Application of information geometry methods in the development of nuclear structure models

Imbrišak, Marko

Doctoral thesis / Doktorski rad

2023

University of Zagreb
Faculty of Science
Department of Physics

Marko Imbrišak

Application of information geometry methods in the development of nuclear structure models

DOCTORAL THESIS

Zagreb, 2023

University of Zagreb
Faculty of Science
Department of Physics

Marko Imbrišak

Application of information geometry methods in the development of nuclear structure models

DOCTORAL THESIS

Supervisor:
dr. sc. Kosuke Nomura

Zagreb, 2023

Sveučilište u Zagrebu
Prirodoslovno-matematički fakultet
Fizički odsjek

Marko Imbrišak

Primjena metoda informacijske geometrije u razvoju modela nuklearne strukture

DOKTORSKI RAD

Mentor:
dr. sc. Kosuke Nomura

Zagreb, 2023

Supervisor information

This thesis was made under the supervision of doc. dr. sc. Kosuke Nomura during the PHD program at the Department of Physics, Faculty of Science, University of Zagreb.

Kosuke Nomura is an associate Professor at the Department of Physics, Faculty of Science, Hokkaido University in Japan. He has been doing research in the field of theoretical nuclear physics, specifically on the low-energy structure and dynamics of the atomic nucleus and related quantum many-body problems. He attained his Masters degree in physics from the University of Tokyo in 2009 and his Ph.D. in physics at the University of Tokyo in 2012. He was a JSPS Fellow at the University of Tokyo (04/2009-03/2012), a JSPS Fellow for Research Abroad at the University of Cologne (04/2012-08/2013), a Marie-Curie Fellow (GANIL, CEA/CNRS, France, 09/2013-08/2015), a postdoc at IPN Orsay (09/2015), a postdoc at the University of Zagreb (10/2015-03/2016 and 01/2018-11/2018), a JSPS Fellow at the University of Tsukuba (04/2016-10/2017), a Tenure-Track Researcher at the Japan Atomic Energy Agency (12/2018-04/2019). He was an assistant Professor and Head of the Laboratory at the University of Zagreb (05/2019-02/2023) moving to the position of an associate professor at the Hokkaido (04/2023-Now).

Curriculum vitae

Marko Imbrišak was born in 1993 in Zagreb (Croatia). He obtained his master's degree in Physics at the Department of Physics, Faculty of Science at the University of Zagreb in 2019, with the thesis titled 'Optimization problem of sloppy nuclear structure models', under the supervision of prof. dr. sc. Tamara Nikšić. In the fall of 2019, he started working in the group led by doc. dr. sc. Kosuke Nomura at the Department of Physics, Faculty of Science at the University of Zagreb under the Tenure Track Pilot Programme titled 'Exotic Nuclear Structure and Dynamics' (TTP-2018-07-3554).

Submitted publications

1. **Imbrišak, M.**; Nomura, K.: *Classical and Bayesian error analysis of the relativistic mean-field model for doubly-magic nuclei* // submitted to Physical Review C

Refereed publications

1. **Imbrišak, M.**; Nomura, K.: *Stability of the manifold boundary approximation method for reductions of nuclear structure models* // Physical Review C, 107 (2023), 3; 034304, 11

2. Tisanić, K.; Smolčić, V.; **Imbrišak, M.**; Bondi, M.; Zamorani, G.; Ceraj, Lana; Vardoulaki, E.; Delhaize, J.: *The VLA-COSMOS 3 GHz Large Project: Average radio spectral energy distribution of active galactic nuclei* // Astronomy & astrophysics, 643 (2020), A51, 11

3. Nikšić, T.; **Imbrišak, M.**; Vretenar, D.: *"Sloppy" nuclear energy density functionals. II. Finite nuclei* // Physical review. C, 95 (2017), 5; 054304

Acknowledgements

The work presented in chapters 4 and 5 is financed within the Tenure Track Pilot Programme of the Croatian Science Foundation and the École Polytechnique Fédérale de Lausanne, and the project TTP-2018-07-3554 Exotic Nuclear Structure and Dynamics, with funds from the Croatian-Swiss Research Programme.

Abstract

Statistical analysis proved to be significant in determining the errors of theoretical models, assessing the risk of extrapolation and the sensitivity of the model to changes, as well as in the description of instabilities in the models. However, the importance of applying the analysis of the unreliability of individual parameters in the form of nuclear energy density functionals has only recently been recognized, where one of the main problems of error analysis of the parameters of those models is the possible sloppiness of some parameters or their combinations. This was precisely the motivation for using methods of reducing the number of model parameters based on information geometry - an interdisciplinary field that applies the ideas of differential geometry to statistical problems that have recently found application in the physical description of various systems in physics, chemistry and biology. Since in nuclear physics the manifold boundary approximation method (MBAM) has recently proved useful for the classification and reduction of complex models such as EDF models of the atomic nucleus, the statistical nature of determining the unreliability by the MBAM method was investigated in more detail using the example of the DD-PC1 functional. Also, using automatic differentiation in the code of the DD-PC1 functional and the MBAM method, a more thorough and precise analysis of the unreliability of the parameters itself is obtained, which enabled the optimization of time and numerical precision. Considering the novelty of the MBAM method itself, stability tests of the conclusions obtained by MBAM geodesics based on differential geometry techniques in the field of classical statistics and Bayesian statistics in the form of the Markov Chain Monte Carlo method were conducted.

Sažetak

Statistička analiza pokazala se značajna pri određivanju pogrešaka teorijskih modela, procjeni rizika ekstrapolacije i osjetljivosti modela na promjene kao i pri opisu nestabilnosti u modelima. Međutim, tek je nedavno prepoznat značaj primjene analize nepouzdanosti parametera nuklearnih energijskih funkcionala gustoće. Jedan od glavnih problema analize pogrešaka parametara tih modela moguća je aljkavost nekih parametara ili njihovih kombinacija. Upravo je to bila motivacija za korištenje metoda redukcije broja parametara modela temeljenih na informacijskoj geometriji - interdisciplinarnom području koje primjenjuje ideje diferencijalne geometrije na statističke probleme koje je odnedavno našlo primjenu u fizikalnom opisu različitih sustava u fizici, kemiji i biologiji. Budući da se u nuklearnoj fizici metoda aproksimacije modela mnogostrukosti s rubom (MBAM) nedavno pokazala korisnom za klasifikaciju i redukciju kompleksnih modela kao što su EDF modeli atomske jezgre, podrobnije je istražena statistička priroda određivanja nepouzdanosti MBAM metodom na primjeru DD-PC1 funkcionala. Također, koristeći automatsku diferencijaciju u kodu DD-PC1 funkcionala te MBAM metode, dobiva se temeljitija i preciznija analiza nepouzdanosti samih parametara čime se omogućila optimizacija vremena i numeričke preciznosti. S obzirom na novost same MBAM metode, provedeni su testovi stabilnosti zaključaka dobivenih MBAM geodezicima temeljeni na tehnikama diferencijalne geometrije u području klasične statistike te Bayesove statistike u vidu Markov Chain Monte Carlo metode.

Prošireni sažetak

Uvod

Atomske jezgre je teško modelirati zbog toga što je nepoznat egzaktan oblik interakcije između nukleona te stoga postoje različiti pristupi ovom problemu, ovisno o položaju jezgre u karti nuklida. Jednostavni modeli koriste se za lake jezgre, dok se za teže koristi model ljusaka. Za pokrivanje cijele karte nuklida koriste se nuklearni energijski funkcionali gustoće (EDF) [1].

Funkcionali gustoće nuklearne energije

Teorija funkcionala gustoće koristi se u fizici čvrstog stanja za rješavanje fenomena više elektrona minimiziranjem funkcionala gustoće elektrona. Nuklearni EDF-ovi aproksimiraju se funkcionalima gustoća i struja osnovnog stanja nukleona i njihovim gradijentima. Poluempirijski EDF-ovi su funkcionali koji slijede mikroskopski motivirani ansatz za nukleonsku gustoću, a njihovi parametri empirijski su prilagođeni za reprodukciju dane jednadžbe stanja.

Korištenje relativističkog Lagrangiana srednjeg polja u modelu mezonske izmjene konačnog dometa poboljšava opis asimetrične nuklearne materije. Utvrđeno je da modeli točkastog vezanja (PC) daju rezultate usporedive s modelima konačnog dometa, a poluempirijski funkcional gustoće energije DD-PC1 je često korišteni funkcional ovog tipa koji uključuje stupnjeve slobode nukleona i primjenjuje se na širok raspon atomskih jezgri.

Relativistički model srednjeg polja

Funkcional relativističkog modela srednjeg polja atomske jezgre izveden je na temelju pretpostavke klase modela točkastog vezanja korištenih u sljedećim poglavljima. Lagrangian modela točkastog vezanja temelji se na skupu bilinearnih struja Diracovog spinora, ψ, koji se koristi za opisivanje nukleona. Rezultirajući Lagrangian podijeljen je na Lagrangian slobodne čestice, bilinearnu struju, bilinearnu strujnu derivaciju i elektromagnetsku komponentu. Interakcijski dijelovi Lagrangiana sastoje se od četiri tipa fermionskih interakcija: izoskalar-skalar, izovektor-vektor, izovektor-skalar i izovektor-vektor.

