p_{1,2,3})] = \frac{1}{2\pi\sigma_1\sigma_2}\exp\left[-\frac{(d_1^{th}-d_1^{exp})^2}{2\sigma_1^2}-\frac{(d_2^{th}-d_2^{exp})^2}{2\sigma_2^2}\right]$$

- MCMC



$$m_v^*/m = 0.79 \pm 0.06,\ E_{sym}(0.05\ \mathrm{fm}^{-3}) = 16.4^{+1.0}_{-0.9}\ \mathrm{MeV},\ 90\%\ \mathrm{CI}$$

Chinese Physics C Vol. 45, No. 6 (2021) 064104

Bayesian inference on isospin splitting of nucleon effective mass from giant resonances in ^{208}Pb*

Zhen Zhang(张振)[1†] Xue-Bin Feng(冯学彬)[1] Lie-Wen Chen(陈列文)[2‡]

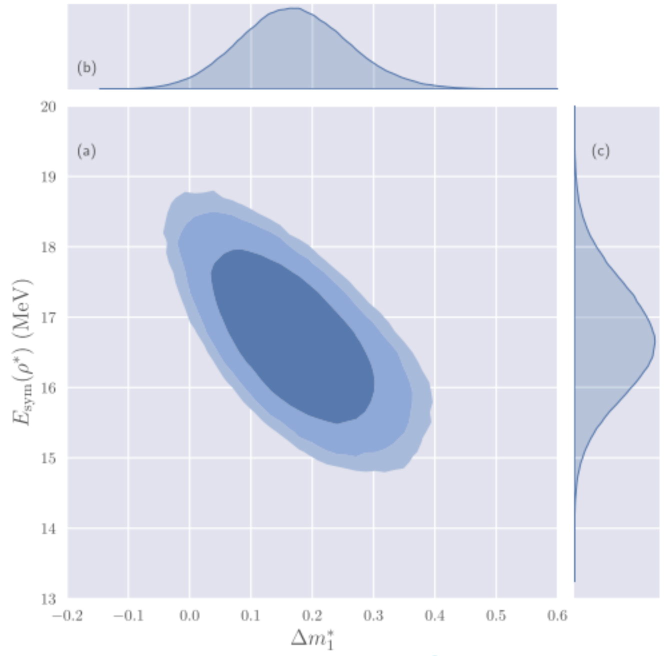
Table 2. Experimental values and uncertainties used for the binding energy E_B [53], charge radius r_C [54], breathing mode energy E_{GMR} [55], neutron $3p_{1/2}-3p_{3/2}$ energy level splitting ϵ_{ls} [56], electric dipole polarizability α_D [49, 50], IVGDR constrained energy [51], and ISGQR peak energy [52] in ^{208}Pb.

	value	σ
E_B/MeV	−1363.43	0.5
r_C/fm	5.5012	0.01
E_{GMR}/MeV	13.5	0.1
ϵ_{ls}/MeV	0.89	0.09
α_D/fm^3	19.6	0.6
E_{GDR}/MeV	13.46	0.1
E_{GQR}/MeV	10.9	0.1

- Data: 7 observables in Pb208.
- Uniform prior; exponential square likelihood.
- 2476 full model calculations as training data.

