Moving beyond Chi-squared in nuclei and neutron stars

2015 J. Phys. G: Nucl. Part. Phys. 42 034004

(http://iopscience.iop.org/0954-3899/42/3/034004)

IOP Publishing Journal of Physics G: Nuclear and Particle Physics
J. Phys. G: Nucl. Part. Phys. **42** (2015) 034004 (13pp) doi:10.1088/0954-3899/42/3/034004

Moving beyond Chi-squared in nuclei and neutron stars

A W Steiner

Institute for Nuclear Theory, University of Washington, Seattle, WA 98195, USA

E-mail: steiner3@uw.edu

Received 1 July 2014, revised 10 September 2014
Accepted for publication 1 October 2014
Published 5 February 2015

Abstract

There are several assumptions made in a standard χ^2 analysis of data, including the frequent assumption that the likelihood function is well approximated by a multivariate Gaussian distribution. This article briefly reviews the standard approach and describes how Bayesian inference can be used to go beyond the assumption that the likelihood is Gaussian. Two separate types of analysis relevant to nuclear physics are used as test cases. The first is the determination of the equation of state of dense matter from neutron star mass and radius data. The second is the use of theoretical nuclear mass models to fit currently available data and predict the value of masses which have not yet been measured. For the problem of predicting nuclear masses, it is demonstrated that approximating the likelihood function with a Gaussian can produce biased predictions of unmeasured masses. Finally, the lessons learned from these fitting problems are used to propose a method for improving constraints on the nuclear symmetry energy.

Keywords: energy density functionals, neutron skin thickness, covariance analysis, equation of state of dense matter, neutron star radii

(Some figures may appear in colour only in the online journal)

1. Introduction

Interpreting data and understanding the accompanying theoretical models often requires a statistical analysis. Frequently, this analysis takes the form of an minimization of a function, χ^2, which quantifies the extent to which a model can reproduce a set of experimental data. Uncertainties and correlations are determined by exploring the properties of the likelihood function near the best-fit parameter set. This process contains several assumptions, one of which is that the likelihood function is nearly Gaussian.

Neutron stars, to a good approximation, all lie on one, universal, mass–radius (M–R) curve. It is of broad interest to determine this universal relation, in part because of the connection to the cold and dense region of the QCD phase diagram [1]. Statistical analyses were not applied to neutron stars until recently [2, 3], in part because of the paucity of mass and radius data. The data is sufficiently sparse that the dense matter equation of state is currently underconstrained. This has motivated the use of Bayesian inference which provides a clearer path forward for underconstrained systems. Systematic uncertainties also play a dominant role, and Bayesian inference is useful for characterizing these uncertainties [4].

The statistical analysis of nuclear mass data, on the other hand, was performed already in the 1920s, just before the discovery of the neutron and the creation of the Bethe–Weisacker mass formula [5]. Historically, works have typically chosen to minimize the root-mean-square deviation either of the mass excess, the binding energy, or the energy per baryon (see the discussion in [6]) with respect to parameters in a Hamiltonian (e.g. the Skyrme model [7] or a covariant mean field model [8]), or a phenomenological nuclear mass formula [9–12]. More recent works have considered other data in the fits, such as charge radii (e.g. [8]) and dipole polarizabilities [13].

In section 2 the basic problem of fitting nuclear data is introduced, along with a discussion of how Bayesian inference can play a role. Section 3 describes the analysis of neutron star data, and shows how the Gaussian approximation fails to accurately describe correlations and biases predictions of the radius of a 1.4 $M_\odot$ neutron star. Section 4 describes predictions of unmeasured nuclear masses and the nuclear symmetry energy, and section 5 highlights the importance of selecting the proper data set.

2. Chi-squared, nuclear masses, and charge radii

The prototypical problem to be discussed is the fitting of a model, i.e. an energy density functional, to nuclear masses and charge radii. A review the basic formalism highlights the various assumptions which are often made. Using a set of N_d data points, D_i, and a model with N_p parameters, p_j, one defines a function, χ^2, by

$$\chi^2 = \sum_i^{N_d} \left[\frac{D_i - P_i(\mathbf{p})}{\sigma_{\mathrm{exp},i}} \right]^2 , \tag{1}$$

where P_i is the model prediction for data point with index i and $\sigma_{\mathrm{exp},i}$ is the experimental uncertainty of the ith data point. Minimizing χ^2 over the N_p dimensional parameter space gives the best fit to the data. Alternatively, one maximizes a likelihood function

$$\mathcal{L} = \exp\left(-\chi^2/2\right), \tag{2}$$

which is equal to a product of Gaussians for each data point. The formulation into a likelihood function makes it clear that an implicit assumption of an independent and Gaussian distribution for each data point has already been made[1]. If this assumption is correct and the deviations of the data from the model predictions follow a Gaussian distribution, then the

[1] The assumption of independence is violated in nuclear mass data; the construction of the atomic mass evaluation is done in such a way that neighboring nuclear masses are non-trivially correlated [14, 15]. Future atomic mass evaluations will attempt to assess these correlations [16]. In any case, these correlations may not be important here because the experimental uncertainties are much smaller than the systematic uncertainties. For examples of fitting correlated data in hadron mass spectroscopy, see [17].

value of χ^2 follows the chi-squared distribution and a 'good fit' is one where $\chi^2 \approx N_d - N_p$. Note that, in an underconstrained system, $N_d - N_p < 0$ and this test is not useful.

