



## From noise to information: The transfer function formalism for uncertainty quantification in reconstructing the nuclear density

P. G. Giuliani <sup>1,2,\*</sup> and J. Piekarewicz <sup>1,†</sup>

<sup>1</sup>*Department of Physics, Florida State University, Tallahassee, Florida 32306, USA*

<sup>2</sup>*Department of Statistics and Probability, and Facility for Rare Isotope Beams, Michigan State University, East Lansing, Michigan 48824, USA*



(Received 31 March 2021; accepted 1 June 2021; published 2 August 2021)

**Background:** The neutron distribution of neutron-rich nuclei provides critical information on the structure of finite nuclei and neutron stars. Parity violating experiments—such as PREX and CREX—provide a clean and largely model-independent determination of neutron densities. Such experiments, however, are challenging and expensive, which is why sound statistical arguments are required to maximize the information gained.

**Purpose:** We introduce a new framework, the “transfer function formalism,” aimed at uncertainty quantification, model selection, and experimental design in the context of neutron densities.

**Methods:** The transfer functions (TFs) are built analytically by expressing the linear response of the objective function (e.g.,  $\chi^2$ ) to small perturbations of the data. Using the TF formalism, we are able to analyze the expected overall uncertainty—quantified in terms of bias and variance—of the mean square radius and interior density of  $^{48}\text{Ca}$  and  $^{208}\text{Pb}$ .

**Results:** Using relativistic mean field models as a proxy for the weak-charge density—and assuming that a total of five measurements could be performed on the weak form factor of  $^{48}\text{Ca}$  and  $^{208}\text{Pb}$ —we identify the optimal models and experimental locations that minimize the uncertainty in the extraction of the radius and interior density. We also explore the use of the TF formalism to understand the influence of prior distributions for the model parameters, as well as the optimization of model hyperparameters not constrained by the data.

**Conclusions:** We establish how the choice of experimental locations and the model that is used can have a significant impact on the final uncertainties of the extracted quantities of interest. For challenging experiments such as CREX and PREX, a proper quantification of such uncertainties is critical. We have demonstrated how the TF formalism provides several advantages for this type of analysis.

DOI: [10.1103/PhysRevC.104.024301](https://doi.org/10.1103/PhysRevC.104.024301)

### I. INTRODUCTION

Nuclear saturation, the existence of an equilibrium density, is a hallmark of the nuclear dynamics. Shortly after Chadwick's discovery of the neutron, the semiempirical mass formula of Bethe and Weizsäcker [1,2] was conceived to predict the binding energy of atomic nuclei. Using only a handful of parameters, the semiempirical mass formula provides a remarkably good description of the masses of stable nuclei by regarding the atomic nucleus as an incompressible quantum drop consisting of  $Z$  protons and  $N$  neutrons ( $A = Z + N$ ). Among the earliest predictions of the semiempirical mass formula was the  $A^{1/3}$  scaling of the nuclear size. Indeed, assuming an incompressible drop at an equilibrium (or “saturation”) density of  $\rho_0 \approx 0.15 \text{ fm}^{-3}$  yields a root-mean-square radius of

$$R(A) = r_0 A^{1/3}, \quad \text{where } r_0 = 3 \sqrt{\frac{3}{4\pi\rho_0}} \approx 1.17 \text{ fm}. \quad (1)$$

While the description of atomic nuclei as an incompressible quantum drop has stood the test of time, we now know that at a finer scale the distribution of nucleons is much more interesting and complex. Shell corrections, deformation, and pairing correlations are all important dynamical effects that impact the spatial distribution in atomic nuclei. To date, the most precise knowledge of the nuclear density comes from mapping the charge distribution of atomic nuclei [3,4]. Starting with the pioneering work of Hofstadter in the late 1950s [5] and continuing to this day, elastic electron scattering has painted the most accurate picture of the atomic nucleus. Our knowledge of the nuclear size, surface thickness, and saturation density all originate from such studies, that have provided some of the most stringent constraints on nuclear properties. For example, the root-mean-square charge radius of  $^{208}\text{Pb}$  is known with exquisite precision:  $R_{\text{ch}} = 5.5012(13) \text{ fm}$  [4].

Electron scattering is an ideal tool to map the charge distribution because the electromagnetic interaction is well known and the coupling (“fine structure”) constant is small. So, in a plane wave impulse approximation, the differential cross section for the elastic scattering of an electron from a spinless

\*pgg15@my.fsu.edu; giulianp@frib.msu.edu

†jpiekarewicz@fsu.edu

target may be written as follows [6]:

$$\left(\frac{d\sigma}{d\Omega}\right) = \left(\frac{d\sigma}{d\Omega}\right)_{\text{Mott}} Z^2 F_{\text{ch}}^2(Q^2), \quad (2)$$

where  $Z$  is the electric charge of the nucleus and  $Q^2 \geq 0$  is the square of the spacelike four-momentum transfer. The Mott cross section represents the scattering of a relativistic (massless) electron from a spinless and structureless target, and is given exclusively in terms of kinematical variables and the fine structure constant. Deviations from the structureless limit are encoded in the charge form factor, which has been normalized to 1 at zero momentum transfer,  $F_{\text{ch}}(Q^2=0)=1$ . The distribution of electric charge in a nucleus—which is carried mostly by the protons—is obtained from the Fourier transform of the charge form factor.

This favorable situation stands in stark contrast to our knowledge of the distribution of weak charge, which is dominated by the neutrons because the weak charge of the proton is small [7,8]. Probing neutron densities has traditionally relied on hadronic experiments involving strongly interacting probes, such as pions, protons, and alpha particles, that are hindered by uncontrolled approximations related to the reaction mechanism, medium modifications to the underlying two-body interaction, and hadronic distortions. For a recent review on this topic see Ref. [9] and references contained therein. For symmetric ( $N=Z$ ) nuclei, the expectation is that both proton and neutron densities will have the same shape, with the proton distribution extending slightly farther out because of the Coulomb repulsion. However, for heavy neutron-rich nuclei—which best illustrate the notion of nuclear saturation—the excess neutrons are pushed out against surface tension, creating a neutron-rich skin. Indeed, the interior baryon density of  $^{208}\text{Pb}$  is expected to be fairly constant and close to  $\rho_0$ . As such, the interior baryon density of  $^{208}\text{Pb}$  may provide the physical observable that is most closely related to  $\rho_0$  [10].

It is also possible to measure weak charge densities with much smaller systematic uncertainties by relying on electroweak probes that offer a clean and model-independent alternative to strongly interacting probes. However, this requires a more challenging and sophisticated class of experiment, such as coherent elastic neutrino-nucleus scattering (CEvNS) or parity violating electron nucleus scattering. The enormous advantage of these electroweak experiments is that the weak  $Z^0$  boson couples preferentially to neutrons because of the small weak charge of the proton [7,8]. For example, in the case of CEvNS the cross section is directly proportional to the square of the weak charge form factor. That is [11,12],

$$\left(\frac{d\sigma}{dT}\right) = \frac{G_F^2}{8\pi} M \left[2 - 2\frac{T}{E} - \frac{MT}{E^2}\right] Q_{\text{wk}}^2 F_{\text{wk}}^2(Q^2), \quad (3)$$

where  $G_F$  is Fermi's constant,  $Q_{\text{wk}} = -N + (1 - 4 \sin^2\theta_W)Z$  is the weak charge of the nucleus written in terms of the weak-mixing angle, and the weak form factor has been normalized to  $F_{\text{wk}}^2(Q^2=0)=1$ . The remaining quantities are all of kinematical origin:  $E$  is the incident neutrino energy,  $T$  the nuclear recoil energy, and  $Q^2=2MT$ . In particular, at forward angles the differential cross section is proportional to the square of

the weak charge of the nucleus  $Q_{\text{wk}}^2 \approx N^2$ . The approximate  $N^2$  scaling is the hallmark of the coherent reaction and the main reason for the identification by Freedman of CEvNS as having favorable cross sections [13], even if it took more than four decades for its experimental confirmation [14,15].

Although CEvNS holds enormous promise in the determination of neutron densities, the parity-violating electron program has become a precision tool in the determination of both hadronic/nuclear structure and electroweak physics. Following the 30-year-old idea by Donnelly, Dubach, and Sick who proposed the use of parity violating electron scattering (PVES) as a clean probe of neutron densities [16], the pioneering Lead Radius EXperiment (PREX) at the Jefferson Laboratory (JLab) extracted the weak radius of  $^{208}\text{Pb}$ , providing for the first time model-independence evidence in favor of a neutron-rich skin [17,18]. To reach the original goal of a  $\pm 0.06$  fm determination of the weak radius of  $^{208}\text{Pb}$ , the follow-up PREX-II campaign has now been completed and has delivered on the promise to determine the neutron radius of  $^{208}\text{Pb}$  with a precision that is about three times better than the original PREX measurement. By combining both experiments the following value for the neutron skin thickness of  $^{208}\text{Pb}$  was reported [19]:  $R_{\text{skin}} = R_n - R_p = (0.283 \pm 0.071)$  fm. This result challenges several experimental measurements and theoretical predictions that systematically underestimate the newly reported value of  $R_{\text{skin}}$  [9]. At the same time, the ongoing CREX campaign will provide the first electroweak determination of the weak radius of  $^{48}\text{Ca}$  [20,21]. Beyond JLab, the Mainz Energy recovery Superconducting Accelerator (MESA), envisioned to start operations by 2023 [22], may be able to determine the weak radius of both  $^{48}\text{Ca}$  and  $^{208}\text{Pb}$  with increased precision [9]. Besides its intrinsic value as a fundamental nuclear-structure observable, the neutron skin thickness of  $^{208}\text{Pb}$ , defined as the difference between the neutron and proton root-mean-square radii  $R_{\text{skin}} \equiv R_n - R_p$ , is strongly correlated to the slope of the symmetry energy at saturation density [23–26]. The symmetry energy at saturation density is a fundamental parameter of the equation of state of neutron-rich matter that impacts the structure, composition, and cooling mechanism of neutron stars [27–33].

A parity violating asymmetry emerges from the difference in the scattering between right- and left-handed polarized electrons. In a plane wave impulse approximation, the parity violating asymmetry from a spinless target may be written as follows [16]:

$$A_{PV}(Q^2) = -\frac{G_F Q^2}{4\pi\alpha\sqrt{2}} \frac{Q_{\text{wk}} F_{\text{wk}}(Q^2)}{Z F_{\text{ch}}(Q^2)}, \quad (4)$$

where  $\alpha$  is the fine structure constant and the nuclear contribution enters as the ratio of the weak to the charge form factor. Given that  $F_{\text{ch}}$  is known from (parity conserving) electron scattering measurements, the parity violating asymmetry determines the weak form factor which, in turn, is dominated by the neutron distribution.

To date, PREX, PREX-II, and CREX have focused on extracting the weak radius  $R_{\text{wk}}$  from a single measurement at a relatively low momentum transfer. Yet additional features

of the weak charge density can be revealed by measuring the parity violating asymmetry at higher momentum transfers. In particular, if  $A_{PV}$  could be measured at several momentum transfers, then the entire weak charge form factor and its associated density could be determined. Such an experimental program may require measurements of  $A_{PV}$  at about six values of  $Q^2$ , a task that may be feasible for  $^{48}\text{Ca}$  [34]. For  $^{208}\text{Pb}$ , such a task is significantly more challenging given that at high momentum transfer the elastic cross section is small because of the strong suppression from the nuclear form factor. Nevertheless, with two experimental points it may be sufficient to gain valuable insights into the weak charge form factor of  $^{208}\text{Pb}$  over a significant range of momentum transfers [10,35]. Regardless, with asymmetries of the order of one part per million [17,18], PVES experiments are both highly expensive and enormously challenging, so robust statistical arguments—above and beyond a compelling physics case—should be made in the quest for an optimal experimental design. Such is the central goal of the present paper.

In this paper, we present a novel statistical analysis—the “transfer function formalism”—inspired from the treatment of noise in signal processing theory [36]. In such a framework, the transfer function is a general function that models a device output for each possible input. In our particular case, we define the transfer function in terms of coefficients that encode the linear part of the response of the fitted model parameters to small changes in the data inputs. We have already implemented an early version of these ideas to estimate the bias and variance of models within the proton puzzle context [37] and in Ref. [38] to estimate the effect of dispersive corrections on the  $^{12}\text{C}$  elastic cross section.

Within the transfer function formalism, the noise is propagated in the measured observable to the uncertainty in the quantity of interest. Given that each single measurement in the data has an associated transfer function, an important feature of the formalism is that we can identify those *critical points*, if any, that are responsible for driving most of the uncertainty. For example, in this paper we are interested in quantifying the statistical error in the extracted weak charge radii of both  $^{48}\text{Ca}$  and  $^{208}\text{Pb}$  from the experimental error in their corresponding weak charge form factor. Values of the form factor with higher transfer functions will propagate their errors more efficiently to the total variance of the calculated weak charge radii. Using the transfer function ( $\mathcal{TF}$ ) formalism, we aim to quantify the ability of seven different models to accurately determine both the interior (saturation) density and mean square radius of the weak charge distribution. Given that the electric charge distribution of both nuclei is accurately known, we are able to validate our formalism against known data before making predictions for the unknown weak charge distribution.

The performance of the seven models is evaluated in terms of bias and variance [39], similar to the approach implemented in [37,40] to extract the charge radius of the proton from electron scattering data. The “bias-variance tradeoff” is an important concept in statistics and machine learning that addresses the complexity of a model. If the model is too simple, it will result in a poor description of the data (underfit = high bias). If the model is too complex, it will be extremely sensitive to the random dispersion in the data (overfit = high

variance). The bias-variance tradeoff is the inevitable conflict that ensues when trying to simultaneously minimize these two critical sources of error.

The rest of this paper is organized as follows. Section II includes a brief review of the main concepts involved in the discussion of nuclear form factors and density distributions. We also discuss statistical concepts related to our proposed formalism, such as Bayesian inference and bias-variance tradeoff. Section III presents a detailed account of the transfer function formalism and how it is implemented in the context of the bias-variance tradeoff. Section IV contains a compilation of our main results. We start this section by testing and validating our method using the experimentally known charge densities of both  $^{48}\text{Ca}$  and  $^{208}\text{Pb}$  as a proxy for the unknown weak charge densities. Finally, Sec. V presents our final remarks and vision for the future. In addition, we provide several Appendices that contain useful information in the form of supporting tables and figures, as well as mathematical proofs of the central concepts that have been developed.

The core idea of the transfer function formalism is that, for small perturbations in the input of a system, the response of the system is perturbed a *proportional* amount. This idea is clearly not new and it has been implemented in many scientific and engineering problems for centuries (consider for example the concept of Green’s functions). On the statistics front, we have found several related concepts such as the adjoint method [41, p. 203], the influence functions [42, p. 45], and the sensitivity of the system response [43, Sec. III F], for example. On the nuclear physics front, we have found studies such as [44] which by using expressions similar to the transfer functions analyze the sensitivity of the model parameters to different observables within the context of optical potentials. However, despite our best efforts, we were not able to find a direct application to model selection, the analysis of the influence of priors, and the description of both bias and variance, such as the one we developed in this work.

## II. THEORETICAL BACKGROUND

### A. Nuclear density and form factor

The electric charge density  $\rho_{\text{ch}}(\mathbf{r})$  and the weak charge density  $\rho_{\text{wk}}(\mathbf{r})$  describe the spatial distribution of electric charge and weak charge in the atomic nucleus, respectively. In the case of  $\rho_{\text{ch}}(\mathbf{r})$ , elastic electron scattering experiments determine the ground state charge density by measuring the differential cross section, which for a spinless nucleus is given by Eq. (2). In the case of the weak charge density, the aim is to extract the weak charge form factor from measuring the parity violating asymmetry given by Eq. (4).

Having extracted the corresponding form factors  $F_{\text{ch}}$  and  $F_{\text{wk}}$  from experiments, the nuclear charge density and weak charge density are obtained through a Fourier transform. To simplify the notation, no subscripts (either “ch” or “wk”) will be included henceforth, except when this omission may create confusion. The density and form factor are related as follows:

$$\rho(\mathbf{r}) = \int \frac{d^3q}{(2\pi)^3} e^{i\mathbf{q}\cdot\mathbf{r}} F(\mathbf{q}), \quad (5)$$

where  $|\mathbf{q}|=q=\sqrt{Q^2}$  in the limit in which the nuclear recoil can be ignored. For a spinless nucleus the density distribution is spherically symmetric so it becomes

$$\rho(r) = \frac{1}{2\pi^2 r} \int_0^\infty F(q) \sin(qr) q dq. \quad (6)$$

Alternatively, the inverse Fourier transform can be written as

$$F(q) = \frac{4\pi}{q} \int_0^\infty \rho(r) \sin(qr) r dr. \quad (7)$$

Note that we have adopted the following normalization condition for both electric and weak distributions:

$$F(q=0) = \int \rho(r) d^3r = 1. \quad (8)$$

Finally, the mean squared radius of the spatial distribution is given by

$$R^2 \equiv \langle r^2 \rangle = \int \rho(r) r^2 d^3r = 4\pi \int_0^\infty \rho(r) r^4 dr. \quad (9)$$

## B. Models, parameters, and errors

Several parametrizations (or models) exist in the literature to describe nuclear densities and their associated form factors [45]. In this paper, we study the performance of seven models in total: Fourier-Bessel [46], Helm [47], symmetrized Fermi function (SF) [48] of two, three, and four parameters, and two hybrid models obtained from combining the SF with a Fourier-Bessel expansion (SF + B) and the SF with a sum of Gaussians (SF + G). Note that we did not consider the original sum of Gaussians model [49] since certain conditions were difficult to implement within the transfer function formalism. Moreover, we found that for the small number (five) of data points here considered, the sum of Gaussians did not provide a good fit to the data. Appendix B describes in detail the seven models employed in this work.

We assume that we have collected  $J$  experimental data points that we write as  $\mathbf{Y} = \{(q_j, y_j, \sigma_j)\}$ , where  $q_j$  is the  $j$ th value of the momentum transfer,  $y_j$  is the value of the form factor at  $q_j$ , and  $\sigma_j$  is the associated experimental error. In turn, we refer to the set of  $K$  calibration parameters of any particular model as  $\boldsymbol{\omega} = \{\omega_k\}$ . Finally, we denote as  $m = m(\boldsymbol{\omega})$  the quantity of interest that we want to estimate from the given experimental data. Such quantity, for example, the mean square radius of the weak-charge distribution, depends on the selection of experimental points through the fitted parameters  $\boldsymbol{\omega}$ .

### 1. Standard fitting protocol

A traditional approach used to estimate the optimal set of parameters  $\boldsymbol{\omega}$  that best describes the observed data is to minimize the sum of the squares of the residuals between the experiment and the model predictions. The residuals are contained in an objective (or cost) function  $\chi^2$  defined as follows:

$$\chi^2 = \sum_{j=1}^J \frac{(F(q_j, \boldsymbol{\omega}) - y_j)^2}{\sigma_j^2}, \quad (10)$$

where  $F(q_j, \boldsymbol{\omega})$  represents the model predictions of the form factor. The optimal set of fitted parameters is obtained by minimizing the objective function and is denoted by  $\boldsymbol{\omega}_0 \equiv \text{argmin}(\chi^2)$ . Fundamental to the quantification of the model uncertainties is the behavior of the objective function in the vicinity of the optimal point  $\boldsymbol{\omega}_0$ . Such a behavior is imprinted in the Hessian matrix of  $\chi^2$  which is computed from its second derivatives evaluated at the optimal value. That is, matrix elements of the  $K \times K$  Hessian matrix are given by

$$\begin{aligned} \mathcal{H}_{i,k} &\equiv \frac{1}{2} \left( \frac{\partial^2 \chi^2}{\partial \omega_i \partial \omega_k} \right)_0 \\ &= \sum_{j=1}^J \frac{1}{\sigma_j^2} \left[ \left( \frac{\partial F(q_j, \boldsymbol{\omega})}{\partial \omega_i} \right) \left( \frac{\partial F(q_j, \boldsymbol{\omega})}{\partial \omega_k} \right) \right. \\ &\quad \left. + [F(q_j, \boldsymbol{\omega}) - y_j] \frac{\partial^2 F(q_j, \boldsymbol{\omega})}{\partial \omega_i \partial \omega_k} \right]_0. \end{aligned} \quad (11)$$

The inverse of the Hessian matrix,  $\mathcal{H}^{-1}$ , often called the error or covariance matrix, is used to estimate uncertainty and correlations associated with the fitted parameters as well as with other quantities [50]. For example, the square of the standard error (or standard deviation) of  $m(\boldsymbol{\omega})$  is given by

$$\Delta m^2 = \nabla m \mathcal{H}^{-1} \nabla m|_{\boldsymbol{\omega}_0}, \quad (12)$$

where  $\nabla m$  is the gradient of  $m$  with respect to the parameters  $\omega_k$ , and all quantities are evaluated at  $\boldsymbol{\omega} = \boldsymbol{\omega}_0$ .

### 2. Bayesian approach

An alternative framework to estimate model parameters and to quantify their statistical properties which has been gaining popularity in the physics community is the Bayesian approach [51,52]. Within this framework, the posterior distribution of model parameters  $\boldsymbol{\omega}$  given the experimental data  $\mathbf{Y}$  is given by Bayes's theorem:

$$P(\boldsymbol{\omega}|\mathbf{Y}) = \frac{P(\mathbf{Y}|\boldsymbol{\omega})P(\boldsymbol{\omega})}{P(\mathbf{Y})}, \quad (13)$$

where  $P(\mathbf{Y}|\boldsymbol{\omega})$  is the likelihood that a given set of model parameters describes the experimental data,  $P(\boldsymbol{\omega})$  is the prior distribution of model parameters, and  $P(\mathbf{Y})$  is the evidence, which can be treated as a normalization constant to enforce  $\int P(\boldsymbol{\omega}|\mathbf{Y}) d\boldsymbol{\omega} = 1$ . The prior distribution encapsulates our prior knowledge (or beliefs) of the distribution of model parameters. Such prior beliefs will be refined as a result of the additional experimental information contained in the likelihood, which ultimately yields an updated distribution of model parameters  $P(\boldsymbol{\omega}|\mathbf{Y})$ .