- PCA+GP+MCMC

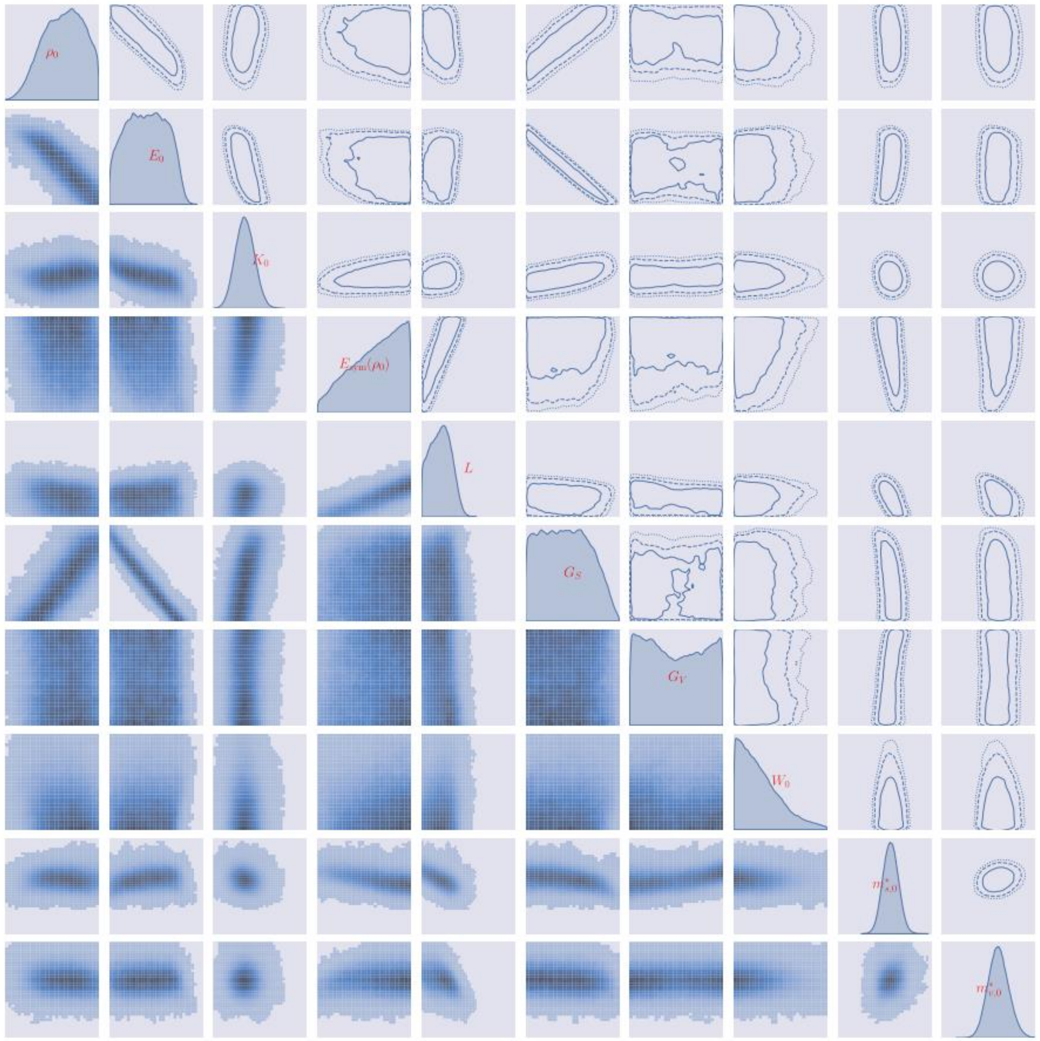
$$m_v^*/m = 0.78^{+0.06}_{-0.05}, \; E_{sym}\left(0.05 \text{ fm}^{-3}\right) = 16.7 \pm 1.3 \text{ MeV, 90\% CI}$$



- Correlation coefficients from 2476 training points.

	α_D	E_{GDR}	E_{GQR}	E_B	r_C	E_{GMR}	ϵ_{ls}
ρ_0	-0.054	0.052	0.146	0.166	-0.662	0.072	0.266
E_0	0.025	-0.009	-0.040	0.767	0.024	-0.123	-0.063
K_0	-0.063	0.062	0.089	0.174	-0.246	0.883	0.094
$E_{\text{sym}}(\rho_0)$	-0.291	0.253	-0.093	0.178	-0.024	-0.156	-0.101
L	0.889	-0.807	0.024	-0.233	-0.194	0.059	-0.189
G_S	0.180	-0.186	-0.199	0.463	0.568	-0.323	-0.236
G_V	0.072	-0.050	0.066	-0.025	-0.066	0.045	0.048
W_0	-0.019	0.019	0.094	-0.074	-0.168	0.067	0.719
$m^*_{s,0}$	0.127	-0.122	-0.923	0.254	0.262	-0.204	-0.542
$m^*_{v,0}$	-0.019	-0.474	0.118	-0.025	0.016	0.032	0.067