If the model is a linear function of the parameters, then the model serves as a linear transformation of the multivariate Gaussian determined by the data and therefore the likelihood function is also Gaussian. For nonlinear models, the likelihood function is approximately Gaussian in the limit $N_d \to \infty$ because of the central limit theorem. If the likelihood function is approximately Gaussian, then it can be parameterized by a best fit vector $\boldsymbol{\mu}$ and a covariance matrix, $\boldsymbol{\Sigma}$

$$\mathcal{L}(\mathbf{p}) \approx A \exp\left[-\frac{1}{2}(\mathbf{p}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{p}-\boldsymbol{\mu})\right], \tag{3}$$

with a normalization factor A fixed so that equation (2) holds[2]. The covariance matrix is often determined by computing the second derivative of $\mathcal{L}$ around the maximum. In terms of the original χ^2 function

$$\Sigma_{ij} = \frac{1}{2}\frac{\partial}{\partial p_i}\frac{\partial}{\partial p_j}\left(\chi^2\right)_{\mathbf{p}=\boldsymbol{\mu}}. \tag{4}$$

The correlation matrix $\boldsymbol{C}$ can be obtained from $\boldsymbol{\Sigma}$ by $C_{ij} = \Sigma_{ij}\big/\sqrt{\Sigma_{ii}\Sigma_{jj}}$. Modern fitting codes often obtain the covariance matrix from the Jacobian of χ^2. The correlation matrix, obtained in this way, is a correct representation of the correlations between model parameters *only if the likelihood function is approximately given by the form in equation (3)*. It has become commonplace to report correlation matrices without verifying that this approximation is actually valid. The implications of this approximation for predicting nuclear masses is discussed in section 4 below.

Alternatively, the covariance matrix can also be obtained directly from a Markov chain Monte Carlo simulation of the likelihood function. Each entry is just the correlation coefficient

$$\Sigma_{ij} = \sum_{i=1}^{N_c}\frac{1}{N_c}\left(p_i - \bar{p}_i\right)\left(p_j - \bar{p}_j\right), \tag{5}$$

where $\bar{p}_i$ is the average value of the ith parameter and N_c is the number of entries in the Markov chain. This latter method is useful even when the likelihood function is not approximately Gaussian.

In the context of nuclear masses, there is evidence that the likelihood functions implied by typical models are often not Gaussian. Explicit calculation shows that the likelihood function from the simple Bethe–Weisacker mass formula is multi-modal because of the pairing contribution. Skyrme models [7], energy density functionals which describe nuclear masses and charge radii, are linear in parameters $t_0, \ldots, t_3$, but not linear in the parameter, α, which characterizes density-dependent two-nucleon force. This nonlinearity could mean that likelihood is not fully Gaussian, though this has not been fully studied. Some evidence for non-Gaussian behavior is present in the posterior probability distributions for the Skyrme-like parameters in figure 1 of [18] which are fit to a large set of data comprising mostly of nuclear masses and charge radii. Covariant mean-field models may also not generate Gaussian likelihoods because of the nonlinearity (which is required to reproduce saturation) in the mean-field equation for the scalar–isoscalar meson (e.g. [19]).

[2] Note that the likelihood function is not renormalized so that its integral is unity. The likelihood function, unlike the prior and posterior distributions below, is not a probability density function. The normalization is important when comparing how two models fit to the data. See also the discussion of Bayes factors below.

If the likelihood function was a multivariate Gaussian, one would expect the minimization of χ^2 to be numerically trivial (so long as the likelihood was evaluated with sufficient numerical accuracy). For example, the Broyden–Fletcher–Goldfarb–Shanno method is guaranteed to converge if the objective function has a quadratic expansion near the extremum. The result that the fit of the Skyrme-like parameter $C_1^{\rho\nabla J}$ depends slightly on the initial guess in [20] (in a fit similar to that in [18] described earlier) means that, in contrast, this minimization is not trivial. This is not surprising as the posterior distribution of this same parameter in figure 1 in [18] appears multi-modal.

All currently available energy density functionals which describe heavier nuclei have systematic uncertainties which are much larger than the $\sigma_{\mathrm{exp},i}$. Thus, it is common to redefine χ^2 as

$$\chi^2 = \sum_i^{N_d} \frac{\left(D_i - P_i\right)^2}{\sigma_{\mathrm{exp},\,i}^2 + \sigma_{\mathrm{sys},\,i}^2}\,, \tag{6}$$

where $\sigma_{\mathrm{sys},i}$ is the systematic uncertainty for each data point. Often $\sigma_{\mathrm{exp},i}$ is small enough that it can be ignored. There are now $N_p + N_d$ parameters: the N_p model parameters and a parameter $\sigma_{\mathrm{sys},i}$ for each data point. Equivalently, one can replace each $\sigma_{\mathrm{sys},\,i}^2$ with a weight, w_i, to be fixed by some recipe. There are several ways to proceed. It is possible to perform a Bayesian analysis with all $N_p + N_d$ parameters, but this is computationally difficult[3]. The alternative is to reduce the number of parameters. A frequent choice is to employ a fixed uncertainty for each type of data point, e.g. σ_{mass} for nuclear masses and σ_{radius} for charge radii. A similar choice is manifest in the literature through the presence of fitting protocols: to choose not to fit data point j is equivalent to assuming that the systematic uncertainty for the model to predict the jth data point, σ_j, is so large that the corresponding term in χ^2 is negligable. Early fits (see review in e.g. [5]) used only doubly-magic nuclei, on the basis that open-shell systems contained correlations which were unlikely to be correctly computed by Hartree–Fock calculations. [21] discusses this issue and suggests fixing the denominators by separately normalizing χ^2 for each type of data (their equation (4)). A slightly different procedure is suggested below. Finally, computing several systematic uncertainties using iteratively-reweighted least squares may be useful. Studies of how relative variations in the χ^2 weights might affect the interpretation of the data, in the context of low-energy nuclear data, are in their infancy.