Once the posterior distribution  $P(\boldsymbol{\omega}|\mathbf{Y})$  is obtained, the average value of any quantity  $m$  and its associated error may be estimated from integrating over the probability distribution. That is,

$$\langle m \rangle = \int m(\boldsymbol{\omega}) P(\boldsymbol{\omega}|\mathbf{Y}) d\boldsymbol{\omega}, \quad (14a)$$

$$\Delta m^2 = \int (m(\boldsymbol{\omega}) - \langle m \rangle)^2 P(\boldsymbol{\omega}|\mathbf{Y}) d\boldsymbol{\omega}, \quad (14b)$$



where  $\langle m \rangle$  denotes the average—or central—value of  $m$ . In the case of the likelihood, it is often assumed that it is related to the  $\chi^2$  function introduced in Eq. (10) as follows:

$$P(\mathbf{Y}|\boldsymbol{\omega}) = e^{-\chi^2(\mathbf{Y},\boldsymbol{\omega})/2}. \quad (15)$$

Hence, reference to the maximum likelihood is equivalent to the minimum value of  $\chi^2$ . For the prior distribution it is common to assume an uncorrelated Gaussian distribution of model parameters, namely,

$$P(\boldsymbol{\omega}) = e^{-\phi^2(\boldsymbol{\omega})/2}, \quad \text{where} \quad (16a)$$

$$\phi^2(\boldsymbol{\omega}) = \sum_{k=1}^K \left( \frac{\omega_k - \omega_k^0}{\sigma_k} \right)^2, \quad (16b)$$

where  $\omega_k^0$  is our prior estimate for the central value of  $\omega_k$  and  $\sigma_k$  is the estimated uncertainty. Small values of  $\sigma_k$  will make the distribution sharply peaked around  $\omega_k^0$  and the fitting procedure more “prior driven.” Conversely, large values of  $\sigma_k$  reflect a large uncertainty in the model parameters so the fitting procedure becomes more “data driven.” Under the prior and likelihood definitions, the posterior distribution takes the following form:

$$P(\boldsymbol{\omega}|\mathbf{Y}) = e^{-\tilde{\chi}^2(\mathbf{Y},\boldsymbol{\omega})/2} = e^{-(\chi^2(\mathbf{Y},\boldsymbol{\omega})+\phi^2(\boldsymbol{\omega}))/2}, \quad (17)$$

where  $\tilde{\chi}^2$  now encodes contributions from both the likelihood and the prior. For an optimal point  $\boldsymbol{\omega}_0 = \text{argmin}(\tilde{\chi}^2)$ , the behavior of  $\tilde{\chi}^2$  around the minimum is encoded in the augmented Hessian matrix  $\tilde{\mathcal{H}}$  defined as

$$\tilde{\mathcal{H}}_{i,k} \equiv \frac{1}{2} \left( \frac{\partial^2 \tilde{\chi}^2}{\partial \omega_i \partial \omega_k} \right)_0 = \mathcal{H}_{i,k} + \delta_{ik} \frac{1}{\sigma_k^2}, \quad (18)$$

where  $\mathcal{H}$  is the Hessian of  $\chi^2$  defined in Eq. (11) and  $\delta_{ik}$  is the Kronecker delta. If the adopted prior includes correlations between the different parameters, then Eq. (16b) will be written as a quadratic form  $\phi^2(\boldsymbol{\omega}) = \boldsymbol{\omega} \Sigma^{-1} \boldsymbol{\omega}$ , where the matrix  $\Sigma$  contains the (prior) covariances between parameters. In such a case Eq. (18) would have to be modified accordingly.

### C. Bias, variance, and MSE

Our objective is to identify which of the seven models defined in Sec. II B will best perform—using a criterion to be precisely defined shortly—in extracting the radius and interior density when faced with real experimental data on the weak charge form factor. Given that the experimental results have yet to be published, we rely on synthetic data generated by a set of five covariant energy density functionals that we refer to as *generators*:  ${}^n F_{\text{true}}(q)$  ( $n=1, \dots, 5$ ). The particular set of accurately calibrated functionals are RMF012, RMF016 (commonly referred to as FSUGarnet), RMF022, RMF028, and RMF032 [33]. The main difference among these generators is the assumed value for the yet to be accurately determined neutron skin thickness of  ${}^{208}\text{Pb}$ ; for example, RMF022 predicts a neutron skin thickness of  $\sim 0.22$  fm. For each data point generated for the weak charge form factor there is an associated error  $\sigma_j$  which resembles realistic experimental uncertainties. Once a generator is

selected, any observable of interest  $m$  can be calculated directly from the synthetic data.

As in Refs. [37,40], we evaluate the performance of each of the seven models using a bias-variance tradeoff criterion. Bias is understood as the discrepancy between the true value of  $m$  (coming from one of the generators  ${}^n F_{\text{true}}$ ) and the extracted value. In contrast, the variance is the spread in the extracted value of  $m$  as given by the square of the standard deviation (SD); see Eqs. (12) and (14b). Thus, we quantify the performance of a model by combining the bias and variance into the mean squared error (MSE), defined as

$$\text{MSE}^2(m, \mathbf{q}, \boldsymbol{\sigma}, n) \equiv \text{Bias}^2 + \text{Variance}. \quad (19)$$

Note that we have highlighted the dependence of the MSE on the quantity  $m$ , the locations of the momentum-transfer points  $\mathbf{q}$ , the associated experimental errors  $\boldsymbol{\sigma}$ , and the generator index  $n$ . The MSE is a good indicator of the score, as it captures the bias vs variance tradeoff often present in predictive models across the fields of statistics and machine learning [39]. Finally, we define the squared average of the MSE by combining the predictions from the  $n$  different generators:

$$\langle \text{MSE} \rangle^2(m, \mathbf{q}, \boldsymbol{\sigma}) = \frac{1}{N} \sum_{n=1}^N \text{MSE}^2(m, \mathbf{q}, \boldsymbol{\sigma}, n). \quad (20)$$

The same formula may be used to obtain the squared average of the bias and variance from the different “truths” (generators).

An abstract representation of these concepts is illustrated in Fig. 1. On the entire function space depicted with the blue surrounding box, the truth region (in green) is assumed to be spanned by the set of all generators, with the blue triangle within this region representing a single member of such family (for example RMF022). The set of possible functions adopted to reproduce the data are also displayed. For example, model 1 (in purple) could be the symmetrized Fermi function whereas model 2 (in orange) could be the Bessel expansion. In turn, the purple and red stars are the members of these respective families that are obtained after fitting the data generated by the blue triangle. The corresponding stars are associated with specific values of their parameters  $\boldsymbol{\omega}$ . Under some metric which depends on our choice for  $m$ , the “distance” from the stars to the triangle will represent the bias. In the example, the bias is larger for model 1. Due to the unavoidable errors in the experimental data, there will be uncertainty in the exact location of both stars. This uncertainty is represented by the dashed contour which size illustrates the variance for each model; in this example the variance is larger for model 2. Once we allow the blue triangle to explore the “truth space,” the combination of the accumulated bias and variance makes the score, as indicated in Eq. (20). The task is to identify the model with the best score, which emerges from a compromise between the bias and variance.

A possible approach to calculate the bias and variance for each model would be to create many noisy realizations of the data to accumulate enough statistics and then apply the standard fitting protocol described in Sec. II B 1 [40]. An alternative approach would be to directly compute the Bayesian integrals highlighted in Sec. II B 2 [35]. In the following

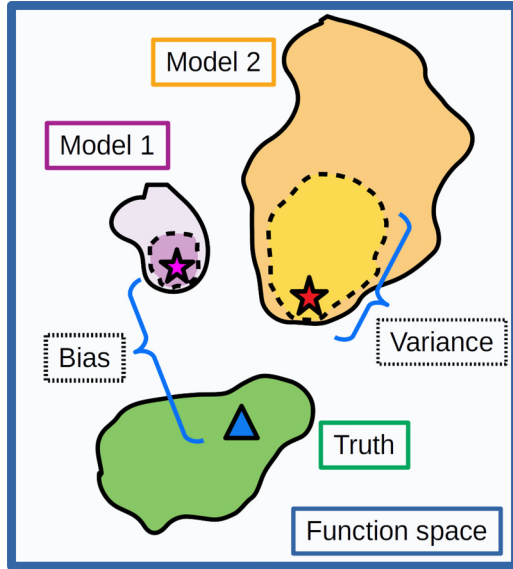


FIG. 1. Abstract representation of the impact of bias and variance on recovering information. The entire function space is represented by the blue enclosing square. The green blob represents the collections of all the truths (generators) such as RMF012 while the blue triangle is one of its members. The purple and yellow blobs represents all the possible members (for different parameters sets) from model 1 and model 2 respectively. The purple and red stars are two particular members of those groups. The bias is shown as the distance between the recovered members (the stars) of each model and the blue triangle. The variance is shown as the dashed contours surrounding each star.

section we present a third option: a new formalism that—under certain assumptions—can speed up these calculations, aid in the identification of “critical” points in the data, provide a highly intuitive picture of the propagation of the uncertainty, and be extended from model selection to model building.

### III. TRANSFER FUNCTION FORMALISM

We want to understand how the uncertainty—both in terms of bias and variance—gets propagated from the experimental data  $\mathbf{Y}$  to the observable of interest  $m$ . To do so, we invoke the “transfer functions,” a central concept in signal processing and control theory [36]: if we can make a linear map connecting an arbitrary change in the input to the associated change in the output, then analyzing the dynamic response of the system becomes straightforward. Note that for nonlinear systems such map is not possible. Nevertheless, if the changes in the input are “small,” then linearizing the system around its equilibrium point might suffice for most practical purposes [36].

In our case, the “system” is the  $\chi^2$  fit in which the inputs are the experimental data  $\mathbf{Y}$  and the output could be either the model parameters  $\boldsymbol{\omega}$  or any quantity  $m$ . Under the transfer function formalism we assume that, once the minimum  $\boldsymbol{\omega}_0$  of  $\chi^2(\boldsymbol{\omega})$  is found, then small changes in the value of the data  $y_j$  will also produce small changes in both the parameters and any observable  $m$ . That is, we assume that the response of the system to the perturbation is linear. To this end, our main

objective is to write

$$\delta m = \sum_{j=1}^J \mathcal{TF}_j^m \delta y_j, \quad (21)$$

where  $\delta m$  is the small change in the observable  $m$  in response to small changes  $\delta y_j$  in the experimental data  $y_j$ . The transfer functions (TF), denoted by  $\mathcal{TF}_j^m$ , encode the changes in  $m$  as a result of a change  $\delta y_j$  in a given individual input  $y_j$ . That is, there is a total of  $J$  transfer functions for each observable  $m$ . The adopted notation uses a subscript for the  $j$ th observation  $y_j$  and a superscript for the responding quantity  $m$ . We can now expand  $\mathcal{TF}_j^m$  in terms of the model’s parameters as follows:

$$\mathcal{TF}_j^m \equiv \frac{\partial m}{\partial y_j} = \sum_{k=1}^K \frac{\partial m}{\partial \omega_k} \frac{\partial \omega_k}{\partial y_j} = \nabla m \cdot \mathcal{TF}_j^\omega, \quad (22)$$

where  $\mathcal{TF}_j^\omega$  is a  $K$ -dimensional vector with its components being the transfer functions connecting a small change in each observation  $y_j$  to the response of the  $k$ th model parameter  $\omega_k$ . That is, in analogy to Eq. (21) we obtain

$$\delta \omega_k = \sum_{j=1}^J \mathcal{TF}_j^{\omega_k} \delta y_j. \quad (23)$$

As we show in Appendix A, the general expression for  $\mathcal{TF}_j^\omega$  is given by

$$\mathcal{TF}_j^\omega = \mathcal{H}^{-1} \nabla F(q_j, \boldsymbol{\omega}) \sigma_j^{-2}, \quad (24)$$

where the gradient  $\nabla F(q_j, \boldsymbol{\omega})$  is taken with respect to the model parameters  $\boldsymbol{\omega}$ . In the following subsections, we use the transfer functions to calculate both the variance and bias of any quantity of interest  $m$ .

One important point to note is that to calculate the transfer functions in Eq. (24) we need to compute first and second derivatives of the model  $F(q_j, \boldsymbol{\omega})$ . This is straightforward for the simple parametric models we use in this work (see Appendix B), but could become an issue if we use more complex models such as the RMF. These complex models, which account for the nuclear dynamics to calculate densities and form factors, use iterative procedures over nonlinear coupled differential equations to produce densities. For such cases, instead of calculating derivatives directly, one could resort to other options available such as performing finite differences [53], automatic differentiation [54], or derivative-free optimization methods such as POUNDERS [55].

#### A. Variance calculation

Equation (21) allows us to write the linear response of any quantity  $\delta m$  to a given set of small changes in the observations  $\delta \mathbf{y}$ . We interpret the errors in the experimental data as independent, Gaussian distributed random variables with mean zero and standard deviation  $\sigma_j$ . Hence, in this scenario, if we identify the perturbations  $\delta y_j$  as these Gaussian independent experimental errors, the variance in  $\delta m$  may be obtained by

adding each term in Eq. (21) in quadrature:

$$\Delta m^2 = \sum_{j=1}^J (\mathcal{TF}_j^m)^2 \sigma_j^2, \quad (25)$$

where the transfer functions  $\mathcal{TF}_j^m$  are evaluated in the model's parameters that are obtained from the original central values of the experimental points  $\mathbf{Y}$  ( $\delta\mathbf{y} = \mathbf{0}$ ). This is analogous to a Taylor series expansion in which the derivatives of the expanded function are evaluated at the unperturbed variable. Note that Eq. (21) may still be used to calculate the variance even in the more general case when there are correlations or the distribution is not Gaussian. However, in this case we would have to perform the appropriate integrals on  $\delta m$ , as a function of  $\delta\mathbf{y}$ , times the joint probability distribution  $P(\delta\mathbf{y})$ .

One of the main advantages of Eq. (25) is that it separates, up to some degree, the contribution from each observation  $y_j$  to the entire variance  $\Delta m^2$ . As we show in Sec. IV A, this separation allows us to identify those data points having undue influence on the variance. This information could be valuable in experimental design through the optimal allocation of resources, such as beam time in scattering experiments (see Sec. IV in Ref. [34]). Note that the Hessian  $\mathcal{H}^{-1}$  in  $\mathcal{TF}_j^m$  effectively mixes all observations, so it is not possible to cleanly isolate the contribution from each data point. Nevertheless, Eq. (25) provides a more efficient and natural way of addressing the influence of each data point as compared to other well-known approaches, such as those represented by Eqs. (12) and (14a). We also note that, in comparing the variance calculated in Eq. (25) to that obtained from the standard approach in Eq. (12), the results are identical in the limit in which the nonlinear part of the Hessian matrix [the terms proportional to second derivatives of  $F$  in Eq. (11)] may be ignored. We give a formal proof of this statement in Appendix A. In cases in which the model parametrizations depend nonlinearly on the model parameters, then the variances will differ.

So, which (if any) of the two approaches is correct in the event that the calculated variances differ from each other? Although the answer is not obvious, the transfer function formalism seems to be in agreement with those analyses in which many realizations of the data are generated via Monte Carlo sampling [40]. The traditional approach in Eq. (12) deviates from the observed Monte Carlo results, an issue generally discussed in statistics under the name of “model misspecification” (for more information on this topic see theorem 5.23 and example 5.25 in Ref. [56]). However, we note that the accuracy of both approaches deteriorates as the errors in the data become large enough for the nonlinearities to become important. In such a case, the Gaussian approximation, namely, the notion that the entire  $\chi^2$  landscape may be described by the second derivatives at the minimum, is no longer valid.

As a final remark, we note that the variance computed as in Eq. (25) changes with the location of the momentum transfers  $q_j$ . This change happens not only because the experimental errors  $\sigma_j$  may change with  $q_j$ , but also because the transfer functions themselves depend on the location of  $q_j$ . Indeed, by exploring the available  $q$  range, we could find the optimal locations that minimize the variance of the quantity of interest. In this way, we can answer a fundamental question

in experimental design: *given the available resources, how do we select the optimal locations of  $q_j$  to minimize the statistical uncertainty?* [35]. When exploring the  $q$  range we must be aware that the fitted parameters  $\omega$  will also change, which in turn will impact the value of each of the transfer functions  $\mathcal{TF}_j^m$  introduced in Eq. (22). This suggests the need to refit the optimal parameters every time a new set of  $q_j$  is considered. As we shall see below, one of the important results of the present formalism is that, under certain assumptions, refitting may be skipped altogether.

## B. Bias calculation and the central function

In this section we study the bias as explained in Sec. II C. That is, the discrepancy between the true value of the observable of interest  $m$  and the one extracted by the model. A traditional way of calculating the bias would be to fit the model parameters to the data  $\mathbf{Y} = \{(q_j, y_j, \sigma_j)\}$  by minimizing Eq. (10), and then calculate  $m(\omega_0)$ . Alternatively, we may compute  $\langle m \rangle$  from Eq. (14a). In both cases the bias is obtained by subtracting the true value  $m_{\text{true}}$ . Regardless of the approach, we must either refit the model parameters or perform the integrals over the posterior distribution for every combination of points  $q_j$  that we want to test. The main reason to explore the behavior of both the bias and the variance as we change the  $q_j$  locations is that we may be interested in finding the optimal locations that minimize the mean squared error defined in Eq. (19). As we will show shortly, once we cast the bias calculation under the TF framework, it is possible to avoid refitting as we explore different sets of  $q_j$  locations.

As indicated in Eq. (25), the main sources that contribute to the variance are the individual data errors  $\sigma_j$ , which get propagated to the quantity of interest through the transfer functions. To write the corresponding expression for the bias in the context of the TF formalism, we must identify the sources that replace  $\sigma_j$  in the variance equation. To do so, we first study how the fitted parameters obtained from minimizing  $\chi^2$  evolve in the parameter space, as the observations  $q_j$  move in their available momentum transfer range. We refer henceforth to the obtained parameters for a given set of locations  $q_j$  as the “empirical” parameters  $\omega_e$ . Note that we employ the specific notation  $\omega_e$ , rather than the more general  $\omega_0$  defined after Eq. (10). The set  $\omega_e$  refers exclusively to parameters obtained directly from data (or pseudodata) without any perturbation, while the set  $\omega_0$  represents the minimum of  $\chi^2$  in any situation, even when we perturb the data by small amounts  $\delta\mathbf{y}$ .

As an example, we show in Fig. 2 how  $\omega_e$  evolves as the location of a single measurement changes. In this case, the model being fitted is the two-parameter ( $\omega = [c, a]$ ) symmetrized Fermi function. We assume that measurements can be made at two different values of the momentum transfer, one fixed at  $q_1 = 0.7 \text{ fm}^{-1}$  and the other one  $q_2$  that is allowed to move along the orange curve in the  $[1.25, 2.2] \text{ fm}^{-1}$  range. The value of the weak form factor of  $^{48}\text{Ca}$  at each of these two points is predicted using the generator RMF012 [33]. The four possible locations of the second point  $q_2$ , labeled respectively as 1,2,3, and 4, are displayed as green circles on the orange curve. Each of these locations, in combination with  $q_1$  determines a single optimized value  $\omega_e$ . The associated

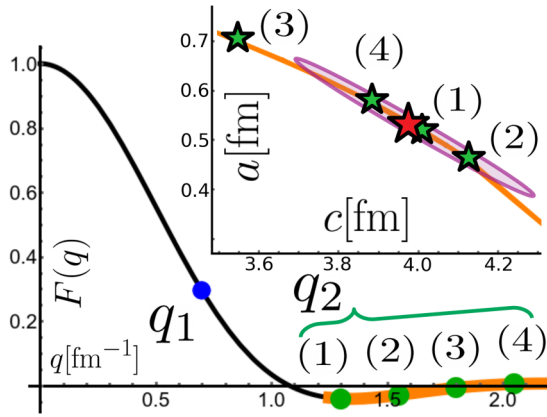


FIG. 2. Evolution of  $\omega_e$  as the second location  $q_2$  is moved in the orange region. The model being fit is the symmetrized Fermi function of two parameters:  $\omega = [c, a]$ . The inset plot shows how the orange curve gets mapped into the parameter space, highlighting four locations in green. Although not easy to observe in the plot the mapping “folds into itself” in the parameter space. The red star represents the central parameters [see the text before Eq. (26)], while the purple ellipse represents the 95% confidence interval for a fit using  $\mathbf{q} = \{0.7, 1.8\} \text{ fm}^{-1}$  but with  $y$  values dictated by the central function [see the text after Eq. (26) and in Fig. 3].

values of  $\omega_e$ , one for each choice of  $q_2$ , are displayed in the inset as the green stars in the parameter space. We expect that as the value of the second point changes, so will the value of any derived quantity  $m$ , which will ultimately result in a change to the bias.