- Univariate and bivariate marginal estimates of the posterior distribution



K_0 from breathing mode energy

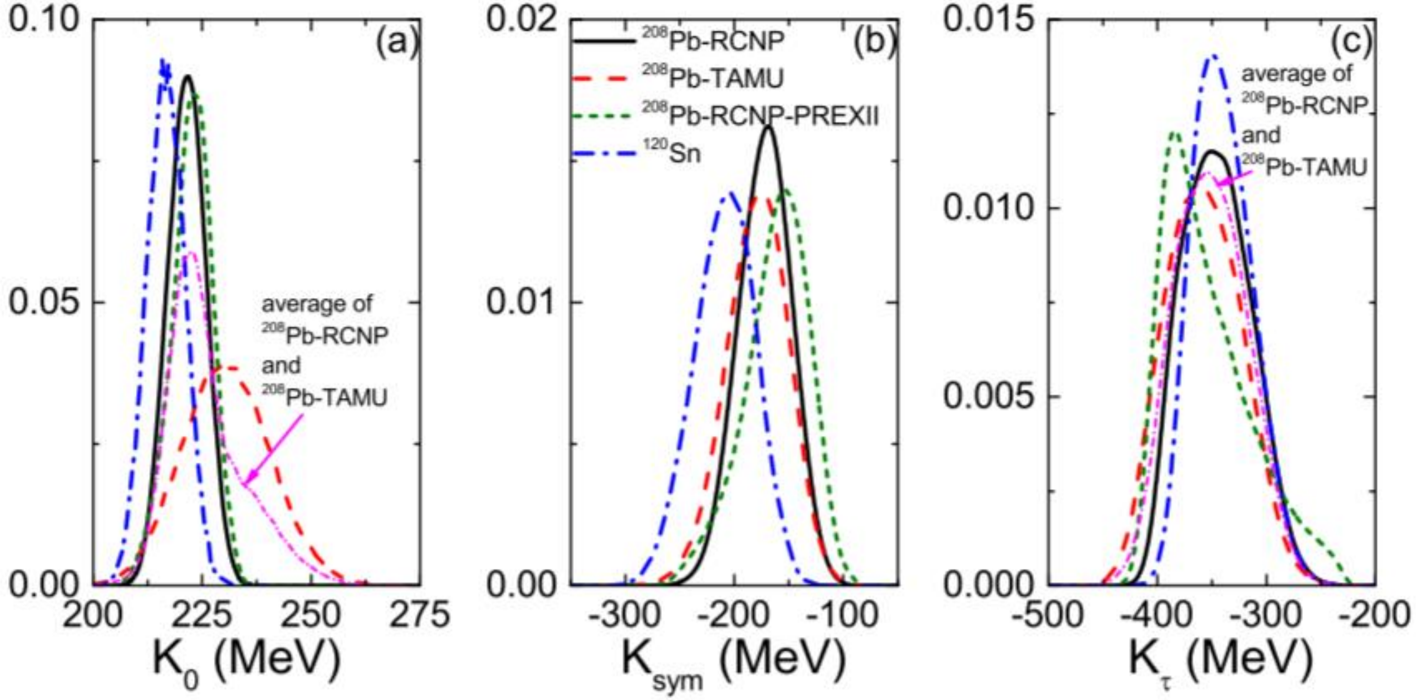
arXiv:2107.10962 [nucl-th]

Bayesian uncertainty quantification for nuclear matter incompressibility

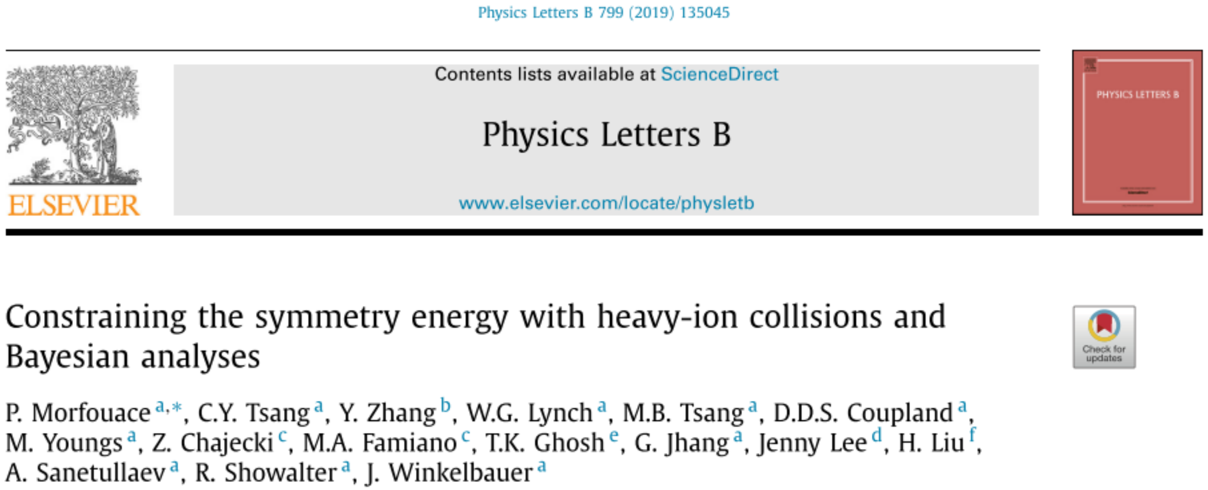
Jun Xu*,[1,2] Zhen Zhang†,[3] and Bao-An Li‡[4]

- Extract K_0 from breathing mode energy.
- **Soft Tin puzzle**: the ISGMR data always favor a smaller K0 value for Sn isotopes than heavy nuclei.
- Maximum a posteriori (MAP) value from Sn120 is about 5 MeV less than that from Pb208.
- Significant overlaps in their posterior distribution.