An implicit assumption in this discussion is the assumption that the systematic uncertainties in the denominators are themselves uncorrelated. If this assumption fails, then one must reformulate the χ^2 function to take into account the uncertainties. This problem has been found particularly relevant for fitting parton distribution functions (see in particular appendix B of [22]). If correlations among parametersand correlations among the systematic uncertainties are not important, and if the χ^2 function is of the form given above, one can estimate parameter uncertainties by varying the parameters, one at a time, until $\chi^2(\mathbf{p}) = N_d - N_p + 1$. However, it seems likely that the systematic uncertainties of masses and radii are indeed correlated. Models with no three-nucleon forces often predict saturation at a higher density and with a larger binding energy than found in laboratory nuclei [23], and these models naturally lead both to smaller masses and smaller radii.

[3] I have found empirically that a Bayesian analysis in this form requires prior distributions for $\sigma_{\mathrm{sys},i}$ that fall off as $\exp(-\sigma^2)$ to to ensure the trivial solution with $\sigma \to \infty$ does not dominate the results. More work in this direction is in progress.

In the context of fits to nuclear masses and charge radii, the model parameters are almost always correlated. If the likelihood function is nearly Gaussian, then one can determine the parameter uncertainty including these correlations by marginalizing over the multivariate Gaussian (which can be done using the covariance matrix and does not require any integration).

3. Bayesian analysis for meutron star masses and radii

For neutron stars, the problem is how to determine the M–R curve from a small sample of poorly constraining observations. (A review is available in [24], only some of which is particularly relevant to this work and reproduced here.) The statistical analysis is complicated for two reasons. The first is that the currently available data has large statistical (and large systematic) uncertainties. There are effectively more parameters than constraints. Second, the M–R curve need not be a function in the mathematical sense, it may fail either or both of the horizontal or vertical line tests. (Given the central energy density, ε_c, the relations $M(\varepsilon_c)$ and $R(\varepsilon_c)$ are both well-defined functions.) An example of how this happens in is [25], where a phase transition induces a new branch in the M–R curve.

Bayesian analysis allows one to solve both of these issues, at the expense of introducing an unknown function, a prior probability distribution function. Also, the two-dimensional nature of the neutron star problem requires that N_d parameters for the neutron star masses, M_i, are also added. Some researchers have been slow to employ Bayesian inference because of the ambiguity in determining prior distributions, but [26] has shown that the prior is not a separate entity but a part of the model which can be checked and verified by comparing to data.

In the case of nuclear mass fitting described above, the likelihood function is a product of N_d Gaussian probability distributions for each data point. The neutron star case is handled similarly (cf equation 31 in [3]): the likelihood function is a product over a data set for each neutron star. Each data set is of the form $\mathcal{D}(M, R)$, a two-dimensional probability distribution for the mass and radius of a neutron star. These probability distributions have complicated shapes, and the likelihood is not Gaussian. As discussed above, the traditional correlation matrix thus does not contain all of the information on correlations present in the problem.

As a demonstration, the correlation matrix is represented in figure 1, for the analysis of neutron star data with the HLPS + model C parameterization from [27] based on [28] and model C from [29]. There are 17 model parameters, 7 from the EOS and 10 neutron star masses. The seven EOS parameters are given in table 1. The results below do not depend sensitively on the values of the saturation density or the energy density grid which is used, so long as the grid covers the full range of energy densities probed in neutron star cores. The four panels each represent the correlation matrices inferred from different subsets of the Markov chain Monte Carlo simulation of the likelihood function. The upper-left panel was computing using only points where the likelihood was within a factor of two of the maximum value at the best fit (this is the result which corresponds to the Gaussian approximation as discussed in section 2). The upper-right panel used points where the likelihood panel was a factor of 8 within the maximum. The lower-right panel used all the Monte Carlo points and thus does not impose any approximation on the likelihood function. The strong correlation between parameters 2 and 3 is related to the correlation between S and L as obtained in [28]. The strong anti-correlation between pressure parameters is also expected. The pressure cannot be too small over a large range in density or the maximum mass will like below the observed value. The pressure cannot also be too large over a large range in density or the radii will be too

