

FLORIDA STATE UNIVERSITY
COLLEGE OF ARTY SCIENCE

OUTSMARTING UNCERTAINTY: STATISTICS STRATEGIES FOR NUCLEAR PHYSICS

By

PABLO GIULIANI

A Dissertation submitted to the
Department of Physics
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

2020

Pablo Giuliani defended this dissertation on N.A..
The members of the supervisory committee were:

Jorge Piekarewicz
Professor Directing Thesis

Simon Capstick
Committee Member

Sergio Almaraz-Calderon
Committee Member

Nicholas Bonesteel
Committee Member

Eric Chicken
University Representative

The Graduate School approval.

ACKNOWLEDGMENTS

My first -and gigantic- thanks goes to my advisor Jorge Piekarewicz. More than an advisor, Jorge was (and is) a mentor for me, someone who not only helped me guide and consolidate my physics research, but also taught me a great deal about what being a researcher and being a professor means. Jorge, thank you very much for your unconditional support these years, for giving me enough freedom to explore subjects I found interesting, for guiding me to canalize those interests into tangible results, and for making these last five years a delightful experience. I am enormously indebted to you, and I look forward to follow your steps and “pay it forward”.

I want to thank my collaborators without whom many of the ideas I develop in this manuscript would not have reached their full potential. From the Statistics Department at FSU I thank Eric Chicken and Antonio Linero (now at UT Austin). Antonio was also a key guide for me in my pursue of knowledge in statistics. From the Statistics Department at TAMU I thank Shuang Zhou, Debdeep Pati, and Anirban Bhattacharya. From the Jefferson Lab Hall A Collaboration I thank Douglas Higinbotham and Paul Gueye. From the LIGO group at Stanford I thank Brian Lantz and Edgard Bonilla. The Transfer Function Formalism I develop here was inspired by the project on Johnson Noise I worked with Brian and Edgard. Many thanks to my committee members as well.

Next, I want to thank my family, with special emphasis to my Mom (Ana Posada), Dad (Fernando Giuliani), and Dog (Maxwell). Mom, thank you so much for all the support you gave me in my transition to the PhD, for helping me to communicate better in my presentations and articles, for being there for me always, “etc., etc., etc.”. Dad, thank you very much for your morning messages every single day, and for sharing your cheers and energy with me. Maxwell, thank you for taking me out to a park to play every single day as well.

Next comes my friends, both Tallahassee-located as well as those scattered around the globe, with a special emphasis to Diogenes Figueroa and Edgard Bonilla. Diogenes, thanks for sharing so many good moments (and coffee) with me in our office at Keen, and for the invaluable research discussions. Edgard, thanks for being an “all-terrain” friend, and for debating, criticizing, and praising many of the ideas I share in this manuscript.

Finally, I want to thank my adoptive family: the Physics Department at FSU. The knowledge, experiences, and friendships I gathered here will last me a lifetime.

TABLE OF CONTENTS

List of Tables	viii
List of Figures	xi
Abstract	1
1 Introduction	2
2 Background	13
2.1 Nuclear Physics	13
2.1.1 Proton's Form Factor and Radius	13
2.1.2 Nuclear Density and Form Factor	15
2.2 Statistics	19
2.2.1 Standard Fitting Protocol	20
2.2.2 Bayesian Approach	21
2.2.3 Bias, Variance, and MSE	23
3 The Transfer Function Formalism	26
3.1 Transfer Functions: Definition	26
3.2 Variance calculation	27
3.3 Bias calculation and the Central Function	29
3.4 Mean Squared Error	34
3.5 Priors under the TF formalism	35
4 The Proton Radius	36
4.1 Parametric model selection: Bias vs Variance trade-off	37
4.1.1 Bias	37
4.1.2 Variance	39
4.1.3 Bias and variance combined: MSE	41
4.1.4 The Best Predictive Model	43
4.1.5 Semi-analytical Calculations as an alternative to Monte Carlo	44
4.1.6 Noise Bias	45
4.1.7 Standard Deviation	47
4.1.8 Maximizing Information Gained	48
4.2 Nonparametric approach: Constrained Gaussian Process	51
4.2.1 Approximating G_E : basis construction	52
4.2.2 Incorporating physical constraints	54
4.2.3 Probabilistic model for fully constrained function estimation	55
4.2.4 Prior specification: Constrained Gaussian Process	56
4.2.5 Posterior sampling and inference	58
4.2.6 Electron-scattering data analysis	59

5 Weak Charge Densities	65
5.1 Bayesian selection of a second measurement of the weak-charge form factor of ^{208}Pb	66
5.1.1 Symmetrized Fermi Form Factor	68
5.1.2 Helm Form Factor	69
5.1.3 Parameter Optimization	71
5.1.4 Selection of the second value of the momentum transfer	73
5.1.5 Modeling the electric form factor	74
5.2 The transfer function formalism applied to uncertainty quantification on weak charge densities reconstruction	81
5.2.1 Reconstruction Bias and the Optimal Function	83
5.2.2 Results: Analyzing charge and weak charge densities	86
5.2.3 Example: ^{48}Ca and ^{208}Pb charge density analysis	87
5.2.4 Model Comparison: ^{48}Ca and ^{208}Pb Weak Charge Density	90
5.2.5 The role of priors	94
6 Dispersive Effects on ^{12}C	97
6.1 The impact of dispersive corrections on nuclear radii	100
7 Conclusions	109
7.1 Going Forward	112
Appendix	
A Nonparametric approach: Constrained Gaussian Process	116
A.1 Theoretical proofs for the constraints on f_ξ and on the independence of \mathcal{C}_ξ on τ	116
A.2 Details regarding the construction of the regression model	117
A.3 Details regarding the constrained model without the constraint $\xi_1 = 1$ (c ₁ GP)	119
A.4 Details regarding the choices of priors and hyperparameters, and the Gibbs sampling steps	121
A.5 Pseudo-data analysis	123
A.6 Details regarding the Electron-Scattering Data Analysis	128
B The Transfer function formalism	137
B.1 Mathematical proofs regarding the TF formalism	137
B.1.1 Transfer Functions structure	137
B.1.2 Comparing the transfer function variance to the standard approach's variance	139
C The TF formalism applied to weak charge densities reconstructions	141
C.1 Model descriptions	141
C.2 Details about the SF+G reconstruction bias analysis	145
C.3 Details about the ^{48}Ca and ^{208}Pb charge examples and model comparison	147
C.4 Details about the ^{48}Ca and ^{208}Pb weak charge analysis	153
C.5 Details about the role of priors	156
Bibliography	159
Biographical Sketch	174

LIST OF TABLES

4.1	The mean a_0 and radius from doing 10^6 Monte Carlo simulations for each interval in Q^2 where Eq. (4.1) was used to generate pseudo data in 0.05 fm^{-2} steps with each data point smeared by a randomly generated, normally distributed point-to-point uncertainty of 0.5%. The results clearly indicate that the linear fits are biased. The input radius was 0.8113 fm (an a_1/a_0 term of 0.1097 fm^{-1}) and $a_0 = 1 \dots \dots \dots$	38
4.2	An expanded version of Table 4.1 where instead of just showing the mean offset of the fit results for a_1/a_0 , the bias, we also indicate the width of the fit results, σ . The point uncertainty is fixed at 0.5%. For the simulated radius of 0.8113 fm , one would expect an unbias fit to give a_1/a_0 of 0.11 fm^2 ; thus, the difference of that value from the mean fitted value of a_1/a_0 is the bias and the width of the distribution is σ . Also shown is the mean square error, MSE, which can be used to quantify the best function for a given interval, taking into account both bias and variance.	39
4.3	Same as Table 4.2, but now with equal number of data points for each range.	39
4.4	For the lowest Q^2 interval, 0.1 to 0.4 fm^{-2} , we compare regressions with a linear function (Eq. 4.2) to the dipole fit function (Eq. 4.4). Keeping the range fixed, a spacing of 0.05 fm^{-2} (7 points), 0.01 fm^{-2} (31 points) and 0.005 fm^{-2} (61 points) was used with various absolute random errors. In several of these cases, the simple linear function provides a better predictive model then the true functional form and is never far from the true function.	44
4.5	High and low regime posterior estimates of the radius and one standard deviation (Δr_p) for cGP, c_0 GP, c_1 GP and uGP with $N = \{n/4, n\}$ and $\nu = 2.5$	60
5.1	Central values for the Gaussian distribution assumed for the prior (see Eq. (5.19)). In all cases, the <i>relative</i> standard deviation has been fixed to one half, for example, $\sigma_c = \bar{c}/2 = 3.342 \text{ fm}$. All quantities are given in fm.	74
5.2	Average values and their corresponding theoretical uncertainties generated from the posterior distribution for a Symmetrized Fermi and Helm form factors. Also shown are the predictions for the charge radius of ^{208}Pb , which should be compared against the experimental value of $R_{\text{exp}}^{208} = (5.5012 \pm 0.0013) \text{ fm}$ [8]. All quantities are given in fm.	78
5.3	Predictions from a set of accurately calibrated relativistic mean-field models [96] for the charge radius, weak-charge radius, and their difference for ^{208}Pb (all in fm). Also shown is the weak-charge form factor at the PREX momentum transfer, the slope of the symmetry energy L (in MeV), and the radius of a $1.4 M_\odot$ neutron star (in km). Some of these quantities may be compared against the following experimental values: $R_{\text{ch}}^{208} = 5.5012(13) \text{ fm}$ [8], $F_{\text{wk}}^{\text{PREX}} = 0.204(28)$, and $R_{\text{wk}}^{208} = 5.826(181) \text{ fm}$ [82].	81

5.4	Natural scales for the uncertainties in the interior density and radius for ^{48}Ca and ^{208}Pb . A subscript name is attached when referring to a specific nucleus. For example, the natural scale for the MSE of the radius of ^{48}Ca is denoted as ΔR_{Ca}	87
5.5	MSE results for the interior (Int) and radius (Rad) for ^{48}Ca and ^{208}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 5.4.	89
6.1	The first column shows the index number of the special momentum transfer $i\pi/R_{\text{cut}}$. The second column shows its form factor value obtained from [101]. The third column shows the value of $\partial R_{\text{ch}}/\partial y_j$. The fourth column shows the percentage change needed in y_j to generate an equivalent change of 1% in the estimated charge radius.	106
6.2	The first column shows the perturbation form $S(q)$ in addition to the empirical perturbation. In the first three cases a strength of $\delta(E_e) = 0.3$ was assumed (30% dispersive effects strength as the extrapolation to 1 GeV in Fig. 6.3). The second column shows the calculated new radius, $R_{\text{ch}}^{\text{stat}}$ (the original radius is 2.4711 fm). The third and fourth columns show the β coefficient and the percentage change in R_{ch} , namely $\delta R_{\text{ch}} \equiv R_{\text{ch}}^{\text{stat}} - R_{\text{ch}}^{\text{disp}}$, respectively.	107
A.1	Posterior estimates of the radius and credible interval for cGP, $c_0\text{GP}$, $c_1\text{GP}$ and uGP with $N = n/4$, $\nu = 2.5$ and $\ell_{\text{opt}} = 20$ for a subset of data of size $n = 500$ in regime i).	125
A.2	Posterior estimates of the radius and credible interval for cGP, $c_0\text{GP}$, $c_1\text{GP}$ and uGP with $N = n/4$, $\nu = 2.5$ and $\ell_{\text{opt}} = 20$ for a subset of data of size $n = 500$ in regime ii).	126
A.3	Posterior estimates of the radius and credible interval for cGP, $c_0\text{GP}$, $c_1\text{GP}$ and uGP with $N = n/4$, $\nu = 2.5$ and $\ell_{\text{opt}} = 1$ for a subset of data of size $n = 500$ in regime iii).	127
A.4	High regime posterior estimates of the radius and credible interval for cGP, $c_0\text{GP}$, $c_1\text{GP}$ and uGP with $N = \{n/4, n\}$ and $\nu = \{2.5, 3, 3.5\}$ and $n = 1422$.	129
A.5	Low regime posterior estimates of the radius and credible interval for cGP, $c_0\text{GP}$, $c_1\text{GP}$ and uGP with $N = \{n/4, n\}$ and $\nu = \{2.5, 3, 3.5\}$ and $n = 500$.	129
C.1	^{48}Ca electric form factor momentum transfer locations q_j and transfer functions $\mathcal{T}\mathcal{F}_j$ 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 5.4.	148
C.2	^{208}Pb electric form factor momentum transfer locations q_j and transfer functions $\mathcal{T}\mathcal{F}_j$ 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 5.4.	148
C.3	Optimal locations \mathbf{q}_m for each model when optimizing the Figure of Merit in Eq (5.26). These q_j values correspond to the solid bars results in Fig. C.2 for the charge density of ^{48}Ca . All values are in units of fm^{-1}	149

C.4	Optimal locations \mathbf{q}_m for each model when optimizing the Figure of Merit in Eq (5.26). These q_j values correspond to the solid bars results in Fig. C.3 for the charge density of ^{208}Pb . All values are in units of fm^{-1} .	149
C.5	Numerical values for the data displayed in Fig. C.2: model comparison for the electric charge density of ^{48}Ca .	150
C.6	Numerical values for the data displayed in Fig. C.3: model comparison for the electric charge density of ^{208}Pb .	150
C.7	Optimal locations \mathbf{q}_m for each model when optimizing the Figure of Merit in Eq (5.26). These q_j values correspond to the solid columns results in Fig. 5.10 for the five generated weak charge densities of ^{48}Ca . All values are in units of fm^{-1} .	153
C.8	Optimal locations \mathbf{q}_m for each model when optimizing the Figure of Merit in Eq (5.26). These q_j values correspond to the solid columns results in Fig. 5.11 for the five generated weak charge densities of ^{208}Pb . All values are in units of fm^{-1} .	153
C.9	Numerical values for the data shown in Fig. 5.10: model comparison for the five generated weak charge densities of ^{48}Ca .	155
C.10	Numerical values for the data shown in Fig. 5.11: model comparison for the five generated weak charge densities of ^{208}Pb .	155
C.11	Absolute value transfer function values for the density at $r = 0 \text{ fm}$ ($ \mathcal{T}\mathcal{F}_j^{\rho(0)}\sigma_j /\Delta\rho_{\text{Ca}}$) for the five locations $\mathbf{q}_0 = [0.9, 1.35, 1.8, 2.24, 2.69] [\text{fm}^{-1}]$ in units of the natural scale $\Delta\rho_{\text{Ca}}$. Each $\mathcal{T}\mathcal{F}$ has been multiplied by their respective σ_j to somehow represent a fraction of the total standard deviation SD. The three different prior options are explored for the SF+G model.	156
C.12	Absolute value of the 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 their respective σ_k (see Eq. (5.28)) to somehow represent a fraction of the total SD. The finite prior is the only one considered.	157
C.13	Prior transfer function values for the density at $r = 0 \text{ fm}$ ($\mathcal{T}\mathcal{F}_k^{\rho(0)}\eta_{\omega_k}/\Delta\rho_{\text{Ca}}$) for the three amplitudes in units of $\Delta\rho_{\text{Ca}}$. Each $\mathcal{T}\mathcal{F}$ has been multiplied by their respective $\tilde{\eta}_k$ to represent their fraction of the prior-induced bias.	158

LIST OF FIGURES

1.1	Radii extracted from electron scattering data as of 2017 [25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 22, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58]. The vertical bands indicate the value and uncertainty of the proton radius from muonic hydrogen [10, 59], and from CODATA-2014 [4]. Figure courtesy of D. Higinbotham.	4
2.1	Electron scattering in the Plane Wave Born Approximation (PWBA) from a target X which could represent a single proton or an entire nucleus. The electron e is shown in blue, the target X as a black circle, and the exchanged photon γ in red. The transferred momentum Q is equal to the difference between the outgoing momentum p' and the incoming momentum p of the electron. The scattering angle θ is measured between p and p' in the laboratory frame.	14
2.2	Proton's form factor G_E as a function of the squared momentum transfer Q^2 . The red points correspond to the Mainz data [109]. The inset plot zooms in the region $Q^2 \leq 0.3 \text{ fm}^{-2}$. The orange curve represents a particular model fitted to the data to perform the extrapolation to $Q^2 \rightarrow 0$	15
2.3	Relationship between the electric charge, protons and the single proton form factors along with their respective densities for ^{12}C . The protons density ρ_{protons} specifies the spatial distribution of the 6 protons inside the ^{12}C nucleus, treating them as point particles (blue circles over the black background in the middle column). The charge form factor F_{ch} , which relates to the <i>charge</i> distribution in the nucleus (left column), is the result of folding the protons form factor F_{proton} with the single proton form factor F_p , which relates to the charge distribution inside the proton (right column, the color circles represent the three quarks).	16
2.4	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) while the blue triangle is one of its members. The purple and yellow blobs represent 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.	24
3.1	Evolution of ω_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 it is difficult to note in the plot, the mapping “folds into itself” in the parameter space. The red star represents the central parameters (see the text preceding Eq. (3.7)), 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 following Eq. (3.7) and in Fig 3.2)	30
3.2 Example of the change in the central parameters driven by deviations between the central and the true form factors. (a) Reconstructed density from ω_c , reconstructed density from ω_e using Eq(3.7), and the true density (red, blue, and black lines, respectively). The inset plot shows the locations in the parameter space of ω_c and the estimated ω_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 5.2.2. (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 (η_2) in the neighborhood of $q_2 = 1.8 \text{ fm}^{-1}$	31
4.1 A graphic representation of the Monte Carlo results showing how the linear fits tend to have a relatively high bias though a low variance, while the quadratic fits tend to have a relatively low bias but a large variance. The black line/point represents the true value.	40
4.2 An illustration of the trade-off between bias and variance when selecting a statistical model. Simple models will have low variance but high bias (under-fitting) while complex models will have low bias but high variance (over-fitting). It is this trade-off that one seeks to balance. While with repeated Monte Carlo simulations it is trivial to find the optimal predictive model for a given set of data, in the real world the true model is typically unknown. We only get to perform a very limited number of experiments and thus, we rely on using real data and statistical methods for model selection [145].	41
4.3 The result of a million simulations and fits of linear fits over the Q^2 range of $0.1 - 0.8 \text{ fm}^{-2}$ and quadratic fits over $0.1 - 1.6 \text{ fm}^{-2}$, both with 31 uniformly spaced data points. Using the mean square error as the metric, neither example is significantly better than the other for exacting the proton radius. This is analogous to a dart game between two equally skilled players: one who hits the bull's eye more often yet has a large spread (low bias but high variance), and another, equally skilled, player who has a tighter cluster of hits but an offset (high bias but low variance).	42
4.4 Line MSE, absolute bias and SD, for the optimal location of points as a function of the number of points available, in the range $0.1-0.8 \text{ fm}^{-2}$. The dashed horizontal lines show the respective values for the configuration of 31 equidistant points showed in Table 4.2.	50
4.5 Parabola MSE, absolute bias and SD, for the optimal location of points as a function of the number of points available, in the range $0.1-0.1.6 \text{ fm}^{-2}$. The dashed horizontal lines show the respective values for the configuration of 31 equidistant points showed in Table 4.2.	51
4.6 Functions $h_0(x)$ (blue), $\psi_0(x)$ (orange) and $\phi_0(x)$ (green) for $N = 10$. The functions ψ_0 and ϕ_0 have been rescaled by a factor of 10 and 50 respectively.	53

4.7	Estimated density plots of MCMC samples of radius r_p for cGP, c_0 GP, c_1 GP and uGP with $N = n/4$ (dotted line), n (solid line) and $\nu = 2.5$ for the high Q^2 regime (a) and for the low Q^2 regime (b). The vertical dashed lines stand for the muonic result of 0.84 fm (red), and the recommended CODATA value of 0.88 fm in 2014 [4] (purple).	61
4.8	Function fit with $\nu = 2.5$ (a), 3 (b), 3.5 (c) and $N = n$ in the high regime. The inset plot in (a) shows the overall fit of the models for $\nu = 2.5$ to the entire data range. The solid curves denote the model predictions while the shaded intervals bounded by dotted lines represent the 95% confidence intervals for the predictions. The red dots denote the experimental data obtained from Mainz with their respective error bars. The red and blue points near the origin at $Q^2 = 0.008 \text{ fm}^{-2}$ represent the estimated we had at the time of these calculations on the lower value the PRad experiment would be able to measure, with two different estimates for the projected uncertainty [157], and arbitrary $G_E(Q^2)$ value.	62
4.9	Function fit with $\nu = 2.5$ (a), 3 (b) 3.5 (c) and $N = n$ in the low regime. The inset plot of (a) shows the overall fit of the models for $\nu = 2.5$ to the entire data range. The solid curves denote the model predictions while the shaded intervals bounded by dotted lines represent the 95% confidence intervals for the predictions. The red dots denote the experimental data obtained from Mainz with their respective error bars. The red and blue points near the origin at $Q^2 = 0.008 \text{ fm}^{-2}$ represent the estimated we had at the time of these calculations on the lower value the PRad experiment would be able to measure, with two different estimates for the projected uncertainty [157], and arbitrary $G_E(Q^2)$ value.	63
5.1	(a) Correlation plot between the half-density radius c and the surface diffuseness a that define the Symmetrized Fermi function. The number of points represents the raw data obtained from the Monte Carlo simulation. Also shown are the respective averages and standard deviations. The solid red line represents the functional form obtained from solving the equation $F_{\text{ch}}(q_1; a, c) = 0.210$. (b) Probability distribution function for the charge radius of ^{208}Pb obtained from the Monte Carlo simulation. The solid orange line marks true experimental value of R_{ch}^{208}	75
5.2	Variability in the charge form factor of ^{208}Pb generated by the distribution of parameters displayed in Fig. 5.1. Also shown with the dark green solid line is the relative uncertainty in the model predictions as a function of the momentum transfer.	76
5.3	(a) Correlation plot between the half-density radius c and the surface diffuseness a that define the Symmetrized Fermi function. The number of points represents the raw data obtained from the Monte Carlo simulation. Also shown are the 39% and 95% confidence ellipses. The solid red and purple lines represent the functional form obtained from solving the equations: $F_{\text{ch}}(q_1; a, c) = 0.210$ $F_{\text{ch}}(q_2; a, c) = -0.061$. (b) Probability distribution function for the charge radius of ^{208}Pb obtained from the Monte Carlo simulation. The black solid line represents a fit to a Gaussian probability distribution and the green vertical line represents the true experimental value of R_{ch}^{208}	77

5.4	(a) Charge form factor, and (b) the corresponding charge density of ^{208}Pb . The two red points on the left-hand panel represent the single input used in the calibration of the Symmetrized Fermi function. The theoretical predictions are displayed by an uncertainty band (in cyan). The experimental data is from Ref. [6].	78
5.5	(a) Charge form factor and (b) the corresponding charge density of ^{208}Pb . The two red points represent the single input used in the calibration of the Helm function. The theoretical predictions are shown by an uncertainty band (in cyan). The experimental data is from Ref. [6].	79
5.6	Charge and weak-charge density of ^{208}Pb as predicted by a collection of accurately calibrated mean-field models. The labels indicate the predicted neutron-skin thickness of ^{208}Pb . The inset shows the fastest Gaussian falloff of the Helm form factor relative to the exponential falloff of Symmetrized Fermi function. The experimental data are from Ref. [6].	81
5.7	(a) Recovered densities using the SF+G model for two different data locations. The generated (“true”) ^{48}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 γ set to 1.4 fm instead of 0.7 fm. The blue band is associated with the scale λ_ρ . (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 λ_F	84
5.8	Results for optimizing the FOM for ^{48}Ca using the Bessel model (a), and for ^{208}Pb using the SF model (b). The original momentum transfer locations 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.	88
5.9	Log plots of the electric form factors recovered in Fig. 5.8. (a) Shows the results for ^{48}Ca using the Bessel model. (b) shows the results for ^{208}Pb using the SF model. For both (a) and (b) the original momentum transfer locations q_0 and their respective fitted form factors are shown in blue, while the optimized locations q_m and their respective fitted form factors are shown in orange. The black curve represents the respective experimental charge form factor [6, 7]. For clarity, the central curves are displayed as dashed lines, while the error bands are shown as colored bands.	88
5.10	Comparison of seven models for recovering the interior density and mean weak radius from the weak form factor data on ^{48}Ca . The three columns to the left of each model show the bias, SD, and MSE for the interior density. The three columns to the right of each model show 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 \mathbf{q}_m , while the borderless columns represent the starting \mathbf{q}_0	91

5.11	Comparison of seven models for recovering the interior density and mean weak radius from the weak form factor data on ^{208}Pb . The three columns to the left of each model show the bias, SD, and MSE for the interior density. The three columns to the right of each model show 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 \mathbf{q}_m , while the borderless columns represent the starting \mathbf{q}_0	92
5.12	(a) Reconstructed weak density of ^{48}Ca for the generator RMF012, and for the model SF+G with three sets of priors. P_0 (in blue), P_1 (in green), and P_2 (in orange). (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 $\Delta\rho_{\text{Ca}} = 0.00015 \text{ fm}^{-3}$	95
6.1	High-order corrections to the one-photon exchange Born approximation in electron/positron-nucleus scattering.	98
6.2	Calculations of Friar and Rosen [100] for dispersion corrections to elastic electron scattering from ^{12}C at 374.5 and 747.2 MeV in the first diffraction minimum $q_{eff} = 1.84 \text{ fm}^{-1}$. Figure adapted from [197].	99
6.3	World data on the energy dependence of dispersive effects in the first diffraction minimum of ^{12}C	100
6.4	^{12}C form factor expanded in the Bessel functions formalism using Offermann [101] coefficients without dispersive corrections. The circles in the q axis show the special values of momentum transfer for the first 9 (red) from experimental data and the second 9 (black) from the extrapolation suggested in [112]. The dashed blue lines enclose the region of the data excluded from the analysis in [101]. The inset plot shows the three test forms for $S(q)$ in addition to the empirical perturbation obtained directly from the data by third degree spline interpolation. The curves in the inset plot are the ones needed to obtain the corrected $F_{\text{ch}}^{\text{stat}}$ from the observed $F_{\text{ch}}^{\text{disp}}$ values.	104
6.5	$\partial a_\nu / \partial y_j$ matrix for the data extracted from Offermann et al. [101].	105
A.1	(a) Functions $h_0(x)$, $\psi_0(x)$ and $\phi_0(x)$ for $N = 10$. The functions ψ_0 and ϕ_0 have been rescaled by a factor of 10 and 50, respectively. (b) Approximation (black) of the dipole function (Red) using the basis functions $h_j(x)$ (blue) on 11 gridpoints between 0 and 1 (black dots). The function $h_7(x)$, which is centered at $x = 0.6$, is highlighted to illustrate its “spike” form. The approximation matches the function so well that the true red curve is hard to see. (c)-(d) Approximation (black) of the same dipole function (red) using the basis functions ϕ_j as in Eq. (A.9) (c), and Eq. (A.10) (d). In both cases, the functions ϕ_j are plotted starting from a neighboring of their respective grid point and matching their value and their slope at the grid point with the complete approximation (black curve). In both cases the function $\phi_7(x)$ is highlighted. When the coefficients of the ϕ_j functions are fitted instead of matched to second derivatives (d), the red curve is hard to see again.	119

A.2	MSEs versus ℓ for $\nu = 2.5$, $\sigma = 0.005$ and $n = 500$. (a) shows the MSEs for data in regime i), (b) for regime ii), and (c) for regime iii). In each plot, the orange line represents the MSEs with the number of grid points $N = n/4$, the blue line represents $N = n/2$, and the green line for $N = n$	124
A.3	Model fits in regimes i)-iii) with noise level $\sigma = 0.005$. (a) is for regime i), (b) for regime ii) and (c) for regime iii). The red line represents the true function values; the green line represents the cGP estimates; the blue line represents the c_0 GP; the orange line represents the c_1 GP; the purple line represents uGP. In each case the 95% point-wise credible intervals are delimited by dashed lines with the respective color.	127
A.4	(a) MSEs over the range $\ell \in [0, 1]$ for cGP model in the high regime with $n = 250$, $N = n/4$, $n/2$, n . (b) MSEs over ℓ for cGP model in the low regime with $N = \{n/4, n/2, n\}$, $n = 500$ and $\nu = 2.5$. In both graphs the green line represents the case with $N = n$, the blue line represents $N = n/2$, and the orange line represents $N = n/4$. (a) (Inset) MSEs over $\ell \in [0.5, 2.5]$ for cGP model with $N = n/4$ on full dataset. (b) (Inset) MSEs over $\ell \in [1, 20]$ for cGP model with $N = n/4$ on the low regime. In both insets $\nu = 2.5$ (solid), $\nu = 3$ (dashed), $\nu = 3.5$ (dot-dashed).	130
A.5	Histogram plots of MCMC samples of the radius r_p for cGP (a), c_0 GP (b), c_1 GP (c) and uGP (d) with $N = n/4$ and $\nu = 2.5$ for the full dataset. The red and purple vertical dashed lines indicate the values of 0.84087 fm and 0.8775 fm, respectively.	131
A.6	Histogram plots of MCMC samples of the radius r_p for cGP (a), c_0 GP (b), c_1 GP (c) and uGP (d) with $N = n$ and $\nu = 2.5$ for the full dataset. The red and purple vertical dashed lines indicate the values of 0.84087 fm and 0.8775 fm, respectively.	132
A.7	Histogram plots of MCMC samples of the radius r_p for cGP (a), c_0 GP (b), c_1 GP (c) and uGP (d) with $N = n/4$ and $\nu = 2.5$ in the low regime case. The red and purple vertical dashed lines indicate the values of 0.84087 fm and 0.8775 fm, respectively.	133
A.8	Histogram plots of MCMC samples of the radius r_p for cGP (a), c_0 GP (b), c_1 GP (c) and uGP (d) with $N = n$ and $\nu = 2.5$ in the low regime. The red and purple vertical dashed lines indicate the values of 0.84087 fm and 0.8775 fm, respectively.	134
A.9	Histogram plots of MCMC samples of ξ_1 ($n_0 G_E(0)$) for c_1 GP and uGP with $N = \{n/4, n\}$ and $\nu = 2.5$ for the full dataset. (a) ($N = n/4$) and (b) ($N = n$) show the results of c_1 GP, (c) ($N = n/4$) and (d) ($N = n$) show the results of uGP.	135
A.10	Histogram plots of MCMC samples of ξ_1 ($n_0 G_E(0)$) for c_1 GP and uGP with $N = \{n/4, n\}$ and $\nu = 2.5$ for the low regime. (a) ($N = n/4$) and (b) ($N = n$) show the results of c_1 GP, (c) ($N = n/4$) and (d) ($N = n$) show the results of uGP.	136
C.1	Bias increase in $\rho(0)$ 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(C.26). The blue and orange points correspond to the two original data sets, while the dashed blue 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, $|\delta\rho(0)| > 5\%$ 146

- C.2 Seven models comparison for recovering the interior density and mean charge radius from the charge form factor data on ^{48}Ca . The bias, SD, and MSE for each model are shown in their respective three left columns for the interior density, and in their respective 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 \mathbf{q}_m , while the borderless ones represent the starting \mathbf{q}_0 . 151
- C.3 Seven models comparison for recovering the interior density and mean charge radius from the charge form factor data on ^{208}Pb . The bias, SD, and MSE for each model are shown in their respective three left columns for the interior density, and in their respective 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 \mathbf{q}_m , while the borderless ones represent the starting \mathbf{q}_0 . 152
- C.4 Bessel model bias, standard deviation (SD), and MSE across the five different generators for the weak charge of ^{48}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 columns refer to the radius. All quantities have been divided by their natural scales defined in Table 5.4. 154
- C.5 SF model bias, standard deviation (SD) and MSE across the five different generators for the weak charge of ^{208}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 columns refer to the radius. All quantities have been divided by their natural scales defined in Table 5.4. 154

ABSTRACT

In this dissertation we explore three topics: 1) the extraction of the proton radius from electron scattering data; 2) the reconstruction of neutron distributions in nuclei from parity violating experiments; and 3) the propagation of dispersive corrections in electron-nucleus scattering to the estimated nuclear densities. We use various statistics tools and strategies including Bayesian analysis, bias-variance trade-off criteria, and a new framework we have developed under the name of “the transfer function formalism”.

The first topic is motivated by the decade old “proton radius puzzle”: a stark disagreement between muonic based and electronic based experiments aimed to extract the electric radius of the proton. Given the long and impassioned discussion about what model should be used to recover the radius, the main question driving us is: *What is the most robust and reliable way, from a statistical point of view, to extrapolate the electron scattering data and extract the radius?*

The second topic is motivated by the combination of the strong interest in weak charge densities by the nuclear physics and astrophysics communities, and by the enormously challenging experiments needed to access neutron distributions inside a nucleus. Given the complexity and cost of such experiments, the main question driving us is: *How to extract the most amount of information in cases with scarce data?*

The third topic is motivated by our lack of knowledge on the magnitude and energy dependence of dispersive effects, a fact that could negatively impact delicate experiments such as those involved with the second topic. In addition, the propagation of dispersive effects from raw data, to the calculated quantities through the fitted model, gave us a perfect benchmark to implement our proposed transfer function formalism. Given the uncertain structure of the dispersive effects, the main question driving us is: *How to propagate small perturbations in the data to calculated quantities in a straightforward way (without re-fitting the model)?*

This dissertation is a contribution to address these questions, and to strengthen multidisciplinary efforts between physics and statistics.

CHAPTER 1

INTRODUCTION

Nuclear physics is a broad field of science which main mission is to understand all manifestations of nuclear phenomena [1]. Regardless of whether probing individual nucleons, atomic nuclei, or neutron stars, a common theme is to characterize these objects in terms of their mass and radius. Shortly after the discovery of the neutron in 1932, Gamow, Weizsäcker, Bethe, and Bacher formulated the “liquid-drop” model to estimate the masses of atomic nuclei [2, 3]. Since then, remarkable advances in experimental techniques have been exploited to determine nucleon and nuclear masses with unprecedented precision. The rest mass of the proton, for example, is known to a few parts in a billion [4]. Similarly, starting with the pioneering work of Hofstadter in the late 1950’s [5], and continuing to this day [6, 7, 8], elastic electron scattering has provided the most accurate and detailed picture of the distribution of charge in nuclear systems.

Although not as impressive as in the case of nuclear masses, the charge radii of atomic nuclei has also been determined with extreme precision. The charge radius of ^{208}Pb , for example, is known to about two parts in 10,000 [8] (or $R_{\text{ch}}^{208} = 5.5012(13) \text{ fm}$). Given such an impressive track record, it came as a shocking surprise that the accepted 2010-CODATA value [9] for the charge radius of the proton obtained from electronic hydrogen and electron scattering, was in stark disagreement with a new result obtained from the Lamb shift in muonic hydrogen [10]. This unforeseen conflict with the structure of the proton has given rise to the “proton radius puzzle” [11, 12, 13].

The value of the charge radius of the proton $r_p = 0.84087(39) \text{ fm}$ determined from muonic hydrogen [10, 11] differed significantly (by $\sim 4\%$ or nearly 7σ) from the recommended 2010 CODATA value of $r_p = 0.8775(51) \text{ fm}$ [9], and even latter (by nearly 6σ) from the 2014 CODATA value of $r_p = 0.8751(61) \text{ fm}$ [14]. Note that these CODATA values were obtained by combining the results from both electron scattering and atomic spectroscopy [11, 4, 13]. The muonic measurement is so remarkably precise because the muon—with a mass that is more than 200 times larger than the electron mass and thus, has a Bohr radius 200 times smaller—is a much more sensitive probe of the internal structure of the proton. Of great relevance to the proton puzzle is the recent (2017)

measurement of the 2S-4P transition frequency in electronic hydrogen that suggests a smaller proton radius of $r_p = 0.8335(95)$ fm [15]. This is in agreement with the results from muonic hydrogen [16]. Although significant, it is still not understood why this present extraction differs from the large number of spectroscopic measurements carried out in electronic hydrogen throughout previous years. The explanation of this disagreement will also have to take into account even newer results (2017) in electronic spectroscopy, such as the measurement of the 1S-3S transition [17], which on the other hand, agrees with the larger 2014 CODATA value.

As in the case of earlier physics puzzles –notably the “solar neutrino problem”– the objective is to explain the discrepancy by exploring three non-mutually-exclusive options: (a) the experiment (at least one of them) is in error; (b) the theoretical models used in the extraction of the proton radius are the culprit (see for example [18] and references contained therein); or (c) there is new physics that affects the muon differently than the electron. Indeed, hints of possible violations to lepton universality are manifested in the anomalous magnetic moment ($g-2$) of the muon [19], and in certain decays of the B-meson into either a pair of electrons or a pair of muons [20].

In an effort to resolve the “proton radius puzzle”, several experiments in both spectroscopy and lepton-proton scattering have been commissioned. Spectroscopy of both electronic and muonic atoms as already initiated by Beyer et al. [16], will continue with a measurement of a variety of transitions to improve both the value of the Rydberg constant and the charge radius of the proton. Note that the Rydberg constant and r_p are known to be highly correlated. Lepton scattering experiments are planned at both the Thomas Jefferson National Accelerator Facility (JLab), and at the Paul Scherrer Institute (PSI). The proton radius experiment (PRad) at JLab which published results in November 2019 [21], collected data in the momentum-transfer range of $Q^2 = (10^{-4} - 10^{-1}) \text{ GeV}^2$, a wide-enough region to allow for comparisons with the most recent Mainz data [22], but also to extend the Mainz data to significantly lower values of Q^2 . Finally, the Muon Proton Scattering Experiment (MUSE), –although not yet in the data gathering stage– will fill an important gap by determining r_p from the scattering of both positive and negative muons of the proton. These experiments will be conducted concurrently with electron scattering measurements in an effort to minimize systematic uncertainties [23].

A significant portion of the physics community has been moving towards accepting the smaller radius consistent with the muonic result (see the new updated 2018 CODATA value of $r_p =$

$8.414(19)$ fm [24]). The PRad results pointed towards a smaller radius as well: $r_p = 0.831(14)$ fm [21]. Once the disagreement between the various electronic spectroscopy experiments is understood, the final piece would come from the MUSE results. Their findings could allow us to know that lepton universality indeed holds and that the cause of the puzzle is not that the proton interacts differently with muons and electrons. If that turns out to be the case, one of the most important questions we should aim to answer will be: *What went wrong with the old experiments and analyses?*

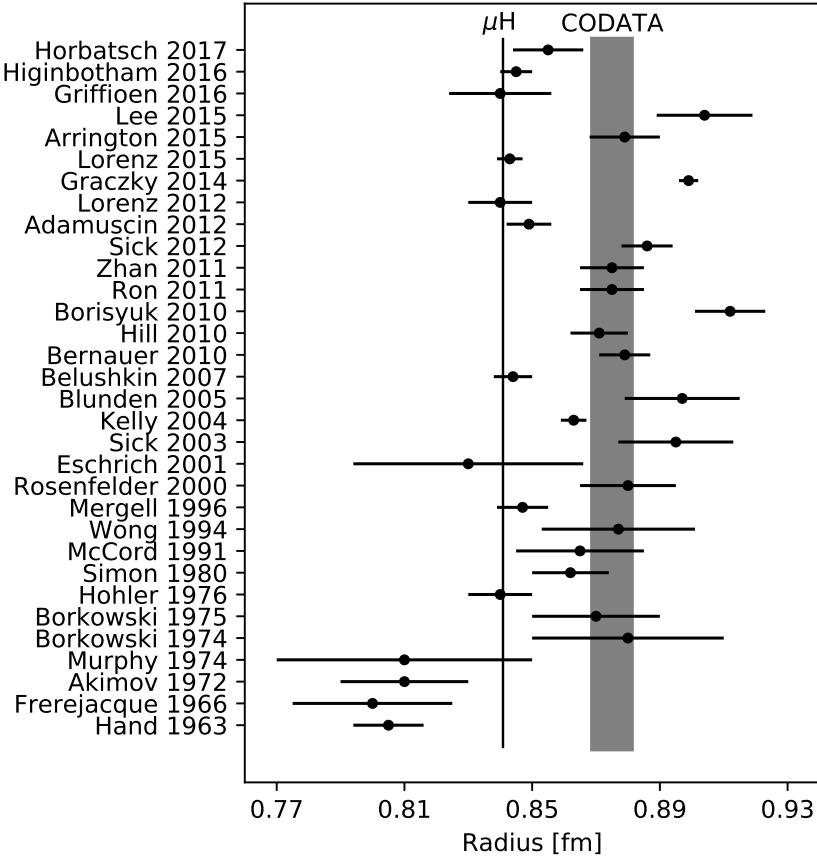


Figure 1.1: Radii extracted from electron scattering data as of 2017 [25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 22, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58]. The vertical bands indicate the value and uncertainty of the proton radius from muonic hydrogen [10, 59], and from CODATA-2014 [4]. Figure courtesy of D. Higinbotham.

In the case of scattering experiments, the proton charge radius is related to the slope of the electric form factor of the proton $G_E(Q^2)$ at the origin, i.e., at $Q^2=0$ (see Sec.2.1.1). To infer the

radius from experimental data, a subtle *extrapolation* to $Q^2 = 0$ is unavoidable. Focusing on the data available before the PRad results, the obtained proton radius from this extrapolation turned out to be quite sensitive to the model used to describe the form factor. To illustrate this, Fig. 1.1 shows some of the various proton radius values that have been extracted from electron scattering data in the last decades. In general, r_p extractions using simple statistical modeling of the low Q^2 data [60, 46, 27, 26] tend towards a smaller radius, while the more complex models with many free parameters tend to extract a large proton radius [22, 40, 28, 30, 32]. Nuclear theory constrained extractions tend to favor a smaller radius [51, 41, 25]. A technical detail that is affecting all these results, seems to be the handling of the normalization of the experimental data.

In a first attempt at mitigating such uncontrolled extrapolations, Higinbotham and collaborators have brought to bear the power of statistical methods into the solution of the problem [26]. In this work, the authors concluded that “*the electron scattering results and the muonic hydrogen results are consistent. It is the atomic hydrogen results that are the outliers*”. In a later work [61], Higinbotham and collaborators studied the performance of several popular models to extract the proton radius from data. They concluded that “*functions with a good χ^2 do not necessarily extrapolate well beyond the data; thus χ^2 alone cannot be used to determine which functions can robustly extract the proton radius*”, where the quantity χ^2 is often used to assess the fit of a model to experimental data (see Sec. 2.2.1). This work [61], identified three functional forms that were robust when extracting the radius from pseudo data. Based on their findings, one of these robust models, the (1,1) rational function, was latter used in the analysis of the PRad data [21].

The proton puzzle has been a physics challenge for over a decade. To solve the puzzle, in addition to the huge investments that have been done in new experiments, and the constant funding of theoretical advances, we strongly believe it is indispensable to involve robust statistical approaches. To give some perspective, the requested funding for the construction of the windowless gas flow target for the PRad experiment was $\sim \$0.5M$ [62], while the requested equipment funding for the MUSE experiment was $\sim \$2M$ [63]. As shown in [61], having accurate experimental data to very low Q^2 values in conjunction with a physically valid theoretical model of the proton form factor $G_E(Q^2)$ (for example, the actual functional form that created the pseudo data in their analysis), might still produce an incorrect proton radius. It is the combination of experiments, theory, and statistics that could get us out of this mess.

In this manuscript, we move in this direction by first exploring parametric extractions of the proton radius with pseudo data and how they are impacted by changing the momentum transfer measurement locations that the experiments could reach. We also explore the radius extraction by a nonparametric approach, guided by theoretical physical considerations on properties of the proton's form factor.

The proton radius puzzle is not the only research area of nuclear physics that has benefited from implementing statistical tools and strategies. In the context of extrapolations on the nuclear landscape, several groups have explored the use of machine learning tools such as Bayesian neural networks [64, 65], Bayesian statistics extrapolation techniques such as Gaussian Processes [64], and Bayesian Model Averaging [66]. In the area of effective field theories, the BUQEYE collaboration (Bayesian Uncertainty Quantification: Errors in Your EFT) [67] has created an uncertainty quantification framework for effective field theories. They have applied this framework, for example, to the nuclear matter equation of state [68]. With the Facility for Rare Isotope Beams (FRIB) very close to completion, the BAND collaboration (Bayesian Analysis for Nuclear Dynamics) has been formed by a group of physicists and statisticians from several institutions [69]. The primary objective of this collaboration is to create frameworks that “[...] *will use advanced statistical methods to produce forecasts for as-yet-unexplored situations that combine nuclear-physics models in an optimal way*”[69]. The study of neutron distributions inside neutron rich nuclei is yet another area which would benefit from statistical inputs. This is one of the topics we explore in this dissertation.

In the case of the proton puzzle, there are abundant data from different experiments. The problem is *how* to extract the important information out of these data sets to a reliable precision. In the case of the study of neutron skins and weak densities on neutron rich nuclei, there is only a single model independent data point from a single experiment with a few more on the way. The reason for this lack of data is that the experiments that can access this nuclear property in a model independent way are extremely expensive and complex. The problems are *how much* information can we extract from that single measurement, and *how to design* these expensive experiments to maximize the information gained. Naturally, statistics arguments are needed to address these issues.

Nuclear saturation, the existence of an equilibrium density, is a hallmark of nuclear dynamics. Using only a handful of parameters, the semi-empirical mass formula of Bethe and Weizsäcker [2, 3]

previously discussed, 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 semi-empirical 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:

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

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 of nucleons in atomic nuclei. To date, the most precise knowledge of the nuclear density comes from mapping the charge distribution of atomic nuclei [7, 8]. Starting with the already mentioned pioneer work of Hofstadter in the late 1950’s [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. As mentioned, the root-mean-square charge radius of ^{208}Pb , for example, is known with exquisite precision [8].

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 target may be written as [70]:

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

where Z is the electric charge of the nucleus and $Q^2 \geq 0$ is the square of the space-like four-momentum transfer. The Mott cross section represents the scattering of a relativistic (massless) electron from a spinless and structureless target, and is calculated 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 contrast to our knowledge of the distribution of weak charge, which is dominated by the neutrons because the weak charge of the proton is very small [71, 72]. 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 of this topic, see [73] 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}Pb is expected to be fairly constant and close to ρ_0 . As such, the interior baryon density of ^{208}Pb may provide the physical observable that is most closely related to ρ_0 [74].

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. Within this context, by “model independent” we mean that the extraction of the data from the experiment does not depend on nuclear models that, for example, describe the strong interaction between the hadronic probes and the nucleus target. The use of electroweak probes 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 [71, 72]. In the case of CEvNS, for example, the cross section is directly proportional to the square of the weak charge form factor. That is [75, 76]:

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

where G_F is Fermi’s constant, $Q_{\text{wk}} = -N + (1 - 4 \sin^2 \theta_W)Z$ is the total weak charge of the nucleus written in terms of the weak-mixing angle θ_W , 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 [77], even if it took more than four decades for its experimental confirmation [78, 79]. Indeed, the challenge in this experimental set-up resides in the fact that we cannot directly detect the scattered neutrino, and we must rely on measuring the recoil of the nucleus. Although still coming short on the necessary sensitivity on the experimental set-up, an analogy could be to measure the scattering of a ping pong ball that has just hit a truck, by measuring the recoil of the truck.

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 3-decade old idea by Donnelly, Dubach, and Sick who proposed the use of Parity Violating Electron Scattering (PVES) as a clean probe of neutron densities [80], the pioneering Lead Radius EXperiment (PREX) at JLab extracted the weak radius of ^{208}Pb , providing for the first time model-independence evidence in favor of a neutron-rich skin [81, 82]. To reach the original goal of a $\pm 0.06\text{ fm}$ determination of the weak radius of ^{208}Pb , the follow-up PREX-II campaign has now been completed and results are expected very soon. The ongoing CREX campaign will provide the first electroweak determination of the weak radius of ^{48}Ca [83, 84]. Beyond JLab, the Mainz Energy recovery Superconducting Accelerator (MESA), envisioned to start operations by 2023 [85], may be able to determine the weak radius of both ^{48}Ca and ^{208}Pb with increased precision [73]. Besides its intrinsic value as a fundamental nuclear-structure observable, knowledge of the neutron skin thickness of ^{208}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 L [86, 87, 88, 89]—a fundamental parameter of the Equation Of State (EOS) of neutron-rich matter that impacts the structure, composition, and cooling mechanism of neutron stars [90, 91, 92, 93, 94, 95, 96].

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 [80]:

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

where α is the fine structure constant, and the nuclear contribution enters as the ratio of the weak to the charge form factor. Given that F_{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_{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 experimental program may required measurements of A_{PV} at about five values of Q^2 , a task that may be feasible for ^{48}Ca [97]. For ^{208}Pb , this is a significantly more challenging task given that at high momentum transfers, the elastic cross section is small because of the strong suppression from the nuclear form factor. Nevertheless, as we explore in this manuscript, with two experimental points it may be sufficient to gain valuable insights into the weak charge form factor of ^{208}Pb over a significant range of momentum transfers (see also [74]). With asymmetries of the order of one part per million [81, 82], PVES experiments are both highly expensive and enormously challenging. Therefore, robust statistical arguments—above and beyond a compelling physics case—should be made in the quest for an optimal experimental design.

In this manuscript we also explore these topics. First, using Bayesian analysis techniques we determine the impact that a single new measurement of the weak form factor of ^{208}Pb could have on the determination of its neutron skin. Second, we consider the possibility of several measurements of the weak form factor of both ^{208}Pb and ^{48}Ca and analyze the performance of different models at extracting their respective densities, as done in [61] within the proton puzzle context. To do so, we introduce a novel statistical analysis—which we call the “transfer function formalism”—inspired by the treatment of noise in signal processing theory [98].

Using the transfer function formalism, we also explore a different though related topic. In all the cross sections we have discussed so far, the plane wave approximation is used to describe to first order the scattering of, for example, an electron from the target nucleus or nucleon. The assumption of this approximation is that the electron interacts weakly (meaning slightly) with the target through the exchange of one boson (either a photon or a Z_0 in the case of PVES). Several second order corrections can be implemented to this process as, for example, two photon exchanges, or radiative corrections in which the electron probe emits a photon that is not absorbed by the

nucleus. Another type of correction, in which the target is virtually excited during the process, is known as dispersive corrections [99]. This dynamical process is very complex and theoretical calculations [100] usually do not match what is observed in experiments [101]. The contribution of these dispersive effects is stronger near the diffraction minima of the form factor of the target $F(q) \sim 0$. Dispersive effects could present a problem when extracting neutron radii from parity violating experiments. This is because these experiments are designed to measure relatively close to the first diffraction minima to have a good balance between the decreasing magnitude of the form factor and the increasing magnitude of the asymmetry signal [102]. In this manuscript we use the transfer function formalism to propagate the influence of the dispersive effects to the extracted charge radius of ^{12}C . This framework can also be used to perform a similar analysis on neutron rich nuclei, like ^{48}Ca and ^{208}Pb .

The analyses we made regarding the proton puzzle, weak charge densities, and dispersive effects are intended not only to review existing data from a careful and robust statistical optic, but to serve as preliminary guidance for the design, implementation, and data analysis of future experiments within these contexts. A possible true discrepancy between muonic and electronic interactions with the proton could indicate new physics beyond the Standard Model [23], impacting many areas well beyond nuclear physics. As mentioned, the results that can be obtained from weak charge experiments do not only pertain to nuclear physics, but strongly correlate with properties of neutron stars [103]. We are entering a new era of multi-messenger astronomy, inaugurated by the simultaneous detection of gravitational waves [104] and short γ -ray bursts [105, 106] associated with a binary neutron star merger. FRIB is to begin operational phase soon, giving us the opportunity to have an unprecedented reach to study extremely neutron rich nuclei [107]. Now more than ever, the experimental designs here on earth, the astronomical observations of the heavens, and the theoretical advances to describe it all must be woven with statistical threads if we want to outsmart uncertainties and be able to explain from the quarks to the cosmos.

We have structured the rest of this dissertation into six additional chapters. Chapter 2, *Background*, provides a review of the concepts that we considered crucial in the discussion of the proton and nuclear form factors, as well as the nuclear density distributions. This chapter also includes a discussion of the main statistical concepts that are an essential component of our work such as traditional χ^2 fits, Bayesian inference, and bias-variance trade-off.

Chapter 3, *The Transfer Function Formalism*, presents a detailed description of the formalism we are proposing, and how it is implemented in the context of the bias-variance trade-off. The application of priors (from Bayesian statistics) is also discussed from the transfer function formalism perspective.

Chapter 4, *The Proton Radius*, discusses the now decade old proton puzzle and some of the efforts that have been made to solve it, and presents the approaches we explore. This chapter includes two main sections. Sec. 4.1 explores the bias vs variance trade-off in parametric models like a straight line and a parabola, as well as the influence of experimental measurement locations on the extracted radius. Sec. 4.2 focus on the implementation of a nonparametric Gaussian Process (GP) approach to the radius extraction.

Chapter 5, *Weak Charge Densities*, is focused on weak charge densities of neutron rich nuclei. This chapter is also structured into two main sections. Sec. 5.1 uses a Bayesian framework to analyze the impact that a second measurement of the weak form factor of ^{208}Pb could have on the reconstruction of its weak density and radius. Sec. 5.2 uses the transfer function formalism to compare several parametric models when extracting the weak charge densities of ^{48}Ca and ^{208}Pb from several form factor measurements. This section describes how to use the transfer function formalism to identify “critical” points in the data (those that drive most of the uncertainty), and to perform model selection and model building.

Chapter 6, *Dispersive Effects*, describes how to propagate the dispersive effects, treated as a perturbation over the charge form factor data of ^{12}C , to the estimation of its electric charge radius. This propagation is done using the transfer function formalism.

Finally, Chapter 7, *Conclusions*, presents our final remarks, conclusions, and visions of the future. This last chapter is followed by 12 appendixes with tables, figures, detailed explanations, and mathematical proofs that support the central concepts we have developed. These appendixes are followed by the bibliography section listing the authors and contributions that were essential to our work.

CHAPTER 2

BACKGROUND

This chapter introduces background concepts that are crucial for the rest of the manuscript. Sec. 2.1 addresses nuclear physics topics such as the proton's form factor, nuclear electric and weak form factors, and their respective density distributions. Sec. 2.2 is focused on statistical concepts such as parameter fitting, Bayesian inference, and bias-variance trade-off.

2.1 Nuclear Physics

2.1.1 Proton's Form Factor and Radius

In the one-photon exchange approximation, the most general expression for the elastic electron-proton cross section consistent with Lorentz and parity invariance is encoded in two Lorentz-scalar functions: the electric G_E and magnetic G_M form factors of the proton. That is,

$$\frac{d\sigma}{d\Omega} = \left(\frac{d\sigma}{d\Omega} \right)_{\text{Mott}} \left(\frac{G_E^2(Q^2) + \tau G_M^2(Q^2)}{1 + \tau} + 2\tau G_M^2(Q^2) \tan^2(\theta/2) \right), \quad (2.1)$$

where the square of the four-momentum transfer is given by:

$$Q^2 \equiv -(p' - p)^2 = 4EE' \sin^2(\theta/2). \quad (2.2)$$

Note that E (E') is the initial (final) energy of the electron, θ is the scattering angle (all in the laboratory frame), $\tau \equiv Q^2/4M^2$, and M is the mass of the proton. Fig. 2.1 shows a diagram of the scattering process. The internal structure of the proton is imprinted in the two form factors, with the electric one describing (in a non-relativistic picture) the distribution of charge, and the magnetic form factor describing the distribution of magnetization. Finally, the Mott cross section introduced in Eq.(2.1) represents the scattering of a massless electron from a spinless and structureless point charge. That is,

$$\left(\frac{d\sigma}{d\Omega} \right)_{\text{Mott}} = \frac{4\alpha^2}{Q^4} \frac{E'^3}{E} \cos^2(\theta/2) = \frac{\alpha^2}{4E^2 \sin^4(\theta/2)} \frac{E'}{E} \cos^2(\theta/2), \quad (2.3)$$

where α is the fine structure constant.

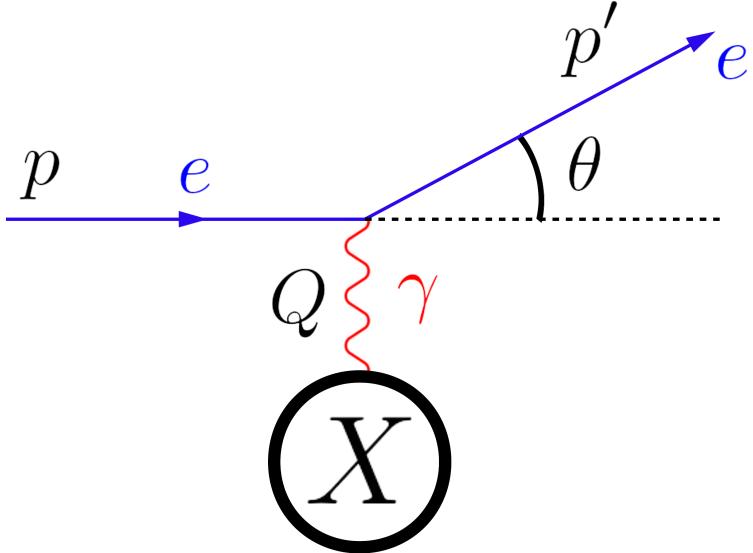


Figure 2.1: Electron scattering in the Plane Wave Born Approximation (PWBA) from a target X which could represent a single proton or an entire nucleus. The electron e is shown in blue, the target X as a black circle, and the exchanged photon γ in red. The transferred momentum Q is equal to the difference between the outgoing momentum p' and the incoming momentum p of the electron. The scattering angle θ is measured between p and p' in the laboratory frame.