Posterior distribution



Esym from heavy ion collisions


Physics Letters B 799 (2019) 135045

Contents lists available at ScienceDirect

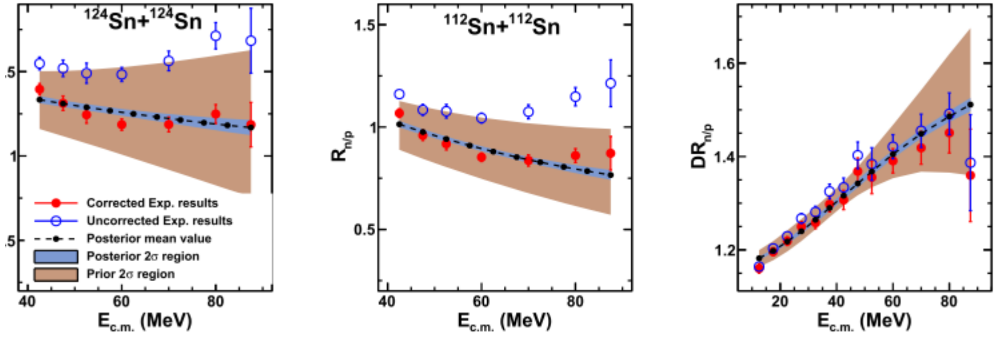
Physics Letters B

www.elsevier.com/locate/physletb

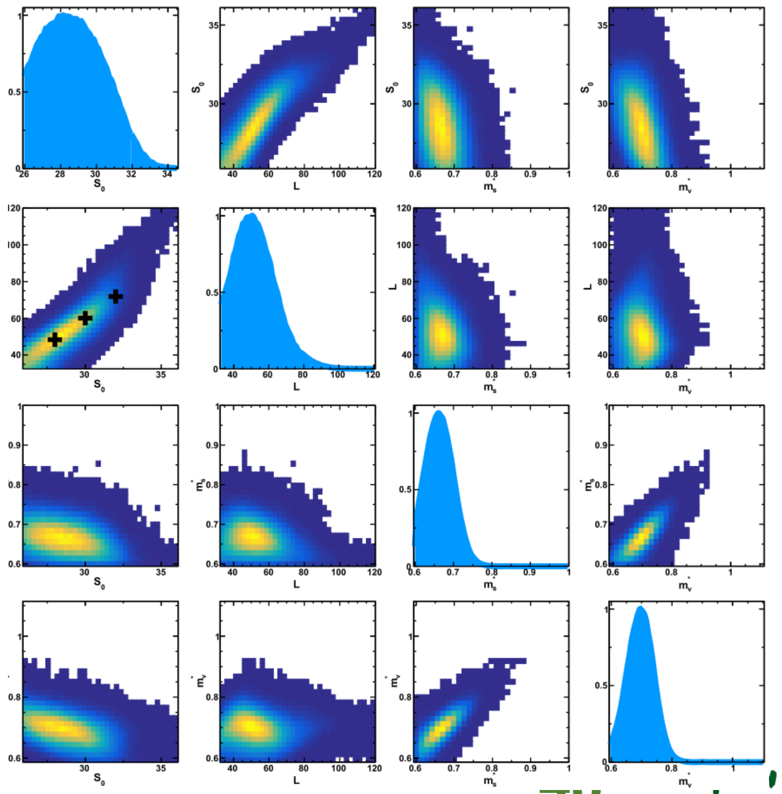
ELSEVIER

Constraining the symmetry energy with heavy-ion collisions and Bayesian analyses

P. Morfouace[a,*], C.Y. Tsang[a], Y. Zhang[b], W.G. Lynch[a], M.B. Tsang[a], D.D.S. Coupland[a], M. Youngs[a], Z. Chajecki[c], M.A. Famiano[c], T.K. Ghosh[e], G. Jhang[a], Jenny Lee[d], H. Liu[f], A. Sanetullaev[a], R. Showalter[a], J. Winkelbauer[a]



- ImQMD transport model with Skyrme interaction.
- 4 parameters
- Observable: n/p single ratios and double spectra in central 112Sn + 112Sn and 124Sn + 124Sn collisions at 120 MeV/u.
- Uniform prior; exponential square likelihood.



EOS from neutron star properties

The Astrophysical Journal, 883:174 (21pp), 2019 October 1
https://doi.org/10.3847/1538-4357/ab3f37

CrossMark

Bayesian Inference of High-density Nuclear Symmetry Energy from Radii of Canonical Neutron Stars

Wen-Jie Xie[1,2] and Bao-An Li[1]

Model

$$\epsilon(\rho, \delta) = \rho[E(\rho, \delta) + M_N] + \epsilon_l(\rho, \delta),$$

$$E_0(\rho) = E_0(\rho_0) + \frac{K_0}{2}\left(\frac{\rho - \rho_0}{3\rho_0}\right)^2 + \frac{J_0}{6}\left(\frac{\rho - \rho_0}{3\rho_0}\right)^3,$$

$$E_{\rm sym}(\rho) = E_{\rm sym}(\rho_0) + L\left(\frac{\rho - \rho_0}{3\rho_0}\right) + \frac{K_{\rm sym}}{2}\left(\frac{\rho - \rho_0}{3\rho_0}\right)^2 + \frac{J_{\rm sym}}{6}\left(\frac{\rho - \rho_0}{3\rho_0}\right)^3.$$