As can be seen in Fig. 2 the four empirical parameters  $\omega_e$  displayed with the green stars do not move too far away from some central value  $\omega_c$  (shown as a red star). From this central location and using the transfer functions, we can describe the entire trajectory of the empirical parameters—particularly how they deviate from the central value  $\omega_c$  to first order ( $\omega_e \approx \omega_c + \delta\omega$ ) for a given set of  $J$  observations. That is,

$$\delta\omega = \sum_{j=1}^J \mathcal{TF}_j^\omega \eta_j, \quad (26)$$

$$\eta_j \equiv y_j - F(q_j, \omega_c). \quad (27)$$

We refer to  $\eta_j$  as the quantity that is now driving the change in the parameters to distinguish it from an arbitrary perturbation  $\delta y_j$ . The transfer functions in Eq. (26) are evaluated at the central parameters  $\omega_c$  and in the “data” created by  $F(q, \omega_c)$ . The main idea that we are exploiting is that the minimum of  $\chi_c^2$  defined as

$$\chi_c^2(\omega) = \sum_{j=1}^J \frac{(F(q_j, \omega) - F(q_j, \omega_c))^2}{\sigma_j^2}, \quad (28)$$

where is  $\omega_c = \text{argmin}(\chi_c^2)$  since  $\chi_c^2(\omega_c) = 0$ . This expression for  $\chi_c^2$  is identical to the one defined in Eq. (10), but with the real observations  $y_j$  replaced by  $F(q_j, \omega_c)$ . From Eq. (26) we can say that, if the values of  $F(q_j, \omega_c)$  are perturbed such that  $F(q_j, \omega_c) \rightarrow F(q_j, \omega_c) + \eta_j = y_j$ , then the central parameters  $\omega_c$  will respond by moving by  $\omega_c \rightarrow \omega_c + \delta\omega \approx \omega_e$ .

If this last approximation,  $\omega_e \approx \omega_c + \delta\omega$ , is accurate enough for our purposes, then we can say that  $m_e = m(\omega_e)$  may be approximated by the central value  $m_c = m(\omega_c)$  plus a small correction  $\delta m$ :

$$m_e = m_c + \delta m = m_c + \sum_{j=1}^J \mathcal{TF}_j^m \eta_j. \quad (29)$$

With these tools at hand, we can write the bias for the quantity of interest  $m$  as follows:

$$\text{Bias}(m) \equiv m_e - m_t = \left[ m_c + \sum_{j=1}^J \mathcal{TF}_j^m \eta_j \right] - m_t, \quad (30)$$

where  $m_t$  is the true value of  $m$  and the  $\mathcal{TF}_j^m$  are evaluated at the central parameters  $\omega_c$ . We note that, if  $m_c - m_t$  is negligible, the bias is completely driven by the  $\eta_j$ , analogous to how the variance in Eq. (25) was driven by the errors  $\sigma_j$ .

Using the same model and generator as in Fig. 2, we display in Fig. 3 an estimate of the bias using the interior density of  $^{48}\text{Ca}$  as the observable of interest. As in Fig. 2, we keep the value of the first point fixed at  $q_1 = 0.7 \text{ fm}^{-1}$  and select the second point at  $q_2 = 1.8 \text{ fm}^{-1}$ , which corresponds to the third point in Fig. 2. Equation (26) is then used to approximate the empirical parameters, which in turn provide an estimate for the empirical density  $\rho(r)_{\text{Emp}}$ , which is depicted as the blue dashed line in Fig. 3(a). The deviation of the empirical density from the central density  $\rho(r)_{\text{Cen}}$  can be understood in terms of the  $\eta_j$ 's: the difference between the central form factor  $F(q, \omega_c)$  and the true form factor  $F(q)_{\text{True}}$  evaluated at  $q_1$  and  $q_2$ . To appreciate these minor differences, we enlarge a window around  $q_2$  and show  $\eta_2 = y_2 - F(q_2, \omega_c)$  in the inset of Fig. 3(b). The hollow blue circle corresponds to  $F(q_2, \omega_c)$  while the filled one corresponds to  $F(q_2)_{\text{True}}$ .

Under the linear approximation assumed in Eq. (26), these  $\eta_j$  will move the central  $\omega_c$  [red star in the inset of Fig. 3(a)] towards the approximated empirical parameters  $\omega_e$  (blue star). Since this is a nonlinear model, Eq. (26) is indeed just an approximation and the change in the parameters in this case was underpredicted. This can be seen when comparing the position of the blue star in Fig. 3 with the green star (3) in Fig. 2. Nevertheless, the empirical density is not too different from the density shown in Fig. 3(a).

We close this section by discussing the selection of  $\omega_c$ . In principle, the precise location of  $\omega_c$  should not have a significant impact on our calculations provided that the actual change  $\delta\omega$  is linear in  $\eta_j$ . In the interest of clarity, and given that the experimental observable is the form factor but we are interested in extracting the spatial density, we distinguish between two main choices for  $\omega_c$ :

- (1) Central function fit: We define  $\omega_c$  as the value that minimizes the  $L^2$  norm between the model  $F(q, \omega)$  and the true function  $F(q)_{\text{true}}$  in the momentum transfer space  $q$ , as in Figs. 2 and 3. The expectation is that  $\omega_c$  should be relatively close to most of the possible obtainable parameters for different locations of the data. We refer to these parameters as  $\omega_{\text{Cen}}$ .
- (2) Optimal fit: We define  $\omega_c$  as the parameters that make the central estimation  $m_c$  as close as possible to the



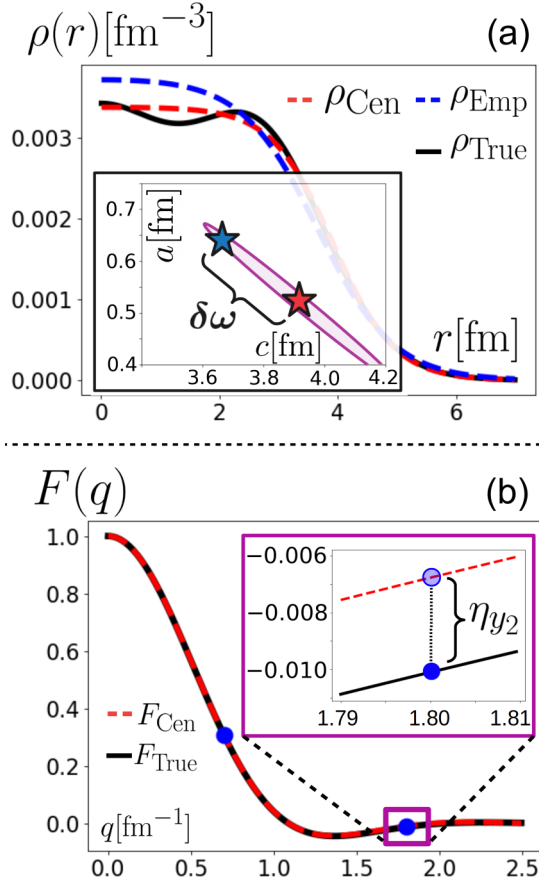


FIG. 3. Example of the change in the central parameters driven by deviations between the central and the true form factors. (a) Reconstructed density from  $\omega_c$ , reconstructed density from  $\omega_e$  using Eq. (26), and the true density (red, blue, and black lines, respectively). The inset plot shows the locations in the parameter space of  $\omega_c$  and the estimated  $\omega_e$  (red and blue stars, respectively). Also shown in purple is the 95% confidence ellipse when fitting the parameters to the central data  $\{[q_1, F(q_1, \omega_c)], [q_2, F(q_2, \omega_c)]\}$  using the errors described in Sec. IV. (b) Central form factor  $F(q, \omega_c)$  and true form factor (red dashed and black lines, respectively). The inset plot shows a close-up of the difference between these two curves ( $\eta_2$ ) in the neighborhood of  $q_2 = 1.8 \text{ fm}^{-1}$ .

true value  $m_t$ . For example, if we are interested in modeling the interior density,  $\omega_c$  should be chosen by fitting the models directly to the spatial density, effectively minimizing the  $L^2$  norm between the model density  $\rho(r, \omega)$  and the true density  $\rho(r)_{\text{true}}$ . Note that this procedure is not feasible in the case of real data given that scattering experiments can only access the form factor directly and not the density. However, the advantage of this option is that, if  $m_c - m_t$  is negligible, then the total bias is dominated by the  $\eta_j$ , making easier the search for the optimal locations. We refer to these parameters as  $\omega_{\text{Opt}}$ , and will use them extensively in Sec. III E.

### C. Mean squared error

Having constructed the bias and variance within the TF formalism, we write the mean squared error (MSE) as

$$\text{MSE}^2 = \left( (m_c - m_t) + \sum_{j=1}^J [\mathcal{TF}_j^{(m)}] \eta_j \right)^2 + \sum_{j=1}^J [\mathcal{TF}_j^{(m)}]^2 \sigma_j^2. \quad (31)$$

Recall that the MSE is the quantity that we aim to optimize in an effort to find a compromise between the bias and the variance. For the specific quantity of interest  $m$ , the MSE will depend on the selected data points  $q_j$  (e.g., the momentum-transfer points), the associated errors  $\sigma_j$ , and the input values  $y_j$  (e.g., the weak form factor), with the last quantity drawn from experimental data or pseudo data generated by mean field models. The equations developed in the TF framework enable us to address the expected MSE for a given set of experimental data and then report which model has the lowest error, as implemented in Ref. [40]. However, if the experiment is still in its design phase, then the TF formalism may be used to optimize the MSE not only with respect to the model, but also relative to the location of the data and the distribution of errors.

Naturally, a unique set of central parameters  $\omega_c$  will be associated with a given model (e.g., Fourier-Bessel) and generator (e.g., RMF012). Thus, unless we suspect that variations under different choices of model and generator are negligible, each MSE should be calculated with its own parameters  $\omega_c$ . Indeed, these parameters  $\omega_c$  are necessary for the numerical calculation of each  $\mathcal{TF}^m$ . Moreover, it is important to note that the  $\mathcal{TF}_j^m$  from the bias term in Eq. (31) are evaluated at the central parameters  $\omega_c$ , while the  $\mathcal{TF}_j^m$  associated with the variance are not. From our construction in Sec. III A, these  $\mathcal{TF}_j^m$  should be evaluated at the parameters associated with the observed data (the empirical parameters  $\omega_e$  defined in Sec. III B). There are two options on how to obtain  $\omega_e$ . One may select  $\omega_e$  directly from Eq. (26) in the event that the linear relationship encoded in the equation provides a good approximation. However, if we suspect that the linear approximation is not accurate, for example when dealing with a strongly nonlinear model, then we should resort to a numerical algorithm informed by the data  $\mathbf{Y}$  every time the data locations change. This will allow to calculate the bias directly from the empirical parameters with no need for Eq. (30). Indeed, to guarantee numerical accuracy, we use this last option for all the nonlinear models that we explore in this paper, while we resort to Eq. (30) for linear models. For example, in the case shown in Fig. 3, the calculated change in  $m = \rho(0)$  from the central value using the TF, underestimates the true change by around 30%. In Sec. III E we describe an important implementation of Eq. (30) that would not be possible with a numerical optimizer and which can be useful even when dealing with nonlinear models.

### D. Priors under the TF formalism

If we have Gaussian priors of the form presented in Eq. (16a), then we can treat each prior term as a pseudo

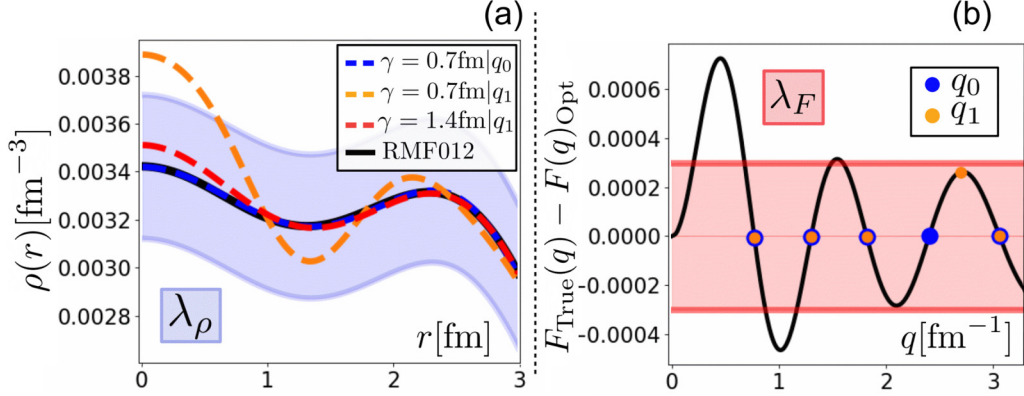


FIG. 4. (a) Recovered densities using the SF + G model for two different data locations. The generated (“true”)  $^{48}\text{Ca}$  weak density is shown in black, while the orange and blue dashed lines are the SF + G obtained densities for the two data sets  $\mathbf{q}_0$  and  $\mathbf{q}_1$ . The red dashed line is obtained by the same SF + G model on  $\mathbf{q}_1$ , but with the hyperparameter  $\gamma$  set to 1.4 fm instead of 0.7 fm. The blue band is associated with the scale  $\lambda_\rho$ . (b) The locations of the two data sets in orange and blue. These data sets only differ on the location of the fourth point  $q_4$ . The black line shows the difference between the true function (RMF012) and the optimal function  $[F(q_j, \omega_{\text{Opt}})]$ . The red band represents the scale  $\lambda_F$ .

observation. These priors act in the same way as true observations in  $\chi^2$ :  $[F(\omega, q_j) - y_j]/(\sigma_j^2)$ , by pulling the value of  $\omega$  in a particular direction in the parameter space. The new  $\tilde{\mathcal{H}}$  defined in Eq. (18) should be used when calculating the observation’s transfer functions defined in Eq. (24).

The effect of the priors will not only be the conversion of  $\mathcal{H}$  to the new  $\tilde{\mathcal{H}}$ , but each prior estimate value  $\omega_k^0$  will have its own transfer function as if it were an observation:

$$\mathcal{T}\mathcal{F}_k^\omega \equiv \frac{\partial \omega}{\partial \omega_k^0} = \tilde{\mathcal{H}}^{-1} I_k \sigma_k^{-2}, \quad (32)$$

where  $I_k$  is the  $k$ th column of the identity matrix of size  $K \times K$  (a vector with 0 in every entry except with a 1 on entry  $k$ ).  $I_k \sigma_k^{-2}$  is analogous to  $\nabla F(\omega, q_j) \sigma_j^{-2}$  when calculating  $\mathcal{T}\mathcal{F}_j^m$ . We use the subindex  $k$  to denote that what we are perturbing is not  $y_j$ , but rather the prior estimate value  $\omega_k^0$ . In the case where the prior contains correlations then Eq. (16b) is written as quadratic form  $\phi^2(\omega) = \omega \Sigma^{-1} \omega$ . In this case  $I_k \sigma_k^{-2}$  will be replaced by the  $k$ th column of the matrix  $\Sigma^{-1}$ .

These transfer functions of the priors “observations” will appear at the same level as regular observations in the variance and bias equations (25) and (30). For the bias part, the associated  $\eta_j$ , which we will call  $\tilde{\eta}_k$ , is defined as the difference between the value of  $[\omega_c]_k$  (the  $k$  entry of the central parameters) and the prior “observation”  $\omega_k^0$ .

### E. Reconstruction bias and the optimal function

In this section, we describe the estimation of a nonintuitive bias which we call the reconstruction bias, that strongly depends on the  $q_j$  locations.

This reconstruction bias is closely related to what we observed in the example in Fig. 3. When dealing with incomplete data (a few  $q_j$  points on the entire form factor curve, for example), the empirical parameters  $\omega_e$  we recover might deviate considerably from the best parameters that reproduce the entire true function  $[\omega_{\text{Cen}}$  in the case of  $F(q)_{\text{True}}$  or  $\omega_{\text{Opt}}$  in the case of  $\rho(r)_{\text{True}}$ ]. As a consequence, the second term inside the

brackets in Eq. (30) could grow substantially. This will result in a significant bias even in flexible models which in principle could reproduce the true function almost perfectly.

For illustration purposes, in this section we use the generator RMF012. Figure 4(a) shows the recovered  $^{48}\text{Ca}$  weak density using the SF + G model with two sets,  $\mathbf{q}_0$  and  $\mathbf{q}_1$ , of five data points each (blue and orange dashed lines). The first data set is  $\mathbf{q}_0 = [0.77, 1.30, 1.82, 2.41, 3.06] \text{ fm}^{-1}$ , while the second one is identical to the first except for the fourth location:  $\mathbf{q}_1 = [0.77, 1.30, 1.82, 2.70, 3.06] \text{ fm}^{-1}$ , as seen in Fig. 4(b).

The blue and orange SF + G model in Fig. 4(a) has its hyperparameter controlling the size of the Gaussians set to  $\gamma = 0.7 \text{ fm}$ , close to the nucleon size (Appendix B shows a detailed description of the SF + G model and its hyperparameters). The orange curve has a clear bias in the interior density. This is the reconstruction bias. It is not the same type of bias, as shown, for example, by the SF model which by definition has a flat interior and cannot reproduce the interior structure of  $^{48}\text{Ca}$ .

To better analyze this phenomenon, we use the optimal function, i.e., the parameter set  $\omega_{\text{Opt}}$  from the SF + G model that creates the density in the  $r$  space that is closest to the true density. By definition, any deviation from  $\omega_{\text{Opt}}$  will result in a stronger bias. We want to understand this increase in bias in terms of the difference  $\delta\omega \equiv \omega - \omega_{\text{Opt}}$ . To simplify our analysis, we just focus on  $\rho(0)$ .

Figure 4(b) shows the difference in momentum space between the optimal function  $F(q, \omega_{\text{Opt}})$  and the true (RMF2012)  $F(q)_{\text{True}}$ . The blue points are situated exactly at the locations where both functions have the same value, while the fourth orange point is at a place where these functions differ.

Similar to what we developed in Sec. III B, we can imagine that our data are currently centered at the optimal function  $F(q, \omega_{\text{Opt}})$  (the minimum of  $\chi^2$  is currently at the optimal parameters). The  $y_j$  values are slightly perturbed from their starting values by small quantities  $\eta_j$  now defined

as

$$\eta_j \equiv F(q_j)_{\text{True}} - F(q_j, \omega_{\text{Opt}}). \quad (33)$$

Using our TF formalism, we can write how much  $\rho(0)$  changes to first order due to these displacements, when compared to the predicted  $\rho(0)$  by the optimal function. Since just  $\eta_{y_4}$  is nonzero, we have

$$\delta\rho(0) = [\mathcal{T}\mathcal{F}_4^{(\rho(0))}]_{\eta_{y_4}}. \quad (34)$$

Therefore, if we move  $q_4$  around in the  $q$  space while leaving the other four  $q_j$  in place, those locations with a high product value  $\mathcal{T}\mathcal{F}_4^{(\rho(0))}\eta_{y_4}$  will create a strong bias. Such is the case shown in Fig. 4(a) by the orange curve. In this particular example, if we want to maintain a bias of less than 5% [ $\delta\rho(0) < 5\%$ ], we can only locate  $q_4$  in around 15% of the possible momentum transfer range  $[0, 3.5] \text{ fm}^{-1}$  (see Appendix C for more details).

Let us call  $\lambda_F$  the expected scale for the size of  $\eta_j$  for our range of  $q_j$  values. Let us call  $\lambda_\rho$  the desired threshold we want for our accuracy in the estimation of  $\rho(0)$ . Figure 4 shows these two scales as the red and blue bands, respectively. Replacing the transfer function by its explicit expression in Eq. (34), we will maintain that threshold in  $\rho(0)$  as long as

$$|\nabla\rho(0)\mathcal{H}^{-1}\nabla F(q_4)/\sigma_4^2| \leq \frac{\lambda_\rho}{\lambda_F}. \quad (35)$$

For our particular problem, we can set  $\lambda_F \approx 3 \times 10^{-4}$  and  $\lambda_\rho \approx 1.5 \times 10^{-4} \text{ fm}^{-3}$  [roughly 5% of  $\rho(0)$ ], which makes the ratio  $\frac{\lambda_\rho}{\lambda_F} = \frac{1}{2} \text{ fm}^{-3}$ . For the SF + G model at  $q_4 = 2.7 \text{ fm}^{-1}$ , the product  $\mathcal{T}\mathcal{F}_4^{(\rho(0))}\eta_4 \approx 1.3 \text{ fm}^{-3}$ , which implies that the reconstruction bias falls outside of our tolerable range  $\lambda_\rho$ .

In an actual experiment, we would not know ahead of time the optimal  $q_j$  locations where  $\mathcal{T}\mathcal{F}_j^{(\rho(0))}\eta_j$  is small. Therefore, we could not use a model like SF + G with such a limited  $q$  range and strong reconstruction bias.

For the SF + G, the situation seems to be mainly driven by the first Gaussian with  $R_1 = 0$ , which scales as  $\gamma^{-3}$  in the  $\rho$  space. Based on this, we decided to double the size of  $\gamma$ , from  $\gamma = 0.7 \text{ fm}$ , to  $\gamma = 1.4 \text{ fm}$  (which is the value we use in Sec. IV B).

Using this new value of  $\gamma$ , the new transfer function product value at  $q_4 = 2.7 \text{ fm}^{-1}$  is  $\mathcal{T}\mathcal{F}_4^{(\rho(0))}\eta_4 \approx 0.3 \text{ fm}^{-3}$  and the reconstruction bias is reduced considerably.<sup>1</sup> This is shown by the red dashed line in Fig. 4(a). Moreover, with this new value of  $\gamma$ ,  $q_4$  can be allocated in around 30% of the possible momentum transfer range  $[0, 3.5] \text{ fm}^{-1}$  while maintaining a bias of less than 5% (see Appendix C for more details).

We close this section with two important remarks regarding this type of analyses. First, they could ultimately serve not only to model selection, but to model building. In many cases,

a hyperparameter (such as  $\gamma$ ) might be fixed to a suboptimal value that hinders rather than helps the extraction of information from experimental data.

Second, these analyses can give an estimate of the impact of the reconstruction bias which is impossible to get by just focusing on the statistical errors in experimental data. Consider the purple 95% confidence ellipse in Figs. 2 and 3 centered at  $\omega_c$ , the red star. This ellipse does not contain the actual estimated parameters from the data, i.e., the green star (3) in Fig. 2 (let us recall that the blue star in Fig. 3 is just the linear approximation). The reverse is also true: the ellipse centered at the true empirical parameters (not shown) will not contain the red star, which reproduces the true weak charge density in Fig. 3 better than the approximated empirical blue density.