Razmatrana klasa modela točkastog vezanja koristi samo članove drugog reda, zanemarujući, npr., šest-fermionske i osam-fermionske vrhove, ali umjesto toga promovira konstante sprezanja u funkcije gustoće nukleona.

Pretpostavke funkcionala DD-PC1

Funkcionali DD-PC definirani su posebnim skupom parametrizacijskih jednadžbi. Ove jednadžbe koriste normaliziranu gustoću simetrične nuklearne materije, predstavljenu varijablom x, za definiranje interakcija izoskalar-skalar, izovektor-vektor i izovektor-skalar.

Funkcional gustoće energije DD-PC1 razvijen je i testiran u proračunima srednje teških i teških jezgri. Veza izovektor-vektor određena je iz empirijskih svojstava asimetrične nuklearne materije, pri čemu su određeni parametri postavljeni na nulu.

Statistička analiza u nuklearnoj fizici

Tijekom prošlog desetljeća, statistička analiza pogrešaka postala je sve više prepoznata u istraživanju EDF-ova zbog svoje sposobnosti kvantificiranja teoretskih pogrešaka, razlikovanja sigurnih i rizičnih ekstrapolacija, pružanja analize osjetljivosti i pružanja uvida u nestabilnosti modela. To uključuje korištenje klasičnog ili Bayesovog zaključivanja za procjenu nesigurnosti parametara i primjenu metoda redukcije modela temeljenih na konceptima informacijske geometrije.

Na primjer, aproksimacijska metoda mnogostrukosti s rubom razvijena je za proučavanje složenih i aljkavih problema koji se javljaju u fizici, kemiji i biologiji. Temelji se na geodezicima u prostoru parametara modela i može se primjenjivati iterativno za pojednostavljenje modela uz zadržavanje preciznosti. Ove su tehnike dosad primijenjene na niz EDF modela, uključujući modele točkastog vezanja ovisne o gustoći koje je teško analizirati korištenjem klasičnih statističkih tehnika.

Informacijska geometrija

Kako bi se pronašao najbolji model, statističke analize često koriste metodu maksimalne vjerodostojnosti. Pritom se pretpostavlja da svako mjerenje slijedi normalnu distribuciju te se minimizira funkcija log-vjerodostojnosti kako bi odredio najbolji model.

Točnost procjena može se provjeriti pomoću Cramer-Rao granice, koja uključuje Fisherovu informacijsku matricu. MBAM metoda može se koristiti za poboljšanje točnosti, čak i u modela sa širokim rasponom vrijednosti matrice kovarijancije. U nuklearnoj fizici, postupak MBAM metode korišten je za stvaranje učinkovitih modela.

Algoritamska diferencijacija i modeli nuklearne strukture

Algoritamsko diferenciranje (AD) moćna je tehnika koja se koristi za učinkovitu i točnu procjenu derivacija numeričkih funkcija. Ova tehnika našla je široku primjenu u poljima kao što su računalna dinamika fluida, atmosferske znanosti, optimizacija inženjerskog dizajna i strojno učenje.

AD iskorištava činjenicu da se numerička izračunavanja formiraju iz konačnog skupa elementarnih operacija s poznatim jakobijanima, što omogućuje izračunavanje jakobijana kompliciranih izraza korištenjem lančanog pravila za kompoziciju funkcija. Postoje dvije glavne varijante AD-unaprijed i AD-unazad.

Python pruža izvrsnu platformu za implementaciju AD tehnika pomoću paketa poput autograda. Autograd je Python paket koji korisnicima omogućuje pisanje složenih funkcija, čak i korištenjem rekurzija. Može rukovati Python kodom koji sadrži jednostavna izračunavanja, while petlje, rekurziju i if naredbe, kao i mnoge funkcije dostupne u bibliotekama numpy i scipy.

Autograd također omogućuje korisnicima da definiraju potrebne derivacije u smislu autogradovih funkcija, a derivacijama višeg reda upravlja se automatski. Uz Python i pakete poput autograda, korisnici mogu jednostavno implementirati AD tehnike i izvesti točne i učinkovite derivacije svojih funkcija.

Stabilnost aproksimacijske metode mnogostrukosti s rubom kod redukcije modela nuklearne strukture

Modeli točkastog vezanja [2; 3] mogu proizvesti rezultate usporedive s onima konačnog dosega, čak i ako su PC interakcije prilagođene nuklearnoj materiji i svojstvima osnovnog stanja konačnih jezgri [4]. Projekt 'Univerzalni funkcional gustoće nuklearne energije' (UNEDF) bio je usmjeren na širok raspon pionirskih razvoja u EDF-u, uključujući nesigurnost kvantifikacija nuklearne teorije [5; 6].

U posljednjem desetljeću, analiza statističkih pogrešaka, koristeći klasičnu ili Bayesovu paradigmu, počela je biti prepoznata u EDF istraživanju radi kvantificiranja teoretskih pogrešaka, razlikovanja sigurne i riskantne ekstrapolacije te zbog uvida u nestabilnosti modela [7–17].

Iako je u početku bila usmjerena na strojno učenje i neuronske mreže [18; 19], informacijska geometrija je interdisciplinarno područje koji uvodi koncepte diferencijalne geometrije na statističke probleme [20; 21]. Nedavno je metoda aproksimacije metode mnogostrukosti s rubom (MBAM) [22–24] razvijena za proučavanje složenih i aljkavih problema koji se javljaju u fizici, kemiji i biologiji [25–27] kako bi se klasificirali ili smanjili složeni modeli, kao što su nuklearni EDF [28–30].

Prepreke uspostavljanju jedinstvenog teorijskog okvira za tretiranje nuklearnog problema s više tijela su složenost nukleon-nukleon interakcije u nuklearnom mediju i sprezanje između stupnjeva slobode jednog nukleona i kolektivnih stupnjeva slobode. Nuklearni EDF-ovi i modeli strukture temeljeni na njima, postali su obećavajući alat za opis svojstava osnovnog stanja i niskoenergijska kolektivna pobuđenja srednje teških i teških jezgri.

Nepoznati točan nuklearni EDF aproksimira se funkcionalima sačinjenima od potencija i gradijenata gustoće i struje nukleona u osnovnom stanju, a koji predstavljaju raspodjelu materije, spina, izospina, impulsa i kinetičke energije. Većina parametara semiempirijskog funkcionala se podešava, u aproksimaciji lokalne gustoće, kako bi se reproducirala mikroskopska jednadžba stanja (EoS), beskonačna simetrična i asimetrična nuklearna materija, a na kraju i neutronska materija. Semiempirijski funkcionali koji su razvijeni tijekom posljednjeg desetljeća [5; 31–40] vrlo su uspješno primijenjeni na proučavanje svojstava strukture, od grupiranja u relativno lake jezgre do stabilnosti superteških sustava.

U prethodnim studijama [28; 29], autori su koristili pojmove iz informacijske geometrije kako bi pokazali da su nuklearni EDF-ovi, općenito, aljkavi [22–25; 41]. Izraz “aljkav" odnosi se na činjenicu da su predviđanja nuklearnih EDF-ova i srodnih modela osjetljiva na samo nekoliko kombinacija parametara (*stiff* kombinacije parametara) i pokazuju eksponencijalno smanjenje osjetljivosti na varijacije preostalih kombinacija parametara (*soft* kombinacije parametara). To znači da su *soft* kombinacije parametara samo labavo ograničene dostupnim podacima te da većina nuklearnih EDF-ova zapravo sadrži modele niže efektivne dimenzionalnosti.

U Ref. [28], koristeći MBAM [24] autori su izveli najučinkovitiji funkcionalni oblik parametara vezanja ovisnih o gustoći reprezentativnog modela nuklearnog EDF-a. Podaci korišteni u ovom izračunu uključivali su skup točaka na mikroskopskoj jednadžbi stanja simetrične nuklearne materije i neutronske materije.

U Ref. [29] smo proširili ovaj izračun korištenjem jednostavnih numeričkih aproksimacija za izračunavanje derivacija observabli s obzirom na parametre modela. Tako smo bili u mogućnosti primijeniti MBAM na realne modele ograničen ne samo nuklearnom jednadžbom stanja nego također opažanjima izmjerenim u konačnim jezgrama. Tijekom naše analize parametarizacija u Ref. [29] primijetili smo da je numerička integracija geodezijske jednadžbe mogla dosegnuti granicu mnogostrukosti u konačnom broju koraka integracije, što ukazuje na divergenciju determinante metrike u određenoj regiji prostora parametara. Ovo iznenađujuće ponašanje motiviralo je istraživanje stabilnosti dobivenih redukcija modela MBAM metodom, budući da se divergentno područje može nenamjerno propustiti korištenjem prevelikih koraka integracije.

U ovom poglavlju primijenili smo informacijsko-teorijske metode kako bismo istražili stabilnost redukcija modela MBAM metodom. U ilustrativnom primjeru modela točkastog vezanja ovisnog o gustoći relativističkog nuklearnog EDF-a, koristeći Monte Carlo simulacije, utvrđeno je da su glavni zaključci dobiveni iz MBAM postupka stabilni na varijacije parametara modela. Nadalje, nalazimo da kraj geodezika MBAM metode nastaje kada determinanta Fisherove informacijske metrike iščezava, čime se učinkovito odvaja prostor parametara na dvije nepovezane regije.