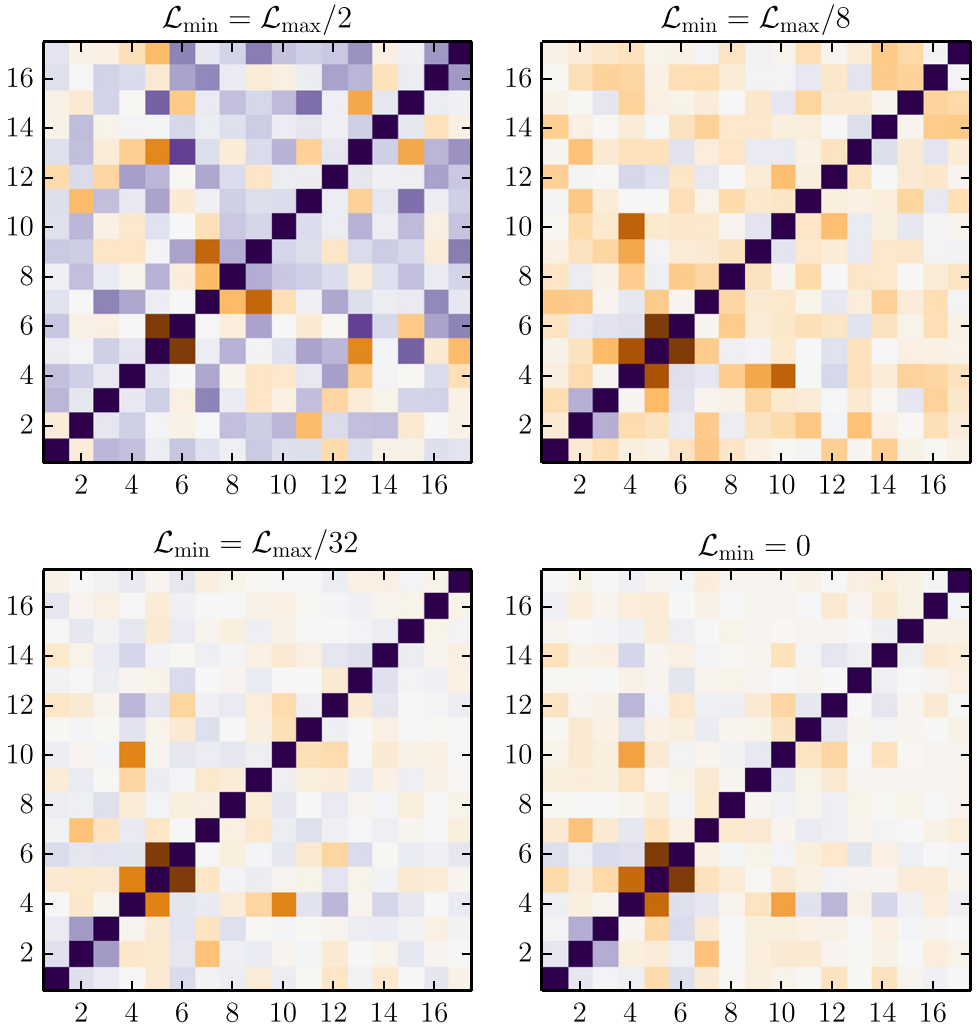

Figure 1. The evolution of the correlation matrix as more Monte Carlo points near the maximum likelihood (best fit) are added. The x and y-axes both enumerate the parameter number as given in table 1. The upper-left shows the correlation matrix in a close neighborhood around the best fit. This is the result as obtained assuming the likelihood is described by a covariance matrix from equation (4). The lower-right panel shows the full covariance with all the Monte Carlo samples (using equation (5)). The upper-right and lower-left panels show the evolution between these two. Strong correlations are in purple, yellow, strong anti-correlations are in orange, and lack of correlation is white. The true correlations in the fit are represented by the lower-right panel even though the upper-left panel is representative of the method most often used to compute these correlations.

large to reproduce the data. Qualitatively, the pressure tends to be either small at moderate densities and large at high densities, or vice-versa. Finally, it is clear that many of the correlations in the Gaussian approximation (upper-left panel) are not representative of the full data set (lower-right panel).

Table 1. The list of parameters for neutron star EOS parameterization described in the text and corresponding to the correlation matrices displayed in figure 1.

Parameter index	Description
1	Compressibility
2	Energy per baryon of neutron matter
3	Derivative of energy per baryon of neutron matter
4	$P_1 \equiv P(\varepsilon = 2\ \mathrm{fm}^{-4}) - P(n_B = 0.16\ \mathrm{fm}^{-3})$
5	$P_2 \equiv P(\varepsilon = 3\ \mathrm{fm}^{-4}) - P(\varepsilon = 2\ \mathrm{fm}^{-4})$
6	$P_3 \equiv P(\varepsilon = 5\ \mathrm{fm}^{-4}) - P(\varepsilon = 3\ \mathrm{fm}^{-4})$
7	$P_4 \equiv P(\varepsilon = 7\ \mathrm{fm}^{-4}) - P(\varepsilon = 5\ \mathrm{fm}^{-4})$
8–17	Individual neutron star masses

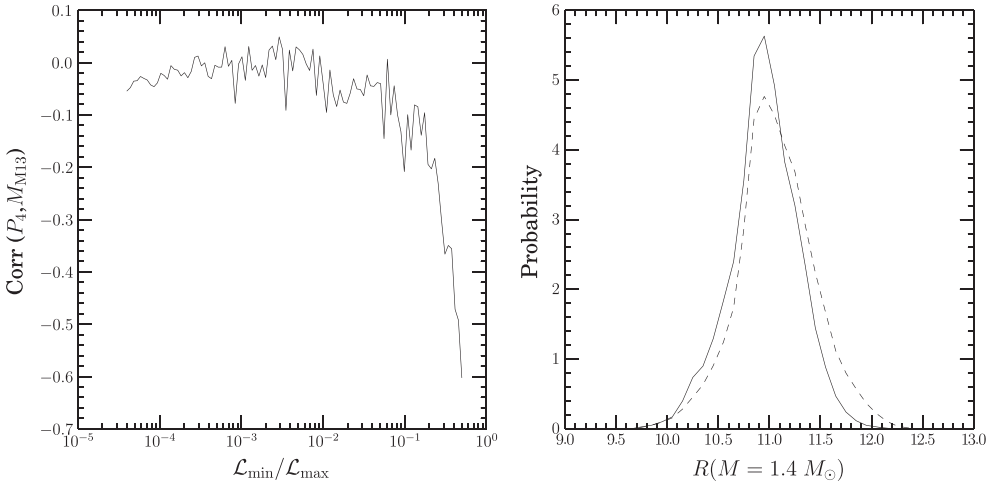

Figure 2. Left panel: the evolution of one correlation coefficient away from the best fit. The pressure at high densities and the mass of one neutron star appear correlated (on the right side of the plot), but in reality are uncorrelated (on the left side of the plot). Right panel: the posterior probability distributions for the radius of a $1.4\,M_\odot$ neutron star, assuming the full likelihood function (solid line) or a Gaussian approximation (dashed line).