Low-energy –or long wavelength– electrons are unable to resolve the internal structure of the proton and are therefore only sensitive to its entire charge. As the momentum transfer increases and the wavelength becomes commensurate with the proton size, finer details may now be resolved. In particular, the charge radius of the proton is defined as:

$$r_p^2 \equiv \langle r_E^2 \rangle = -6 \frac{dG_E}{dQ^2} \Big|_{Q^2=0}. \quad (2.4)$$

For a recent discussion on the subtleties of the definition of the proton radius, see [108].

Once the cross section has been measured, the electric form factor $G_E(Q^2)$ can be extracted from Eq (2.1) using, for example, the Rosenbluth separation technique [110]. Fig. 2.2 shows the 1422 experimental points of $G_E(Q^2)$ by the Mainz Collaboration [109]. Despite considerable efforts to determine $G_E(Q^2)$ at extremely low values of Q^2 , an extrapolation to $Q^2 = 0$ is unavoidable in order to recover the proton radius from Eq (2.4). This is usually done by fitting a model to the experimental data and extracting the model’s prediction for the slope at zero.

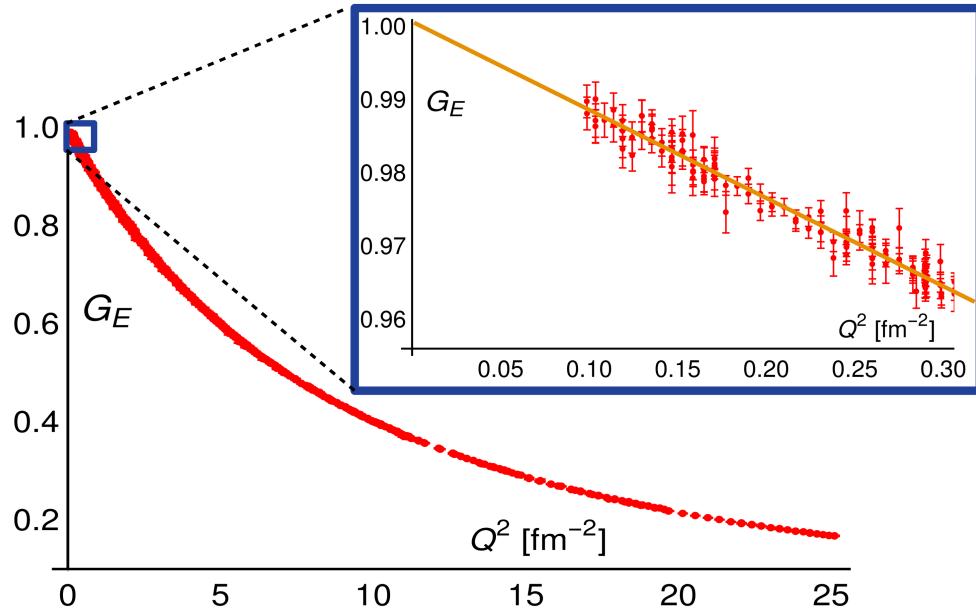


Figure 2.2: Proton’s form factor G_E as a function of the squared momentum transfer Q^2 . The red points correspond to the Mainz data [109]. The inset plot zooms in the region $Q^2 \leq 0.3 \text{ fm}^{-2}$. The orange curve represents a particular model fitted to the data to perform the extrapolation to $Q^2 \rightarrow 0$.

2.1.2 Nuclear Density and Form Factor

Protons can combine with neutrons to form a nucleus. We are interested in describing the distributions of these nucleons inside different nuclei. Since the electric charge of the neutron is zero, the nuclear electric charge density $\rho_{\text{ch}}(\mathbf{r})$, i.e., the distribution of electric charge in the nucleus, strongly correlates with the nuclear proton distribution. Since the weak charge of the proton is strongly suppressed by the weak mixing angle θ_W , $Q_W^p = 1 - 4\sin^2\theta_W = 0.0719(45)$ [111], the nuclear weak charge density $\rho_{\text{wk}}(\mathbf{r})$ strongly correlates to the weak charge in the atomic nucleus.

Under the approximation that the neutron does not contribute to the electric charge distribution in the nucleus, the relationship between $\rho_{\text{ch}}(r)$ and the protons density distribution $\rho_{\text{nuc}}(r)$ is then given in the Fourier transform space (see also Eq (2.9)):

$$F_{\text{ch}}(q) = F_{\text{protons}}(q)G_E(q), \quad (2.5)$$

where $F_{\text{ch}}(q)$ is the electric form factor of the nucleus (the Fourier transform of $\rho_{\text{ch}}(r)$), $F_{\text{protons}}(q)$ is the Fourier transform of the proton distribution $\rho_{\text{protons}}(r)$, $G_E(q)$ the form factor of a single

proton introduced in the previous section which relates to the electric density inside the proton $\rho_p(r)$, and $|\mathbf{q}|=q=\sqrt{Q^2}$ in the limit in which the nuclear recoil can be ignored (\mathbf{q} being the space component of the four vector Q) . Therefore, $\rho_{ch}(r)$ could be considered as resulting from folding the distribution $\rho_{protons}(r)$ of the protons inside the nucleus with the finite extension of the protons $\rho_p(r)$ [112]. Fig. 2.3 shows an example of this folding for the ^{12}C nucleus.

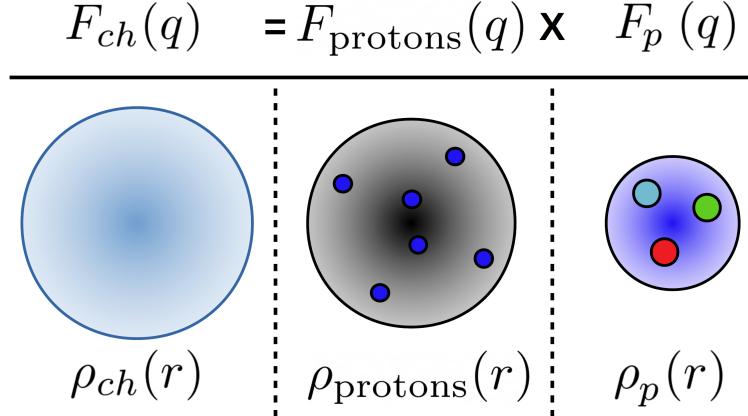


Figure 2.3: Relationship between the electric charge, protons and the single proton form factors along with their respective densities for ^{12}C . The protons density ρ_{protons} specifies the spatial distribution of the 6 protons inside the ^{12}C nucleus, treating them as point particles (blue circles over the black background in the middle column). The charge form factor F_{ch} , which relates to the *charge* distribution in the nucleus (left column), is the result of folding the protons form factor F_{proton} with the single proton form factor F_p , which relates to the charge distribution inside the proton (right column, the color circles represent the three quarks).

An analogous equation to (2.5) can be built between the neutron distribution $\rho_{\text{neutrons}}(r)$ and the weak charge distribution in the nucleus $\rho_{\text{wk}}(r)$ if the weak charge of the proton is ignored.

Once the relationship between both charge densities (electric and weak) and the corresponding nucleons distributions (protons and neutrons) is established, experiments that measure the former can provide information about the latter. In the case of $\rho_{ch}(\mathbf{r})$, as we discussed in the previous chapter elastic electron scattering experiments determine the ground state charge density by measuring the differential cross section, which for a spinless nucleus is given by:

$$\left(\frac{d\sigma}{d\Omega} \right) = \left(\frac{d\sigma}{d\Omega} \right)_{\text{Mott}} Z^2 F_{ch}^2(Q^2), \quad (2.6)$$

where Z is the number of protons and therefore the total electric charge of the nucleus, $F_{ch}^2(Q^2)$ is the nucleus electric form factor as a function of the momentum transfer Q^2 defined in (2.2), and

once again the Mott cross section represents the scattering of a massless electron from a spinless and structureless point charge, as introduced in Eq.(2.1). Fig. 2.1 shows a diagram of this scattering process.

In the case of the weak charge density, one path is to measure the parity violating asymmetry, which as we discussed in the previous chapter for a spinless target, and to first approximation, is given by [80]:

$$A_{PV}(Q^2) = -\frac{G_F Q^2}{4\pi\alpha\sqrt{2}} \frac{Q_{\text{wk}} F_{\text{wk}}(Q^2)}{Z F_{\text{ch}}(Q^2)}, \quad (2.7)$$

where α is the fine structure 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 θ_W , the number of protons Z and the number of neutrons N , and the nuclear contribution enters as the ratio of the weak $F(Q^2)_{\text{wk}}$ to the charge $F(Q^2)_{\text{ch}}$ form factor.

Another way to infer the nuclear weak charge density is through the coherent elastic neutrino-nucleus scattering (CEvNS). As we discussed in the previous chapter the cross section in this case is directly proportional to the square of the weak charge form factor [75, 76]:

$$\left(\frac{d\sigma}{dT} \right) = \frac{G_F^2}{8\pi} M \left[2 - 2 \frac{T}{E} - \frac{MT}{E^2} \right] Q_{\text{wk}}^2 F_{\text{wk}}^2(Q^2), \quad (2.8)$$

where G_F is Fermi's constant, and the weak form factor is normalized to $F_{\text{wk}}^2(Q^2=0)=1$. All the remaining quantities are of kinematical origin: E is the incident neutrino energy, T the nuclear recoil energy, and $Q^2=2MT$.

Having extracted the corresponding form factors F_{ch} and F_{wk} from experiment, 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 in the rest of this manuscript, except when this omission may create confusion. The density and form factor are related as follows:

$$\rho(\mathbf{r}) = \int \frac{d^3 q}{(2\pi)^3} e^{i\mathbf{q}\cdot\mathbf{r}} F(\mathbf{q}), \quad (2.9)$$

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:

$$\rho(r) = \frac{1}{2\pi^2 r} \int_0^\infty F(q) \sin(qr) q dq. \quad (2.10)$$

Alternatively, the inverse Fourier transform can be written as:

$$F(q) = \frac{4\pi}{q} \int_0^\infty \rho(r) \sin(qr) r dr. \quad (2.11)$$

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

$$F(q=0) = \int \rho(r) d^3r = 1. \quad (2.12)$$

The estimated density $\rho(r)$ and form factor $F(q)$ contain important information that is useful to many direct nuclear physics applications such as the calibration of the parameters of nuclear models (see for example [113]). In the case of the form factor, experimental values can also help us better understand complex dynamic properties of the nucleons inside the nucleus. The virtual excitation of the nucleus, a correction to the PWBA shown in Fig. 2.1, enters through what is called dispersive corrections. The effects of these corrections on the extracted charge radius of ^{12}C (see Eq (2.13)) is explored in Chapter 6.

In the case of the density $\rho(r)$, shell oscillations in the interior of several nuclei can be used to test the validity of nuclear models. Indeed, for ^{208}Pb the small bump in $\rho_{\text{ch}}(r)$ close to $r \sim 0$ due to the filling of the $3S$ proton state is over predicted by most density functional calculations [114]. Shell oscillations and the interior electric and weak density of ^{208}Pb and ^{48}Ca are explored in Chapter 5

From the density $\rho(r)$ the mean squared electric or weak radius can also be calculated as:

$$R^2 \equiv \langle r^2 \rangle = \int \rho(r) r^2 d^3r = 4\pi \int_0^\infty \rho(r) r^4 dr. \quad (2.13)$$

Once the electric charge and weak charge radius of the nucleus are calculated, the corresponding protons and neutrons distributions radii are obtained by taking into account the finite size of the nucleons. For example, the protons distribution's radius R_{protons} is related to the electric charge radius R_{ch} as:

$$R_{\text{ch}}^2 = R_{\text{nuc}}^2 + R_p^2 \quad (2.14)$$

with $R_p = 0.8414(19)$ fm the single proton radius [115]. An analogous equation can be written to relate the weak charge radius R_{wk} with the radius of the neutrons distribution R_{neutrons} . The extraction of the weak and electric radius of ^{208}Pb and ^{48}Ca is explored in Chapter 5.

Learning about the difference between the protons and neutrons radii for neutron rich nuclei have become an important goal in the nuclear physics community [1]. For example, since the proton radius of ^{208}Pb (R_p^{208}) is known with enormous accuracy [8], PREX effectively determined the neutron skin thickness of ^{208}Pb [81, 82]:

$$R_{\text{skin}}^{208} \equiv R_{\text{neutrons}}^{208} - R_{\text{protons}}^{208} = 0.33_{-0.18}^{+0.16} \text{ fm}. \quad (2.15)$$

The determination of the neutron skin thickness of ^{208}Pb is of great significance for multiple reasons. First, as an observable sensitive to the *difference* between the neutron and proton densities, it plays a critical role in constraining the isovector sector of the nuclear energy density functional [116, 117, 118, 95, 96]. Second, a very strong correlation has been found between the slope of the symmetry energy at saturation density (L) and R_{skin}^{208} [86, 87, 88, 89, 95]. This provides a powerful connection between a fundamental parameter of the equation of state (EOS) and a laboratory observable. Note that L is closely related to the pressure of pure neutron matter at saturation density. Third, constraining the EOS of neutron-rich matter provides critical guidance on the interpretation of heavy-ion experiments involving nuclei with large neutron-proton asymmetries [119, 120, 121, 122, 123, 124]. Finally, even though there is a difference in length scales of 18 orders of magnitude, the neutron skin thickness of ^{208}Pb and the radius of a neutron star share a common dynamical origin [90, 91, 92, 93, 125, 94, 95]. Although in general neutron star properties are sensitive to the high-density component of the EOS, it is the pressure in the neighborhood of twice nuclear matter saturation density that sets the overall scale for stellar radii [126]. Thus, whether pushing against surface tension in a nucleus or against gravity in a neutron star, it is the pressure in this neighborhood that determines both the thickness of the neutron skin and the radius of a neutron star. The symmetry energy L not only impacts the size, but also the structure (in terms of the tidal polarizability Λ [103]), composition, and cooling mechanism of neutron stars [90, 91, 92, 93, 94, 95, 96]. All these factors play a crucial role in our interpretation of multimesenger astronomy observations of neutron star mergers.

2.2 Statistics

In this manuscript we are interested in *local* models to describe the proton's form factor outlined in Sec. 2.1.1, and the nuclear densities and form factors outlined in Sec. 2.1.2. By *local* we mean

models that are fitted to experimental form factor values belonging to a single entity (either the proton or a particular nucleus). These are models such as the Symmetrized Fermi Function (SF) [127] and Fourier Bessel expansion [128] for nuclei, and the dipole function [129] for the proton’s form factor. In contrast, *global* models like FSUGarnet [96] are fitted using properties from many different nuclei.

Let us recall that the quantity that can be accessed experimentally is the form factor $F(q)$ at different momentum transfer values q . 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 σ_j is the associated experimental error. 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 or the proton radius, depends on the selection of experimental points through the fitted parameters $\boldsymbol{\omega}$.

2.2.1 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 χ^2 defined as follows:

$$\chi^2 = \sum_{j=1}^J \frac{(F(q_j, \boldsymbol{\omega}) - y_j)^2}{\sigma_j^2}, \quad (2.16)$$

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 $\boldsymbol{\omega}_0 \equiv \text{argmin}(\chi^2)$. Once the minimum of χ^2 is obtained, the central value of any quantity m is usually computed by just evaluating it at that parameter set:

$$\text{Reported value of } m = m(\boldsymbol{\omega}_0). \quad (2.17)$$

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 behavior is imprinted in the Hessian matrix of $\chi^2/2$ 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:

$$\mathcal{H}_{i,k} \equiv \frac{1}{2} \left(\frac{\partial^2 \chi^2}{\partial \omega_i \partial \omega_k} \right)_0 = \sum_{j=1}^J \frac{1}{\sigma_j^2} \left[\left(\frac{\partial F(q_j, \boldsymbol{\omega})}{\partial \omega_i} \right) \left(\frac{\partial F(q_j, \boldsymbol{\omega})}{\partial \omega_k} \right) + (F(q_j, \boldsymbol{\omega}) - y_j) \frac{\partial^2 F(q_j, \boldsymbol{\omega})}{\partial \omega_i \partial \omega_k} \right]_0. \quad (2.18)$$

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 [130]. For example, the square of the standard error (or standard deviation) of $m(\boldsymbol{\omega})$ is given by:

$$\Delta m^2 = \nabla m \mathcal{H}^{-1} \nabla m \Big|_{\boldsymbol{\omega}_0}, \quad (2.19)$$

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

2.2.2 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 [131, 132]. Within this framework, the posterior distribution of model parameters $\boldsymbol{\omega}$ given the experimental data \mathbf{Y} , is obtained from Bayes' theorem:

$$P(\boldsymbol{\omega}|\mathbf{Y}) = \frac{P(\mathbf{Y}|\boldsymbol{\omega})P(\boldsymbol{\omega})}{P(\mathbf{Y})}, \quad (2.20)$$

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,

$$\langle m \rangle = \int m(\boldsymbol{\omega}) P(\boldsymbol{\omega}|\mathbf{Y}) d\boldsymbol{\omega}, \quad (2.21a)$$

$$\Delta m^2 = \int (m(\boldsymbol{\omega}) - \langle m \rangle)^2 P(\boldsymbol{\omega}|\mathbf{Y}) d\boldsymbol{\omega}, \quad (2.21b)$$

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 χ^2 function introduced in Eq.(2.16) as follows:

$$P(\mathbf{Y}|\boldsymbol{\omega}) = e^{-\chi^2(\mathbf{Y}, \boldsymbol{\omega})/2}. \quad (2.22)$$

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

$$P(\boldsymbol{\omega}) = e^{-\phi^2(\boldsymbol{\omega})/2}, \text{ where} \quad (2.23a)$$

$$\phi^2(\boldsymbol{\omega}) = \sum_{k=1}^K \left(\frac{\omega_k - \omega_k^0}{\sigma_k} \right)^2, \quad (2.23b)$$

where ω_k^0 is our prior estimate for the central value of ω_k and σ_k is the estimated uncertainty. Small values of σ_k will make the distribution sharply peaked around ω_k^0 and the fitting procedure more “prior driven”. Conversely, large values of σ_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:

$$P(\boldsymbol{\omega}|\mathbf{Y}) = e^{-\tilde{\chi}^2(\mathbf{Y}, \boldsymbol{\omega})/2} = e^{-(\chi^2(\mathbf{Y}, \boldsymbol{\omega}) + \phi^2(\boldsymbol{\omega}))/2}, \quad (2.24)$$

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

$$\tilde{\mathcal{H}}_{i,k} \equiv \frac{1}{2} \left(\frac{\partial^2 \tilde{\chi}^2}{\partial \omega_i \partial \omega_k} \right)_0 = \mathcal{H}_{i,k} + \delta_{ik} \frac{1}{\sigma_k^2}, \quad (2.25)$$

where \mathcal{H} is the Hessian of χ^2 defined in Eq.(2.18) and δ_{ik} is the Kronecker delta. If the adopted prior includes correlations between the different parameters, then Eq.(2.23b) will be written as a quadratic form:

$$\phi^2(\boldsymbol{\omega}) = (\boldsymbol{\omega} - \boldsymbol{\omega}^0) \Sigma^{-1} (\boldsymbol{\omega} - \boldsymbol{\omega}^0), \quad (2.26)$$

where the matrix Σ contains the (prior) covariances between parameters. In such a case, Eq.(2.25) would have to be modified accordingly. In Sec. 4.2 we study the proton radius extraction using a constrained Gaussian Process. The prior we implement is precisely of the form in Eq (2.26), where Σ is built by the Gaussian Process kernel correlation function.

2.2.3 Bias, Variance, and MSE

In Sec. 4.1 and Sec. 5.2 one of our objectives is to quantify the performance of different models at extracting information from data. In Sec. 4.1 we focus on extracting the proton radius from generated form factor pseudo data, the models used being, for example, a straight line or a parabola. On Sec. 5.2 we focus on extracting both the electric and weak charge interior density and mean squared radius of ^{48}Ca and ^{208}Pb from experimental scattering data and generated form factor pseudo data, the models used being, for example, the Bessel Fourier expansion [128] and the Symmetrized Fermi Function [127]. The pseudo data in each case are created by a set of N functions we call *generators*: ${}^nF_{\text{true}}(q)$ ($n = 1, \dots, N$). For each data point generated there is an associated error σ_j which resembles realistic experimental uncertainties. Let us recall that we refer to any of these quantities of interest as m .

As in Refs.[61], we evaluate the performance of each model using a bias-variance trade-off criterion. Bias is understood as the discrepancy between the true value of m (coming from experiments or from one of the generators ${}^nF_{\text{true}}$) 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.(2.19) and (2.21b). Thus, we quantify the performance of a model by combining the bias and variance into the Mean Squared Error (MSE), defined as:

$$\text{MSE}^2(m, \mathbf{q}, \boldsymbol{\sigma}) \equiv \text{Bias}^2 + \text{Variance}. \quad (2.27)$$

Note that we have highlighted the dependence of the MSE on the quantity m , the locations of the momentum-transfer points \mathbf{q} and the associated errors $\boldsymbol{\sigma}$. The MSE will also depend on the y_j values of the data, obtained either from experiment or from a generator with index n , if any. The MSE is a good indicator of the score, as it captures the bias vs variance trade-off often present in predictive models across the fields of statistics and machine learning [133]. Finally, if we are dealing with pseudo data from multiple generators, we define the squared average of the MSE by combining the predictions from the n different generators:

$$\langle \text{MSE} \rangle^2(m, \mathbf{q}, \boldsymbol{\sigma}) = \frac{1}{N} \sum_{n=1}^N \text{MSE}^2(m, \mathbf{q}, \boldsymbol{\sigma}, n). \quad (2.28)$$

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

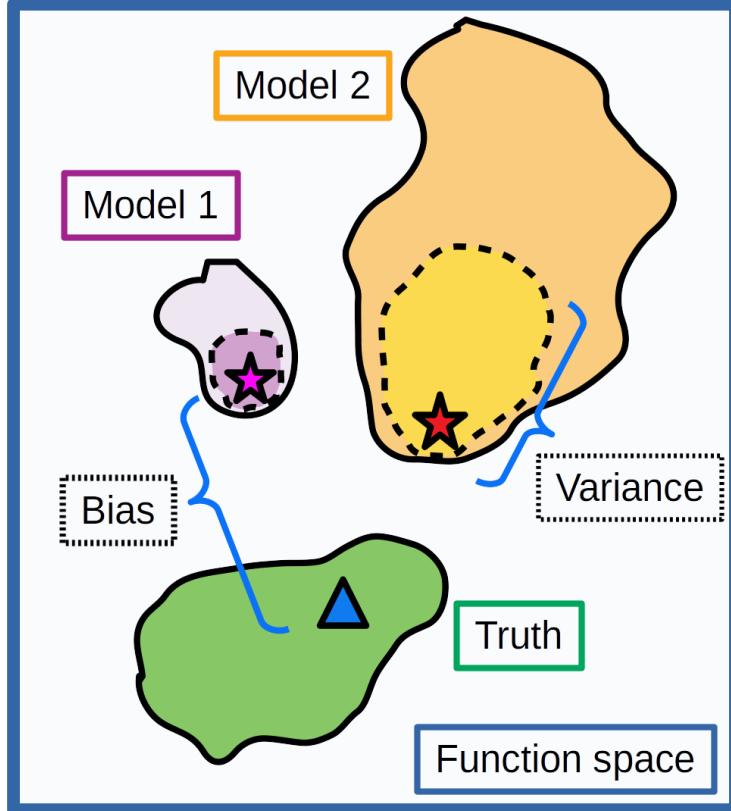


Figure 2.4: Abstract representation of the impact of bias and variance on recovering information. The entire function space is represented by the blue enclosing square. The green blob represents the collections of all the truths (generators) while the blue triangle is one of its members. The purple and yellow blobs represent 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.

An abstract representation of these concepts is illustrated in Fig.2.4. 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 N generators, with the blue triangle within this region representing a single member of such family (for example, FSU Garnet [96] in Sec. 5.2). If we are dealing with experimental data, the truth region will just contain a single member. 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 ω . 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.(2.28). 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 several noisy realizations of the data to accumulate enough statistics and then apply the standard fitting protocol described in Sec.2.2.1, as done in [61] . An alternative approach would be to directly compute the Bayesian integrals highlighted in Sec.2.2.2. In the following chapter, we present a third option: a new formalism that—under certain assumptions—speeds up these calculations, can aid in the identification of “critical” points in the data, provides a highly intuitive picture of the propagation of the uncertainty, and can be extended from model selection to model building. We call this option *the transfer function formalism*.

CHAPTER 3

THE TRANSFER FUNCTION FORMALISM

This chapter introduces the Transfer Function (TF) formalism. Sec. 3.1 defines the transfer functions and shows their analytical expressions. Sec. 3.2 and Sec. 3.3 illustrate how to calculate the bias and variance within the TF formalism, while Sec. 3.4 combines both calculations to form the MSE. Finally, Sec. 3.5 describes the implementation of Bayesian priors under this formalism.

3.1 Transfer Functions: Definition

We want to understand how the uncertainty -both in terms of bias and variance- gets propagated from the data \mathbf{Y} to the quantity of interest m . To do so, we borrow the idea of transfer functions from signal processing and control theory [98]: if we can make a linear map to connect an arbitrary input change in our system to the corresponding output change, then analyzing the dynamic response of the system becomes a straightforward process¹.

In our case, the “system” is the standard χ^2 fit in which the input are the data \mathbf{Y} and the output could be either the parameters $\boldsymbol{\omega}$ or any quantity m . Under the transfer function formalism we will assume that, once the minimum $\boldsymbol{\omega}_0$ of $\chi^2(\boldsymbol{\omega})$ is found, small changes in the data values y_j will produce proportional (linear) changes in the parameters and in any m . Then, our first objective is to write:

$$\delta m = \sum_{j=1}^J \mathcal{T}\mathcal{F}_j^m \delta y_j, \quad (3.1)$$

where δm is this small change in the fitted quantity m when each data point observation y_j is altered by a small amount δy_j . The transfer functions (TF) are the objects connecting the input change δy_j and the output change δm , denoted by $\mathcal{T}\mathcal{F}_j^m$. This notation identifies the j th input observation y_j in the sub index and the affected quantity m as a super index. We can explicitly

¹For nonlinear systems such map is not possible. Nevertheless, if input changes are small, linearizing the system around its equilibrium point might suffice for practical purposes [98].

expand \mathcal{TF}_j^m as:

$$\begin{aligned}\mathcal{TF}_j^m &\equiv \frac{\partial m}{\partial y_j} = \sum_{k=1}^K \frac{\partial m}{\partial \omega_k} \frac{\partial \omega_k}{\partial y_j} = \nabla m \cdot \mathcal{TF}_j^\omega, \\ \mathcal{TF}_j^\omega &= \{\mathcal{TF}_j^{\omega_1}, \mathcal{TF}_j^{\omega_2}, \dots, \mathcal{TF}_j^{\omega_K}\},\end{aligned}\quad (3.2)$$

where $\mathcal{TF}_j^{\omega_k}$ are the transfer functions connecting a small change in each observation y_j with a response in parameter ω_k :

$$\delta \omega_k = \sum_{j=1}^J \mathcal{TF}_j^{\omega_k} \delta y_j. \quad (3.3)$$

The general expression for \mathcal{TF}_j^ω is:

$$\mathcal{TF}_j^\omega = \mathcal{H}^{-1} \nabla F(q_j, \omega) \sigma_j^{-2}, \quad (3.4)$$

where the gradient $\nabla F(q_j, \omega)$ is taken with respect to the parameters ω . Appendix B.1 shows the derivation of Eq. (3.4). We could then interpret that \mathcal{TF}_j^m is the “sandwich”, or covariance, between the linear response of our quantity of interest ∇m , and the linear response of our model at the observation location $\nabla F(q_j, \omega) \sigma_j^{-2}$, through \mathcal{H}^{-1} :

$$\mathcal{TF}_j^m = \nabla m \mathcal{H}^{-1} \nabla F(q_j, \omega) \sigma_j^{-2}. \quad (3.5)$$

The core idea of the transfer function formalism is that for small perturbations in the input of a system, its response 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). In the statistics front, we have found several related concepts such as the adjoint method (page 203 [134]), the influence functions (page 45 [135]), and the sensitivity of the system response (Sec. III F in [136]), for example. However, despite our best efforts, we could not find a direct application -like the one we developed in this work- to model selection and the description of both bias and variance.

In the following sections, we use the transfer functions to calculate both the bias and variance in our quantity m .

3.2 Variance calculation

Eq (3.1) allows to write *any* small change in the quantity δm for a given set of small changes in the observations $\delta \mathbf{y}$. We can interpret the errors in the data σ_j as the standard deviation of the

random independent Gaussian variables centered at each y_j . In this scenario, the perturbations δy_j can be considered to be random independent Gaussian variables with mean 0 and standard deviation σ_j^2 . This allows to calculate the variance in δm by adding each term in Eq. (3.1) in quadrature:

$$\Delta m^2 = \sum_{j=1}^J \left[\mathcal{T}\mathcal{F}_j^m \right]^2 \sigma_j^2, \quad (3.6)$$

where the $\mathcal{T}\mathcal{F}_j^m$ are evaluated in the parameters obtained when all the δy are zero, i.e., the parameters associated with the original data points \mathbf{Y} .

One of the main advantages of this expression is that it separates –up to some degree³– the contribution that each observation y_j is making to the total variance Δm^2 . This separation, as shown in Sec. 5.2.3, allows to identify those data points having an excessive influence in the variance. This information could help in experimental design to distribute resources such as beam time in the case of scattering experiments (see Sec IV in [97] for example).

The variance calculated in Eq (3.6) can be compared to the variance obtained from the standard approach in Eq (2.19). When the nonlinear part of the Hessian matrix in Eq (2.18) (the terms proportional to second derivatives of F) can be ignored, these two variances are identical. In all other cases, they will differ. Appendix B.1 presents a proof of this statement.

When the calculated variances differ, the result of the transfer functions (Eq (3.6)) seems to be more in agreement with what is obtained in analyses such as [61], where the data are sampled from some central values in a Monte Carlo approach. For a better understanding of the discrepancy with the usual approach, see Theorem 5.23 and example 5.25 in [137].

The accuracy of these two approaches will deteriorate when the errors in the data are so big that the nonlinear part of model F becomes more important. In this case, the Gaussian approximation, i.e., that the entire χ^2 landscape can be described by the second derivatives at the minimum, is no longer valid.

As a final remark, we comment on the fact that the calculated variance from Eq. (3.6) will change depending on where we set the locations q_j . This change happens not only because the

²Eq (3.1) can still be used to calculate the variance in a more general case when there are either correlations or the distribution is not Gaussian. It would be necessary to perform the appropriate integrals on $\delta m(\delta y)$ times the joint probability distribution $P(\delta y)$.

³Since the Hessian \mathcal{H}^{-1} in $\mathcal{T}\mathcal{F}_j^m$ effectively mixes all observations, it is impossible to truly isolate the contribution from each data point. Nevertheless, through Eq. (3.6) it is easier to address the influence of each single data point than with the other approaches, namely Eqs (2.19) and (2.21a).

errors σ_j might vary with the locations q_j , but also because the transfer functions themselves will have different values at different q_j . By exploring the entire available q range, we could find the optimal locations that minimize the total variance in our quantity of interest which is a way of answering the question: *where should we measure to minimize the (statistical) uncertainty?*. When exploring the q range, we must also be aware that the fitted parameters ω will change as well, which will also impact the numerical value of each \mathcal{TF}_j^m (see Eq. (3.4)). Therefore, we should perform a fit every time we want to try a new set of q_j . One of the results discussed in the following section, is that under certain assumptions we can actually skip that re-fitting step.

3.3 Bias calculation and the Central Function

In this section, we study the bias as explained in Sec. 2.2.3, i.e., the discrepancy between the true value of m and the value extracted by the model. A traditional way of calculating the bias would be to either fit our parameters to the data $\mathbf{Y} = \{(q_j, y_j, \sigma_j)\}$ and either calculate $m(\omega_0)$ (by minimizing Eq. (2.16)), or $\langle m \rangle$ in Eq. (2.21a) and then subtract it to m_{true} . In either case, parallel with our discussion of the variance, we have to re-fit the function's parameters or perform the integrals in Eq. (2.21a) for every combination of points q_j we want to test. The reason to explore the bias and variance behavior as we change the q_j locations is that we might be interested in finding the optimal places that minimize the MSE in Eq. (2.27). As we will see, once we cast the bias calculation under the TF framework, it is possible to avoid performing a re-fit for every different set of q_j locations.

The sources of the variance in Eq. (3.6) are the individual data errors σ_j , which get propagated through the transfer functions. To write a bias equation in terms of the TF we must identify what are the sources equivalent to σ_j for each location q_j . To do so, let us study first how the fitted parameters (by minimizing χ^2) change when we move the q_j locations. We will refer to these fitted parameters as the “empirical” parameters ω_e . Note that we employ the specific notation ω_e , rather than the more general ω_0 defined after Eq.(2.16). The former refers exclusively to an ω obtained directly from data (or pseudo data), while the latter represents the minimum of χ^2 in any situation, even when we perturb the data by small amounts δy .

Fig. 3.1 shows an example of how ω_e in the parameter space changes as the second location q_2 moves. The model being fit is the Symmetrized Fermi function of two parameters ($\omega = [c, a]$) (see

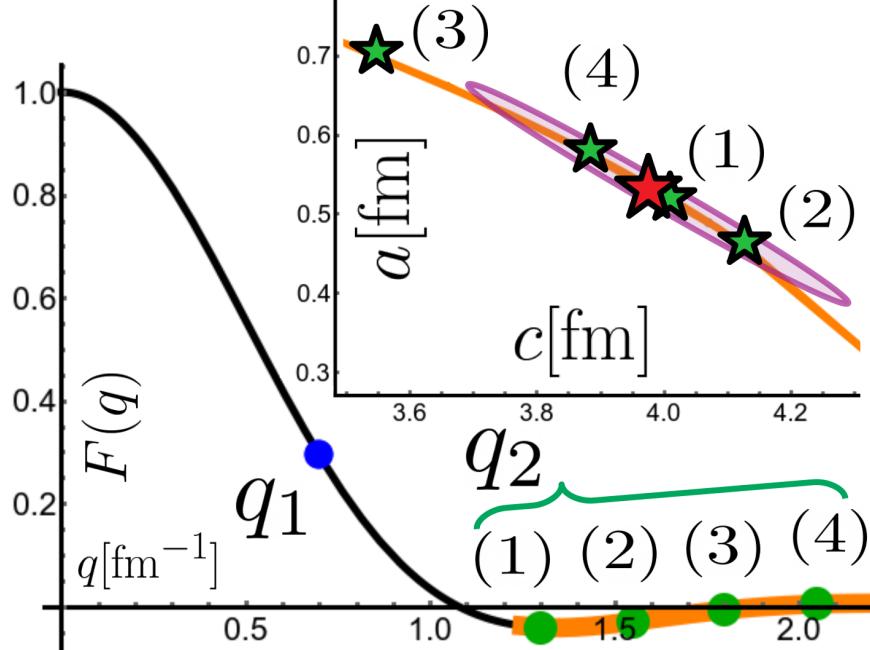


Figure 3.1: Evolution of ω_e as the second location q_2 is moved in the orange region. The model being fit is the Symmetrized Fermi function of two parameters: $\omega = [c, a]$. The inset plot shows how the orange curve gets mapped into the parameter space, highlighting four locations in green. Although it is difficult to note in the plot, the mapping “folds into itself” in the parameter space. The red star represents the central parameters (see the text preceding Eq. (3.7)), while the purple ellipse represents the 95% confidence interval for a fit using $\mathbf{q} = \{0.7, 1.8\}$ fm $^{-1}$ but with y values dictated by the central function (see the text following Eq. (3.7) and in Fig 3.2)

Sec. 5.1.1 or Appendix C.1). There are two observations, $q_1 = 0.7$ fm $^{-1}$, and q_2 which moves in the orange curve with range [1.25, 2.2] fm $^{-1}$. Both of these point’s y_j values are determined by the ^{48}Ca weak form factor created by the generator RMF012 (see Sec. 2.2.3). In the orange curve we highlight four locations labeled 1, 2, 3, and 4, displayed as green circles. Each of these locations, in combination with q_1 , will correspond to a value of ω_e , shown as green stars in the parameter space in the inset plot. We expect the value of any quantity m computed at the different q_2 locations to change, and this, in turn, will change the value of the bias of the model.

As can be seen in Fig. 3.1, the empirical parameters ω_e (green stars), do not move too far away from some *central* value ω_c (shown as a red star). From this central position and using the transfer functions, we can describe the *trajectory* of the empirical parameters and how they deviate from

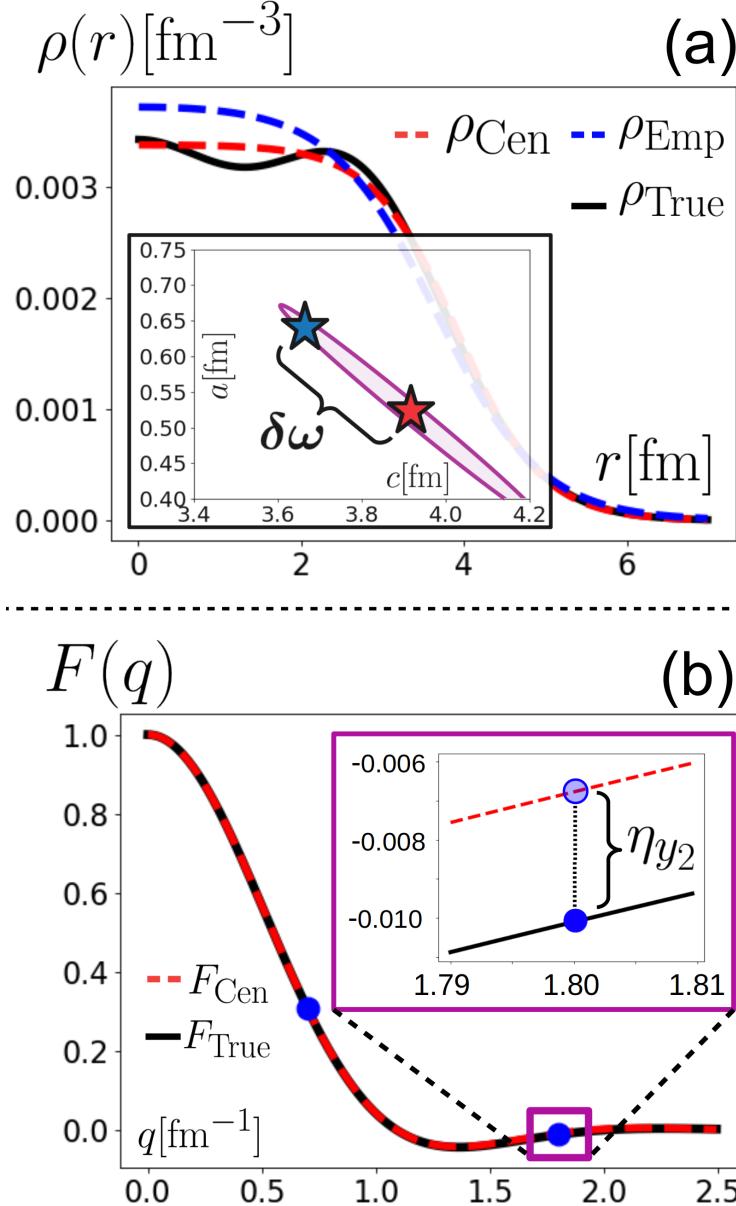


Figure 3.2: Example of the change in the central parameters driven by deviations between the central and the true form factors. (a) Reconstructed density from ω_c , reconstructed density from ω_e using Eq(3.7), and the true density (red, blue, and black lines, respectively). The inset plot shows the locations in the parameter space of ω_c and the estimated ω_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 5.2.2. (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 (η_2) in the neighborhood of $q_2 = 1.8 \text{ fm}^{-1}$.

the central value ω_c to first order $\omega_c + \delta\omega \approx \omega_e$ for a given set of J observations:

$$\delta\omega = \sum_{j=1}^J \mathcal{T}\mathcal{F}_j^\omega \eta_j, \quad (3.7)$$

$$\eta_j \equiv y_j - F(q_j, \omega_c). \quad (3.8)$$

We call η_j the quantity that is now driving the change in the parameters to distinguish it from an arbitrary perturbation δy_j . The transfer functions in Eq (3.7) are evaluated at the central parameters ω_c and in the “data” created by $F(q, \omega_c)$. The main idea we are exploiting is that the minimum of a χ_c^2 defined as:

$$\chi_c^2(\omega) = \sum_{j=1}^J \frac{(F(q_j, \omega) - F(q_j, \omega_c))^2}{\sigma_j^2}, \quad (3.9)$$

is $\omega_c = \text{argmin}(\chi_c^2)$, since $\chi_c^2(\omega_c) = 0$. This expression for χ_c^2 is identical to the one defined in Eq. (2.16), but with the real observations y_j replaced by $F(q_j, \omega_c)$. From Eq. (3.7), we can say that if the $F(q_j, \omega_c)$ are perturbed in such a way that $F(q_j, \omega_c) \rightarrow F(q_j, \omega_c) + \eta_j = y_j$, the central parameters ω_c will move a certain distance such that $\omega_c \rightarrow \omega_c + \delta\omega \approx \omega_e$.

If this last approximation, $\omega_c + \delta\omega \approx \omega_e$, is enough for our purposes, we can say that m_e , i.e., the empirical value of m obtained with the empirical parameters from the actual data, can be approximated by the central value m_c obtained with the central parameters, plus a small correction δm :

$$m_e = m_c + \delta m = m_c + \sum_{j=1}^J \mathcal{T}\mathcal{F}_j^m \eta_j. \quad (3.10)$$

With these tools, we can write the bias term for m as:

$$\text{Bias } (m) \equiv m_e - m_t = \left[m_c + \sum_{j=1}^J \mathcal{T}\mathcal{F}_j^m \eta_j \right] - m_t, \quad (3.11)$$

where m_t is the true value of m and the $\mathcal{T}\mathcal{F}_j^m$ are evaluated on the central parameters ω_c .

Using the same model and generator as in Fig. 3.1, Fig. 3.2 shows an example of this bias calculation where our quantity m is the interior density. As in Fig. 3.1, the first observation is at $q_1 = 0.7 \text{ fm}^{-1}$, while the second one at $q_2 = 1.8 \text{ fm}^{-1}$ (the third green point in Fig. 3.1). Using Eq. (3.7) we approximate the empirical parameters and thus, the empirical density $\rho(r)_{\text{Emp}}$, which is shown as a blue dashed line in Fig. 3.2 (a).

The deviation of this approximated empirical density from the central density $\rho(r)_{\text{Cen}}$ can be understood in terms of the η_j 's: the difference between the central form factor $F(q, \omega_c)$, and the true form factor $F(q)_{\text{True}}$ values at q_1 and q_2 . Since these are tiny deviations, we enlarge a window around q_2 and show η_2 in the inset plot of Fig. 3.2 (b). The hollow blue circle corresponds to $F(q_2, \omega_c)$, while the filled blue circle corresponds to $F(q_2)_{\text{True}}$.

Under the linear approximation (3.7), these η_j will move the central ω_c (red star in the inset of Figure (a)) towards the approximated empirical parameters ω_e (blue star). Since this is a nonlinear model, Eq (3.7) 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. 3.2 with the green star (3) in Fig. 3.1. Nevertheless, the respective empirical density is not too different from the one shown in Fig. 3.2 (a).

The precise location of ω_c should not have a big impact on our calculations if the approximation that the actual change $\delta\omega$ is linear in the η_j holds. Given that we measure in the form factor space although we are interested in the density space, we distinguish between two main options for choosing ω_c :

1) Central Function Fit: We define ω_c as the value that minimizes the L2 norm between our model $F(q, \omega)$ and the true function $F(q)_{\text{true}}$ in the momentum transfer space q . This is the case in Figures 3.1 and 3.2. The expectation is that ω_c should be relatively close to most of the possible obtainable parameters for different locations of the data. We refer to these parameters as ω_{Cen} .

2) Optimal Fit: We define ω_c as the parameters that make the central estimation m_c as close as possible to the true value m_t . For example, if we are interested in modeling the interior density, ω_c should be chosen by fitting our models directly in the density space $\rho(r)$ ⁴. The advantage of this option is that, if $m_c - m_t$ is indeed negligible, the total bias is dominated by the η_j , making the analysis of optimal locations easier. We refer to these parameters as ω_{Opt} , and will use them extensively in Sec 5.2.1.

⁴This procedure is, of course, not physically possible in the case of real data since scattering experiments can only access the form factor directly, not the density.

3.4 Mean Squared Error

Having constructed the bias and variance within the TF formalism, we write the Mean Squared Error (MSE) as:

$$\text{MSE}^2 = \left((m_c - m_t) + \sum_{j=1}^J [\mathcal{T}\mathcal{F}_j^{(m)}] \eta_j \right)^2 + \sum_{j=1}^J [\mathcal{T}\mathcal{F}_j^{(m)}]^2 \sigma_j^2. \quad (3.12)$$

The value of the MSE will naturally depend on the quantity m in which we are interested, as well as on the data points locations $\mathbf{q} = \{q_j\}$ and errors $\boldsymbol{\sigma} = \{\sigma_j\}$. It will also depend on the input y values, which could either be an experimental data set or pseudo data generated by mean field models.

This equation allows us to address the expected MSE for a given set of already observed data and analyze which model should have the lowest error as the study in [61]. However, if the entirety of the data have not been measured yet, this equation can be used to optimize the MSE not only with respect to the model, but with respect to the data locations and errors.

If we are working with pseudo data, for a fixed model each generator will be associated with a different set of central parameters $\boldsymbol{\omega}_c$. Therefore, unless we suspect this variation is negligible, in Eq (2.28) each MSE should be calculated with its own parameters $\boldsymbol{\omega}_c$. These parameters $\boldsymbol{\omega}_c$ are necessary for the numerical calculation of each $\mathcal{T}\mathcal{F}^m$.

Furthermore, it is important to note that the $\mathcal{T}\mathcal{F}_j^m$ from the bias term in Eq (3.12) are evaluated at the central parameters $\boldsymbol{\omega}_c$, while the $\mathcal{T}\mathcal{F}_j^m$ associated with the variance are not. From our discussion in Sec 3.2, these $\mathcal{T}\mathcal{F}_j^m$ should be evaluated at the parameters associated with the observed data (the empirical parameters $\boldsymbol{\omega}_e$ defined in Sec 3.3).

There are two options to obtain $\boldsymbol{\omega}_e$. If we believe that the linear relationship (3.7) is a good approximation, we can obtain $\boldsymbol{\omega}_e$ from that equation while we calculate the bias. If we believe that this approximation might not be enough (we are dealing with a strongly non-linear model), the parameters should be fitted directly with a numerical algorithm from 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 (3.11).

To guarantee numerical accuracy, we use this last option for all the nonlinear models we explore in Sec. 5.2, while we use Eq (3.11) for the linear models. For example, in the case shown in Fig. 3.2,

the calculated change in $m = \rho(0)$ from the central value using the TF, underestimates the true change by around 30%.

In Sec. 5.2.1 and Sec. 5.2.5 we describe two important application of Eq (3.11) that would not be possible with a numerical fitter. These applications can be very useful even when dealing with nonlinear models.

3.5 Priors under the TF formalism

If we have Gaussian priors of the form presented in Eq (2.23a), then we can treat each prior term as a pseudo observation. These priors act in the same way as true observations in χ^2 : $(F(\boldsymbol{\omega}, q_j) - y_j)/(\sigma_j^2)$, by pulling the value of $\boldsymbol{\omega}$ in a particular direction in the parameter space. The new $\tilde{\mathcal{H}}$ defined in Eq (2.25), should be used when calculating the observation's transfer functions defined in Eq (3.4).

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 ω_k^0 will have its own transfer function as if it were an observation:

$$\mathcal{T}\mathcal{F}_k^\omega \equiv \frac{\partial \boldsymbol{\omega}}{\partial \omega_k^0} = \tilde{\mathcal{H}}^{-1} I_k \sigma_k^{-2}, \quad (3.13)$$

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). Let us recall that K denotes the total number of parameters of the model, while k acts as an index in that particular parameter set. $I_k \sigma_k^{-2}$ is the analogous of $\nabla F(\boldsymbol{\omega}, q_j) \sigma_j^{-2}$ when calculating $\mathcal{T}\mathcal{F}_j^m$. We use the sub index k to denote that what we are perturbing is not y_j , but rather the prior estimate value ω_k^0 . In the case, where the prior contains correlations then Eq (2.23b) is written as quadratic form $\phi^2(\boldsymbol{\omega}) = \boldsymbol{\omega} \Sigma^{-1} \boldsymbol{\omega}$. In this case $I_k \sigma_k^{-2}$ will be replaced by the k -th column of the matrix Σ^{-1} .

These transfer functions of the priors “observations” will appear at the same level as regular observations in the variance and bias equations (3.6) and (3.11). For the bias part, the associated η_j -which we will call $\tilde{\eta}_k$ - is defined as the difference between the value of $[\boldsymbol{\omega}_c]_k$ (the k entry of the central parameters) and the prior “observation” ω_k^0 .

CHAPTER 4

THE PROTON RADIUS

In this chapter, we focus our attention on the decade old proton puzzle [11, 12, 13] discussed in the introduction. Let us recall that the puzzle arose as a 7σ discrepancy between the combined estimation of the proton radius from electronic scattering experiments and electronic spectroscopy (see [9] for the accepted 2010-CODATA value based on these experiments), with the extremely precise value determined from muonic hydrogen [10, 11].

Within the broad context of explanations that have been proposed to explain the puzzle, our contribution in this chapter is rather modest, as our main goal is to address how best to extract the charge radius of the proton from previous electron scattering data. The view we adopt here is that, from the electron scattering side, the puzzle was not in the experimental data, but rather in the extraction of the proton radius from the scattering data.

The first section of this chapter presents a follow-up analysis we performed on the results in [61], focusing on the bias vs variance trade-off, discussed in Sec. 2.2.3, applied to linear vs quadratic extrapolation on the proton’s form factor pseudo data. Using a numerical version (as an early stage) of the transfer function formalism described in Chapter 3, we also discuss how the locations of the data can greatly impact the uncertainty in the obtained radius.

The second part of this chapter focuses to our nonparametric analysis on the Mainz data [109]. The recent analyses of electron scattering data that agree with a smaller proton radius have been called into question [138]. Moreover, much controversy has been generated around the optimal (“parametric”) model that should be used to fit the electric charge form factor of the proton—ranging from monopole, to dipole, to polynomial fits, to Pade’ approximants, among many others. There has been also several attempts in performing extractions that rely on analytical properties of the form factor, see for example [139, 140].

In an effort to eliminate the reliance on specific functional forms, we introduce a method that does not assume a particular parametric form for the form factor. Such a nonparametric approach aims to “let the data set speak for itself” without introducing any preconceived biases. Although the

nonparametric approach does not assume a particular form for the form factor, several constraints justified by physical considerations are imposed. In essence, a *nonparametric Bayesian* curve fitting procedure that incorporates various physical constraints is used to provide robust predictions and uncertainty estimates for the charge radius of the proton.

4.1 Parametric model selection: Bias vs Variance trade-off

A general criticism of the small radius extractions of the proton radius is the presence of statistical bias [141, 142], with an implication that bias needs to be avoided in order to successfully extract the true radius from the data. The use of Monte Carlo methods to find bias and then reject simple proton radius extraction methods originated with the classic Monte Carlo study of Borkowski *et al.* [52] where linear extrapolations were flatly rejected in favor of quadratic extrapolations. Interestingly, that work ignored the variance of the more complex function.

We show in this chapter that when using a Monte Carlo study to test a model's ability to extract the proton radius, we need to consider not only bias but also variance, and find an appropriate balance between the two. After all, it is better to have a slightly biased watch to tell the hour, than a broken one that is unbiased because it overestimates and underestimates the time symmetrically.

We will also illustrate that one must consider the range, quantity and precision of the data when determining the best predictive statistical model and show that simple biased statistical models can have a higher predictive validity than unbiased more complex models [143].

4.1.1 Bias

In everyday language, bias is often used as a pejorative term. In the context of regression, it is simply an offset of the central estimated value from the true central value, as we discussed in Sec. 2.2.3. If the central value is obtained as the mean of a distribution (for example as in Eq. (2.21a), in contrast to Eq. (2.17)), then it is not a property of a single realization but can be determined by repeated sampling or integration. In the context of the proton radius extractions, bias was nicely illustrated by Borkowski *et al.* [52] and we will describe their procedure in the following paragraphs.

interval fm ⁻²	linear fit		quadratic fit	
	a_0	Radius [fm]	a_0	Radius [fm]
0.1 – 0.4	1.000	0.79	1.000	0.81
0.1 – 0.8	0.999	0.78	1.000	0.81
0.1 – 1.2	0.997	0.77	1.000	0.81
0.1 – 1.6	0.996	0.76	1.000	0.81

Table 4.1: The mean a_0 and radius from doing 10^6 Monte Carlo simulations for each interval in Q^2 where Eq. (4.1) was used to generate pseudo data in 0.05 fm^{-2} steps with each data point smeared by a randomly generated, normally distributed point-to-point uncertainty of 0.5%. The results clearly indicate that the linear fits are biased. The input radius was 0.8113 fm (an a_1/a_0 term of 0.1097 fm^{-1}) and $a_0 = 1$.

Form factor pseudo-data are first systematically generated from 0.1 fm^{-2} to 0.4 , 0.8 , 1.2 , and 1.6 fm^{-2} in steps of 0.05 fm^{-2} using the standard dipole function:

$$G_D(Q^2) = (1 + Q^2/(18.23\text{ fm}^{-2}))^{-2}, \quad (4.1)$$

where the cutoff parameter of 18.23 fm^{-2} corresponds to a radius of 0.8113 fm .

Next, to mimic real data, the pseudo data points are each randomly shifted using a normal distribution with mean zero and a sigma of 0.5%. Then the pseudo data sets are fit using both linear and quadratic functions:

$$f_{\text{linear}}(Q^2) = a_0 + a_1 Q^2, \quad (4.2)$$

$$f_{\text{quadratic}}(Q^2) = a_0 + a_1 Q^2 + a_2 Q^4. \quad (4.3)$$

These functions are written so that they are linear in the fit coefficients, and this allows the χ^2 minimization to be performed exactly while also allowing the normalization to float. To obtain the physical $G_E(0) = 1$ behavior, one simply divides these functions by the normalization term, a_0 , to find the slope of $G_E(0)$ given by a_1/a_0 which can be used in Eq. (2.4) to determine the radius.

This procedure was repeated with 10^6 sets of pseudo data to precisely determine the mean of the extracted radii for these two functions. Since the standard dipole was used as the input function, one would expect an unbiased function to return a radius of 0.8113 fm . Table 4.1 reproduces the original result [52]. As the table shows, the mean values of a_0 and r_p show a clear bias. Based on this study, the authors of the original work erroneously concluded that the linear models should always be rejected in favor of the lower-bias quadratic function.

Data Points	Interval fm ⁻²	linear fit						quadratic fit					
		a_0	Radius [fm]	a_1/a_0 [fm ²]	Bias [fm ²]	σ [fm ²]	MSE [fm ²]	a_0	Radius [fm]	a_1/a_0 [fm ²]	Bias [fm ²]	σ [fm ²]	MSE [fm ²]
7	0.1 – 0.4	1.000	0.795	-0.105	-0.004	0.018	0.019	1.000	0.807	-0.108	-0.001	0.109	0.109
15	0.1 – 0.8	0.999	0.783	-0.102	-0.008	0.006	0.010	1.000	0.809	-0.109	-0.001	0.028	0.028
22	0.1 – 1.2	0.998	0.771	-0.099	-0.011	0.003	0.011	1.000	0.809	-0.109	-0.001	0.014	0.014
31	0.1 – 1.6	0.996	0.760	-0.096	-0.013	0.002	0.014	1.000	0.808	-0.109	-0.001	0.009	0.009

Table 4.2: An expanded version of Table 4.1 where instead of just showing the mean offset of the fit results for a_1/a_0 , the bias, we also indicate the width of the fit results, σ . The point uncertainty is fixed at 0.5%. For the simulated radius of 0.8113 fm, one would expect an unbias fit to give a_1/a_0 of 0.11 fm²; thus, the difference of that value from the mean fitted value of a_1/a_0 is the bias and the width of the distribution is σ . Also shown is the mean square error, MSE, which can be used to quantify the best function for a given interval, taking into account both bias and variance.

Data Points	Interval fm ⁻²	linear fit						quadratic fit					
		a_0	Radius [fm]	a_1/a_0 [fm ²]	Bias [fm ²]	σ [fm ²]	MSE [fm ²]	a_0	Radius [fm]	a_1/a_0 [fm ²]	Bias [fm ²]	σ [fm ²]	MSE [fm ²]
31	0.1 – 0.4	1.000	0.795	-0.105	-0.004	0.010	0.011	1.000	0.809	-0.109	-0.001	0.063	0.063
31	0.1 – 0.8	0.999	0.783	-0.102	-0.007	0.004	0.009	1.000	0.810	-0.109	-0.000	0.021	0.021
31	0.1 – 1.2	0.997	0.771	-0.099	-0.011	0.003	0.011	1.000	0.809	-0.109	-0.001	0.012	0.012
31	0.1 – 1.6	0.996	0.760	-0.096	-0.013	0.002	0.014	1.000	0.808	-0.109	-0.001	0.009	0.009

Table 4.3: Same as Table 4.2, but now with equal number of data points for each range.