Klasična i Bayesova analiza pogreške relativističkog modela srednjeg polja za dvostruko magične jezgre

Nuklearni EDF-ovi naširoko su korišteni okvir za opisivanje fenomena nuklearne strukture. Budući da se izmjena teških mezona ne može razriješiti pri niskim energijama, samosuglasni relativistički okvir srednjeg polja može se formulirati u smislu PC nukleonske interakcije. Ovaj pristup daje rezultate usporedive s pristupom vezanja mezonske izmjene za konačne jezgre [2; 3]. Rezultirajući 'najprikladniji model' [kao što je DD-PC1 funkcional, vidi, npr. 42] zahtijeva fino podešavanje ovisnosti o gustoći interakcije izoskalar-skalar i izovektor-vektor termina za nuklearnu materiju i svojstva osnovnog stanja konačnih jezgri.

Pitanje kvantifikacije nesigurnosti u nuklearnim EDF-ovima nedavno je privuklo pozornost, fokusirajući se na proučavanje procjena pogreške statističkom analizom [10; 11], procjena sustavnih pogrešaka [8; 9] i analiza korelacije [9; 43]. Međutim, statistička analiza je izazovnija za PC modele jer je utvrđeno da oni pokazuju eksponencijalni raspon osjetljivosti na varijacije parametara [42]. Utvrđeno je da je ovo ponašanje značajka aljkavih modela [44].

Nedavni napredak u razumijevanju ponašanja aljkavih modela [23; 25; 26; 45] doveo je do novih pristupa analizi kao što je MBAM metoda [24]. MBAM je već korišten za sustavnu konstrukciju učinkovitih funkcionala nuklearne gustoće sve nižih dimenzija i manjeg utjecaja aljkavosti. To je ilustrirano na funkcionalu DD-PC1 procijenjenom na pseudo-podacima za beskonačnu simetričnu nuklearnu materiju [28].

U prethodnom poglavlju istražena je ukupna stabilnost MBAM procedure primijenjene u redukciji modela nuklearne strukture koristeći metode informacijske geometrije i Monte Carlo simulacija. Za razliku od jednostavnog slučaja beskonačne nuklearne materije, gdje bismo morali riješiti samo jednostavnu iterativnu proceduru za dobivanje Diracove mase i energije vezanja, konačne jezgre zahtijevaju pažljiv opis nuklearnog problema više tijela.

Općenito govoreći, statistička analiza može se izvesti ili u Bayesovom okviru - korištenjem razrađenih Monte Carlo simulacija ili u “klasičnom" okviru, dobivenom izračunavanjem Fisherove informacijske matrice (FIM) i njezinog inverza (matrice kovarijance) iz odabrane statistike modela [vidi, npr., 9]. Potonji bi pristup trebao, u načelu, oduzimati manje vremena od izvođenja velike Monte Carlo simulacije. Međutim, kada se računa FIM, potrebno je izvrijedniti prve derivacije odabranog statističkog modela, bilo numerički ili analitički. Pokušaj jednostavnog proširenja postojećih implementacija RMF kodova napisanih u Fortranu [46–49] uveo bi nesigurnosti zbog upotrebe numeričke diferencijacije. Stoga smo odlučili implementirati jednostavnu verziju RMF koda u Pythonu, u kojemu već postoje dobro testirane biblioteke za algoritamsku diferencijaciju (AD).

Proširujući naš prethodni rad [50], koji se bavio primjenom informacijske geometrije na EDF u slučaju nuklearne materije, ovdje smo predstavili statističku analizu jednostavnog postupka za određivanje RMF energije vezanja za skup dvostruko magičnih jezgri s Woods-Saxon potencijalom. Usporedili smo procjene pogreške između bržeg postupka koji koristi FIM i numerički zahtjevnije Bayesove MCMC metode.

Čak i u složenom slučaju konačnih jezgri, nesigurnosti EDF parametara mogu se pouzdano procijeniti korištenjem FIM-a u kombinaciji s algoritamskom diferencijacijom. Predloženi pristup analizi grešaka ima prednost u izbjegavanju dugotrajnog uzorkovanja prostora parametara, kakav zahtijevaju Bayesove statističke tehnike.

List of Figures

List of Tables

Contents

Chapter 1

Introduction

The quantitative description of nuclear structure is one of the most complex quantum-mechanical many-body problems for which the exact solutions are unavailable, and the exact nucleon-nucleon interaction is unknown. Therefore, many possible approaches to describing nuclei exist in the literature, and this choice is strongly impacted by the position of the modeled nucleus in the chart of nuclides. For example, very light nuclei are often described by exactly solvable ab initio models based on the assumed nucleon-nucleon interaction. In contrast, heavier nuclei are usually characterized by the shell model based on the effective interaction whose matrix elements need to be adjusted for different regions of the chart of nuclides. If a single model that can encompass the entire chart of nuclides is needed, the only available choice is the class of models based on the nuclear energy density functionals [EDFs, 1].

1.1 Nuclear energy density functionals

The density functional theory is a theory first developed to treat many-electron phenomena in solid-state physics, wherein the solutions to the many-body problem are transformed to a minimization of the functional of the electron density. Being an approximation method, the exact nuclear energy density functional (EDF) is unknown. It is therefore approximated by functionals of powers of ground-state nucleon densities and currents and their gradients, representing distributions of matter, spin, isospin, momentum, and kinetic energy. A particular

EDF need not be related to the microscopic interactions between nucleons, and many are therefore motivated empirically.

The semi-empirical EDFs represent an intermediate class of functionals that follow a microscopically motivated ansatz for nucleonic density. Their parameters are empirically adjusted to reproduce a given equation of state. Among the various possible nuclear EDFs, using the relativistic mean-field Lagrangian in the finite-range meson-exchange model has been found to improve the description of asymmetric nuclear matter [52]. This result has been possible using density-dependent meson-nucleon couplings [53]. In the last decade, many successful semi-empirical EDFs have been developed [5; 31; 33–40; 54], and successfully applied to study a diversity of structure properties, from clustering in relatively light nuclei to the stability of superheavy systems, and from bulk and spectroscopic properties of stable nuclei to the physics of exotic nuclei at the particle drip lines.

Although approximating the finite-range effects of the interactions by contact interaction, a simpler class of point-coupling (PC) models has been found to produce comparable results to the finite-range models [2; 3], even if the point-coupling interactions are being adjusted to nuclear matter and ground-state properties of finite nuclei [4]. The semi-empirical energy density functional DD-PC1 [54] is an often-used point-coupling functional that includes nucleon degrees of freedom and considers only second-order interaction terms, and that applies to a wide range of atomic nuclei [55].

1.2 The relativistic mean-field model

In this section, the functional of the relativistic mean-field model of the atomic nucleus is described. This section covers the derivation based on assuming the class of the point-coupling models used in the following chapters.

The relativistic Lagrangian governing point-coupling models is based on a set of bilinear currents of the Dirac spinor, ψ, used to describe nucleons

$$\bar{\psi} O_\tau \Gamma \psi, O_\tau \in \{1, \tau_i\}, \Gamma \in \{1, \gamma_\mu, \gamma_5, \gamma_5\gamma_\mu, \sigma_{\mu\nu}\}. \tag{1.1}$$

where τ_i are the Pauli matrices for isospin, and Γ are the Dirac matrices. The resulting Lagrangian may be divided into the free-particle, $\mathcal{L}_{free}$, bilinear current, $\mathcal{L}_{4f}$, bilinear current derivative, $\mathcal{L}_{der}$, and the electromagnetic, $\mathcal{L}_{em}$, components [54]

$$\mathcal{L} = \mathcal{L}_{free} + \mathcal{L}_{4f} + \mathcal{L}_{der} + \mathcal{L}_{em}. \tag{1.2}$$

The interacting parts of the Lagrangian are composed of the four types of fermion interactions: the isoscalar-scalar, $(\bar{\psi}\psi)^2$, isovector-vector, $(\bar{\psi}\gamma_\mu\psi)(\bar{\psi}\gamma^\mu\psi)$, isovector-scalar, $(\bar{\psi}\vec{\tau}\psi)\cdot(\bar{\psi}\vec{\tau}\psi)$, and the isovector-vector type, $(\bar{\psi}\vec{\tau}\gamma_\mu\psi)\cdot(\bar{\psi}\vec{\tau}\gamma^\mu\psi)$.

In the point-coupling class of models, the interacting terms are added to the Lagrangian by multiplying the bilinear currents by their respective couplings (δ_S, α_S, α_V, α_{TS} and α_{TV}) that are dependent on the baryon density, $\hat{\rho}$, defined as

$$\hat{\rho}u^\mu = \bar{\psi}\gamma^\mu\psi, \tag{1.3}$$

where u^μ is the four-velocity $u^\mu = (1 - v^2)^{-1/2}(1, \vec{v})$. The considered class of point-coupling models uses only second-order terms, disregarding, e.g., six-fermion and eight-fermion vertices, but instead promotes the coupling constants to functions of nucleon density [54]. These models consider the same building blocks as the meson-exchange models, wherein the single-particle properties are tied to the three meson fields: the isoscalar-scalar σ meson, the isoscalar-vector ω meson, and the isovector-vector ρ meson, without the isovector-scalar term [4].