This result is further demonstrated in the left panel of figure 2, where the correlation coefficient between the high density pressure parameter, P_4 and the mass of the neutron star in the M13 globular cluster is shown as a function of the range of likelihoods considered, $\mathcal{L}_{\rm min}/\mathcal{L}_{\rm max}$. The two parameters appear anti-correlated near the best fit. In reality, they are not correlated. This lack of correlation is expected, as the neutron star in M13 is typically lower in mass, and thus less sensitive to the pressure at higher densities. The right panel of figure 2 shows that the Gaussian approximation to the likelihood also gives a different range for the radius of a $1.4\,M_\odot$ neutron star implying slightly larger radii.

In Bayesian inference, marginalization is often used to determine model parameters. The posterior probability distribution for the ith parameter is determined from

$$P(p_i) \propto \int \mathrm{d}p_1 \mathrm{d}p_2 \ldots \mathrm{d}_{p_{i-1}} \mathrm{d}_{p_{i+1}} \ldots \mathrm{d}p_{N_p} \mathrm{d}M_1 \ldots \mathrm{d}M_{N_d} \mathcal{L}(\mathbf{p}, \mathbf{M}) \mathcal{P}(\mathbf{p}, \mathbf{M}), \tag{7}$$

where $\mathcal{L}$ is the likelihood and $\mathcal{P}$ is the prior distribution. This integral is typically evaluated using a Monte Carlo method. One may directly parameterize the M–R curve and compare it to the data or parameterize the EOS of dense matter, $P(\varepsilon)$ and then use the TOV equations to compute the M–R curve at each point in the Monte Carlo integration. Each parameterization (given a fixed data set) is a different model. A particular advantage of this formalism, as applied in [3, 24, 27, 29–31], is that one need not assume any correlations between the properties of matter near the saturation density and matter in the inner core of the neutron star. Matter in the core may be best described, for example, by quark degrees of freedom and have little similarity with neutrons and protons at the saturation density.

Models can be compared with each other by computing the Bayes factor. The Bayes factor of model A with respect to model B is the ratio, $B_{AB} \equiv E_A/E_B$, where E is the evidence

$$E = \int \mathrm{d}\mathbf{p}\, \mathrm{d}\mathbf{M}\, \mathcal{L}(\mathbf{p}, \mathbf{M}) \mathcal{P}(\mathbf{p}, \mathbf{M}) \tag{8}$$

Note that the two models need not have the same number of parameters. The Bayes factor is equivalent to betting odds: $B_{AB} = 10$ implies that model A is ten times more likely than model B.

Several different models can be compared with Bayes factors, a strategy employed in [24]. This work used Bayes factors and an analytical model of the neutron star atmosphere to show that different values of the 'hydrogen column density' were favored. The hydrogen column density is a parameter which characterizes the extent to which x-rays are absorbed between the neutron star and the detector. [24] also found a model where some neutron stars have Helium atmospheres was favored. Recent work in [32], including more observational data, has confirmed a lower hydrogen column density for the neutron star in the ω Cen globular cluster and a Helium atmosphere for the neutron star in NGC 6397, as partially predicted in [24].

4. Predictions, the nuclear symmetry energy, and the neutron skin thickness of lead

Given any function $\mathcal{F}(\mathbf{p})$ of the model parameters[4], one can use the best fit to compute the predicted value, $\mathcal{F}(\mathbf{p} = \boldsymbol{\mu})$. To compute the uncertainty in the prediction, one can compute a probability distribution

$$P(F) = \int \mathrm{d}\mathbf{p} \mathcal{L}(\mathbf{p}) \delta[\mathcal{F}(\mathbf{p}) - F] \mathcal{P}(\mathbf{p}). \tag{9}$$

This integral is easier if $\mathcal{L}$ can be accurately described by a multivariate Gaussian, because one can directly sample the likelihood function using the Cholesky decomposition of the covariance matrix. When the Gaussian approximation is inadequate, as in the neutron star problem discussed above, and the likelihood function must be indirectly sampled, e.g. through Markov chain Monte Carlo. In order to make the Monte Carlo more efficient, the Gaussian approximation can be used as a proposal distribution in a Metropolis–Hastings step.

Figure 3 demonstrates how the Gaussian approximation of the likelihood function discussed in section 2 affects predictions for nuclear masses. The nuclear mass model of [12] is fit to the nuclear masses from the UNEDF collaboration [33]. Equation (9) is used to predict the binding energy of ^{137}Sn. Uniform prior distributions are used to allow a straightforward

[4] Because this section is more general, the explicit reference to the neutron star masses, $\mathbf{M}$, is removed.