4.1.2 Variance

While the linear fit does exhibit a bias, bias is not the only quantity that must be considered when selecting an appropriate statistical model. In particular, along with the offset from the true value, the bias, one must also consider the distribution of the outcomes, the variance, where variance is the square of the standard deviation, σ , of the statistical distribution. Table 4.2 shows a more complete picture of the simulation results where the σ of the results is shown along with the bias. These results are graphically represented in Fig. 4.1.

For all Q^2 intervals, the linear fits provide significantly smaller σ than the more complex quadratic fits. Therefore, picking the function to use based solely on variance would also be incorrect. This seemingly simple example has turned into a nearly textbook illustration of the trade-off between variance and bias with the linear fits having a relatively high bias with a low variance, while the quadratic fits have a low bias and high variance. Of course, to calculate the bias we need to know the true value which often is not the case in a real experiment.

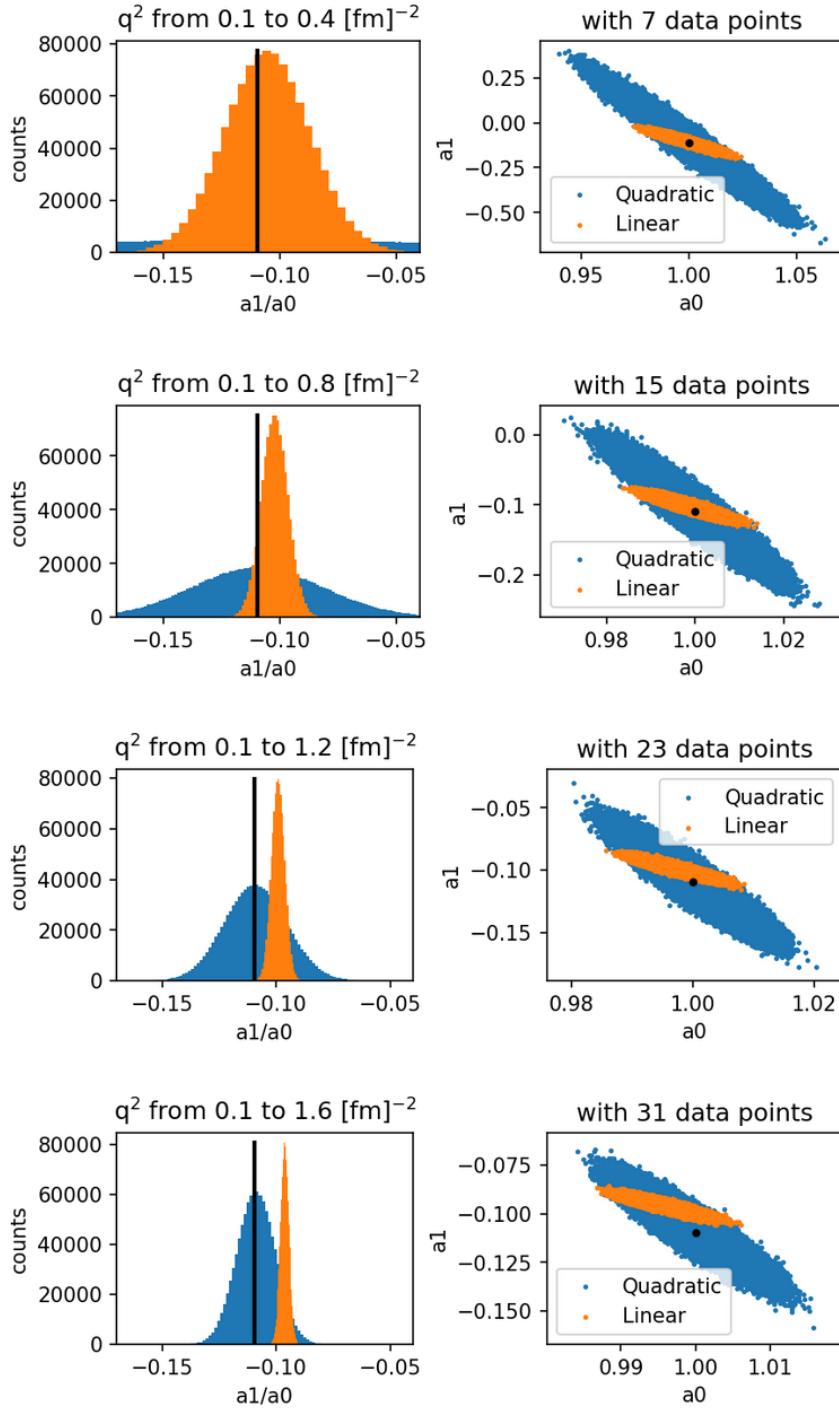


Figure 4.1: A graphic representation of the Monte Carlo results showing how the linear fits tend to have a relatively high bias though a low variance, while the quadratic fits tend to have a relatively low bias but a large variance. The black line/point represents the true value.

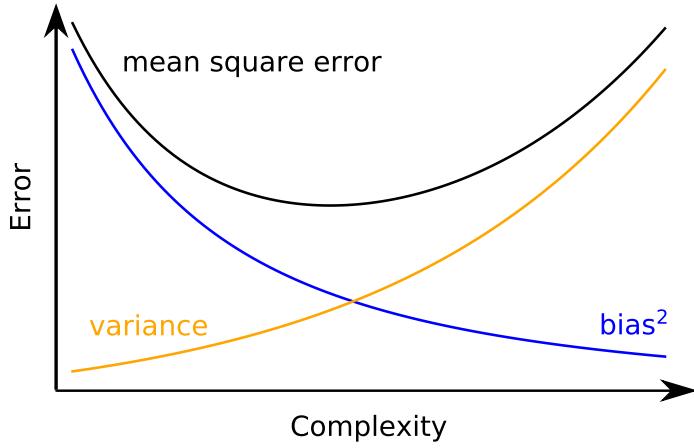


Figure 4.2: An illustration of the trade-off between bias and variance when selecting a statistical model. Simple models will have low variance but high bias (under-fitting) while complex models will have low bias but high variance (over-fitting). It is this trade-off that one seeks to balance. While with repeated Monte Carlo simulations it is trivial to find the optimal predictive model for a given set of data, in the real world the true model is typically unknown. We only get to perform a very limited number of experiments and thus, we rely on using real data and statistical methods for model selection [145].

4.1.3 Bias and variance combined: MSE

As was noted by George Box, all models are wrong, thus, the goal is to find the most useful one [144]. As discussed in Sec. 2.2.3, as the complexity of a model increases (for example by increasing the number of basis functions it uses), the expected bias tends to decrease while the variance increase. For any given set of statistical models, the goal is to find the ones with the optimal balance between bias and variance, as illustrated in Fig. 4.2. Thus, to quantify the goodness of the fits, we choose the Mean Square Error (MSE) we defined in Eq. (2.27).

Using the MSE values in Table 4.2, one can now quantify that for this example the $0.1\text{--}0.8 \text{ fm}^{-2}$ interval is the preferred range for the linear model while the $0.1\text{--}1.6 \text{ fm}^{-2}$ interval is the preferred range for the quadratic model.

It is interesting to repeat the Monte Carlo simulation for equal number of data points within each range particularly because for elastic scattering, cross sections are significantly higher at lower values of Q^2 and thus, it is easy to obtain more low Q^2 data. This is shown in Table 4.3 and now the picture is even grayer as the MSE of the linear fit is nearly equal to the quadratic. Therefore, assuming that the standard dipole was the true generating function, an experiment with

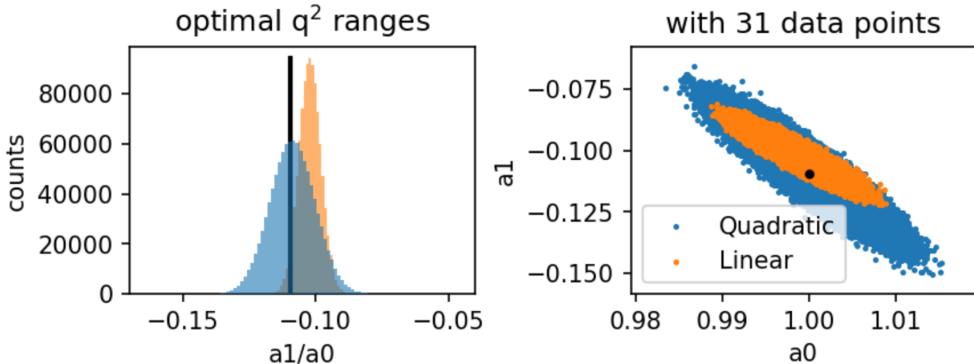


Figure 4.3: The result of a million simulations and fits of linear fits over the Q^2 range of $0.1 - 0.8 \text{ fm}^{-2}$ and quadratic fits over $0.1 - 1.6 \text{ fm}^{-2}$, both with 31 uniformly spaced data points. Using the mean square error as the metric, neither example is significantly better than the other for exacting the proton radius. This is analogous to a dart game between two equally skilled players: one who hits the bull's eye more often yet has a large spread (low bias but high variance), and another, equally skilled, player who has a tighter cluster of hits but an offset (high bias but low variance).

31 data points and an uncertainty of 0.005 per point over a range of 0.1 to 0.8 fm^{-2} and a different experiment over a range of 0.1 to 1.6 fm^{-2} would have an equal probability of reproducing the correct radii if all other things were equal (given that the modeler in the first group adjusts a line and the one in the second, a parabola). This is shown in Fig. 4.3 where the linear fit is clearly biased but has a small variance compared to the unbiased, large variance quadratic fit.

The choice of the parsimonious modeler to use the low Q^2 data would likely be driven by the recognition of the fact that as Q^2 increases the extraction of the charge form factor is complicated by the growing influence of the magnetic form factor. The choice to use a larger Q^2 range would likely be driven by a desire to form a more complete picture of the proton's structure. For example, the parsimonious modeler may only be interested in the proton radius while another modeler may be interested in higher order moments [146]. Thus, the tension in the extractions of the proton radius from electron scattering data could be about the fact that research groups using the low Q^2 are generally getting a systematically different result than the groups doing fits which include high Q^2 data. This perhaps points to a systematic problem with our knowledge of the magnetic form factor and/or the functional form of the form factors.

4.1.4 The Best Predictive Model

Selecting between a linear or quadratic regression of the more complex standard dipole function may seem a bit contrived, as one might naively think that just using the generating function itself would always yield the best results. In Sec. 4.1.5, we prove that for non linear fits there is an induced bias even when using the generating function as a model to fit the data. This induced bias grows proportional to the square of the errors in the data, and might only be noticeable if these errors are big enough.

If the variance is taken into account, we can also show that this is not always the case with a simple set-up: we use the lowest Q^2 range, $0.1 - 0.4 \text{ fm}^{-2}$, and replace the quadratic function with the generating function and a floating normalization term:

$$f_{\text{DipoleFit}}(Q^2) = n_0(1 - b_1 Q^2/2)^{-2}, \quad (4.4)$$

where n_0 is the normalization factor and the radius is given by $\sqrt{-6b_1}$. Pseudo data are generated for absolute random errors of 0.01, 0.005 and 0.003 with three different spacings: 0.05 fm^{-2} spacing with 7 points, 0.01 fm^{-2} spacing with 31 points, and 0.005 fm^{-2} spacing with 61 points. The results of fitting these pseudo data sets are shown in Table 4.4. For a given row in the Table, the linear fit has the greater bias and the dipole fit always has the greater variance; bringing the mean square error very close for all the test cases. This example highlights that it is not just the number of points that matter, but the size of the uncertainties and the range of the data that will also be critical parameters in model selection. In the cases where the MSE of the line was smaller than the MSE of the dipole, the line was a better “predictive” model for the proton radius, while the dipole was a better “descriptive” model of the form factor as a whole.

Of course, for real data, nature hides the true generating function from us, so perhaps it is reassuring to know that a reasonable approximation can reveal the underlying physics just as well as, if not better than, the true function. To be clear, the lesson is not that one function is better than another; it is that for a given set of data, the scientist is challenged to use the most appropriate model (either descriptive or predictive) for the task at hand. Further details on the general mathematics behind these example problems can be found in [143].

Data Points	Random Error	linear fit						dipole fit					
		a_0	Radius [fm]	a_1/a_0	Bias [fm 2]	σ [fm 2]	MSE [fm 2]	a_0	Radius [fm]	a_1/a_0	Bias [fm 2]	σ [fm 2]	MSE [fm 2]
7	0.01	0.9995	0.7941	-0.1051	-0.0046	0.0359	0.0361	1.0001	0.8109	-0.1096	-0.0001	0.0378	0.0378
7	0.005	0.9995	0.7945	-0.1052	-0.0045	0.0174	0.0180	1.0000	0.8113	-0.1097	0.0000	0.0194	0.0194
7	0.003	0.9996	0.7952	-0.1054	-0.0043	0.0108	0.0116	1.0000	0.8113	-0.1097	0.0000	0.0114	0.0114
31	0.01	0.9996	0.7945	-0.1052	-0.0045	0.0186	0.0191	1.0000	0.8113	-0.1097	0.0000	0.0207	0.0207
31	0.005	0.9995	0.7952	-0.1054	-0.0043	0.0093	0.0102	1.0000	0.8113	-0.1097	0.0000	0.0103	0.0103
31	0.003	0.9995	0.7952	-0.1054	-0.0043	0.0056	0.0070	1.0000	0.8113	-0.1097	0.0000	0.0062	0.0062
61	0.01	0.9995	0.7949	-0.1053	-0.0044	0.0135	0.0142	1.0000	0.8113	-0.1097	0.0000	0.0150	0.0150
61	0.005	0.9995	0.7952	-0.1054	-0.0043	0.0069	0.0081	1.0000	0.8113	-0.1097	0.0000	0.0073	0.0073
61	0.003	0.9995	0.7952	-0.1054	-0.0043	0.0040	0.0059	1.0000	0.8113	-0.1097	0.0000	0.0045	0.0045

Table 4.4: For the lowest Q^2 interval, 0.1 to 0.4 fm^{-2} , we compare regressions with a linear function (Eq. 4.2) to the dipole fit function (Eq. 4.4). Keeping the range fixed, a spacing of 0.05 fm^{-2} (7 points), 0.01 fm^{-2} (31 points) and 0.005 fm^{-2} (61 points) was used with various absolute random errors. In several of these cases, the simple linear function provides a better predictive model than the true functional form and is never far from the true function.

4.1.5 Semi-analytical Calculations as an alternative to Monte Carlo

What follows are calculations involving an early stage of the transfer functions techniques we developed in Sec. 3. We made these calculations taking the random errors realizations in the data, which we label ϵ_j , as the variables instead of the observations y_j themselves. Instead of writing an explicit expression for the transfer function from the data to the quantity of interest m (as in Eq. (3.5)), we leave it explicitly as $\partial m / \partial y_j$ (or in this case $\partial m / \partial \epsilon_j$), and estimate it numerically at the moment of performing the calculation. Although less precise than writing analytical derivatives, this approach could be necessary for more complex models in which, for example, the parameters are fitted through iteratively solving coupled differential equations and in which not even automatic differentiation is possible.

We use these semi-analytical procedure to calculate the bias and variance induced by the noise in our estimation of the slope at zero. This framework can be used to reproduce the graph shown in Fig. 4.1 without the repeated Monte Carlo procedure of creating noisy data sets and fitting the models. This method can give us explicitly the influence each data point is having on our bias and variance and, thus, can be used to identify the points that maximize the information gained.

We postulate that the bias in our estimation of the slope at zero of $G_E(Q^2)$ has two sources. The first one is due to the noise in the data, i.e., experimental errors, which we call σ , and will increase as the noise increases. The second is due to the possible fact that we are not using the “true” function to fit the data, as was shown in Table 4.4 when the line had a clear bias even when

the noise was very small. We therefore split the bias as:

$$bias = bias_\sigma + bias_0, \quad (4.5)$$

where $bias_\sigma$ denotes the bias that scales with the noise, while $bias_0$ will always be present even when $\sigma \rightarrow 0$, where σ represents the size of the noise. Note that we have ignored the first type of bias in our calculations in Chapter 3 since, as we shortly show, it scales as σ^2 .

We proceed to show how to calculate the first source, while the second can only be calculated if we know the “true” value of the slope. Let us define our quantity of interest, the slope at zero, as $m(X, Y, \epsilon)$, where X and Y are the lists of n data points and ϵ is a particular realization of the noise, which we assume to be Gaussian with mean zero and standard deviation σ common to every point. Here, $m(X, Y, \epsilon)$ can be a closed expression for the slope given the data, as is the case for linear models, or it can be treated as a numerical routine that returns the slope when fitting a non linear model, like the dipole.

4.1.6 Noise Bias

Following the spirit of the Monte Carlo study we want to find the average value of $m(X, Y, \epsilon)$, $\langle m(X, Y) \rangle$, once all the possible realizations of the noise have been taken into account, weighted correctly by their Gaussian distribution. The result is given by the integral:

$$\langle m(X, Y) \rangle \equiv \int P(\epsilon) m(X, Y, \epsilon) d\epsilon, \quad (4.6)$$

where $d\epsilon = d\epsilon_1 d\epsilon_2 \dots d\epsilon_n$, and $P(\epsilon)$ denotes the assumed Gaussian probability distribution for the noise given by:

$$P(\epsilon) = \left(\frac{1}{(\sqrt{2\pi}\sigma^2)} \right)^n e^{(\epsilon_1^2 + \epsilon_2^2 + \dots + \epsilon_n^2)/2\sigma^2}. \quad (4.7)$$

The integral in Eq. (4.6) is effectively taking into account all possible noise realizations, weighted by their corresponding probabilities.

Since in most cases we do not have an available expression for $m(X, Y, \epsilon)$, we can expand it using a multivariate Taylor expansion and evaluate each term directly under the integral sign in Eq. (4.6). The first three terms in this expansion are:

$$m(X, Y, \epsilon) \approx m_0 + \left[\vec{\nabla}_\epsilon m \right]_0 \cdot \vec{\epsilon} + \frac{1}{2} \vec{\epsilon}^T \cdot [\mathcal{H}_m]_0 \cdot \vec{\epsilon}, \quad (4.8)$$

where $m_0 = m(X, Y, \epsilon = 0)$ is the value of the slope when the noise is zero. We denote $\left[\vec{\nabla}_\epsilon m \right]_0$ the gradient of m taking the errors ϵ as variables, $\vec{\nabla}_\epsilon m \equiv \left(\frac{\partial m}{\partial \epsilon_1}, \frac{\partial m}{\partial \epsilon_2}, \dots, \frac{\partial m}{\partial \epsilon_n} \right)$, evaluated at $\epsilon = 0$. Finally, $[\mathcal{H}_m]_0$ denotes the Hessian matrix of m , $\mathcal{H}_m{}_{i,j} \equiv \frac{\partial^2 m}{\partial \epsilon_i \partial \epsilon_j}$, again taking the ϵ as variables and evaluating at $\epsilon = 0$. This second term represents the second order response of m to changes in the observations, an expression that can be also calculated analytically from the TF defined in Chapter 3.

Eq. (4.8) states that for some small realization of the errors ϵ_i , the value of m is approximately its value when $\epsilon = 0$, plus a linear correction in ϵ by the gradient, and finally a second order correction proportional to ϵ^2 that involves the Hessian. Both of these quantities, the gradient and the Hessian, are evaluated at $\epsilon = 0$ and are therefore, just a list and a matrix of fixed numbers, respectively. These two groups of numbers can be obtained numerically by taking finite differences on m as an approximation to derivatives.

Once we have expanded the slope function m , we proceed to calculate the integral (4.6):

$$\langle m(X, Y) \rangle \approx \int \left(m_0 + \left[\vec{\nabla}_\epsilon m \right]_0 \cdot \vec{\epsilon} + \frac{1}{2} \vec{\epsilon}^T \cdot [\mathcal{H}_m]_0 \cdot \vec{\epsilon} \right) P(\epsilon) d\epsilon. \quad (4.9)$$

The linear correction, being proportional to ϵ , would integrate to zero. The term m_0 does not depend on the noise and since $P(\epsilon)$ integrates to 1, it would appear just as it is in the final result. The last term with the Hessian is a quadratic form that would look like $\epsilon_1^2 H_{11} + 2\epsilon_1\epsilon_2 H_{12} + \dots + \epsilon_n^2 H_{nn}$, where the constants $H_{i,j}$ are the (i, j) elements of the Hessian. Since our noise additions are mutually independent (ϵ_i and ϵ_j do not correlate), the integration of terms involving mixing of different ϵ_i will also yield zero. The integration over terms that involve the same ϵ_i are, by definition, the variance of the noise, σ^2 . Therefore, we have:

$$\begin{aligned} \langle m(X, Y) \rangle &\approx m_0 + \frac{1}{2} \int (\vec{\epsilon}^T \cdot [\mathcal{H}_m]_0 \cdot \vec{\epsilon}) P(\epsilon) d\epsilon \\ &= m_0 + \frac{1}{2} \sigma^2 \left[\frac{\partial^2 m}{\partial \epsilon_1^2} + \frac{\partial^2 m}{\partial \epsilon_2^2} + \dots + \frac{\partial^2 m}{\partial \epsilon_n^2} \right]_0. \end{aligned} \quad (4.10)$$

This equation shows that there will be a noise-induced bias in the average estimation of our quantity m , which grows proportionally to σ^2 . This bias will not be present if the function that estimates m in terms of the data is linear in the observations Y (and therefore linear in the noise ϵ), since second derivatives of m will be zero from the start. This means that, regardless of the size of the noise, linear fits like a straight line, or any polynomial, will not gain a noise-induced bias, and the distributions for m will always be centered at the zero noise value.

On the other hand, for non linear fits, there will be a bias that will grow quadratically in the noise. This can apply to the actual function that generated the data, like the dipole in our study. Once the noise gets too high, the true generating function might not be the most reliable fit to the data, bias-wise. Nevertheless, the effect will usually be very small, since it is proportional to σ^2 .

The term accompanying σ^2 , the trace of the Hessian matrix, can be interpreted as the Laplacian of m at zero noise. Consistent with our findings, the Laplacian operator of $f(x)$ at a point a can be related to the rate of change of the average of f , $\langle f \rangle$, at a small sphere centered at a compared to $f(a)$ [147].

4.1.7 Standard Deviation

To calculate the standard deviation (SD) $\Delta m = \sqrt{\langle m^2 \rangle - \langle m \rangle^2}$ of our distribution for m , we must compute the expected value of m^2 , since we already have $\langle m \rangle$. We can calculate this quantity up to second order in σ , using expression (4.8):

$$\begin{aligned} \langle m(X, Y)^2 \rangle &\approx \int \left[m_0 + \left[\vec{\nabla}_\epsilon m \right]_0 \cdot \vec{\epsilon} + \frac{1}{2} \vec{\epsilon}^T \cdot [\mathcal{H}_m]_0 \cdot \vec{\epsilon} \right]^2 P(\epsilon) d\epsilon = \\ &\quad \int \left[m_0^2 + \cancel{2m_0 \left[\vec{\nabla}_\epsilon m \right]_0 \cdot \vec{\epsilon}}^0 + \left(\left[\vec{\nabla}_\epsilon m \right]_0 \cdot \vec{\epsilon} \right)^2 + \right. \\ &\quad \left. 2m_0 \frac{1}{2} \vec{\epsilon}^T \cdot [\mathcal{H}_m]_0 \cdot \vec{\epsilon} + \mathcal{O}(\epsilon^4) \right] P(\epsilon) d\epsilon. \end{aligned} \quad (4.11)$$

The first term does not depend on ϵ and can go out of the integral. The second term is proportional to ϵ and integrates to zero. We already know how to integrate the Hessian term and we have neglected its square, which will be order ϵ^4 . The third term, the gradient squared, reads:

$$\left(\left[\vec{\nabla}_\epsilon m \right]_0 \cdot \vec{\epsilon} \right)^2 = \left(\frac{\partial m}{\partial \epsilon_1} \Big|_0 \epsilon_1 + \frac{\partial m}{\partial \epsilon_2} \Big|_0 \epsilon_2 + \dots + \frac{\partial m}{\partial \epsilon_n} \Big|_0 \epsilon_n \right)^2. \quad (4.12)$$

Once we expand the square, we have two types of terms: those of the form $\left(\frac{\partial m}{\partial \epsilon_i} \Big|_0 \right)^2 \epsilon_i^2$, and those with mixed ϵ , $\left(\frac{\partial m}{\partial \epsilon_i} \frac{\partial m}{\partial \epsilon_j} \Big|_0 \right) \epsilon_i \epsilon_j$, $i \neq j$. Since our noise components are independent, only the first type of terms will give non zero results under integration: σ^2 . Therefore, we have:

$$\langle m(X, Y)^2 \rangle \approx m_0 + m_0 \sigma^2 \text{Tr}[\mathcal{H}_m]_0 + \sigma^2 \left[\left(\frac{\partial m}{\partial \epsilon_1} \right)^2 + \left(\frac{\partial m}{\partial \epsilon_2} \right)^2 \dots + \left(\frac{\partial m}{\partial \epsilon_n} \right)^2 \right]_0. \quad (4.13)$$

The first coefficient accompanying σ^2 , is the trace of the Hessian matrix, which will cancel the same exact term that appears in the square of the mean value for m : $\langle m \rangle^2 \approx (m_0 + \frac{1}{2} \sigma^2 \text{Tr}[\mathcal{H}_m]_0)^2 \approx m_0^2 + m_0 \sigma^2 \text{Tr}[\mathcal{H}_m]_0$.

The second coefficient accompanying σ^2 , is the magnitude of the gradient of m , evaluated at zero noise. Our result for the SD, Δm , if we only keep terms up to first order in σ for $\langle m \rangle^2$, then reads:

$$\Delta m = \sqrt{\langle m^2 \rangle - \langle m \rangle^2} = \sigma \|\vec{\nabla}_\epsilon m\|_0. \quad (4.14)$$

Therefore, our calculations show that no matter how small the noise in the data, there would be an appreciable, proportional, standard deviation in our estimate of m . This is true unless the gradient is zero, meaning that our quantity m is not sensitive to change in the data to first order. Eq. (4.14) is consistent with the analytical expression in Eq. (3.6).

All the calculations presented here can be generalized to include correlations between the ϵ_i and to include different sizes of noise for each point: σ_j , instead of a single σ . The only requirement is that the noise is considered Gaussian and, therefore, the integrals would still be easy to solve.

It is also possible to extend the Taylor expansion of m in Eq. (4.8) to include more orders in the noise ϵ . Since the Taylor expansion will contain some combination of different ϵ_i , the Gaussian integrals can still be performed in closed form, term by term.

4.1.8 Maximizing Information Gained

Some fundamental questions that arise in experimental designs and theory of information are: for a given relation $f(x) = y$ between two quantities, what are the most relevant values of x such that, when y is measured, the relation f is constrained the most? How much new information one gains by adding more points (see for example [148]), or by reducing the statistical uncertainties on the existing points? In other words, what are the (x, y) values that, considered as a set, contain the most amount of information regarding f , for a given set of experimental uncertainties? With respect to the proton radius extraction that we have studied in this chapter, a question of interest would be: for a given range in Q^2 , what specific values Q_i^2 give us the largest amount of information regarding the slope at zero, or the proton radius.

Using the semi-analytical framework described, we can give a quantitative answer to this question. Eqs. (4.10) and (4.13) show explicitly the contribution each point makes to the noise bias, and the SD, respectively. Therefore, we could directly identify how a reduction in uncertainty on each particular point would propagate to our estimate of the radius.

Alternatively, if we are fixed on some level of uncertainty, we could estimate where the n measurements should be made such that, again, our information gain is maximal. To test these

ideas, we minimized the MSE defined in Eq. (2.27) as a function of the location of $n Q^2$ points, and studied how much the uncertainty in the slope at zero is reduced as n increases.

Our y points were again generated by the dipole function defined in Eq. (4.1), and we studied the fit of a line in the range $0.1 - 0.8 \text{ fm}^{-2}$ and the fit of a parabola in the range $0.1 - 1.6 \text{ fm}^{-2}$, both of which have proven to be optimal in our simulations in Sec. 4.1.3.

Since the slope estimate from both models is linear in the observations y , we know that there will be no noise-induced bias, and the only source of bias will be a constant term, see Eq. (4.5). By “constant” we mean that it will not grow with the noise, but it would change depending on the points Q_i^2 used. This is easier to see with the analytical version of the TF, as we described in Sec. 3.3.

Therefore, we seek to minimize:

$$\text{MSE}^2(\mathbf{Q}^2) = \text{bias}^2(\mathbf{Q}^2) + \Delta m^2 = (m_{\text{true}} - m_0(\mathbf{Q}^2))^2 + \sigma^2 \|\vec{\nabla}_\epsilon m(\mathbf{Q}^2)\|_0^2, \quad (4.15)$$

where $\mathbf{Q}^2 = (Q_1^2, \dots Q_n^2)$ are the n locations of the “observed” points. We have omitted the Y dependence on our quantities since we are calculating Y directly from our points \mathbf{Q}^2 , and $m_{\text{true}} = -0.1097$ is the fixed value of the slope at zero, in fm^2 .

Now, the bias quantity $(m_{\text{true}} - m_0(\mathbf{Q}^2))$ will indeed depend on \mathbf{Q}^2 . Is a number that is usually hidden to us by nature, since we do not know the true value. Nevertheless, this exercise is helpful in showing that some points have more information than others, and we will see how both terms evolve as n grows.

Figures 4.4 and 4.5 show the results obtained when the points Q_i^2 are not distributed uniformly but are rather moved to positions that, respecting the assigned range, minimize the MSE given by Eq. (4.15). It is interesting to note that, as we expected from our analysis in Sec. 4.1.3, the MSE of the parabola is primarily driven by the standard deviation contribution, while the line is more balanced, although towards higher bias.

The main result that we can extract from these two graphs is that it takes only 10 points in the case of the line and 13 in the case of the parabola, to obtain the same value of MSE compared to 31 uniformly spaced points (see Table 4.3). If we use 31 points but locate them in optimal positions, we obtain an MSE of 0.0068 fm^2 for the line and an MSE of 0.0056 fm^2 for the parabola. The first one represents an improvement of 21%, while the second represents an improvement of 31%.

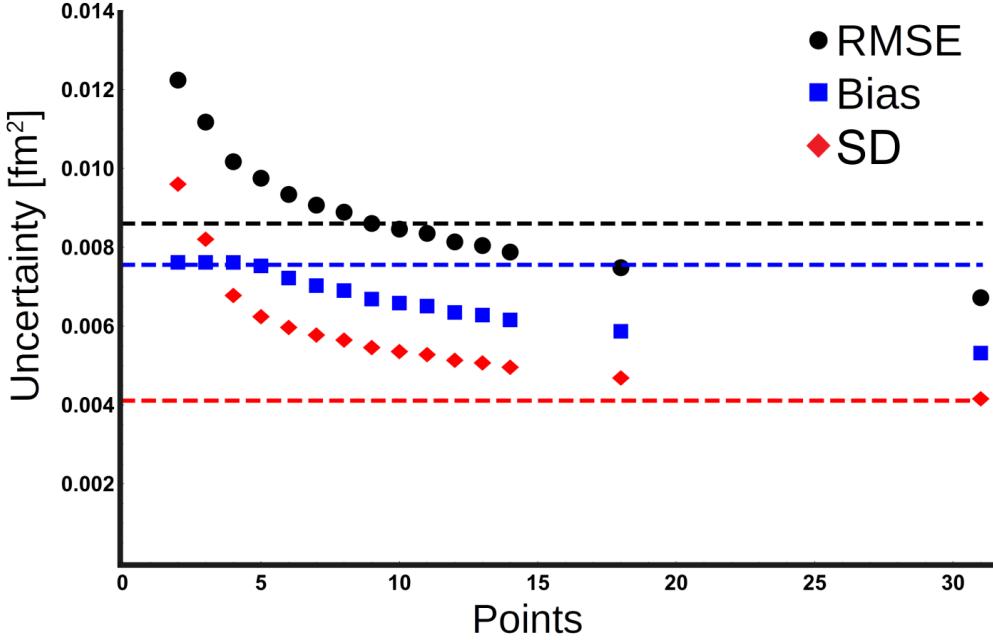


Figure 4.4: Line MSE, absolute bias and SD, for the optimal location of points as a function of the number of points available, in the range $0.1\text{-}0.8 \text{ fm}^{-2}$. The dashed horizontal lines show the respective values for the configuration of 31 equidistant points showed in Table 4.2.

As we have already mentioned, calculating (and optimizing with respect to) the bias, requires to know the true value, which is usually not the case in interesting problems. Nevertheless, the standard deviation we have defined can be calculated with no more than the observed values and errors, and therefore, it is a quantity we can estimate when facing the real data. As can be seen in Fig. 4.5, the bias of the parabola does not seem to change much for any configuration of points. Nevertheless, the SD, and therefore the MSE, can be improved substantially.

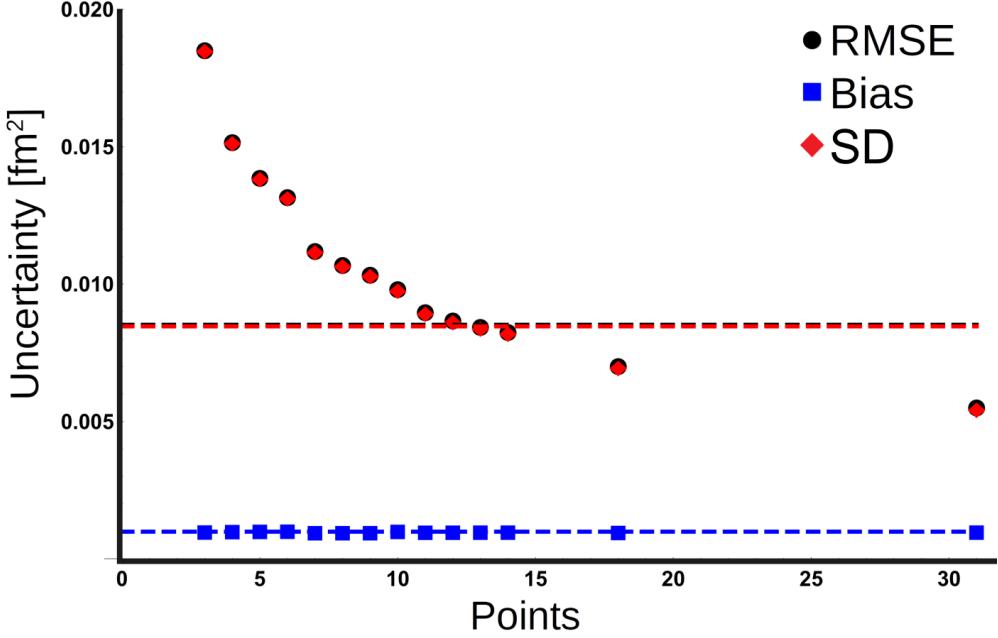


Figure 4.5: Parabola MSE, absolute bias and SD, for the optimal location of points as a function of the number of points available, in the range $0.1\text{-}0.1.6 \text{ fm}^{-2}$. The dashed horizontal lines show the respective values for the configuration of 31 equidistant points showed in Table 4.2.

4.2 Nonparametric approach: Constrained Gaussian Process

In this section we take a different route to extrapolate $G_E(Q^2)$ to $Q^2 = 0$ and obtain the proton radius r_p . Instead of comparing different parametric models, we develop a novel and powerful nonparametric method flexible enough to fit the data without any parametric assumptions on the form factor. The Bayesian estimation is guided by imposing the following constraints that are the cornerstone of the nonparametric approach:

$$G_E(Q^2=0) = 1, \quad (4.16a)$$

$$G'_E(Q^2) \equiv \frac{dG_E}{dQ^2} < 0, \quad (4.16b)$$

$$G''_E(Q^2) \equiv \frac{d^2G_E}{d(Q^2)^2} > 0. \quad (4.16c)$$

The first equation (4.16a), is model independent since it is directly related to the charge of the proton. The other two equations (4.16b)-(4.16c), which we will call the shape constraints, are not guaranteed by the definitions above, but are rather deduced from the analytic properties of the form factor (see for example [149, 150]). Most of the parametric models that have been used for the form factor respect these shape constraints for the parameters estimated (see for example [61]).

In contrast with what we did in Sec. 4.1, we work directly with the 1422 experimental data points from the Mainz collaboration [109] and avoid changing their locations. However, we explore the effect that the change in the Q^2 range used has on the calculated results. We also test our nonparametric model against pseudo data generated by a dipole function as the one in Eq. (4.1), but with a radius of 0.84 fm (see Appendix A.5 for a detailed description).

4.2.1 Approximating G_E : basis construction

Having introduced the electric form factor of the proton, we now proceed to build the nonparametric model that will allow us to extrapolate $G_E(Q^2)$ to $Q^2=0$. Our main goal is to incorporate the general constraints given in Eq. (4.16) into the estimation procedure without making parametric assumptions on the functional form of $G_E(Q^2)$. The available experimental data will guide the shape of such nonparametric curve, ultimately allowing us to estimate r_p . In this section we describe the fully constrained model while in Sec. 4.2.6 we explore the impact of relaxing the constraints.

To facilitate the implementation of the nonparametric approach, we select a maximum value of Q^2 , Q_{\max}^2 , up to where the analysis is performed, a selection that has been shown to impact the estimation of r_p , as we discussed and showed in Sec 4.1 (see Tables 4.2 and 4.3). Once the momentum-transfer range has been selected, $0 \leq Q^2 \leq Q_{\max}^2$, we define the dimensionless scaled variable x as $x = Q^2/Q_{\max}^2$. Following this definition, the range of x is the unit interval $x \in [0, 1]$.

We start by defining a working grid formed by a collection of $N+1$ equally spaced points $x_j = j/N$ in the closed interval $[0, 1]$, in which j runs from 0 to N . We adopt the notation of [151] to define a set of basis functions:

$$h_j(x) = \begin{cases} 1 - N|x - x_j|, & \text{if } |x - x_j| \leq 1/N \\ 0, & \text{otherwise.} \end{cases} \quad (4.17)$$

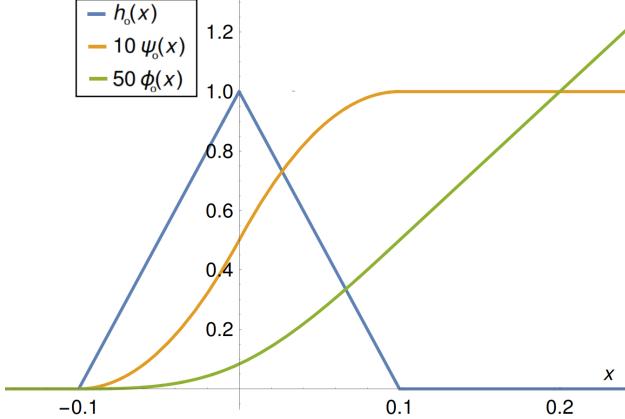


Figure 4.6: Functions $h_0(x)$ (blue), $\psi_0(x)$ (orange) and $\phi_0(x)$ (green) for $N = 10$. The functions ψ_0 and ϕ_0 have been rescaled by a factor of 10 and 50 respectively.

and their corresponding integrals as follows:

$$\psi_j(x) = \int_0^x h_j(t) dt, \quad (4.18)$$

$$\phi_j(x) = \int_0^x dt \int_0^t h_j(s) ds. \quad (4.19)$$

Figure 4.6 shows the form of these functions for $j = 0$. We characterize our regression model in terms of $(N+3)$ free parameters ξ_j that will be obtained from a suitable fit to the experimental data. That is,

$$f(x) \approx f_\xi(x) \equiv \xi_1 + \xi_2 x + \sum_{j=0}^N \xi_{j+3} \phi_j(x), \quad (4.20)$$

where $f(x)$ is the form factor in the rescaled variable x . Under this scheme, the terms ξ_1 and ξ_2 relate to the value of the modeled function $f(x)$ and its derivative at zero, respectively, while the terms ξ_{j+3} relate to the values of the second derivative of the modeled function on the grid points. Note that during the regression process all of these coefficients remain as free parameters.

We explain in detail in Appendix A.2 the construction of this approximation for $f(x)$ and the relation between the terms ξ_{j+3} and the values of its second derivative. Moreover, this approximation has a nice physical underpinning. If we regard $f(t)$ as the one-dimensional trajectory of a particle as a function of time t , then the approximation:

$$f(t) \approx f(0) + t f'(0) + \sum_{j=0}^N f''(t_j) \phi_j(t), \quad (4.21)$$

may be explained as follows. At time $t = 0$ the particle starts at a position $f(0)$ with an initial velocity $f'(0)$. As time evolves, corrections to the straight-line trajectory are implemented by the different ϕ_j in proportion to $f''(t_j)$, that can be thought as “acceleration spikes” that stir the particle into the correct trajectory.

4.2.2 Incorporating physical constraints

The great virtue of the nonparametric approach adopted here is that no assumption is made about the functional form of the electric form factor. However, if the calibration parameters ξ_j defined in Eq. (4.20) are left unrestricted, the resulting model for $f_\xi(x)$ is likely to violate the physical constraints imposed in Eq. (4.16). In the notation assumed in this section, these constraints are given by: (a) $f_\xi(0) = 1$, (b) $f'_\xi(x) < 0$, and (c) $f''_\xi(x) > 0$, where (b) and (c) hold for all x in $[0, 1]$. In this section, we discuss the model formulation with all the constraints. To satisfy these constraints, the model parameters must obey the following linear relations:

$$\xi_1 = 1, \quad (4.22a)$$

$$\xi_2 + \sum_{j=0}^N c_j \xi_{j+3} \leq 0, \quad (4.22b)$$

$$\xi_{j+3} \geq 0, \text{ for } j = 0, 1, \dots, N, \quad (4.22c)$$

where $c_j = \psi_j(1)$ is the area under the triangle formed by the function $h_j(x)$, except for $j = 0$ (c_0), and for $j = N$ (c_N). These last two values are equal to half the area of the triangle. The proof of this statement is shown in Appendix A.1.

To incorporate the constraints in Eq. (4.22), let us call \mathcal{C}_ξ the set of all the ξ_j that satisfies the relationships (4.22). In formal notation that would be:

$$\mathcal{C}_\xi \equiv \left\{ \xi \in \mathbb{R}^{N+3} : \xi_1 = 1, \xi_2 + \sum_{j=0}^N c_j \xi_{j+3} \leq 0, \xi_{j+3} \geq 0, j = 0, \dots, N \right\}. \quad (4.23)$$

The proton radius introduced in Eq. (2.4) is expressed directly in terms of ξ_2 as:

$$r_p = \frac{\sqrt{-6\xi_2}}{Q_{\max}}, \quad (4.24)$$

where Q_{\max} enters to account for the rescaling of Q^2 into the dimensionless variable $x = Q^2/Q_{\max}^2$.

Note that the value of ξ_1 is fixed at 1 and r_p only depends on the value of ξ_2 in the constrained set \mathcal{C}_ξ . We provide a detailed discussion of a partially constrained model with the condition $\xi_1 = 1$

removed in Appendix A.3. The rest of the discussion in the following sections follows a fully constrained model.

4.2.3 Probabilistic model for fully constrained function estimation

The observed experimental data consists of n pairs of the form (x_i, g_i) , where $x_i = Q_i^2/Q_{\max}^2$ and g_i is equal to the form factor $G_E(Q_i^2)$ up to some experimental noise. Specifically, we assume that the n experimental measurements g_i have normally distributed experimental errors ϵ_i . That is, $g_i = G_E(Q_i^2) + \epsilon_i$, where we assume that each ϵ_i is a normally distributed variable with zero mean and a common standard deviation σ (as in Eq. (4.7), an unknown hyperparameter that we learn from the data).

Let $Y = (y_1, \dots, y_n)^T$ with $y_i := g_i - \xi_1 = g_i - 1$ (the subtraction of the independent term ξ_1 is made in order to build an homogeneous matrix equation), and set $\varepsilon = (\epsilon_1, \dots, \epsilon_n)^T$. Also, we define a basis matrix Φ (a $n \times (N+2)$ matrix) with i th row $(x_i, \phi_0(x_i), \dots, \phi_N(x_i))$. With these ingredients, we express our model in vectorized notation as:

$$Y = \Phi\xi + \varepsilon, \quad \varepsilon \sim \mathcal{N}_n(0, \sigma^2 I_n), \quad \xi \in \mathcal{C}_\xi, \quad (4.25)$$

where \mathcal{C}_ξ is defined in Eq. (4.23), and I_n denotes the unity matrix of size n . The notation $v \sim \mathcal{N}_n(\mu, \Sigma)$ means that the random variable v follows a multivariate Gaussian distribution with mean μ and covariance matrix Σ .

As we discussed in Sec. 2.2.2, we operate in a Bayesian framework [152] and express pre-experimental uncertainty in ξ through a prior distribution $P(\xi)$. As in Eq. (2.20), the prior for ξ is combined with the data likelihood $P(Y|\xi)$ to obtain the posterior distribution for ξ given the observed values Y :

$$P(\xi|Y) = \frac{P(Y|\xi)P(\xi)}{P(Y)}. \quad (4.26)$$

This posterior distribution of the parameters $P(\xi|Y)$ can then be used to make inference on r_p including point estimates and uncertainty quantification through credible intervals. Since we assume Gaussian distributed noise ε_i for the observational points y_i , our likelihood term $P(Y|\xi)$ will follow Eq. (2.22) and have the form $Y \sim \mathcal{N}_n(\Phi\xi, \sigma^2 I_n)$, which represents an exponential decay in the square of the difference between our observed data and our model prediction: $\chi^2 = \sum_1^n (Y_i - f_\xi(x_i))^2 / \sigma^2$. The choice of a suitable prior $P(\xi)$ is critical for a valid inference on r_p . A flexible representation

for f can be reproduced through the coefficients ξ . These coefficients relate to f through its derivatives (see A.2). In the unconstrained setting, a natural choice of prior for ξ can be induced through a Gaussian process prior on f . On the other hand, the prior for ξ should be supported on the restricted space \mathcal{C}_ξ so that any prior draw follows the constraints for ξ . We combine these two features to propose a flexible constrained Gaussian prior for ξ and describe this procedure in the following section.

4.2.4 Prior specification: Constrained Gaussian Process

A Gaussian process (GP) [153] is a distribution of functions on the functions space such that the collection of random variables obtained by evaluating the random function at a finite set of points is multivariate Gaussian (see [153]). We use the notation $f \sim \text{GP}(\mu, \tau^2 K)$ to denote that the function f follows a Gaussian process with mean function μ and covariance function $\tau^2 K$, where τ is a hyperparameter controlling the overall scale of K . This means that any finite collection of points $f_1(x_1), \dots, f_N(x_N)$ at locations x_1, \dots, x_N has a joint normal Gaussian distribution given by:

$$\left(f_1(x_1), \dots, f_N(x_N) \right) \sim \mathcal{N}(\mu, \Sigma), \quad (4.27)$$

where $\mu = (\mu(x_1), \dots, \mu(x_N))$ and $\Sigma_{ij} = \tau^2 K(x_i, x_j)$. The mean function $\mu(x)$ controls around *where* the points are distributed, and the covariance function $K(x, x')$ controls *how* the points are distributed, τ being an expected size for the deviations. The covariance function $K(x, x')$ would allow to introduce a notion of correlation between different momenta Q^2 in the form factor, allowing our estimation for the radius (related to the slope at $Q^2 = 0$) to borrow information for the entire Q^2 range.

The model parameters ξ_j are related to first and second derivatives of the form factor G_E , or equivalently to its rescaled version f at the various grid points x_j . Since Gaussian processes are closed under linear operations such as taking derivatives [153], they represent an optimal choice in estimating the form factor. If $f \sim \text{GP}(0, \tau^2 K)^1$, then any finite number of observations $f(x_1), \dots, f(x_N)$ follows the distribution specified by Eq. (4.27). Therefore, a collection of random variables that involves derivatives $f'(0), f''(x_0), \dots, f''(x_N)$ also follows a Gaussian distribution with a covariance matrix Γ that involves up to four mixed partial derivatives of the covariance function $K(x, x')$ (see

¹The selection $\mu(x) = 0$ is done to avoid centering the GP around any parametric form.

Theorem 2.2.2 in [154]). That is,

$$\Gamma = \begin{bmatrix} \frac{\partial^2 K}{\partial x \partial x'}(0, 0) & \frac{\partial^3 K}{\partial x \partial x'^2}(0, x_0) & \cdots & \frac{\partial^3 K}{\partial x \partial x'^2}(0, x_N) \\ \frac{\partial^3 K}{\partial x^2 \partial x'}(x_0, 0) & \frac{\partial^4 K}{\partial x^2 \partial x'^2}(x_0, x_0) & \cdots & \frac{\partial^4 K}{\partial x^2 \partial x'^2}(x_0, x_N) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^3 K}{\partial x^2 \partial x'}(x_N, 0) & \frac{\partial^4 K}{\partial x^2 \partial x'^2}(x_N, x_0) & \cdots & \frac{\partial^4 K}{\partial x^2 \partial x'^2}(x_N, x_N) \end{bmatrix}_{(N+2) \times (N+2)}. \quad (4.28)$$

For illustration purposes, consider the first row of the matrix Γ . It specifies how the derivative of the function at zero, ξ_2 , correlates with all the other ξ_j . The correlation between ξ_2 and the other ξ_j for $j > 2$ is controlled by the mixed partial third derivative of K at x_j .

If the model parameters ξ_j are left unconstrained, then a natural prior, induced from a GP prior on the unknown function f , would be a finite-dimensional Gaussian prior $\xi \sim \mathcal{N}_{N+2}(0, \tau^2 \Gamma)$ with Γ as in Eq. (4.28). However, since the various shape constraints on the function impose a corresponding set of constraints on the model parameters, we adopt a truncated Gaussian prior on ξ :

$$p(\xi) = \frac{1}{M_\xi} (2\pi)^{-(N+2)/2} |\Gamma|^{-1/2} (\tau^2)^{-(N+2)/2} e^{-\frac{\xi^T \Gamma^{-1} \xi}{2\tau^2}} \mathbb{1}_{C_\xi}(\xi), \quad (4.29)$$

where the ‘‘indicator function’’ $\mathbb{1}_{C_\xi}(\xi)$ filters the ξ_j such that only the allowed combinations are those that satisfy the constraints listed in Eq. (4.22): $\mathbb{1}_{C_\xi}(\xi) = 1$ if $\xi \in C_\xi$, and $\mathbb{1}_{C_\xi}(\xi) = 0$ otherwise. In the above expression, M_ξ is a constant of proportionality required to make $p(\xi)$ a density distribution, i.e., $p(\xi)$ must integrate to one. As is commonly done [155], we place an (improper) objective prior on τ^2 . We shall denote $p(\xi)$ by $\mathcal{N}_{N+2}(0, \tau^2 \Gamma) \mathbb{1}_{C_\xi}(\xi)$ and refer to it as the constrained Gaussian Process (cGP) prior for ξ .

To fully specify the cGP prior, we still need to define the covariance function $K(x, x')$ that determines the matrix Γ . Following common practice, we chose K to be a stationary Matérn kernel [153] with smoothness parameter $\nu = 5/2$ and length-scale $\ell > 0$. Such a kernel only depends on the relative distance between the coordinates $r \equiv |x - x'|$ and can be written in closed form as follows:

$$K(x, x') \equiv k_{\nu=5/2, \ell}(r) = \left(1 + \frac{\sqrt{5}r}{\ell} + \frac{5r^2}{3\ell^2} \right) \exp \left(-\frac{\sqrt{5}r}{\ell} \right). \quad (4.30)$$

In our analysis we also explored the values $\nu = 3$ and $\nu = 7/2$. The more general definition for the Matérn kernel is shown in Appendix A.4. The optimal value for the correlation length ℓ was chosen by a cross-validation scheme outlined in Appendix A.5 and Appendix A.6.

4.2.5 Posterior sampling and inference

Our objective is to obtain a reasonable sized distribution of values for the coefficients ξ_j that will allow to report a distribution for the inferred proton radius r_p . Given the complex nature of the model space associated with the allowed values of ξ , an analytic expression of M_ξ is not available. However, we show in Appendix A.1 that M_ξ does not depend on τ . Hence, provided Γ is fixed, we can exploit this fact and use a Markov Chain Monte Carlo (MCMC) algorithm to sample the posterior distribution. The model along with priors on various components are represented in a hierarchical fashion as follows:

$$\begin{aligned} Y \mid \xi, \sigma^2, \tau^2 &\sim \mathcal{N}_n(\Phi\xi, \sigma^2 I_n), \\ \xi &\sim \mathcal{N}_{N+2}(\xi; 0, \tau^2 \Gamma) \mathbb{1}_{\mathcal{C}_\xi}(\xi), \quad p(\tau^2) \propto \frac{1}{\tau^2}, \quad p(\sigma^2) \propto \frac{1}{\sigma^2}, \end{aligned} \quad (4.31)$$

in which we have made the common non-informative prior choice for the observational noise standard deviation σ^2 . For the hierarchical model above, the joint posterior distribution of the model parameters is given by:

$$P(\xi, \tau^2, \sigma^2 \mid Y) \propto \left\{ (\sigma^2)^{-n/2} e^{-\frac{\|Y - \Phi\xi\|^2}{2\sigma^2}} \right\} \left\{ (\tau^2)^{-(N+2)/2} e^{-\xi^T \Gamma^{-1} \xi / (2\tau^2)} \mathbb{1}_{\mathcal{C}_\xi}(\xi) \right\} (\tau^2)^{-1} (\sigma^2)^{-1}. \quad (4.32)$$

The final normalizing constant of the posterior distribution is intractable and hence, we resort to MCMC [152] to sample from the posterior distribution of the model parameters. More specifically, we use Gibbs sampling to iteratively sample from the full conditional distribution of (i) $\xi \mid \tau^2, \sigma^2, Y$ ², (ii) $\tau^2 \mid \xi, \sigma^2, Y$, and (iii) $\sigma^2 \mid \xi, \tau^2, Y$. The conditional posterior of ξ in (i), is a truncated multivariate normal distribution which is sampled using the method proposed in [156]. The conditional posteriors of σ^2 and τ^2 in (ii) and (iii), are inverse-gamma (IG) distributions and hence easy to sample from. The details of the algorithm are provided in Appendix A.4.

After discarding initial burn-in samples, let $\xi_j^{(1)}, \dots, \xi_j^{(T)}$ be T successive iterate values of ξ_j from the Gibbs sampling algorithm, for $j = 2, \dots, N + 3$. Our point estimate for r_p based on the posterior samples is:

$$\hat{r}_p = T^{-1} \sum_{t=1}^T \frac{\sqrt{-6\xi_2^{(t)}}}{Q_{\max}}. \quad (4.33)$$

The 68% confidence interval (one sigma) for r_p is also computable from our sampling algorithm.

²Note that in Bayesian notation $\xi \mid \tau^2, \sigma^2, Y$ means the conditional posterior distribution of ξ given τ^2, σ^2 , and Y .

4.2.6 Electron-scattering data analysis

Before dealing with the real data set we analyzed pseudo data generated by a known dipole function to test the performance of our method and to observe the impact that the different hyperparameters and data range had on the extracted radius (see Appendix A.5 for a detailed description). In this section, we present our analysis of the real electron-proton scattering data from Mainz [109].

We tried four different variants of our model by imposing or relaxing the different constraints on ξ in Eq. (4.22). These are:

1. **cGP**: denotes the proposed constrained GP model as described in Eq. (4.25). The curve is restricted to be convex and the value at $Q^2 = 0$ is fixed at 1 ($\xi_1 = 1$).
2. **c₀GP**: denotes the model in Eq. (4.25) with the only constraint being (4.16a), the value at zero ($\xi_1 = 1$). The parameters ξ_2, \dots, ξ_{N+3} are left unconstrained in this model and therefore, the curve is not necessarily monotonic and convex.
3. **c₁GP**: denotes the model with only shape constraints (4.16b) and (4.16c), which implies that the function is non-increasing and convex, but the value at zero is not fixed (ξ_1 is left unconstrained).
4. **uGP**: denotes the completely unconstrained GP, all the parameters $\xi_1, \xi_2, \dots, \xi_{N+3}$ are free.

We note that although the condition $G_E(0) = 1$ (4.16a) is ultimately related to the charge of the proton, systematic errors can have an appreciable impact on the fulfillment of this constraint in the experimental data. It has become a customary practice (see for example [61]), to represent the observed values as $f(Q^2) = n_0 G_E(Q^2)$, where n_0 is a floating normalization parameter, $f(Q^2)$ are the observed values, and $G_E(Q^2)$ is the true proton form factor. We can identify in our framework the choice $n_0 = 1$ with the requirement that our model estimate for the form factor has the fixed value of 1 at $Q^2 = 0$ (cGP and c₀GP). Instead, leaving n_0 as an adjustable parameter corresponds to the relaxation of the constraint at zero (c₁GP and uGP).

We conducted the real analysis in two regimes: low $Q^2 < 1.36 \text{ fm}^{-2}$ (the first 500 data points) and high $Q^2 < 25.12 \text{ fm}^{-2}$ (the full data set). The low regime was chosen based on the results in the pseudo-data analysis in which we observed that in this range the models gave a more accurate estimate of the slope of the assumed dipole function. On the other hand, even though in the high Q^2 regime we observed some biasing toward lower estimates of the slope, we considered also the full data analysis. As we discussed at the beginning of this chapter, it is well known that due to the

difficulty of measuring the form factor for smaller values of the momentum, the experimental data might be significantly biased for $Q^2 \approx 0$. Also, the noise structure could not satisfy the assumptions we made on the pseudo data analysis: data could be not independent and identically distributed and all the points might not share the same variance. Thus, incorporating the whole range of values could help the analysis to overcome that experimental bias. Finally, having the two extremes (low and high regimes) is beneficial for comparison.