Using the notation outlined above, the components of the Lagrangian are expanded as

$$\mathcal{L}_{free} = \bar{\psi}(i\gamma_\mu\partial^\mu - M)\psi, \tag{1.4}$$

$$\begin{aligned}\mathcal{L}_{4f} \;=\; & -\frac{1}{2}\alpha_S(\hat{\rho})(\bar{\psi}\psi)(\bar{\psi}\psi) - \frac{1}{2}\alpha_V(\hat{\rho})(\bar{\psi}\gamma_\mu\psi)(\bar{\psi}\gamma^\mu\psi) \\ & -\frac{1}{2}\alpha_{TS}(\hat{\rho})(\bar{\psi}\vec{\tau}\psi)\cdot(\bar{\psi}\vec{\tau}\psi) - \frac{1}{2}\alpha_{TV}(\hat{\rho})(\bar{\psi}\vec{\tau}\gamma_\mu\psi)\cdot(\bar{\psi}\vec{\tau}\gamma^\mu\psi),\end{aligned} \tag{1.5}$$

$$\mathcal{L}_{der} = -\frac{1}{2}\delta_S(\hat{\rho})(\partial_\nu\bar{\psi}\psi)(\partial^\nu\bar{\psi}\psi) + O(\partial^2), \tag{1.6}$$

$$\mathcal{L}_{em} = eA^\mu\bar{\psi}\frac{1+\tau_3}{2}\gamma_\mu\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \tag{1.7}$$

where the derivative Lagrangian $\mathcal{L}_{der}$ was expanded out to the first derivative terms.

The equation of motion for the nucleons can be found by varying the Lagrangian by $\bar{\psi}$, yielding an equation

$$[\gamma_\mu(i\partial^\mu - V^\mu) - (m + S)]\psi = 0, \tag{1.8}$$

where the following abbreviations were introduced

$$S = \Sigma_S + \vec{\tau} \cdot \vec{\Sigma}_{TS} + \Sigma_{rS}, \tag{1.9}$$

$$V^\mu = \Sigma^\mu + \vec{\tau} \cdot \vec{\Sigma}_S + \Sigma_{rS}, \tag{1.10}$$

where the components were further separated into isoscalar-vector, Σ^μ, isoscalar-scalar Σ_S, isovector-vector, Σ_T^μ, isovector-scalar, Σ_{TS}, isoscalar-scalar derivative components Σ_{rS} and Σ_r^μ:

$$\Sigma^\mu = \alpha_V(\bar{\psi}\gamma^\mu\psi) - eA^\mu\frac{1+\tau_3}{2}, \tag{1.11}$$

$$\vec{\Sigma}_T^\mu = \alpha_{TV}(\bar{\psi}\vec{\tau}\gamma^\mu\psi), \tag{1.12}$$

$$\Sigma_S = \alpha_S(\bar{\psi}\psi) - \delta_S\Box(\bar{\psi}\psi), \tag{1.13}$$

$$\vec{\Sigma}_{TS} = \alpha_{TS}(\bar{\psi}\vec{\tau}\psi), \tag{1.14}$$

$$\Sigma_{rS} = -\frac{\partial\delta_S}{\partial\hat{\rho}}(\partial_\nu j^\mu)u_\mu(\partial^\nu(\bar{\psi}\psi)), \tag{1.15}$$

$$\begin{aligned}\Sigma_r^\mu \quad &= \tfrac{u^\mu}{2}\Big(\tfrac{\partial\alpha_S}{\partial\hat{\rho}}(\bar{\psi}\psi)(\bar{\psi}\psi) + \tfrac{\partial\alpha_{TS}}{\partial\hat{\rho}}(\bar{\psi}\vec{\tau}\psi)\cdot(\bar{\psi}\vec{\tau}\psi)\\ &+\tfrac{\partial\alpha_V}{\partial\hat{\rho}}(\bar{\psi}\gamma^\nu\psi)(\bar{\psi}\gamma_\nu\psi) + \tfrac{\partial\alpha_{TV}}{\partial\hat{\rho}}(\bar{\psi}\vec{\tau}\gamma^\nu\psi)\cdot(\bar{\psi}\vec{\tau}\gamma_\nu\psi)\\ &+\tfrac{\partial\delta_S}{\partial\hat{\rho}}(\partial^\nu(\bar{\psi}\psi))(\partial_\nu(\bar{\psi}\psi))\Big).\end{aligned} \tag{1.16}$$

In the rest frame, the only contribution to the currents is from the densities. Therefore, in the ground state of the nucleus, Φ, the following densities can be introduced

$$\rho_S = \langle\Phi|\bar{\psi}\psi|\Phi\rangle = \rho_s^p + \rho_s^n, \tag{1.17}$$

$$\rho = \langle\Phi|\bar{\psi}\gamma^0\psi|\Phi\rangle = \rho^p + \rho^n, \tag{1.18}$$

$$\rho_{s3} = \langle\Phi|\bar{\psi}\tau_3\psi|\Phi\rangle = \rho_s^p - \rho_s^n, \tag{1.19}$$

$$\rho_{tv} = \langle\Phi|\bar{\psi}\tau_3\gamma^0\psi|\Phi\rangle = \rho^p - \rho^n. \tag{1.20}$$

In the rest-frame, the single-particle energies, ϵ_i, can therefore be found by solving the Dirac equation,

$$(\alpha \cdot \mathbf{p} + \beta m^*(r) + V(r))\,\psi_i = \epsilon_i \psi, \tag{1.21}$$

where the potential V and the effective mass, m^* are given by

$$V(r) = \alpha_V \rho + \alpha_{TV} \tau_3 \rho_{TV} + eA^0 + \frac{1}{2}\left(\frac{\partial \alpha_S}{\partial \rho}\rho_S^2 + \frac{\partial \alpha_V}{\partial \rho}\rho_V^2 + \frac{\partial \alpha_{TV}}{\partial \rho}\rho_{TV}^2\right) \tag{1.22}$$

$$m^*(r) = m + \alpha_S \rho_S + \delta_S \nabla^2 \rho_S. \tag{1.23}$$

The associated energy density functional in the rest frame is

$$E_{RMF} = \int d^3x \sum_i \psi_i^\dagger (\alpha{\cdot}\mathbf{p}{+}\beta m)\psi_i {+} \frac{1}{2}\left(\alpha_S \rho_S^2 + \alpha_V \rho_V^2 + \alpha_{TV}\rho_{TV}^2 + \delta_S \rho_S \nabla^2 \rho_S - (\nabla A^0)^2 + eA^0 \rho^p\right). \tag{1.24}$$

1.3 Pairing

Pairing is a necessary ingredient for studying nuclei with open shells and is therefore necessary to describe nuclei that are not doubly magic [56].

In the constant gap approximation [57], each single-particle state is occupied according to the occupation probability v_i^2, calculated using the BCS formula

$$v_i^2 = \frac{1}{2}\left(1 - \frac{\epsilon_i - \lambda}{\sqrt{(\epsilon_i - \lambda)^2 + \Delta^2}}\right), \tag{1.25}$$

where λ is the chemical potential and Δ the gap parameter. The chemical potential is determined separately for protons and neutrons by finding a solution of the equations for the chemical potentials for protons and neutrons

$$\sum_i v_{i,p}^2(\lambda_p) = Z \tag{1.26}$$

$$\sum_i v_{i,n}^2(\lambda_n)) = N, \tag{1.27}$$

so that the total number of neutrons and protons is conserved. The pairing energy can then be computed from a simple expression

$$E_{pair} = -G \sum_i (v_i u_i)^2, \tag{1.28}$$

where u_i is the complement of the occupation probabilities, $u_i^2 = 1 - v_i^2$, and G is a constant determined from the self-consistency condition

$$\Delta = G \sum_i u_i v_i. \tag{1.29}$$

Since the sum necessary for computing the pairing energy diverges, one often introduces cutoff energy [46; 56].

1.4 The assumptions of the DD-PC1 functional

The density-dependent point coupling (DD-PC) functionals are defined by the following parametrization [4]

$$\alpha_s(\rho) \quad = a_s + (b_s + c_s)e^{-d_s x}, \tag{1.30}$$

$$\alpha_v(\rho) \quad = a_v + b_v e^{-d_v x}, \tag{1.31}$$

$$\alpha_{tv}(\rho) \quad = b_{tv} e^{-d_{tv} x}, \tag{1.32}$$

where the density has been normalized by the saturation density in symmetric nuclear matter, ρ_{sat}, as

$$x = \frac{\rho}{\rho_{sat}}. \tag{1.33}$$

An optimal energy density functional of the DD-PC class, DD-PC1, has been determined from a fit to the masses of 64 axially deformed nuclei and tested in calculations of properties of spherical and deformed medium-heavy and heavy nuclei [4]. The density dependence of the isovector-vector coupling has been inferred from empirical properties of asymmetric nuclear matter with the a_{TV} and c_{TV} parameters being set to zero to calculations of asymmetric nuclear matter [58] - an already established procedure for density-dependent couplings [59].