Data

Table 1
The Radius $R_{1.4}$ Data Used in This Work

Radius $R_{1.4}$ (km) (90% CFL)	Source	Reference
$11.9^{+1.4}_{-1.4}$	GW170817	Abbott et al. (2018)
$10.8^{+2.1}_{-1.6}$	GW170817	De et al. (2018)
$11.7^{+1.1}_{-1.1}$	QLMXBs	Lattimer & Steiner (2014)
11.9 ± 0.8, 10.8 ± 0.8, 11.7 ± 0.8	Imaginary case 1	This work
11.9 ± 0.8	Imaginary case 2	This work

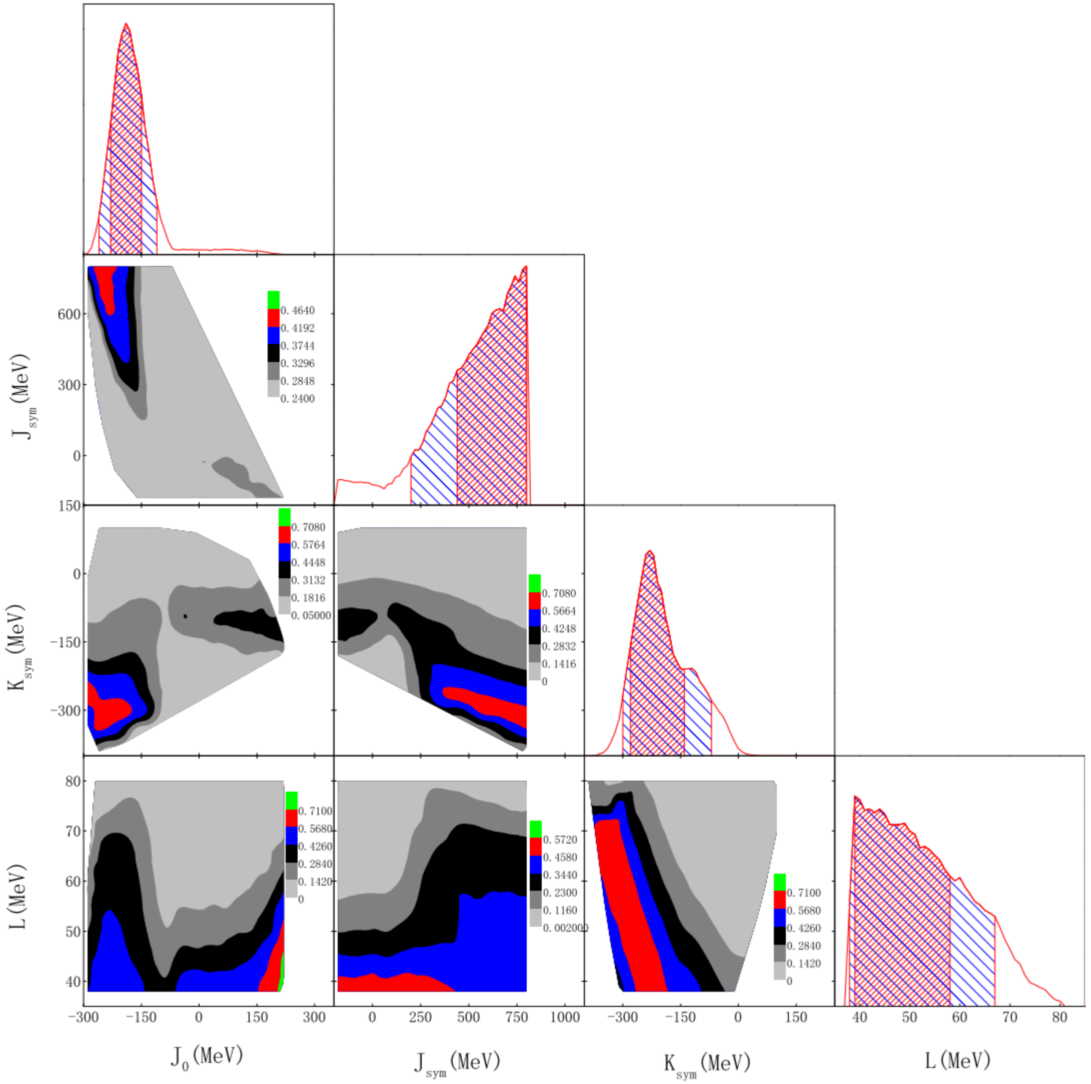
Prior

Prior Ranges of the Six EOS Parameters Used

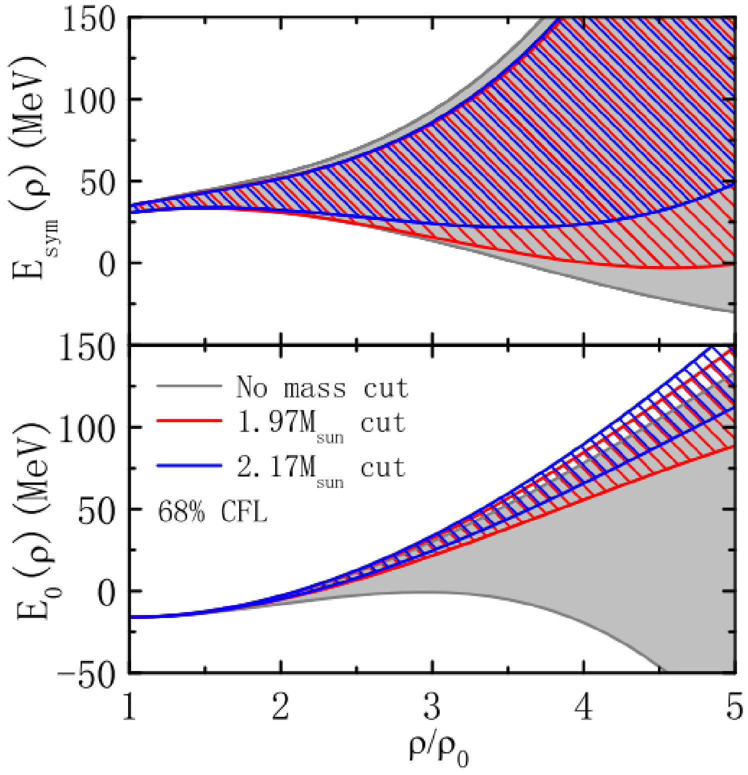
Parameters	Lower Limit	Upper Limit (MeV)
K_0	220	260