The errors  $\sigma_j$  and the deviations  $\eta_j$  are two unrelated scales. Confidence ellipses are usually related to the errors  $\sigma_j$  but the reconstruction bias is related to the  $\eta_j$ . There is no reason for the ellipse obtained from the true data to contain  $\omega_{\text{Opt}}$ , i.e., the set of parameters in our model that best describe the real curve that generated that data. However, this is often the assumed scenario when extracting information from experiments.

#### IV. RESULTS: ANALYZING CHARGE AND WEAK CHARGE DENSITIES

In this section, we discuss in detail the process used to select the optimal models and the impact that varying the locations of the selected momentum transfers  $q_j$  will have on the extracted densities of both  $^{48}\text{Ca}$  and  $^{208}\text{Pb}$ . In particular, we are interested in describing the root-mean-square radius and interior density of the charge and weak charge distributions.

The calculation of the MSE for the charge radius is straightforward as it involves a single, well-defined quantity. For the interior density, we allocate 30 grid points between  $r=0 \text{ fm}$  and  $r=3 \text{ fm}$  for  $^{48}\text{Ca}$ , and between  $r=0 \text{ fm}$  and  $r=5 \text{ fm}$  for  $^{208}\text{Pb}$ . The MSE for the interior density is then constructed by averaging in quadrature the single MSE for each individual point. That is,

$$\text{MSE}[\text{Interior}]^2 \equiv \frac{1}{30} \sum_{i=1}^{30} \text{MSE}[\rho(r_i)]^2. \quad (36)$$

We then combine both the radius and interior MSE into a single quantity known as the figure of merit (FOM):

$$\text{FOM}^2 \equiv \left( \frac{\text{MSE}[\text{Radius}]}{\Delta R} \right)^2 + \left( \frac{\text{MSE}[\text{Interior}]}{\Delta \rho} \right)^2, \quad (37)$$

where  $\Delta\rho$  and  $\Delta R$  are natural scales associated with each quantity; roughly 5%–10% and 1% for the interior density and radius, respectively; see Table I. By adjusting these scales, the FOM could be made more sensitive to the radius or the interior density. Note that for both  $^{48}\text{Ca}$  and  $^{208}\text{Pb}$  the densities have been normalized to 1 rather than to the number of nucleons.

To simplify the analysis, we assume that the errors in the experimental data for both the charge and weak-charge form factor depend only on their assumed value at the selected momentum transfers. For example, following [35], we assume a constant value of  $\sigma(q)=0.005$  for  $^{208}\text{Pb}$ .

<sup>1</sup>To be rigorous, we should now move the other  $q_j$  values to the locations where the new optimal model is equal to the true function. Since they are almost in the same location, we decided to keep them in the same place to simplify the discussion.

TABLE I. Natural scales for the uncertainties in the interior density and radius for  $^{48}\text{Ca}$  and  $^{208}\text{Pb}$ .

	$\Delta\rho$ ( $\text{fm}^{-3}$ )	$\Delta R$ (fm)
$^{48}\text{Ca}$	0.000 15	0.04
$^{208}\text{Pb}$	0.000 08	0.06

For the case of  $^{48}\text{Ca}$  we adopt the prescription given in Ref. [34] for the errors at the selected five momentum transfers  $[q_1, q_2, q_3, q_4, q_5]$ . When required, a simple function of the form  $\sigma(q) = \text{Max}(0.000\,57, 0.0081 - 0.003q)$  is used to interpolate between the selected  $q$  values. Once the data set and the selected model are specified, the FOM depends solely on the location of the momentum transfers. In the following subsections we illustrate how the value of the FOM can be minimized by optimizing such locations and how the TF formalism is the ideal tool for interpreting the results and further reducing the uncertainties, for example, by identifying critical measurements for error reduction.

### A. Electric charge densities

To test the new formalism we start by analyzing the well known experimentally determined electric charge density of  $^{48}\text{Ca}$  and  $^{208}\text{Pb}$  [3,45]. We illustrate the power and flexibility of the transfer functions formalism by describing the  $^{48}\text{Ca}$  data using a Fourier–Bessel expansion and a SF model in the case of  $^{208}\text{Pb}$ . Following the prescription of Ref. [34], we assign the starting values of the momentum transfer for  $^{48}\text{Ca}$  at  $q_0 = [0.9, 1.35, 1.8, 2.24, 2.69] \text{ fm}^{-1}$ . Similarly, for the case of  $^{208}\text{Pb}$ , we fixed the two starting locations as in Ref. [35] at  $q_0 = [0.5, 0.8] \text{ fm}^{-1}$ .

We show in Fig. 5 results obtained before and after optimizing the location of the momentum-transfer points. We display the original locations in blue and their shift to their optimal locations in orange, where the FOM is minimized subject to the following constraints:  $q_j \leq 3.5 \text{ fm}^{-1}$  for  $^{48}\text{Ca}$  and  $q_j \leq 2 \text{ fm}^{-1}$  for  $^{208}\text{Pb}$ . Beyond these limits, we assume

TABLE II. MSE results for the interior and radius for  $^{48}\text{Ca}$  and  $^{208}\text{Pb}$  for their respective original locations  $q_0$  and optimal locations  $q_m$ . Each quantity has been divided by its respective natural scale as defined in Table I.

	$^{48}\text{Ca}$		$^{208}\text{Pb}$	
	Interior	Radius	Interior	Radius
MSE ( $q_0$ )	1.27	1.37	0.27	0.77
MSE ( $q_m$ )	1.26	0.94	0.32	0.63

that the experimental challenge to measure such small cross sections can not be met. This may be better appreciated by displaying the form factor in a logarithmic plot, as in Fig. 6; note that the cross section is proportional to the *square* of the form factor. Note that the minimization of the FOM was done by running the Python NUMPY optimization library with the “TNC” method for ten different seeds including the original  $q_0$  choice. In Table II, we show results for the MSE for both nuclei in terms of their natural scales. The MSE in the interior was not substantially reduced for  $^{48}\text{Ca}$  and it even increased by  $\sim 20\%$  for  $^{208}\text{Pb}$ , as can be seen by the slightly larger error bands in Fig. 5. On the other hand, the MSE for the radius was improved by  $\sim 30\%$  for  $^{48}\text{Ca}$  and by  $\sim 20\%$  for  $^{208}\text{Pb}$ . These results are driven by our selection of scales which favored an improvement in the radius rather than in the interior density. Also, the radius is an easier quantity to constrain than the interior density. Note, however, that even minimizing the FOM with only the interior term does not significantly improve the interior density.

Finally, listed in Tables III and IV in Appendix D are the numerical values of the  $\mathcal{TF}_j$  times the respective error  $\sigma_j$  for the density at  $r=0$  fm and the radius for two sets of locations of the momentum transfer, namely, original  $q_0$  and optimal  $q_m$ . These individual values illustrate how much each measurement is currently impacting the variance in the radius and in the density at  $r=0$ . Note that in Eq. (25) each term  $\mathcal{TF}_j\sigma_j$  is added in quadrature. Therefore, the final variance is not linear on each component. Indeed, the quadrature equation

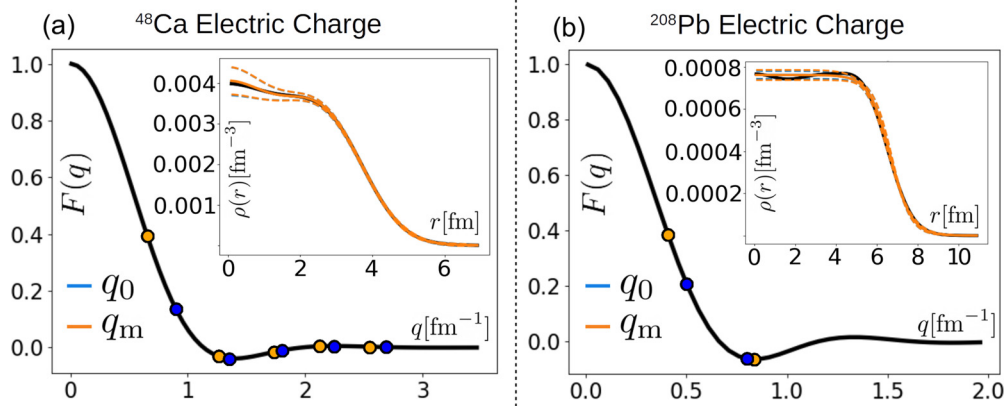


FIG. 5. Optimization of the FOM for  $^{48}\text{Ca}$  using a Fourier-Bessel expansion (a) and for  $^{208}\text{Pb}$  using the symmetrized Fermi function (b). The original locations of the momentum transfer  $q_0$  (blue points) are displaced to  $q_m$  (orange points) to minimize the FOM. The inset plots show the reconstructed charge densities with their respective error bands.



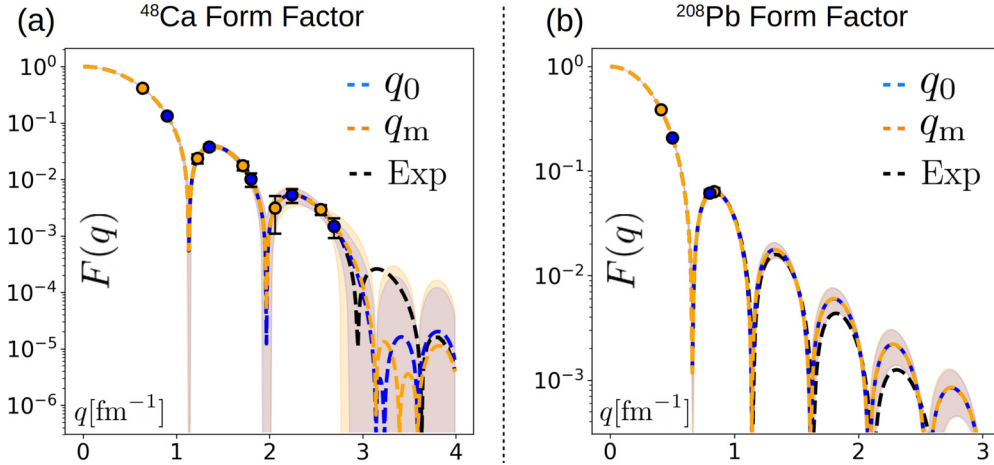


FIG. 6. Same as in Fig. 5, but now for the absolute value of the form factor using a logarithmic scale. The black curve represents the respective experimental charge form factor [3,45] and, for clarity, the central curves are displayed as dashed lines, whereas the error bands are shown as colored bands.

will enhance the effect of bigger numbers with respect to their smaller counterparts. For example, in the case  $^{48}\text{Ca}$  with the optimized set, the variance in  $\rho(0)$  is dominated by the observations at  $q_2$  and  $q_3$ , whereas for the radius the variance is largely driven by the form factor at  $q_1$ . A similar analysis for  $^{208}\text{Pb}$  reveals that the variance in  $\rho(0)$  is driven by  $q_2$ , whereas the measurement at  $q_1$  dominates the variance in the radius. Given that the radius is obtained from the slope of the form factor at zero momentum transfer and the interior density is controlled by the large- $q$  behavior of the form factor, the previous results are fully consistent with our expectations. Note that as the errors in the observations change, these statements might no longer hold true. Our main conclusion is that to reduce the final variance on each quantity  $m$  within this hypothetical experimental design—and to first approximation—these are the critical data locations that should be targeted for error reduction.

### B. Weak charge densities

We now proceed to compare the performance of each of the seven models mentioned in Sec. II B in reproducing the interior density and radius of the weak charge distribution of  $^{48}\text{Ca}$  and  $^{208}\text{Pb}$ . Appendix D presents the corresponding analysis for the charge densities.

Given that there is no experimental information on the weak charge form factors of  $^{48}\text{Ca}$  and  $^{208}\text{Pb}$ , we use Eq. (20) to calculate the squared average MSE from the five different generators obtained from Ref. [33], namely, RMF012, RMF016, RMF022, RMF028, and RMF032. As in the previous section, we start with five fixed locations  $q_0$  and then optimize these values to  $q_m$  to minimize the average FOM. We apply the same restrictions as in the example of the charge density:  $q_j \leq 3.5 \text{ fm}^{-1}$  for  $^{48}\text{Ca}$  and  $q_j \leq 2 \text{ fm}^{-1}$  for  $^{208}\text{Pb}$ . Note that since the goal is to minimize the average mean square error of all five generators, the resulting optimal values  $q_m$  only depend on the choice of the model. The starting locations for the momentum transfer in the case of  $^{48}\text{Ca}$  are once again fixed at  $q_0 = [0.90, 1.35,$

$1.8, 2.24, 2.69] \text{ fm}^{-1}$ , whereas for  $^{208}\text{Pb}$  they are now chosen at  $q_0 = [0.63, 0.94, 1.26, 1.57, 1.88] \text{ fm}^{-1}$ . Note that these values correspond to the special choice of  $q_\nu \equiv \nu\pi/R_{\text{cut}}$  for  $\nu \in [2, 6]$ , with the cutoff radius  $R_{\text{cut}} = 7 \text{ fm}$  for  $^{48}\text{Ca}$  [34] and  $R_{\text{cut}} = 10 \text{ fm}$  for  $^{208}\text{Pb}$ .

In Figs. 7 and 8 we display the performance of the seven models employed in the text to describe the weak charge of  $^{48}\text{Ca}$  and  $^{208}\text{Pb}$ , respectively. Shown in each figure are the resulting bias, standard deviation (SD), and MSE for the interior density (three bars on the left of each panel) and the weak charge radius (three bars on the right of each panel). The corresponding figures for the electric charge density are shown in Figs. 12 and 13 in Appendix D. We note that for a fixed model we obtained very similar results regardless of the particular RMF generator; as an example see Figs. 14 and 15 in Appendix E. This suggests that the conclusions that we draw within each model are robust, at least within the (RMF) family of generators considered in this study.

We want to highlight two main points from these results. First, changing the data locations (i.e., the selection of the various momentum transfers) has a significant impact on the performance of the model. For example, in the description of the weak charge density of  $^{208}\text{Pb}$  the performance of the Helm model improves by nearly a factor of two. Second, we observe large variations in the model performance when the original fixed locations  $q_0$  are adopted. Indeed, for  $^{48}\text{Ca}$  the Bessel-Fourier expansion outperforms the SF + B model by about a factor of 2. Such large discrepancy is often mitigated by selecting the optimal locations  $q_m$  for each model; see the dramatic improvement in the description of the weak charge radius of  $^{208}\text{Pb}$  when the optimal locations are adopted.

Based on these two points, we can conclude that the optimal model will strongly depend on the data structure, regarding both locations and errors. We should expect that variance driven models (like the Bessels) will outperform bias driven models (like the SF) in cases where the data errors are small. Nevertheless, for the number of experimental measurements and errors assumed in this example, we conclude

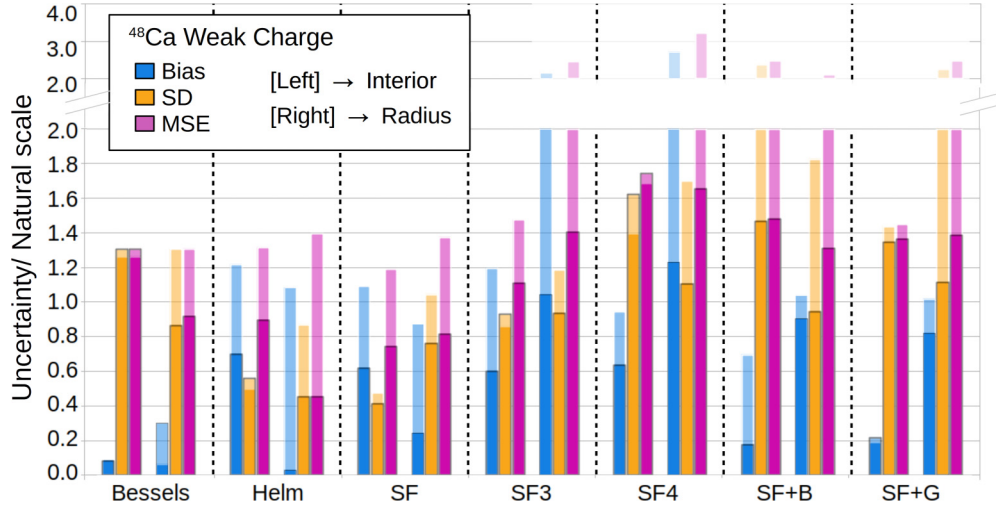


FIG. 7. Comparison of the performance of the seven models used in the text for recovering the interior density and weak charge radius of  $^{48}\text{Ca}$  as generated from a form factor “measured” at five points. The three columns to the left of each model show the bias, SD, and MSE for the interior density, whereas the three columns to the right display the bias, SD, and MSE for the radius. All quantities have been divided by their natural scales:  $\Delta\rho(\text{Ca}) = 0.00015 \text{ fm}^{-3}$  and  $\Delta R(\text{Ca}) = 0.04 \text{ fm}$ . The solid columns represent the optimal locations  $q_m$ , whereas the light borderless columns were obtained from the starting  $q_0$  points.

that the Helm and SF models are best suited for the simultaneous extraction of the radius and interior density of both nuclei. This could be expected for the case of  $^{208}\text{Pb}$  given that both the Helm and SF models are characterized by a flat interior density that provides our closest connection to the saturation density of infinite nuclear matter [10]. It might come as a surprise that these flat models outperform more flexible models like the Bessel-Fourier expansion which better describes the interior shell oscillations of  $^{48}\text{Ca}$ . The reason behind this finding is that we aim to minimize the MSE, which involves a combination of the bias and the variance.

On *average* (mean value), the Bessel model will provide a more genuine representation of the interior oscillations of the weak charge density of  $^{48}\text{Ca}$ . However, the noise level as quantified by the variance is so high that the expected deviation is large enough to make more desirable a flat description with smaller error bands. Figure 9 shows, for the weak charge density of  $^{48}\text{Ca}$  generated by RMF012, the extracted density by the Bessel and SF models. The Bessel expansion uses the original  $q_0$  locations proposed in [34], while the SF model uses the optimized locations  $q_m$  presented in Table IX in Appendix E.

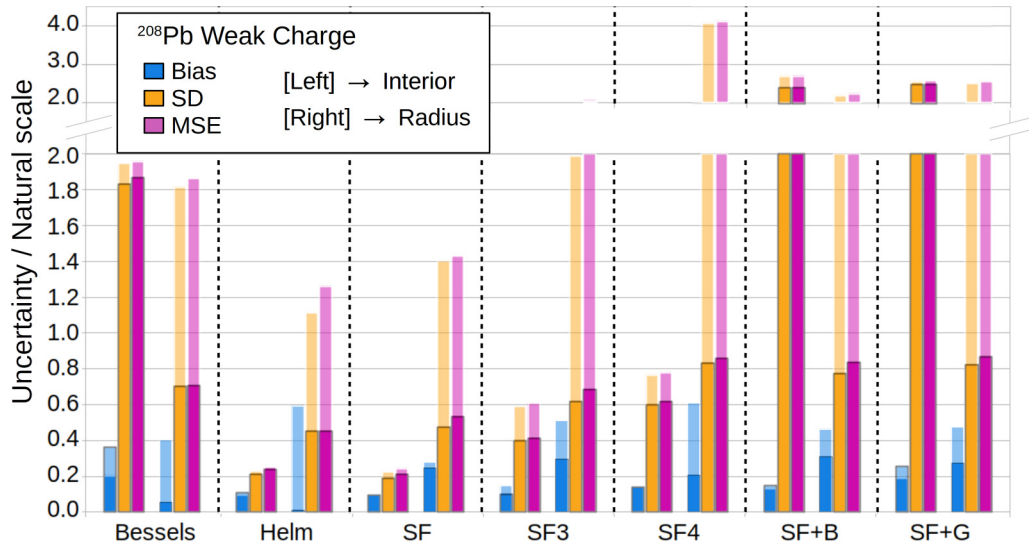


FIG. 8. Comparison of the performance of the seven models used in the text for recovering the interior density and weak charge radius of  $^{208}\text{Pb}$  as generated from a form factor “measured” at five points. The three columns to the left of each model show the bias, SD, and MSE for the interior density, whereas the three columns to the right display the bias, SD, and MSE for the radius. All quantities have been divided by their natural scales:  $\Delta\rho(\text{Pb}) = 0.00008 \text{ fm}^{-3}$  and  $\Delta R(\text{Pb}) = 0.06 \text{ fm}$ . The solid columns represent the optimal locations  $q_m$ , whereas the light borderless columns were obtained from the starting  $q_0$  points.

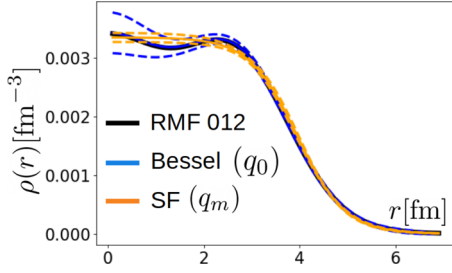


FIG. 9. Extracted weak density of  $^{48}\text{Ca}$  from the RMF012 generator using the Bessel and SF models. The Bessel model uses the original momentum transfers  $q_0$ , while the SF model uses its corresponding optimized values  $q_m$ . For the data and errors assumed, our analysis supports flat models like the SF or Helm, over the more flexible ones such as the Bessel expansion.