The analysis started with conducting pilot experiments with subsets of the data of size $n = 250$ randomly selected from the range of the potential values (Q^2) for the high regime, and with the full 500 points in the low regime. The pilot experiments provided us with a better idea of the roles of the different hyperparameters of our model in the real data set, N , ℓ and ν , before eventually analyzing this real data set. Let us recall that the Q^2 values are rescaled to $[0, 1]$ before the analysis. Overall, we used 500 MCMC iterations after discarding a burning of 100 samples to form the posterior summary estimates of the radius.

		High Regime n=(1422)		Low Regime (n=500)	
Model	N	r_p (fm)	Δr_p (fm)	r_p (fm)	Δr_p (fm)
cGP	$n/4$	0.844	0.002	0.853	0.002
	n	0.845	0.001	0.855	0.002
c_0 GP	$n/4$	0.836	0.005	0.841	0.008
	n	0.845	0.004	0.840	0.006
c_1 GP	$n/4$	0.842	0.004	0.873	0.006
	n	0.831	0.003	0.872	0.005
uGP	$n/4$	0.847	0.011	0.857	0.017
	n	0.858	0.008	0.862	0.01

Table 4.5: High and low regime posterior estimates of the radius and one standard deviation (Δr_p) for cGP, c_0 GP, c_1 GP and uGP with $N = \{n/4, n\}$ and $\nu = 2.5$.

We conducted a cross validation procedure to select the optimal scale-length parameter ℓ for each regime, the details of which are shown in Appendix A.6. Our analysis guided us to choose $\ell_{\text{opt}} = 0.5$ for the full data set and $\ell_{\text{opt}} = 10$ on the low Q^2 set.

Having chosen the correlation length, we performed the MCMC iterations for the four models, selecting the number of grid points $N = n/4$ and $N = n$ in order to compare results. Table 4.2.6 shows the posterior medians of r_p of the four models and the one sigma value in the high and low regime, respectively for $\nu = 2.5$ while Fig. 4.7 shows the density plots (posterior distribution $P(r_p)$) in the high regime (a) and low regime (b).

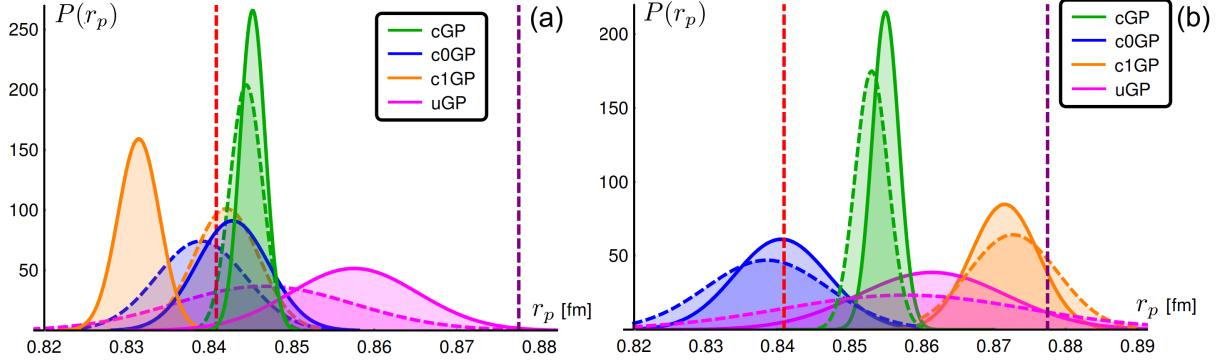


Figure 4.7: Estimated density plots of MCMC samples of radius r_p for cGP, c_0 GP, c_1 GP and uGP with $N = n/4$ (dotted line), n (solid line) and $\nu = 2.5$ for the high Q^2 regime (a) and for the low Q^2 regime (b). The vertical dashed lines stand for the muonic result of 0.84 fm (red), and the recommended CODATA value of 0.88 fm in 2014 [4] (purple).

We observed that, as a general trend, as the constraints are removed the model becomes more sensitive to the data range used and the choice of the number of grid points N and ν (the analysis and results for $\nu = 3$ and $\nu = 3.5$ are given in Appendix A.6). For example, in the high regime for all values of N and ν cGP estimations of the radius are in all cases around 0.843 fm, while on the other extreme the unconstrained model uGP estimations range between 0.76 (for $\nu = 3$) and 0.86 fm. Incorporating the constraints also strongly affects the one sigma deviation of each model: cGP one sigma intervals were between 0.001 and 0.005 fm wide, while uGP one sigma intervals were as wide as 0.014 fm. This is another manifestation of the fact that more complex models tend to have stronger variances, as represented in Fig. 4.2. In our case, uGP would be considered more complex since its prior is less restrictive than the other three, while cGP would be considered the simplest model with the most constrained prior.

For the choice $\nu = 2.5$ we can see in Fig. 4.7 that as N increases, the estimate of c_1 GP moves to a lower value of r_p while the estimates of all the other models increase to a higher value of r_p . This effect is less prominent in the low regime and overall cGP is the most robust with respect to changing the number of grid points. In all the cases, as N increases, the variability in the estimation decreases (the estimated σ is slightly smaller than those for $N = n/4$), giving more precise results. As can be observed in Table 4.2.6, in going from high regime to low regime, all the models with the exception of c_0 GP, gave a larger estimate of the radius, c_1 GP being the one that showed the

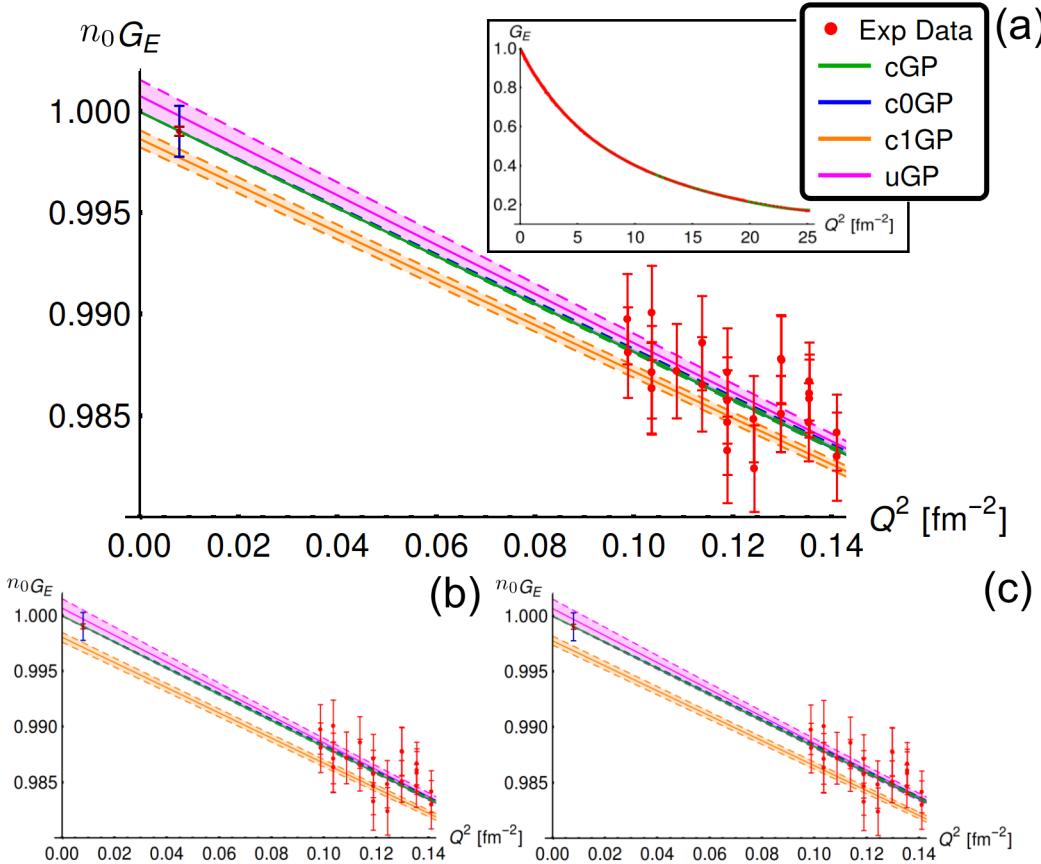


Figure 4.8: Function fit with $\nu = 2.5$ (a), 3 (b), 3.5 (c) and $N = n$ in the high regime. The inset plot in (a) shows the overall fit of the models for $\nu = 2.5$ to the entire data range. The solid curves denote the model predictions while the shaded intervals bounded by dotted lines represent the 95% confidence intervals for the predictions. The red dots denote the experimental data obtained from Mainz with their respective error bars. The red and blue points near the origin at $Q^2 = 0.008 \text{ fm}^{-2}$ represent the estimated we had at the time of these calculations on the lower value the PRad experiment would be able to measure, with two different estimates for the projected uncertainty [157], and arbitrary $G_E(Q^2)$ value.

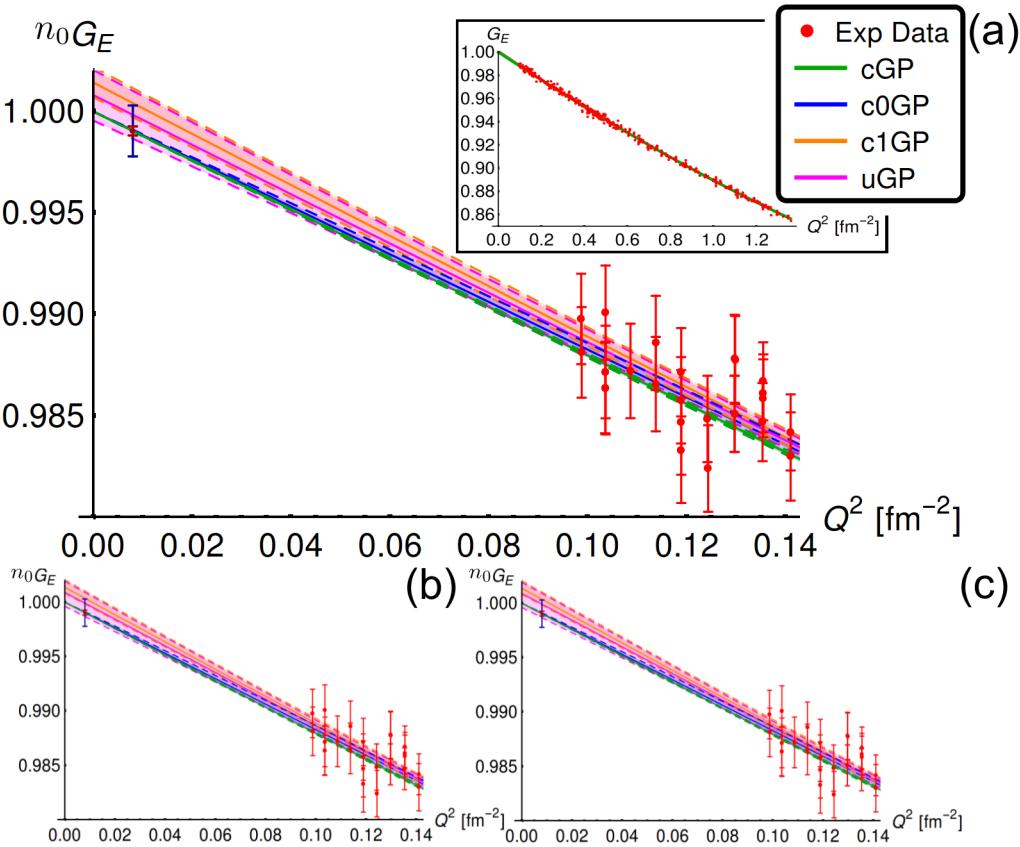


Figure 4.9: Function fit with $\nu = 2.5$ (a), 3 (b) 3.5 (c) and $N = n$ in the low regime. The inset plot of (a) shows the overall fit of the models for $\nu = 2.5$ to the entire data range. The solid curves denote the model predictions while the shaded intervals bounded by dotted lines represent the 95% confidence intervals for the predictions. The red dots denote the experimental data obtained from Mainz with their respective error bars. The red and blue points near the origin at $Q^2 = 0.008 \text{ fm}^{-2}$ represent the estimated we had at the time of these calculations on the lower value the PRad experiment would be able to measure, with two different estimates for the projected uncertainty [157], and arbitrary $G_E(Q^2)$ value.

biggest change. uGP is the only model that includes both 0.84 and 0.88 fm in its support in both regimes.

Appendix A.6 presents a more detailed description of each individual posterior histogram of the MCMC samples from GP models for the radius and the MCMC samples of $n_0 G_E(0)$ from c₁GP and uGP. Note that n_0 is defined as a floating normalization factor, while $G_E(0) = 1$ is a guaranteed property by the definition of G_E . We observed that the sample centers of $n_0 G_E(0)$ deviate from 1 by a very small amount ($|n_0 G_E(0) - 1| \lesssim 0.0014$) for both models in both regimes. It is remarkable how such a small deviation in the case of c₁GP can produce such drastic changes when r_p results are compared with the fully constrained model cGP. For example, in the low regime for $N = n/4$ cGP estimates $r_p = 0.853$ fm while c1GP, having a value of 1.0014 at zero, estimates $r_p = 0.873$, a result that highlights the impact that a floating normalization can have on the extraction of the radius.

Figures 4.8 and 4.9 show the function fits for the high and low regime respectively, with $\nu = 2.5, 3, 3.5$ for $N = n$. The overall fit is good for all the methods in both regimes, the real differences appear as $Q^2 \rightarrow 0$. For this reason, we show the full fit in each regime only for $\nu = 2.5$ in the inset of the respective top plot, the full fits for the other values of ν being visually indistinguishable.

Overall we found relatively small variability in the function fits across different values of ν in both regimes, not enough to change the estimation of the radius by more than 0.01 fm within any of the models. Due to the constraint at the origin, both posterior medians of cGP and c0GP agree as $Q^2 \rightarrow 0$ with very narrow credible intervals, while c₁GP and uGP are either below or above and start going close to the other GP models estimates as Q^2 grows. As expected, the shape constraints help reduce the variability of the models, which is evidenced by the smaller credible intervals of c1GP in comparison with uGP, specially in the low regime. In the low regime, it seems that without the normalization constraint the extrapolations are likely to attain values at $Q^2 = 0$ larger than 1, which in turn, pushes the estimate of the radius to larger values, as can be seen in Fig. 4.7. In the low regime, as a general trend, we observed wider credible intervals for all the models.

Parts of this chapter are adapted from the author's publications in Physical Review C 99.5 (2019): 055202, and in the arXiv preprint: arXiv:1812.05706 (2018). Copyright held by APS.

CHAPTER 5

WEAK CHARGE DENSITIES

In this chapter, we focus on the extraction of weak charge densities of neutron rich nuclei, specifically ^{48}Ca and ^{208}Pb . As already discussed, the interest in the neutron distribution of these nuclei spans 18 orders of magnitude in physics scales: from constraining the isovector sector of nuclear energy density functionals, to important implications on the structure of neutron stars [158]. To test our procedures we also discuss in this chapter the experimentally well-known electric charge density of both nuclei.

Let us recall that nuclear densities (both electric and weak) are obtained from experimental measurements of cross sections as in Eqs. (2.6) and (2.8), or from the asymmetry in Eq. (2.7). The respective form factor can be extracted from these measurements, and with it, we can obtain the density using Eq. (2.10). Since these types of experiments that can access weak densities (PVES and CEvNS) are extremely complex and expensive, we must make compelling statistical arguments to maximize the information gained from each experimental measurement.

The first part of this chapter is focused on the selection of a second measurement of the weak form factor of ^{208}Pb using PVES. The motivation for this study comes from the already existing single momentum transfer measurement by the PREX collaboration [82]. A second experimental project was approved, PREX-II, and it is expected to report results soon. To reach the original goal of a ± 0.06 fm determination of the weak radius of ^{208}Pb , the second experiment was performed at the same original location of momentum transfer of PREX. As we discuss in Sec. 5.1, a second measurement at a different location could serve to better constrain two parameter models that describe the ^{208}Pb density. We perform a Bayesian analysis to identify the optimal location of this possible new measurement.

In the second part of this chapter, we explore the possibility of having more than two measurements on the weak form factors of ^{48}Ca and ^{208}Pb . Inspired by the work done in [61], we analyze the performance of seven models when extracting the densities from five form factor measurements. We use the transfer function formalism described in Chapter 3 to judge the models in terms of bias

and variance, to identify the optimal locations for the measurements that will reduce the uncertainty in the density reconstructions the most, and to improve the existing models in terms of hyperparameters and priors.

5.1 Bayesian selection of a second measurement of the weak-charge form factor of ^{208}Pb

As mentioned earlier, the pioneering PREX experiment measured the weak form factor of ^{208}Pb at the single momentum transfer of $q=0.475 \text{ fm}^{-1}$ [81, 82]. PREX-II targeted the same location to achieve the originally expected uncertainty in that measurement. The main goal of this section is to assess the impact that a measurement in a second location could have in determining the overall weak-form factor of ^{208}Pb .

The need for a measurement in a second location may be justified using simple arguments based on the nuclear mass formula of Bethe and Weizsäcker [2, 3]. Given that the nuclear force saturates, Bethe and Weizsäcker modeled the atomic nucleus as an incompressible liquid drop. The nearly uniform density found in the interior of a heavy nucleus is one of the most successful predictions of the model. However, the liquid drop is finite so a penalty must be assessed for the formation of the nuclear surface. In this way, the density of a heavy nucleus is largely characterized by two parameters: a *radius* that accounts for the distance over which the density is nearly uniform, and a *surface thickness* that controls the transition from high to low density. The corresponding nuclear form factor—which is the physical observable that is actually probed in the experiment—is obtained from the Fourier transform of the density distribution as in Eq (2.11) (for spherical nuclei). As such, the radius and the surface thickness leave a very distinct imprint on the form factor. Indeed, the nuclear form factor displays a nearly universal behavior characterized by diffractive oscillations controlled by the radius that are in turn modulated by an exponential falloff controlled by the surface thickness [159].

Clearly, a single measurement of the form factor can only constrain a linear combination of the radius and the surface thickness. This hinders the model-independent determination of the mean-square radius of the distribution as one must rely on theoretical models to lift the “degeneracy”. On the other hand, a second measurement of the form factor allows the experimental determination

of these two critical parameters. Note that we use the term “theoretical model” to refer to global models such as FSU Garnet [96] (see the first paragraph in Sec. 2.2).

In this chapter, we assess the impact of a second measurement of the weak-charge form factor of ^{208}Pb by introducing, or rather re-introducing, a highly convenient two-parameter characterization of the density distribution: the *Symmetrized Fermi function* (see Ref. [160] and references contained therein). For heavy nuclei with a radius parameter that is significantly larger than the surface thickness, the Symmetrized Fermi (SF) function is practically indistinguishable from the standard Fermi (or Woods-Saxon) parametrization. However, unlike the standard Fermi function that displays a “cusp” at the center of the nucleus, the SF parametrization is an even function, having its derivative equal to zero in the nucleus center. In addition, the SF offers a unique advantage over the standard Fermi function: the form factor associated to the SF function can be computed exactly in closed form. That this elegant result remains largely unknown to the nuclear physics community comes as a surprise (although see Ref. [161]). As stated in Ref. [160]: “The Symmetrized Fermi function has been known to some experts, but the least one can say is that it is not *well known* generally. None of the text books on nuclear physics refers to it”. On the other hand, a well-known parametrization of the nuclear form factor—with a density that is also known in closed form—was introduced by Helm almost 6 decades ago [162]. However, the Helm form factor has a Gaussian falloff rather than the more realistic exponential falloff displayed by the SF form factor.

Although presently plagued by large and uncontrolled uncertainties, hadronic reactions provide the only viable tool to probe the neutron-rich skin of exotic nuclei. Indeed, one of the main science drivers of the Facility for Rare Isotope Beams (FRIB) is the study of nuclei with neutron skins three or four times thicker than is currently possible [163]. Thus, in order to reduce the hadronic uncertainties, it is imperative to find model-independent probes. As we show below, two electroweak measurements of the weak form factor of ^{208}Pb are sufficient to constrain the half-density radius and the surface difussness of the corresponding Symmetrized Fermi distribution. In this manner, besides providing the first model-independent (once again, global models) picture of the neutron distribution of ^{208}Pb , these two electroweak measurements could supply critical calibrating anchors for future hadronic experiments at FRIB.

In the following sections, we develop the formalism required to assess the role of a second electroweak measurement of the weak form factor of ^{208}Pb . We introduce the SF and Helm functions underscoring that both have density distributions and form factors that are known in closed analytic form. We also discuss the simple Bayesian approach that we implement to constrain the two parameters that define the SF and Helm form factors. This approach follows directly from our discussion in Sec. 2.2.2.

5.1.1 Symmetrized Fermi Form Factor

The Woods-Saxon, or Fermi, function was introduced more than six decades ago to describe nucleon-nucleus scattering [164]. Given that the nucleon-nucleon interaction is of short range relative to the overall size of the nucleus, the mean-field potential that the nucleon scatters from resembles the underlying nuclear density. This density is fairly described in terms of a two-parameter Fermi shape. The conventional Fermi function is defined as follows:

$$f_{\text{F}}(r) = \frac{1}{1 + e^{(r-c)/a}}, \quad (5.1)$$

where c is the “half-density radius” and a the “surface diffuseness”.

Practically identical to the Fermi function in the relevant nuclear domain of $c \gg a$, the *Symmetrized Fermi* function is defined as:

$$f_{\text{SF}}(r) \equiv f_{\text{F}}(r) + f_{\text{F}}(-r) - 1 = \left(\frac{1}{1 + e^{(r-c)/a}} - \frac{1}{1 + e^{(r+c)/a}} \right) = \frac{\sinh(c/a)}{\cosh(r/a) + \cosh(c/a)}. \quad (5.2)$$

Although practically indistinguishable from the conventional Fermi function, the Symmetrized Fermi function enjoys a distinct advantage over it: it can be *evaluated in closed analytic form* [160]. That is,

$$\begin{aligned} F_{\text{SF}}(q) &= \int e^{-i\mathbf{q}\cdot\mathbf{r}} \rho_{\text{SF}}(r) d^3r = 4\pi \int_0^\infty \frac{\sin(qr)}{qr} \rho_{\text{SF}}(r) r^2 dr \\ &= \frac{3}{qc((qc)^2 + (\pi qa)^2)} \left(\frac{\pi qa}{\sinh(\pi qa)} \right) \left[\frac{\pi qa}{\tanh(\pi qa)} \sin(qc) - qc \cos(qc) \right]. \end{aligned} \quad (5.3)$$

where

$$\rho_{\text{SF}}(r) \equiv \rho_0 f_{\text{SF}}(r); \quad \rho_0 \equiv \frac{3}{4\pi c(c^2 + \pi^2 a^2)}, \quad (5.4)$$

and we have adopted the normalization convention in Eq. (2.12).

The fact that all the moments of the distribution can be evaluated exactly, is one of the many appealing features of an analytic expression for the form factor. Indeed, for low momentum transfers a Taylor series expansion of the form factor yields

$$F_{\text{SF}}(q) = 1 - \frac{q^2}{3!} R^2 + \frac{q^4}{5!} R^4 - \frac{q^6}{7!} R^6 + \dots \quad (5.5)$$

where the first three moments of the SF distribution are given by

$$R^2 \equiv \langle r^2 \rangle = \frac{3}{5} c^2 + \frac{7}{5} (\pi a)^2, \quad (5.6a)$$

$$R^4 \equiv \langle r^4 \rangle = \frac{3}{7} c^4 + \frac{18}{7} (\pi a)^2 c^2 + \frac{31}{7} (\pi a)^4, \quad (5.6b)$$

$$R^6 \equiv \langle r^6 \rangle = \frac{1}{3} c^6 + \frac{11}{3} (\pi a)^2 c^4 + \frac{239}{15} (\pi a)^4 c^2 + \frac{127}{5} (\pi a)^6. \quad (5.6c)$$

Unlike the conventional Fermi function, these expressions—and indeed all the moments of the SF distribution—are exact as they do not rely on a power series expansion in terms of the “small” parameter $\pi a/c$. Also interesting and highly insightful is the behavior of the SF form factor in the limit of high momentum transfers. Indeed, in this limit the SF form factor takes a remarkably simple form

$$F_{\text{SF}}(q) \rightarrow -6 \frac{\pi a}{\sqrt{c^2 + \pi^2 a^2}} \frac{\cos(qc + \delta)}{qc} e^{-\pi qa}, \quad (5.7)$$

$$\text{where } \tan \delta \equiv \frac{\pi a}{c}. \quad (5.8)$$

This expression encapsulates many of the insights developed more than three decades ago in the context of the conventional Fermi function. Namely, that for large momentum transfers the oscillations in the form factor are controlled by the half-density radius c and the exponential falloff by the diffuseness parameter a (or rather πa) [165, 159]. Again, it should be underscored that this expression is exact in the limit of high momentum transfers.

5.1.2 Helm Form Factor

Another simple yet realistic distribution that captures the main features of the form factor of heavy nuclei is the Helm function. Although much better known than the Symmetrized Fermi form factor, in the interest of completeness we provide a short summary of its most important properties. The Helm form factor was introduced over six decades ago to analyze elastic scattering of electrons from nuclei [162]. The Helm form factor is defined as the product of two fairly simple form factors:

one associated with a uniform (“box”) density, and the other accounting for a *Gaussian* falloff [162, 166, 82]. That is,

$$F_{\text{H}}(q) = F_{\text{B}}(q)F_{\text{G}}(q) = 3 \frac{j_1(qR_0)}{qR_0} e^{-q^2\sigma^2/2}, \quad (5.9)$$

where

$$F_{\text{B}}(q) = \int e^{-i\mathbf{q}\cdot\mathbf{r}} \rho_{\text{B}}(r) d^3r = \int e^{-i\mathbf{q}\cdot\mathbf{r}} \left(\frac{3\Theta(R_0-r)}{4\pi R_0^3} \right) d^3r = 3 \frac{j_1(qR_0)}{qR_0}, \quad (5.10\text{a})$$

$$F_{\text{G}}(q) = \int e^{-i\mathbf{q}\cdot\mathbf{r}} \rho_{\text{G}}(r) d^3r = \int e^{-i\mathbf{q}\cdot\mathbf{r}} \left(\frac{e^{-r^2/2\sigma^2}}{(2\pi\sigma^2)^{3/2}} \right) d^3r = e^{-q^2\sigma^2/2}. \quad (5.10\text{b})$$

Here $j_1(x)$ is the spherical Bessel function of order one:

$$j_1(x) = \frac{\sin(x)}{x^2} - \frac{\cos(x)}{x}, \quad (5.11)$$

and $\Theta(x)$ is the Heaviside theta function:

$$\Theta(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{otherwise.} \end{cases} \quad (5.12)$$

Analogous to the SF form factor, the Helm form factor is defined entirely in terms of two constants: the box (or “diffraction”) radius R_0 and the surface thickness σ . Although slightly more complicated than the form factor, a closed-form expression for the Helm density also exists. It is given by:

$$\rho_{\text{H}}(r) = \frac{1}{2} \rho_0 \left[\operatorname{erf} \left(\frac{r+R_0}{\sqrt{2}\sigma} \right) - \operatorname{erf} \left(\frac{r-R_0}{\sqrt{2}\sigma} \right) \right] + \quad (5.13)$$

$$\frac{1}{\sqrt{2\pi}} \left(\frac{\sigma}{r} \right) \rho_0 \left[\exp \left(-\frac{(r+R_0)^2}{2\sigma^2} \right) - \exp \left(-\frac{(r-R_0)^2}{2\sigma^2} \right) \right], \quad (5.14)$$

$$\text{where } \rho_0 \equiv \frac{3}{4\pi R_0^3}, \quad (5.15)$$

where $\operatorname{erf}(x)$ is the error function:

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-z^2} dz. \quad (5.16)$$

As in the case of the Symmetrized Fermi function, the Helm form factor has been normalized to $F_{\text{H}}(q=0)=1$. Finally, the first three moments of the Helm distribution are given by the following

simple expressions:

$$R^2 \equiv \langle r^2 \rangle = \frac{3}{5} R_0^2 + 3\sigma^2, \quad (5.17a)$$

$$R^4 \equiv \langle r^4 \rangle = \frac{3}{7} R_0^4 + 6R_0^2\sigma^2 + 15\sigma^4, \quad (5.17b)$$

$$R^6 \equiv \langle r^6 \rangle = \frac{1}{3} R_0^6 + 9R_0^4\sigma^2 + 63R_0^2\sigma^4 + 105\sigma^6. \quad (5.17c)$$

5.1.3 Parameter Optimization

We now proceed to discuss the determination of the two model parameters that define the SF and Helm form factors from a measurement of the experimental weak-charge form factor of ^{208}Pb . Evidently, without further theoretical assumptions, it is impossible to constrain both model parameters from our current knowledge, namely, a single measurement of the form factor; see Eqs. (5.6) and (5.17). In the particular case of PREX—where the weak form factor was extracted at a single q -point—constraints on the surface thickness σ of the Helm model were obtained by analyzing the theoretical predictions of several mean-field models. This lead to a *theoretical* uncertainty in the determination of σ of about 10% [82], which was ultimately incorporated into the final estimate of the weak-charge radius of ^{208}Pb . Our aim is to demonstrate that measuring the weak form factor at a suitable second point minimizes the reliance on theoretical models.

Naturally, the selection of the first q -point should match the PREX momentum transfer of $q_1 = 0.475 \text{ fm}^{-1}$. Given this single data point, how accurately can we constrain the weak form factor of ^{208}Pb and in particular its mean-square radius? To answer this question from a rigorous statistical perspective, we must construct the *likelihood function* defined in Eq. (2.22), where χ^2 takes the following form:

$$\chi^2(\mathbf{Y}, \boldsymbol{\omega}) = \frac{(F_{\text{SF}}(q_1, \boldsymbol{\omega}) - F_{\text{exp}}(q_1))^2}{\Delta F_{\text{exp}}^2(q_1)}. \quad (5.18)$$

We identify the experimental observation $\mathbf{Y} = y_1 = F_{\text{exp}}(q_1)$ and $\sigma_1 = \Delta F_{\text{exp}}$ defines the experimental error. In this case $\boldsymbol{\omega}$ denotes either the two model parameters of the Symmetrized Fermi function $F_{\text{SF}}(q)$ $\boldsymbol{\omega} \equiv \{a, c\}$, or the two model parameters of the Helm form factor $\boldsymbol{\omega} \equiv \{R_0, \sigma\}$. Note that the likelihood function $P(\mathbf{Y}|\boldsymbol{\omega})$ represents the probability that a given set of SF parameters describes the observed outcome, namely, the measured experimental form factor. Often, however,

one may refine the probability distribution by including our own biases and intuition. For example, as indicated in Ref. [82], mean-field predictions of the weak-charge density of ^{208}Pb suggest a Helm surface thickness of $\sigma = (1.02 \pm 0.09)$ fm. Such biases may then be incorporated into the *prior* probability $P(\boldsymbol{\omega})$ that represents the best estimate of the model parameters prior to the realization of the experiment(s). As in Eq. (2.23a) for the prior we adopt a fairly broad Gaussian distribution (see Table 5.1) centered around the predictions of an accurately-calibrated set of mean-field models [96]. That is,

$$P(\boldsymbol{\omega}) = \prod_{i=1}^2 \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-(\omega_i - \bar{\omega}_i)^2/2\sigma_i^2}. \quad (5.19)$$

To provide the connection between our own theoretical biases (encoded in the prior) and the experimental measurement (encoded in the likelihood), we rely on Bayes' theorem and write the posterior distribution as in Eq. (2.24):

$$P(\boldsymbol{\omega}|\mathbf{Y}) = \frac{P(\mathbf{Y}|\boldsymbol{\omega})P(\boldsymbol{\omega})}{P(\mathbf{Y})}, \quad (5.20)$$

where $P(\mathbf{Y})$ is known as the *marginal likelihood* [132]. The *posterior* probability density $P(\boldsymbol{\omega}|\mathbf{Y})$, namely the updated beliefs about $\boldsymbol{\omega}$ given the result of the measurement, represents the probability of the model parameters given the experimental evidence, namely, the form factor. In principle, the implementation of Bayes' theorem requires the calculation of the marginal likelihood. In practice, however, the calculation of this term—as well as any other normalization factor independent of $\boldsymbol{\omega}$ —may be bypassed through the use of Monte Carlo methods. Another method would be the use of the Gaussian approximation in which we assume that $P(\boldsymbol{\omega}|\mathbf{Y})$ is sufficiently close to a Gaussian distribution and can be characterized just by a maximum $\boldsymbol{\omega}_0$ and a covariance matrix $\tilde{\mathcal{H}}^{-1}$ (see Eq. (2.25)). In this case, integrals like Eq. (2.21a) and Eq. (2.21b) can be done analytically.

Our goal when selecting the second point is to lift the degeneracy among the many models that are consistent with the single PREX measurement. To do so, we search for a region in q that maximizes the variability among the theoretical predictions. Perhaps not surprising, this happens near the first diffraction maximum (i.e., near the first maximum in $|F(q)|$ away from $q=0$). Once the two values of the momentum transfer (q_1 and q_2) have been selected, the determination of the model parameters follows from a likelihood function suitably augmented relative to Eq. (5.18).

That is, the augmented objective function $\chi^2(\mathbf{Y}, \boldsymbol{\omega})$ now becomes:

$$\chi^2(\mathbf{Y}, \boldsymbol{\omega}) = \frac{(F_{\text{SF}}(q_1; \boldsymbol{\omega}) - F_{\text{exp}}(q_1))^2}{\Delta F_{\text{exp}}^2(q_1)} + \frac{(F_{\text{SF}}(q_2; \boldsymbol{\omega}) - F_{\text{exp}}(q_2))^2}{\Delta F_{\text{exp}}^2(q_2)}, \quad (5.21)$$

where we now identify $\mathbf{Y} = [y_1, y_2] = [F_{\text{exp}}(q_1), F_{\text{exp}}(q_2)]$ and $\Delta F_{\text{exp}}^2(q_2) = \sigma_2^2$ the experimental error on the second location.

5.1.4 Selection of the second value of the momentum transfer

We now proceed to examine the accuracy that may be attained from only two experimental measurements in the determination of the weak form factor of ^{208}Pb and more specifically, in its weak-charge radius. To test the soundness and reliability of the proposed method, we rely on a form factor that is known with exquisite precision over many orders of magnitude in momentum transfer: the *charge* form factor of ^{208}Pb [6]. That is, we simulate the impact of a second measurement by selecting the charge form factor of ^{208}Pb as a proxy for the weak form factor.

The selection of the second value of the momentum transfer q_2 , is motivated by our desire to lift the degeneracy between the many models that satisfy the constraint imposed by the single PREX measurement. To do so, we search in a region of q that maximizes the variability among the theoretical predictions. Namely, we define the one-point likelihood in Eq. (5.18) in terms of the charge form factor of ^{208}Pb at $q_1 = 0.5 \text{ fm}^{-1}$ which is given by $F_{\text{ch}}(q_1) = 0.210$. This choice is motivated by the existing PREX [81] measurement that resulted in a weak form factor comparable to $F_{\text{ch}}(q_1)$, namely, $F_{\text{wk}}(q_1) = 0.204 \pm 0.028 \text{ (exp)} \pm 0.001 \text{ (model)}$ [82]. For the experimental uncertainty, we assume $\Delta F_{\text{exp}}(q_1) = 0.005$. Although significantly larger than the typical error of the charge form factor, attaining such level of precision may be difficult for the weak-charge form factor—at least as presently envisioned by PREX-II. However, such precision may be achieved at the planned “Mainz Energy-Recovering Superconducting Accelerator” (MESA) facility. Finally, we adopt a fairly conservative choice for the Gaussian prior $P(\boldsymbol{\omega})$ defined in Eq. (5.19). Whereas for the central values $\bar{\omega}_i$ we rely on the predictions of an accurately calibrated model [96], for the *relative* standard deviation we adopt a fixed value of one half to account for the significant spread displayed by a large collection of mean-field models [88, 89]. The (prior) central values for both the SF and Helm distributions have been tabulated in Table 5.1. Note that we have tested other choices and found our results robust against the choice of prior.

SF	Helm
$\bar{c} = 6.683$	$\bar{R}_0 = 6.788$
$\bar{a} = 0.494$	$\bar{\sigma} = 0.880$

Table 5.1: Central values for the Gaussian distribution assumed for the prior (see Eq. (5.19)). In all cases, the *relative* standard deviation has been fixed to one half, for example, $\sigma_c = \bar{c}/2 = 3.342$ fm. All quantities are given in fm.

With the definition of the posterior distribution now firmly in place, we proceed to generate the distribution of SF parameters using a standard Metropolis Monte Carlo algorithm [167]. We underscore that the distribution of parameters is generated exclusively by the information that is presently known, namely, a single experimental measurement of the form factor that is encoded in the likelihood, and a set of theoretical biases that is embedded in the prior. Fig. 5.1 (a) shows the correlation plot between the half-density radius c and the surface diffuseness a that define the SF distribution. As expected from just a single measurement, the model parameters are highly spread. The solid red line represents the functional relation implied by an ideal—*i.e.*, error free and unbiased—single measurement of the form factor: $F_{\text{ch}}(q_1; a, c) = 0.210$. In turn, such a distribution of parameters generates the highly-asymmetric probability distribution function for the charge radius of ^{208}Pb showed in Fig. 5.1 (b). For comparison, the experimental value is $R_{\text{ch}}^{208} = 5.5012(13)$ fm [8], and is marked by the orange solid line.

As already mentioned and now clearly shown in Fig. 5.2, there is an infinite number of pairs (c, a) that passes through the single experimental point (indicated by the red circle). This plot is highly informative because it identifies the region of largest variability among the many models that satisfy the experimental constraint. This variability is quantified in terms of the relative variance in the model predictions (dark green solid line) which is maximized in the region around 0.8 fm^{-1} . This situation is highly favorable because the momentum transfer is large enough for the parity-violating asymmetry to be “sizable” (as the asymmetry scales with q^2) but not overly large for the nuclear form factor to be highly suppressed. Based on the maximum variability criterion, we select the second location to be $q_2 = 0.8 \text{ fm}^{-1}$ (light green solid line in Fig. 5.2).

5.1.5 Modeling the electric form factor

Having identified the second q -point, we are now ready to answer the central question: *How accurately can we describe the form factor of ^{208}Pb by measuring only two points?* To answer this

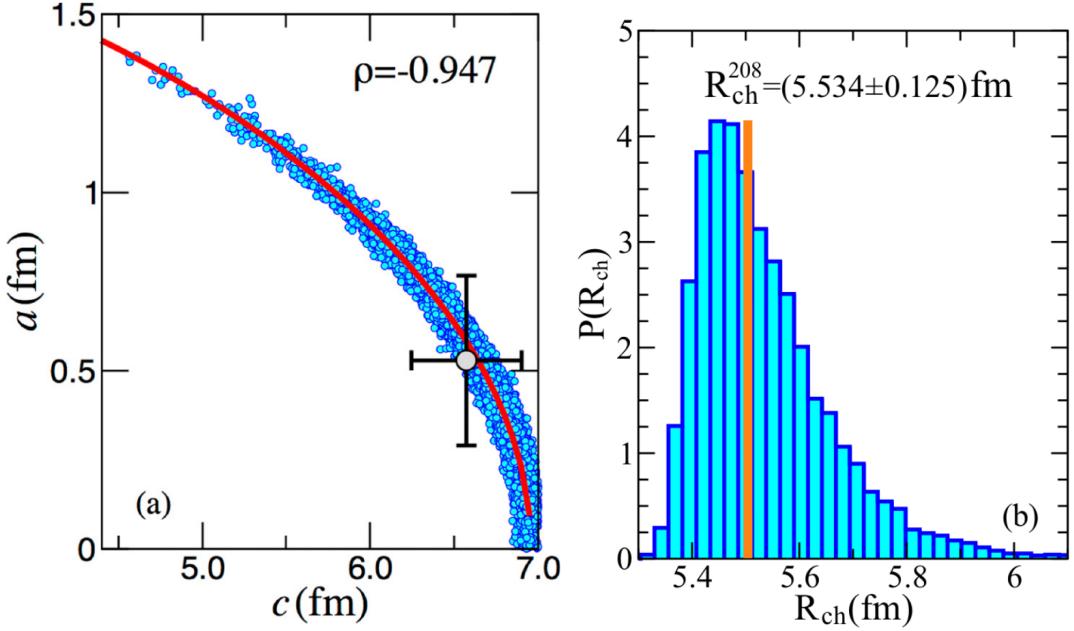


Figure 5.1: (a) Correlation plot between the half-density radius c and the surface diffuseness a that define the Symmetrized Fermi function. The number of points represents the raw data obtained from the Monte Carlo simulation. Also shown are the respective averages and standard deviations. The solid red line represents the functional form obtained from solving the equation $F_{\text{ch}}(q_1; a, c) = 0.210$. (b) Probability distribution function for the charge radius of ^{208}Pb obtained from the Monte Carlo simulation. The solid orange line marks true experimental value of R_{ch}^{208} .

question, we repeat the same procedure that we have just implemented but now with a likelihood function augmented by a second value of the charge form factor at the newly identified momentum transfer of $q_2 = 0.8 \text{ fm}^{-1}$, namely, $F_{\text{ch}}(q_2) = -0.0614$. For simplicity, we assume the same prior distribution and the same experimental error as before, that is, $\Delta F_{\text{exp}}(q_1) = \Delta F_{\text{exp}}(q_2) = 0.005$.

The improvement in our knowledge of the underlying Symmetrized Fermi function is readily apparent in Fig. 5.3. This figure is the analog of Fig. 5.1, although note the disparity in scales between the two figures. The distribution of parameters is displayed on the left-hand panel of Fig. 5.3 together with the 39% (in yellow) and 95% (in blue) confidence ellipsoids. Note that now the posterior distribution is constrained by two functional relations: $F_{\text{ch}}(q_1; a, c) = 0.210$ (solid red line, and $F_{\text{ch}}(q_2; a, c) = -0.061$ (solid purple line) which together provide nearly “orthogonal” constraints that set stringent limits on both the half-density radius c and the surface diffuseness a (see Table 5.2). Note that the theoretical errors quoted in Table 5.2, namely, about 1% for c and

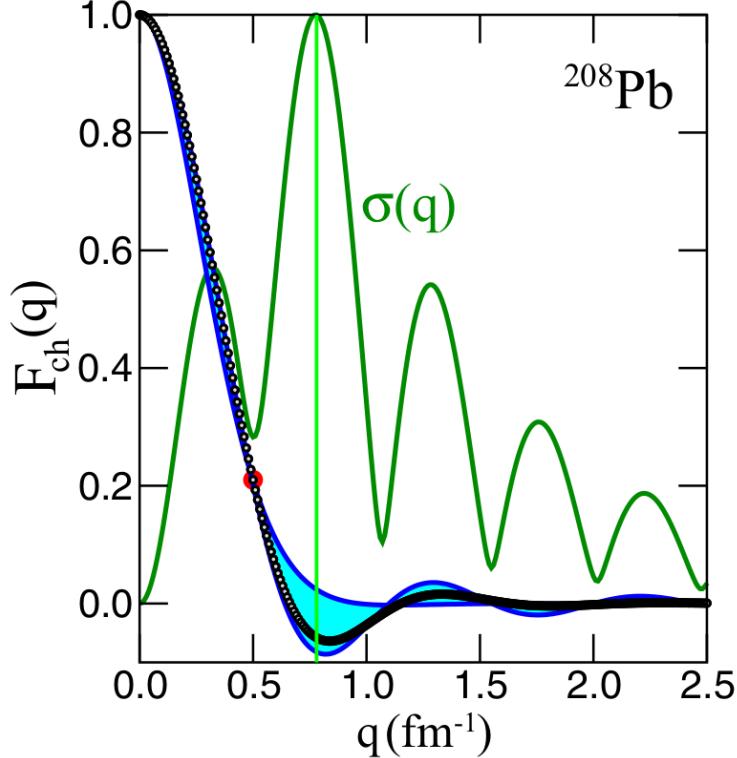


Figure 5.2: Variability in the charge form factor of ^{208}Pb generated by the distribution of parameters displayed in Fig. 5.1. Also shown with the dark green solid line is the relative uncertainty in the model predictions as a function of the momentum transfer.

13% for a , were obtained by assuming that the form factor at both momentum transfers may be determined with an experimental error of 0.005. Finally, the right-hand panel in Fig. 5.3 shows the probability distribution function for the charge radius of ^{208}Pb (see Eq. (5.6)), along with the best Gaussian fit. Thus, the theoretical prediction for the charge radius of ^{208}Pb obtained from the knowledge of only two experimental points is $R_{\text{ch}}^{208} = 5.504(45)$ fm, which compares very favorably with the corresponding experimental value of $R_{\text{ch}}^{208} = 5.5012(13)$ fm [8].

Having calibrated the parameters of the Symmetrized Fermi function using two experimental points and a fairly unconstrained prior distribution, we are now in a position to examine the overall agreement between the theoretical predictions and the experimental data for the *entire* charge form factor and density. The form factor comparison is shown in Fig. 5.4a using both linear and logarithmic scales. The two isolated red points represent the two experimental measurements that were used to calibrate the model parameters (c and a). In turn, the dense collection of black

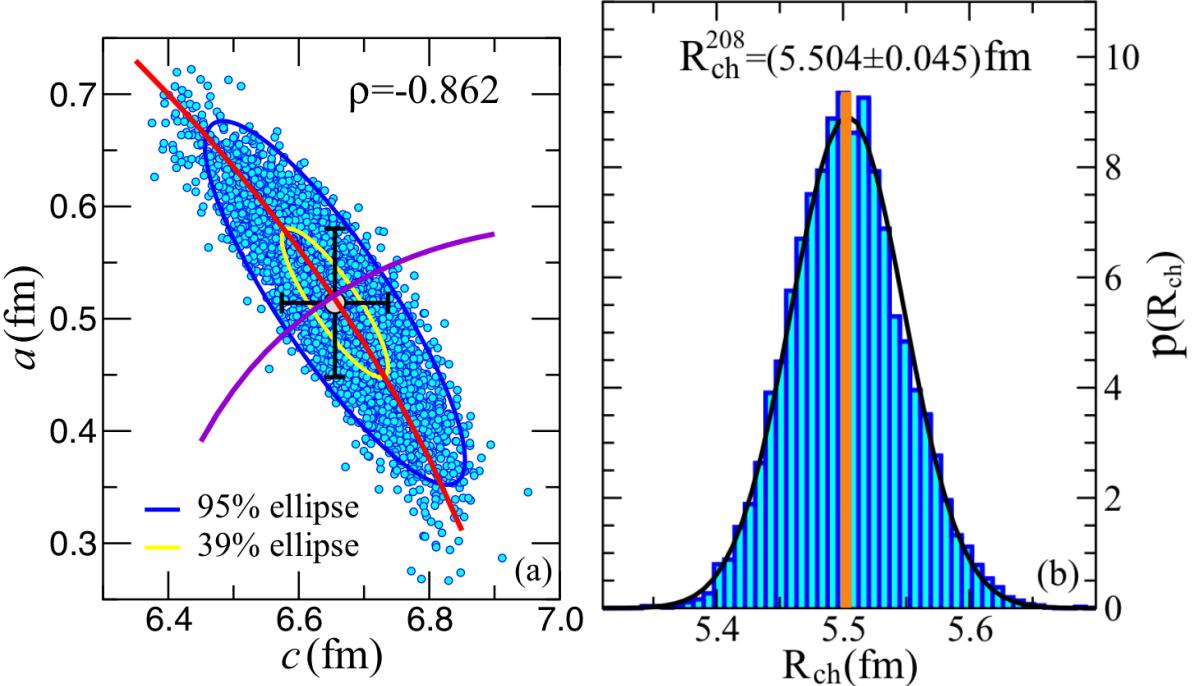


Figure 5.3: (a) Correlation plot between the half-density radius c and the surface diffuseness a that define the Symmetrized Fermi function. The number of points represents the raw data obtained from the Monte Carlo simulation. Also shown are the 39% and 95% confidence ellipses. The solid red and purple lines represent the functional form obtained from solving the equations: $F_{\text{ch}}(q_1; a, c) = 0.210$ $F_{\text{ch}}(q_2; a, c) = -0.061$. (b) Probability distribution function for the charge radius of ^{208}Pb obtained from the Monte Carlo simulation. The black solid line represents a fit to a Gaussian probability distribution and the green vertical line represents the true experimental value of R_{ch}^{208} .

points represents the full experimental form factor [6] that we aim to reproduce. Our theoretical predictions are shown with a blue solid line together with the theoretical-uncertainty band shown in cyan. On a linear scale, it is difficult to discern the agreement (or lack-thereof) between theory and experiment. Moreover, on a linear scale it is also difficult to appreciate the diffractive oscillations modulated by an exponential envelope that are the hallmark of the nuclear form factor for heavy nuclei. Thus, we display on the inset in Fig. 5.4 (a), the absolute value of the form factor using a logarithmic scale. The diffractive oscillations (controlled by c) and the exponential envelope (controlled by a), are now easily discernible. We observe a fairly good agreement between theory and experiment over several diffractive maxima up to momentum transfers well beyond the value of the second point ($q_2 = 0.8 \text{ fm}^{-1}$). However, at the largest momentum transfers displayed in

SF	Helm
$c = 6.655 \pm 0.081$	$R_0 = 6.785 \pm 0.057$
$a = 0.514 \pm 0.066$	$\sigma = 0.913 \pm 0.116$
$R_{\text{ch}} = 5.504 \pm 0.045$	$R_{\text{ch}} = 5.492 \pm 0.041$

Table 5.2: Average values and their corresponding theoretical uncertainties generated from the posterior distribution for a Symmetrized Fermi and Helm form factors. Also shown are the predictions for the charge radius of ^{208}Pb , which should be compared against the experimental value of $R_{\text{exp}}^{208} = (5.5012 \pm 0.0013) \text{ fm}$ [8]. All quantities are given in fm.

the figure, *i.e.*, $q \gtrsim 2.5 \text{ fm}^{-1}$, there is a clear deterioration in the model predictions. Finally, the associated charge density of ^{208}Pb is shown in Fig. 5.4 (b). Although the recreated density provides an excellent description of the experimental data at large distances as evinced in the inset, it fails to account for the experimental “dip” in the nuclear interior, since by construction the SF model presents a flat interior. Note that in contrast, accurately-calibrated mean-field models tend to *overestimate* the dip in the nuclear interior which is sensitive to shell effects [168]; see Fig. 5.6.

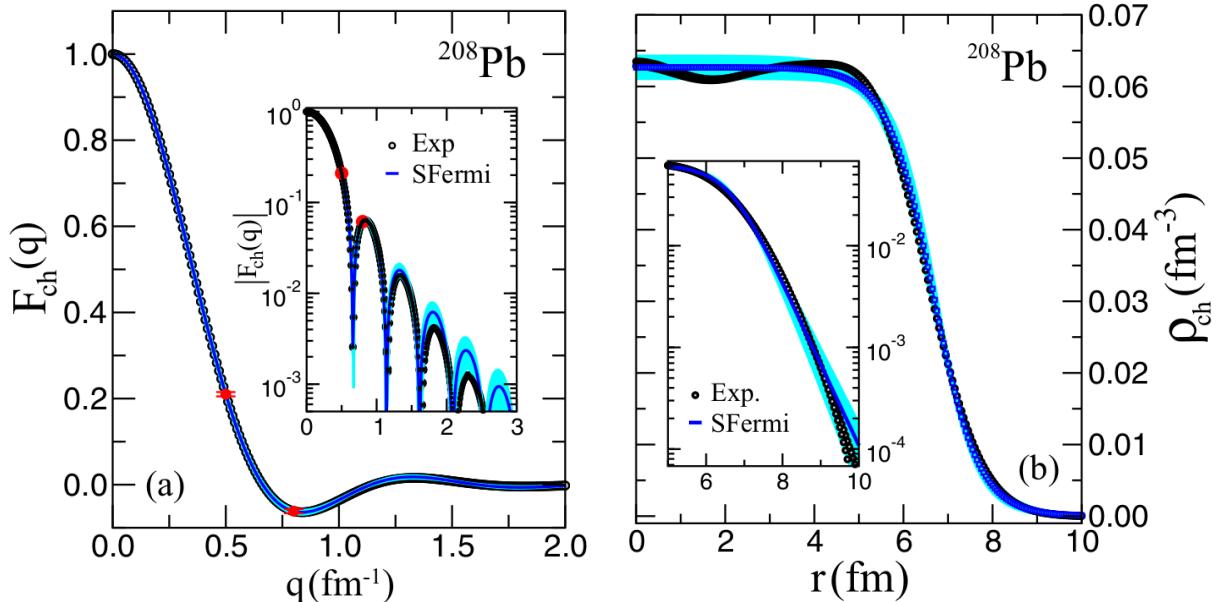


Figure 5.4: (a) Charge form factor, and (b) the corresponding charge density of ^{208}Pb . The two red points on the left-hand panel represent the single input used in the calibration of the Symmetrized Fermi function. The theoretical predictions are displayed by an uncertainty band (in cyan). The experimental data is from Ref. [6].

For completeness, we present in Fig. 5.5 the form factor and corresponding spatial density of

^{208}Pb —but now using the Helm representation. Here too the agreement with experiment is fairly good. Naturally, the great virtue of the Symmetrized Fermi and Helm parametrizations is that both the spatial density and the form factor, are known in closed analytic form. However, a distinct advantage of the former over the latter is that it displays an *exponential* rather than a Gaussian falloff at large distances. Nevertheless, it seems that the Helm parametrization stays closer to the experimental result for the high q range.

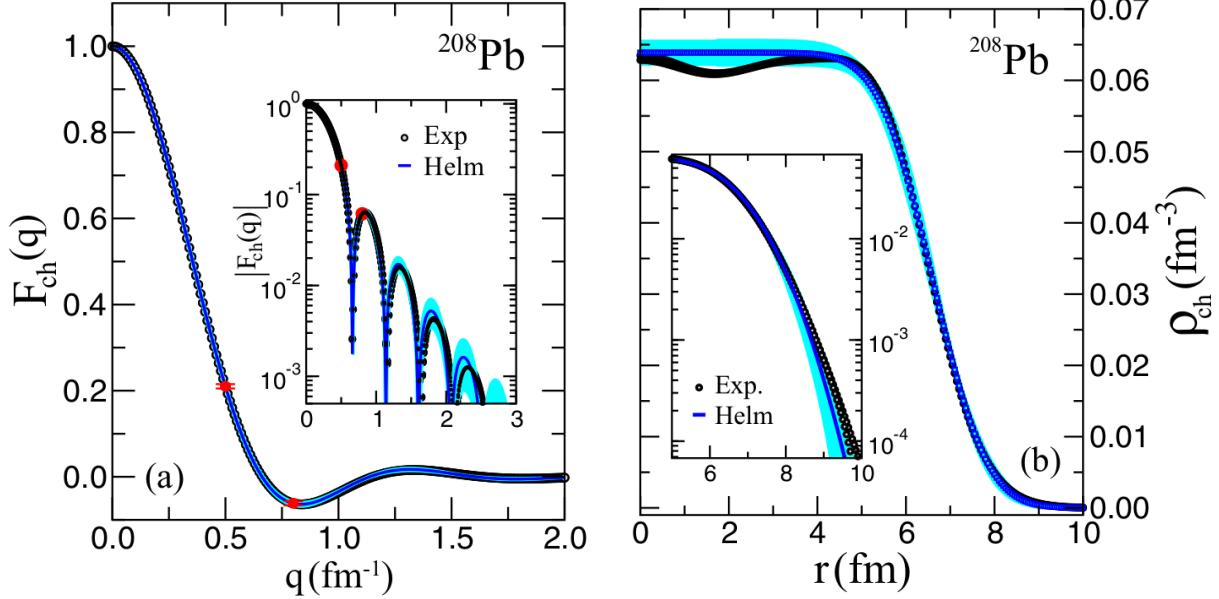


Figure 5.5: (a) Charge form factor and (b) the corresponding charge density of ^{208}Pb . The two red points represent the single input used in the calibration of the Helm function. The theoretical predictions are shown by an uncertainty band (in cyan). The experimental data is from Ref. [6].

As mentioned earlier, the main goal of this section is to assess the impact of a second electroweak measurement of the weak form factor of ^{208}Pb using exclusively statistical methods and physical insights. In particular, we aim to quantify the experimental precision required in the determination of the weak-charge (or neutron) radius of ^{208}Pb to have a strong impact on both nuclear structure and astrophysics. Using the charge form factor of ^{208}Pb as a proxy, we found that by measuring two suitable points with relative small yet attainable errors, the charge radius of ^{208}Pb could be reproduced accurately with a precision of about 0.04 fm. It is therefore natural to ask how mean-

ingful would a measurement of the *weak-charge* radius of ^{208}Pb to this precision be on constraining the density dependence of the symmetry energy.

To address this question we show in Table 5.3 predictions for several relevant quantities computed with a recent set of accurately calibrated relativistic mean field models. These models are constrained by the same isoscalar sector but differ in a single isovector assumption, namely, the choice of the neutron skin thickness of ^{208}Pb [96]. Although the set of models is relatively small, note that the theoretical spread in R_{wk}^{208} is nearly five times as large as the assumed (0.04 fm) experimental precision. The imprint of the isovector sector is illustrated in Fig. 5.6. Indeed, whereas the charge density remains practically unchanged, significant differences emerge in the weak-charge density, as the latter is dominated by the neutron distribution. Shown in the inset on a logarithmic scale are the SF and Helm fits to the RMF electric charge density that exhibits the more realistic exponential falloff of the SF density. As a figure of merit, we can establish that if R_{wk}^{208} is determined within the assumed ± 0.04 fm uncertainty, one could constrain the slope of the symmetry energy to about 15 MeV and the radius of a $1.4 M_{\odot}$ neutron star to within 1.2 km. As a comparison, the Neutron Star Interior Composition Explorer (NICER) telescope stationed in the international space station was able to determine the radius of a $1.44 M_{\odot}$ neutron star to be $13.02^{+1.24}_{-1.06}$ Km. Of course, the estimates related to the ^{208}Pb neutron skin are based on a very limited set of RMF models that suffer from their own limitations and theoretical biases. Yet, our conclusions appear consistent with other studies that incorporate a very large ensemble of reasonable nuclear energy density functionals [89, 169].