J_0	−800	400
$K_{\rm sym}$	−400	100
$J_{\rm sym}$	−200	800
L	30	90
$E_{\rm sym}(\rho_0)$	28.5	34.9

EOS from neutron star properties

Posterior distribution



Predictions



Bayesian inference on neutron star properties

PHYSICAL REVIEW LETTERS **121,** 062701 (2018)

Neutron Star Tidal Deformabilities Constrained by Nuclear Theory and Experiment

Yeunhwan Lim[1,*] and Jeremy W. Holt[1,2,†]

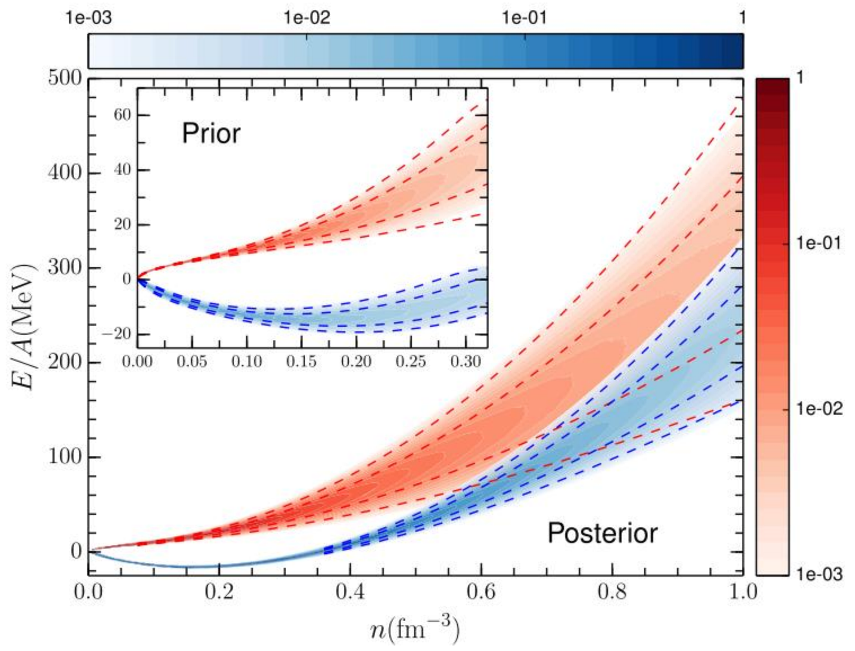
- **Model:**

$$\mathcal{E}(n,x) = \frac{1}{2m}\tau_n + \frac{1}{2m}\tau_p + (1-2x)^2 f_n(n) + [1-(1-2x)^2] f_s(n)$$