1.5 Statistical analysis in nuclear physics

In the last decade, statistical error analysis, either employing classical or Bayesian inference, has started to be recognized [11] in EDF research for its ability to quantify theoretical errors, distinguish safe and risky extrapolations, provide sensitivity analysis, and offer insight into model instabilities [7–9; 11–13; 15]. For example, a recent paper [13] estimated the confidence intervals of the mean-field single-nucleon energies in the chain of superheavy nuclei using the Monte Carlo techniques. They have, however, been restricted to the Woods-Saxon nuclear mean-field model to reduce computational complexity. The Bayesian approach for estimating parameter uncertainties is to assume a prior distribution of parameters and then sample the posterior distribution. Such a procedure has been applied, e.g., to study the liquid drop model and the Skyrme functional [15].

The density-dependent point-coupling models are difficult to analyze using classical statistical techniques. This behavior has been tied to an exponential range of sensitivity to parameter variations, prompting the application of model reduction methods based on concepts of information geometry [28; 29]. Information geometry is an interdisciplinary field that introduces differential geometry concepts to statistical problems [20; 21]. While its initial applications centered around machine learning and neural networks [18], it has recently started being applied to various questions in physics. For example, the Manifold Boundary approximation method (MBAM) [22–24] has been developed to study complex and sloppy problems occurring in physics, chemistry, and biology [25; 26] to either classify or reduce complex models, including EDFs [28; 29]. The method is based on geodesics in the space of model parameters (model manifold), equipped with the Fisher information metric (FIM). Its eigenvalues are computed at the best-fitting model point, and the geodesic equation is solved in the direction of the FIM eigenvector corresponding to the largest uncertainty eigenvalue. After a long-enough integration, the geodesic either reaches the end of the model manifold, breaking the integration procedure, or the components of the eigenvector stabilize. In either case, the method establishes the single parameter (or combination of parameters) that is the biggest contributor to the uncertainty of the model. This constraint is then used to simplify the model, and the procedure can be applied iteratively as long as the resulting model describes the initial dataset to the wanted precision.

Chapter 2

Information geometry

To find the best model, statistical analyses often use the maximum likelihood method. This assumes that each measurement follows a normal distribution. They maximize the log-likelihood function over a range of possible values to determine the best model. The accuracy of the estimates can be checked with the Cramer-Rao bound, which involves the Fisher information matrix. The manifold boundary approximation method can be used to improve accuracy, even with a wide range of covariance matrix values. In nuclear physics, the MBAM procedure was used to create effective models with tightly constrained parameters.

2.1 Maximum likelihood method

Model selection is usually performed using the maximum likelihood method, with the assumption that at the a-th measurement the data (x^a, y^a) can be described using a normal distribution, denoted by $\mathcal{N}$, by a model function $f(x^a, \mathbf{p}) \equiv f^a(\mathbf{p})$ as $y^a \sim \mathcal{N}\left(f^a(\mathbf{p}), (\sigma^a)^2\right)$. Here, σ^a is the uncertainty of each measurement, and $\mathbf{p}$ is chosen from an appropriate parameter space, denoted by $\mathcal{M}$. Finding the best-fitting model is equivalent to maximizing the following log-likelihood function $l(\mathbf{p})$ over $\mathbf{p} \in \mathcal{M}$,

$$l(\mathbf{p}) = \sum_a \ln \phi \left(\frac{y^a - f^a(\mathbf{p})}{\sigma^a} \right) , \tag{2.1}$$

with ϕ a Gaussian probability density. To simplify the notations, we shall use indices from the beginning of the Latin alphabet for measurements, and the Greek letters for derivatives

$\frac{\partial}{\partial \mathbf{p}^\mu}$, shortened to ∂_μ. In order to assess the accuracy of parameter estimates, one can utilize the Cramer-Rao bound. This involves expanding the log-likelihood to the second order by means of the Hessian. A comprehensive explanation of this approach can be found in Amari's publication [21] regarding information applications. Using the Hessian of the log-likelihood, we can compute the quantity

$$g_{\mu\nu}(\mathbf{p}) = \sum_a \frac{\partial_\mu f^a \partial_\nu f^a}{(\sigma^a)^2} \ . \tag{2.2}$$

which is referred to as the Fisher information matrix (FIM).

2.1.1 Information geometry

Information geometry can aid in the interpretation of this basic picture. The function $l(\mathbf{p})$ serves as a link between the manifolds, $\mathcal{M}$ and $\mathcal{N}$. Furthermore, the differential form, i.e., $dl = \partial_\mu l d\mathbf{p}^\mu$, forms a basis for the cotangent bundle on $\mathcal{N}$, labeled as $T^*\mathcal{N}$, while the FIM serves as a metric on $\mathcal{N}$. When dealing with mathematical equations in the context of Special and General relativity, it is common to use the same index repeatedly to indicate summation with respect to that index (Einstein's convention). This convention will be followed throughout this discussion. To create a metric on the parameter space, denoted by $\mathcal{M}$, we use the functional form of the log-likelihood. This involves computing the expectation value with respect to $\mathcal{N}$ [21]:

$$g \equiv E[dl \otimes dl].$$

The pullback operation, l^*, then induces a metric $g(\mathbf{p}) \in (T^*\mathcal{M})^2$ on $\mathcal{M}$, as

$$g(\mathbf{p}) = g_{\mu\nu} d\mathbf{p}^\mu \otimes d\mathbf{p}^\nu = E[\partial_\mu l \partial_\nu l] d\mathbf{p}^\mu \otimes d\mathbf{p}^\nu = l^* g.$$

In this procedure, we provide the model manifold $\mathcal{M}$ with a tangent bundle containing the basis $\partial_\mu \in T\mathcal{M}$ and a cotangent bundle consisting of the dual basis $d\mathbf{p}^\mu \in T^*\mathcal{M}$. As the normal family is included in the exponential family, $\mathcal{M}$ is a submanifold embedded within $\mathcal{N}$ and forms part of the curved exponential family [20].

Differential geometry involves the study of the tangent spaces of nearby points in a given space $\mathcal{M}$. These tangent spaces are interconnected through the covariant derivative, which is denoted as ∇_X and involves an arbitrary direction X. To put it simply, the covariant derivative operates on a tangent vector $Y \in T\mathcal{M}$ using the following formula:

$$\nabla_X(Y) = \nabla_X(Y^\mu \partial_\mu) = X^\nu \partial_\nu (Y^\mu) \partial_\mu + \Gamma^\kappa_{\mu\nu} X^\mu Y^\nu \partial_\kappa .$$

This formula essentially describes how the covariant derivative works on the tangent vector Y by taking into account the direction X. It involves a combination of partial derivatives of Y and the Christoffel symbols $\Gamma^\kappa_{\mu\nu}$, which represent the curvature of the space $\mathcal{M}$. The quantity $\Gamma^\kappa_{\mu\nu}$ stands for the Christoffel symbol when the metric-compatible connection with the condition

$$\nabla_X(g) = 0$$

is chosen (for details, see, e.g., Ref. [60]). For the FIM, the Christoffel symbols are given by

$$\Gamma^\kappa_{\mu\nu}(\mathbf{p}) = g^{\kappa\rho} \sum_a \frac{\partial_\rho f^a \partial_{\mu\nu} f^a}{(\sigma^a)^2} , \tag{2.3}$$

where $g^{\kappa\rho} = (g^{-1})_{\kappa\rho}$ denotes the inverse of the metric.

We also perform calculations of the Riemann curvature tensor and scalar curvature along the geodesic path. The Riemann curvature tensor is utilized for vectors X, Y, and Z in T(M) using the following formula:

$$R(X,Y)Z = [\nabla_X, \nabla_Y]Z - \nabla_{[X,Y]}Z.$$

The components of the Riemann tensor are expressed as

$$R_{\mu\nu\rho\kappa} = \sum_{ab} P^{ab} \left(\partial_{\mu\rho} \frac{f^a}{\sigma^a} \partial_{\nu\kappa} \frac{f^b}{\sigma^b} - \partial_{\mu\kappa} \frac{f^a}{\sigma^a} \partial_{\nu\rho} \frac{f^b}{\sigma^b} \right) , \tag{2.4}$$

where P^{ab} denotes the projection operator

$$P^{ab} = \delta^{ab} - g^{\mu\nu} \partial_\mu \frac{f^a}{\sigma^a} \partial_\nu \frac{f^b}{\sigma^b} . \tag{2.5}$$