However, we note that the density dependence of the symmetry energy can be directly constrained from the measured weak form factor at the momentum transfer of the PREX experiment. Indeed, as shown in Table 5.3, the correlation between L and $F_{\text{wk}}^{\text{PREX}}$ remains strong, at least for the limited set of models used in this section. Connecting L directly to $F_{\text{wk}}^{\text{PREX}}$ would avoid ambiguities associated with the determination of both the half-density radius and the surface diffuseness from a single measurement.

Model	R_{ch}^{208}	R_{wk}^{208}	$R_{\text{wk}}^{208} - R_{\text{ch}}^{208}$	$F_{\text{wk}}^{\text{PREX}}$	L	$R(1.4M_{\odot})$
RMF-012	5.504	5.636	0.132	0.239	48.254	12.400
RMF-016	5.499	5.667	0.168	0.234	50.961	12.839
RMF-022	5.496	5.722	0.226	0.226	63.524	13.609
RMF-028	5.495	5.790	0.295	0.216	112.644	14.234
RMF-032	5.489	5.822	0.333	0.212	125.626	14.718

Table 5.3: Predictions from a set of accurately calibrated relativistic mean-field models [96] for the charge radius, weak-charge radius, and their difference for ^{208}Pb (all in fm). Also shown is the weak-charge form factor at the PREX momentum transfer, the slope of the symmetry energy L (in MeV), and the radius of a $1.4 M_{\odot}$ neutron star (in km). Some of these quantities may be compared against the following experimental values: $R_{\text{ch}}^{208} = 5.5012(13)$ fm [8], $F_{\text{wk}}^{\text{PREX}} = 0.204(28)$, and $R_{\text{wk}}^{208} = 5.826(181)$ fm [82].

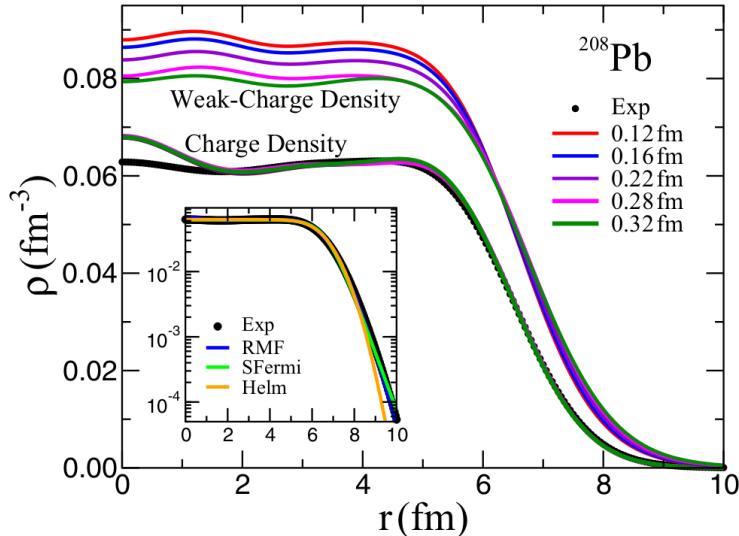


Figure 5.6: Charge and weak-charge density of ^{208}Pb as predicted by a collection of accurately calibrated mean-field models. The labels indicate the predicted neutron-skin thickness of ^{208}Pb . The inset shows the fastest Gaussian falloff of the Helm form factor relative to the exponential falloff of Symmetrized Fermi function. The experimental data are from Ref. [6].

5.2 The transfer function formalism applied to uncertainty quantification on weak charge densities reconstruction

In this section, we use the transfer function formalism to analyze the extraction of the electric and weak charge densities of ^{48}Ca and ^{208}Pb from several form factor data points. Various parametrizations (or models) exist in the literature to describe nuclear densities and their associ-

ated form factors [170]. We study the performance of seven models: Fourier Bessel [128], Helm [171], Symmetrized Fermi Function (SF) [127] 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 Gaussian (SF+G). Note that we did not consider the original Sum of Gaussians model [172] since certain conditions were difficult to implement within the transfer function formalism. Moreover, we found that for the small (5) number of data points here considered, the Sum of Gaussians did not provide a good fit to the data. Appendix C.1 describes in detail the seven models we employ. Note that we already discussed the SF and Helm models in Sec 5.1.

Our objective is to identify which of these seven models will best perform 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. These are the *generators* discussed in Sec. 2.2.3: ${}^nF_{\text{true}}(q)$ ($n = 1, \dots, 5$). The particular set of accurately calibrated functionals are: RMF012, RMF016 (commonly referred to as “FSUGarnet”), RMF022, RMF028 and RMF032 [96]. The main difference among these generators is the assumed value for the yet to be accurately determined neutron skin thickness of ^{208}Pb ; for example, RMF022 predicts a neutron skin thickness of ~ 0.22 fm. The predictions for the ^{208}Pb weak density of these five generators are shown in Fig. 5.6. As discussed in Sec. 2.2.3, for each data point generated for the weak charge form factor there is an associated error σ_j which resembles realistic experimental uncertainties. Once a generator is selected, any observable of interest m (for example, the radius or interior density) can be calculated directly from the synthetic data.

As in Refs.[61], we evaluate the performance of each of the seven models using a bias-variance trade-off criterion. As we discussed in Sec. 2.2.3 the “bias-variance trade-off” is an important concept in statistics and in 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 trade-off is the inevitable conflict that ensues when trying to simultaneously minimize these two critical sources of error. As we discussed in Chapter 3, we can calculate both the bias and variance through the transfer function formalism.

Given that each single measurement in the data has an associated transfer function, an important feature of the formalism is that one can identify those *critical points*, if any, that are responsible for driving most of the uncertainty. For example, we are interested in quantifying the statistical error in the extracted weak charge radii of both ^{48}Ca and ^{208}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. 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.

Before doing the full analysis on the electric and weak charge densities of ^{48}Ca and ^{208}Pb , we explore the estimation of a non intuitive bias which we call the reconstruction bias, that strongly depends on the q_j locations. This type of bias is not present in all the models, but when it is present it can completely hinder the ability of the model to extract the information in a reliable way.

5.2.1 Reconstruction Bias and the Optimal Function

The reconstruction bias is closely related to what we observed in the example in Fig. 3.2. When dealing with incomplete data (a few q_j points on the entire form factor curve, for example), the empirical parameters ω_e we recover might deviate considerably from the best parameters that reproduce the entire true function: ω_{Cen} in the case of $F(q)_{\text{true}}$, or ω_{Opt} in the case of $\rho(r)_{\text{true}}$ (see Sec. 3.3). As a consequence, the second term inside the brackets in Eq (3.11) 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. Fig. 5.7 (a) shows the recovered ^{48}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 data set 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. 5.7 (b).

The blue and orange SF+G model in Fig. 5.7 has its hyperparameter controlling the size of the Gaussians set to $\gamma = 0.7 \text{ fm}$, close to the nucleon size (Appendix C.1 shows a detailed description of the SF+G model and its hyperparameters). The orange curve has a clear bias in the interior

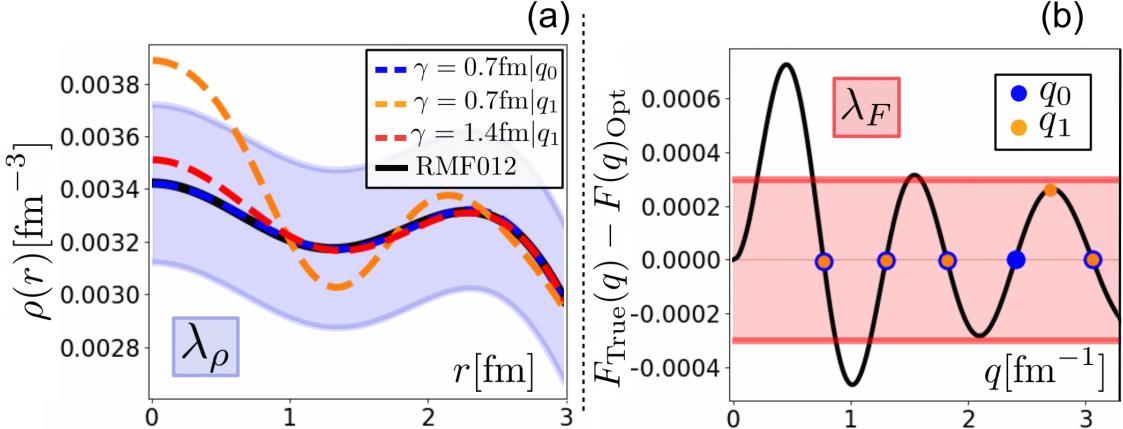


Figure 5.7: (a) Recovered densities using the SF+G model for two different data locations. The generated (“true”) ^{48}Ca weak density is shown in black, while the orange and blue dashed lines are the SF+G obtained densities for the two data sets \mathbf{q}_0 and \mathbf{q}_1 . The red dashed line is obtained by the same SF+G model on \mathbf{q}_1 , but with the hyperparameter γ set to 1.4 fm instead of 0.7 fm. The blue band is associated with the scale λ_ρ . (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 λ_F .

density. This is the reconstruction bias. It is not the same type of bias showed by, for example, the SF model which by definition has a flat interior and can not reproduce the interior structure of ^{48}Ca . To better analyze this phenomenon, we use the optimal function, i.e., the parameter set ω_{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 ω_{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 $\rho(0)$.

Fig. 5.7 (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 3.3, we can imagine that our data are currently centered at the optimal function $F(q, \omega_{\text{Opt}})$ (the minimum of χ^2 is currently at the optimal parameters). The y_j values are slightly perturbed from their starting values by small quantities η_j now defined

as:

$$\eta_j \equiv F(q_j)_{\text{True}} - F(q_j, \omega_{\text{Opt}}). \quad (5.22)$$

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 η_{y_4} is nonzero, we have:

$$\delta\rho(0) = [\mathcal{T}\mathcal{F}_4^{(\rho(0))}] \eta_{y_4}. \quad (5.23)$$

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 showed in Fig. 5.7 (a) by the orange curve. In this particular example, if we want to maintain a bias of less than five percent ($\delta\rho(0) < 5\%$), we can only locate q_4 in around 15% of the possible momentum transfer range $[0 - 3.5] \text{ fm}^{-1}$ (see Appendix C.2 for more details).

Let us call λ_F the expected scale for the size of η_j for our range of q_j values. Let us call λ_ρ the desired threshold we want for our accuracy in the estimation of $\rho(0)$. Fig. 5.7 (a) and (b) shows graphically these two scales as the red and blue bands, respectively. Replacing the transfer function by its explicit expression in Eq. (5.23), we will maintain that threshold in $\rho(0)$ as long as:

$$|\nabla\rho(0)\mathcal{H}^{-1}\nabla F(q_4)/\sigma_4^2| \leq \frac{\lambda_\rho}{\lambda_F}. \quad (5.24)$$

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 λ_ρ . In a real experiment, we would not know beforehand 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 γ^{-3} in the ρ space. Based on this, we decided to double the size of γ , from $\gamma = 0.7 \text{ fm}$, to $\gamma = 1.4 \text{ fm}$ (which is the value we use in Sec. 5.2.4).

Using this new value of γ , 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 considerably¹. This is shown by

¹To 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.

the red dashed line in Fig. 5.7 (a). Moreover, with this new value of γ , q_4 can be allocated in around 30% of the possible momentum transfer range $[0 - 3.5]$ fm $^{-1}$ while maintaining a bias of less than five percent (see Appendix C.2 for more details).

We close this subsection 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, an hyperparameter (such as γ) might be fixed to a sub-optimal 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 Fig. 3.1 and Fig. 3.2 (a) centered at ω_c , the red star. This ellipse does not contain the actual estimated parameters from the data, i.e., the green star (3) in Fig. 3.1 (let us recall that the blue star in Fig. 3.2 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. 3.2 better than the approximated empirical blue density.

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

5.2.2 Results: Analyzing charge and weak charge densities

In this subsection, we discuss the process to select the optimal models and the impact that different locations q_j have on the extracted densities of the ^{48}Ca and ^{208}Pb nuclei. We are interested in describing both the mean radius and the interior density of the charge and weak charge distributions.

The MSE calculation for the charge radius is straightforward since it is a single quantity. For the interior density we allocate 30 grid points between $r = 0$ fm and $r = 3$ fm for ^{48}Ca and $r = 5$ fm for ^{208}Pb . The interior MSE is then constructed by averaging in quadrature the single MSE for

each point:

$$\text{MSE[Interior]}^2 \equiv \frac{1}{30} \sum_{i=1}^{30} \text{MSE}[\rho(r_i)]^2. \quad (5.25)$$

We combine both the radius and interior MSE into a single quantity known as the Figure of Merit (FOM):

$$\text{FOM}^2 \equiv \left(\frac{\text{MSE[Radius]}}{\Delta R} \right)^2 + \left(\frac{\text{MSE[Interior]}}{\Delta \rho} \right)^2, \quad (5.26)$$

where $\Delta \rho$ and ΔR are the expected MSE size (natural scales) in each quantity and for each nucleus (roughly 5% – 10% and 1% uncertainty in their respective densities and radius). Table 5.4 shows their numerical values. By changing these scales, the FOM will be more driven by the radius or the interior.

	$\Delta \rho [\text{fm}^{-3}]$	$\Delta R [\text{fm}]$
^{48}Ca	0.00015	0.04
^{208}Pb	0.00008	0.06

Table 5.4: Natural scales for the uncertainties in the interior density and radius for ^{48}Ca and ^{208}Pb . A subscript name is attached when referring to a specific nucleus. For example, the natural scale for the MSE of the radius of ^{48}Ca is denoted as ΔR_{Ca}

To simplify our analysis, we assume that the errors in the data (both electric and weak) depend only on their respective q_j value. Following our error estimate in Sec. 5.1, we assume a constant value of $\sigma(q) = 0.005$ for the ^{208}Pb data. Following the estimated errors for five momentum transfer values $[q_1, q_2, q_3, q_4, q_5]$ reported in [97], we create a truncated linear interpolation between these q values of $\sigma(q) = \text{Max}[0.00057, 0.0081 - 0.003q]$ for the ^{48}Ca data.

Once the data set and the model used are specified, the FOM depends solely on the measurement locations q_j . In the following subsection, we show how the value of the FOM can be minimized by changing the locations of the q_j , and how the TF formalism aids in interpreting the results.

5.2.3 Example: ^{48}Ca and ^{208}Pb charge density analysis

To test our framework, we first analyze the experimentally well constrained electric charge density [6, 7]. We illustrate some aspects of the transfer functions formalism by analyzing two examples: the electric ^{48}Ca data with the Bessel model, and the electric ^{208}Pb data with the SF model. Following [97], we assign five possible q_j starting data locations for ^{48}Ca : $\mathbf{q}_0 = [0.9, 1.35, 1.8, 2.24, 2.69] \text{ fm}^{-1}$. To compare with our results from Sec. 5.1, we assign two possible starting locations for ^{208}Pb :

$\mathbf{q}_0 = [0.5, 0.8] \text{ fm}^{-1}$. In Sec. 5.1, the first momentum transfer value was fixed at 0.5 fm^{-1} , and the best location we found for the second one was precisely 0.8 fm^{-1} .

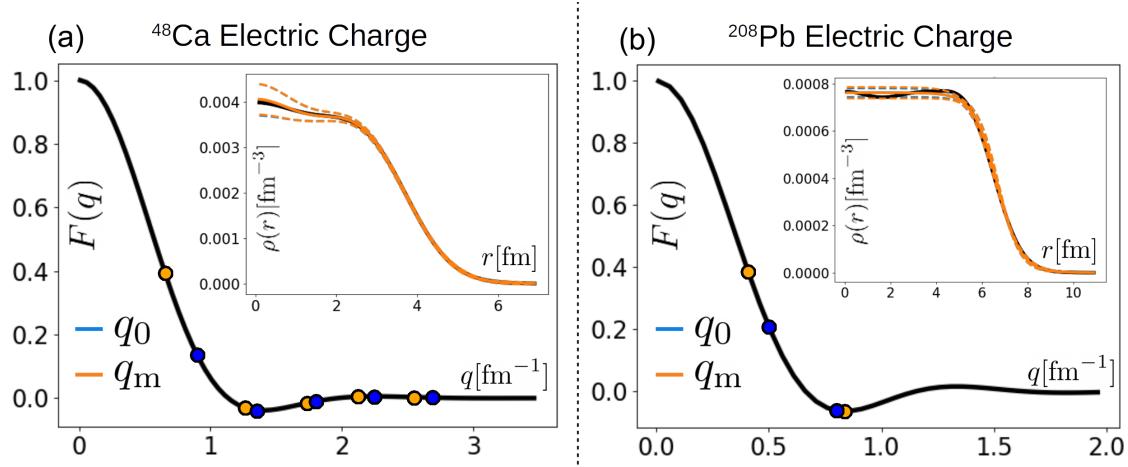


Figure 5.8: Results for optimizing the FOM for ^{48}Ca using the Bessel model (a), and for ^{208}Pb using the SF model (b). The original momentum transfer locations q_0 (blue points) are displaced to q_m (orange points) to minimize the FOM. The inset plots show the reconstructed charge densities with their respective error bands.

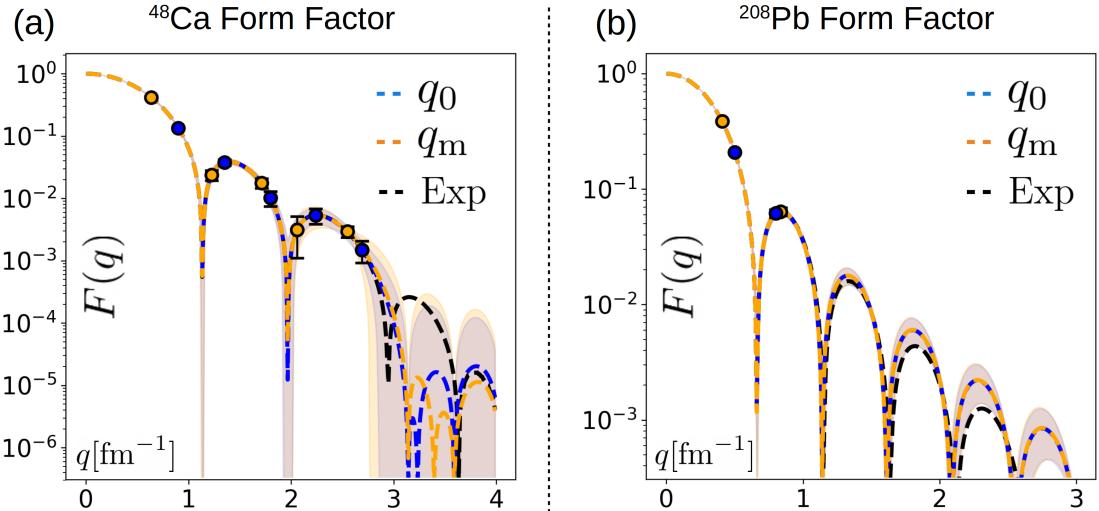


Figure 5.9: Log plots of the electric form factors recovered in Fig. 5.8. (a) Shows the results for ^{48}Ca using the Bessel model. (b) shows the results for ^{208}Pb using the SF model. For both (a) and (b) the original momentum transfer locations q_0 and their respective fitted form factors are shown in blue, while the optimized locations q_m and their respective fitted form factors are shown in orange. The black curve represents the respective experimental charge form factor [6, 7]. For clarity, the central curves are displayed as dashed lines, while the error bands are shown as colored bands.

Fig. 5.8 shows the results of both nuclei when the original data locations \mathbf{q}_0 (in blue) are relocated to optimal places \mathbf{q}_m (in orange) where the FOM is minimized with the restriction that $q_j \leq 3.5 \text{ fm}^{-1}$ for ^{48}Ca and $q_j \leq 2 \text{ fm}^{-1}$ for ^{208}Pb . The minimization is done by running the python numpy optimization library with the ‘TNC’ method for 10 different seeds including the respective original positions \mathbf{q}_0 .

Interestingly, for ^{208}Pb the second point optimal location is no longer 0.8 fm^{-1} , but rather 0.84 fm^{-1} (see Table C.2 in Appendix C.3). The main reason for this difference when compared to the results in Sec. 5.1 is that we are allowing the first point to move as well: is no longer fixed at 0.5 fm^{-1} . This example illustrates a very important concept: each optimal location is not independent, its value will rely on the position of the rest of the data. This in turn implies that, when planning an experiment, minimizing the uncertainty on each step as we are measuring will probably be sub-optimal.

	^{48}Ca		^{208}Pb	
	Int	Rad	Int	Rad
MSE (q_0)	1.27	1.37	0.27	0.77
MSE (q_m)	1.26	0.94	0.32	0.63

Table 5.5: MSE results for the interior (Int) and radius (Rad) for ^{48}Ca and ^{208}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 5.4.

Table 5.5 shows the numerical values of the MSE for both nuclei in units of their natural scales. The MSE in the interior was not substantially reduced for ^{48}Ca . For ^{208}Pb , it actually increased by $\sim 20\%$, as can be seen in the error bands in Fig. 5.8. The MSE in the radius on the other hand, was improved by $\sim 30\%$ for ^{48}Ca , and by $\sim 20\%$ for ^{208}Pb . These results are explained by the scales we selected for the FOM which favored radius improvements, and also by the fact that the radius is an easier quantity to constraint than the interior density².

Tables C.1 and C.2 in Appendix C.3 present the numerical values of the $\mathcal{T}\mathcal{F}_j$ times the respective error σ_j for the density at $r = 0 \text{ fm}$ and the radius for both data locations \mathbf{q}_0 and \mathbf{q}_m . These values can provide an idea of how much each measurement is currently impacting the variance of the density at $r = 0$ (as a representative of the entire interior) and radius³.

²Minimizing the FOM with only the interior term included does not significantly improve the interior density.

³Note that in Eq(3.6) each term $\mathcal{T}\mathcal{F}_j\sigma_j$ is added in quadrature. Therefore, the final variance is not linear on each component. The quadrature equation will enhance the effect of bigger numbers with respect to their smaller counterparts.

In ^{48}Ca for the final set \mathbf{q}_m , the $\rho(0)$ variance is mainly driven by observations q_2 and q_3 , while the radius variance is driven by observation q_1 . For the ^{208}Pb nucleus in the final set \mathbf{q}_m , the $\rho(0)$ variance is driven by q_2 while the radius variance is driven by q_1 . As the errors in the observations change, these statements might not hold. Our main conclusion is that to reduce the final variance in each quantity m within this hypothetical experiment and to first approximation, these are the critical data locations that should be targeted for error reduction.

Fig. 5.9 shows the reproduced form factors of both nuclei across several orders of magnitude. It seems that, once we move beyond the measurement locations in ^{48}Ca , both Bessel expansions start to deviate considerably from the experimental curve. Both SF fits in the case of ^{208}Pb seem more robust in that sense. It is also interesting how, for ^{208}Pb , both q_0 and q_m reproduce basically the same form factor, but the errors for the radius gets propagated $\sim 20\%$ less for the second set.

5.2.4 Model Comparison: ^{48}Ca and ^{208}Pb Weak Charge Density

We now proceed to compare the performance of each of the seven models (see Appendix C.1) when recovering the interior density and radius of the weak charge of ^{48}Ca and ^{208}Pb . Appendix C.3 presents the analogous analysis for the charge densities.

We use Eq (2.28) to calculate the squared average MSE from the five different generators (RMF012, RMF016, RMF022, RMF028 and RMF032 [113]). For all the models, we start with five fixed locations \mathbf{q}_0 and find the \mathbf{q}_m that minimizes the average FOM. We apply the same restrictions as in the electric example: $q_j \leq 3.5 \text{ fm}^{-1}$ for ^{48}Ca and $q_j \leq 2 \text{ fm}^{-1}$ for ^{208}Pb . For a fixed model, the obtained \mathbf{q}_m will be the same for all generators since we are minimizing Eq (2.28). For example, for the ^{48}Ca data, the Bessel and SF model will have different sets of \mathbf{q}_m , but each individual set will be the same across all RMF.

The starting locations for ^{48}Ca are, once again, $\mathbf{q}_0 = [0.9, 1.35, 1.8, 2.24, 2.69] \text{ fm}^{-1}$, and for ^{208}Pb are now $\mathbf{q}_0 = [0.63, 0.94, 1.26, 1.57, 1.88] \text{ fm}^{-1}$. Following [97], these values correspond to the Bessel special momentum transfer values $q_\nu \equiv \nu\pi/R_{\text{cut}}$ for $\nu \in [2, 6]$. The corresponding R_{cut} values are 7 fm for ^{48}Ca , and 10 fm for ^{208}Pb .

Fig. 5.10 and Fig. 5.11 show the resulting square average bias, SD ($\sqrt{\text{Variance}}$), and MSE for the seven models for the weak charge of ^{48}Ca and ^{208}Pb , respectively. Figures C.2 and C.3 in Appendix C.3 show a similar model analysis for the electric charge counterparts of both nuclei.

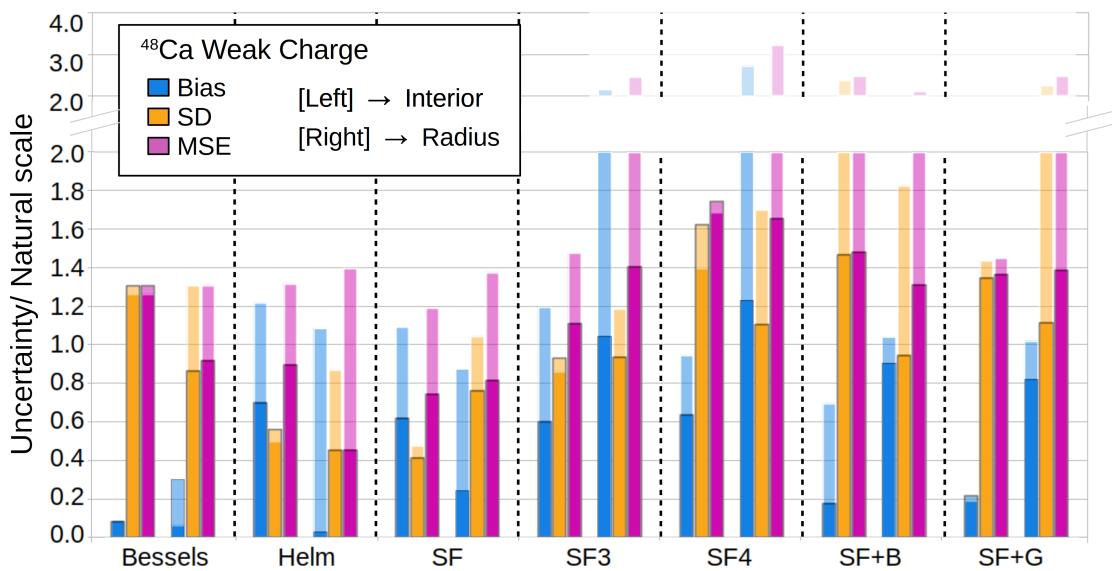


Figure 5.10: Comparison of seven models for recovering the interior density and mean weak radius from the weak form factor data on ^{48}Ca . The three columns to the left of each model show the bias, SD, and MSE for the interior density. The three columns to the right of each model show the bias, SD, 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 , while the borderless columns represent the starting q_0 .

In the weak charge case, we observed that for a fixed model the individual results among different generators (the RMF family) were very similar (see an example on Fig. C.4 and Fig. C.5 in Appendix C.4). Therefore, we are confident that the conclusions we draw regarding each model are robust, at least within the family of generators we considered in this study.

We want to highlight two main points from these results. First, changing the data locations usually has a significant impact within each model performance, for example, a reduction of a factor of two in the case of the Helm model with the ^{208}Pb nucleus. Second, there are large variations in performance across models for the fixed locations \mathbf{q}_0 , for example, a factor of two between the Bessels and SF+B in the ^{48}Ca nucleus. This difference is mitigated in some cases once each model settles on its own optimal locations \mathbf{q}_m , such as when comparing the SF and SF3 models in the ^{208}Pb nucleus.

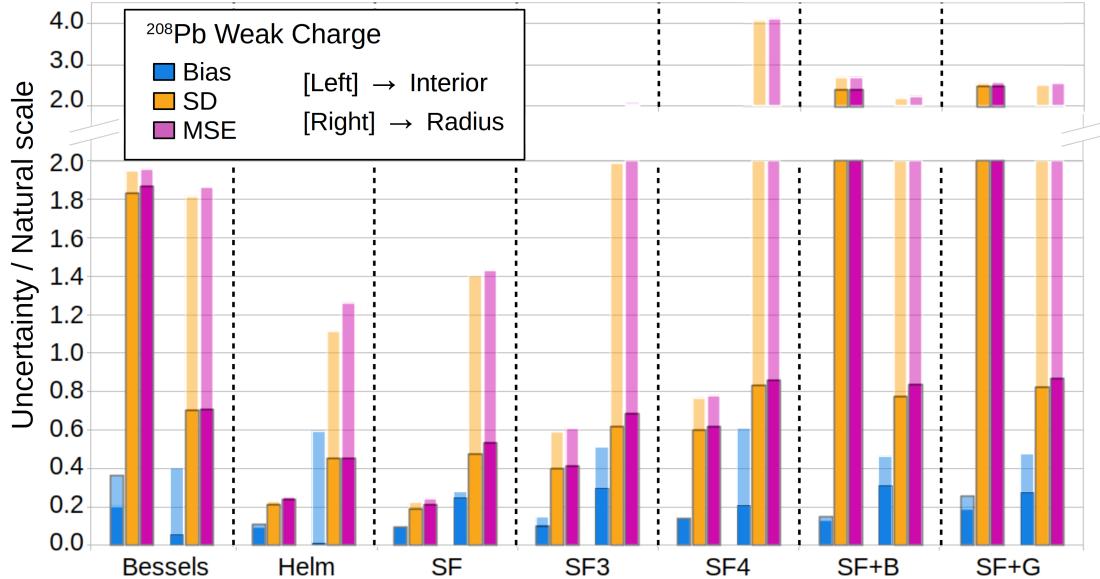


Figure 5.11: Comparison of seven models for recovering the interior density and mean weak radius from the weak form factor data on ^{208}Pb . The three columns to the left of each model show the bias, SD, and MSE for the interior density. The three columns to the right of each model show the bias, SD, 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 \mathbf{q}_m , while the borderless columns represent the starting \mathbf{q}_0 .

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 variance driven models

(like the Bessels) to outperform bias driven models (like the SF), in cases where the data errors are small.

For the data range and errors we have assumed in this study, the Helm and SF models are the best suited for extracting the radius and interior density of both nuclei ^{48}Ca and ^{208}Pb . This could be expected for the ^{208}Pb nucleus since both Helm and SF models are characterized by a flat interior which resembles the saturation density achieved by heavy nuclei.

On the other hand, it might come as a surprise that these flat models outperform more flexible models like the Bessel expansion which are capable of describing the interior oscillations of the ^{48}Ca nucleus. The reason behind this result is that we are combining both bias and variance to judge our models. On *average* (mean value), the Bessel model will recover a more genuine representation of the true oscillating interior weak density of ^{48}Ca . However, the noise levels in the data are so high that the expected deviation is large enough to make more reliable a flat description with smaller bands (SD).

As an example, consider the RMF012 generator displayed as the black curve in Fig. 5.7 (a). In this generator, the ^{48}Ca weak density starts as an upward oscillation from the average density at $r = 0$ and then drops to a downward oscillation around $r \approx 1.2$. We could ask how probable it is to conclude the opposite ($\rho(0) \leq \rho(1.2)$) while using the optimal locations \mathbf{q}_m for the Bessel model.

Let us define a new quantity of interest as $m \equiv \rho(1.2) - \rho(0)$ and investigate when $m \geq 0$. When using the fitted parameters for the RMF012 \mathbf{q}_m data, the value obtained is $m = -2.3 \times 10^{-4} \text{ fm}^{-3}$. The variance of this quantity can be calculated using the TF formalism from Eq (3.6) as:

$$\Delta m^2 = \sum_{j=1}^5 \left([\nabla \rho(1.2) - \nabla \rho(0)] \mathcal{H}^{-1} \nabla F(q_j, \omega) \sigma_j^{-2} \right)^2 \sigma_j^2, \quad (5.27)$$

where we used Eq (3.2) and Eq (3.4) to write the explicit form of $\mathcal{T}\mathcal{F}_j^m$. We obtain a value of $\Delta m = 2.2 \times 10^{-4} \text{ fm}^{-3}$. The third measurement $q_3 = 1.73 \text{ fm}^{-1}$ impacts this variance the most, followed by appreciable contributions from $q_4 = 2.12$ and $q_5 = 2.55 \text{ fm}^{-1}$ (see Table C.3 in Appendix C.3 for the values of \mathbf{q}_m).

Under our linear assumptions we find that $\rho(0) \leq \rho(1.2)$ in about 15% of the experimental realizations. Therefore, if the noise in the experimental data (mainly in q_3 , q_4 and q_5) cannot be reduced, we will conclude the wrong shell oscillation structure of ^{48}Ca in one of every six experiments.

For the interior density of ^{208}Pb , which strongly relates to the value of saturation density ρ_0 , the best score was achieved by the SF model with a total MSE of $0.21 \times \Delta\rho_{\text{Pb}} = 1.68 \times 10^{-5} \text{ fm}^{-3}$ ($\Delta\rho_{\text{Pb}}$ is defined in Table 5.4). Let us focus on RMF012 as an example. In particular, the total variance in $\rho(0)$ is driven mainly by the third observation with a value of $|\mathcal{T}\mathcal{F}_3^{\rho(0)}\sigma_3| = 0.14 \times \Delta\rho_{\text{Pb}}$. Taking $\rho(0)$ as a representative of the interior density, this means that $q_3 = 0.77 \text{ fm}^{-1}$ is a critical point. This is the point that should be targeted for error reduction to improve the uncertainty in the saturation density ρ_0 .

The radius of ^{48}Ca and ^{208}Pb could be determined with good accuracy by the Helm model: an MSE of $0.45 \times \Delta R_{\text{Ca}} = 0.018 \text{ fm}$ for ^{48}Ca and an MSE of $0.45 \times \Delta R_{\text{Pb}} = 0.027 \text{ fm}$ for ^{208}Pb (both values of ΔR are defined in Table 5.4). Let us focus again on RMF012 as an example. For ^{48}Ca , the total variance in R is uniformly distributed within the first three observations $q_1 = 0.51$, $q_2 = 0.63$ and $q_3 = 0.77 \text{ fm}^{-1}$ with values of $|\mathcal{T}\mathcal{F}_j^{\rho(0)}\sigma_j| \approx 0.2 \times \Delta R_{\text{Ca}}$. For ^{208}Pb , the total variance in R is driven by the first two close points $q_1 = 0.37$ and $q_2 = 0.40 \text{ fm}^{-1}$ with values of $|\mathcal{T}\mathcal{F}_j^{\rho(0)}\sigma_j| = 0.29 \times \Delta R_{\text{Pb}}$. To improve the weak charge radius uncertainty, these are the corresponding observations for each nuclei that should be targeted for error reduction.

5.2.5 The role of priors

The incorporation of priors allows us to include information from previous experiments when fitting the data of a new experiment. Priors also play the important role of serving as leverage to reduce the variance of a model in exchange of increasing our bias [173, 134]. This can be particularly beneficial for models such as the Bessel expansion or SF+G. The MSE of these models is completely variance driven given the level of noise in our data.

In this subsection, we briefly explore the influence that the prior can have on the performance of SF+G on extracting the weak charge density of ^{48}Ca . We use the RMF012 generator. The measured locations are the original momentum transfer values $\mathbf{q}_0 = [0.9, 1.35, 1.8, 2.24, 2.69] \text{ fm}^{-1}$. Let us recall that the implementation of priors was defined in Sec 2.2.2 and extended to the TF formalism in Sec 3.5.

The SF+G model consists of the two-parameter Symmetrized Fermi function plus three Gaussian. These Gaussians are centered in three locations at the interior density: $[R_1, R_2, R_3] = [0, 1.3, 2.6] \text{ fm}$. We analyze three prior options for the amplitude of the Gaussian (A_1, A_2, A_3) , while we leave the two intrinsic parameters of the SF model unconstrained. First, we consider a

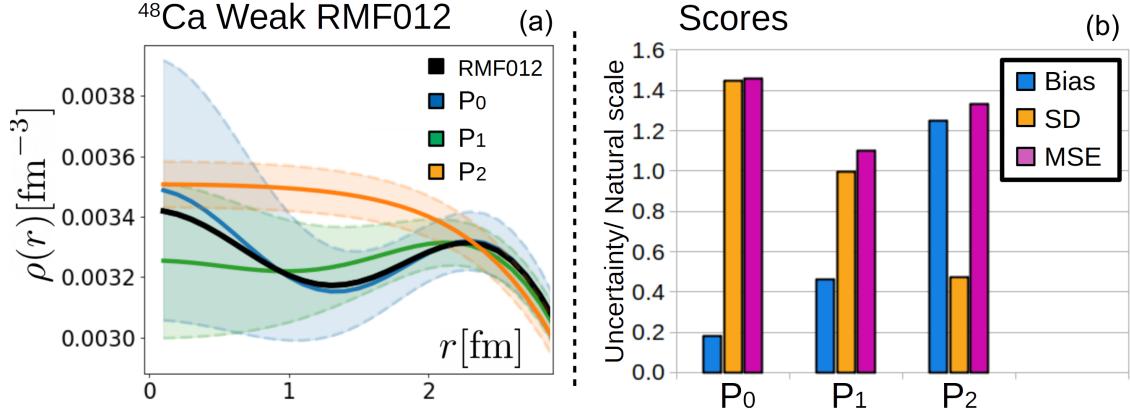


Figure 5.12: (a) Reconstructed weak density of ^{48}Ca for the generator RMF012, and for the model SF+G with three sets of priors. P_0 (in blue), P_1 (in green), and P_2 (in orange). (b) Scores on the interior density in terms of bias, SD, and MSE for the three choices of prior. The numerical values have been divided by the natural scale of $\Delta\rho_{\text{Ca}} = 0.00015 \text{ fm}^{-3}$.

nonexistent prior which we call P_0 . This ‘‘prior’’ effectively reproduces the original unconstrained SF+G model. Second, we consider a fairly uninformative prior such that the deviation from a flat density at the peak of each Gaussian is order $\sim 0.0003 \text{ fm}^{-3}$. These deviations represent oscillations of roughly 10%. We call this prior P_1 . It is given by the following parameter centers and standard deviations:

$$\boldsymbol{\omega}^0 = [0, 0, 0], \quad (5.28)$$

$$\boldsymbol{\sigma} = [0.005, 0.027, 0.08].$$

Finally, we consider an extremely restrictive prior which we call P_2 . This prior forces the value of the amplitudes $A_i = 0$, effectively reproducing the original SF model.

Fig. 5.12 (a) shows the reconstructed weak charge density for these choices of prior: P_0 (in blue), P_1 (in green) and P_2 (in orange). Fig. 5.12 b shows the scores in the interior density for the three choices of prior. This plot seems to resemble perfectly the behavior of bias, variance, and MSE shown in Fig. 4.2.

Indeed, these results are an interesting example of the bias vs variance trade off: the model without prior (P_0) reproduces the true curve almost perfectly, but presents a huge error band, while the model with the strongest prior (P_2), or effectively fewer parameters, has the biggest bias

but the smallest error bands⁴. The model with our proposed prior (P_1) sets in the middle ground, resulting in the best overall MSE. Its MSE score is even better than any of the average MSE scores of the models studied in Sec. 5.2.4 (for the fixed \mathbf{q}_0 locations).

We can analyze the behaviour of the MSE directly from the transfer function formalism as the prior strength changes. Let us focus on $\rho(0)$. Stronger priors will constrain more $\tilde{\mathcal{H}}^{-1}$, reducing the magnitude of the transfer functions $\mathcal{T}\mathcal{F}_j^{\rho(0)}$ of each data point. This effectively reduces the propagation of experimental uncertainty σ_j towards the calculated variance in $\rho(0)$.

The trade off 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. This is expected to increase the total bias in $\rho(0)$. This change can be written to first approximation as:

$$\delta\rho(0) = \sum_{k=1}^3 \mathcal{T}\mathcal{F}_k^{\rho(0)} \tilde{\eta}_k, \quad (5.29)$$

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

Appendix C.5 presents tables with the numerical values of the transfer functions for $\rho(0)$, and analyzes their meaning in more depth. The important fact is that, as the prior strength increases from P_0 to P_2 , the numerical value of the $\mathcal{T}\mathcal{F}_j^{\rho(0)}$ for each observation q_j tends to decrease sometimes even 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 $\mathcal{T}\mathcal{F}_k^{\rho(0)}$ become stronger. This allows each prior center ω_k^0 to push away the value of $\rho(0)$ from what the data suggest, effectively increasing the bias.

This 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 to overlook new discoveries or make erroneous conclusions. A more in depth analysis is required to optimize the prior strength and structure for each particular problem to reduce the MSE the most for a set of different truths.

⁴Note that, since we measure in the form factor space, the reconstructed density in the interior does not have to be well constrained if there are not enough data. This is why the orange curve which we expect to fail to reproduce the wiggles, also fails at reproducing the average interior density.

Parts of this chapter are adapted from the author's publication in Physical Review C 94.3 (2016): 034316. Copyright held by APS.

CHAPTER 6

DISPERSIVE EFFECTS ON ^{12}C

In this chapter, we focus our attention on calculating the impact that dispersive corrections – the virtual excitation of a nucleus in a scattering process– could have on the extracted electric charge radius of ^{12}C . The nuclear charge radius is related to the experimentally measured cross sections through Eqs. (2.6), (2.9), and (2.13). To propagate the dispersive corrections, treated as a perturbation to the cross section, we employ the transfer function formalism described in Chapter 3. We start this section with a brief description of the dispersive effects.

During the 80s and 90s, higher order corrections to the first Born approximation were extensively studied through dedicated elastic and quasi-elastic scattering experiments using unpolarized electron and positron beams (see [174, 175, 176, 101, 99, 177] and references therein), following the seminal paper from [178]. These effects scale as $S_{HOB} = V_C/E_e$ where S_{HOB} is the Scaling factor to account for Higher Order corrections to the Born approximation, V_C is the Coulomb potential of the target nucleus and E_e is the incident energy of the lepton probe [177]. Incidentally, they are expected to be small in the medium to intermediate energy regime, and have been neglected in the analysis of GeV energy data: V_C reaches a maximum of about 26 MeV for ^{208}Pb with a corresponding value of $S_{HOB} = 0.52\%$ for a 5 GeV beam.

In the 1st order approximation, the scattering cross section is evaluated using plane wave functions for the incoming and outgoing electrons. This approach is also known as the Plane Wave Born approximation (PWBA) or simply the Born Approximation (Fig. 6.1). This is the approach behind Eq. (2.6). Coulomb corrections originate from the Coulomb field of the target nucleus that causes an acceleration (deceleration) of the incoming (outgoing) electrons and a Coulomb distortion of the plane waves: these effects are treated within a Distorted Wave Born Approximation (DWBA) analysis for inelastic scattering or elastic/quasi-elastic scattering on heavy nuclei [177]. Two other corrections are required to properly evaluate the scattering cross section: radiative corrections due to energy loss processes and dispersive effects due to virtual excitations of the nucleus at the moment of the interaction (Fig. 6.1).

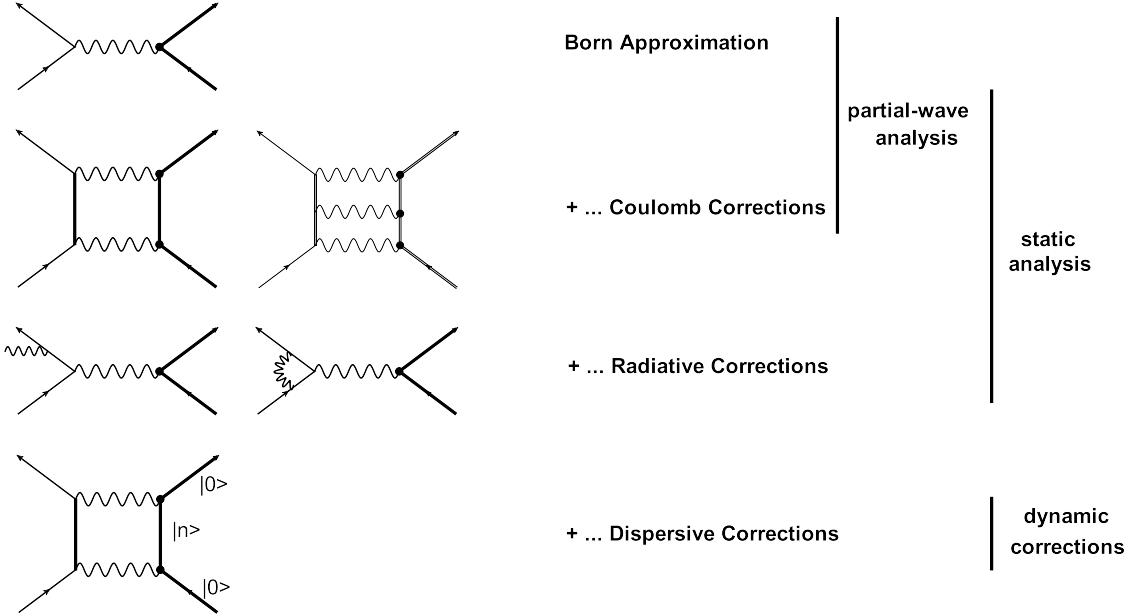


Figure 6.1: High-order corrections to the one-photon exchange Born approximation in electron/positron-nucleus scattering.

Within the last decade, a renewed interest in going beyond the PWBA arose from a discrepancy between unpolarized and polarized elastic scattering data on the measurement of the proton form factor ratio $\mu G_E^p/G_M^p$ which can be attributed to the contribution of two-photon exchanges [179, 180, 181, 182, 183, 184, 185, 186]. These effects have been investigated with a series of dedicated experiments [187, 188, 189, 190] (also see reviews [191, 192, 193] and references therein), including their impact on the measurement of form factors for nucleons and light ($A \leq 3$) nuclei. They include both Coulomb corrections [177, 194], excited intermediate states and treatment of the off-shell nucleons through dispersion relations as a function of the 4-momentum transfer.

Coulomb corrections have historically been labeled as *static* corrections to the Born approximation as depicted in Fig. 6.1. While these effects contribute to a few percents [177, 191, 192, 194], *dynamic* corrections known as dispersive effects are enhanced in the diffraction minima, where the first-order (Born approximation) cross section has a zero. For example, these corrections have been observed to contribute around 20% in the first diffraction minimum of ^{12}C at 690 MeV [101, 99, 195].

Experimentally, in order to extract the magnitude of the dispersive effects, the momentum transfer q is modified to account for the Coulomb effects into an effective momentum transfer q_{eff} (see [177, 194, 196] for the validity of this so-called Effective Momentum Approximation). The

latter is obtained by modifying the incident (E_e) and scattered ($E_{e'}$) energies of the incoming and outgoing electrons [177]:

$$q = 4E_e E_{e'} \sin^2(\theta/2) \rightarrow q_{eff} = 4E_{e,eff} E_{e',eff} \sin^2(\theta/2) \quad (6.1)$$

with $E_{e,eff} = E_e \left(1 - \frac{|V_C|}{E_e}\right)$ and $E_{e',eff} = E_{e'} \left(1 - \frac{|V_C|}{E_{e'}}\right)$. $|V_C|$ is the average (magnitude of the) Coulomb potential of the target nucleus, which turns out to be the potential at its surface: $|V_C| = KZ/R$, with R defined in Eq. (2.13), Z the total electric charge of the nucleus and $K = 1/4\pi\epsilon_0$. The corresponding experimentally measured cross section can then be compared to the theoretical cross section calculated using a static charge density [101].

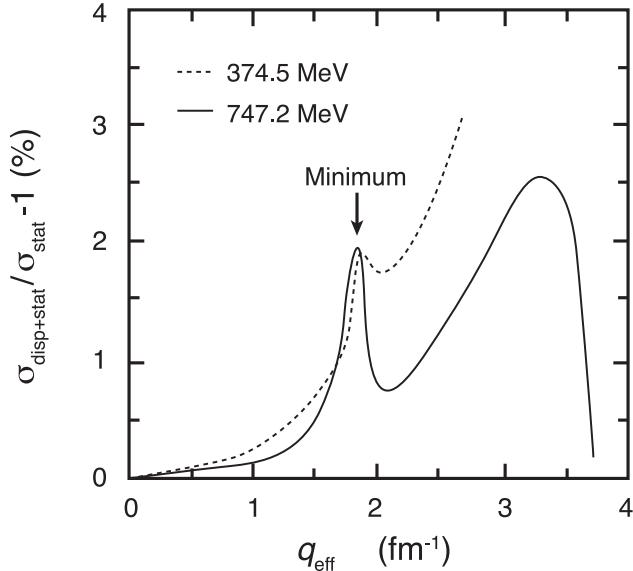


Figure 6.2: Calculations of Friar and Rosen [100] for dispersion corrections to elastic electron scattering from ^{12}C at 374.5 and 747.2 MeV in the first diffraction minimum $q_{eff} = 1.84 \text{ fm}^{-1}$. Figure adapted from [197].

Theoretical calculations for dispersive effects in elastic electron scattering for p-shell, spin-0 targets such as ^{12}C were performed in the mid-70s by Friar and Rosen [100]. Fig. 6.2 shows these calculations for two beam energies as a function of q_{eff} (see Eq. (6.1)). The inclusion of dispersive corrections $\sigma_{stat+disp}$ is compared to the cross section σ_{stat} obtained from a static charge distribution (see Eq. (2.6)). At the diffraction minimum they predict a constant (no energy dependence) 2% size. These results do not agree with what has been observed in experiments, as is shown in the world data in Fig. 6.3.

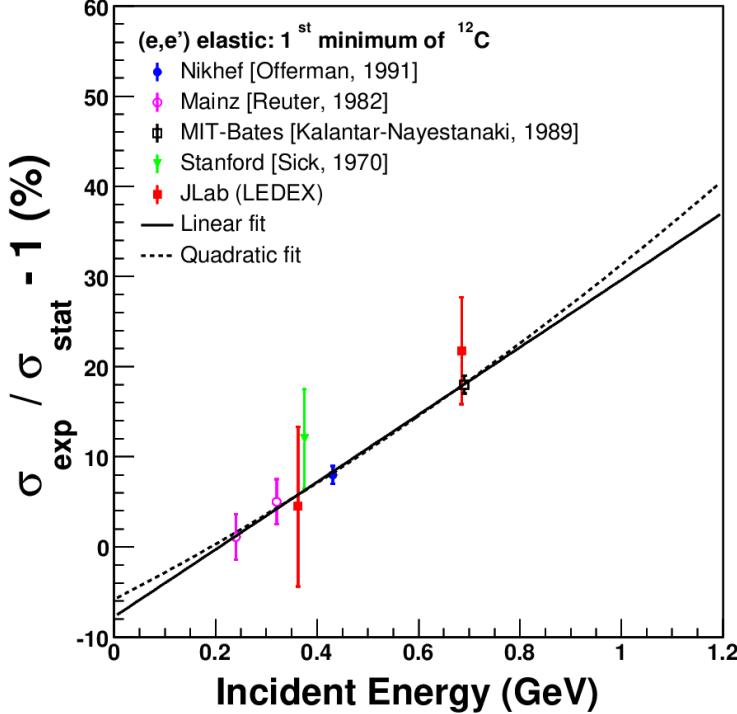


Figure 6.3: World data on the energy dependence of dispersive effects in the first diffraction minimum of ^{12}C .

Fig. 6.3 shows that dispersive effects have a clear energy dependence, and a magnitude of more than 20% at energies of 685 MeV. From a first order (solid line) and a second order (dashed line) polynomial fits (see Fig. 6.3), extrapolations indicate deviations at 1 GeV of around 30%. One pseudo-data point from the average of the fit functions is also shown at 1 GeV with a 3% error bar (which is a reasonable systematic error for an elastic peak cross section measurement at JLab for this energy).

6.1 The impact of dispersive corrections on nuclear radii

We now proceed to estimate the effects of dispersive corrections on nuclear densities reconstruction, with a specific aim in the extracted radius.

For simplicity, we write the dispersive cross section $\sigma_{disp} = \sigma_{stat+disp}$ as a function of the static cross section σ_{stat} :

$$\sigma_{disp} = \sigma_{stat}[1 + \delta_{disp}(E_e)] \quad (6.2)$$

with $\delta_{disp}(E_e)$ the higher order correction to the Born Approximation. Our convention throughout this chapter is to label any quantity with the subscript *disp*, such as the cross section σ_{disp} , that has been directly obtained from experimental measurements and is affected by the contribution from dispersive effects. Analogously, the subscript *stat*, such as σ_{stat} , is attached to any quantity that could be obtained by removing the contribution from dispersive effects, thus correcting the experimental observation. In that sense, σ_{stat} will be the expected cross section from the Born Approximation. Eq. (6.2) states that the observed experimental cross sections σ_{disp} could be modeled by a small multiplicative perturbation added to the static σ_{stat} cross section.

Note that modeling the dispersive effects as in Eq. (6.2) might at first appear to be in contradiction with Eq. (2.6). Since Eq. (2.6) implies that the cross section reaches zero if $F(Q^2) = 0$, Eq. (6.2) would force the dispersive effects to be zero at that minimum. Nevertheless, experimentally this is not the case since Eq. (2.6) is just the first order approximation to the scattering process. Other corrections (such as the static analysis displayed in Fig. 6.1 and implemented in phase shift analyses) will fill up vanishing cross sections [198]. Fig. 6.3 shows that, from those corrections, dispersive effects have take values up to 20% at the exact minimum, and these contributions are energy dependent.

In the Plane Wave Born Approximation, the nuclear charge density distribution $\rho_{ch}(r)$ is the Fourier transform of the nuclear form factor and for spherically symmetric charge distributions the relation is given by Eq. (2.10). As we discussed in Sec. 2.1.2, the density $\rho_{ch}(r)$ can thus be extracted from the experimentally measured $F_{ch}(q)$. For that goal, we use the Fourier Bessel (FB) model [112], one of the seven models we tested Sec. 5.2.

Using the zero'th spherical Bessel function $j_0(r) = \sin(qr)/qr$ we expand the charge density as:

$$\rho_{ch}^{FB}(r) = \begin{cases} \sum_\nu a_\nu j_0\left(\frac{\nu\pi r}{R_{cut}}\right) & \text{for } r \leq R_{cut} \\ 0 & \text{for } r > R_{cut} \end{cases} \quad (6.3)$$

with R_{cut} the cut-off radius chosen such as the charge distribution is zero beyond that value ($R_{cut} = 8$ fm for ^{12}C [101]) and the coefficients a_ν related to the form factor as $a_\nu = q_\nu^2 F_{ch}(q_\nu)/2\pi R_{cut}$, where $q_\nu = \nu\pi/R_{cut}$ is obtained from the ν -th zero of the Bessel function j_0 . We identify the coefficients \mathbf{a} as the free parameters ω described in Sec. 2.2 and adopt the notation a_ν from [112] through this chapter.

The form factor is given by:

$$F_{ch}^{FB}(q) = \sum_{\nu} a_{\nu} G_{\nu}(q), \text{ where} \quad (6.4)$$

$$G_{\nu}(q) \equiv 4\pi(-1)^{\nu} j_0(qR_{\text{cut}}) \frac{R_{\text{cut}}}{q^2 - q_{\nu}^2}. \quad (6.5)$$

As mentioned in Sec. 2.1.2, we ignore the contribution of the neutrons to the electric charge distribution of the nucleus. The charge radius $\langle r_{\text{ch}}^2 \rangle^{1/2}$ can then be obtained from the nuclear charge density distribution from Eq. (2.13).

Using the Bessel expansion of ρ_{ch} from Eq. (6.3) leads to:

$$\langle r_{\text{ch}}^2 \rangle = 4\pi \int_0^{R_{\text{cut}}} \sum_{\nu} a_{\nu} j_0\left(\frac{\nu\pi r}{R_{\text{cut}}}\right) r^4 dr \quad (6.6)$$

Evaluating the integral of each Bessel function gives:

$$\int_0^{R_{\text{cut}}} j_0\left(\frac{\nu\pi r}{R_{\text{cut}}}\right) r^4 dr = \frac{(-1)^{\nu} R_{\text{cut}}^5 (6 - \nu^2 \pi^2)}{\nu^4 \pi^4}. \quad (6.7)$$

We obtain the radius as:

$$\langle r_{\text{ch}}^2 \rangle = 4\pi \sum_{\nu} a_{\nu} \frac{(-1)^{\nu} R_{\text{cut}}^5 (6 - \nu^2 \pi^2)}{\nu^4 \pi^4} \quad (6.8)$$

Therefore, all the coefficients a_{ν} of the Fourier Bessel expansion play a role in estimating the radius of the charge density distribution, decreasing in importance as $1/\nu^2$. If the measured cross sections used to extract the value of the form factor $F_{\text{ch}}(q)$ are indeed modified by the dispersive corrections, then the change would propagate through the fitted coefficients a_{ν} to the estimate of the charge radius $R_{\text{ch}} \equiv \langle r_{\text{ch}}^2 \rangle^{1/2}$. We are interested in estimating this propagation.