To better illustrate our result for  $^{48}\text{Ca}$ , consider that the RMF012 generator in Fig. 9 predicts a weak charge density that is slightly enhanced at  $r=0$  relative to the average interior density and then drops from the average around  $r \approx 1.2$  fm. We could then ask, how probable is it to conclude the opposite [i.e.,  $\rho(0) \leq \rho(1.2)$ ] even after adopting the optimal locations  $q_m$  for the Bessel model? To answer that question, let us define the new quantity of interest  $m \equiv \rho(1.2) - \rho(0)$  and investigate the probability that  $m \geq 0$ . Once the optimal  $q_m$  values are adopted, an average value of  $m = -2.3 \times 10^{-4} \text{ fm}^{-3}$  is obtained, suggesting that  $\rho(0) > \rho(1.2)$ , in agreement with the predictions from the RMF012 generator. But what about the variance in this result? The variance of this quantity can be calculated using the TF formalism from Eq. (25) as

$$\Delta m^2 = \sum_{j=1}^5 ([\nabla\rho(1.2) - \nabla\rho(0)] \mathcal{H}^{-1} \nabla F(q_j, \omega) \sigma_j^{-2})^2 \sigma_j^2, \quad (38)$$

where we have used Eqs. (22) and (24) to write the explicit form of  $\mathcal{TF}_j^m$ . Following this procedure, we obtain a standard deviation of  $\Delta m = 2.2 \times 10^{-4} \text{ fm}^{-3}$ . We note that the third measurement at  $q_3 = 1.73 \text{ fm}^{-1}$  has the largest impact on the the variance, followed by appreciable contributions from  $q_4 = 2.12 \text{ fm}^{-1}$  and  $q_5 = 2.55 \text{ fm}^{-1}$ ; see Table V in Appendix D for the values of  $q_m$ . Hence, under the assumption that the errors are Gaussian distributed random variables, we can infer that  $\rho(0) \leq \rho(1.2)$  in  $\sim 15\%$  of the experimental realizations. That is, if the experimental noise (primarily in  $q_3$ ,  $q_4$ , and  $q_5$ ) cannot be significantly reduced, we will conclude the incorrect oscillation structure in the interior density of  $^{48}\text{Ca}$  in one out of six experiments.

For the interior density of  $^{208}\text{Pb}$ , the best overall score was achieved by the SF model with a total MSE of  $0.21 \times \Delta\rho(\text{Pb}) = 1.68 \times 10^{-5} \text{ fm}^{-3}$ , where  $\Delta\rho(\text{Pb})$  is defined in Table I. Using again RMF012 as an example of a generator, we observe that the total variance in  $\rho(0)$  is mainly driven by the third observation, having a value of  $|\mathcal{TF}_3^{\rho(0)} \sigma_3| = 0.14 \times \Delta\rho(\text{Pb})$ . Taking  $\rho(0)$  as a representative value of the interior density, this implies that the measurement at  $q_3 = 0.77 \text{ fm}^{-1}$  should be primarily targeted for error reduction in order to improve the uncertainty in the saturation density  $\rho_0$ . We underscore that the interior density of  $^{208}\text{Pb}$  is a genuine

experimental observable that provides the closest connection to the saturation density of infinite nuclear matter. For a recent analysis on how a measurement of the interior density of  $^{208}\text{Pb}$  could constrain  $\rho_0$  see Ref. [10].

In the case of the weak charge radii of both nuclei, we found that they can be accurately determined using the Helm model: an MSE of  $0.45 \times \Delta R(\text{Ca}) = 0.018 \text{ fm}$  for  $^{48}\text{Ca}$  and of  $0.45 \times \Delta R(\text{Pb}) = 0.027 \text{ fm}$  for  $^{208}\text{Pb}$ , with both values of  $\Delta R$  listed in Table I. In the case of  $^{48}\text{Ca}$ , and relying again on RMF012, the total variance in  $R$  is uniformly distributed among the first three observations at  $q_1 = 0.51$ ,  $q_2 = 0.63$ , and  $q_3 = 0.77 \text{ fm}^{-1}$ , with values of  $|\mathcal{TF}_j^{\rho(0)} \sigma_j| \approx 0.2 \times \Delta R(\text{Ca})$ . Instead, for  $^{208}\text{Pb}$  we found that the total variance in  $R$  is driven by the two points closest to the origin, namely,  $q_1 = 0.37$  and  $q_2 = 0.40 \text{ fm}^{-1}$ , with values of  $|\mathcal{TF}_j^{\rho(0)} \sigma_j| = 0.29 \times \Delta R(\text{Pb})$ . To improve the uncertainty in the weak charge radii, the observations at these “low- $q$ ” points should be targeted for error reduction. These results are hardly surprising given that the weak charge radius is defined in terms of the slope of the associated form factor at the origin. It is worth noting that the weak charge radius of  $^{208}\text{Pb}$ , when combined with the corresponding (electric) charge radius into a neutron skin, provides a stringent constraint on the slope of the symmetry energy  $L$ , and ultimately on the radius of neutron stars [28]. In particular, a 1% determination of the weak charge radius of  $^{208}\text{Pb}$  translates into an uncertainty of about 40 MeV in the slope of the symmetry energy [57].

### C. The role of priors

The incorporation of priors lies at the heart of Bayesian statistics. Priors allow us to include physical biases and intuition as well as information from previous experiments. Moreover, priors play the important role of serving as leverage to reduce the variance of a model at the expense of increasing its bias [41,58]. This can be particularly beneficial for models such as the Bessel-Fourier expansion or SF + G, whose MSE is largely driven by the variance given the level of noise in the generated data.

In this section, we briefly explore the impact of an informed prior on the performance of the SF + G model as it pertains to the weak charge density of  $^{48}\text{Ca}$ . As we did earlier, we use the RMF012 generator to produce synthetic data. The proposed locations of the measurements are the original five values of the momentum transfer:  $q_0 = [0.9, 1.35, 1.8, 2.24, 2.69] \text{ fm}^{-1}$ . Note that the implementation of priors was discussed in Sec. II B 2 and extended to the TF formalism in Sec. III D. The SF + G model consists of the two-parameter symmetrized Fermi function plus three Gaussians “bumps” to account for shell oscillations in the interior. The Gaussians are centered in the interior at three different locations:  $[R_1, R_2, R_3] = [0, 1.3, 2.6] \text{ fm}$ . We analyze three prior options for the amplitude of the Gaussians ( $A_1, A_2, A_3$ ), while we leave the two intrinsic parameters of the SF model unconstrained. First, we consider a null prior that we refer to as  $P_0$ . Such “prior” effectively reproduces the original unconstrained SF + G model. Second, we consider a fairly uninformed prior, defined in such a way that the deviation from a flat density at the peak of each Gaussian is of the order of  $\sim 0.0003 \text{ fm}^{-3}$ ,

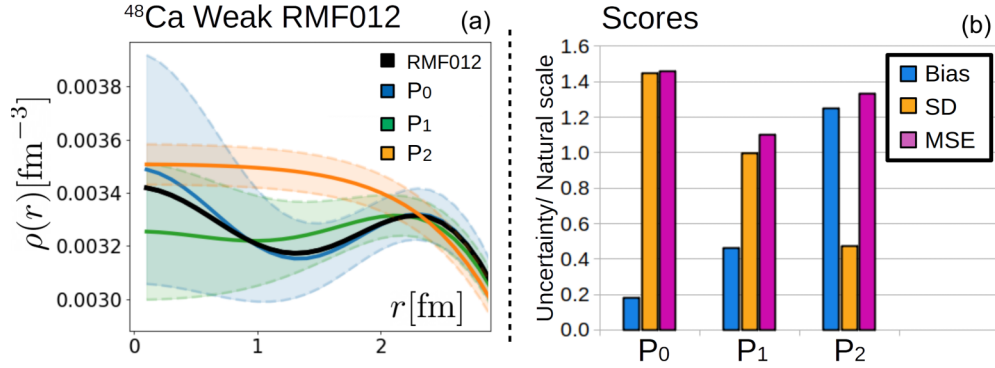


FIG. 10. (a) Reconstructed weak density of  $^{48}\text{Ca}$  for the generator RMF012, the model SF + G and three sets of priors.  $P_0$  (in blue),  $P_1$  (in green), and  $P_2$ . (b) Scores on the interior density in terms of bias, SD, and MSE for the three choices of prior. The numerical values have been divided by the natural scale of  $\Delta\rho_{\text{Ca}} = 0.00015 \text{ fm}^{-3}$ .

or roughly 10% of the average interior density. We call this prior  $P_1$  and is given by the following parameter centers and standard deviations:

$$\omega^0 = [0, 0, 0] \quad \text{and} \quad \sigma = [0.005, 0.027, 0.08]. \quad (39)$$

Finally, we consider an extremely restrictive prior ( $P_2$ ) that forces the value of all three Gaussian amplitudes to zero ( $A_1 = A_2 = A_3 = 0$ ), effectively reproducing the original flat SF model without oscillations. For an example on the incorporation of priors see Appendix F.

The reconstructed weak charge density of  $^{48}\text{Ca}$  for the three choices of priors is displayed in Fig. 10(a), with  $P_0$  (in blue) displaying the largest variance,  $P_1$  (in green), and  $P_2$  (in orange) displaying the smallest variance but the largest bias. The overall performance for each choice is quantified in Fig. 10(b). These results are an interesting example of the bias vs variance tradeoff: the model without prior ( $P_0$ ) reproduces the true (RMF012) curve almost perfectly, but displays a huge error band, whereas the model with the most restrictive prior ( $P_2$ ), or effectively with the fewer number of parameters, has the largest bias but the smallest error bands (variance). Note that  $P_0$  and  $P_2$  have almost the same overall MSE score. Also note that since what can be measured is the form factor as a function of the momentum transfer, the reconstructed spatial density in the interior does not have to be well constrained if there are not enough data. This is the main reason that the orange curve ( $P_2$ ) that fails to reproduce the interior oscillations, also fails to reproduce the average interior density. The model with the  $P_1$  prior provides the best overall MSE score. Indeed, its MSE score is even better than any of the average MSE scores of the models studied in Sec. IV B, for fixed  $q_0$  locations.

We can analyze the behavior of the MSE directly from the transfer function formalism as the prior is modified. To do so, let us focus on the interior density  $\rho(0)$ . Stronger priors constrain more effectively  $\hat{H}^{-1}$ , thereby reducing the impact of the transfer functions  $\mathcal{TF}_j^{\rho(0)}$  of each data point. This effectively reduces the propagation of experimental uncertainty  $\sigma_j$  towards the calculated variance in  $\rho(0)$ . The tradeoff is due to the fact that the inclusion of a strong prior will push away the central value of  $\rho(0)$  from what the central values of the data ( $y_j$ ) suggest, resulting in an increase of the total bias in  $\rho(0)$ .

Such a change can be written to first approximation as

$$\delta\rho(0) = \sum_k^3 \mathcal{TF}_k^{\rho(0)} \tilde{\eta}_k, \quad (40)$$

where the  $\tilde{\eta}_k$  are now defined as the difference between the parameter's value without priors and the new prior centers  $\omega_k^0$ .

Appendix F includes tables with the numerical values of the transfer functions for  $\rho(0)$  and clarifies their meaning in more detail. The important fact is that, as the prior strength increases from  $P_0$  to  $P_2$ , the numerical value of the  $\mathcal{TF}_j^{\rho(0)}$  for each observation  $q_j$  tends to decrease, sometimes by an order of magnitude. This leads to a dramatic decrease in the total variance in the interior density. On the other hand, as the prior strength increases, the prior transfer functions  $\mathcal{TF}_k^{\rho(0)}$  become stronger. This allows each prior center  $\omega_k^0$  to push away the value of  $\rho(0)$  from what the data suggest, effectively increasing the bias.

The example highlights how a well chosen prior could be crucial to reduce uncertainties. However, if the prior strength is excessively high, there is the risk of overlooking new discoveries or making erroneous conclusions. A more in-depth analysis is required to optimize the prior strength and structure for each particular problem in order to effectively reduce the MSE for a set of given truths.

## V. CONCLUSIONS AND FUTURE DIRECTIONS

In this paper, we proposed a novel statistical framework—the transfer function (TF) formalism—and applied it to the extraction of nuclear densities from the associated form factors obtained from electron-scattering data. From this new perspective, we explored model selection and model building, the impact of data locations and errors, the role of priors, and the bias vs variance tradeoff. Given the importance of the PREX and CREX campaigns at JLab in constraining the density dependence of the symmetry energy and in bridging *ab initio* descriptions to density functional theory, we focused our analysis on  $^{48}\text{Ca}$  and  $^{208}\text{Pb}$ . In particular, the two observables of interest explored in this work were the mean square radii and interior densities of both neutron-rich nuclei. We evaluated the performance of seven models in faithfully



reproducing these two observables, from noisy experimental data on the electric form factor and noisy pseudodata generated from a variety of relativistic mean field models for the case of the weak-charge form factor. The performance of the various models was quantified in terms of the mean squared error (MSE) defined as a combined score obtained from incorporating both the bias and the variance.

For both the charge and weak charge densities we showed that, for the adopted noise level assumed in the data, the best performance was obtained with the simpler SF and Helm models that are characterized by a flat interior density. More complex models such as the SF + G or a Fourier-Bessel expansion did not perform as well. Whereas both of these more complex models are able to reproduce the interior shell oscillations of both nuclei, they are hindered by a very high variance, which ultimately results in a high MSE score. In this regard, we suggest that it will be difficult for any of the models used in this paper—at least in their present form—to faithfully reproduce the shell oscillation of both nuclei, particularly in the case of  $^{48}\text{Ca}$  where the oscillation structure is expected to be more pronounced. Indeed, when using the Fourier-Bessel expansion as in [34], we estimated that there is a 15% chance of predicting the wrong oscillating structure in the interior of  $^{48}\text{Ca}$ , namely, peaks become valleys and valleys become peaks.

In the context of experimental design, we illustrated how to use the TF formalism to identify those critical observations that are driving most of the uncertainty in our estimations. The identification of those critical points could help in the design of future experiments to allocate more resources (e.g., beam time) to those critical locations to maximize the information gained from such experiments. Within our assumed experimental error distributions and momentum transfer range, and within the family of RMF generators we used, our final recommendation would be to use the SF or Helm models to extract the weak charge densities and radius of  $^{48}\text{Ca}$  and  $^{208}\text{Pb}$ . The optimized locations  $q_m$  for these models, which are the optimal experimental measurements, are displayed in Tables IX and X of Appendix E.

Finally, we explored the impact of priors on the extracted weak charge density of  $^{48}\text{Ca}$  under the SF + G model. As the influence of the prior increased, so did the bias while the variance was reduced, as expected from the bias vs variance tradeoff.

As mentioned during the derivation of the transfer function formalism, knowledge of first and second derivatives of the model  $F(q_j, \omega)$  is needed to compute the transfer functions in Eq. (24) and perform all the analysis we have showed. If analytic expressions for these derivatives are not available (for example, if we are using a computationally complex model), then other options to estimate derivatives could be used such as performing finite differences [53], automatic differentiation [54], or derivative-free optimization methods like POUNDERS [55].

Going forward, there are several directions that are worth exploring. First, it would be interesting to integrate the TF formalism directly into model building. We believe questions such as *what makes a model better than others?*, could be tackled from the TF perspective. Answering *which* model is

better at extracting data has become a central question in nuclear physics, for example in the context of the proton puzzle. Yan *et al.* investigated this question and provided fundamental insights to the analysis by the PRaD Collaboration. This seminal work—which inspired a great portion of the development of the TF formalism—identified the models optimally suited to extract the proton radius, but did not elaborate on *what* made those model successful. We believe the TF formalism could be used to make significant advances in that direction. As shown in this paper, the TF formalism seems to be ideal to identify the delicate interplay between signal and noise. Understanding the TF distribution of successful models could help not only in identifying but also in creating, some sort of “optimal” model. This technique could be applied beyond density reconstruction from scattering data as implemented in this paper, to more general problems that involve the calibration of model parameters from experimental data.

Another fruitful direction of investigation is the role of priors and hyperparameters. Hyperparameters, such as  $\gamma$  and the Gaussian locations for the SF + G, or the Bessel cut off radius  $R_{\text{cut}}$  and number of coefficients, can drastically impact the performance of a model. In the context of the TF formalism, we could ask questions like: *Given six observations, is it better in terms of an overall MSE score to have five or six adjustable Fourier-Bessel coefficients? How does the answer scale with the number of data points?* We believe it is possible to create a framework using the TF formalism that can tackle this type of questions in a robust and direct manner. This would allow to conduct a more informed search in the hyperparameter space of each model instead of just by trial and error. Our work showed that the incorporation of priors can have a dramatic effect on a model’s performance. After all, priors are essential ingredients of the Bayesian formalism as they encode prior beliefs before additional experimental evidence becomes available. A more in depth study should be carried out to identify how to optimize the hyperparameters that define the priors. To reach robust conclusions, such a research project should include more generator functions from other nuclear model families. We are confident that the TF formalism can guide this optimization procedure as well.

Finally, a third possible application of the TF formalism is related to the recent use in nuclear physics of Bayesian frameworks for combining different competing models to improve over the predictions of single models [59]. Within the context of nuclear densities, using the MSE score should allow us to test the circumstances under which the mixing of several models outperforms the predicting power of a single model. It would be interesting to explore in the future the generalization of the TF formalism to Bayesian model mixing.

## ACKNOWLEDGMENTS

We are grateful to Edgard Bonilla for his help and critical observations during this project. We thank Prof. Antonio Linero for his guidance and key support. We thank Diogenes Figueroa for many useful conversations. We thank Prof. Douglas Higinbotham for introducing us to the bias vs variance analysis and for his encouragement at the beginning of the project. Finally, we thank Ana Posada for a careful read of the

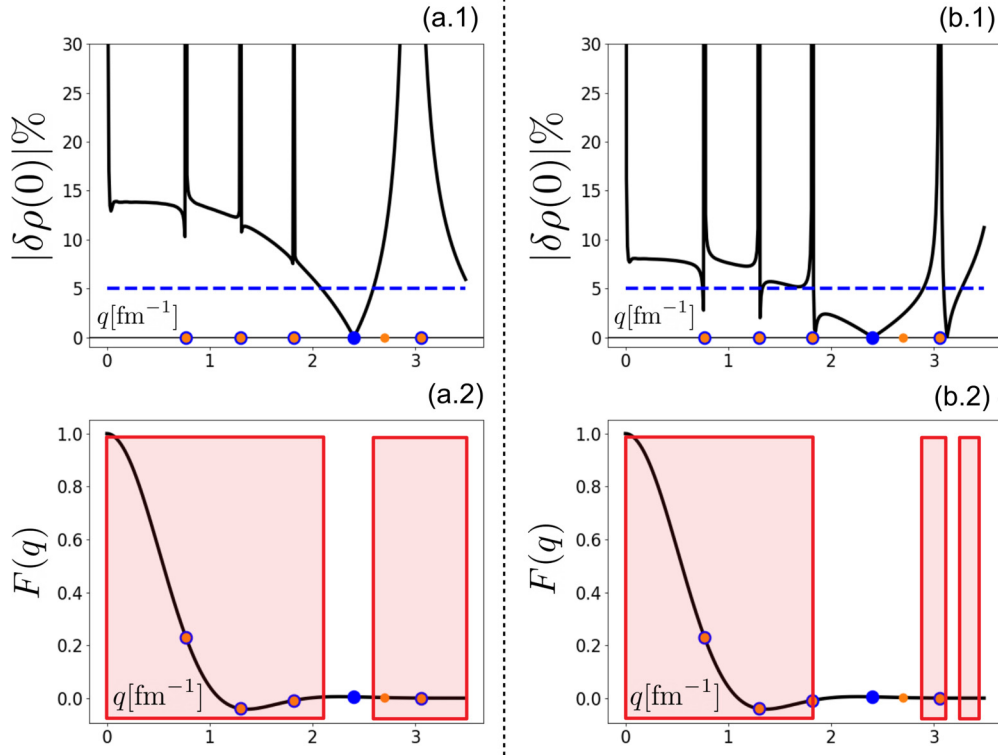


FIG. 11. Bias increase in  $\rho(0)$ , when using the SF + G model, as a function of the location of the fourth measurement  $q_4$  for  $\gamma = 0.7$  fm (a) and  $\gamma = 1.4$  fm (b). (a.1) and (b.1) show in black the calculated  $\delta\rho(0)$  (as an absolute percentage) using Eq. (C1). The blue and orange points correspond to the two original data sets while the blue dashed line represents the threshold  $|\delta\rho(0)| = 5\%$ . (a.2) and (b.2) show the true form factor in black as well as the measurement locations for both sets. The red rectangles encompass the regions where, if  $q_4$  is located,  $|\delta\rho(0)| > 5\%$ .

manuscript. This material is based upon work supported by the U.S. Department of Energy Office of Science, Office of Nuclear Physics under Grant. No. DE-FG02-92ER40750

## APPENDIX A: MATHEMATICAL PROOFS ON THE TF FORMALISM

### 1. Transfer functions structure

This subsection presents a formal proof of the structure of the transfer functions [Eq. (24)], namely that the first-order change coefficients on the parameters  $\omega$  due to a perturbation

TABLE III.  $^{48}\text{Ca}$  electric form factor momentum transfer locations  $q_j$  and transfer functions  $\mathcal{TF}_j$  absolute values of the density at  $r = 0$  fm and radius for two data sets: original  $\mathbf{q}_0$  and optimized  $\mathbf{q}_m$ . Both transfer functions have been normalized by their respective natural scales defined in Table I.

$q_j$ (fm $^{-1}$ )	$[q_0]$	0.90	1.35	1.80	2.24	2.69
	$[q_m]$	0.65	1.26	1.73	2.12	2.55
$ \mathcal{TF}_j^{\rho(0)}\sigma_j $	$[q_0]$	0.84	1.02	1.44	1.07	0.65
	$[q_m]$	0.65	1.18	1.40	0.90	0.72
$ \mathcal{TF}_j^R\sigma_j $	$[q_0]$	0.88	0.80	0.58	0.32	0.12
	$[q_m]$	0.82	0.34	0.26	0.15	0.03

on observation  $y_j$  are

$$\mathcal{TF}_j^\omega = \mathcal{H}^{-1}\nabla F_j\sigma_j^{-2}, \quad (\text{A1})$$

where  $\mathcal{H}^{-1}$  is the inverse of the Hessian matrix of  $\chi^2/2$  defined in (11) and  $\nabla F_j$  is the gradient with respect to the parameters  $\omega$  of the function  $F$  being fit evaluated at observation  $y_j$ .