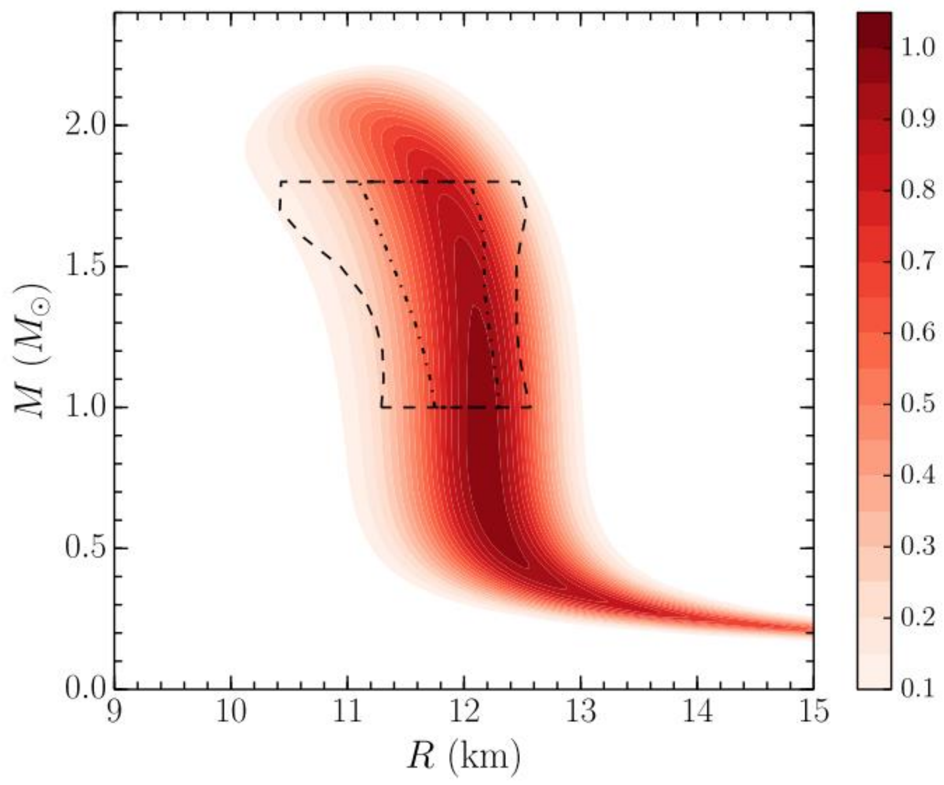
$$f_s(n) = \sum_{i=0}^{3} a_i n^{(2+i/3)}$$

$$f_n(n) = \sum_{i=0}^{3} b_i n^{(2+i/3)}$$

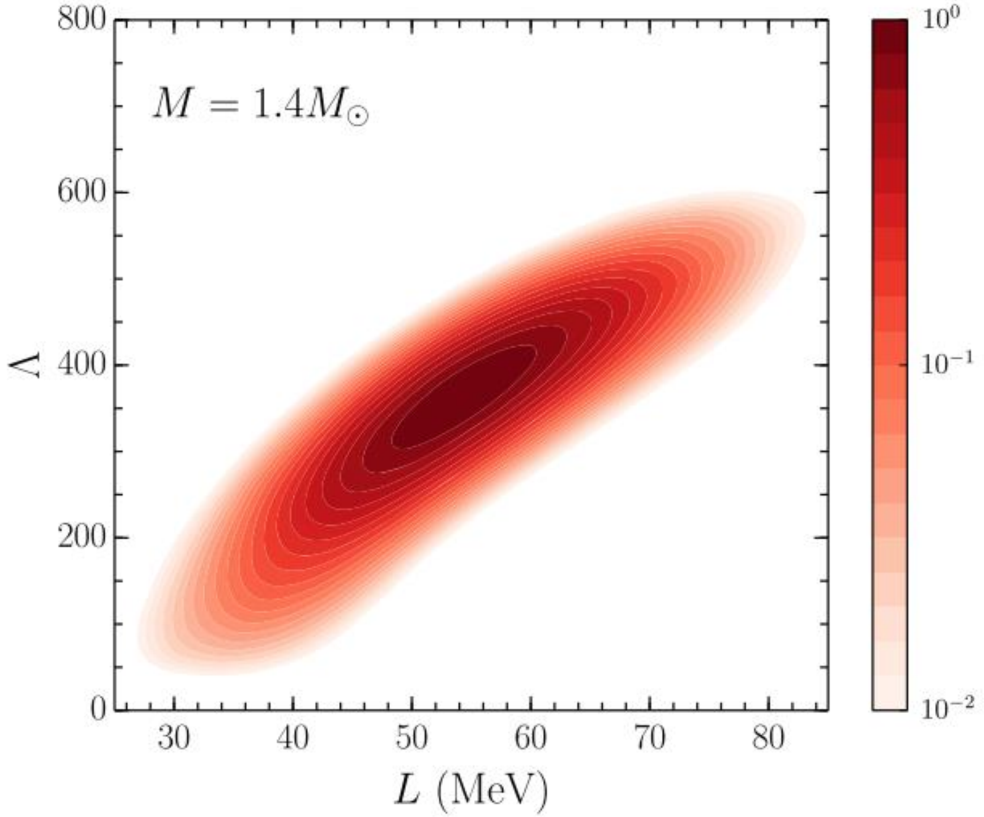
- Gaussian likelihood
- Fit empirical knowledge and ChEFT predictions



Bayesian inference on neutron star properties



- Posterior M-R distribution



- Posterior tidal deformability -L distribution

Summary

- Probabilistic view of model parameters and predictions
- Bayesian method is powerful in uncertainty quantification, correlation analysis
- Model everywhere, Bayesian everywhere.
- How about model dependence?