The charge radius is a function of the M parameters of our model, in this case the M independent Bessel coefficients a_{ν} , which in turn depend on the N experimentally extracted form factor values y_j . We assume that these parameters are fitted to the data by minimizing χ^2 defined in Eq. (2.16).

Therefore, through the coefficients a_{ν} the charge radius is a function of the experimental points and one can write a small change in R_{ch} due to a given small change in the observations $(\delta y_1, \delta y_2, \dots, \delta y_N)$ as:

$$\delta R_{\text{ch}} = \sum_{j=1}^N \frac{\partial R_{\text{ch}}}{\partial y_j} \delta y_j = \sum_{j=1}^N \left(\sum_{\nu=1}^M \frac{\partial R_{\text{ch}}}{\partial a_{\nu}} \frac{\partial a_{\nu}}{\partial y_j} \right) \delta y_j. \quad (6.9)$$

The coefficients mediating these changes are the Transfer Functions (TF) defined in Chapter 3 (see Eq. (3.5)). In terms of the TF, Eq. (6.9) takes the form:

$$\delta R_{ch} = \sum_{j=1}^N \mathcal{T}\mathcal{F}_j^{R_{ch}} \delta y_j = \sum_{j=1}^N [\nabla R_{ch} \mathcal{H}^{-1} \nabla F(q_j, \mathbf{a}) \sigma_j^{-2}] \delta y_j, \quad (6.10)$$

where $\nabla F(q_j, \mathbf{a})$ is the gradient of the form factor model in Eq (6.4), \mathcal{H} is the Hessian matrix of χ^2 defined in Eq (2.18), and σ_j are the experimental errors. For M independent coefficients a_ν , one has $M + 1$ Bessel functions due to the normalization constraint in Eq. (2.12). The a_{M+1} can be explicitly written by solving the constraint:

$$\begin{aligned} 4\pi \int \rho(r) r^2 dr &= 1, \\ \sum_{\nu=1}^{M+1} (-1)^{\nu+1} \frac{4\pi R_{cut}}{q_\nu^2} a_\nu &= 1, \\ a_{M+1} &= (-1)^M \left(1 - \sum_{\nu=1}^M (-1)^{\nu+1} \frac{4\pi R_{cut}}{q_\nu^2} a_\nu \right) \frac{(M+1)^2 \pi}{4R_{cut}^3}. \end{aligned} \quad (6.11)$$

Following Eq. (6.8), and taking into account the normalization condition, the partial derivative of R_{ch} with respect to a coefficient a_ν is given by:

$$\frac{\partial R_{ch}}{\partial a_\nu} = \frac{1}{2R_{ch}} 4\pi \frac{(-1)^\nu R_{cut}^5 (6 - \nu^2 \pi^2)}{\nu^4 \pi^4} + \frac{\partial R_{ch}}{\partial a_{M+1}} \frac{\partial a_{M+1}}{\partial a_\nu}, \quad (6.12)$$

from which the gradient ∇R_{ch} can be constructed. The last term has to be included since R_{ch} depends on the $M + 1$ coefficients and a_{M+1} depends linearly on the rest of the a_ν , making the calculation straightforward from Eq. (6.11).

Meanwhile, to build $\nabla F(q_j, \mathbf{a})$ we compute:

$$\frac{\partial F(q_j, \mathbf{a})}{\partial a_\nu} = G_\nu(q_j) + G_{M+1}(q_j) \frac{\partial a_{M+1}}{\partial a_\nu} \quad (6.13)$$

where $G_\nu(q)$ is given by Eq. (6.5).

We treat the set of changes in the observations, δy_j , as modifications in the data to take into account the dispersive corrections and calculate the change in R_{ch} from Eq. (6.9). In the following discussion, we apply this framework to the data set presented by Offermann et al. [101] under the convention that $\delta R_{ch} = R_{ch}^{stat} - R_{ch}^{disp}$ and $\delta y_j = y_j^{stat} - y_j^{disp}$, since we want to estimate the change in the radius once the corrections for the dispersive effects have been implemented.

In their work in [101], the authors used 18 Bessel functions to fit cross section experimental data from ^{12}C . To show our method, we use the values of their first 9 coefficients a_ν , $\nu \in \{1, 9\}$ from their Table X second column (without dispersive corrections) to generate 9 values y_ν of the form factor according to the relation $a_\nu = F(q_\nu)q_\nu^2/2\pi R_{\text{cut}}$ at those 9 special q_ν values with $R_{\text{cut}} = 8$ fm. For the error associated with each "observation" y_ν , we use the adapted error Δy_ν from their reported percentage error in Δa_ν . For the remaining 9 points $\nu \in \{10, 18\}$, we center the observations y_ν at zero and add an error band associated with the form factor of the proton as the authors did following the recommendation in [112]. The q_ν values are displayed as red and black points in Fig. 6.4. Since the normalization condition must be respected, only 17 from the 18 coefficients a_ν are independent. We identify therefore, the number of data points $N = 18$, and the number of parameters $M = 17$.

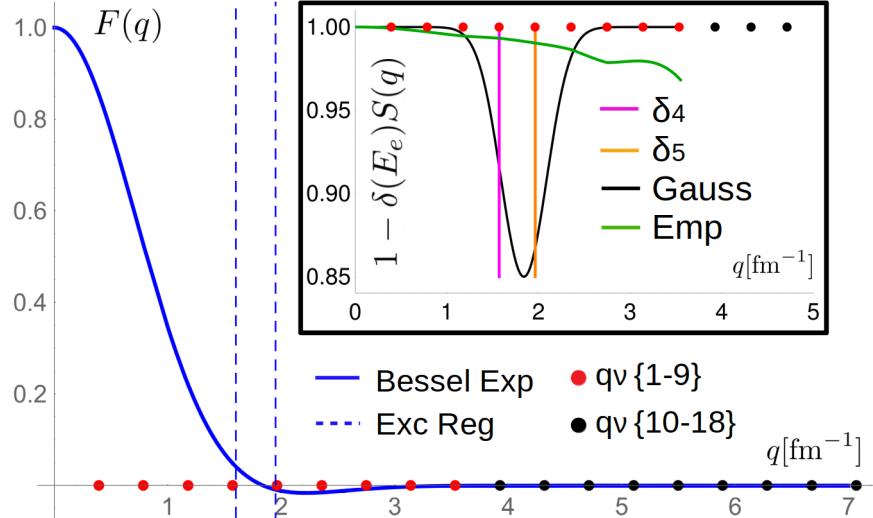


Figure 6.4: ^{12}C form factor expanded in the Bessel functions formalism using Offermann [101] coefficients without dispersive corrections. The circles in the q axis show the special values of momentum transfer for the first 9 (red) from experimental data and the second 9 (black) from the extrapolation suggested in [112]. The dashed blue lines enclose the region of the data excluded from the analysis in [101]. The inset plot shows the three test forms for $S(q)$ in addition to the empirical perturbation obtained directly from the data by third degree spline interpolation. The curves in the inset plot are the ones needed to obtain the corrected $F_{\text{ch}}^{\text{stat}}$ from the observed $F_{\text{ch}}^{\text{disp}}$ values.

Fig. 6.5 shows the matrix $\partial a_\nu / \partial y_j$ (the TF defined in Chapter 3) for the 18 observations y_j and $17 + 1$ coefficients a_ν . Even though we are not treating a_{18} as an independent variable since we

solved the constraint explicitly, we can still calculate how much its value changes when any of the observations y_j changes. It can be seen that as ν increases, a_ν becomes more dependent on y_ν and less sensitive to other values y_j ($j \neq \nu$). In principle, if the 18 coefficients were independent, each a_ν will only be sensitive to their corresponding y_ν [112], but the normalization constraint introduces mixing.

In the third column of Table 6.1, there are the numerical values of $\partial R_{\text{ch}}/\partial y_j$ for the first 9 observations y_j . Each one of these numbers, when multiplied by a small change in their associated observation, will yield the corresponding small change in R_{ch} as in Eq. (6.9). The fourth column shows the percentage change needed in observation y_j to create a 1% change in the radius. Even though the values $\partial R_{\text{ch}}/\partial y_j$ are roughly the same size for all the observations, this fourth column shows that R_{ch} is more sensitive to *percentage* changes in the first observations.

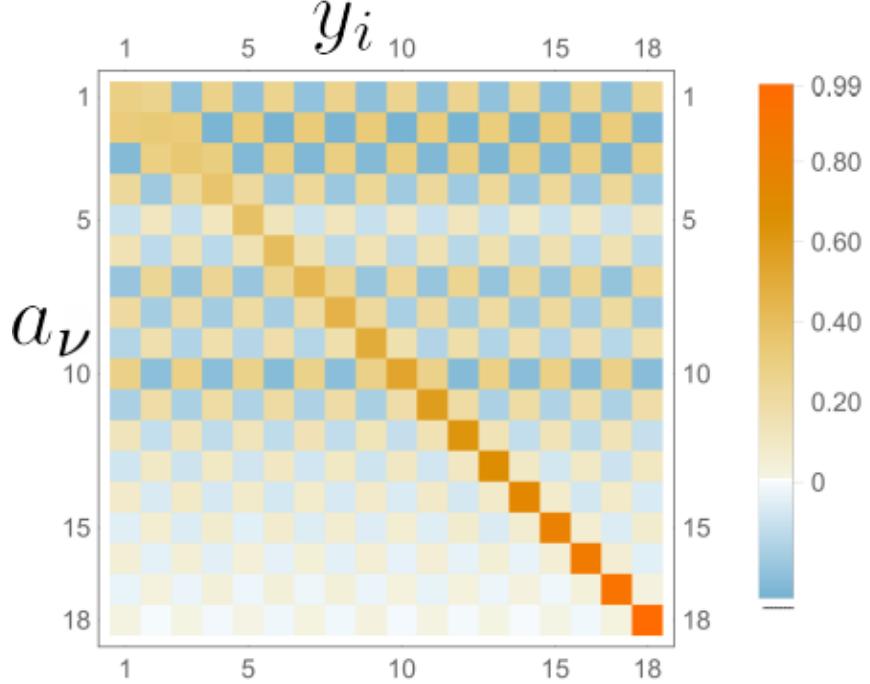


Figure 6.5: $\partial a_\nu / \partial y_j$ matrix for the data extracted from Offermann et al. [101].

Estimating the exact values of δy_j is a complicated task beyond the scope of this dissertation. This is because the change in the cross section as shown in Eq. (6.2) depends on the energy, but the momentum transfer q is a function of both the energy and the angle θ . Therefore, for the same fixed value of q we could have different pairs of (E, θ) which will be impacted differently.

Location j	y_j	$\partial R_{\text{ch}}/\partial y_j$ [fm]	δR_{ch} = 1%
1	0.854	-9.214	0.3
2	0.526	-2.595	1.8
3	0.221	+4.782	2.3
4	0.049	-5.547	9.1
5	-0.0098	+5.901	43
6	-0.0151	-6.094	27
7	-0.00754	+6.210	53
8	-0.00235	-6.285	168
9	-0.00039	+6.337	994

Table 6.1: The first column shows the index number of the special momentum transfer $i\pi/R_{\text{cut}}$. The second column shows its form factor value obtained from [101]. The third column shows the value of $\partial R_{\text{ch}}/\partial y_j$. The fourth column shows the percentage change needed in y_j to generate an equivalent change of 1% in the estimated charge radius.

In order to simplify our discussion, we assume now that we can separate the total effect of the dispersive effects on the form factor values as:

$$F_{\text{disp}}(q) = F(q)_{\text{stat}} \left[1 + \frac{1}{2} \delta(E_e) S(q) \right], \quad (6.14)$$

with $\delta_{\text{disp}} = \delta(E_e) S(q)$ from Eq. (6.2) where $\delta(E_e)$ controls the overall strength of the perturbation and $S(q)$ controls the impact this change would have on different q values. The factor of 1/2 comes from assuming that $\delta(E_e)$ is small and propagating the change from Eqs. (2.6) and (6.2).

Since the variable q depends on both E_e and θ , a separation such as Eq. (6.14) might not be completely accurate. As it can be seen in the calculations of Friar and Rosen (Fig. 6.2), a change in E_e clearly affects the overall shape of the dispersion corrections as a function of q . Nevertheless, Eq. (6.14) is simple enough to allow us to provide an estimate for the impact of such a change in inferred nuclear properties of the nucleus. In particular, we can now write the change in the charge radius as:

$$R_{\text{ch}}^{\text{disp}} = R_{\text{ch}}^{\text{stat}} \left[1 + \beta \delta(E_e) \right]. \quad (6.15)$$

where β is a proportionality coefficient fixed once $S(q)$ is specified. For a given fixed strength $\delta(E_e)$, the change in the radius will depend on the shape of $S(q)$, which is encoded in β .

Table 6.2 shows the results for three different test perturbations $S(q)$, in addition to an empirical one obtained from comparing columns 2 and 3 of Table X in [101], for the central values of the form

$S(q)$	$R_{\text{ch}}^{\text{stat}}$ [fm]	β	δR_{ch}
δ_4	2.512	-0.055	1.65 %
δ_5	2.480	-0.012	0.35 %
Gaussian	2.495	-0.032	0.98 %
Empirical	2.477	-	0.25 %

Table 6.2: The first column shows the perturbation form $S(q)$ in addition to the empirical perturbation. In the first three cases a strength of $\delta(E_e) = 0.3$ was assumed (30% dispersive effects strength as the extrapolation to 1 GeV in Fig. 6.3). The second column shows the calculated new radius, $R_{\text{ch}}^{\text{stat}}$ (the original radius is 2.4711 fm). The third and fourth columns show the β coefficient and the percentage change in R_{ch} , namely $\delta R_{\text{ch}} \equiv R_{\text{ch}}^{\text{stat}} - R_{\text{ch}}^{\text{disp}}$, respectively.

factor. For the test perturbations, the central values of the form factor were modified assuming a constant high value of $\delta(E_e) = 30\%$, so that our analysis could serve as an upper bound.

The three test forms for $\delta(E_e)S(q)$ consist of δ_4 , δ_5 and Gaussian. The first two represent an up-shift of 15% on the value of $F(q_\nu)$ for $\nu = 4$ and $\nu = 5$ alone respectively, while the Gaussian represents a Gaussian up-shift of amplitude 15% at its peak, centered at the diffraction minimum $q = 1.84 \text{ fm}^{-1}$ and with a standard deviation of 0.25 fm^{-1} . The functional forms of the three $S(q)$ are shown in the inset of Fig. 6.4 as well as the empirical perturbation, while the outset plot shows the Bessel expanded form factor and the special values of the momentum transfer q_ν . An overall up-shift in the form factor was chosen based on the calculations shown on Fig. 6.2, which predict an up-shift in the observed cross sections due to the dispersive effects: $\sigma_{\text{disp}} \geq \sigma_{\text{stat}}$.

In all three test cases for $S(q)$, the change on the radius did not exceed 2%, which is still a substantial increase compared to Offermann result [101] of a 0.28% increase. This contrast suggests that our overall strength $\delta(E_e) = 30\%$ was too large for the data locations in Offermann work [101]. This might indicate that the effects *on the available data* of the dispersive corrections are roughly at least a factor of five smaller outside the immediate vicinity of the diffraction minimum.

The empirical perturbation was obtained as $\delta_{\text{emp}}(q_\nu) = [F_{\text{disp}}^*(q_\nu) - F_{\text{stat}}^*(q_\nu)]/F_{\text{stat}}^*(q_\nu)$, where $F_{\text{disp}}^*(q_\nu)$ ($F_{\text{stat}}^*(q_\nu)$) represents the form factor values obtained from the second (third) column in Table X of [101]. Since no amplitude $\delta(E_e)$ was involved in the empirical perturbation, the value of β cannot be defined and we have that:

$$F_{\text{disp}}(q_\nu) = F(q_\nu)_{\text{stat}}[1 + \delta_{\text{emp}}(q_\nu)]. \quad (6.16)$$

When using the empirical perturbation for the δy_j , we obtain an effect of 0.25% in the radius, very close to the actual 0.26% (reported as 0.28% when using rounded values for the radii) in [101]. This empirical perturbation was only calculated at the special values q_ν and interpolated using a third degree spline and therefore, is not discarded that its strength can reach a peak of $\sim 30\%$ in the excluded region around the diffraction minimum $1.6 < q < 1.95 \text{ fm}^{-1}$. Indeed, the authors excluded these data to perform their analysis and avoid as much as possible the dispersive effects. This region is located between the two vertical blue dashed lines in Fig 6.4.

The situation could be more complex for parity-violating experiments [82, 81, 199] from which the measured asymmetry is used to extract a neutron skin. These experiments typically occur near diffractive minima to maximize their sensitivity to the physics, where also dispersive corrections contribute the most. It is clear one should take dispersive effects into account. However, to our knowledge there are no known measurements of dispersive effects using polarized beams and/or target. Therefore, measurements of the energy dependence for dispersive effects using polarized elastic scattering on various nuclear targets ($A > 1$) should be performed to provide an accurate information about the size of these effects in and outside minima of diffraction.

This chapter is adapted from the author's publication: Gueye, P. (2020). Dispersive corrections in elastic electron-nucleus scattering: an investigation in the intermediate energy regime and their impact on the nuclear matter. European Physical Journal. A, 56(JLAB-PHY-18-2707; DOE/OR/23177-4450; arXiv: 1805.12441). With kind permission of The European Physical Journal (EPJ)

CHAPTER 7

CONCLUSIONS

Bigger error bars: a decade ago they could have been enough to change the course of events that the muonic proton radius extraction experiment [10] produced from a “puzzle” to a “refinement”. At that time, the 2010 CODATA recommended proton radius was $r_p = 0.8775(51)$ fm [9], a value obtained from combining both electron scattering and electronic spectroscopy results [11, 4, 13]. This radius is completely incompatible (7σ away) with the single extremely precise muonic spectroscopy value of $r_p = 0.84184(67)$ fm [10]. If at least some of the most precise experimental results (see Fig. 1.1) that were combined to create the 2010 CODATA value had appreciable bigger error bars, the muonic results would have appeared as a welcoming improvement, rather than a shocking surprise. Claiming that so many experiments underestimated their uncertainties by such a huge amount sounds implausible, but with the new results in electronic spectroscopy [15], and the reported smaller radius from the PRad collaboration [21], not many other possible explanations seem to remain on the table. If the MUSE experiment proves that lepton universality holds within this context, and their results prove to be consistent with the muonic spectroscopy radius, then a serious revision of previous experiments should be performed to understand what went wrong.

There are extremely important physics-driven considerations such as the impact of two photon exchange [200], that could have affected many of the past analyses. From our perspective, there is a comparably important statistical-driven consideration as well: *how is the confidence interval (error bar) constructed*. This issue encompasses several key points, including the model used to fit the data, the propagation of statistical uncertainty from the experimental observations, and, if present, the tension generated by competing models. These key points are not only crucial in the proton puzzle context, but also apply to the other physics challenges we have considered in this manuscript.

With respect to the model selection, we have observed that in many cases the decision to use a given model is based on arguments that do not necessarily correlate to the ability of the model to actually obtain the desired information in a robust way. For example, a nine and a ten degree

polynomial expansions used by some authors might fit and interpolate the proton’s form factor data very well, but this does not make them good radius extractors [201]. As another example, consider a parametrization based on physical considerations and the analytical structure of the proton form factor as a function with a complex argument, such as the z expansion done in [140]. If the number of parameters in the expansion is chosen based on the χ^2 overall fit, the resulting function will most likely overfit the data and provide an incorrect radius. As a third example, consider the nonparametric constrained Gaussian process we proposed in Chapter 4. We chose this approach to create a very flexible model which constraints are derived only from known physical considerations. Once again, these properties do not guarantee a reliable performance when extracting the radius and, as we have shown, depending on the implemented constraints our analysis might favor a bigger or a smaller proton. As it has become customary in the last decade, theoretical predictions are expected to be accompanied by uncertainty estimations [202]. We strongly believe that studies as the one performed by Higinbotham and collaborators [61], in which the proposed models are tested with pseudo data on performing the very specific task they are meant to do, should also become the new norm.

With respect to the propagation of *statistical* uncertainty from experimental observation, in this manuscript referred to as the experimental errors σ_j , we reiterate one of our main points throughout the manuscript: *calculating just the variance does not seem to be enough*. As we discussed in Chapters 4 and 5, the possible bias of our estimation must also be taken into account. It is this combination of bias and variance, i.e., the mean squared error, that gives a better reflection on how a particular model is performing. When analyzing real data such as the single weak form factor measurement of ^{208}Pb made by PREX [82], we usually lack the “true” value to compare to our prediction and calculate the bias, a situation that can be tackled by pseudo data analysis [61]. If we ignore this important source of error, and report that the uncertainty in our quantity of interest m just comes from propagating experimental errors σ_j , we run the risk of giving a false sense of certainty.

In some situations, it is easier to realize that the bias of a model might play an important role in our estimations. One of the examples discussed in Chapter 4, was the fit of a straight line to a clearly curved data set generated by a dipole function. In other cases, it could be not as obvious, as we showed in Sec. 5.2.1. Fig. 5.7 in that same section, shows that the SF+G model –that is

perfectly capable of reproducing the internal structure of the weak charge density of ^{48}Ca – fails to do so if fitted in a particular set of observations. In the limit where the experimental errors σ_j on those observations go to zero, by only using the associated variance of our prediction to report the uncertainty we would reproduce the wrong weak density, and even worse, we will be convinced that our results are extremely precise.

With respect to the tension generated by competing models, we believe care should be taken when estimating a combined uncertainty. If we have reason to believe that the spread in predictions were caused by *statistical* fluctuations, then combining each model’s single prediction in an appropriate way could yield a correctly reduced overall uncertainty. Such is the goal of Bayesian Model Averaging used in the context of nuclear mass models [66]. If, on the other hand, each model prediction seems to be very precise (low variance), but the central values are scattered on a scale bigger than the respective variances, the combined uncertainty should take into account that, a priori, those predictions are incompatible. This is the case, for example, in the proton puzzle context when comparing the clearly different results obtained by groups that focused on low Q^2 data with simpler models vs the results obtained by groups that focused on the entire Q^2 data with more complex models. If we cannot rule out one of the clusters in favor of another, i.e., to our current knowledge both approaches seem valid, we should report an uncertainty that encompasses both results. This is clearly not a desirable situation when analyzing data in hopes of discarding or proving a theory, but it is a better alternative than to be confident about a wrong result.

In some cases, a puzzle might be outside our current technological and theoretical reach in such a way that, no matter how carefully we analyze the current data, the answer is just not there. Such was the situation more than four centuries ago with the discussion about the nature of light, with the competing “particle” and “wave” explanations. Resolving the dichotomy would have to wait until the developments of the late XIX and early XX centuries with the electromagnetic and quantum mechanical description of light. In other cases, such as the proton puzzle, for example, a partial solution could have been achieved by rigorous statistical analysis, as we have already argued. With the enormous amount of new discoveries coming our way, we must be ready to avoid any preventable puzzle and outsmart uncertainties when possible.

Such a case could already be developing in the context of the Equation Of State (EOS) of nuclear matter. The observation of the binary neutron star merger in 2017 [104], pointed towards a

less stiff EOS, i.e., smaller neutron stars. Meanwhile, although the error bars are significantly large, the single PREX measurement of the weak form factor of ^{208}Pb seems to point towards stiffer EOS, i.e., bigger neutron skins and neutron stars. If PREX-II confirms these results with a much more constrained error bar, there will be tension between the astrophysical and the laboratory results. This discrepancy could point towards a phase transition in the interior of neutron stars [103]. We believe that strong statistical arguments will be needed to ensure that the discrepancy is real, and not the product of an underestimated error bar.

We are entering an unprecedeted age for nuclear physics, fueled, among many developments, by the birth of multi messenger astronomy and the construction of FRIB with unparalleled reach for exotic nuclear matter. We hope this dissertation is a small step in the right direction to contribute to the multidisciplinary efforts involving statistics that are indispensable to capitalize on the enormous investment that is being made in our field.

7.1 Going Forward

While we await the exciting results from the MUSE collaboration, we are interested in improving the performance of the constrained Gaussian process model we developed in Chapter 4. The possible bias in the analysis of the high Q^2 region should be explored and corrected –an objective that could be accomplished by developing a better procedure for estimating the hyperparameters (ν , N , l). To do so, a similar procedure as the one developed in [61] could be applied, in which the many “models” under the selection process are variants of the nonparametric approach with different constraints and hyperparameter values. We trust that lessons learned from comparing parametric models will help us improve the robustness of our nonparametric model. We have already started this exploration project with our collaborators at the statistics department at Texas A&M.

In terms of the parametric approaches to the proton radius extraction, assuming that the PRad results [21] got the correct proton radius, we strongly recommend that a serious statistical analysis be performed to identify what actions, if any, could have been taken in previous analyses to report more realistic error bars. The proposed project, which should nourish from both physics and statistics contributions, would be crucial to increase our chances of avoiding similar pitfalls in future research areas. Yet, even if the resolution of the proton puzzle is found elsewhere (e.g., technical difficulties such as the normalization or two photon exchange effects, or even new physics beyond

the Standard Model), the advances along this direction would have not been in vain. The proton puzzle as well as many other developments have allowed us to realize the importance of *enhancing the interaction between nuclear experiment and theory through information and statistics* [203].

With respect to weak charge densities, one of the most exciting awaited new developments is the announcement of the PREX-II campaign results, to happen very soon. As we already discussed, the direction to move forward will strongly depend on the central value and error bar the collaboration reports. If their results are in accordance with the gravitational wave detection, this will serve as a compelling case about the expected correlation between the sizes of neutron skins and those of neutron stars [204]. On the other hand, if the new PREX-II results seem to disagree to a reasonable statistical level with the gravitational waves observation, then we will have to start the exciting path of exploring possible explanations, including phase transitions in the EOS.

Beyond the PREX-II reports, we expect many other interesting results coming from the CREX campaign, which will provide the first electroweak determination of the weak radius of ^{48}Ca [83, 84]. New experiments such as those which are to begin operations by 2023 [85] at the Mainz Energy recovery Superconducting Accelerator (MESA), may be able to determine the weak radius of both ^{48}Ca and ^{208}Pb with increased precision [73]. FRIB will soon start operations as well, having as one of the short-term goals “*First constraints on EOS and symmetry energy term at various densities and temperatures from nucleus-nucleus collisions and from measurements of neutron-skin thicknesses*” [205]. On the astrophysical side, LIGO is to enter its fourth observational run by 2022. During this run, it is expected to observe several neutron star mergers per year [206], which could provide more invaluable information on the EOS at higher densities and pressure.

To better capitalize on these investments and orchestrate the unprecedented abundance of new data, statistics and information analyses will be highly valuable. We would be interested in tackling issues such as the connection through theoretical models between different observations, coordinated experimental designs to maximize the information gained, and robust procedures to extract said information from raw data. Several efforts are already being made in those directions, as for example, the analysis of the information content of new observables [148]. Based on this type of analyses, we could answer questions such as: “*Should one attempt a single high-precision measurement or would it be better to attempt two (or more) measurements with less precision*” [148]. As we move forward, we could face peculiar, but interesting, scenarios. For example, most of our

knowledge about neutron distributions inside nuclei (which are dictated by the nuclear strong force) could come from gravitational wave observations of neutron stars (objects that are 10^{18} times bigger than nuclei) which merged millions of light years away.

During the development of the tools and frameworks required to achieve these goals, we believe that the transfer function formalism proposed in Chapter 3 can be a very useful addition in several battles against uncertainty.

First of all, it would be interesting to integrate the TF formalism directly into model building. We believe questions like *what makes a model better than others at extracting information?*, could be tackled from this optics. As discussed in this dissertation, answering *which* model is better at extracting data has become a crucial question in nuclear physics, specifically within the context of the proton puzzle. The work by Higinbotham and collaborators [61] investigated this question and identified the three best models to extract the proton radius, but did not seem to elaborate much on *what* made those models successful. We believe the TF formalism could be used to advance in that direction. As shown in Chapter 5, having a bigger \mathcal{TF}_j at a data location makes the model more sensitive to the noise σ_j and to the η_j . But if all \mathcal{TF}_j are too small, the model loses sensitivity to the signal amidst the noise. Understanding the TF distribution of successful models could help not only to identify, but to create some sort of “optimal” model. Once this model is determined it will be possible to find the optimal locations and resources distribution, such as beam time, for each particular experiment. This could not only be applied to density reconstruction from scattering data, or the proton radius extraction, but to several other nuclear physics problems which involve model fitting from data.

Another straightforward direction to explore more in depth is the role of priors and hyperparameters. Hyperparameters, such as γ and the Gaussian locations for the SF+G, or the Bessel cut off radius R_{cut} and basis numbers, can change drastically a model’s performance. We could explore questions like: *given six observations of the weak form factor of ^{48}Ca , which is better in terms of MSE, to have 5 or 6 adjustable Bessels coefficients? How does this answer scale with the number of data points?* We believe it could be possible to create a framework using the TF formalism that can tackle this type of questions in a robust, direct, and straightforward manner. This would allow to conduct a more informed search in the hyperparameter space of each model instead of just trial and error.

We showed that the incorporation of priors in reconstructing densities can have a dramatic effect on a model’s performance (Sec. 5.2.5). A more in depth study should be carried out to identify optimal centers for the parameter’s priors ω_k^0 and their respective uncertainties σ_k for each model. To validate such a research project, more generator functions from other nuclear model families should be incorporated. We are confident that the TF formalism can also guide this optimization procedure.

Finally, a third application 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 [66]. Within our context of nuclear densities, using the MSE score should allow to test under which circumstances the mixing of several models outperforms the predictability power of a single model. This situation connects to what we discussed about competing models in the proton radius context. It would be interesting to explore the generalization of the TF formalism to Bayesian model mixing.

We are planning on exploring some of these directions with collaborators from the BAND team [69] at Michigan State University. We believe that the ideas we have developed and implemented in this dissertation can be translated into a software for the nuclear community. We have already begun working on it. This new resource could select, for example, the optimal model for analyzing the data using the bias vs variance criterion. Another possible task could be related to experimental design: *where to measure to maximize the information gained*. If the program is flexible enough, it could take into account experimental budgets and constraints, existing models, and the quantities of interest that the user wants to obtain. The code could also show the user the observations that are currently driving most of the uncertainty, so that more resources are assigned to these locations in future experiments. These objectives align very well with those of the BAND collaboration: to create “*Cyberinfrastructure Framework—designed to facilitate principled uncertainty quantification (UQ) with multiple nuclear models*” [69]. We anticipate tangible results within the next few years.

APPENDIX A

NONPARAMETRIC APPROACH: CONSTRAINED GAUSSIAN PROCESS

A.1 Theoretical proofs for the constraints on f_ξ and on the independence of \mathcal{C}_ξ on τ

This appendix presents the equivalence between the shape constraints on our model f_ξ and the inequalities on the coefficients ξ . Denote by \mathcal{C}_f , the function subspace of all the f_ξ defined in Eq (A.10) that obey the required constraints: $f_\xi(0) = 1$, $f'_\xi(x) < 0$ and $f''_\xi(x) > 0$.

We show below that the constraints that define \mathcal{C}_f can be *equivalently* represented as linear restrictions on ξ . We state Proposition 1 which provides an explicit characterization of the stated linear constraints.

Proposition 1 $f_\xi \in \mathcal{C}_f$ if and only if $\xi \in \mathcal{C}_\xi$.

Recall \mathcal{C}_ξ is defined as:

$$\mathcal{C}_\xi \equiv \left\{ \xi \in \mathbb{R}^{N+3} : \xi_1 = 1, \xi_2 + \sum_{j=0}^N c_j \xi_{j+3} \leq 0, \xi_{j+3} \geq 0, j = 0, \dots, N \right\}, \quad (\text{A.1})$$

Proof 1 We first check the convexity constraint. By taking second order derivative we have $f''_\xi(x) = \sum_{j=0}^N \xi_{j+3} h_j(x)$, by the non-negativity of h_j for all $x \in [0, 1]$ and any $j = 0, \dots, N$, the set $\{f''_\xi(x) \geq 0, \forall x \in [0, 1]\}$ is equivalent to $\{\xi_{j+3} \geq 0, j = 0, \dots, N\}$. To impose the non-increasing constraint, we need to check the following:

$$f'_\xi(x) = \xi_2 + \sum_{j=0}^N \xi_{j+3} \psi_j(x) \leq 0, \forall x \in [0, 1].$$

Observe that this is equivalent to:

$$\xi_2 \leq - \max_{x \in [0, 1]} \left(\sum_{j=0}^N \xi_{j+3} \psi_j(x) \right) = - \sum_{j=0}^N c_j \xi_{j+3}. \quad (\text{A.2})$$

(A.2) follows since ψ_j defined in (A.4) is a non-decreasing function of x and $\max_{x \in [0, 1]} \psi_j(x) = \psi_j(1) =: c_j$ for $j = 0, \dots, N$. This concludes the proof of the proposition. \diamond

Now we proceed to show why the normalizing constant M_ξ of the truncated prior distribution of ξ is independent of τ .

Proposition 2 *The normalizing constant M_ξ associated with the truncated prior distribution of ξ is a constant in $[0, 1]$ that does not depend on τ .*

Proof 2 *By definition*

$$M_\xi = \int_{C_\xi} (\tau^2)^{-(N+2)/2} (|\Gamma|)^{(-1/2)} e^{-\frac{1}{2\tau^2} \xi^T \Gamma^{-1} \xi} d\xi.$$

By change of variable $\xi' = \xi/\tau$, observe that the truncated region $C_{\xi'}$ is the same as C_ξ as long as $\tau > 0$. Hence, $M_\xi \in [0, 1]$ does not depend on τ . \diamond

A.2 Details regarding the construction of the regression model

This appendix provides a detailed construction of the regression model employed in Section 4.2.1. Note that we defined a working grid formed by a collection of $N+1$ equally spaced points $x_j = j/N$ in the closed interval $[0, 1]$ and a set of basis functions:

$$h_j(x) = \begin{cases} 1 - N|x - x_j|, & \text{if } |x - x_j| \leq 1/N; \\ 0, & \text{otherwise,} \end{cases} \quad (\text{A.3})$$

and their corresponding integrals:

$$\psi_j(x) = \int_0^x h_j(t) dt, \quad (\text{A.4})$$

$$\phi_j(x) = \int_0^x dt \int_0^t h_j(s) ds. \quad (\text{A.5})$$

The basis functions $h_j(x)$ can be used to approximate any continuous function $f(x)$ by linearly interpolating between the grid points. That is,

$$f(x) \approx \sum_{j=0}^N f(x_j) h_j(x). \quad (\text{A.6})$$

To illustrate the quality of the approximation, we used a grid of size $N = 10$ to display in Fig. A.1(b) the results for a dipole function of the form:

$$f(x) = \left(1 + \frac{\bar{r}_p^2 x}{12}\right)^{-2}, \quad (\text{A.7})$$

where $x = Q^2/Q_{\max}^2$, $Q_{\max} = 25.01 \text{ fm}^{-1}$, and $\bar{r}_p = r_p Q_{\max} = 4.21$. The apex of each triangle, namely, the scale factor multiplying each basis function $h_j(x)$, is the value of the dipole function at the j th grid point, or $f(x_j)$. The approximation is so accurate that the underlying exact dipole function (shown in red) is difficult to discern. As we show below, for the purpose of extracting the proton radius it is better not to approximate directly the electric form factor $G_E(Q^2)$ using the basis functions $h_j(x)$, but rather the smoother set of related functions $\phi_j(x)$ defined in Eq. (A.5). To do so, we invoke the fundamental theorem of calculus for any twice differentiable function $f(x)$ defined on the closed interval $[0, 1]$. That is,

$$f(x) = f(0) + xf'(0) + \int_0^x dt \int_0^t f''(s)ds. \quad (\text{A.8})$$

If we now approximate $f''(s)$ under the integral sign using the basis functions $h_j(x)$ we obtain:

$$f(x) \approx f(0) + xf'(0) + \sum_{j=0}^N f''(x_j) \int_0^x dt \int_0^t h_j(s)ds = f(0) + xf'(0) + \sum_{j=0}^N f''(x_j) \phi_j(x). \quad (\text{A.9})$$

This approximation to the exact dipole is shown in Fig. A.1(c) together with the underlying behavior of $\phi_j(x)$. In this case, the approximation to the exact dipole is not as accurate as in Fig. A.1(b). However, in a regression problem the function is unknown, including its first and second derivatives. Hence, we characterize our regression model in terms of $(N+3)$ free parameters ξ_j that will be obtained from a suitable fit to the experimental data. That is,

$$f(x) \equiv f_\xi(x) \approx \xi_1 + \xi_2 x + \sum_{j=0}^N \xi_{j+3} \phi_j(x). \quad (\text{A.10})$$

As displayed in Fig. A.1(d), once this scheme is adopted, there is excellent agreement with the real dipole function. Clearly, one great advantage of Eq. (A.10) is that values for the floating normalization and mean-squared radius are directly encoded in ξ_1 and ξ_2 . Moreover, this approximation has a nice physical interpretation which we highlighted already in Sec. 4.2.1. Regard $f(t)$ as the one-dimensional trajectory of a particle as a function of time t , as shown in Eq. (4.21). At time $t = 0$, the particle starts at a position $f(0)$ with an initial velocity $f'(0)$. As time evolves, corrections to the straight-line trajectory are implemented by the different ϕ_j in proportion to $f''(t_j)$. These corrections can be thought as “acceleration spikes” that stir the particle into the correct trajectory.

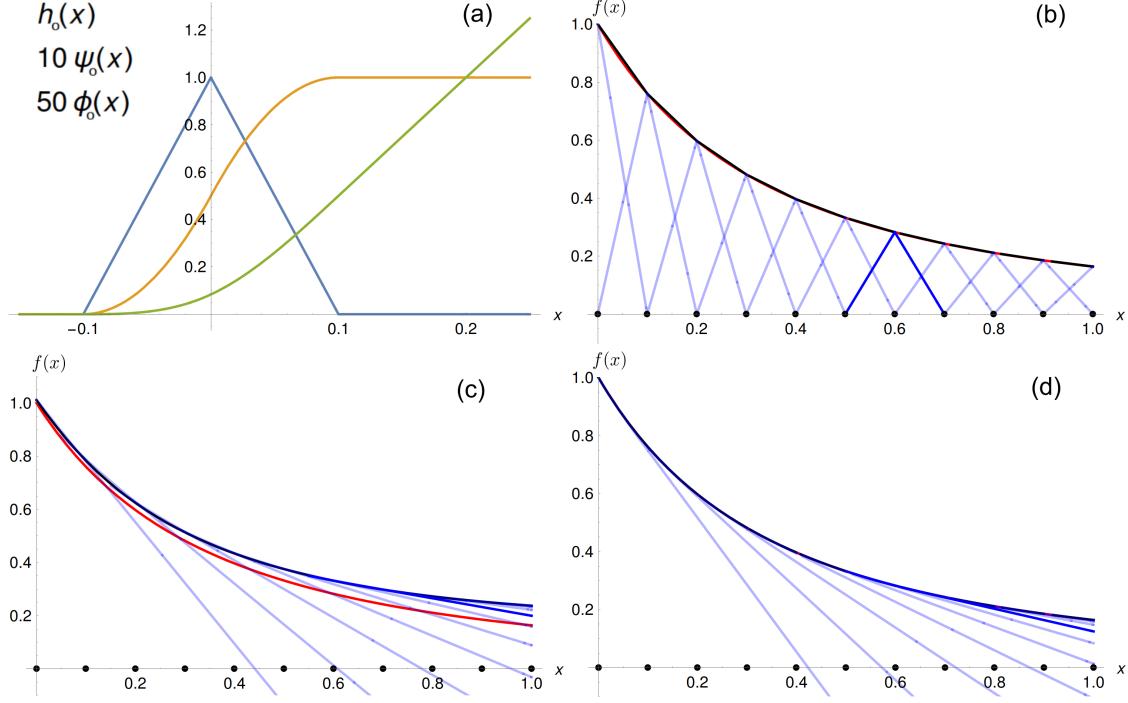


Figure A.1: (a) Functions $h_0(x)$, $\psi_0(x)$ and $\phi_0(x)$ for $N = 10$. The functions ψ_0 and ϕ_0 have been rescaled by a factor of 10 and 50, respectively. (b) Approximation (black) of the dipole function (Red) using the basis functions $h_j(x)$ (blue) on 11 gridpoints between 0 and 1 (black dots). The function $h_7(x)$, which is centered at $x = 0.6$, is highlighted to illustrate its “spike” form. The approximation matches the function so well that the true red curve is hard to see. (c)-(d) Approximation (black) of the same dipole function (red) using the basis functions ϕ_j as in Eq. (A.9) (c), and Eq. (A.10) (d). In both cases, the functions ϕ_j are plotted starting from a neighboring of their respective grid point and matching their value and their slope at the grid point with the complete approximation (black curve). In both cases the function $\phi_7(x)$ is highlighted. When the coefficients of the ϕ_j functions are fitted instead of matched to second derivatives (d), the red curve is hard to see again.

A.3 Details regarding the constrained model without the constraint $\xi_1 = 1$ (c_1 GP)

This appendix provides a detailed information regarding the c_1 GP model variant. As mentioned before, in order to account for a possible systematic error in the experimental data we could consider adding an unknown multiplicative parameter n_0 to G_E , a normalization constant. Assuming $f(x) =$

$n_0 G_E(x)$ and expanding as in Eq. (A.10), we get:

$$\begin{aligned} n_0 G_E(x) &\approx n_0 G_E(0) + n_0 x G'_E(0) + \sum_{j=0}^N n_0 G''_E(x_{j+3}) \phi_j(x) \\ &= \tilde{\xi}_1 + \tilde{\xi}_2 x + \sum_{j=0}^N \tilde{\xi}_{j+3} \phi_j(x). \end{aligned} \quad (\text{A.11})$$

With the assumption $G_E(0) = 1$, $\tilde{\xi}_1$ captures all the information about n_0 . Consider the constraint set

$$\mathcal{C}_{\tilde{\xi}} \equiv \left\{ \tilde{\xi} \in \mathbb{R}^{N+3} : \tilde{\xi}_1 \in \mathbb{R}, \tilde{\xi}_2 + \sum_{j=0}^N c_j \tilde{\xi}_{j+3} \leq 0, \tilde{\xi}_{j+3} \geq 0, j = 0, \dots, N \right\} \quad (\text{A.12})$$

where $\tilde{\xi} = \{\tilde{\xi}_j, j = 1, \dots, N+3\}$. Then the proton radius is expressed in terms of both ξ_1 and ξ_2 as:

$$r_p = \frac{1}{Q_{\max}} \sqrt{-6 \frac{\tilde{\xi}_2}{\tilde{\xi}_1}}.$$

By dividing by ξ_1 we are able eliminate the effect on the radius estimation from the floating systematic error term.

Now we discuss the partially constrained model that only incorporates the shape constraints. Let $\tilde{Y} = (\tilde{y}_1, \dots, \tilde{y}_n)^T$ with $\tilde{y}_i \equiv g_i$, and define the corresponding basis matrix $\tilde{\Phi}$ (an $n \times (N+3)$ matrix) with i th row $(1, x_i, \phi_0(x_i), \dots, \phi_N(x_i))$. Similar to the model in Sec. 4.2.5, we now have:

$$\tilde{Y} = \tilde{\Phi} \tilde{\xi} + \varepsilon, \quad \varepsilon \sim \mathcal{N}_n(0, \sigma^2 I_n), \quad \tilde{\xi} \in \mathcal{C}_{\tilde{\xi}}. \quad (\text{A.13})$$

Again, the random variables $f(0), f'(0), f''(x_0), \dots, f''(x_N)$ follow a Gaussian distribution, with the following covariance matrix:

$$\tilde{\Gamma} = \begin{bmatrix} K(0, 0) & \frac{\partial K}{\partial x'}(0, 0) & \frac{\partial^2 K}{\partial x'^2}(0, x_0) & \cdots & \frac{\partial^2 K}{\partial x'^2}(0, x_N) \\ \frac{\partial K}{\partial x}(0, 0) & \frac{\partial^2 K}{\partial x \partial x'}(0, 0) & \frac{\partial^3 K}{\partial x \partial x'^2}(0, x_0) & \cdots & \frac{\partial^3 K}{\partial x \partial x'^2}(0, x_N) \\ \frac{\partial^2 K}{\partial x^2}(x_0, 0) & \frac{\partial^3 K}{\partial x^2 \partial x'}(x_0, 0) & \frac{\partial^4 K}{\partial x^2 \partial x'^2}(x_0, x_0) & \cdots & \frac{\partial^4 K}{\partial x^2 \partial x'^2}(x_0, x_N) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 K}{\partial x^2}(x_N, 0) & \frac{\partial^3 K}{\partial x^2 \partial x'}(x_N, 0) & \frac{\partial^4 K}{\partial x^2 \partial x'^2}(x_N, x_0) & \cdots & \frac{\partial^4 K}{\partial x^2 \partial x'^2}(x_N, x_N) \end{bmatrix}_{(N+3) \times (N+3)}. \quad (\text{A.14})$$

The joint posterior distribution of the model parameters with partial constraints is:

$$P(\tilde{\xi}, \tau^2, \sigma^2 | \tilde{Y}) \propto \left\{ (\sigma^2)^{-n/2} e^{-\frac{\|\tilde{Y} - \tilde{\Phi} \tilde{\xi}\|^2}{2\sigma^2}} \right\} \left\{ (\tau^2)^{-(N+3)/2} e^{-\tilde{\xi}^T \tilde{\Gamma}^{-1} \tilde{\xi} / (2\tau^2)} \mathbb{1}_{\mathcal{C}_{\tilde{\xi}}}(\tilde{\xi}) \right\} (\tau^2)^{-1} (\sigma^2)^{-1}. \quad (\text{A.15})$$

Therefore the estimation of the proton radius based on the posterior samples of $\tilde{\xi}_1^{(t)}$ and $\tilde{\xi}_2^{(t)}$, with $t = 1, \dots, T$ is:

$$\tilde{r}_p = T^{-1} \sum_{t=1}^T \frac{\sqrt{-6\tilde{\xi}_2^{(t)}/\tilde{\xi}_1^{(t)}}}{Q_{\max}}. \quad (\text{A.16})$$

Note that Proposition 1 and 2 in Appendix A.1 still hold for the c₁GP model. To see how Proposition 2 holds, the normalizing constant is:

$$M_{\tilde{\xi}} = \int_{C_{\tilde{\xi}}} (\tau^2)^{-(N+3)/2} (|\tilde{\Gamma}|)^{(-1/2)} e^{-\frac{1}{2\tau^2} \tilde{\xi}^T \tilde{\Gamma}^{-1} \tilde{\xi}} d\tilde{\xi}.$$

Since $\tilde{\xi}_1 \in \mathbb{R}$, by the change of variable $\tilde{\xi}' = \tilde{\xi}/\tau$, it is easy to see $C_{\tilde{\xi}'} = C_{\tilde{\xi}}$. Thus, the integration does not depend on τ as well.

A.4 Details regarding the choices of priors and hyperparameters, and the Gibbs sampling steps

Choice of ν : Assuming f to be smooth in addition to being convex, the minimum possible smoothness required is twice differentiability. In an unconstrained Gaussian process regression, Corollary 3.1 and 3.2 of [207] show that the point-wise posterior credible intervals contain the true function with at least the nominal coverage probability provided that the prior smoothness is set to be less than or equal to the smoothness of the underlying function. We conjecture that this will continue to hold in the case of function estimation using a constrained Gaussian process, which motivated the following choice of ν . It is well-known that the reproducing kernel Hilbert space of Gaussian process endowed with Matérn covariance kernel with smoothness ν consists of Hölder class of smoothness $\nu + 0.5$. Hence, the choice of $\nu = 2.5$ (corresponding to twice-differentiable functions) ensures that the posterior credible intervals will not underestimate the uncertainty in estimating r_p . The choices of $\nu = 3$ and $\nu = 3.5$ were made in order to assess the impact of this hyperparameter on the estimation of r_p . The following is the most general definition of the Matérn covariance kernel:

$$k_{\nu}(r) \equiv \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu}r}{\ell} \right)^{\nu} K_{\nu} \left(\frac{\sqrt{2\nu}r}{\ell} \right), \quad (\text{A.17})$$

where K_{ν} is the modified Bessel function of the second kind and $\Gamma(z)$ is the gamma function. For the choice $\nu = 5/2$ this equation equals Eq (4.30).

Choice of τ : τ controls the prior signal to noise ratio. An objective choice is the non-informative prior $p(\tau^2) \propto 1/\tau^2$.

Choice of ℓ : The parameter ℓ is typically called the *length-scale* parameter of a Gaussian process. It controls the rate of decay of the covariance kernel with the inter-site distances. Typically one chooses ℓ so that the correlation between two points far apart in the covariate space is negligible. Empirically, one can use a variogram plot of the data to estimate the value of ℓ . Instead, we used a cross-validation approach to estimate the value of ℓ . We varied ℓ in the range $\{0.05, 0.1, 0.5, 1, 2, \dots, 20\}$, where the overall scale is the same after scaling Q^2 to $[0, 1]$, and we used 5-fold cross-validation (repeating the cross validation five times) to choose the optimal value of ℓ that minimizes the predictive mean squared error.

Choice of σ : Based on the error values in the experimental data, we noticed that the estimated error (standard deviation) is no larger than 0.01, which motivated our choice of values for σ for the pseudo data analysis, since it is computationally less expensive than adding a prior distribution to σ . For the real data analysis, we allowed σ to vary by putting an objective prior such that $p(\sigma^2) \propto 1/\sigma^2$.

Choice of N : The number of grid points N in Eq. (A.3) directly influences the approximation power of the function estimation method. The role of the grid points is to project a Gaussian process onto a regular grid. The function at any intermediate value is then obtained using linear interpolation (do not confuse with the linear interpolation made by the functions $h_j(x)$ in Sec. A.2). It is unreasonable to set N to a value larger than the sample size n , since that may lead to overfitting. In order to conduct a thorough analysis on the full dataset by constrained GP and following [151] we considered $N = \{n/4, n/2, n\}$ in our data analysis.

Gibbs sampling: Given the above choice of hyperparameters, the joint posterior distribution defined in Sec. 4.2.5:

$$P(\xi, \tau^2, \sigma^2 | Y) \propto \left\{ (\sigma^2)^{-n/2} e^{-\frac{\|Y - \Phi\xi\|^2}{2\sigma^2}} \right\} \left\{ (\tau^2)^{-(N+2)/2} e^{-\xi^T \Gamma^{-1} \xi / (2\tau^2)} \mathbb{1}_{\mathcal{C}_\xi}(\xi) \right\} (\tau^2)^{-1} (\sigma^2)^{-1}, \quad (\text{A.18})$$

can be updated by Gibbs sampling,

- Update $[\xi | \tau^2, \sigma^2, \Phi, Y] \sim N(\mu_\xi, \Sigma_\xi) \mathbb{1}_{\mathcal{C}_\xi}(\xi)$, with $\Sigma_\xi = (\Phi^T \Phi / \sigma^2 + \Gamma^{-1} / \tau^2)^{-1}$ and $\mu_\xi = \Sigma_\xi^{-1} \Phi^T Y / \sigma^2$,
- Update $[\tau^2 | \xi, \sigma^2, \Phi, Y] \sim \text{IG}(a_\tau, b_\tau)$, with $a_\tau = (N+2)/2$ and $b_\tau = \xi^T \Gamma^{-1} \xi / 2$,

- Update $[\sigma^2 \mid \xi, \tau^2, \Phi, Y] \sim \text{IG}(a_\sigma, b_\sigma)$, with $a_\sigma = n/2$ and $b_\sigma = \|Y - \Phi\xi\|^2/2$,

where IG denotes an Inverse Gamma distribution. Note that the above Gibbs sampling procedure is applicable to all proposed GP models associated with different constraint sets, and Φ, Y, N, ξ, Γ vary in different cases.

A.5 Pseudo-data analysis

Before analyzing the real data, we tested the GP methods on synthetically generated datasets. To mimic the real dataset, we used the Q^2 from the electron-proton scattering data obtained from Mainz [208, 209, 210], and generated the pseudo G_E data using the “dipole function” given by:

$$G_E(Q^2) = \left(1 + \frac{r_p^2 Q^2}{12}\right)^{-2},$$

r_p being the pseudo-radius of the proton. The “dipole function” is a good proxy for the electric form factor and serves as the ground truth for conducting the simulation study. In the following numerical examples we set $r_p = 0.84$ fm. We extracted $n = 500$ sample points of Q^2 from the Mainz dataset in three regimes: i) low $Q^2(\leq 1.36 \text{ fm}^{-2})$, ii) medium $Q^2(\leq 4.85 \text{ fm}^{-2})$, and iii) high $Q^2(\leq 25.12 \text{ fm}^{-2})$. To generate noisy observations of G_E , we added independent and identically distributed zero-mean Gaussian noise with standard deviations in the set $\{0, 0.002, 0.005, 0.01\}$, where $\sigma = 0$ means no noise is added to G_E . The interval for σ is chosen to contain the typical observed errors in the Mainz data [208, 209, 210]. The four models (cGP, c_0 GP, c_1 GP, uGP) are used with smoothness parameter $\nu = 2.5$. Following [151], the number of basis functions (the same as the number of grid point) is set to $N = \{n/4, n/2, n\}$.

To select the optimal length-scale parameter ℓ we developed a cross validation procedure. By dividing the total data into 80% training and 20% testing datasets, the cross validation procedure seeks to minimize the MSE defined as the average of squared deviations from the 20 % testing set and the model predictions.

Fig. A.2 shows the MSE as a function of ℓ for the three regimes. In low and medium Q^2 regimes, the MSEs dropped fast for relatively small values of ℓ , and then stayed flat as ℓ increased. On the other hand, in the high Q^2 regime the MSEs dropped first and then increased slightly as ℓ increased. For very small values of ℓ (~ 0.1) we obtained much higher MSEs in all cases, since a very small

value of ℓ reduces the correlation of the constrained Gaussian process between neighboring points and fails to borrow information from neighbors for an accurate extrapolation.

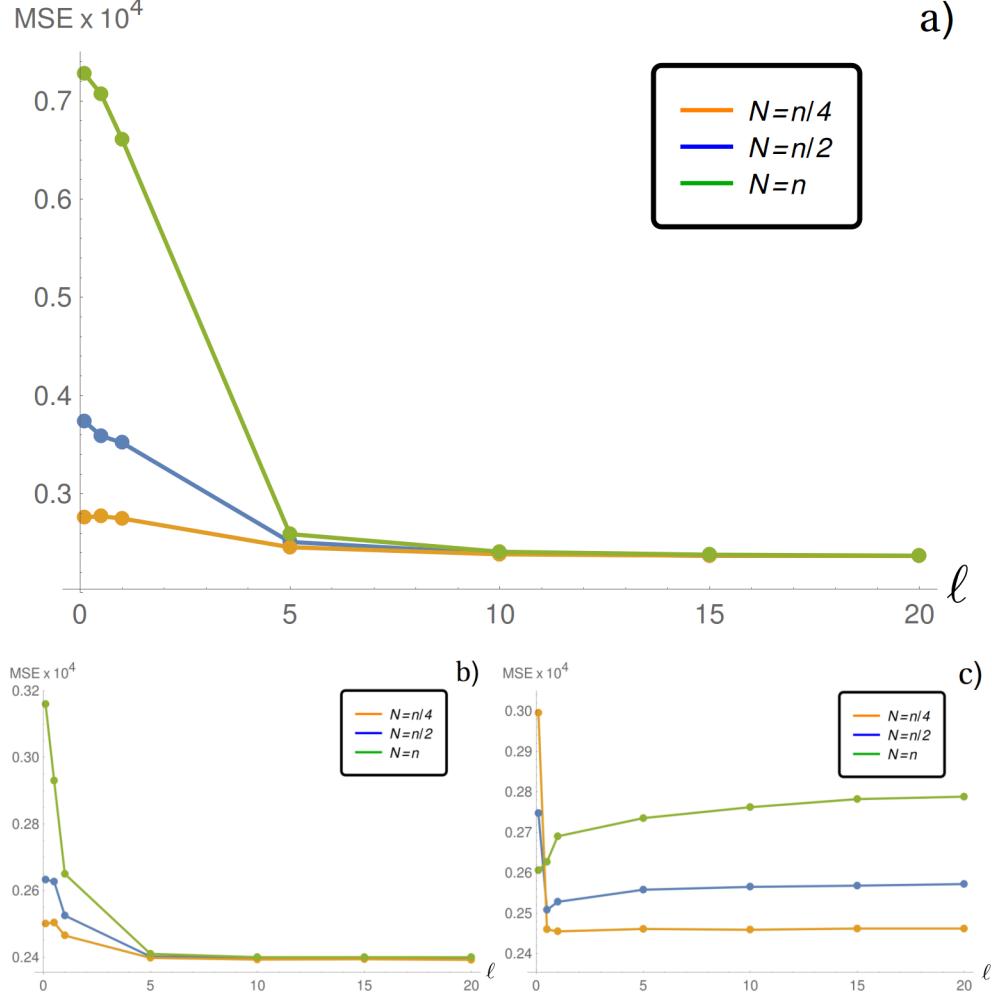


Figure A.2: MSEs versus ℓ for $\nu = 2.5$, $\sigma = 0.005$ and $n = 500$. (a) shows the MSEs for data in regime i), (b) for regime ii), and (c) for regime iii). In each plot, the orange line represents the MSEs with the number of grid points $N = n/4$, the blue line represents $N = n/2$, and the green line for $N = n$.