Let us assume that we are at the minimum  $\omega_0$  of the unperturbed  $\chi^2/2$ . At this point, the condition of a minimum implies that the first derivative of  $\chi^2/2$  with respect to all  $\omega_k$

TABLE IV.  $^{208}\text{Pb}$  electric form factor momentum transfer locations  $q_j$  and transfer functions  $\mathcal{TF}_j$  absolute values of the density at  $r = 0$  fm and radius for two data sets: original  $\mathbf{q}_0$  and optimized  $\mathbf{q}_m$ . Both transfer functions have been normalized by their respective natural scales defined in Table I.

$q_j$ (fm $^{-1}$ )	$[q_0]$	0.50	0.80
	$[q_m]$	0.41	0.84
$ \mathcal{TF}_j^{\rho(0)}\sigma_j $	$[q_0]$	0.14	0.18
	$[q_m]$	0.17	0.25
$ \mathcal{TF}_j^R\sigma_j $	$[q_0]$	0.59	0.48
	$[q_m]$	0.54	0.31

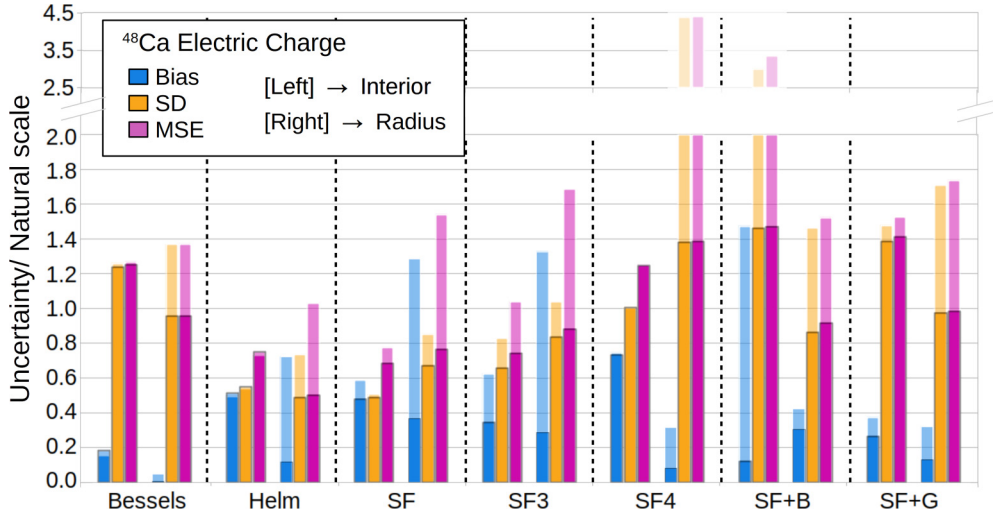


FIG. 12. Seven models comparison for recovering the interior density and mean charge radius from the charge form factor data on  $^{48}\text{Ca}$ . The bias, SD, and MSE for each model are shown in their respective three left columns for the interior density and in the three right columns for the radius. All quantities have been divided by their natural scales:  $\Delta\rho_{\text{Ca}} = 0.00015 \text{ fm}^{-3}$  and  $\Delta R_{\text{Ca}} = 0.04 \text{ fm}$ . The solid columns represent the optimal locations  $q_m$ , while the borderless ones represent the starting  $q_0$ .

( $K$  in total) should be zero:

$$\frac{1}{2} \frac{\partial \chi^2}{\partial \omega_k} \Big|_{(\omega_0, y_0)} \equiv G_k(\omega, \mathbf{y})|_{(\omega_0, y_0)} = 0, \quad (\text{A2})$$

where we use the notation  $\mathbf{y} \equiv (y_1, \dots, y_n)$  to refer to the group of all  $J$  observations, and the subscript "0" to refer to the unperturbed variables. We call  $G_k(\omega, \mathbf{y})$  the first derivative of  $\chi^2/2$  with respect to parameter  $\omega_k$ . The  $G_k$  are the following functions of both the parameters and the observations:

$$G_k(\omega, \mathbf{y}) = \sum_j \frac{(F_j - y_j)}{\sigma_j^2} \frac{\partial F_j}{\partial \omega_k}. \quad (\text{A3})$$

Now, if we perturb observation  $y_j$  by a small amount  $\delta y_j$  the minimum of  $\chi^2/2$  will move accordingly. If we want to preserve all  $K$  equations (A2) (there is one equation for every parameter), then the values of all  $\omega_k$  should change a small amount as well  $\delta \omega_k$  to compensate. Quantitatively, this means (to first order)

$$\delta y_j \frac{\partial G_k}{\partial y_j} = - \sum_i^K \delta \omega_i \frac{\partial G_k}{\partial \omega_i}. \quad (\text{A4})$$

We can arrange all  $K$  equations into a matrix form:

$$\frac{\partial G_1}{\partial y_j} \delta y_j = - \left( \frac{\partial G_1}{\partial \omega_1} \delta \omega_1 + \frac{\partial G_1}{\partial \omega_2} \delta \omega_2 + \dots + \frac{\partial G_1}{\partial \omega_K} \delta \omega_K \right),$$

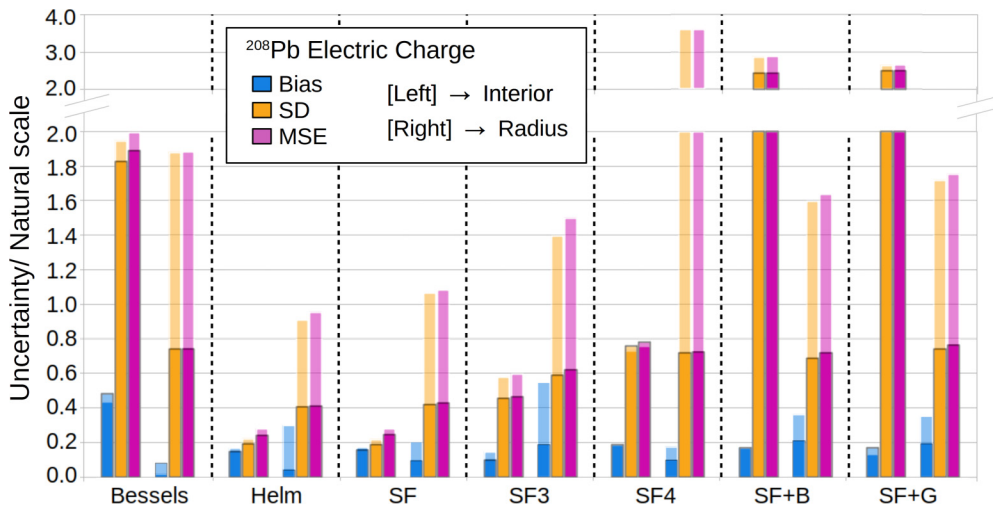


FIG. 13. Seven models comparison for recovering the interior density and mean charge radius from the charge form factor data on  $^{208}\text{Pb}$ . The bias, SD, and MSE for each model are shown in their respective three left columns for the interior density and in the three right columns for the radius. All quantities have been divided by their natural scales:  $\Delta\rho_{\text{Pb}} = 0.00008 \text{ fm}^{-3}$  and  $\Delta R_{\text{Pb}} = 0.06 \text{ fm}$ . The solid columns represent the optimal locations  $q_m$ , while the borderless ones represent the starting  $q_0$ .

TABLE V. Optimal locations  $q_m$  for each model when optimizing the figure of merit in Eq. (37). These  $q_j$  values correspond to the solid bars results in Fig. 12 for the charge density of  $^{48}\text{Ca}$ . All values are in units of  $\text{fm}^{-1}$ .

$q_j$ ( $\text{fm}^{-1}$ )	$q_1$	$q_2$	$q_3$	$q_4$	$q_5$
Bessels	0.65	1.26	1.73	2.12	2.55
Helm	0.68	0.71	1.39	2.09	2.10
SF	0.44	0.62	1.24	2.09	2.79
SF3	0.56	0.90	1.20	1.84	2.50
SF4	0.50	1.10	1.92	2.12	2.59
SF+B	0.67	1.31	1.80	2.43	3.20
SF+G	0.63	1.25	1.78	2.17	2.80

$$\begin{aligned} \frac{\partial G_2}{\partial y_j} \delta y_j &= - \left( \frac{\partial G_2}{\partial \omega_1} \delta \omega_1 + \frac{\partial G_2}{\partial \omega_2} \delta \omega_2 + \dots + \frac{\partial G_2}{\partial \omega_K} \delta \omega_K \right), \\ &\vdots \\ \frac{\partial G_K}{\partial y_j} \delta y_j &= - \left( \frac{\partial G_K}{\partial \omega_1} \delta \omega_1 + \frac{\partial G_K}{\partial \omega_2} \delta \omega_2 + \dots + \frac{\partial G_K}{\partial \omega_K} \delta \omega_K \right), \end{aligned}$$

where, since the  $G_k$  were already first derivatives of  $\chi^2/2$ , we can recognize the Hessian matrix  $\frac{\partial G_i}{\partial \omega_k} = \mathcal{H}_{i,k}$ . We also recognize  $\frac{\partial G_k}{\partial y_j} = -\frac{\partial F_j}{\partial \omega_k} \sigma_j^{-2}$ . We therefore have

$$-\nabla F_j \sigma_j^{-2} \delta y_j = -\mathcal{H} \delta \omega, \quad [\mathcal{H}^{-1} \nabla F_j \sigma_j^{-2}] \delta y_j = \delta \omega,$$

where, since the perturbation  $\delta y_j$  can be made arbitrary small, we must conclude that the quantity in brackets is what were looking for. These are the linear coefficients connecting a small change in  $y_j$  with the small change in every parameter  $\omega_k$ , proving Eq. (A1).

This was a constructive proof of the transfer function structure. Another approach would be to use Newton's minimization method to find the location of the new minimum of  $\chi^2/2$  once we make the perturbation  $y_j \rightarrow y_j + \delta y_j$ . Newton's method involves the same ingredients shown in Eq. (A1), namely the gradient and Hessian of the scalar objective function ( $\chi^2/2$ ). The reason why the gradient does not involve all observation but just the  $j$ th component is because in Eq. (A3) only the term multiplying  $\delta y_j$  survives. Everything else gets canceled by the definition of the minimum.

Finally, a third approach results by invoking the implicit function theorem on the minimum conditions of  $\chi^2/2$ . Those are  $K$  equations and each one is a function of the  $J$  observations and  $K$  parameters, where  $K \leq J$ . Under these conditions, there exists a map in a vicinity around  $(\omega_0, y_0)$  from the bigger set of variables (the observations) to the smaller set (the parameters). The coefficients of this linear map (first derivatives) are precisely given by (A1) (see [60, pp. 41–42]).

## 2. Comparing the transfer function variance with the standard approach's variance

This subsection presents a formal proof of the statement discussed at the end of Sec. III A. When the Hessian matrix

TABLE VI. Optimal locations  $q_m$  for each model when optimizing the figure of merit in Eq. (37). These  $q_j$  values correspond to the solid bars results in Fig. 13 for the charge density of  $^{208}\text{Pb}$ . All values are in units of  $\text{fm}^{-1}$ .

$q_j$ ( $\text{fm}^{-1}$ )	$q_1$	$q_2$	$q_3$	$q_4$	$q_5$
Bessels	0.37	0.86	1.20	1.60	1.80
Helm	0.41	0.43	0.83	0.88	1.31
SF	0.42	0.43	0.83	0.87	0.88
SF3	0.23	0.41	0.88	1.07	1.16
SF4	0.25	0.36	0.70	1.22	1.51
SF+B	0.38	0.87	1.27	1.75	1.97
SF+G	0.36	0.90	1.29	1.68	1.91

(11) only involves the linear part of the model,

$$\mathcal{H}_{i,k} = \sum_{j=1}^J \frac{1}{\sigma_j^2} \left[ \left( \frac{\partial F(q_j, \omega)}{\partial \omega_i} \right) \left( \frac{\partial F(q_j, \omega)}{\partial \omega_k} \right) \right], \quad (\text{A5})$$

then the variance calculated using the standard approach (12) and the variance calculated using the transfer function formalism (25) are identical, namely

$$\Delta m^2 = \nabla m \mathcal{H}^{-1} \nabla m, \quad (\text{A6})$$

$$\Delta m^2 = \sum_j^J [\nabla m \mathcal{H}^{-1} \nabla F(q_j, \omega) \sigma_j^{-2}]^2 \sigma_j^2. \quad (\text{A7})$$

Let us recall that we have  $J$  observations and  $K$  parameters with  $K \leq J$ . To prove this statement, we first observe that the matrix  $\mathcal{H}$  is built by the sum of  $J$  tensor products between the gradients  $\nabla F_j \sigma_j^{-1}$  with themselves:

$$\mathcal{H} = \sum_j^J [\nabla F_j \sigma_j^{-1}] \otimes [\nabla F_j \sigma_j^{-1}]. \quad (\text{A8})$$

Therefore,  $\mathcal{H}$  can be written as the product of a matrix  $\mathcal{F}$  and its transpose as:

$$\mathcal{H} = \mathcal{F} \mathcal{F}^T, \quad (\text{A9})$$

where  $\mathcal{F}$  is a  $K \times J$  matrix which columns are the gradients of  $F$ :

$$\mathcal{F} = (\nabla F_1 \sigma_1^{-1} \quad \nabla F_2 \sigma_2^{-1} \quad \dots \quad \nabla F_J \sigma_J^{-1}). \quad (\text{A10})$$

We decompose  $\mathcal{F}$  and  $\mathcal{F}^T$  into their  $QR$  decomposition (see Theorem 2 in [61]):

$$\mathcal{F} = R^T Q^T, \quad \mathcal{F}^T = QR, \quad (\text{A11})$$

where  $Q$  and  $R$  are a  $J \times K$  and a  $K \times K$  matrix, respectively. The  $Q$  matrix is an orthonormal matrix meaning:  $Q^T Q = I_{K,K}$ , where  $I_{K,K}$  is the  $K \times K$  identity matrix. Under these conditions, we have that

$$\begin{aligned} \nabla m \mathcal{H}^{-1} \nabla m &= \nabla m [R^T Q^T QR]^{-1} \nabla m \\ &= \nabla m R^{-1} (R^T)^{-1} \nabla m = \|(R^T)^{-1} \nabla m\|^2. \end{aligned} \quad (\text{A12})$$

Therefore, under the standard approach Eq. (A6) is calculating the norm squared of the vector  $(R^T)^{-1} \nabla m$  in  $\mathbb{R}^K$ . Let



TABLE VII. Numerical values for the data displayed in Fig. 12: model comparison for the electric charge density of  $^{48}\text{Ca}$ .

	$q_0$						$q_m$					
	Interior			Radius			Interior			Radius		
	Bias	SD	MSE	Bias	SD	MSE	Bias	SD	MSE	Bias	SD	MSE
Bessels	0.15	1.26	1.27	0.05	1.37	1.37	0.19	1.24	1.25	0.00	0.96	0.96
Helm	0.50	0.54	0.73	0.72	0.73	1.03	0.51	0.55	0.75	0.12	0.49	0.50
SF	0.59	0.51	0.78	1.28	0.85	1.54	0.48	0.49	0.68	0.37	0.67	0.77
SF3	0.62	0.83	1.04	1.33	1.04	1.68	0.34	0.66	0.74	0.29	0.84	0.88
SF4	0.75	1.00	1.25	0.31	4.38	4.40	0.73	1.01	1.25	0.08	1.38	1.38
SF+B	1.47	3.00	3.35	0.42	1.46	1.52	0.12	1.46	1.47	0.30	0.86	0.92
SF+G	0.37	1.48	1.52	0.32	1.71	1.74	0.26	1.39	1.41	0.13	0.97	0.98

us now work with Eq. (A7) and obtain a similar structure but in  $\mathbb{R}^J$ .

Given how  $\mathcal{F}$  is defined, we realize that

$$\mathcal{F} \cdot e_j = \nabla F_j \sigma_j^{-1}, \quad (\text{A13})$$

where  $e_j = (0, 0, \dots, 1, \dots, 0)$  is the  $j$ th vector in the canonical base of  $\mathbb{R}^J$  with all entries as 0 except for an entry of 1 in position  $j$ .

Therefore, in Eq. (A7), we can replace  $\nabla F_j \sigma_j^{-1}$  by  $\mathcal{F} \cdot e_j$  and obtain

$$\begin{aligned} & \sum_j^J [\nabla m \mathcal{H}^{-1} \mathcal{F} \cdot e_j]^2 \\ &= \sum_j^J [\nabla m R^{-1} (R^T)^{-1} R^T Q^T \cdot e_j]^2 \\ &= \sum_j^J [(\nabla m R^{-1} Q^T) \cdot e_j]^2 = \|\mathcal{Q} (R^T)^{-1} \nabla m\|^2, \quad (\text{A14}) \end{aligned}$$

where the last term is the norm of the vector  $\mathcal{Q} (R^T)^{-1} \nabla m$  calculated in  $\mathbb{R}^J$ . But, since the matrix  $\mathcal{Q}$  is orthonormal, it preserves norms when taking vectors from  $\mathbb{R}^K$  to  $\mathbb{R}^J$ . We must conclude that this expression is also the norm of the vector  $(R^T)^{-1} \nabla m$  in  $\mathbb{R}^K$ , which proves that Eqs. (A6) and (A7) are identical.

TABLE VIII. Numerical values for the data displayed in Fig. 13: model comparison for the electric charge density of  $^{208}\text{Pb}$ .

	$q_0$						$q_m$					
	Interior			Radius			Interior			Radius		
	Bias	SD	MSE	Bias	SD	MSE	Bias	SD	MSE	Bias	SD	MSE
Bessels	0.44	1.95	2.00	0.02	1.88	1.88	0.48	1.82	1.89	0.08	0.74	0.74
Helm	0.17	0.22	0.28	0.30	0.91	0.96	0.14	0.19	0.24	0.04	0.41	0.41
SF	0.17	0.22	0.28	0.20	1.06	1.08	0.16	0.19	0.24	0.09	0.42	0.43
SF3	0.15	0.58	0.60	0.55	1.40	1.50	0.10	0.45	0.46	0.19	0.59	0.62
SF4	0.19	0.73	0.76	0.17	3.59	3.59	0.18	0.76	0.78	0.09	0.72	0.72
SF+B	0.17	2.86	2.86	0.36	1.60	1.64	0.17	2.43	2.44	0.21	0.69	0.72
SF+G	0.13	2.64	2.64	0.35	1.72	1.75	0.17	2.50	2.51	0.19	0.74	0.76

## APPENDIX B: MODEL DESCRIPTIONS

This section presents detailed information about the seven models we studied in this work. For each model, we provide its analytic form (if available) as a function of its parameters in both coordinate space  $\rho(r)$  and momentum space  $F(q)$ ; a normalization condition (if any), that restricts the parameters; the mean squared radius  $R$  as a function of the parameters when available; and the values of the model's hyperparameters, if any, used in this work.

*Fourier-Bessel.* Under this formalism [46], the density is written as

$$\rho_{FB}(r) = H(R_{\text{cut}} - r) \sum_{\nu=1}^N a_{\nu} j_0(q_{\nu} r), \quad (\text{B1})$$

where  $j_0$  denotes the zeroth-order spherical Bessel function of the first kind,  $a_{\nu}$  are the free parameters,  $q_{\nu} = \nu\pi/R_{\text{cut}}$ , and  $R_{\text{cut}}$  is such that  $\rho(r) = 0$  for  $r > R_{\text{cut}}$ . This last condition is enforced by the Heviside theta function  $H$ .

The form factor can be expressed analytically as

$$F_{FB}(q) = \sum_{\nu=1}^N a_{\nu} G_{\nu}(q), \quad \text{where} \quad (\text{B2})$$

$$G_{\nu}(q) \equiv 4\pi (-1)^{\nu} j_0(q R_{\text{cut}}) \frac{R_{\text{cut}}}{q^2 - q_{\nu}^2}. \quad (\text{B3})$$

TABLE IX. Optimal locations  $q_m$  for each model when optimizing the figure of merit in Eq. (37). These  $q_j$  values correspond to the solid columns results in Fig. 7 for the five generated weak charge densities of  $^{48}\text{Ca}$ . All values are in units of  $\text{fm}^{-1}$ .

$q_j$ ( $\text{fm}^{-1}$ )	$q_1$	$q_2$	$q_3$	$q_4$	$q_5$
Bessels	0.71	1.32	1.81	2.28	2.65
Helm	0.51	0.63	0.77	1.44	2.93
SF	0.74	0.92	1.36	1.67	2.50
SF3	0.54	1.32	1.69	1.87	2.42
SF4	0.44	1.41	2.10	2.82	2.88
SF+B	0.61	1.26	1.78	2.32	3.03
SF+G	0.57	1.18	1.69	2.06	2.68

The normalization condition translates to

$$F_{FB}(0) = \sum_{\nu=1}^N (-1)^{\nu+1} \frac{4\pi R_{\text{cut}}}{q_\nu^2} a_\nu = 1. \quad (\text{B4})$$

The mean square radius is obtained as

$$R^2 = 4\pi \sum_{\nu} a_\nu \frac{(-1)^\nu R_{\text{cut}}^5 (6 - \nu^2 \pi^2)}{\nu^4 \pi^4}. \quad (\text{B5})$$

In this work, we use  $R_{\text{cut}} = 7$  fm when analyzing  $^{48}\text{Ca}$ , and  $R_{\text{cut}} = 10$  fm when analyzing  $^{208}\text{Pb}$ . We use a total of five adjustable  $a_\nu$  for both nuclei which translates to six Bessels ( $N = 6$ ) due to the normalization condition.