Bayesian model selection

- Bayesian evidence/marginal conditional

$$p(\mathcal{D} \mid \mathcal{M}_i) \equiv \int_{\Omega_M} p(\mathcal{D} \mid \theta, \mathcal{M}_i) p(\theta \mid \mathcal{M}_i) \mathrm{d}\theta$$

- Posterior odds

$$\frac{p(\mathcal{M}_i \mid \mathcal{D})}{p(\mathcal{M}_j \mid \mathcal{D})} = \frac{p(\mathcal{D} \mid \mathcal{M}_i)}{p(\mathcal{D} \mid \mathcal{M}_j)} \frac{p(\mathcal{M}_i)}{p(\mathcal{M}_j)}$$

A larger ratio means the model M_i is more likely to be true.

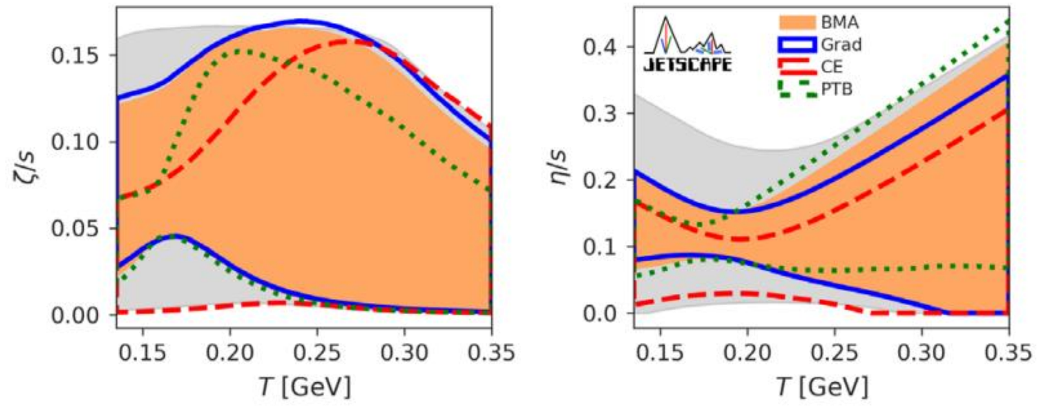
Bayesian model averaging

- Average model predications according to how likely each model is.
- The BMA posterior density

$$p\left(\mathcal{O} \mid \mathcal{D}\right) = \sum_{k=1}^{K} p\left(\mathcal{O} \mid \mathcal{D}, \mathcal{M}_k\right) P\left(\mathcal{M}_k \mid \mathcal{D}\right)$$

- Posterior model weights

$$P\left(\mathcal{M}_k \mid \mathcal{D}\right) = \frac{p\left(\mathcal{D} \mid \mathcal{M}_k\right) p\left(\mathcal{M}_k\right)}{\sum_{\ell=1}^{K} p\left(\mathcal{D} \mid \mathcal{M}_\ell\right) p\left(\mathcal{M}_\ell\right)}$$



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]

PHYSICAL REVIEW LETTERS **126,** 242301 (2021)

Editors' Suggestion

Phenomenological Constraints on the Transport Properties of QCD Matter with Data-Driven Model Averaging

D. Everett,[1] W. Ke,[2,3] J.-F. Paquet,[4] G. Vujanovic,[5] S. A. Bass,[4] L. Du,[1] C. Gale,[6] M. Heffernan,[6] U. Heinz,[1] D. Liyanage,[1] M. Luzum,[7] A. Majumder,[5] M. McNelis,[1] C. Shen,[5,8] Y. Xu,[4] A. Angerami,[9] S. Cao,[5] Y. Chen,[10,11] J. Coleman,[12] L. Cunqueiro,[13,14] T. Dai,[4] R. Ehlers,[13,14] H. Elfner,[15,16,17] W. Fan,[4] R. J. Fries,[18,19] F. Garza,[18,19] Y. He,[20] B. V. Jacak,[2,3] P. M. Jacobs,[2,3] S. Jeon,[6] B. Kim,[18,19] M. Kordell II,[18,19] A. Kumar,[5] S. Mak,[12] J. Mulligan,[2,3] C. Nattrass,[13] D. Oliinychenko,[3] C. Park,[6] J. H. Putschke,[5] G. Roland,[10,11] B. Schenke,[21] L. Schwiebert,[22] A. Silva,[13] C. Sirimanna,[5] R. A. Soltz,[5,9] Y. Tachibana,[5] X.-N. Wang,[20,2,3] and R. L. Wolpert[12]