The MSE behavior as a function of ℓ is fundamentally different for the high regime in comparison with the low and medium regimes. The reason for this difference is the way the experimental points are distributed across the entire Q^2 range. In the Mainz data with range 25.12 fm^{-2} , the Q^2 values are collected more often with small values and only a few are collected with large values: 70% of the Q^2 values are less than 5 fm^{-2} but only around 5% of Q^2 are greater than 17 fm^{-2} . Therefore, when re-scaling to the $[0,1]$ interval in the low regime, the Q^2 values are more evenly distributed,

while in the high regime most of the Q^2 values are concentrated around 0 and only a few are close to 1. The selected length-scale parameter ℓ depends on the dispersion of the Q^2 values, thus, in the high Q^2 regime the cross validation procedure tends to select a smaller value of ℓ so that the correlation between two points with long distance is relatively small. On the other hand, in the low Q^2 regime the cross validation procedure tends to select a larger value of ℓ that leads to a stronger correlation between any two points of Q^2 . The selected optimal values for ℓ are $\ell = 20$ for both low and medium regime, and $\ell = 1$ for the high regime. Since the scale for ℓ is on the re-scaled variable x , $\ell \gg 1$ can be interpreted as an indicative that the whole range of Q^2 that was considered is highly correlated. We observed that in all three regimes using a smaller number of the basis functions (N) led to smaller values of MSEs on the 20% held out Q^2 values.

Tables A.1-A.3 show the posterior summaries of the estimates of the radius r_p and 95% credible intervals, the lower and upper bound denoted respectively by CI_l and CI_u . Since we know the generated G_E values as well as the generated radius r_p , we are able to evaluate the results of the GP methods with different constraints and in different regimes.

σ	0	0.002	0.005	0.01
cGP	r_p	0.8400	0.8384	0.8402
	CI_l	0.8393	0.8340	0.8340
	CI_u	0.8403	0.8429	0.8488
c_0 GP	r_p	0.8400	0.8359	0.8364
	CI_l	0.8393	0.8272	0.8243
	CI_u	0.8402	0.8440	0.8536
c_1 GP	r_p	0.8388	0.8391	0.8435
	CI_l	0.8375	0.8301	0.8302
	CI_u	0.8401	0.8474	0.8612
uGP	r_p	0.8389	0.8363	0.8336
	CI_l	0.8376	0.8182	0.8089
	CI_u	0.8403	0.8555	0.8601

Table A.1: Posterior estimates of the radius and credible interval for cGP, c_0 GP, c_1 GP and uGP with $N = n/4$, $\nu = 2.5$ and $\ell_{\text{opt}} = 20$ for a subset of data of size $n = 500$ in regime i).

We found that in the presence of noise, imposing all the constraints (cGP) reduces the uncertainty in the estimation while maintaining accuracy, while only imposing the constraint at zero (c_0 GP), tends to give accurate results but with wider credible intervals. If we only consider the shape constraints (c_1 GP), the estimates of the radius are somewhat variable as the noise level increases, becoming more biased for the higher Q^2 regimes. The unconstrained model (uGP) leads

σ	0	0.002	0.005	0.01
cGP	r_p	0.8399	0.8337	0.8327
	CI _l	0.8394	0.8297	0.8276
	CI _u	0.8404	0.8381	0.8413
c_0 GP	r_p	0.8399	0.8376	0.8341
	CI _l	0.8394	0.8311	0.8276
	CI _u	0.8404	0.8459	0.8413
c_1 GP	r_p	0.8385	0.8317	0.8247
	CI _l	0.8373	0.8232	0.8152
	CI _u	0.8399	0.8397	0.8354
uGP	r_p	0.8385	0.8344	0.8306
	CI _l	0.8372	0.8208	0.8126
	CI _u	0.8398	0.8487	0.8544

Table A.2: Posterior estimates of the radius and credible interval for cGP, c_0 GP, c_1 GP and uGP with $N = n/4$, $\nu = 2.5$ and $\ell_{\text{opt}} = 20$ for a subset of data of size $n = 500$ in regime ii).

to the widest credible intervals in general and to reasonably good estimates when the noise level is small. Comparing results across different regimes, we found that in the medium and high Q^2 regimes our methods tend to give lower estimates at the origin as the noise level increases. The trend of obtaining lower estimates of r_p in higher regimes could be caused by several factors. A possible factor could be that the model is only able to borrow information from one side when estimating over the boundary (the origin), but the model hyperparameters are selected according to the overall model fitting. As we mention in Chapter 7, we shall explore this trend in a future work to improve the GP model.

Fig. A.3 shows the model fitting in all regimes as $Q^2 \rightarrow 0$, and the inset plot at the right upper corner shows the respective model fit in the entire range. We can see that without the restriction $n_0 G_E(0) = 1$, the estimates of c_1 GP and uGP are off from the truth (red line) for small values of Q^2 . On the contrary, and as expected, cGP and c_0 GP agree with the truth as $Q^2 \rightarrow 0$. Without the shape constraints ($f'_\xi(x) < 0$ and $f''_\xi(x) > 0$) we found in the high Q^2 regime that the estimates of c_0 GP and uGP are not even convex near higher values of Q^2 .

We conclude that incorporating all the physical constraints is crucial for providing a realistic estimate of the radius. It is also evident that the low Q^2 regime data informs about the radius more reliably than the high Q^2 regime data, at least under the assumption of additive independent and identically distributed errors and the fulfillment of all the constraints by the data. However, in the real-data scenario with unknown errors, and possibly with some violation of the constraints,

σ	0	0.002	0.005	0.01
cGP	r_p	0.8386	0.8312	0.8213
	CI _l	0.8249	0.8252	0.8131
	CI _u	0.8454	0.8373	0.8295
c_0 GP	r_p	0.8169	0.8355	0.8294
	CI _l	0.8151	0.8241	0.8126
	CI _u	0.8188	0.8478	0.8451
c_1 GP	r_p	0.8303	0.8176	0.8071
	CI _l	0.8019	0.8077	0.7963
	CI _u	0.8409	0.8290	0.8191
uGP	r_p	0.8400	0.8229	0.8202
	CI _l	0.8387	0.8030	0.7918
	CI _u	0.8404	0.8438	0.8448

Table A.3: Posterior estimates of the radius and credible interval for cGP, c_0 GP, c_1 GP and uGP with $N = n/4$, $\nu = 2.5$ and $\ell_{\text{opt}} = 1$ for a subset of data of size $n = 500$ in regime iii).

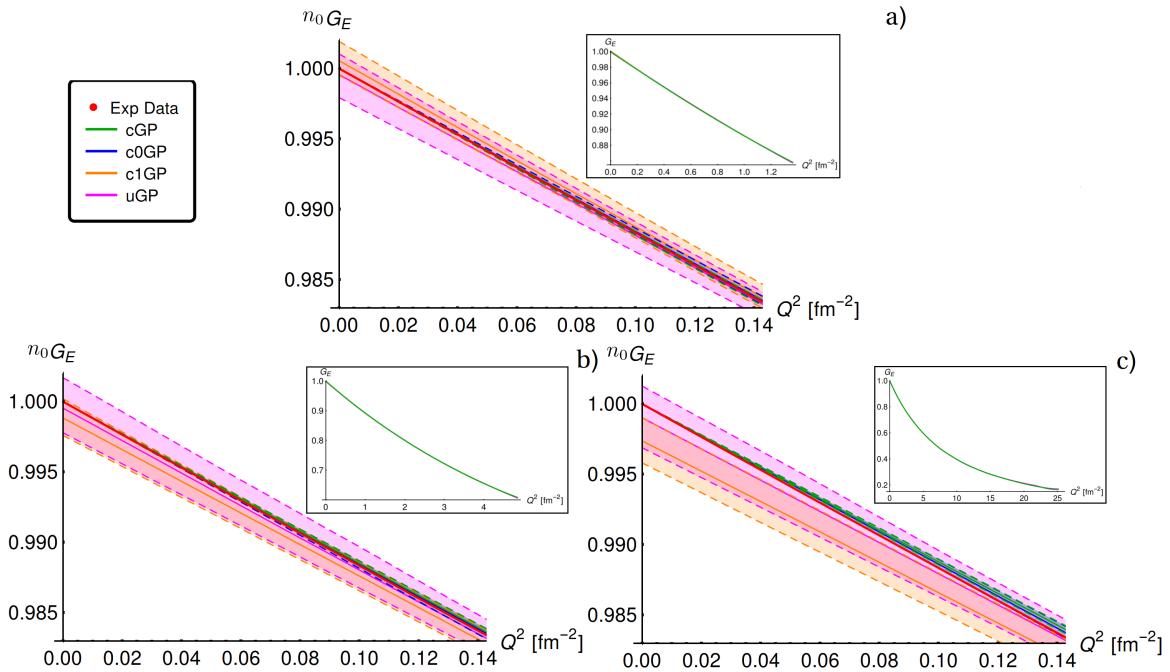


Figure A.3: Model fits in regimes i)-iii) with noise level $\sigma = 0.005$. (a) is for regime i), (b) for regime ii) and (c) for regime iii). The red line represents the true function values; the green line represents the cGP estimates; the blue line represents the c_0 GP; the orange line represents the c_1 GP; the purple line represents uGP. In each case the 95% point-wise credible intervals are delimited by dashed lines with the respective color.

specially of the first one ($f(0) = 1$), it might be important to consider the full dataset to take into account all sources of variation in the analysis.

A.6 Details regarding the Electron-Scattering Data Analysis

Fig. A.4 (a) shows the 5-fold cross-validation MSEs for the high regime over $\ell \in \{0.05, 0.1, 0.5, 1\}$ of cGP for $n = 250$ with $N = \{n/4, n/2, n\}$. Since it is evident that smaller values of ℓ (≤ 0.5) cause the MSE to increase, we focused on $\ell \geq 0.5$ in the following analysis, following also our observations from the MSEs results in the high Q^2 regime of pseudo generated data. From the results used to plot Fig. A.4 (a) and Fig. A.2, we noted that choosing a smaller number of grid points leads to more accurate predictions in terms of MSE when $\ell > 0.5$. We therefore chose the number of grid points $N = n/4$ and considered different smoothness parameters $\nu = \{2.5, 3, 3.5\}$ to perform the finer cross validation procedure to select the optimal value of ℓ in the grid $\ell \in \{0.5, 1, 1.5, 2, 2.5\}$ (Fig. A.4 (a) (Inset)). We observed that the MSEs increased as ℓ increased from 0.5 to 2.5 for all ν and $\nu = 2.5$ gave relatively lower MSEs in this case. The results of our analysis guide us to choose $\ell_{\text{opt}} = 0.5$ for the full data set analysis.

Fig. A.4 (b) shows the 5-fold cross-validation MSEs for the low regime. Again, based on the results in the pseudo data analysis indicating that a larger value of ℓ is preferred, we conducted the cross validation for cGP model over the parameter set $\ell \in \{1, 5, 10, 15, 20\}$. Fig. A.4 (b) shows that as ℓ increases, the MSEs drops fast first and then stays stable for large values of ℓ (≥ 10). Also, in the low regime cGP with a smaller number of grid points ($N = n/4$) gave lower MSEs, which is similar to the full dataset case. This analysis led us to choose $\ell_{\text{opt}} = 10$ on the low Q^2 set.

Tables A.4 and A.5 show the posterior summaries of the four models for the estimates of the radius r_p and 95% credible intervals, the lower and upper bound denoted respectively by CI_l and CI_u

The following figures show the detailed histograms for the 400 MCMC samples for the four models discussed in the real data analysis section. In all cases $\nu = 2.5$ was used. Fig. A.5 and A.6 show the results for the high regime with $N = n/4$, and $N = n$ respectively. Fig. A.7 and A.8 show the results for the low regime with $N = n/4$, and $N = n$ respectively. Fig. A.9 and A.10 show the samples of ξ_1 ($n_0 G_E(0)$) for c₁GP and uGP for high and low regimes, respectively.

ν	2.5	2.5	3	3	3.5	3.5
N	n/4	n	n/4	n	n/4	n
cGP	r_p	0.8435	0.8452	0.8413	0.8431	0.8425
	CI _l	0.8396	0.8426	0.8265	0.8406	0.8301
	CI _u	0.8481	0.8476	0.8524	0.8457	0.8511
c_0 GP	r_p	0.8355	0.8448	0.8226	0.8431	0.8319
	CI _l	0.8254	0.8373	0.8045	0.8328	0.8167
	CI _u	0.8467	0.8519	0.8415	0.8527	0.8467
c_1 GP	r_p	0.8423	0.8311	0.8295	0.8259	0.8347
	CI _l	0.8346	0.8266	0.7993	0.8217	0.8111
	CI _u	0.8507	0.8369	0.8447	0.8303	0.8461
uGP	r_p	0.8474	0.8577	0.7665	0.8563	0.8253
	CI _l	0.8256	0.8419	0.7374	0.8398	0.7969
	CI _u	0.8683	0.8742	0.7938	0.8744	0.8530

Table A.4: High regime posterior estimates of the radius and credible interval for cGP, c_0 GP, c_1 GP and uGP with $N = \{n/4, n\}$ and $\nu = \{2.5, 3, 3.5\}$ and $n = 1422$.

ν	2.5	2.5	3	3	3.5	3.5
N	n/4	n	n/4	n	n/4	n
cGP	r_p	0.8529	0.8550	0.8543	0.8561	0.8550
	CI _l	0.8488	0.8514	0.8503	0.8529	0.8511
	CI _u	0.8576	0.8587	0.8591	0.8593	0.8597
c_0 GP	r_p	0.8399	0.8408	0.8411	0.8432	0.8399
	CI _l	0.8213	0.8269	0.8143	0.8309	0.8168
	CI _u	0.8533	0.8516	0.8598	0.8547	0.8584
c_1 GP	r_p	0.8725	0.8719	0.8721	0.8731	0.8735
	CI _l	0.8613	0.8626	0.8628	0.8660	0.8640
	CI _u	0.8857	0.8815	0.8820	0.8799	0.8836
uGP	r_p	0.8573	0.8618	0.8612	0.8667	0.8593
	CI _l	0.8212	0.8424	0.8321	0.8467	0.8249
	CI _u	0.8898	0.8830	0.8897	0.8851	0.8899

Table A.5: Low regime posterior estimates of the radius and credible interval for cGP, c_0 GP, c_1 GP and uGP with $N = \{n/4, n\}$ and $\nu = \{2.5, 3, 3.5\}$ and $n = 500$.

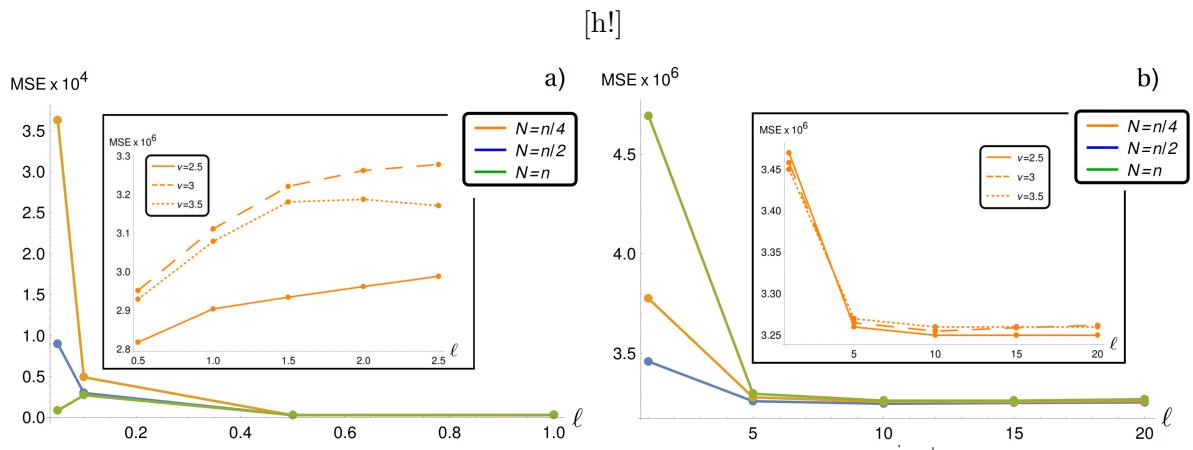


Figure A.4: (a) MSEs over the range $\ell \in [0, 1]$ for cGP model in the high regime with $n = 250$, $N = n/4$, $n/2$, n . (b) MSEs over ℓ for cGP model in the low regime with $N = \{n/4, n/2, n\}$, $n = 500$ and $\nu = 2.5$. In both graphs the green line represents the case with $N = n$, the blue line represents $N = n/2$, and the orange line represents $N = n/4$. (a) (Inset) MSEs over $\ell \in [0.5, 2.5]$ for cGP model with $N = n/4$ on full dataset. (b) (Inset) MSEs over $\ell \in [1, 20]$ for cGP model with $N = n/4$ on the low regime. In both insets $\nu = 2.5$ (solid), $\nu = 3$ (dashed), $\nu = 3.5$ (dot-dashed).

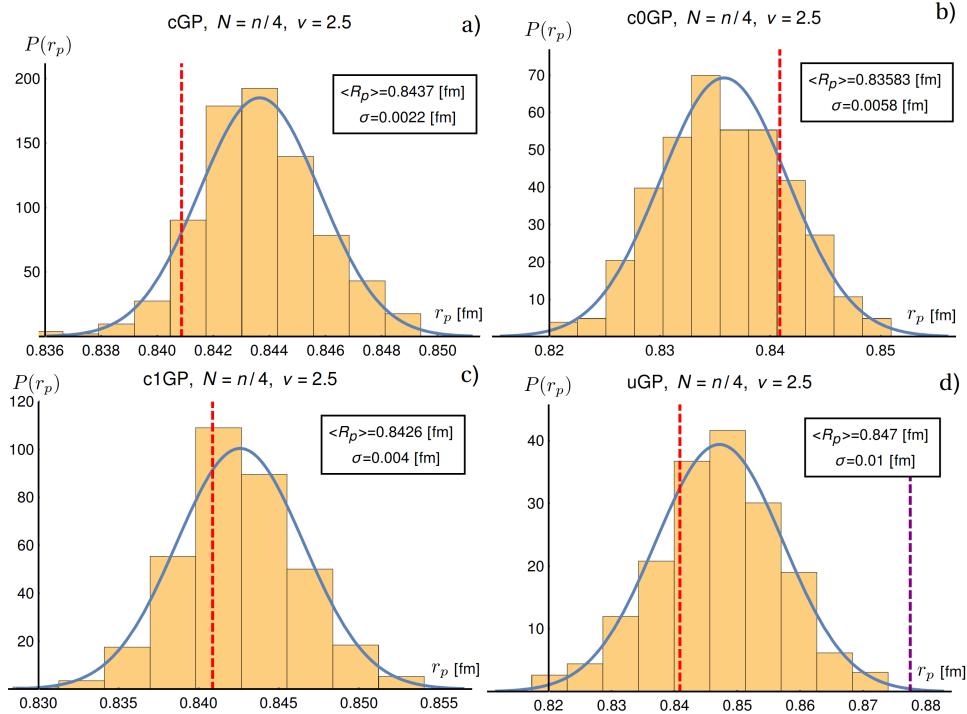


Figure A.5: Histogram plots of MCMC samples of the radius r_p for cGP (a), c₀GP (b), c₁GP (c) and uGP (d) with $N = n/4$ and $\nu = 2.5$ for the full dataset. The red and purple vertical dashed lines indicate the values of 0.84087 fm and 0.8775 fm, respectively.

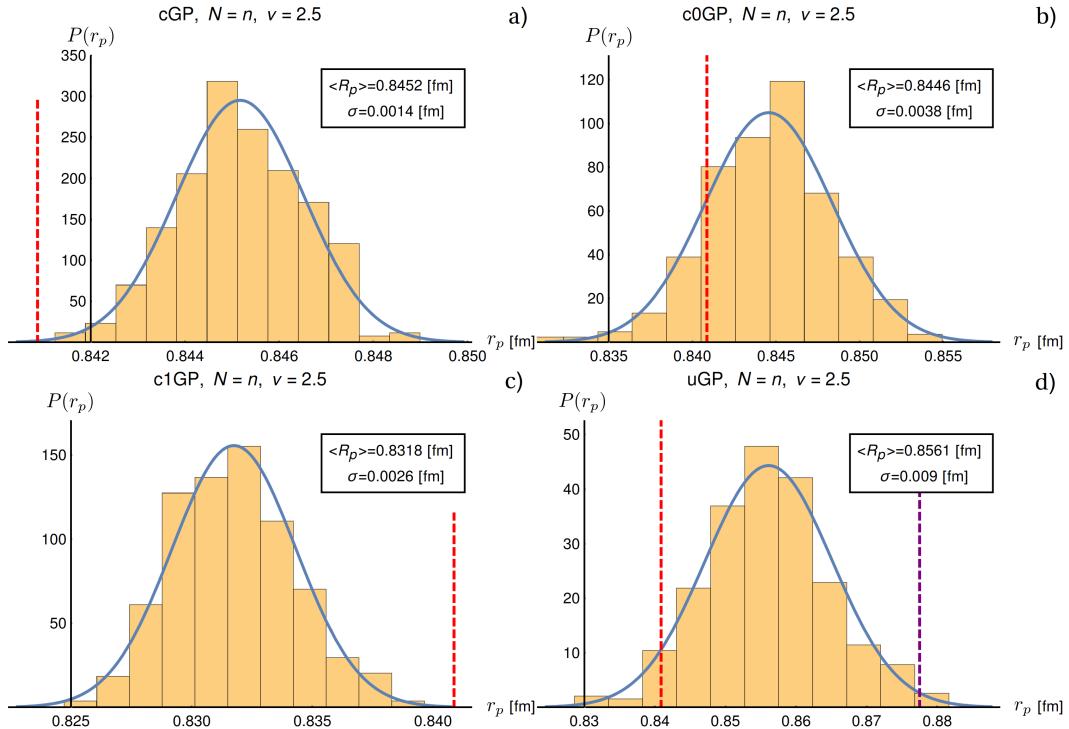


Figure A.6: Histogram plots of MCMC samples of the radius r_p for cGP (a), c₀GP (b), c₁GP (c) and uGP (d) with $N = n$ and $\nu = 2.5$ for the full dataset. The red and purple vertical dashed lines indicate the values of 0.84087 fm and 0.8775 fm, respectively.

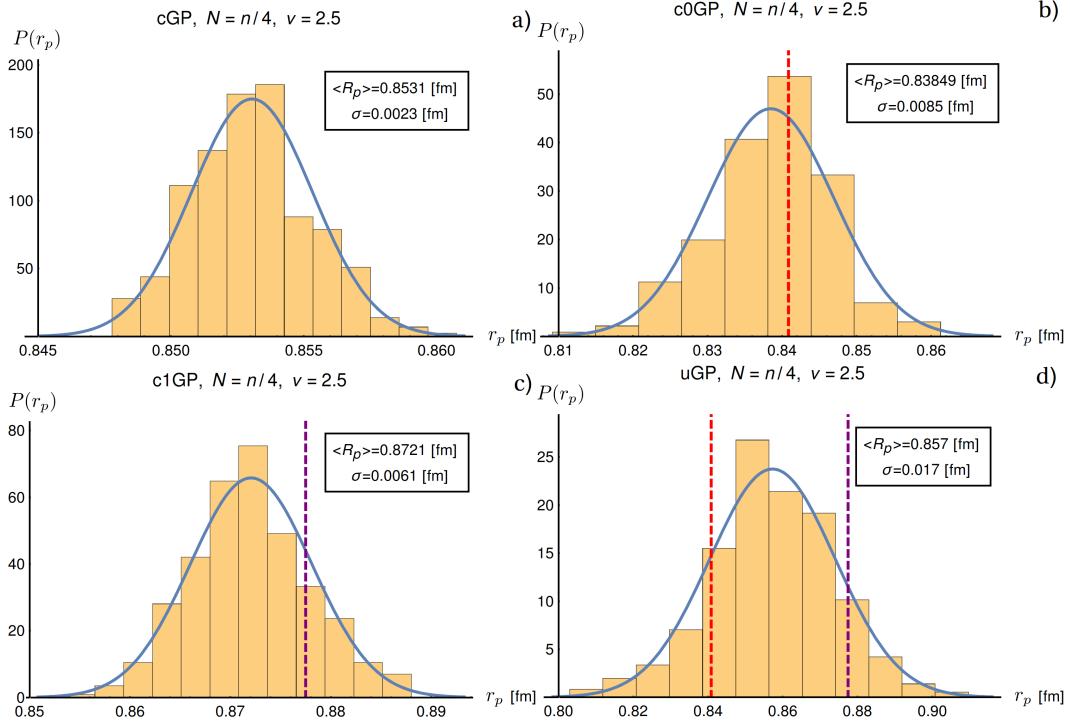


Figure A.7: Histogram plots of MCMC samples of the radius r_p for cGP (a), c_0 GP (b), c_1 GP (c) and uGP (d) with $N = n/4$ and $\nu = 2.5$ in the low regime case. The red and purple vertical dashed lines indicate the values of 0.84087 fm and 0.8775 fm, respectively.

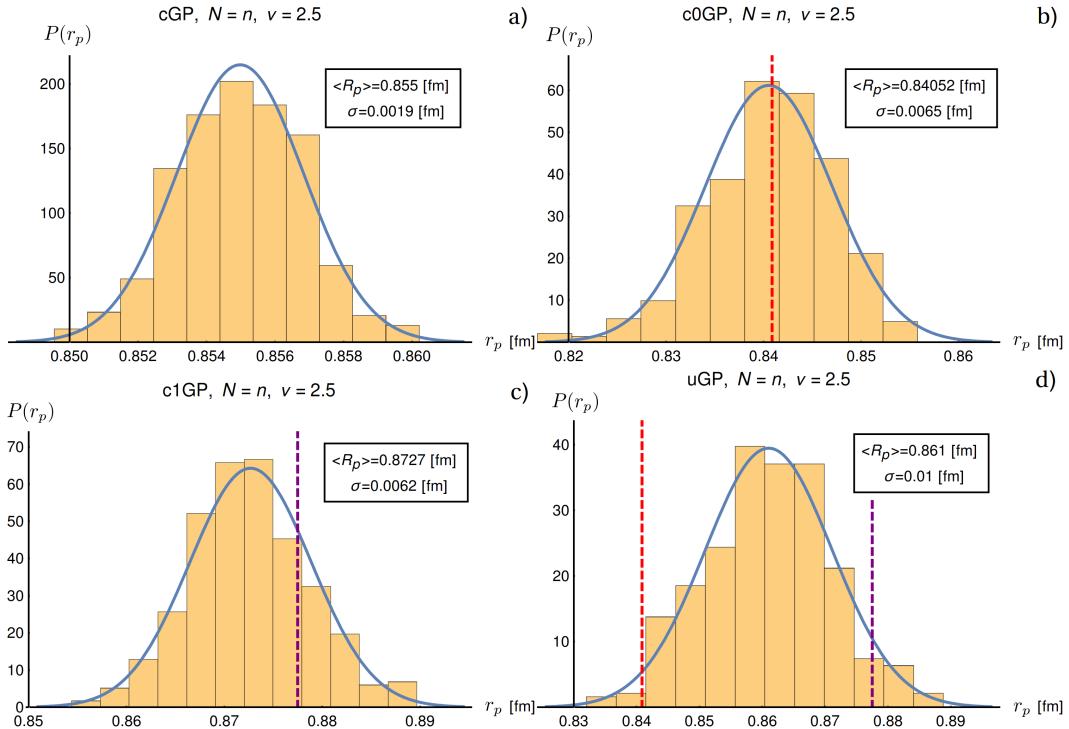


Figure A.8: Histogram plots of MCMC samples of the radius r_p for cGP (a), c₀GP (b), c₁GP (c) and uGP (d) with $N = n$ and $\nu = 2.5$ in the low regime. The red and purple vertical dashed lines indicate the values of 0.84087 fm and 0.8775 fm, respectively.

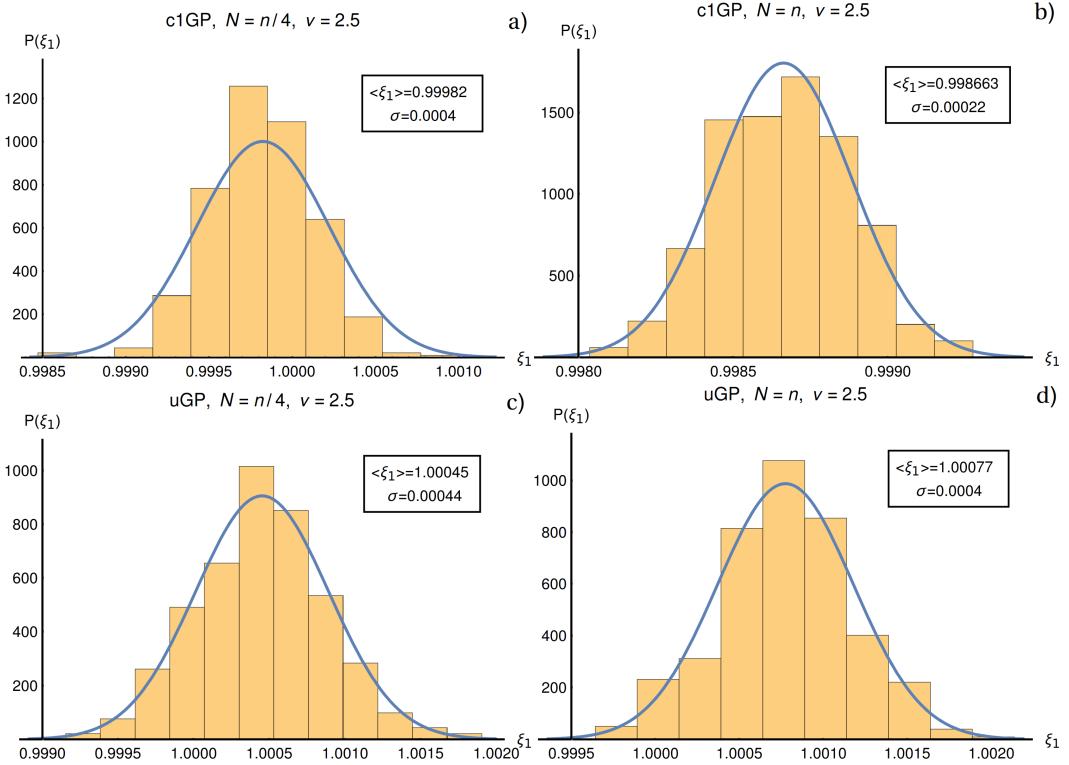


Figure A.9: Histogram plots of MCMC samples of ξ_1 ($n_0 G_E(0)$) for c1GP and uGP with $N = \{n/4, n\}$ and $\nu = 2.5$ for the full dataset. (a) ($N = n/4$) and (b) ($N = n$) show the results of c1GP, (c) ($N = n/4$) and (d) ($N = n$) show the results of uGP.

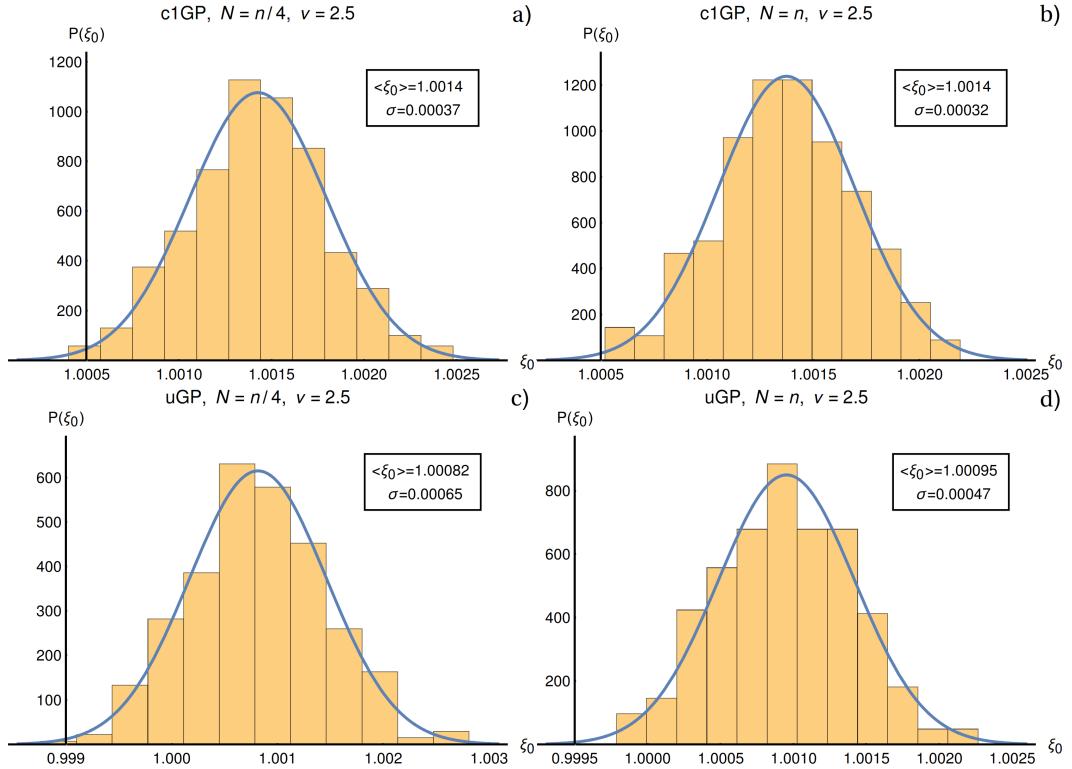


Figure A.10: Histogram plots of MCMC samples of ξ_1 ($n_0 G_E(0)$) for c1GP and uGP with $N = \{n/4, n\}$ and $\nu = 2.5$ for the low regime. (a) ($N = n/4$) and (b) ($N = n$) show the results of c1GP, (c) ($N = n/4$) and (d) ($N = n$) show the results of uGP.

APPENDIX B

THE TRANSFER FUNCTION FORMALISM

B.1 Mathematical proofs regarding the TF formalism

B.1.1 Transfer Functions structure

This appendix presents a formal proof on the structure of the transfer functions (Eq (3.4)), namely that the first order change coefficients on the parameters ω due to a perturbation on observation y_j are:

$$\mathcal{T}\mathcal{F}_j^\omega = \mathcal{H}^{-1} \nabla F_j \sigma_j^{-2}, \quad (\text{B.1})$$

where \mathcal{H}^{-1} is the inverse of the Hessian matrix of $\chi^2/2$ defined in (2.18) and ∇F_j is the gradient with respect to the parameters ω of the function F being fit, which is evaluated at observation y_j .

Let us assume that we are at the minimum ω_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 ω_k (K in total) should be zero:

$$\frac{1}{2} \frac{\partial \chi^2}{\partial \omega_k} \Big|_{(\omega_0, \mathbf{y}_0)} \equiv G_k(\omega, \mathbf{y}) \Big|_{(\omega_0, \mathbf{y}_0)} = 0, \quad (\text{B.2})$$

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 ω_k . The G_k are the following functions of both the parameters and the observations:

$$G_k(\omega, \mathbf{y}) = \sum_{j=1}^J \frac{(F_j - y_j)}{\sigma_j^2} \frac{\partial F_j}{\partial \omega_k}. \quad (\text{B.3})$$

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

$$\delta y_j \frac{\partial G_k}{\partial y_j} = - \sum_{i=1}^K \delta \omega_i \frac{\partial G_k}{\partial \omega_i}. \quad (\text{B.4})$$

We can arrange all K equations into a matrix form:

$$\begin{aligned}\frac{\partial G_1}{\partial y_j} \delta y_j &= -\left(\frac{\partial G_1}{\partial \omega_1} \delta \omega_1 + \frac{\partial G_1}{\partial \omega_2} \delta \omega_2 \dots + \frac{\partial G_1}{\partial \omega_K} \delta \omega_K\right) \\ \frac{\partial G_2}{\partial y_j} \delta y_j &= -\left(\frac{\partial G_2}{\partial \omega_1} \delta \omega_1 + \frac{\partial G_2}{\partial \omega_2} \delta \omega_2 \dots + \frac{\partial G_2}{\partial \omega_K} \delta \omega_K\right) \\ &\vdots & \vdots & \vdots \\ \frac{\partial G_K}{\partial y_j} \delta y_j &= -\left(\frac{\partial G_K}{\partial \omega_1} \delta \omega_1 + \frac{\partial G_K}{\partial \omega_2} \delta \omega_2 \dots + \frac{\partial G_K}{\partial \omega_K} \delta \omega_K\right),\end{aligned}$$

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:

$$\begin{aligned}-\nabla F_j \sigma_j^{-2} \delta y_j &= -\mathcal{H} \boldsymbol{\delta \omega}, \\ [\mathcal{H}^{-1} \nabla F_j \sigma_j^{-2}] \delta y_j &= \boldsymbol{\delta \omega},\end{aligned}$$

where, since the perturbation δy_j can be made arbitrary small, we must conclude that the quantity in brackets is what we are looking for. These are the linear coefficients connecting a small change in y_j with the small change in every parameter ω_k , proving Eq (B.1).

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 (B.1), 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 (B.3) only the term multiplying δy_j survives. Everything else gets cancelled 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 $(\boldsymbol{\omega}_0, \mathbf{y}_0)$ 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 (B.1) (see [211] pages 41-42).

B.1.2 Comparing the transfer function variance to the standard approach's variance

This appendix presents a formal proof of the statement discussed at the end of Sec 3.2. When the Hessian matrix (2.18) only involves the linear part of the model:

$$\mathcal{H}_{i,k} = \sum_{j=1}^J \frac{1}{\sigma_j^2} \left[\left(\frac{\partial F(q_j, \boldsymbol{\omega})}{\partial \omega_i} \right) \left(\frac{\partial F(q_j, \boldsymbol{\omega})}{\partial \omega_k} \right) \right], \quad (\text{B.5})$$

then, the variance calculated using the standard approach (2.19), and the variance calculated using the transfer function formalism (3.6), are identical, namely:

$$\Delta m^2 = \nabla m \mathcal{H}^{-1} \nabla m, \quad (\text{B.6})$$

$$\Delta m^2 = \sum_{j=1}^J \left[\nabla m \mathcal{H}^{-1} \nabla F(q_j, \boldsymbol{\omega}) \sigma_j^{-2} \right]^2 \sigma_j^2. \quad (\text{B.7})$$

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:

$$\mathcal{H} = \sum_{j=1}^J [\nabla F_j \sigma_j^{-1}] \otimes [\nabla F_j \sigma_j^{-1}]. \quad (\text{B.8})$$

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

$$\mathcal{H} = \mathcal{F} \mathcal{F}^T, \quad (\text{B.9})$$

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

$$\mathcal{F} = \begin{pmatrix} \nabla F_1 \sigma_1^{-1} & \nabla F_2 \sigma_2^{-1} & \cdots & \nabla F_J \sigma_J^{-1} \end{pmatrix}. \quad (\text{B.10})$$

We decompose \mathcal{F} and \mathcal{F}^T into their QR decomposition (see Theorem 2 in [212]):

$$\mathcal{F} = R^T Q^T, \quad (\text{B.11})$$

$$\mathcal{F}^T = QR,$$

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:

$$\begin{aligned}\nabla m \mathcal{H}^{-1} \nabla m &= \nabla m \left[R^T Q^T Q R \right]^{-1} \nabla m = \\ \nabla m R^{-1} (R^T)^{-1} \nabla m &= \|(R^T)^{-1} \nabla m\|^2.\end{aligned}\tag{B.12}$$

Therefore, under the standard approach Eq (B.6) is calculating the norm squared of the vector $(R^T)^{-1} \nabla m$ in \mathbb{R}^K . Let us now work with Eq (B.7) and obtain a similar structure but in \mathbb{R}^J .

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

$$\mathcal{F} \cdot e_j = \nabla F_j \sigma_j^{-1},\tag{B.13}$$

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 (B.7), we can replace $\nabla F_j \sigma_j^{-1}$ by $\mathcal{F} \cdot e_j$ and obtain:

$$\begin{aligned}\sum_{j=1}^J \left[\nabla m \mathcal{H}^{-1} \mathcal{F} \cdot e_j \right]^2 &= \\ \sum_{j=1}^J \left[\nabla m R^{-1} (R^T)^{-1} R^T Q^T \cdot e_j \right]^2 &= \\ \sum_{j=1}^J \left[(\nabla m R^{-1} Q^T) \cdot e_j \right]^2 &= \|(Q(R^T)^{-1} \nabla m)\|^2,\end{aligned}\tag{B.14}$$

where the last term is the norm of the vector $Q(R^T)^{-1} \nabla m$ calculated in \mathbb{R}^J . But, since the matrix 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 Eq (B.6) and Eq (B.7) are identical.

APPENDIX C

THE TF FORMALISM APPLIED TO WEAK CHARGE DENSITIES RECONSTRUCTIONS

C.1 Model descriptions

This appendix presents detailed information about the seven models we compared in Sec. 5.2. 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 Sec. 5.2.

Fourier Bessel

Although already explained in Chapter 6, for completeness and to provide more details we now discuss the Bessel expansion.

Under this formalism [128], the density is written as:

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

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

The form factor can be expressed analytically as:

$$F_{FB}(q) = \sum_{\nu=1}^N a_{\nu} G_{\nu}(q), \text{ where} \quad (\text{C.2})$$

$$G_{\nu}(q) \equiv 4\pi(-1)^{\nu} j_0(qR_{\text{cut}}) \frac{R_{\text{cut}}}{q^2 - q_{\nu}^2}. \quad (\text{C.3})$$

The normalization condition translates to:

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

The mean square radius is obtained as:

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

In Sec. 5.2, we use $R_{\text{cut}} = 7$ fm when analyzing ^{48}Ca , and $R_{\text{cut}} = 10$ fm when analyzing ^{208}Pb . We use a total of 5 adjustable a_{ν} for both nuclei which translates into six Bessels ($N = 6$) due to the normalization condition.

Helm Density

Although already explained in Sec. 5.1, for completeness and to provide more details we now discuss the Helm form factor.

The Helm density [171] is:

$$\begin{aligned} \rho_H(r) = & \frac{1}{2} \rho_{0H} \left[\operatorname{erf}\left(\frac{r + R_0}{\sqrt{2}\sigma}\right) - \operatorname{erf}\left(\frac{r - R_0}{\sqrt{2}\sigma}\right) \right] + \\ & \frac{1}{\sqrt{2\pi}} \left(\frac{\sigma}{r} \right) \rho_{0H} \left\{ \exp\left[-\frac{(r + R_0)^2}{2\sigma^2}\right] - \exp\left[-\frac{(r - R_0)^2}{2\sigma^2}\right] \right\}, \end{aligned} \quad (\text{C.6})$$

where:

$$\rho_{0H} \equiv \frac{3}{4\pi R_0^3}, \quad (\text{C.7})$$

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

$$\operatorname{erf}(x) \equiv \frac{2}{\sqrt{\pi}} \int_0^x e^{-z^2} dz. \quad (\text{C.8})$$

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.

The form factor is:

$$F_H(q) = 3 \frac{j_1(qR_0)}{qR_0} e^{-q^2 \sigma^2 / 2}, \quad (\text{C.9})$$

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

The radius is given by:

$$R^2 = \frac{3}{5} R_0^2 + 3\sigma^2. \quad (\text{C.10})$$

Symmetrized Fermi Function

Although already explained in Sec. 5.1, for completeness and to provide more details we now discuss the Symmetrized Fermi function.

The Symmetrized Fermi function [127] is constructed as: $f_{SF}(r) \equiv f_{SF}(r) + f_{SF}(-r) - 1$, where f_{SF} is the traditional Fermi Function [213]. Its density and form factor are expressed as:

$$\rho_{SF}(r) = \rho_{0SF} \frac{\sinh(c/a)}{\cosh(r/a) + \cosh(c/a)}, \quad (\text{C.11})$$

$$F_{SF}(q) = \frac{3}{qc[(qc)^2 + (\pi qa)^2]} \left[\frac{\pi qa}{\sinh(\pi qa)} \right] \left[\frac{\pi qa}{\tanh(\pi qa)} \sin(qc) - qc \cos(qc) \right], \quad (\text{C.12})$$

where the normalization constant is:

$$\rho_{0SF} = \frac{3}{4\pi c(c^2 + \pi^2 a^2)}, \quad (\text{C.13})$$

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

$$R^2 = \frac{3}{5}c^2 + \frac{7}{5}(\pi a)^2. \quad (\text{C.14})$$

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:

$$\rho_{SF3}(r) = \rho_{0SF3}(1 + wr^2) \frac{\sinh(c/a)}{\cosh(r/a) + \cosh(c/a)}, \quad (\text{C.15})$$

$$\rho_{SF4}(r) = \rho_{0SF4}(1 + wr^2 + ur^4) \frac{\sinh(c/a)}{\cosh(r/a) + \cosh(c/a)}. \quad (\text{C.16})$$

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 ρ_{0SF3} and ρ_{0SF4} , the form factors and the radius expressions 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 behaviour in the interior, so it cannot describe shell oscillations (small bumps and valleys around the saturation density background). As a different approach to adding parameters as in Eq. (C.15) and Eq. (C.16), we propose to use the following hybrid models.

SF+B:

$$\rho_{SF_B}(r) = \left(1 - \sum_{\nu=1}^{N_B} (-1)^{\nu+1} \frac{4\pi R_{\text{cut}}}{q_{\nu}^2} a_{\nu} \right) \rho_{SF}(r) + H(R_{\text{cut}} - r) \sum_{\nu=1}^{N_B} a_{\nu} j_0(q_{\nu} r), \quad (\text{C.17})$$

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}Ca and $R_{\text{cut}} = 5$ fm for ^{208}Pb . We use a total of three adjustable coefficients a_ν , 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:

$$\rho_{SFG}(r) = \left(1 - \sum_{i=1}^{N_G} A_i\right) \rho_{SF}(r) + \frac{1}{2\pi^{3/2}\gamma^3} \sum_{i=1}^{N_G} A_i g(r, R_i), \quad (\text{C.18})$$

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

$$g(r, R_i) = \frac{1}{1 + 2R_i^2/\gamma^2} (e^{-(r-R_i)^2/\gamma^2} + e^{-(r+R_i)^2/\gamma^2}). \quad (\text{C.19})$$

The amplitudes of the Gaussians A_i act as adjustable parameters. In our case, they are not restricted to be positive in contrast to [172]. 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 γ , which represents the common width of the Gaussians, is set to $\gamma = 1.4$ fm for both nuclei (see Sec 5.2.1 for a discussion of this value). The center of each Gaussian is denoted by R_i , and we chose the following values:

$$\mathbf{R} = [0, 1.3, 2.6] \text{ fm for } ^{48}\text{Ca}, \quad (\text{C.20})$$

$$\mathbf{R} = [0, 2, 4] \text{ fm for } ^{208}\text{Pb}.$$

The main idea behind these expansions is that the gross behaviour 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_{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 $\rho_{SFB}(r)$ will present a discontinuity at $r = R_{\text{cut}}$. This ‘‘kink’’, while nonphysical, might not preclude the entire model from describing nuclear densities.

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

$$F_{SFB}(q) = \left(1 - \sum_{\nu=1}^{N_B} (-1)^{\nu+1} \frac{4\pi R_{\text{cut}}}{q_\nu^2} a_\nu\right) F_{SF}(q) + \sum_{\nu=1}^{N_B} a_\nu G_\nu(q), \quad (\text{C.21})$$

$$F_{SFG}(q) = \left(1 - \sum_{i=1}^{N_G} A_i\right) F_{SF}(q) + \sum_{i=1}^{N_G} A_i \tilde{g}(q, R_i), \quad (\text{C.22})$$

where $G_\nu(q)$ is defined in Eq (C.3), and $\tilde{g}(q, R_i)$ is defined as:

$$\tilde{g}(q, R_i) = e^{-q^2 \gamma^2 / 4} \frac{1}{1 + 2R_i^2 / \gamma^2} \left(\cos(qR_i) + \frac{2R_i^2}{\gamma^2} j_0(qR_i) \right). \quad (\text{C.23})$$

The SF+B radius is given by:

$$R^2 = \left(1 - \sum_{\nu=1}^{N_B} (-1)^{\nu+1} \frac{4\pi R_{\text{cut}}}{q_\nu^2} a_\nu\right) \left(\frac{3}{5}c^2 + \frac{7}{5}(\pi a)^2\right) + 4\pi \sum_{\nu=1}^{N_B} a_\nu \frac{(-1)^\nu R_{\text{cut}}^5 (6 - \nu^2 \pi^2)}{\nu^4 \pi^4}. \quad (\text{C.24})$$

The SF+G radius is given by:

$$R^2 = \left(1 - \sum_{i=1}^{N_G} A_i\right) \left(\frac{3}{5}c^2 + \frac{7}{5}(\pi a)^2\right) + \sum_{i=1}^{N_G} A_i \frac{3\gamma^4 + 4R_i^4 + 12\gamma^2 R_i^2}{2(\gamma^2 + 2R_i^2)}. \quad (\text{C.25})$$

C.2 Details about the SF+G reconstruction bias analysis

This appendix provides a more detailed view at the reconstruction bias for the SF+G model discussed in Sec 5.2.1. 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:

$$\delta\rho(0) = \left[\mathcal{T}\mathcal{F}_4^{(\rho(0))} \right] \eta_{y4}. \quad (\text{C.26})$$

Fig. C.1 ((a.1) and (b.1)) shows the 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. C.1 (a.1), it can be observed that for most locations the calculated $|\delta\rho(0)|\%$ exceeds the 5% threshold (the dashed blue 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 5.2.2.

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

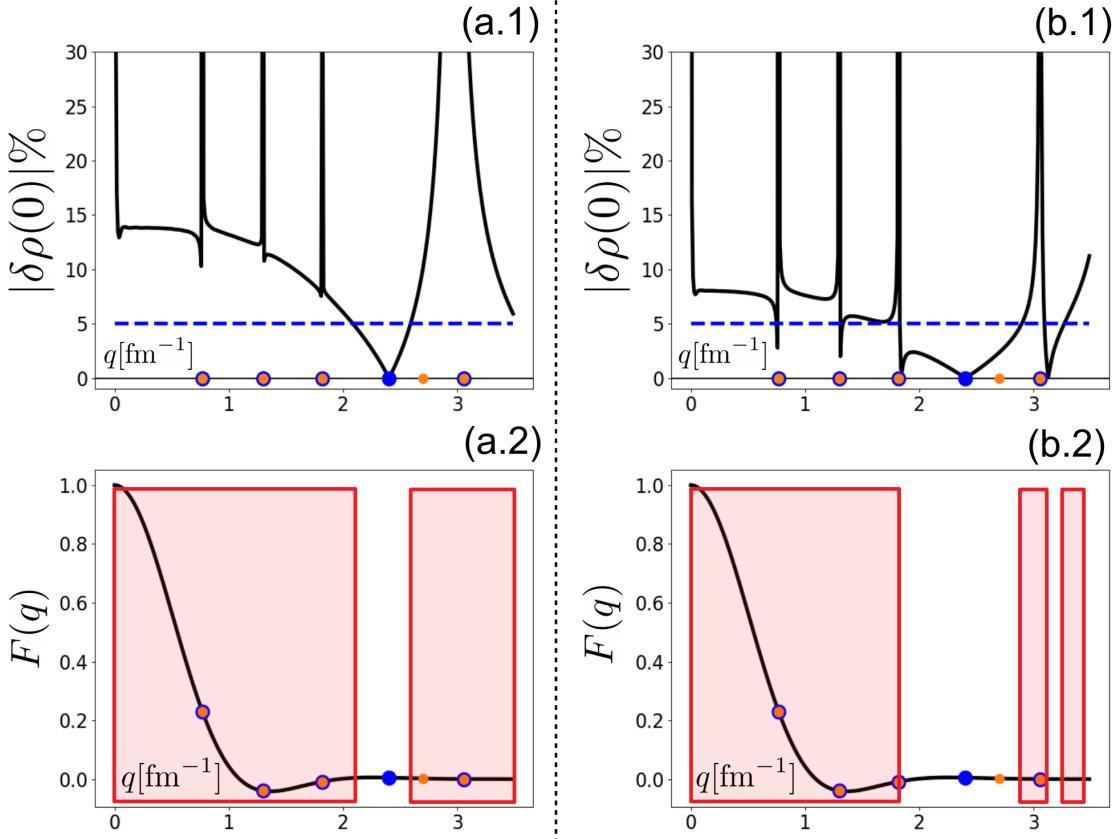


Figure C.1: Bias increase in $\rho(0)$ as a function of the location of the fourth measurement q_4 for $\gamma = 0.7 \text{ fm}$ (a), and $\gamma = 1.4 \text{ fm}$ (b). (a.1) and (b.1) show in black the calculated $\delta\rho(0)$ (as an absolute percentage) using Eq(C.26). The blue and orange points correspond to the two original data sets, while the dashed blue 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, $|\delta\rho(0)| > 5\%$.

observations, this results in a non-invertible Hessian for χ^2 , which blows up the transfer function $\mathcal{TF}_4^{\rho(0)}$. It is interesting to note that, even though there is not an observation located at $q = 0 \text{ fm}^{-1}$, a “spike” can still be observed. This is because by definition all model form factors must respect $F(0) = 1$. Therefore, a measure at $q = 0 \text{ fm}^{-1}$ provides no new information, resulting in a non-invertible Hessian.

Fig. C.1 ((a.2) and (b.2)) shows the true form factor (RMF012) in black and the observation locations as the blue and orange points, being (a.2) for $\gamma = 0.7 \text{ fm}$, and (b.2) for $\gamma = 1.4 \text{ fm}$. 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 5.2.1, locating q_4 at the blue point $q = 2.41 \text{ fm}^{-1}$ 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. 5.7, for the same $q_4 = 2.7 \text{ fm}^{-1}$, the SF+G with $\gamma = 0.7 \text{ fm}$ (dashed orange curve) fails to be within the λ_ρ band, while the SF+G with $\gamma = 1.4 \text{ fm}$ (dashed red line) stays within. In Fig. C.1 (a.2), the orange point is covered by the red rectangle. This is not the case for (b.2).

C.3 Details about the ^{48}Ca and ^{208}Pb charge examples and model comparison

This appendix presents tables and details relevant to the ^{48}Ca and ^{208}Pb charge density example developed in Sec 5.2.3. It includes the results of applying the same analysis on model comparison developed in Sec 5.2.4 to the electric charge densities of both nuclei instead of their weak densities counterparts.

Let us first consider the charge examples from Sec 5.2.3. Tables C.1 and C.2 show for the ^{48}Ca and ^{208}Pb examples, respectively, the original locations \mathbf{q}_0 as well as locations \mathbf{q}_m that minimize the FOM defined in Eq (5.26). Tables C.1 and C.2 also show the numerical values of the \mathcal{TF} times the respective error σ_j for the density at $r = 0 \text{ fm}$ 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 (5.25)), what goes in

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

Table C.1: ^{48}Ca electric form factor momentum transfer locations q_j and transfer functions $\mathcal{T}\mathcal{F}_j$ 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 5.4.

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

Table C.2: ^{208}Pb electric form factor momentum transfer locations q_j and transfer functions $\mathcal{T}\mathcal{F}_j$ 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 5.4.

the FOM Eq (5.26). 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 5.4.

By adding in quadrature each element in the rows of Tables C.1 and C.2 (the transfer functions values), the total variance in $\rho(0)$ or R can be calculated from Eq (3.6). 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}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.

Fig. C.2 and Fig. C.3 show, for ^{48}Ca and ^{208}Pb , respectively, the results of comparing the seven models specified in Appendix C.1 to recover the interior density and the radius. The numerical values shown in these figures are written in Tables C.5 and C.6. Tables C.3 and C.4 show the associated optimal locations \mathbf{q}_m for every model for ^{48}Ca and ^{208}Pb , respectively.

q_j [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

Table C.3: Optimal locations \mathbf{q}_m for each model when optimizing the Figure of Merit in Eq (5.26). These q_j values correspond to the solid bars results in Fig. C.2 for the charge density of ^{48}Ca . All values are in units of fm $^{-1}$.

q_j [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

Table C.4: Optimal locations \mathbf{q}_m for each model when optimizing the Figure of Merit in Eq (5.26). These q_j values correspond to the solid bars results in Fig. C.3 for the charge density of ^{208}Pb . All values are in units of fm $^{-1}$.

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 5.2.4, the starting locations are $\mathbf{q}_0 = [0.9, 1.35, 1.8, 2.24, 2.69]$ fm $^{-1}$ for ^{48}Ca and $\mathbf{q}_0 = [0.63, 0.94, 1.26, 1.57, 1.88]$ fm $^{-1}$ for ^{208}Pb .

We observe similar trends as the one shown in Figures 5.10 and 5.11 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 two by the SF in the ^{48}Ca radius. There are also significant variations in performance among the models, for example, a factor of three between the SF+B and SF+G when compared to the Helm model when extracting the interior density of ^{208}Pb for \mathbf{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 \mathbf{q}_m for the seven models in Tables C.3 and C.4. Let us recall that we limited the maximum value of any q_j to be less than 3.5

	q_0 Interior			Radius			q_m 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

Table C.5: Numerical values for the data displayed in Fig. C.2: model comparison for the electric charge density of ^{48}Ca .

	q_0 Interior			Radius			q_m 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

Table C.6: Numerical values for the data displayed in Fig. C.3: model comparison for the electric charge density of ^{208}Pb .

fm^{-1} for ^{48}Ca and less than 2 fm^{-1} for ^{208}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 this case, a more detailed analysis could be done to optimize a modified version of the FOM in which we took into account the experimental budget. It is very likely that the FOM is not extremely sensitive to the exact locations of q_m . Therefore, an adjustment of each q_j could result in a notorious reduction of the experimental budget while the FOM deteriorates just a small amount.