The Ricci scalar (or scalar curvature) is computed simply as

$$R_{\mu\nu\rho\kappa} g^{\mu\rho} g^{\nu\kappa} . \tag{2.6}$$

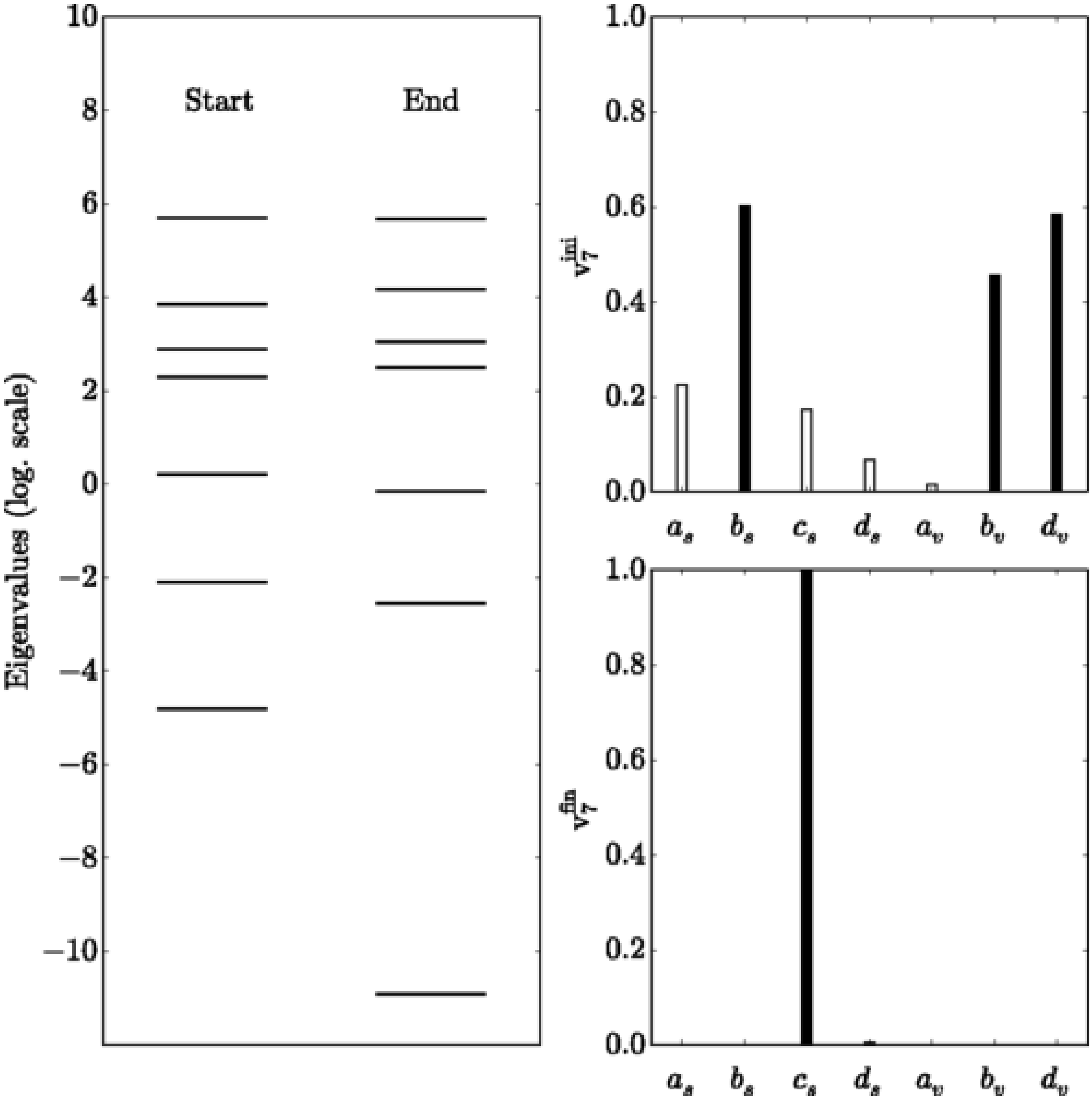

Figure 2.1: The initial (best-fit point) and final (at the boundary of the model manifold) eigenspectrum of the FIM for the DD-PC1 functional, with seven parameters in the isoscalar channel (left panel), and the initial and final eigenvectors that correspond to the smallest eigenvalues (panels on the right). Figure adapted from [28].

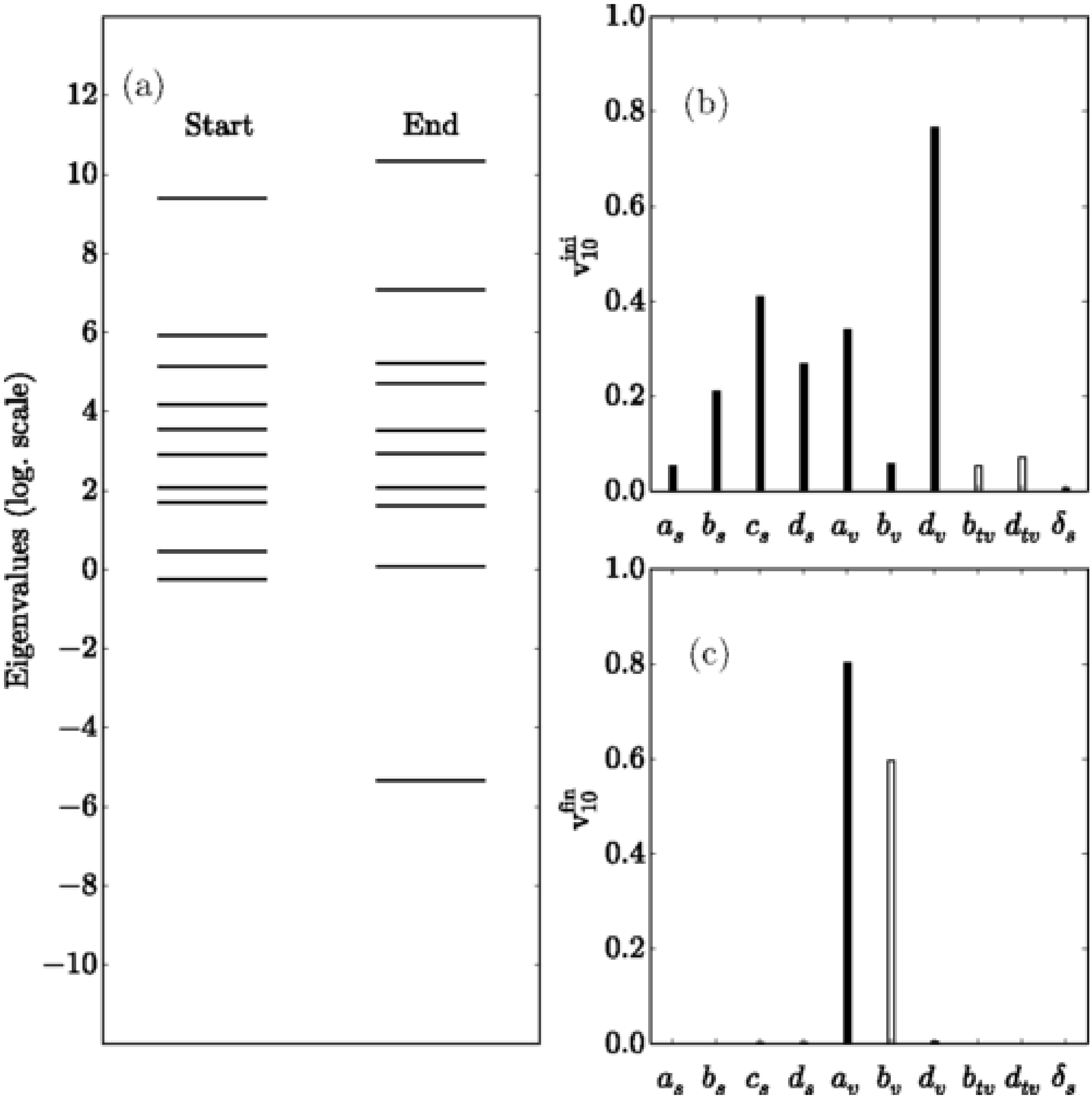

Figure 2.2: The initial (best-fit point) and final (at the boundary of the model manifold) eigenspectrum of the FIM for the ten-parameter functional (panel (a)). The eigenvectors that correspond to the initial and final smallest eigenvalues are shown in panels (b) and (c). Figure adapted from [29].

2.2 The manifold boundary approximation method

When dealing with large parameter uncertainties, it is often helpful to utilize model reduction procedures to enhance accuracy in parameter estimation. By implementing the MBAM method, parameters can be effectively constrained across a range of physical disciplines. Specifically, this method calculates the geodesic through solving the geodesic equation from the best-fitting point in the model manifold. This approach can lead to improved accuracy in parameter estimation, even when dealing with a wide spectrum of covariance matrix values. In the cases where the covariance matrix, and therefore the corresponding FIM, have a spectrum spanning many orders of magnitude [45], model reduction procedures can improve parameter estimates. The method computes the geodesic by solving the geodesic equation,

$$\nabla_{\dot{\mathbf{p}}}\dot{\mathbf{p}} = 0,$$

by starting from the best-fitting (bf) point in the model manifold,

$$\mathbf{p}_{\mathrm{bf}} \equiv \mathbf{p}^{\mu}_{\mathrm{bf}}\partial_{\mu}.$$

Note that the dot on $\mathbf{p}$ represents the differentiation with respect the affine parametrization of the geodesic. The geodesic equation, written in parameter components as

$$\ddot{\mathbf{p}}^{\kappa} + \Gamma^{\kappa}_{\mu\nu}\dot{\mathbf{p}}^{\mu}\dot{\mathbf{p}}^{\nu} = 0, \tag{2.7}$$

is solved with the $\dot{\mathbf{p}}$ initial conditions pointing in the direction of the FIM eigenvector, v^0, corresponding to its smallest eigenvalue. The largest eigenvalue of the covariance matrix is the main cause of uncertainty in the model parameters. To determine the parameter or combination of parameters that contribute the most to the uncertainty, we track the behavior of v^0 along the geodesic. Once identified, we eliminate this parameter from the model, resulting in a simpler model with lower parameter uncertainties. This process can be repeated until the simplified model accurately describes the data set.