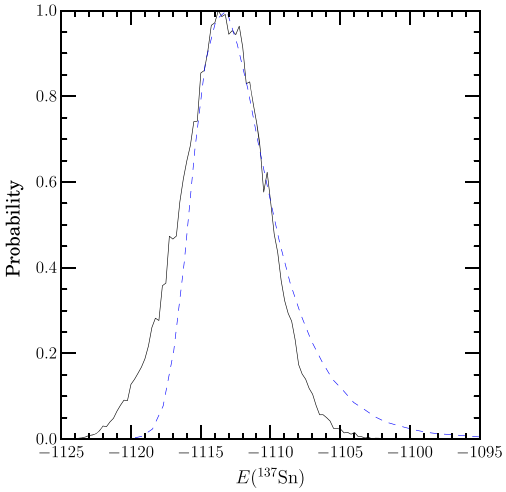

Figure 3. Predicted probability distribution for the mass of ^{137}Sn given the nuclear mass model of [12]. The black solid line gives the full result from equation (9) using a Monte Carlo simulation without any approximation and the blue dashed line gives the result using a Gaussian approximation for $\mathcal{L}$. The latter underestimates the predicted binding energy.

comparison to what would be obtained in the frequentist method. Two results are presented, with and without the Gaussian approximation of the likelihood function. It is clear that the Gaussian approximation systematically predicts less binding.

A quantity of central interest in low-energy nuclear physics is the nuclear symmetry energy, the energy cost to create an isospin asymmetry. Given the energy per baryon of neutron matter as a function of the baryon density, $E_{\rm neut}(n_B)$, and the energy per baryon of nuclear matter $E_{\rm nuc}(n_B)$, the symmetry energy can be defined as the difference $S(n_B) \equiv E_{\rm neut}(n_B) - E_{\rm nuc}(n_B)$. Of particular interest is the value of the symmetry energy at the nuclear saturation density, $S \equiv S(n_B = n_0)$, and its derivative, $L \equiv 3n_0 S'(n_B = n_0)$.

Arguably, two of the best ways to obtain constraints on the nuclear symmetry energy are from nuclear data and neutron star observations described in sections 2 and 3 above. Nuclear masses offer a strong constraint on a linear combination of S and L, but do not determine the two quantities separately [34]. Neutron star radius measurements and measurements of the neutron skin thickness of lead both strongly constrain L and are weakly correlated with S [35].

In part motivated by the connection between the neutron skin thickness in lead, δR, and the parameter L, several recent experiments have measured the skin thickness. The PREX experiment, which uses parity violating electron scattering, found $\delta R = 0.33^{+0.16}_{-0.18}$ fm [36]. This weak probe of nuclear structure should have smaller systematic uncertainties than those from strongly interacting probes. A more precise (but more model-dependent) result of $\delta R = 0.15^{+0.04}_{-0.06}$ fm was recently obtained from photoproduction of pions at the MAMI electron beam facility in [37]. Using probability distributions determined by neutron star radius measurements, several authors have determined the parameter L [3, 29, 30, 38]. [29] predicted the neutron skin thickness would be less than 0.20 fm, consistent with both the recent PREX and MAMI measurements. In this work, fits employed four seperate models (named A, B, C, and D) and different interpretations of the neutron star data in order to attempt to assess the systematic uncertainty (see figure 4 of [29]). It remains to be seen if the prediction of the

neutron skin thickness of lead from neutron star observations in [29] will prove true with more accurate experiments, such as PREX II.

In general, uncontrolled systematic uncertainties are particularly pernicious for predicted values. *There is no method which guarantees that reality lies between any computed confidence interval which is obtained from the data*. In the context of energy density functionals, there is a long history of using several nearly equivalent models of the same data in order to estimate the systematic uncertainties, as done in the neutron star analysis above, as earlier done in [39–41] and as discussed in [21][5]. Studies of inter-model dependence are important, even when one is not focused on predicting a particular observable, because they help diagnose cases where a model may fit the data accurately with an incorrect physical mechanism.

5. Completeness versus accuracy

The purpose of making a model is two-fold: the first is to provide insight regarding the physical mechanisms which underlie the observed data, and second is to make predictions. These two purposes are sometimes at odds: the ability of a model to make predictions must often strike a balance between attempting to describe the most complete set of currently available data, and describing a small set of data with high accuracy. This qualitative picture is behind the practice (in Bayesian inference) of using different prior distributions for parameter estimation than those used for model comparison.

Much of the recent work surrounding the description of nuclear structure observables has focused on generating an energy density functional. The Kohn–Sham theorem suggests that, if the correct energy density functional was found, one could accurately describe all nuclei at the level of single reference Hartree–Fock–Bogoliubov (see e.g. [33]). In many works, the purpose is mainly that of completness: one describes as large of a data set as is possible in order to obtain the best energy density functional. Part of the promise of chiral effective theory is that we might not have to sacrifice completeness for accuracy (see review in [42]), and progress is being made in this direction in lighter nuclei. On the other hand, very accurate descriptions of neutron-rich calcium isotopes have been obtained with coupled cluster methods [43, 44] with interactions which might otherwise have difficulty describing nuclear matter at the saturation density [45].

In the context of nuclear masses, this tension between completeness and accuracy is demonstrated in figure 4, where two mass models from [12, 46] are used to predict the mass of ^{137}Sn. In the left panel, the full experimental mass data set from [47, 48]. The mass models predict rather small statistical uncertainties, but the systematic uncertainties are at least 3 MeV. In the right panel, the models are fit to a more limited data set in the Sn region, and while the statistical uncertainties are nearly unchanged the systematic uncertainties appear smaller. Of course, two models are insufficient to make any conclusive statements about the magnitude of systematic uncertainties, but this plot illustrates the basic point.

Constraints on the nuclear symmetry energy will be improved by understanding this tension between completeness and accuracy. It might be possible to obtain models which more accurately describe the symmetry energy by fitting to only carefully selected nuclear data and avoiding complications of attempting to describe nuclei which are currently not well-described by Hartree–Fock–Bogoliubov. In more detail, the best constraints on the symmetry

[5] These studies can be viewed as an inexpensive and rough way of performing a hierarchical Bayesian analysis. In particular, they are only representative of the systematic uncertainty so long as they faithfully represent the space of reasonable models.