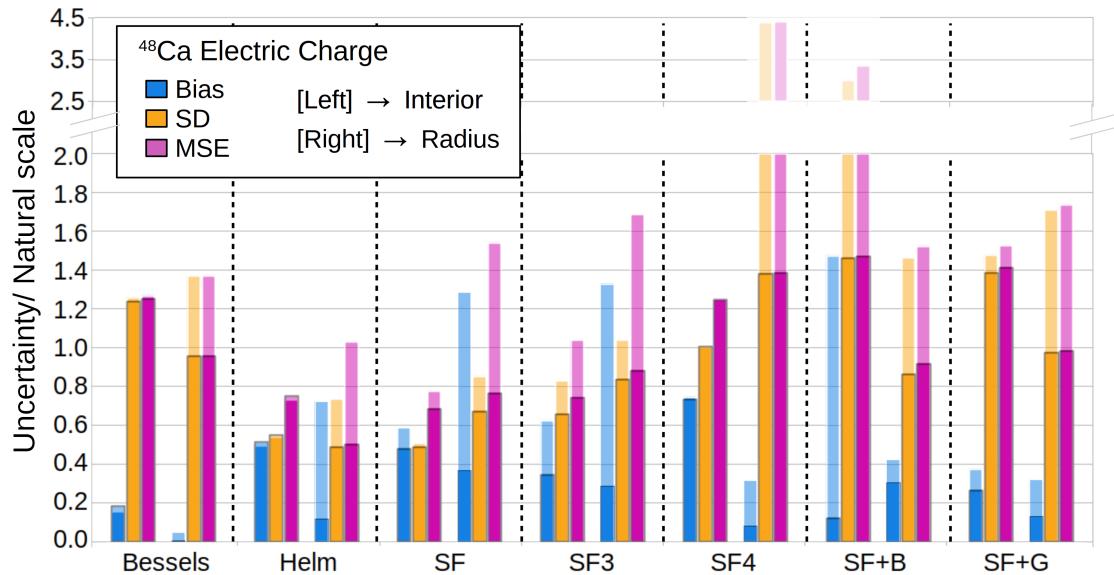


Figure C.2: Seven models comparison for recovering the interior density and mean charge radius from the charge form factor data on ^{48}Ca . The bias, SD, and MSE for each model are shown in their respective three left columns for the interior density, and in their respective three right columns for 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 \mathbf{q}_m , while the borderless ones represent the starting \mathbf{q}_0 .

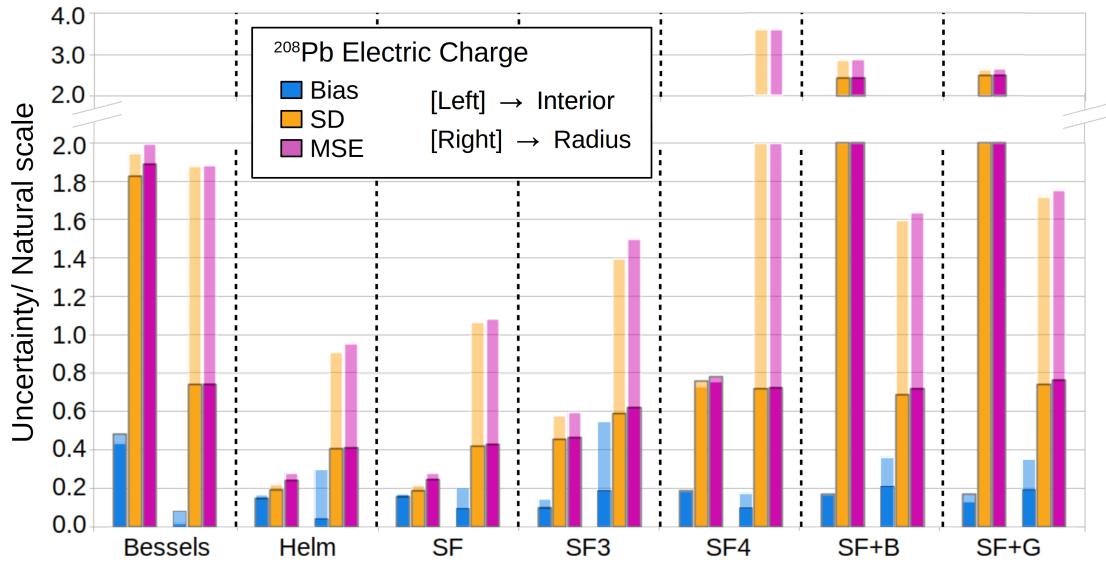


Figure C.3: Seven models comparison for recovering the interior density and mean charge radius from the charge form factor data on ^{208}Pb . The bias, SD, and MSE for each model are shown in their respective three left columns for the interior density, and in their respective three right columns for 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 \mathbf{q}_m , while the borderless ones represent the starting \mathbf{q}_0 .

C.4 Details about the ^{48}Ca and ^{208}Pb weak charge analysis

This appendix presents tables and details relevant to the ^{48}Ca and ^{208}Pb weak charge density model comparison developed in Sec 5.2.4. The numerical values associated with Figures 5.10 and 5.11 are shown in Tables C.9 and C.10, respectively. Tables C.7 and C.8 show the associated optimal locations \mathbf{q}_m for every model for ^{48}Ca and ^{208}Pb , respectively.

$q_j [\text{fm}^{-1}]$	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

Table C.7: Optimal locations \mathbf{q}_m for each model when optimizing the Figure of Merit in Eq (5.26). These q_j values correspond to the solid columns results in Fig. 5.10 for the five generated weak charge densities of ^{48}Ca . All values are in units of 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

Table C.8: Optimal locations \mathbf{q}_m for each model when optimizing the Figure of Merit in Eq (5.26). These q_j values correspond to the solid columns results in Fig. 5.11 for the five generated weak charge densities of ^{208}Pb . All values are in units of fm^{-1} .

As mentioned in Appendix C.3, it is interesting to analyze the distribution of these optimal locations \mathbf{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 seem 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 also anticipate that the FOM will not be extremely sensitive

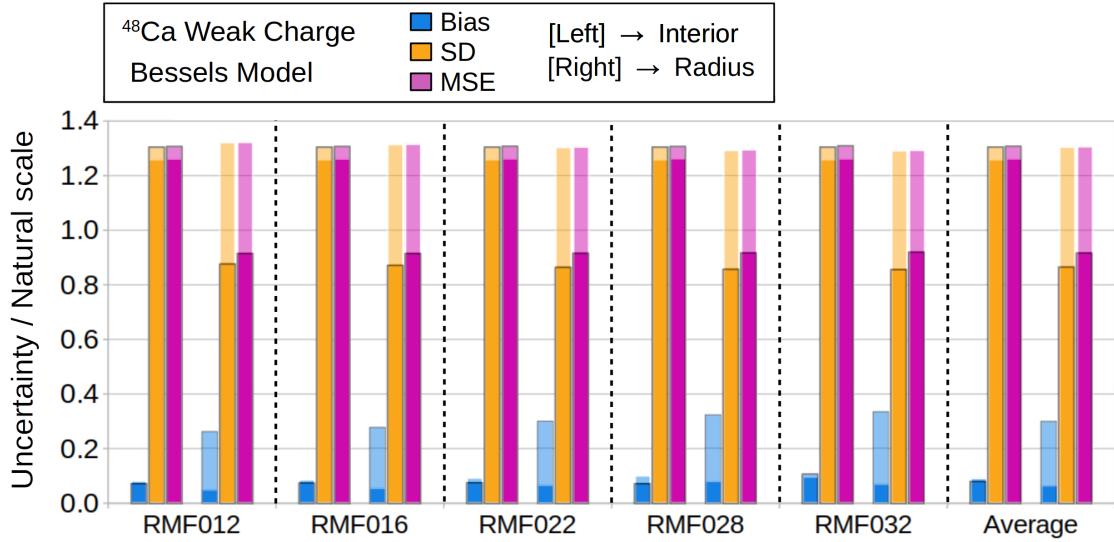


Figure C.4: Bessel model bias, standard deviation (SD), and MSE across the five different generators for the weak charge of ^{48}Ca . The last group of columns shows the square average of the five generators. Within each generator, the first three columns refer to the interior density while the second three columns refer to the radius. All quantities have been divided by their natural scales defined in Table 5.4.

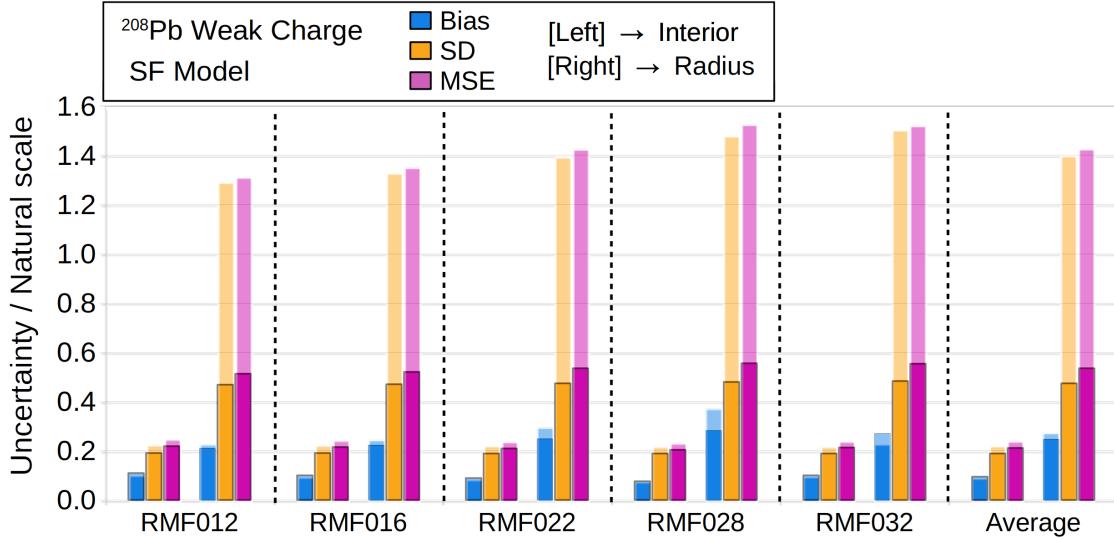


Figure C.5: SF model bias, standard deviation (SD) and MSE across the five different generators for the weak charge of ^{208}Pb . The last group of columns shows the square average of the five generators. Within each generator, the first three columns refer to the interior density while the second three columns refer to the radius. All quantities have been divided by their natural scales defined in Table 5.4.

	q_0 Interior			Radius			q_m 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

Table C.9: Numerical values for the data shown in Fig. 5.10: model comparison for the five generated weak charge densities of ^{48}Ca .

	q_0 Interior			Radius			q_m 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

Table C.10: Numerical values for the data shown in Fig. 5.11: model comparison for the five generated weak charge densities of ^{208}Pb .

to the exact locations of \mathbf{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}Pb , the Helm model seems to have a better FOM combination than the SF, but it could be that the \mathbf{q}_m associated with the SF model are more experimentally accessible than those locations associated with the Helm model.

Finally, Figures C.4 and C.5 show, as an example, the detailed results of the Bessels model for ^{48}Ca and the SF model for ^{208}Pb , respectively, across the five different generators used in Sec 5.2.4. The last columns in each figure show the squared average (see Eq 2.28) of each quantity (bias, SD, and MSE), which corresponds to the single values shown in Figures 5.10 and 5.11 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.

C.5 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 we discussed in Sec 5.2.5. 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 (5.28), and 3) P_2 : a very restrictive prior which practically behaves as the SF model without Gaussians.

	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

Table C.11: Absolute value transfer function values for the density at $r = 0$ fm ($|\mathcal{T}\mathcal{F}_j^{\rho(0)}\sigma_j|/\Delta\rho_{\text{Ca}}$) 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 $\Delta\rho_{\text{Ca}}$. Each $\mathcal{T}\mathcal{F}$ has been multiplied by their respective σ_j to somehow represent a fraction of the total standard deviation SD. The three different prior options are explored for the SF+G model.

Table C.11 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 (5.26), is calculated using the entire interior density. However, in this section we are focusing on $\rho(0)$ as a representative of the entire interior density. 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 C.11 in quadrature¹. In this sense, each number in the table reflects how much that particular data point uncertainty σ_j contributes to the total band that surrounds $\rho(0)$ in Fig. 5.12, in units of the natural scale $\Delta\rho_{\text{Ca}}$ from Table 5.4.

It can be observed that as the prior strength increases, the influence of each data point uncertainty σ_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 (2.25)).

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 (2.23a), we can see that each Gaussian prior has an associated center ω_k^0 , and an associated “error” σ_k which are analogous to the true observations

¹Let us recall that $r = 0$ fm is only one of the 30 grid points in Eq (5.25). The total interior MSE will therefore, receive contributions not only from $\rho(0)$.

y_j and true error σ_j of the data in Eq (2.16). Although $\tilde{\mathcal{H}}s$ 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, ω_k^0 could fluctuate around its center by as much as σ_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 since it is as if the parameters associated with the Gaussians were not there in the first place (also, numerically $\mathcal{T}\mathcal{F}_k^{\rho(0)}\sigma_k \rightarrow 0$ when $\sigma_k \rightarrow 0$). Table C.12 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 σ_k in terms of the natural scale $\Delta\rho_{\text{Ca}}$. 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 C.11.

	A_1	A_2	A_3
P_1	0.89	0.24	0.58

Table C.12: Absolute value of the prior transfer function values for the density at $r = 0$ 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 their respective σ_k (see Eq. (5.28)) to somehow represent a fraction of the total SD. The finite prior is the only one considered.

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 ω_{P_0} move to either ω_{P_1} or ω_{P_2} (the fitted parameters when using prior P_1 or P_2). If we are currently at ω_{P_0} and we add a prior term to χ^2 (converting it to $\tilde{\chi}^2$ defined in Eq (2.24)) in such a way that the centers ω_k^0 fall exactly at the value of their respective parameters $\omega_{P_0}\{k\}$, then the value of the 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 parameters is $\omega_{P_0}\{1\} = 0.5$ fm. Then, if we add to the total χ^2 a prior term of the form:

$$\frac{(\omega_1 - 0.5\text{fm})^2}{\sigma_1^2}, \quad (\text{C.27})$$

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 ω_k^0 (in the example 0.5 fm), and move it to the original location

of the prior we want to enforce (either ω_{P_1} or ω_{P_2}). This change which we call $\tilde{\eta}_k$ will now produce a change in the value of the parameters and therefore, in $\rho(0)$:

$$\delta\rho(0) = \mathcal{T}\mathcal{F}_k^{\rho(0)}\tilde{\eta}_{\omega_k}, \quad (\text{C.28})$$

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 C.13 shows in units of $\Delta\rho_{\text{Ca}}$, 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 $\Delta\rho_{\text{Ca}}$), we must add all numbers in each row. Due to the fact that in this example these numbers alternate sign, 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.

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

Table C.13: Prior transfer function values for the density at $r = 0$ fm ($\mathcal{T}\mathcal{F}_k^{\rho(0)}\eta_{\omega_k}/\Delta\rho_{\text{Ca}}$) for the three amplitudes in units of $\Delta\rho_{\text{Ca}}$. Each $\mathcal{T}\mathcal{F}$ has been multiplied by their respective $\tilde{\eta}_k$ to represent their fraction of the prior-induced bias.

Finally, note that Eq(C.28) 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 re-fitting the parameters with the new priors. Even though the numerical accuracy is not perfect, this type of analyses can help in better estimating the bias vs variance trade off when including priors.

BIBLIOGRAPHY

- [1] D. Geesaman et al. *Reaching for the Horizon; The 2015 Long Range Plan for Nuclear Science.* 2015.
- [2] C. F. von Weizsäcker. *Z. Physik*, 96:431, 1935.
- [3] H. A. Bethe and R. F. Bacher. Nuclear physics a. stationary states of nuclei. *Rev. Mod. Phys.*, 8(2):82, Apr 1936.
- [4] Peter J. Mohr, David B. Newell, and Barry N. Taylor. CODATA Recommended Values of the Fundamental Physical Constants: 2014. *Rev. Mod. Phys.*, 88(3):035009, 2016.
- [5] Robert Hofstadter. Electron scattering and nuclear structure. *Rev. Mod. Phys.*, 28:214–254, 1956.
- [6] H. De Vries, C. W. De Jager, and C. De Vries. Nuclear charge and magnetization density distribution parameters from elastic electron scattering. *Atom. Data Nucl. Data Tabl.*, 36:495–536, 1987.
- [7] G. Fricke, C. Bernhardt, K. Heilig, L. A. Schaller, L. Schellenberg, E. B. Shera, and C. W. de Jager. Nuclear ground state charge radii from electromagnetic interactions. *Atom. Data and Nucl. Data Tables*, 60:177, 1995.
- [8] I. Angeli and K.P. Marinova. Table of experimental nuclear ground state charge radii: An update. *At. Data Nucl. Data Tables*, 99:69 – 95, 2013.
- [9] Peter J Mohr, Barry N Taylor, and David B Newell. The 2010 codata recommended values of the fundamental physical constants. *Web version*, 6:96, 2011.
- [10] Randolph Pohl et al. The size of the proton. *Nature*, 466:213–216, 2010.
- [11] Randolph Pohl, Ronald Gilman, Gerald A. Miller, and Krzysztof Pachucki. Muonic hydrogen and the proton radius puzzle. *Ann. Rev. Nucl. Part. Sci.*, 63:175–204, 2013.
- [12] Jan C. Bernauer and Randolph Pohl. The proton radius problem. *Sci. Am.*, 310(2):18–25, 2014.
- [13] Carl E. Carlson. The Proton Radius Puzzle. *Prog. Part. Nucl. Phys.*, 82:59–77, 2015.
- [14] Peter J Mohr, David B Newell, and Barry N Taylor. Codata recommended values of the fundamental physical constants: 2014. *Journal of Physical and Chemical Reference Data*, 45(4):043102, 2016.

- [15] Axel Beyer, Lothar Maisenbacher, Arthur Matveev, Randolph Pohl, Ksenia Khabarova, Alexey Grinin, Tobias Lamour, Dylan C Yost, Theodor W Hänsch, Nikolai Kolachevsky, et al. The rydberg constant and proton size from atomic hydrogen. *Science*, 358(6359):79–85, 2017.
- [16] Axel Beyer et al. The rydberg constant and proton size from atomic hydrogen. *Science*, 358(6359):79–85, 2017.
- [17] Hélène Fleurbaey, Sandrine Galtier, Simon Thomas, Marie Bonnaud, Lucile Julien, François Biraben, François Nez, Michel Abgrall, and Jocelyne Guéna. New measurement of the 1 s- 3 s transition frequency of hydrogen: contribution to the proton charge radius puzzle. *Physical review letters*, 120(18):183001, 2018.
- [18] D. Robson. Solution to the proton radius puzzle. *Int. J. Mod. Phys.*, E23(12):1450090, 2015.
- [19] G. W. Bennett et al. Final Report of the Muon E821 Anomalous Magnetic Moment Measurement at BNL. *Phys. Rev.*, D73:072003, 2006.
- [20] R. Aaij et al. Test of lepton universality using $B^+ \rightarrow K^+\ell^+\ell^-$ decays. *Phys. Rev. Lett.*, 113:151601, Oct 2014.
- [21] W Xiong, A Gasparian, H Gao, D Dutta, M Khandaker, N Liyanage, E Pasyuk, C Peng, X Bai, L Ye, et al. A small proton charge radius from an electron–proton scattering experiment. *Nature*, 575(7781):147–150, 2019.
- [22] J. C. Bernauer et al. High-precision determination of the electric and magnetic form factors of the proton. *Phys. Rev. Lett.*, 105:242001, 2010.
- [23] R. Gilman et al. Studying the Proton "Radius" Puzzle with μ p Elastic Scattering. 2013.
- [24] Eite Tiesinga, Peter J Mohr, David B Newell, and Barry N Taylor. The 2018 codata recommended values of the fundamental physical constants. *Web Version*, 8, 2019.
- [25] Marko Horbatsch, Eric A. Hessels, and Antonio Pineda. Proton radius from electron-proton scattering and chiral perturbation theory. *Phys. Rev.*, C95(3):035203, 2017.
- [26] Douglas W. Higinbotham, Al Amin Kabir, Vincent Lin, David Meekins, Blaine Norum, and Brad Sawatzky. Proton radius from electron scattering data. *Phys. Rev.*, C93(5):055207, 2016.
- [27] Keith Griffioen, Carl Carlson, and Sarah Maddox. Consistency of electron scattering data with a small proton radius. *Phys. Rev.*, C93(6):065207, 2016.
- [28] Gabriel Lee, John R. Arrington, and Richard J. Hill. Extraction of the proton radius from electron-proton scattering data. *Phys. Rev.*, D92(1):013013, 2015.

- [29] John Arrington and Ingo Sick. Evaluation of the proton charge radius from e-p scattering. *J. Phys. Chem. Ref. Data*, 44:031204, 2015.
- [30] Krzysztof M. Graczyk and Cezary Juszczak. Proton radius from Bayesian inference. *Phys. Rev.*, C90:054334, 2014.
- [31] I. T. Lorenz, Ulf-G. Meißner, H. W. Hammer, and Y. B. Dong. Theoretical Constraints and Systematic Effects in the Determination of the Proton Form Factors. *Phys. Rev.*, D91(1):014023, 2015.
- [32] I. T. Lorenz and Ulf-G. Meißner. Reduction of the proton radius discrepancy by 3σ . *Phys. Lett.*, B737:57–59, 2014.
- [33] I. T. Lorenz, H. W. Hammer, and Ulf-G. Meissner. The size of the proton - closing in on the radius puzzle. *Eur. Phys. J.*, A48:151, 2012.
- [34] C. Adamuscin, S. Dubnicka, and A. Z. Dubnickova. New value of the proton charge root mean square radius. *Prog. Part. Nucl. Phys.*, 67:479–485, 2012.
- [35] Ingo Sick. Problems with proton radii. *Prog. Part. Nucl. Phys.*, 67:473–478, 2012.
- [36] X. Zhan et al. High-Precision Measurement of the Proton Elastic Form Factor Ratio $\mu_p G_E/G_M$ at low Q^2 . *Phys. Lett.*, B705:59–64, 2011.
- [37] G. Ron et al. Low Q^2 measurements of the proton form factor ratio $\mu u_p G_E/G_M$. *Phys. Rev.*, C84:055204, 2011.
- [38] Dmitry Borisuk. Proton charge and magnetic rms radii from the elastic ep scattering data. *Nucl. Phys.*, A843:59–67, 2010.
- [39] Richard J. Hill and Gil Paz. Model independent extraction of the proton charge radius from electron scattering. *Phys. Rev.*, D82:113005, 2010.
- [40] J. C. Bernauer et al. Electric and magnetic form factors of the proton. *Phys. Rev.*, C90(1):015206, 2014.
- [41] M. A. Belushkin, H. W. Hammer, and U. G. Meissner. Dispersion analysis of the nucleon form-factors including meson continua. *Phys. Rev.*, C75:035202, 2007.
- [42] Peter G. Blunden and Ingo Sick. Proton radii and two-photon exchange. *Phys. Rev.*, C72:057601, 2005.
- [43] J. J. Kelly. Simple parametrization of nucleon form factors. *Phys. Rev.*, C70:068202, 2004.
- [44] Ingo Sick. On the RMS radius of the proton. *Phys. Lett.*, B576:62–67, 2003.

- [45] Ivo M. Gough Eschrich et al. Measurement of the Sigma- charge radius by Sigma- electron elastic scattering. *Phys. Lett.*, B522:233–239, 2001.
- [46] R. Rosenfelder. Coulomb corrections to elastic electron proton scattering and the proton charge radius. *Phys. Lett.*, B479:381–386, 2000.
- [47] P. Mergell, Ulf G. Meissner, and D. Drechsel. Dispersion theoretical analysis of the nucleon electromagnetic form-factors. *Nucl. Phys.*, A596:367–396, 1996.
- [48] Chun Wa Wong. Deuteron radius and nuclear forces in free space. *Int. J. Mod. Phys.*, E3:821–908, 1994.
- [49] M. McCord, H. Crannell, L. W. Fagg, J. T. O’Brien, D. I. Sober, J. W. Lightbody, X. K. Maruyama, and P. A. Treado. Preliminary results of a new determination of the rms charge radius of the proton. *Nucl. Instrum. Meth.*, B56/57:496–499, 1991.
- [50] G.G. Simon, C. Schmitt, F. Borkowski, and V.H. Walther. Absolute electron Proton Cross-Sections at Low Momentum Transfer Measured with a High Pressure Gas Target System. *Nucl.Phys.*, A333:381–391, 1980.
- [51] G. Hohler, E. Pietarinen, I. Sabba Stefanescu, F. Borkowski, G. G. Simon, V. H. Walther, and R. D. Wendling. Analysis of Electromagnetic Nucleon Form-Factors. *Nucl. Phys.*, B114:505–534, 1976.
- [52] F. Borkowski, G. G. Simon, V. H. Walther, and R. D. Wendling. On the determination of the proton RMS-radius from electron scattering data. *Z. Phys.*, A275(1):29–31, 1975.
- [53] F. Borkowski, P. Peuser, G. G. Simon, V. H. Walther, and R. D. Wendling. Electromagnetic form-factors of the proton at low four-momentum transfer. *Nucl. Phys.*, A222:269–275, 1974.
- [54] F. Borkowski, P. Peuser, G. G. Simon, V. H. Walther, and R. D. Wendling. Electromagnetic Form-Factors of the Proton at Low Four-Momentum Transfer. *Nucl. Phys.*, B93:461–478, 1975.
- [55] Yu. K. Akimov et al. Small angle scattering of electrons by protons. *Sov. Phys. JETP*, 35:651–654, 1972. [Zh. Eksp. Teor. Fiz.62,1231(1972)].
- [56] J. J. Murphy, Y. M. Shin, and D. M. Skopik. Proton form factor from 0.15 to 0.79 fm-2. *Phys. Rev.*, C9:2125–2129, 1974. [Erratum: Phys. Rev. **C10**, 2111 (1974)].
- [57] Daniel Frerejacque, David Benaksas, and Darrell James Dickey. PROTON FORM-FACTORS FROM PROTON OBSERVATION. *Phys. Rev.*, 141:1308–1312, 1966.
- [58] L. N. Hand, D. G. Miller, and Richard Wilson. Electric and Magnetic Formfactor of the Nucleon. *Rev. Mod. Phys.*, 35:335, 1963.

- [59] Aldo Antognini et al. Proton Structure from the Measurement of $2S - 2P$ Transition Frequencies of Muonic Hydrogen. *Science*, 339:417–420, 2013.
- [60] M. Horbatsch and E. A. Hessels. Evaluation of the strength of electron-proton scattering data for determining the proton charge radius. *Phys. Rev.*, C93(1):015204, 2016.
- [61] Xuefei Yan, Douglas W. Higinbotham, Dipangkar Dutta, Haiyan Gao, Ashot Gasparian, Mahbub A. Khandaker, Nilanga Liyanage, Eugene Pasyuk, Chao Peng, and Weizhi Xiong. Robust extraction of the proton charge radius from electron-proton scattering data. *Phys. Rev. C*, 98(2):025204, 2018.
- [62] PRad Colaboration. Jefferson Science Associates, LLC. <http://www.jsallc.org/IF/Demo/Gasparian-App1.pdf>.
- [63] J. Arrington et al. A Proposal for the Paul Scherrer Institute π M1 beam line: Studying the Proton “Radius” Puzzle with μp Elastic Scattering. http://www.physics.rutgers.edu/~rgilman/elasticmup/muse_prop_2012.pdf.
- [64] Léo Neufcourt, Yuchen Cao, Witold Nazarewicz, Frederi Viens, et al. Bayesian approach to model-based extrapolation of nuclear observables. *Physical Review C*, 98(3):034318, 2018.
- [65] Raditya Utama, Wei-Chia Chen, and Jorge Piekarewicz. Nuclear charge radii: density functional theory meets bayesian neural networks. *Journal of Physics G: Nuclear and Particle Physics*, 43(11):114002, 2016.
- [66] Léo Neufcourt, Yuchen Cao, Samuel A Giuliani, Witold Nazarewicz, Erik Olsen, Oleg B Tarasov, et al. Quantified limits of the nuclear landscape. *Physical Review C*, 101(4):044307, 2020.
- [67] The BUQEYE collaboration. <https://buqeye.github.io/>.
- [68] C Drischler, JA Melendez, RJ Furnstahl, and DR Phillips. Quantifying uncertainties and correlations in the nuclear-matter equation of state. *arXiv preprint arXiv:2004.07805*, 2020.
- [69] The band collaboration. <https://bandframework.github.io>.
- [70] J. D. Walecka. *Electron scattering for nuclear and nucleon structure*. Cambridge University Press, New York, 2001.
- [71] D. Androic et al. First Determination of the Weak Charge of the Proton. *Phys. Rev. Lett.*, 111:141803, 2013.
- [72] D. Androic et al. Precision measurement of the weak charge of the proton. *Nature*, 557(7704):207–211, 2018.

- [73] M. Thiel, C. Sfienti, J. Piekarewicz, C. J. Horowitz, and M. Vanderhaeghen. Neutron skins of atomic nuclei: per aspera ad astra. *J. Phys.*, G46(9):093003, 2019.
- [74] C.J. Horowitz, J. Piekarewicz, and Brendan Reed. Measuring nuclear density with parity violating electron scattering. 7 2020.
- [75] Kate Scholberg. Prospects for measuring coherent neutrino-nucleus elastic scattering at a stopped-pion neutrino source. *Phys. Rev.*, D73:033005, 2006.
- [76] Junjie Yang, Jesse A. Hernandez, and J. Piekarewicz. Electroweak probes of ground state densities. *Phys. Rev. C*, 100(5):054301, 2019.
- [77] Daniel Z. Freedman. Coherent Neutrino Nucleus Scattering as a Probe of the Weak Neutral Current. *Phys. Rev.*, D9:1389–1392, 1974.
- [78] D. Akimov et al. Observation of Coherent Elastic Neutrino-Nucleus Scattering. *Science*, 357(6356):1123–1126, 2017.
- [79] D. Akimov et al. First Constraint on Coherent Elastic Neutrino-Nucleus Scattering in Argon. *Phys. Rev. D*, 100(11):115020, 2019.
- [80] T.W. Donnelly, J. Dubach, and Ingo Sick. ISOSPIN DEPENDENCES IN PARITY VIOLATING ELECTRON SCATTERING. *Nucl. Phys.*, A503:589, 1989.
- [81] S. Abrahamyan, Z. Ahmed, H. Albataineh, K. Aniol, D. S. Armstrong, et al. Measurement of the Neutron Radius of ^{208}Pb Through Parity-Violation in Electron Scattering. *Phys. Rev. Lett.*, 108:112502, 2012.
- [82] C. J. Horowitz, Z. Ahmed, C. M. Jen, A. Rakhman, P. A. Souder, et al. Weak charge form factor and radius of ^{208}Pb through parity violation in electron scattering. *Phys. Rev.*, C85:032501, 2012.
- [83] CREX: Parity-violating measurement of the weak charge distribution of ^{48}Ca . <http://hallaweb.jlab.org/parity/prex/c-rex/c-rex.pdf>.
- [84] C. J. Horowitz, K. S. Kumar, and R. Michaels. Electroweak Measurements of Neutron Densities in CREX and PREX at JLab, USA. *Eur. Phys. J.*, A50:48, 2014.
- [85] Dominik Becker et al. The P2 experiment. 2018.
- [86] B. Alex Brown. Neutron radii in nuclei and the neutron equation of state. *Phys. Rev. Lett.*, 85:5296, 2000.
- [87] R. J. Furnstahl. Neutron radii in mean-field models. *Nucl. Phys.*, A706:85–110, 2002.

- [88] M. Centelles, X. Roca-Maza, X. Viñas, and M. Warda. Nuclear symmetry energy probed by neutron skin thickness of nuclei. *Phys. Rev. Lett.*, 102:122502, 2009.
- [89] X. Roca-Maza, M. Centelles, X. Viñas, and M. Warda. Neutron skin of ^{208}Pb , nuclear symmetry energy, and the parity radius experiment. *Phys. Rev. Lett.*, 106:252501, 2011.
- [90] C. J. Horowitz and J. Piekarewicz. Neutron star structure and the neutron radius of ^{208}pb . *Phys. Rev. Lett.*, 86:5647, 2001.
- [91] C. J. Horowitz and J. Piekarewicz. The neutron radii of lead and neutron stars. *Phys. Rev.*, C64:062802, 2001.
- [92] J. Carriere, C. J. Horowitz, and J. Piekarewicz. Low mass neutron stars and the equation of state of dense matter. *Astrophys. J.*, 593:463, 2003.
- [93] Andrew W. Steiner, Madappa Prakash, James M. Lattimer, and Paul J. Ellis. Isospin asymmetry in nuclei and neutron stars. *Phys. Rept.*, 411:325, 2005.
- [94] J. Erler, C. J. Horowitz, W. Nazarewicz, M. Rafalski, and P.-G. Reinhard. Energy density functional for nuclei and neutron stars. *Phys. Rev.*, C87(4):044320, 2013.
- [95] Wei-Chia Chen and J. Piekarewicz. Building relativistic mean field models for finite nuclei and neutron stars. *Phys. Rev.*, C90:044305, 2014.
- [96] Wei-Chia Chen and J. Piekarewicz. Searching for isovector signatures in the neutron-rich oxygen and calcium isotopes. *Phys. Lett.*, B748:284, 2015.
- [97] Z. Lin and C. J. Horowitz. Full weak-charge density distribution of ^{48}Ca from parity-violating electron scattering. *Phys. Rev.*, C92(1):014313, 2015.
- [98] Karl Johan Åström and Richard M Murray. *Feedback systems: an introduction for scientists and engineers*. Princeton university press, 2010.
- [99] P. Gueye et al. Dispersive effects from a comparison of electron and positron scattering from C-12. *Phys. Rev.*, C57:2107–2110, 1998.
- [100] James Lewis Friar and M. Rosen. Dispersion corrections to elastic electron scattering by c-12 and o-16. 2. on the use of the closure approximation. *Annals Phys.*, 87:289–326, 1974.
- [101] E. A. J. M. Offermann, L. S. Cardman, C. W. de Jager, H. Miska, C. de Vries, and H. de Vries. Energy dependence of the form-factor for elastic electron scattering from C-12. *Phys. Rev.*, C44:1096–1117, 1991.
- [102] CJ Horowitz, Steven J Pollock, Paul A Souder, and Robert Michaels. Parity violating measurements of neutron densities. *Physical Review C*, 63(2):025501, 2001.

- [103] FJ Fattoyev, J Piekarewicz, and Charles J Horowitz. Neutron skins and neutron stars in the multimessenger era. *Physical Review Letters*, 120(17):172702, 2018.
- [104] Benjamin P Abbott, Rich Abbott, TD Abbott, Fausto Acernese, Kendall Ackley, Carl Adams, Thomas Adams, Paolo Addesso, RX Adhikari, VB Adya, et al. Gw170817: observation of gravitational waves from a binary neutron star inspiral. *Physical Review Letters*, 119(16):161101, 2017.
- [105] A Goldstein, P Veres, E Burns, MS Briggs, R Hamburg, D Kocevski, CA Wilson-Hodge, RD Preece, S Poolakkil, OJ Roberts, et al. An ordinary short gamma-ray burst with extraordinary implications: Fermi-gbm detection of grb 170817a. *The Astrophysical Journal Letters*, 848(2):L14, 2017.
- [106] V Savchenko, C Ferrigno, E Kuulkers, A Bazzano, E Bozzo, S Brandt, J Chenevez, TJ-L Courvoisier, R Diehl, A Domingo, et al. Integral detection of the first prompt gamma-ray signal coincident with the gravitational-wave event gw170817. *The Astrophysical Journal Letters*, 848(2):L15, 2017.
- [107] AB Balantekin, J Carlson, DJ Dean, GM Fuller, RJ Furnstahl, M Hjorth-Jensen, RVF Janssens, Bao-An Li, W Nazarewicz, FM Nunes, et al. Nuclear theory and science of the facility for rare isotope beams. *Modern Physics Letters A*, 29(11):1430010, 2014.
- [108] Gerald A Miller. Defining the proton radius: A unified treatment. *Physical Review C*, 99(3):035202, 2019.
- [109] Data files with the data from the Mainz collaboration [208, 209, 210] obtained from D. Higinbotham in a private communication.
- [110] MN Rosenbluth. High energy elastic scattering of electrons on protons. *Physical Review*, 79(4):615, 1950.
- [111] Jefferson Lab Qweak Collaboration et al. Precision measurement of the weak charge of the proton. *Nature*, 557(7704):207, 2018.
- [112] B. Dreher, J. Friedrich, K. Merle, H. Rothhaas, and G. Lührs. The determination of the nuclear ground state and transition charge density from measured electron scattering data. *Nuclear Physics A*, 235(1):219 – 248, 1974.
- [113] Wei-Chia Chen and J Piekarewicz. Building relativistic mean field models for finite nuclei and neutron stars. *Physical Review C*, 90(4):044305, 2014.
- [114] Xavier Roca-Maza, M Centelles, F Salvat, and X Vinas. Theoretical study of elastic electron scattering off stable and exotic nuclei. *Physical Review C*, 78(4):044332, 2008.

- [115] CODATA18. Committee on data international science council, 2018 codata recommended values, 2018.
- [116] P.-G. Reinhard and W. Nazarewicz. Information content of a new observable: The case of the nuclear neutron skin. *Phys. Rev.*, C81:051303(R), 2010.
- [117] P. G. Reinhard, J. Piekarewicz, W. Nazarewicz, B.K. Agrawal, N. Paar, et al. Information content of the weak-charge form factor. *Phys.Rev.*, C88:034325, 2013.
- [118] W. Nazarewicz, P. G. Reinhard, W. Satula, and D. Vretenar. Symmetry energy in nuclear density functional theory. *Eur. Phys. J.*, A50:20, 2014.
- [119] M. B. Tsang et al. Isospin Diffusion and the Nuclear Symmetry Energy in Heavy Ion Reactions. *Phys. Rev. Lett.*, 92:062701, 2004.
- [120] Lie-Wen Chen, Che Ming Ko, and Bao-An Li. Studying stiffness of nuclear symmetry energy through isospin diffusion in heavy-ion collisions. *Phys. Rev. Lett.*, 94:032701, 2005.
- [121] Andrew W. Steiner and Bao-An Li. Isospin diffusion in heavy-ion collisions and the neutron skin thickness of lead. *Phys. Rev.*, C72:041601, 2005.
- [122] D. V. Shetty, S. J. Yennello, and G. A. Souliotis. Density dependence of the symmetry energy and the nuclear equation of state: A Dynamical and Statistical model perspective. *Phys. Rev.*, C76:024606, 2007.
- [123] M. B. Tsang et al. Constraints on the density dependence of the symmetry energy. *Phys. Rev. Lett.*, 102:122701, 2009.
- [124] Bao-An Li, Lie-Wen Chen, and Che Ming Ko. Recent Progress and New Challenges in Isospin Physics with Heavy-Ion Reactions. *Phys. Rept.*, 464:113–281, 2008.
- [125] Bao-An Li and Andrew W. Steiner. Constraining the radii of neutron stars with terrestrial nuclear laboratory data. *Phys. Lett.*, B642:436–440, 2006.
- [126] James M. Lattimer and Maddapa Prakash. Neutron Star Observations: Prognosis for Equation of State Constraints. *Phys. Rept.*, 442:109, 2007.
- [127] DWL Sprung and J Martorell. The symmetrized fermi function and its transforms. *Journal of Physics A: Mathematical and General*, 30(18):6525, 1997.
- [128] B Dreher, J Friedrich, K Merle, H Rothhaas, and G Lührs. The determination of the nuclear ground state and transition charge density from measured electron scattering data. *Nuclear Physics A*, 235(1):219–248, 1974.

- [129] F Borkowski, GG Simon, VH Walther, and RD Wendling. On the determination of the proton rms-radius from electron scattering data. *Zeitschrift für Physik A Atoms and Nuclei*, 275(1):29–31, 1975.
- [130] Philip R Bevington and D Keith Robinson. *Data reduction and error analysis for the physical sciences; 3rd ed.* McGraw-Hill, New York, NY, 2003.
- [131] P. C. Gregory. *Bayesian Logical Data Analysis for the Physical Sciences*. Cambridge University Press, Cambridge, UK, 2005.
- [132] James V. Stone. *Bayes' Rule: A Tutorial Introduction to Bayesian Analysis*. Sebtel Press, Sheffield, UK, 2013.
- [133] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media, 2009.
- [134] Timothy John Sullivan. *Introduction to uncertainty quantification*, volume 63. Springer, 2015.
- [135] Peter J Huber. *Robust statistics*, volume 523. John Wiley & Sons, 2004.
- [136] Dan G Cacuci. *Sensitivity and uncertainty analysis, volume I: Theory*, volume 1. CRC press, 2003.
- [137] Aad W Van der Vaart. *Asymptotic statistics*. Cambridge university press, 2000.
- [138] Jan C. Bernauer and Michael O. Distler. Avoiding common pitfalls and misconceptions in extractions of the proton radius. In *ECT* Workshop on The Proton Radius Puzzle Trento, Italy, June 20-24, 2016*, 2016.
- [139] Richard J Hill and Gil Paz. Model-independent extraction of the proton charge radius from electron scattering. *Physical Review D*, 82(11):113005, 2010.
- [140] Gabriel Lee, John R Arrington, and Richard J Hill. Extraction of the proton radius from electron-proton scattering data. *Physical Review D*, 92(1):013013, 2015.
- [141] Ingo Sick and Dirk Trautmann. Proton rms-radii from low-q power expansions? *Phys. Rev.*, C95(1):012501, 2017.
- [142] Ingo Sick. Proton charge radius from electron scattering. *Atoms*, 6, 2018.
- [143] Galit Shmueli. To explain or to predict? *Statist. Sci.*, 25(3):289–310, 08 2010.
- [144] George E. P. Box. Science and statistics. *Journal of the American Statistical Association*, 71(356):791–799, 1976.

- [145] Trevor Hastie, Robert Tibshirani, and Jerome H. Friedman. *The elements of statistical learning: data mining, inference, and prediction, 2nd Edition*. Springer series in statistics. Springer, 2009.
- [146] J. M. Alarcón and C. Weiss. Nucleon form factors in dispersively improved Chiral Effective Field Theory II: Electromagnetic form factors. *Phys. Rev.*, C97(5):055203, 2018.
- [147] Daniel F Styer. The geometrical significance of the laplacian. *American Journal of Physics*, 83(12):992–997, 2015.
- [148] Wei-Chia Chen and J Piekarewicz. Analytic insights on the information content of new observables. *Physical Review C*, 102(4):042801, 2020.
- [149] JM Alarcón and C Weiss. Accurate nucleon electromagnetic form factors from dispersively improved chiral effective field theory. *arXiv preprint arXiv:1803.09748*, 2018.
- [150] JM Alarcón and C Weiss. Nucleon form factors in dispersively improved chiral effective field theory. ii. electromagnetic form factors. *Physical Review C*, 97(5):055203, 2018.
- [151] Hassan Maatouk and Xavier Bay. Gaussian process emulators for computer experiments with inequality constraints. *Mathematical Geosciences*, 49(5):557–582, 2017.
- [152] Andrew Gelman, John B Carlin, Hal S Stern, David B Dunson, Aki Vehtari, and Donald B Rubin. *Bayesian data analysis*. CRC press, 2013.
- [153] C.E. Rasmussen and C.K.I. Williams. *Gaussian Processes for Machine Learning*. Adaptative computation and machine learning series. University Press Group Limited, 2006.
- [154] Robert J Adler. *The geometry of random fields*, volume 62. Siam, 1981.
- [155] Harold Jeffreys. An invariant form for the prior probability in estimation problems. *Proc. R. Soc. Lond. A*, 186(1007):453–461, 1946.
- [156] ZI Botev. The normal law under linear restrictions: simulation and estimation via minimax tilting. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79(1):125–148, 2017.
- [157] Ashot H. Gasparian. The New Proton Radius Experiment at Jefferson Lab. *JPS Conf. Proc.*, 13:020052, 2017.
- [158] Junjie Yang, Jesse A Hernandez, and J Piekarewicz. Electroweak probes of ground state densities. *Physical Review C*, 100(5):054301, 2019.
- [159] R. D. Amado. ANALYTIC INSIGHTS INTO INTERMEDIATE-ENERGY HADRON NUCLEUS SCATTERING. *Adv. Nucl. Phys.*, 15:1–42, 1985.

- [160] D. W. Sprung and J. Martorell. The symmetrized Fermi function and its transforms. *J. Phys. A*, 30:6525–6534, 1997.
- [161] CL Jiang, AM Stefanini, H Esbensen, KE Rehm, L Corradi, E Fioretto, P Mason, G Montagnoli, F Scarlassara, R Silvestri, et al. Fusion hindrance for ca+ ca systems: Influence of neutron excess. *Physical Review C*, 82(4):041601, 2010.
- [162] Richard H. Helm. Inelastic and Elastic Scattering of 187-Mev Electrons from Selected Even-Even Nuclei. *Phys. Rev.*, 104:1466–1475, 1956.
- [163] A. B. Balantekin et al. Nuclear Theory and Science of the Facility for Rare Isotope Beams. *Mod. Phys. Lett.*, A29(11):1430010, 2014.
- [164] Roger D. Woods and David S. Saxon. Diffuse Surface Optical Model for Nucleon-Nuclei Scattering. *Phys. Rev.*, 95:577–578, 1954.
- [165] R. D. Amado, J. P. Dedonder, and F. Lenz. An Explicit Formula for Hadron - Nucleus Elastic Scattering in the Eikonal Approximation. *Phys. Rev.*, C21:647–661, 1980.
- [166] S. Mizutori, J. Dobaczewski, G. A. Lalazissis, W. Nazarewicz, and P. G. Reinhard. Nuclear skins and halos in the mean field theory. *Phys. Rev.*, C61:044326, 2000.
- [167] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21:1087–1092, 1953.
- [168] F. J. Fattoyev, C. J. Horowitz, J. Piekarewicz, and G. Shen. Relativistic effective interaction for nuclei, giant resonances, and neutron stars. *Phys. Rev.*, C82:055803, 2010.
- [169] J. Piekarewicz, B.K. Agrawal, G. Colò, W. Nazarewicz, N. Paar, et al. Electric dipole polarizability and the neutron skin. *Phys. Rev.*, C85:041302(R), 2012.
- [170] H De Vries, CW De Jager, and C De Vries. Nuclear charge-density-distribution parameters from elastic electron scattering. *Atomic data and nuclear data tables*, 36(3):495–536, 1987.
- [171] Richard H Helm. Inelastic and elastic scattering of 187-mev electrons from selected even-even nuclei. *Physical Review*, 104(5):1466, 1956.
- [172] Ingo Sick. Model-independent nuclear charge densities from elastic electron scattering. *Nuclear Physics A*, 218(3):509–541, 1974.
- [173] Christopher M Bishop. *Pattern recognition and machine learning*. Springer, 2006.
- [174] L. S. Cardman, J. W. Lightbody, S. Penner, S. P. Fivozinsky, X. K. Maruyama, W. P. Trower, and S. E. Williamson. The charge distribution of 12 C. *Phys. Lett.*, 91B:203–206, 1980.

- [175] E.A.J.M. Offermann, L.S. Cardman, H.J. Emrich, G. Fricke, C.W. deJager, H. Miska, D. Rychel, and H. deVries. Search for Dispersive Effects in Elastic Electron Scattering From ^{12}C . *Phys. Rev. Lett.*, 57:1546–1549, 1986.
- [176] V. Breton et al. High accuracy comparison of electron and positron scattering from nuclei. *Phys. Rev. Lett.*, 66:572–575, 1991.
- [177] P. Gueye et al. Coulomb distortion measurements by comparing electron and positron quasielastic scattering off C-12 and Pb-208. *Phys. Rev.*, C60:044308, 1999.
- [178] I. Sick and J. S. McCarthy. Elastic electron scattering from c-12 and o-16. *Nucl. Phys.*, A150:631–654, 1970.
- [179] Pierre A. M. Guichon and M. Vanderhaeghen. How to reconcile the Rosenbluth and the polarization transfer method in the measurement of the proton form-factors. *Phys. Rev. Lett.*, 91:142303, 2003.
- [180] P. G. Blunden, W. Melnitchouk, and J. A. Tjon. Two photon exchange and elastic electron proton scattering. *Phys. Rev. Lett.*, 91:142304, 2003.
- [181] Michail P. Rekalo and Egle Tomasi-Gustafsson. Polarization observables in the processes $p + p \rightarrow \zeta \Theta^+ + \Sigma^+$ and $n + p \rightarrow \zeta \Theta^+ + \Lambda^0$, for any spin and parity of the Θ^+ hyperon in the threshold region. *Eur. Phys. J.*, A22:119–124, 2004.
- [182] Y. C. Chen, A. Afanasev, S. J. Brodsky, C. E. Carlson, and M. Vanderhaeghen. Partonic calculation of the two photon exchange contribution to elastic electron proton scattering at large momentum transfer. *Phys. Rev. Lett.*, 93:122301, 2004.
- [183] Andrei V. Afanasev and N. P. Merenkov. Large logarithms in the beam normal spin asymmetry of elastic electron-proton scattering. *Phys. Rev.*, D70:073002, 2004.
- [184] J. Arrington. Extraction of two-photon contributions to the proton form-factors. *Phys. Rev.*, C71:015202, 2005.
- [185] P. G. Blunden, W. Melnitchouk, and J. A. Tjon. Two-photon exchange in elastic electron-nucleon scattering. *Phys. Rev.*, C72:034612, 2005.
- [186] P. G. Blunden and W. Melnitchouk. Dispersive approach to two-photon exchange in elastic electron-proton scattering. *Phys. Rev.*, C95(6):065209, 2017.
- [187] D. Adikaram et al. Towards a resolution of the proton form factor problem: new electron and positron scattering data. *Phys. Rev. Lett.*, 114:062003, 2015.
- [188] I. A. Rachek et al. Measurement of the two-photon exchange contribution to the elastic $e^\pm p$ scattering cross sections at the VEPP-3 storage ring. *Phys. Rev. Lett.*, 114(6):062005, 2015.

- [189] B. S. Henderson et al. Hard Two-Photon Contribution to Elastic Lepton-Proton Scattering: Determined by the OLYMPUS Experiment. *Phys. Rev. Lett.*, 118(9):092501, 2017.
- [190] D. Rimal et al. Measurement of two-photon exchange effect by comparing elastic $e^\pm p$ cross sections. *Phys. Rev.*, C95(6):065201, 2017.
- [191] Carl E. Carlson and Marc Vanderhaeghen. Two-Photon Physics in Hadronic Processes. *Ann. Rev. Nucl. Part. Sci.*, 57:171–204, 2007.
- [192] J. Arrington, P. G. Blunden, and W. Melnitchouk. Review of two-photon exchange in electron scattering. *Prog. Part. Nucl. Phys.*, 66:782–833, 2011.
- [193] A. Afanasev, P. G. Blunden, D. Hasell, and B. A. Raue. Two-photon exchange in elastic electron–proton scattering. *Prog. Part. Nucl. Phys.*, 95:245–278, 2017.
- [194] Andreas Aste, Cyrill von Arx, and Dirk Trautmann. Coulomb distortion of relativistic electrons in the nuclear electrostatic field. *Eur. Phys. J.*, A26:167–178, 2005.
- [195] R. Gilman, D. W. Higinbotham, X. Jiang, A. Sarty, and S. Strauch. LEDEX: Low Energy Deutron EXperiments E05-044 and E05-103 at Jefferson Lab, 2015.
- [196] Marco Traini. Coulomb distortion in quasielastic (e, e-prime) scattering on nuclei: Effective momentum approximation and beyond. *Nucl. Phys.*, A694:325–336, 2001.
- [197] EAJM Offermann, LS Cardman, HJ Emrich, G Fricke, CW de Jager, H Miska, D Rychel, and H de Vries. Search for dispersive effects in elastic electron scattering from c 12. *Physical review letters*, 57(13):1546, 1986.
- [198] Bogdan Povh, Klaus Rith, Christoph Scholz, and Frank Zetsche. Particles and nuclei. *An Introduction to the Physical Concepts, Berlin and Heidelberg: Springer-Verlag (Italian Translation:(1998), Particelle e nuclei. Un'introduzione ai concetti sici, Torino: Bollati Boringhieri editore)*, 1995.
- [199] S. Abrahamyan et al. New Measurements of the Transverse Beam Asymmetry for Elastic Electron Scattering from Selected Nuclei. *Phys. Rev. Lett.*, 109:192501, 2012.
- [200] Oleksandr Tomalak, Barbara Pasquini, and Marc Vanderhaeghen. Two-photon exchange contribution to elastic e–proton scattering: Full dispersive treatment of $\pi^- n$ states and comparison with data. *Physical Review D*, 96(9):096001, 2017.
- [201] Douglas W Higinbotham, Vincent Lin, David Meekins, Blaine Norum, Brad Sawatzky, et al. Proton radius from electron scattering data. *Physical Review C*, 93(5):055207, 2016.
- [202] Editors. Uncertainty estimates. *Phys Rev A*, 83(4):040001, 2011.

- [203] D. G. Ireland and W. Nazarewicz. Enhancing the interaction between nuclear experiment and theory through information and statistics. *J. Phys.*, G42(3):030301, 2015.
- [204] Charles J Horowitz and Jorge Piekarewicz. Neutron radii of 208 pb and neutron stars. *Physical Review C*, 64(6):062802, 2001.
- [205] FRIB Scientific and Technical Merit. https://frib.msu.edu/_files/pdfs/frib_scientific_and_technical_merit_lite_0.pdf.
- [206] LIGO Scientific Collaboration et al. Gravitational wave astronomy with ligo and similar detectors in the next decade. *arXiv preprint arXiv:1904.03187*, 2019.
- [207] Yun Yang, Anirban Bhattacharya, and Debdeep Pati. Frequentist coverage and sup-norm convergence rate in gaussian process regression. *arXiv preprint arXiv:1708.04753*, 2017.
- [208] Jan C Bernauer and A1 Collaboration. High-precision determination of the electric and magnetic form factors of the proton. In *AIP Conference Proceedings*, volume 1388, pages 128–134. AIP, 2011.
- [209] JC Bernauer, MO Distler, J Friedrich, Th Walcher, P Achenbach, C Ayerbe Gayoso, R Böhm, Damir Bosnar, L Debenjak, L Doria, et al. Electric and magnetic form factors of the proton. *Physical Review C*, 90(1):015206, 2014.
- [210] JC Bernauer, P Achenbach, C Ayerbe Gayoso, R Böhm, Damir Bosnar, L Debenjak, MO Distler, L Doria, A Esser, H Fonvieille, et al. Bernauer et al. reply. *Physical Review Letters*, 107(11):119102, 2011.
- [211] Michael Spivak. *Calculus on manifolds: a modern approach to classical theorems of advanced calculus*. CRC press, 2018.
- [212] Colin R Goodall. 13 computation using the qr decomposition. 1993.
- [213] Roger D Woods and David S Saxon. Diffuse surface optical model for nucleon-nuclei scattering. *Physical Review*, 95(2):577, 1954.

BIOGRAPHICAL SKETCH

Pablo Giuliani

EDUCATION

Florida State University, Tallahassee, FL

PhD. in Physics, 2015-2020

Advisor: Jorge Piekarewicz

Universidad Simon Bolivar, Caracas, Venezuela

BcS. in Physics, 2009-2016

Advisor: Claudia Colonnello Olivares

PUBLICATIONS

- Power of two: Assessing the impact of a second measurement of the weak-charge form factor of ^{208}Pb . Piekarewicz, J., Linero, A. R., Giuliani, P., & Chicken, E. Physical Review C, 94(3), 034316. (2016).
- Reexamining the proton-radius problem using constrained Gaussian processes. Zhou, S., Giuliani, P., Piekarewicz, J., Bhattacharya, A., & Pati, D. Phys. Review C 99, 055202. (2019).
- Bias-Variance Trade-off and Model Selection for Proton Radius Extractions. Douglas W. Higinbotham, Pablo Giuliani, Randall E. McClellan, Simon Sirca, and Xuefei Yan. arXiv preprint: arXiv:1812.05706. (2018)
- Dispersive corrections in elastic electron-nucleus scattering: an investigation in the intermediate energy regime and their impact on the nuclear matter. The Jefferson Lab Hall A Collaboration. European Physical Journal. A, 56(JLAB-PHY-18-2707; DOE/OR/23177-4450; arXiv: 1805.12441). (2020).
- Electromechanic modelling for Johnson Noise in Advanced LIGO. Edgard Luis Bonilla Carrasquel and Pablo Giuliani and Brian Lantz and Aaron Buikema. Classical and Quantum Gravity. (2020)
- From noise to information: The transfer function formalism for uncertainty quantification on nuclear densities reconstruction. Pablo Giuliani and Piekarewicz, J. In preparation.

CONFERENCES AND SCHOOLS ATTENDED

FRIB summer school 2020 “Dense matter in astrophysics”

FRIB summer school 2019 “Machine learning applied to nuclear physics”

FRIB summer school 2018 “Neutron star mergers”

Gordon Research Conference on Photonuclear Reactions 2018:

- Contributed Poster: *Revisiting the proton-radius problem using constrained Gaussian processes*