*Helm density.* The Helm density [47] is

$$\begin{aligned} \rho_H(r) = & \frac{1}{2} \rho_{0H} \left[ \text{erf}\left(\frac{r+R_0}{\sqrt{2}\sigma}\right) - \text{erf}\left(\frac{r-R_0}{\sqrt{2}\sigma}\right) \right] \\ & + \frac{1}{\sqrt{2\pi}} \left(\frac{\sigma}{r}\right) \rho_{0H} \\ & \times \left\{ \exp\left[-\frac{(r+R_0)^2}{2\sigma^2}\right] - \exp\left[-\frac{(r-R_0)^2}{2\sigma^2}\right] \right\}, \end{aligned} \quad (\text{B6})$$

where

$$\rho_{0H} \equiv \frac{3}{4\pi R_0^3}, \quad (\text{B7})$$

where  $R_0$  and  $\sigma$  are the adjustable parameters, and  $\text{erf}(x)$  is the error function:

$$\text{erf}(x) \equiv \frac{2}{\sqrt{\pi}} \int_0^x e^{-z^2} dz. \quad (\text{B8})$$

The Helm form factor is built by the product of two form factors: a uniform ‘‘box’’ density inspired by the fact that nuclear matter in nuclei saturates, and a Gaussian falloff which takes into account the finite size of the nucleons:

$$F_H(q) = 3 \frac{j_1(qR_0)}{qR_0} e^{-q^2 \sigma^2 / 2}, \quad (\text{B9})$$

where  $j_1$  is the spherical Bessel function of first order,  $j_1(x) = \frac{\sin(x)}{x^2} - \frac{\cos(x)}{x}$ .

TABLE X. Optimal locations  $q_m$  for each model when optimizing the figure of merit in Eq. (37). These  $q_j$  values correspond to the solid columns results in Fig. 8 for the five generated weak charge densities of  $^{208}\text{Pb}$ . All values are in units of  $\text{fm}^{-1}$ .

$q_j$ ( $\text{fm}^{-1}$ )	$q_1$	$q_2$	$q_3$	$q_4$	$q_5$
Bessels	0.37	0.84	1.21	1.61	1.80
Helm	0.37	0.40	0.84	1.23	1.24
SF	0.37	0.40	0.77	0.84	1.36
SF3	0.38	0.43	0.87	1.03	1.17
SF4	0.32	0.52	0.68	0.81	1.09
SF+B	0.37	0.84	1.25	1.70	1.94
SF+G	0.32	0.82	1.23	1.65	1.85

The radius is given by

$$R^2 = \frac{3}{5} R_0^2 + 3\sigma^2. \quad (\text{B10})$$

*Symmetrized Fermi function.* The symmetrized Fermi function [48] is constructed as  $f_{SF}(r) \equiv f_{SF}(r) + f_{SF}(-r) - 1$ , where  $f_{SF}$  is the traditional Fermi function [62]. Its density and form factor are expressed as

$$\rho_{SF}(r) = \rho_{0SF} \frac{\sinh(c/a)}{\cosh(r/a) + \cosh(c/a)}, \quad (\text{B11})$$

$$\begin{aligned} F_{SF}(q) = & \frac{3}{qc[(qc)^2 + (\pi qa)^2]} \left[ \frac{\pi qa}{\sinh(\pi qa)} \right] \\ & \times \left[ \frac{\pi qa}{\tanh(\pi qa)} \sin(qc) - qc \cos(qc) \right], \end{aligned} \quad (\text{B12})$$

where the normalization constant is

$$\rho_{0SF} = \frac{3}{4\pi c(c^2 + \pi^2 a^2)}, \quad (\text{B13})$$

where the parameters  $a$  and  $c$  represent the surface diffuseness and half-density radius, respectively. The radius  $R$  is given by

$$R^2 = \frac{3}{5} c^2 + \frac{7}{5} (\pi a)^2. \quad (\text{B14})$$

*Symmetrized Fermi function of three and four parameters.* Based on the two parameter symmetrized Fermi density, we can build three-parameter and four-parameter densities as

$$\rho_{SF3}(r) = \rho_{0SF3} (1 + wr^2) \frac{\sinh(c/a)}{\cosh(r/a) + \cosh(c/a)}, \quad (\text{B15})$$

$$\rho_{SF4}(r) = \rho_{0SF4} (1 + wr^2 + ur^4) \frac{\sinh(c/a)}{\cosh(r/a) + \cosh(c/a)}. \quad (\text{B16})$$

The parameters  $w$  and  $u$  are introduced to add flexibility near  $r = 0$  to the densities and to reproduce oscillations. Due to the size of their expressions, the normalization constants  $\rho_{0SF3}$  and  $\rho_{0SF4}$ , the form factors and the radius equations are not included here but can be calculated analytically.

*Symmetrized Fermi function plus Bessels (SF + B) and plus Gaussians (SF + G).*

By construction the two parameter symmetrized Fermi function  $\rho_{SF}(r)$  exhibits a flat behavior in the interior, so it

TABLE XI. Numerical values for the data displayed in Fig. 7: model comparison for the five generated weak charge densities of  $^{48}\text{Ca}$ .

	$q_0$						$q_m$					
	Interior			Radius			Interior			Radius		
	Bias	SD	MSE	Bias	SD	MSE	Bias	SD	MSE	Bias	SD	MSE
Bessels	0.09	1.26	1.26	0.06	1.31	1.31	0.08	1.30	1.31	0.30	0.87	0.92
Helm	1.22	0.50	1.32	1.08	0.87	1.39	0.70	0.56	0.90	0.03	0.45	0.45
SF	1.09	0.47	1.19	0.87	1.04	1.37	0.62	0.41	0.74	0.24	0.76	0.81
SF3	1.20	0.86	1.47	2.14	1.18	2.44	0.60	0.93	1.11	1.04	0.94	1.40
SF4	0.94	1.39	1.68	2.71	1.70	3.22	0.63	1.62	1.74	1.23	1.10	1.65
SF+B	0.70	2.37	2.47	1.04	1.83	2.10	0.17	1.47	1.48	0.90	0.95	1.31
SF+G	0.19	1.44	1.45	1.02	2.25	2.47	0.22	1.35	1.36	0.82	1.11	1.38

cannot describe shell oscillations (small bumps and valleys around the saturation density background). As a different approach to adding parameters as in Eqs. (B15) and (B16), we propose to use the following hybrid models.

SF + B:

$$\rho_{SF+B}(r) = \left(1 - \sum_{\nu=1}^{N_B} (-1)^{\nu+1} \frac{4\pi R_{\text{cut}}}{q_\nu^2} a_\nu\right) \rho_{SF}(r) + H(R_{\text{cut}} - r) \sum_{\nu=1}^{N_B} a_\nu j_0(q_\nu r),$$

where  $j_0$  is the spherical Bessel function of the first kind and  $H$  the Heaviside theta function. We set  $R_{\text{cut}} = 3.3$  fm for  $^{48}\text{Ca}$  and  $R_{\text{cut}} = 5$  fm for  $^{208}\text{Pb}$ .  $R_{\text{cut}}$  in this context specifies the region where we believe the interior oscillations are important. We use a total of three adjustable coefficients  $a_\nu$  which in this case correspond to three Bessels ( $N_B = 3$ ) since the normalization is enforced automatically. We have therefore a total of five parameters including  $a$  and  $c$  from the SF model.

SF + G:

$$\rho_{SF+G}(r) = \left(1 - \sum_{i=1}^{N_G} A_i\right) \rho_{SF}(r) + \frac{1}{2\pi^{3/2}\gamma^3} \sum_{i=1}^{N_G} A_i g(r, R_i), \quad (\text{B17})$$

where  $g(r, R_i)$  is defined as

$$g(r, R_i) = \frac{1}{1 + 2R_i^2/\gamma^2} (e^{-(r-R_i)^2/\gamma^2} + e^{-(r+R_i)^2/\gamma^2}). \quad (\text{B18})$$

The amplitudes of the Gaussians  $A_i$  act as adjustable parameters. In our case they are not restricted to be positive in contrast with [49]. We use a total of three Gaussians ( $N_G = 3$ ), giving us five adjustable parameters including  $a$  and  $c$  from the SF model. The hyperparameter  $\gamma$ , which represents the common width of the Gaussians, is set to  $\gamma = 1.4$  fm for both nuclei (see Sec. III E for a discussion of this value). The center of each Gaussian is denoted by  $R_i$  and we chose the following values:

$$\begin{aligned} \mathbf{R} &= [0, 1.3, 2.6] \text{ fm for } ^{48}\text{Ca}, \\ \mathbf{R} &= [0, 2, 4] \text{ fm for } ^{208}\text{Pb}. \end{aligned} \quad (\text{B19})$$

The main idea behind these expansions is that the principal behavior of the nuclear density is modeled by the symmetrized Fermi density, while the fine details are modeled by either a sum of Bessels or Gaussians. Both  $R_{\text{cut}}$  and the  $R_i$  are chosen in such a way that they cover the region where we expect the oscillations around a flat density to be important.

The Bessel approach has the disadvantage that the total density  $\rho_{SF+B}(r)$  will present a discontinuity at  $r = R_{\text{cut}}$ . This ‘‘kink,’’ while nonphysical, might not preclude the entire model from describing nuclear densities.

TABLE XII. Numerical values for the data displayed in Fig. 8: model comparison for the five generated weak charge densities of  $^{208}\text{Pb}$ .

	$q_0$						$q_m$					
	Interior			Radius			Interior			Radius		
	Bias	SD	MSE	Bias	SD	MSE	Bias	SD	MSE	Bias	SD	MSE
Bessels	0.20	1.95	1.96	0.40	1.82	1.86	0.36	1.83	1.87	0.05	0.70	0.71
Helm	0.10	0.23	0.25	0.59	1.11	1.26	0.11	0.21	0.24	0.01	0.45	0.45
SF	0.09	0.22	0.24	0.28	1.40	1.43	0.09	0.19	0.21	0.25	0.47	0.53
SF3	0.15	0.59	0.61	0.51	1.99	2.06	0.10	0.40	0.41	0.29	0.62	0.69
SF4	0.14	0.77	0.78	0.61	4.08	4.13	0.14	0.60	0.62	0.21	0.83	0.86
SF+B	0.13	2.70	2.70	0.46	2.19	2.24	0.15	2.40	2.41	0.31	0.78	0.84
SF+G	0.19	2.56	2.57	0.47	2.50	2.54	0.26	2.48	2.49	0.27	0.82	0.87

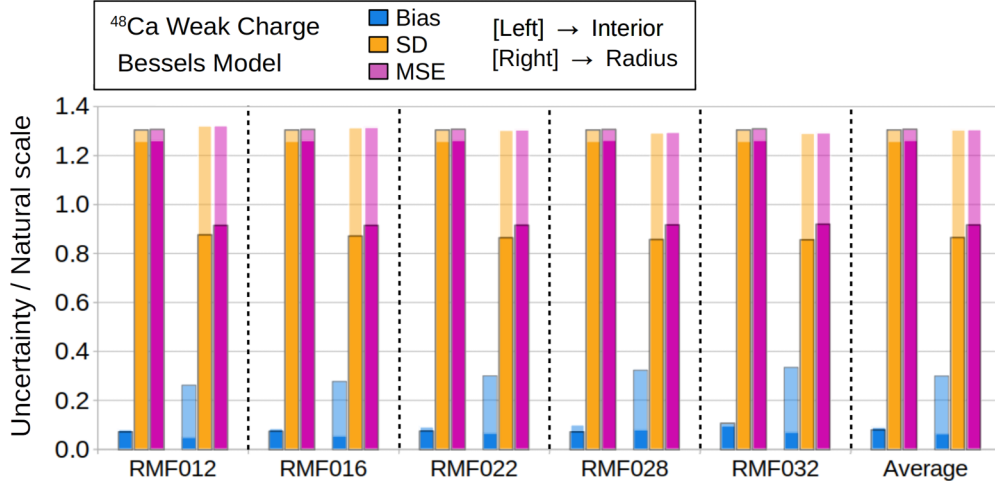


FIG. 14. Bessel model bias, standard deviation (SD), and MSE across the five different generators for the weak charge of  $^{48}\text{Ca}$ . The last group of columns shows the square average of the five generators. Within each generator, the first three columns refer to the interior density while the second three refer to the radius. All quantities have been divided by their natural scales defined in Table I.

The form factors for both densities can be expressed in analytic form:

$$F_{SF B}(q) = \left(1 - \sum_{\nu=1}^{N_B} (-1)^{\nu+1} \frac{4\pi R_{\text{cut}}}{q_\nu^2} a_\nu\right) F_{SF}(q) + \sum_{\nu=1}^{N_B} a_\nu G_\nu(q), \quad (\text{B20})$$

$$F_{SF G}(q) = \left(1 - \sum_{i=1}^{N_G} A_i\right) F_{SF}(q) + \sum_{i=1}^{N_G} A_i \tilde{g}(q, R_i), \quad (\text{B21})$$

where  $G_\nu(q)$  is defined in Eq. (B3), and  $\tilde{g}(q, R_i)$  is defined as

$$\tilde{g}(q, R_i) = e^{-q^2 \gamma^2 / 4} \frac{1}{1 + 2R_i^2 / \gamma^2} \left( \cos(qR_i) + \frac{2R_i^2}{\gamma^2} j_0(qR_i) \right). \quad (\text{B22})$$

The SF + B radius is given by

$$R^2 = \left(1 - \sum_{\nu=1}^{N_B} (-1)^{\nu+1} \frac{4\pi R_{\text{cut}}}{q_\nu^2} a_\nu\right) \left( \frac{3}{5} c^2 + \frac{7}{5} (\pi a)^2 \right) + 4\pi \sum_{\nu=1}^{N_B} a_\nu \frac{(-1)^\nu R_{\text{cut}}^5 (6 - \nu^2 \pi^2)}{\nu^4 \pi^4}. \quad (\text{B23})$$

The SF + G radius is given by

$$R^2 = \left(1 - \sum_{i=1}^{N_G} A_i\right) \left( \frac{3}{5} c^2 + \frac{7}{5} (\pi a)^2 \right) + \sum_{i=1}^{N_G} A_i \frac{3\gamma^4 + 4R_i^4 + 12\gamma^2 R_i^2}{2(\gamma^2 + 2R_i^2)}.$$

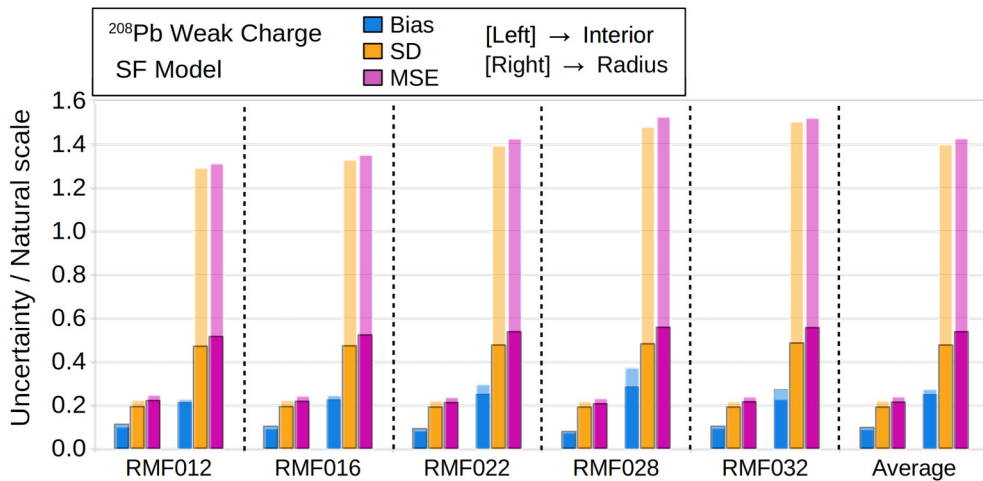


FIG. 15. SF model bias, standard deviation (SD), and MSE across the five different generators for the weak charge of  $^{208}\text{Pb}$ . The last group of columns shows the square average of the five generators. Within each generator, the first three columns refer to the interior density while the second three refer to the radius. All quantities have been divided by their natural scales defined in Table I.



TABLE XIII. Absolute value transfer function values for the density at  $r = 0$  fm ( $|\mathcal{T}\mathcal{F}_j^{\rho(0)}\sigma_j|/\Delta\rho_{Ca}$ ) for the five locations  $\mathbf{q}_0 = [0.9, 1.35, 1.8, 2.24, 2.69]$  fm $^{-1}$  in units of the natural scale  $\Delta\rho_{Ca}$ . Each  $\mathcal{T}\mathcal{F}$  has been multiplied by its respective  $\sigma_j$  to somehow represent a fraction of the total standard deviation SD. The three different prior options are explored for the SF+G model.

	$q_1$	$q_2$	$q_3$	$q_4$	$q_5$
$P_0$	1.06	0.46	2.38	0.35	1.10
$P_1$	0.68	0.78	0.64	0.47	0.20
$P_2$	0.40	0.16	0.16	0.09	0.20

### APPENDIX C: DETAILS ABOUT THE SF + G RECONSTRUCTION BIAS ANALYSIS

This section provides a more detailed view at the reconstruction bias for the SF + G model discussed in Sec. III E. Let us recall that we are analyzing how much the placement of the fourth measurement  $q_4$  impacts the change in  $\rho(0)$  through the transfer function formalism:

$$\delta\rho(0) = [\mathcal{T}\mathcal{F}_4^{\rho(0)}]_{\eta_{y_4}}. \quad (C1)$$

Figures 11(a.1) and 11(b.1) show this calculated  $\delta\rho(0)$  in black (as an absolute percentage) as  $q_4$  moves along the possible momentum transfer range  $[0, 3.5]$  fm $^{-1}$  while the other  $q_j$  remain in position. The SF + G model with the hyperparameter  $\gamma = 0.7$  fm is displayed in (a.1) while  $\gamma = 1.4$  fm is used in (b.1). The  $x$  axis of both plots shows the two original data sets of  $\mathbf{q}_0 = [0.77, 1.30, 1.82, 2.41, 3.06]$  fm $^{-1}$  (blue), and  $\mathbf{q}_1 = [0.77, 1.30, 1.82, 2.7, 3.06]$  fm $^{-1}$  (orange), which only differ in their  $q_4$  value.

In Fig. 11(a.1) it can be observed that for most locations the calculated  $|\delta\rho(0)|\%$  exceeds the 5% threshold (the blue dashed line). On the other hand, for (b.1) the black curve is overall lower, showing a wider region where  $|\delta\rho(0)| \leq 5\%$ . This overall reduction is the reason why we chose  $\gamma = 1.4$  fm to perform the analysis in Sec. IV.

The “spikes” where  $|\delta\rho(0)|$  grows abruptly appear when we aim to move  $q_4$  to the other  $q_j$  locations, effectively measuring twice in the same spot. Since we have five parameters and five observations, this results in a non-invertible Hessian for  $\chi^2$ , which blows up the transfer function  $\mathcal{T}\mathcal{F}_4^{\rho(0)}$ . It is interesting to note that, even though there is not an observation located at  $q = 0$  fm $^{-1}$ , a “spike” can still be observed. This is because by definition all model form factors must respect  $F(0) = 1$ . Therefore a measure at  $q = 0$  fm $^{-1}$  provides no new information, resulting in a noninvertible Hessian as well.

Figures 11(a.2) and 11(b.2) show the true form factor (RMF012) in black and the observation locations as the blue and orange points, where (a.2) is for  $\gamma = 0.7$  fm and (b.2) is for  $\gamma = 1.4$  fm. The red squares represent “forbidden zones”: regions in the  $q$  space where if we place the fourth location  $q_4$  we obtain a bias in  $\rho(0)$  bigger than 5%. For (a.2) these zones occupy around 85% of the total  $q$  range, while for (b.2) they occupy around 70%.

By the construction in Sec. III E, locating  $q_4$  at the blue point  $q = 2.41$  fm $^{-1}$  will result in almost zero bias since the optimal form factor and the true form factor have the same

value ( $\eta_4 = 0$ ). As  $q_4$  moves away from this location, we expect the bias to increase. A “bad” model will present a bias that grows too quickly, while a “good” model will be more tolerable. As can be observed in Fig. 4, for the same  $q_4 = 2.7$  fm $^{-1}$ , the SF + G with  $\gamma = 0.7$  fm (orange dashed curve) fails to be within the  $\lambda_\rho$  band, while the SF + G with  $\gamma = 1.4$  fm (red dashed line) stays within. In Fig. 11(a.2) the orange point is covered by the red rectangle. This is not the case for (b.2).

It should be noted that changing the hyperparameter  $\gamma$  from 0.7 to 1.4 fm can impact the locations where  $\eta_j = 0$  since now the optimal function will be different. Nevertheless, in the case we studied here the change in these locations was negligible, a fact that simplified our discussion.

### APPENDIX D: DETAILS ABOUT THE $^{48}\text{Ca}$ AND $^{208}\text{Pb}$ CHARGE EXAMPLES AND MODEL COMPARISON

This Appendix presents tables and details relevant to the  $^{48}\text{Ca}$  and  $^{208}\text{Pb}$  charge density example developed in Sec. IV A. It includes the results of applying the same analysis on model comparison developed in Sec. IV B to the electric charge densities of both nuclei instead of their weak densities counterparts.

Tables III and IV show for the  $^{48}\text{Ca}$  and  $^{208}\text{Pb}$  examples, respectively, the original locations  $\mathbf{q}_0$  as well as the locations  $\mathbf{q}_m$  that minimize the FOM defined in Eq. (37). Tables III and IV also show the numerical values of the  $\mathcal{T}\mathcal{F}$  times the respective error  $\sigma_j$  for the density at  $r = 0$  fm and the radius for both data sets  $\mathbf{q}_0$  and  $\mathbf{q}_m$ . Let us recall that it is the total interior density [the 30 grid points in Eq. (36)] that goes in the FOM, Eq. (37). We are using  $\rho(0)$  as a representative of the total interior density. All of the transfer function values have been divided by the natural scales defined in Table I.