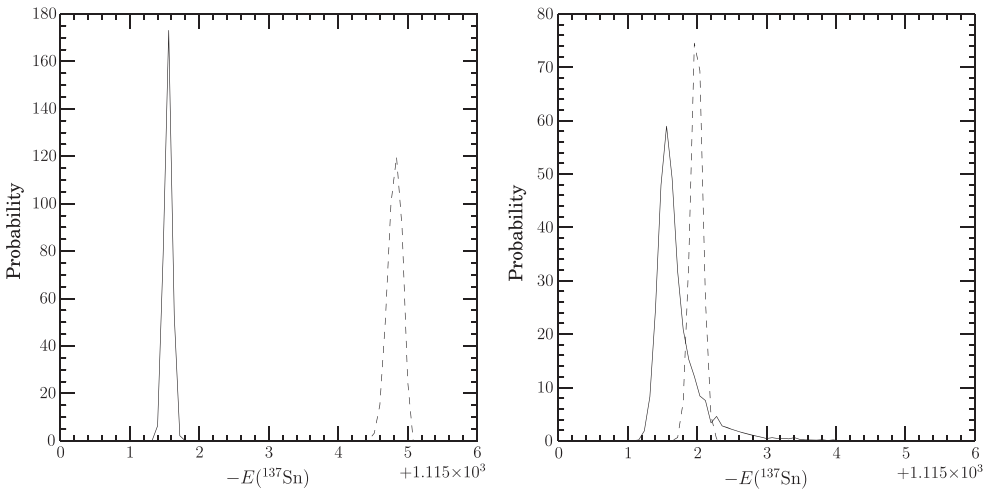

Figure 4. A demonstration of the competition between completeness and accuracy. Solid lines use the model from [12] and dashed lines use the mass model from [46]. Left panel: the predicted binding energy of ^{137}Sn when fitting all measured nuclear masses. Right panel: the binding energy of ^{137}Sn when fitting only Sn, In, and Sb isotopes.

energy will come from an fitting several models to a data set which is specifically optimized to generate the smallest combination of (i) the uncertainty in the fit and (ii) the systematic uncertainty implied by the variation between models. This optimization of this data set should include varying the individual weights, w_i, in the χ^2 function to ensure the highest possible accuracy for the symmetry energy. In addition to nuclear masses and charge radii described above, giant resonance data will likely be helpful in determining the symmetry energy. When the systematic uncertainties of neutron star mass and radius observations are sufficiently small, they will also provide a powerful constraint.

6. Discussion

The analysis of neutron star data has a low computational requirement; a single solution of the TOV equations is much faster than the accurate determination of several nuclear masses and charge radii from an energy density functional. This low computational requirement has enabled the use of Bayesian analysis to attempt to understand currently available neutron star data. A similar statistical analysis for the matching of energy density functionals to low-energy nuclear data is not yet possible. Nevertheless, the analogy permits recommendations on future directions for fits to energy density functionals.

- Likelihood functions should be published and made widely accessible[6]. If the likelihood function is sufficiently accurate for predictions, then publishing the best fit and the covariance matrix is sufficient. Otherwise, the likelihood function can either be tabulated or represented by a Markov chain of several points. In the neutron star case, this latter procedure was used in [27]; a list of points sampled from the likelihood function was provided for each model. This list can be directly used by other authors to provide

[6] Thanks to David W Hogg for arguing that likelihoods ought to be published in his research blog.

predictions based on the same astronomical data. Also, the differences between the likelihood functions between models can be used to obtain an estimate on the systematic uncertainty as described above.

- The classical χ^2 procedure can be applied in cases where it is inappropriate, clouding our ability to properly interpret the data. Important methods to alleviate this issue include examining the residuals from the best fit model [21], and systematically refitting with part of the data set removed (as e.g. done in the neutron star context in [49]). In addition, one critical issue is the potential for correlations between our systematic uncertainties. More work needs to be done to understand how these correlations might potentially affect results. These correlations can be explicitly modeled using the method described in [22].
- A method for obtaining modern constraints on the nuclear symmetry energy is proposed in the previous section. Generalizing this method, the ability of a model to make predictions should drive the data set which is included in χ^2 fits and the associated weights that are used. This can only be effectively done when the systematic uncertainties of the model are assessed, either by a reasonable physical argument or a study of the variation among several models. The exception to this is when a more complete description of the data is likely to provide physical insight. The ability to make more accurate predictions will be useful, in particular, to those communities (like neutron star astronomers) who need nuclear data for their models.

Acknowledgments

The author would like to thank S Liuti for pointing out [22], G Bertsch and S Reddy for several helpful discussions, and the DOE for support under grant DE-FG02–00ER41132. Experimental values for the nuclear masses, the theoretical mass formulas, and a solver for the stellar structure equations were obtained from O2scl [50].