2.2.1 MBAM method and density functionals

The MBAM procedure was previously used in nuclear physics [28] to create effective models with fewer dimensions. It involved identifying best-fit parameters, reaching the model manifold boundary, creating a new model with one less parameter, and fitting it to the data. The procedure was successful in eliminating model sloppiness. The analysis was extended to include ground-state properties of finite nuclei, leading to similar results for a ten-parameter model [29].

In the context of the DD-PC1 functional, the MBAM procedure was applied [28] by constructing effective models of successively lower dimension until sloppiness can eventually be eliminated, and all linearly independent parameter combinations are tightly constrained. They applied the MBAM procedure in four distinct steps. In the first step, they identified the best-fit parameters as well as the accompanying Hessian matrix of the χ^2 function. In the second step, they integrated the geodesic equation using the parameter values at the best-fit point and the eigendirection with the smallest eigenvalue as initial conditions until the boundary of the model manifold was reached. In the third step, the model limit associated with the model boundary was evaluated to produce a new model with one less parameter, while, in the fourth step, the new model was fit to the data and used as a starting point for the next iteration. The manifold boundary corresponds to a limit in which one or more parameters tend to limit values, and they found that the parameter c_s tended to zero. In Fig. 2.1, they plot the initial and final (at the boundary) eigenspectrum of the FIM in the left panel and the initial and final eigenvectors corresponding to the smallest eigenvalues (panels on the right). At the boundary of the model manifold, only the component c_s determines the decoupled eigendirection with the eigenvalue of the FIM approaching zero. The analysis was later extended [29] by including the data on ground-state properties of finite nuclei, which enabled the application of the MBAM procedure to the ten-parameter DD-PC1 model, with the initial and final eigenspectrum of the FIM shown in Fig. 2.2.

Chapter 3

Algorithmic differentiation and nuclear structure models

Algorithmic differentiation (AD) is a powerful technique used for the efficient and accurate evaluation of derivatives of numeric functions. It has found widespread applications in fields like computational fluid dynamics, atmospheric sciences, engineering design optimization, and machine learning. AD exploits the fact that numerical computations are formed from a finite set of elementary operations with known derivatives, making it possible to compute derivatives of complicated expressions using the chain rule for the composition of functions. There are two main variants of AD - the forward and reverse mode ADs.

Python provides an excellent platform for implementing AD techniques using packages like autograd. Autograd is a Python package that enables users to write complex functions, even using recursions. It can handle Python code containing simple computations, while loops, recursion, and if statements, as well as list indexing operations and many functions available in the numpy and scipy libraries. Autograd also allows users to define the required derivatives in terms of autograd functions, and higher-order derivatives are handled automatically. With Python and packages like autograd, users can easily implement AD techniques and derive accurate and efficient derivatives of their functions.

3.1 The concept of algorithmic differentiation

Algorithmic differentiation (AD, also called automatic differentiation) is a family of techniques developed for the efficient and accurate evaluation of derivatives of numeric functions [61]. AD has already found applications in fields like computational fluid dynamics, atmospheric sciences, engineering design optimization, and machine learning [62].

The AD technique exploits the fact that the numerical computations are formed from a finite set of elementary operations with known derivatives [63; 64], reducing the computation of derivatives of complicated expressions to the chain rule for the composition of functions. The computational realization of AD comes in two different variants, the forward and reverse mode ADs.

For a given function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, with the appropriate domain, $\{\mathbf{e}_i\}$, and codomain, $\{\mathbf{E}_j\}$ bases, the forward mode vector jacobian product (VJP) is a function of two variables in $\mathbb{R}^n$ that produces codomain vectors, $\mathbf{J}_F = J_F^{f,i}\mathbf{E}_i : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^m$

$$\mathbf{J}_F^f(\mathbf{x},\mathbf{r}) = \mathbf{E}_i J_F^f(\mathbf{x},\mathbf{r})^i = \mathbf{E}_i \frac{\partial f^i}{\partial x^j}\bigg|_x r^j, \tag{3.1}$$

while the reverse mode VJP $J_R : \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^m$ evaluates as a domain vector $J_R^f(\mathbf{x},\mathbf{r}) \in \mathbb{R}^n$ and is defined as

$$\mathbf{J}_R^f(\mathbf{x},\mathbf{r}) = J_R^f(\mathbf{x},\mathbf{r})_j \mathbf{e}^j = r^i \frac{\partial f^i}{\partial x^j}\bigg|_x \mathbf{e}^j, \tag{3.2}$$

where $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{r} \in \mathbf{R}^m$. For example, a simple composition of functions $y = f_1(f_2(x))$ is simply differentiated by the chain rule as

$$\frac{\partial y^i}{\partial x^j} = \frac{\partial f_1^i}{\partial r^k}(f_2(x)))\frac{\partial f_2^k}{\partial x^j}(x). \tag{3.3}$$

In the forward mode, the differentiation is realized as a composition of jacobians of the form

$$\mathbf{J}_F^{f_1 \circ f_2}(\mathbf{x},\mathbf{r}) = \mathbf{J}_F^{f_1}\left(f_2(\mathbf{x}), \mathbf{J}_F^{f_2}(\mathbf{x},\mathbf{r})\right), \tag{3.4}$$

while in the reverse mode, the differentiation yields the following composition of jacobians:

$$\mathbf{J}_R^{f_1 \circ f_2}(\mathbf{x},\mathbf{r}) = \mathbf{J}_R^{f_2}\left(\mathbf{x}, \mathbf{J}_R^{f_1}(f_2(\mathbf{x}),\mathbf{r})\right). \tag{3.5}$$

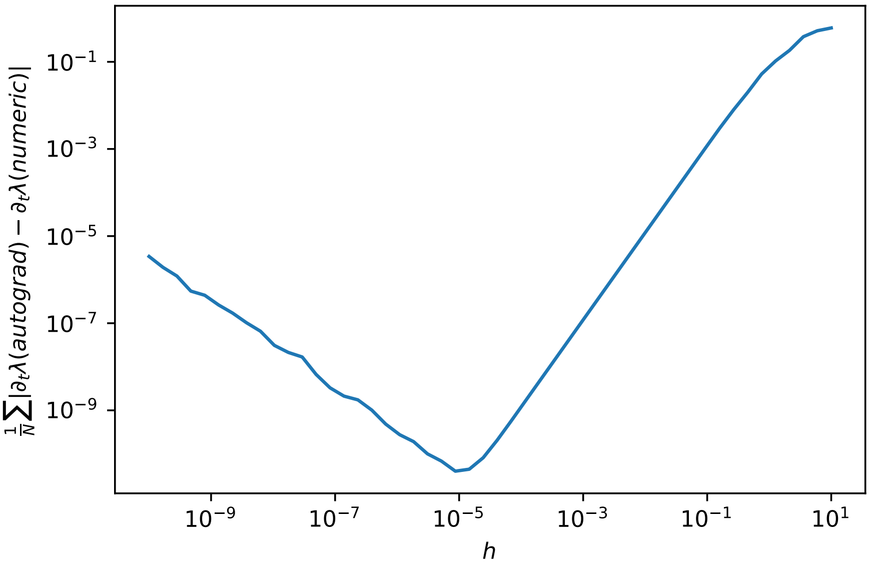

Figure 3.1: Error estimates of the derivative of the eigenvalues, λ, of the matrix $M(t) = A + tB$ as a function of the numeric differentiation step, h. The figure shows the mean of absolute differences between the solution obtained using automatic differentiation and the numerical solution dependent on h.