By adding in quadrature each element in the rows of Tables III and IV (the transfer functions values times the respective errors), the total variance in  $\rho(0)$  or  $R$  can be calculated from Eq. (25), in units of the natural scale. In this sense, each number in the table represents the contribution of that measurement to the total variance in that quantity. This allows us to identify, for example, that the variance in  $R$  in the case of  $\mathbf{q}_m$  for  $^{48}\text{Ca}$  is completely driven by the first observation  $q_1 = 0.9$  fm $^{-1}$ , while the contribution of the last point  $q_5 = 2.69$  fm $^{-1}$  is negligible. From an experimental design point of view, this means that if our main goal is to reduce the uncertainty in the radius, we must allocate the resources accordingly and reduce the error bar on  $q_1$  rather than reducing the error bars in the other locations.

Figures 12 and 13 show, for  $^{48}\text{Ca}$  and  $^{208}\text{Pb}$ , respectively, the results for comparing the seven models defined in Sec. II B to recover the interior charge density and charge radius. The numerical values shown in these Figures are written in Tables VII and VIII. Tables V and VI show the associated optimal locations  $\mathbf{q}_m$  for every model for  $^{48}\text{Ca}$  and  $^{208}\text{Pb}$ , respectively.

We judge each model by their individual bias, standard deviation, and MSE both in the original locations  $\mathbf{q}_0$  as well as the optimized locations  $\mathbf{q}_m$  where the FOM is minimized. As we did in Sec. IV B, the starting locations

are  $q_0 = [0.9, 1.35, 1.8, 2.24, 2.69] \text{ fm}^{-1}$  for  $^{48}\text{Ca}$  and  $q_0 = [0.63, 0.94, 1.26, 1.57, 1.88] \text{ fm}^{-1}$  for  $^{208}\text{Pb}$ .

We observe trends similar to the one displayed by Figs. 7 and 8 when comparing models for the weak charge densities. Changing the data locations can result in an important reduction of the MSE, for example by a factor of 2 by the SF in the  $^{48}\text{Ca}$  radius. There are also significant variations in performance among the models, for example a factor of 3 between the SF + B and SF + G when compared to the Helm model when extracting the interior density of  $^{208}\text{Pb}$  for  $q_0$ . Once again, for the data range and errors we have assumed, the Helm and SF model seem to outperform all of the other options in both nuclei.

Finally, it is interesting to analyze the distribution of optimal locations  $q_m$  for the seven models in Tables V and VI. Let us recall that we limited the maximum value of any  $q_j$  to be less than  $3.5 \text{ fm}^{-1}$  for  $^{48}\text{Ca}$  and less than  $2 \text{ fm}^{-1}$  for  $^{208}\text{Pb}$ . Although there is no clear pattern, both the SF and Helm models seem to have overall smaller values of  $q_j$ , while the models involving Bessels seem to be more in the high end. Depending on the experiment details and constraints, some regions of the  $q$  space would be easier to access than others. In that case, a more detailed analysis could be done to optimize a modified version of the FOM in which we take into account the experimental budget. It is very likely that the FOM would not be extremely sensitive to the exact locations of  $q_m$ . Therefore, an adjustment of each  $q_j$  could result in a substantial reduction of the experimental budget while the FOM deteriorates just a small amount.

#### APPENDIX E: DETAILS ABOUT THE $^{48}\text{Ca}$ AND $^{208}\text{Pb}$ WEAK CHARGE ANALYSIS

This Appendix presents tables and details relevant to the  $^{48}\text{Ca}$  and  $^{208}\text{Pb}$  weak charge density model comparison developed in Sec. IV B. The numerical values associated with Figs. 7 and 8 are shown in Tables XI and XII, respectively. Tables IX and X show the associated optimal locations  $q_m$  for every model for  $^{48}\text{Ca}$  and  $^{208}\text{Pb}$ , respectively.

As done in Appendix D, we can analyze the distribution of these optimal locations  $q_m$ . We observe a similar structure from their charge counterparts: no clear pattern but both SF and Helm models seem to have overall smaller values of  $q_j$ , while the models involving Bessels seems to be more in the high end.

Since parity violating experiments are extremely expensive and challenging, the experimental constraints and budget should definitely be considered in a more detailed analysis to optimize a modified version of the FOM. We again anticipate that the FOM will not be extremely sensitive to the exact locations of  $q_m$ , and that an adjustment of each  $q_j$  could result in a huge impact on the budget while the FOM deteriorates just a small amount. This analysis could also change our conclusions regarding the optimal models. For example, for the weak results of  $^{208}\text{Pb}$ , the Helm model seems to have a better FOM combination than the SF, but it could be that the  $q_m$  from the SF model are more experimentally accessible than those from the Helm model.

TABLE XIV. Absolute value prior transfer function values for the density at  $r = 0 \text{ fm}$  ( $|\mathcal{T}\mathcal{F}_k^{\rho(0)}\sigma_k|/\Delta\rho_{\text{Ca}}$ ) for the three amplitudes in units of the natural scale  $\Delta\rho_{\text{Ca}}$ . Each  $\mathcal{T}\mathcal{F}$  has been multiplied by its respective  $\sigma_k$  [see Eq. (39)] to somehow represent a fraction of the total SD. The finite prior is the only one considered.

	$A_1$	$A_2$	$A_3$
$P_1$	0.89	0.24	0.58

Finally, Figs. 14 and 15 show, as an example, the detailed results of the Bessels model for  $^{48}\text{Ca}$  and the SF model for  $^{208}\text{Pb}$ , respectively, across the five different generators used in Sec. IV B. The last columns on each figure show the squared average [see Eq. (20)] of each quantity (bias, SD, and MSE), which corresponds to the single values displayed in Figs. 7 and 8 in the main text. As can be seen, the variations among different generators are very small. We interpret this as a sign that our conclusions are robust at least within the family of generators we considered in this study.

#### APPENDIX F: DETAILS ABOUT THE ROLE OF PRIORS

This section presents tables with the numerical values of the transfer functions of both the data and the priors related to what was discussed in Sec. IV C. Let us recall that we are using the SF + G model with three scenarios for the prior. In terms of their strength these are (1)  $P_0$ : no prior (unconstrained SF + G); (2)  $P_1$ : the specified prior in Eq. (39); and (3)  $P_2$ : a very restrictive prior which makes the SF + G practically behave as the SF model without Gaussians.

Table XIII shows the numerical value of  $|\mathcal{T}\mathcal{F}_j^{\rho(0)}\sigma_j|/\Delta\rho_{\text{Ca}}$  for the five different locations and for the three different priors. Let us recall that the FOM in Eq. (37) is calculated using the entire interior density, but in this section we are focusing in  $\rho(0)$  as a representative. For each scenario, the total contribution of the data in the variance on the density at  $r = 0$ , namely  $\Delta\rho(0)^2$ , is obtained by adding the numbers in Table XIII in quadrature.<sup>2</sup> In this sense, each number in the table reflects how much that particular data point uncertainty  $\sigma_j$  contributes to the total band that surrounds  $\rho(0)$  in Fig. 10, in units of the natural scale  $\Delta\rho_{\text{Ca}}$  from Table I.

It can be observed that, as the prior strength increases, the influence of each data point uncertainty  $\sigma_j$  tends to decrease, in some cases even by an order of magnitude. This reflects the fact that models constrained by priors, which are in some sense less complex, will present a smaller variance. From the transfer function point of view, this is driven by a more constrained  $\tilde{\mathcal{H}}$  [see Eq. (18)].

To get a complete picture in terms of standard deviation, we must also look at the variance introduced by the addition of the prior itself. In Eq. (16a), we can see that each Gaussian prior has an associated center  $\omega_k^0$ , and “error”  $\sigma_k$  which are analogous to the true observations  $y_j$  and true error  $\sigma_j$  in

<sup>2</sup>The position  $r = 0 \text{ fm}$  is only one of the 30 grid points in Eq. (36). The total interior MSE will receive contributions not only from  $\rho(0)$ .

TABLE XV. Prior transfer function values for the density at  $r = 0$  fm ( $\mathcal{T}\mathcal{F}_k^{\rho(0)}\eta_{\omega_k}/\Delta\rho_{Ca}$ ) for the three amplitudes in units of  $\Delta\rho_{Ca}$ . Each  $\mathcal{T}\mathcal{F}$  has been multiplied by its respective  $\tilde{\eta}_k$  to represent its fraction of the prior-induced bias.

	$A_1$	$A_2$	$A_3$
$P_1$	5.1	-0.17	-2.7
$P_2$	10	0.64	-11

the data in Eq. (10). Although  $\tilde{\mathcal{H}}$  is more constrained, we now have to take into account the transfer functions ( $\mathcal{T}\mathcal{F}_k^{\rho(0)}$ ) associated with the fact that in principle  $\omega_k^0$  could fluctuate around its center by as much as  $\sigma_k$  (to 1 sigma).

The scenario without prior  $P_0$  does not present these transfer functions. The scenario with the extremely constrained prior  $P_2$  does not present the transfer functions either since it is as if the parameters associated with the Gaussians were not there in the first place (also, we have observed that numerically  $\mathcal{T}\mathcal{F}_k^{\rho(0)}\sigma_k \rightarrow 0$  when  $\sigma_k \rightarrow 0$ ). Table XIV shows in the case of the intermediate prior  $P_1$  the numerical value of these  $\mathcal{T}\mathcal{F}_k^{\rho(0)}$  times their respective uncertainty  $\sigma_k$  in terms of the natural scale  $\Delta\rho_{Ca}$ . The total variance in  $\rho(0)$  for the model with this prior is calculated by adding in quadrature these values plus the ones associated with the data in Table XIII.

We can also describe the induced bias by the inclusion of a prior, i.e., how the estimated central value of  $\rho(0)$  is impacted by the new prior. This description is done in terms of the prior transfer functions  $\mathcal{T}\mathcal{F}_k^{\rho(0)}$  by analyzing how the parameters  $\omega_{P_0}$  (those obtained in the absence of a prior) move to either  $\omega_{P_1}$  or  $\omega_{P_2}$  (the fitted parameters when using prior  $P_1$  or  $P_2$ ). The reasoning is similar to the discussion about the  $\eta_j$  and how they moved the fitted parameters  $\omega$  away from the optimal value  $\omega_{Opt}$ .

Suppose that we are currently at  $\omega_{P_0}$  and we add a prior term to  $\chi^2$  [converting it to  $\tilde{\chi}^2$  defined in Eq. (17)] in such a way that the centers  $\omega_k^0$  fall exactly at the value of their respective parameters  $\omega_{P_0}\{k\}$ ; then the central value of the fitted parameters will not change by the addition of that prior.

As a concrete example, let us assume that when the data are fitted the value of the first unconstrained parameter of (A1) is  $\omega_{P_0}\{1\} = 0.005$ . Then, if we add to the total  $\chi^2$  a prior term of the form

$$\frac{(\omega_1 - 0.005)^2}{\sigma_1^2}, \quad (F1)$$

and fit the parameters again, we will obtain the same values for all the parameters (the Hessian would be more restricted, but the center location will be intact). Now, let us imagine that we ‘‘perturb’’ the value of the center  $\omega_k^0$  (in the example 0.005), and move it to the original location of the prior we want to enforce (either  $\omega_{P_1}$  or  $\omega_{P_2}$ ). This change, which we call  $\tilde{\eta}_k$ , will now produce a change in the value of the parameters and therefore, in  $\rho(0)$ :

$$\delta\rho(0) = \mathcal{T}\mathcal{F}_k^{\rho(0)}\tilde{\eta}_{\omega_k}, \quad (F2)$$

where  $\tilde{\eta}_k \equiv \omega_k^0 - \omega_{P_0}\{k\}$ , the difference between the parameter’s value without priors and the new prior centers. Table XV shows in units of  $\Delta\rho_{Ca}$ , the predicted change  $\delta\rho(0)_k = \mathcal{T}\mathcal{F}_k^{\rho(0)}\tilde{\eta}_{\omega_k}$  driven by the inclusions of the prior in the three amplitudes of the Gaussians for  $P_1$  and  $P_2$ .

To obtain the total predicted change in  $\rho(0)$  (in units of  $\Delta\rho_{Ca}$ ), we must add all numbers in each row. Due to the fact that they alternate signs for this example, the total change in the case with the stronger prior  $P_2$  ends up being smaller than the intermediate strength. However, we can appreciate that as the prior strength increases, the influence of its transfer functions increases by a factor between 2 and 4. We interpret this as an increase in the bias of the model as compared to its priorless counterpart.

Finally, note that Eq. (F2) is just a linear approximation. Nonlinear models will deviate from the predictions of this linear approximation if the parameters change considerably. In this particular case, the predictions on  $\delta\rho(0)$  are within  $\approx 40\%$  of the true change  $\rho(0)$  obtained when refitting the parameters with the new priors. Even though the numerical accuracy is not perfect, these types of analysis can help in better estimating the bias vs variance tradeoff when including priors.

- 
- [1] C. F. von Weizsäcker, *Z. Phys.* **96**, 431 (1935).  
[2] H. A. Bethe and R. F. Bacher, *Rev. Mod. Phys.* **8**, 82 (1936).  
[3] G. Fricke, C. Bernhardt, K. Heilig, L. A. Schaller, L. Schellenberg, E. B. Shera, and C. W. de Jager, *At. Data Nucl. Data Tables* **60**, 177 (1995).  
[4] I. Angeli and K. Marinova, *At. Data Nucl. Data Tables* **99**, 69 (2013).  
[5] R. Hofstadter, *Rev. Mod. Phys.* **28**, 214 (1956).  
[6] J. D. Walecka, *Electron Scattering for Nuclear and Nucleon structure* (Cambridge University Press, New York, 2001).  
[7] D. Androic *et al.*, *Phys. Rev. Lett.* **111**, 141803 (2013).  
[8] D. Androic *et al.*, *Nature (London)* **557**, 207 (2018).  
[9] M. Thiel, C. Sfienti, J. Piekarewicz, C. J. Horowitz, and M. Vanderhaeghen, *J. Phys. G* **46**, 093003 (2019).  
[10] C. J. Horowitz, J. Piekarewicz, and B. Reed, *Phys. Rev. C* **102**, 044321 (2020).  
[11] K. Scholberg, *Phys. Rev. D* **73**, 033005 (2006).  
[12] J. Yang, J. A. Hernandez, and J. Piekarewicz, *Phys. Rev. C* **100**, 054301 (2019).  
[13] D. Z. Freedman, *Phys. Rev. D* **9**, 1389 (1974).  
[14] D. Akimov *et al.*, *Science* **357**, 1123 (2017).  
[15] D. Akimov *et al.*, *Phys. Rev. D* **100**, 115020 (2019).  
[16] T. Donnelly, J. Dubach, and I. Sick, *Nucl. Phys. A* **503**, 589 (1989).  
[17] S. Abrahamyan, Z. Ahmed, H. Albatineh, K. Aniol, D. S. Armstrong *et al.*, *Phys. Rev. Lett.* **108**, 112502 (2012).  
[18] C. J. Horowitz, Z. Ahmed, C. M. Jen, A. Rakhman, P. A. Souder, M. M. Dalton, N. Liyanage, K. D. Paschke, K. Saenboonruang, R. Silwal, G. B. Franklin, M. Friend, B. Quinn, K. S. Kumar, D. McNulty, L. Mercado, S. Riordan, J. Wexler, R. W. Michaels, and G. M. Urciuoli, *Phys. Rev. C* **85**, 032501(R) (2012).

- [19] D. Adhikari *et al.* (PREX Collaboration), *Phys. Rev. Lett.* **126**, 172502 (2021).
- [20] CREX: Parity-violating measurement of the weak charge distribution of  $^{48}\text{Ca}$ , <http://hallaweb.jlab.org/parity/prex/c-rex/c-rex.pdf>
- [21] C. J. Horowitz, K. S. Kumar, and R. Michaels, *Eur. Phys. J. A* **50**, 48 (2014).
- [22] D. Becker *et al.*, *Eur. Phys. J. A* **54**, 208 (2018).
- [23] B. A. Brown, *Phys. Rev. Lett.* **85**, 5296 (2000).
- [24] R. J. Furnstahl, *Nucl. Phys. A* **706**, 85 (2002).
- [25] M. Centelles, X. Roca-Maza, X. Viñas, and M. Warda, *Phys. Rev. Lett.* **102**, 122502 (2009).
- [26] X. Roca-Maza, M. Centelles, X. Viñas, and M. Warda, *Phys. Rev. Lett.* **106**, 252501 (2011).
- [27] C. J. Horowitz and J. Piekarewicz, *Phys. Rev. Lett.* **86**, 5647 (2001).
- [28] C. J. Horowitz and J. Piekarewicz, *Phys. Rev. C* **64**, 062802(R) (2001).
- [29] J. Carriere, C. J. Horowitz, and J. Piekarewicz, *Astrophys. J.* **593**, 463 (2003).
- [30] A. W. Steiner, M. Prakash, J. M. Lattimer, and P. J. Ellis, *Phys. Rep.* **411**, 325 (2005).
- [31] J. Erler, C. J. Horowitz, W. Nazarewicz, M. Rafalski, and P.-G. Reinhard, *Phys. Rev. C* **87**, 044320 (2013).
- [32] W.-C. Chen and J. Piekarewicz, *Phys. Rev. C* **90**, 044305 (2014).
- [33] W.-C. Chen and J. Piekarewicz, *Phys. Lett. B* **748**, 284 (2015).
- [34] Z. Lin and C. J. Horowitz, *Phys. Rev. C* **92**, 014313 (2015).
- [35] J. Piekarewicz, A. R. Linero, P. Giuliani, and E. Chicken, *Phys. Rev. C* **94**, 034316 (2016).
- [36] K. J. Åström and R. M. Murray, *Feedback Systems: An Introduction for Scientists and Engineers* (Princeton University Press, Princeton, 2010).
- [37] D. W. Higinbotham, P. Giuliani, R. E. McClellan, S. Sirca, and X. Yan, [arXiv:1812.05706](https://arxiv.org/abs/1812.05706).
- [38] P. Guèye *et al.* (Jefferson Lab Hall A Collaboration), *Eur. Phys. J. A* **56**, 126 (2020).
- [39] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Springer-Verlag, New York, 2009).
- [40] X. Yan, D. W. Higinbotham, D. Dutta, H. Gao, A. Gasparian, M. A. Khandaker, N. Liyanage, E. Pasyuk, C. Peng, and W. Xiong, *Phys. Rev. C* **98**, 025204 (2018).
- [41] T. J. Sullivan, *Introduction to Uncertainty Quantification*, Texts in Applied Mathematics Vol. 63 (Springer International, Cham, 2015).
- [42] P. J. Huber, *Robust Statistics*, Wiley Series in Probability and Statistics Vol. 523 (John Wiley & Sons, New York, 2004).
- [43] D. G. Cacuci, *Sensitivity and Uncertainty Analysis, Volume I: Theory* (CRC, Boca Raton, FL, 2003).
- [44] M. Catacora-Rios, G. B. King, A. E. Lovell, and F. M. Nunes, [arXiv:2012.06653](https://arxiv.org/abs/2012.06653).
- [45] H. De Vries, C. De Jager, and C. De Vries, *At. Data Nucl. Data Tables* **36**, 495 (1987).
- [46] B. Dreher, J. Friedrich, K. Merle, H. Rothhaas, and G. Lührs, *Nucl. Phys. A* **235**, 219 (1974).
- [47] R. H. Helm, *Phys. Rev.* **104**, 1466 (1956).
- [48] D. Sprung and J. Martorell, *J. Phys. A: Math. Gen.* **30**, 6525 (1997).
- [49] I. Sick, *Nucl. Phys. A* **218**, 509 (1974).
- [50] P. R. Bevington and D. K. Robinson, *Data Reduction and Error Analysis for the Physical Sciences*, 3rd ed. (McGraw-Hill, New York, 2003).
- [51] P. C. Gregory, *Bayesian Logical Data Analysis for the Physical Sciences* (Cambridge University Press, Cambridge, UK, 2005).
- [52] J. V. Stone, *Bayes' Rule: A Tutorial Introduction to Bayesian Analysis* (Sebtel, Sheffield, UK, 2013).
- [53] J. J. Moré and S. M. Wild, *ACM Trans. Math. Software* **38**, 1 (2012).
- [54] *Recent Advances in Algorithmic Differentiation*, edited by S. Forth, P. D. Hovland, E. Phipps, J. Utke, and A. Walther (Springer-Verlag, Berlin, 2012).
- [55] S. M. Wild, in *Advances and Trends in Optimization with Engineering Applications*, edited by T. Terlaky, M. F. Anjos, and S. Ahmed (SIAM, Philadelphia, 2017), pp. 529–540.
- [56] A. W. Van der Vaart, *Asymptotic Statistics* (Cambridge University Press, Cambridge, 2000).
- [57] C. J. Horowitz, E. F. Brown, Y. Kim, W. G. Lynch, R. Michaels *et al.*, *J. Phys. G* **41**, 093001 (2014).
- [58] C. M. Bishop, *Pattern Recognition and Machine Learning* (Springer-Verlag, New York, 2006).
- [59] L. Neufcourt, Y. Cao, S. A. Giuliani, W. Nazarewicz, E. Olsen, O. B. Tarasov *et al.*, *Phys. Rev. C* **101**, 044307 (2020).
- [60] M. Spivak, *Calculus on Manifolds: A Modern Approach to Classical Theorems of Advanced Calculus* (CRC, Boca Raton, FL, 2018).
- [61] C. R. Goodall, in *Handbook of Statistics*, Vol. 9 (Elsevier, 1993), Chap. 13, pp. 467–508.
- [62] R. D. Woods and D. S. Saxon, *Phys. Rev.* **95**, 577 (1954).