References

[1] Lattimer J M and Prakash M 2001 *Astrophys. J.* **550** 426
[2] Özel F and Psaltis D 2009 *Phys. Rev.* D **80** 103003
[3] Steiner A W, Lattimer J M and Brown E F 2010 *Astrophys. J.* **722** 33
[4] Kass R E and Raftery A E 1995 *J. Am. Stat. Assoc.* **90** 773
[5] Lunney D, Pearson J M and Thibault C 2003 *Rev. Mod. Phys.* **75** 1021
[6] Toivanen J, Dobaczewski J, Kortelainen M and Mizuyama K 2008 *Phys. Rev.* C **78** 034306
[7] Skyrme T H R 1959 *Nucl. Phys.* **9** 615
[8] Lalazissis G A, König J and Ring P 1997 *Phys. Rev.* C **55** 540–3
[9] v Weizsäcker C F 1935 *Z. Phys.* **96** 431
[10] Bethe H A and Bacher R F 1936 *Rev. Mod. Phys.* **8** 82–229
[11] Duflo J and Zuker A P 1995 *Phys. Rev.* C **52** 23
[12] Dieperink A E L and van Isacker P 2009 *Eur. Phys. J.* A **42** 269
[13] Piekarewicz J, Agrawal B K, Colò G, Nazarewicz W, Paar N, Reinhard P G, Roca-Maza X and Vretenar D 2012 *Phys. Rev.* C **85** 041302
[14] Audi G, Davies W G and Lee-Whiting G E 1986 *Nucl. Instrum. Methods Phys. Res.* A **249** 443
[15] Audi G, Wapstra A H and Thibault C 2003 *Nucl. Phys.* A **729** 337
[16] Kondev F 2014 private communication
[17] Michael C 1994 *Phys. Rev.* D **49** 2616–9
[18] Schunck N, McDonnell J D, Higdon D, Sarich J and Wild S 2014 arXiv:1406.4374
[19] Roca-Maza X, Paar N and Colò G 2015 *J. Phys. G: Nucl. Part. Phys.* **42** 034033
[20] Wild S M, Sarich J and Schunck N 2015 *J. Phys. G: Nucl. Part. Phys.* **42** 034031
[21] Dobaczewski J, Nazarewicz W and Reinhard P G 2014 *J. Phys. G: Nucl. Part. Phys.* **41** 074001

[22] Stump D, Pumplin J, Brock R, Casey D, Huston J, Kalk J, Lai H L and Tung W K 2001 *Phys. Rev.* D **65** 014012
[23] Coester F, Cohen S, Day B and Vincent C M 1970 *Phys. Rev.* C **1** 769–76
[24] Lattimer J M and Steiner A W 2014 *Eur. Phys. J.* A **50** 40
[25] Alford M G, Han S and Prakash M 2013 *Phys. Rev.* D **88** 083013
[26] Gelman A and Shalizi C R 2012 *Br. J. Math. Stat. Psychol.* **66** 8
[27] Steiner A W, Gandolfi S, Fattoyev F J and Newton W G 2014 arXiv:1403.7546
[28] Hebeler K, Lattimer J M, Pethick C J and Schwenk A 2013 *Astrophys. J.* **773** 11
[29] Steiner A W, Lattimer J M and Brown E F 2013 *Astrophys. J. Lett.* **765** 5
[30] Steiner A W and Gandolfi S 2012 *Phys. Rev. Lett.* **108** 081102
[31] Lattimer J M and Steiner A W 2014 *Astrophys. J.* **784** 123
[32] Heinke C O *et al* 2014 *Mon. Not. R. Astron. Soc.* **444** 443
[33] Kortelainen M, McDonnell J, Nazarewicz W, Reinhard P G, Sarich J, Schunck N, Stoitsov M V and Wild S M 2012 *Phys. Rev.* C **85** 024304
[34] Farine M, Pearson J M and Rouben B 1978 *Nucl. Phys.* A **304** 317
[35] Horowitz C J and Piekarewicz J 2001 *Phys. Rev. Lett.* **86** 5647–50
[36] Abrahamyan S *et al* (PREX collaboration) 2012 *Phys. Rev. Lett.* **108** 112502
[37] Tarbert C M *et al* (Crystal Ball at MAMI and A2 Collaboration) 2014 *Phys. Rev. Lett.* **112** 242502
[38] Hebeler K, Lattimer J M, Pethick C J and Schwenk A 2010 *Phys. Rev. Lett.* **105** 161102
[39] Furnstahl R J 2002 *Nucl. Phys.* A **706** 85
[40] Stone J R, Miller J C, Koncewicz R, Stevenson P D and Strayer M R 2003 *Phys. Rev.* C **68** 034324
[41] Steiner A W, Prakash M, Lattimer J M and Ellis P J 2005 *Phys. Rep.* **411** 325
[42] Epelbaum E, Hammer H W and Meißner U G 2009 *Rev. Mod. Phys.* **81** 1773–825
[43] Hagen G, Hjorth-Jensen M, Jansen G R, Machleidt R and Papenbrock T 2012 *Phys. Rev. Lett.* **109** 032502
[44] Wienholtz F *et al* 2013 *Nature* **498** 346–9
[45] Hagen G, Papenbrock T, Hjorth-Jensen M and Dean D J 2014 *Rep. Prog. Phys.* **77** 096302
[46] Steiner A W 2012 *Phys. Rev.* C **85** 055804
[47] Audi G, Wang M, Wapstra A H, Kondev F G, MacCormick M, Xu X and Pfeiffer B 2012 *Chin. Phys.* C **36** 1287
[48] Wang M, Audi G, Wapstra A H, Kondev F G, MacCormick M, Xu X and Pfeiffer B 2012 *Chin. Phys.* C **36** 1603
[49] Guillot S, Servillat M, Webb N A and Rutledge R E 2013 *Astrophys. J.* **772** 7
[50] Steiner A W 2014 O2scl: object-oriented scientific computing library (record ascl:1408.019) http://adsabs.harvard.edu/abs/2014ascl.soft08019S