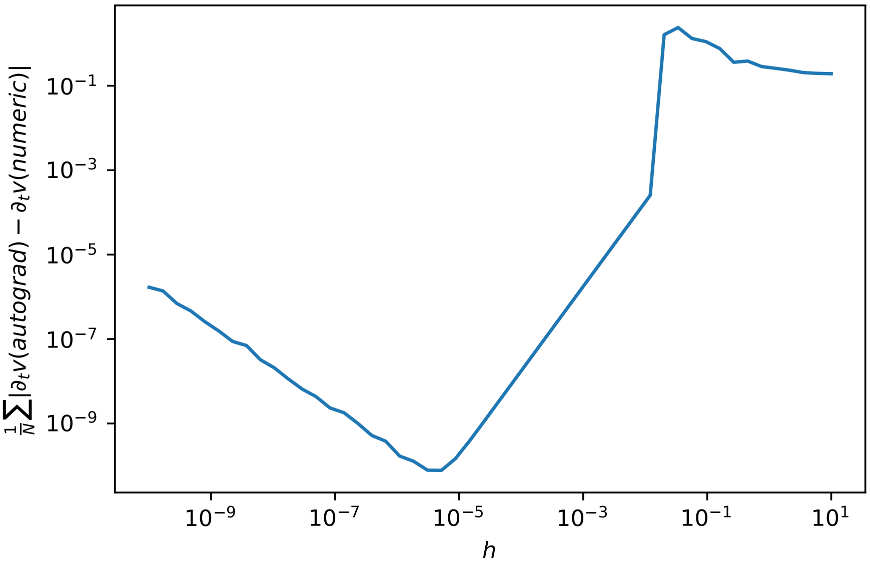

Figure 3.2: Error estimates of the derivative of the eigenvectors, v, of the matrix $M(t) = A + tB$ as a function of the numeric differentiation step, h. The figure shows the mean of the absolute difference between the solution obtained using automatic differentiation and the numerical solution dependent on h.

The use of VJPs enables clear and effective implementations of derivatives of simple functions that can be extended by the user if needed by simply writing a jacobian of an arbitrarily complicated function as either a forward-mode or reverse-mode VJP.

3.2 Python implementation of AD

The autograd package[1] is a Python package that implements the algorithmic differentiation techniques. Autograd has been developed in order to bring automatic differentiation techniques to Python that can handle Python code containing both simple computations, such as for loops, while loops, recursion, and if statements, as well as list indexing operations and many functions available in the numpy and scipy libraries [65].

The autograd package gives the user the opportunity to write complicated functions, even using recursions. For example, in the application to nuclear structure codes, it was necessary to define the associated Laguerre polynomials, $L_n^\alpha(x)$, and their derivatives using recurrence relations using the expression

$$L_n^\alpha(x) = \begin{cases} 1, & n = 0 \\ 1 + \alpha - x, & n = 1 \\ \frac{(2n-1+\alpha-x)}{n} L_{n-1}^\alpha(x) - \frac{(n+\alpha-1)}{n} L_{n-2}^\alpha(x), & n > 1 \end{cases} . \tag{3.6}$$

Only the first derivative needed to be defined in terms of autograd functions, and the higher-order derivatives are handled automatically by autograd.

---

[1] `https://github.com/HIPS/autograd`

The Python implementation of the associated Laguerre polynomials is given below.

```
import autograd.numpy as np
def Laguerre(x,n,alpha):
    if n==0:
        return 1
    if n==1:
        return 1+alpha-x
    return ((2*n-1+alpha-x)*Laguerre(x,n-1,alpha)\
           -(n+alpha-1)*Laguerre(x,n-2,alpha))/(n)
```

An example of a situation where extending autograd was necessary is the case of the derivatives of the eigenvalue problem. The function that needed to be extended is the numpy function eigh, giving the list of eigenvalues and the associated eigenvectors as two parameters. First, the function was wrapped to produce a single matrix as an output, with eigenvalues on the diagonal merged with the transition matrix V. For a $N \times N$ matrix, labeled H, this reshaping of outputs yields a function

$$\mathrm{eigh}(H)_{ij} = \begin{cases} \lambda_i \delta_{ij}, & j < N \\ V_{i,j-N}, & j \geq N \end{cases} . \tag{3.7}$$

Its derivative is then implemented as a VJP, $u^{mn}\partial_i eigh(H)_{mn}$, based on the derivative formula

$$\partial_i \mathrm{eigh}(H)_{mn} = \begin{cases} \sum\limits_{kl} V_{kn}\partial_i H_{kl} V_{ln} \delta_{jm}, & n < N \\ \sum\limits_{j\neq m}\sum\limits_{kl} V_{km}\partial_i H_{kl} V_{lj} V_{(n-N)j}\frac{1}{E_m - E_j}, & N \leq n \end{cases} , \tag{3.8}$$

where the sums have been written with no Einstein summation implied. The AD implementation of the *eigh* derivative was tested on random 5x5 matrices, A and B, merged into a function of the form $f(t) = A + tB$ compared to the numeric derivatives using a finite symmetric step size h. The test of the implementation of eigenvalue derivatives is shown in Fig. 3.1. In contrast, their corresponding eigenvector derivatives are shown in Fig. 3.2. The errors are minimal around the floating-point precision of $\sim 10^{-5}$ in both figures, as expected from numerical implementations. The Python implementation of this procedure is realized by the *@primitive* decorator and the use of function *defvjp* that connects the function of the jacobian VJP to the original function. Since the matrices for which this function is applied are assumed to be hermitian, the imaginary parts of *eigh* were removed from the implementation:

```
import autograd as au
import autograd.numpy as np
from autograd.extend import primitive,defvjp
@primitive
def eighReal(theta,H):
    h = H(theta)
    e,V = np.linalg.eigh(h)
    return np.array([np.diag(e),np.real(V)])
def eighdkReal(Eval,theta,H):
    evals, vects = np.diag(Eval[0]),Eval[1]
    dh = au.jacobian(H)(theta)
    dEs = np.array([[([vects[:,j].T@dh[:,:,i]@vects[:,j] if j==m else 0.
                        for i in range(len(theta))])
                    for j in range(len(evals)) ]
                    for m in range(len(evals))])
    Vs = np.array([np.sum([[vects[:,n]*(vects[:,m].T@dh[:,:,i]@vects[:,n])\
                            /(evals[m]-evals[n])
                            for  i in range(len(theta))]
                        for n in range(len(evals)) if m!=n],axis=0).T
                     for m in range(len(evals))]).transpose(1,0,2)
    def out(u):
        o=np.tensordot(u,np.array([dEs,np.real(Vs)]),[[0,1,2],[0,1,2]])
        return np.real(o)
    return out
defvjp(eighReal,eighdkReal)
```

Chapter 4

Stability of the manifold boundary approximation method for reducing the nuclear structure models

The framework of nuclear energy density functionals has been employed to describe nuclear structure phenomena for a wide range of nuclei. Recently, statistical properties of a given nuclear model, such as parameter confidence intervals and correlations, have received much attention, particularly in the situations where one needs to fit complex models. We apply information-theoretic methods to investigate stability of model reductions by the manifold boundary approximation method (MBAM). In an illustrative example of the density-dependent point-coupling model of the relativistic energy density functional, utilizing Monte Carlo simulations, it is found that main conclusions obtained from the MBAM procedure are stable under variation of the model parameters. Furthermore, we find that the end of the geodesic occurs when the determinant of the Fisher information metric vanishes, thus effectively separating the parameter space into two disconnected regions.

This chapter has been published in our recent paper [50].

4.1 Introduction

The nuclear energy density functional (EDF) framework is a promising, unified theoretical approach for a global description of nuclear structure phenomena. One of the successful EDFs has been the one that is based on the relativistic mean-field Lagrangian in the finite-range meson-exchange model [53], with the density-dependent meson-nucleon couplings providing an improved description of asymmetric nuclear matter [52]. Moreover, it has been found that simpler, point-coupling models [2; 3] produce comparable results to the finite-range ones, even if the point-coupling interactions are being adjusted to nuclear matter and ground-state properties of finite nuclei [4]. These density-dependent point-coupling models, however, have been shown to exhibit an exponential range of sensitivity to parameter variations, prompting the application of model reduction methods based on concepts of information geometry [28; 29].

Information geometry is an interdisciplinary field that introduces differential geometry concepts to statistical problems [20; 21] with its initial applications centered around machine learning and neural networks [18; 19]. Recently, the manifold boundary approximation method (MBAM) [22–24] has been developed to study complex and sloppy problems occurring in physics, chemistry and biology [25–27] in order to either classify or reduce complex models, such as the nuclear EDFs [28–30].

The complexity of nucleon-nucleon interaction in the nuclear medium, coupling between single-nucleon and collective degrees of freedom, and finite-size effects present obstacles to numerous attempts to establish a single theoretical framework to treat the nuclear many-body problem. The nuclear EDFs, and structure models based on them, have become a promising tool for the description of ground-state properties and low-energy collective excitation spectra of medium-heavy and heavy nuclei. A variety of structure phenomena have been successfully described using the nuclear EDF framework with a high level of global precision and accuracy over the entire chart of nuclides, and at a very moderate computational cost.

The unknown exact nuclear EDF is approximated by functionals of powers and gradients of ground-state nucleon densities and currents, representing distributions of matter, spin, isospin, momentum, and kinetic energy. A generic density functional is not necessarily microscopic, i.e., it is related to the underlying inter-nucleon interactions, but some of the most successful functionals are entirely empirical. However, one can also follow the middle way between fully microscopic and entirely empirical EDFs, and consider semi-empirical functionals that start from a microscopically motivated ansatz for the nucleonic density dependence of the energy of a system of protons and neutrons. Most of the parameters of such a functional are adjusted, in a local density approximation, to reproduce a given microscopic equation of state (EoS) of infinite symmetric and asymmetric nuclear matter, and eventually neutron matter. The remaining, usually few, terms that do not contribute to the energy density at the nuclear matter level, are then adjusted to selected ground-state data of an arbitrarily large set of spherical and/or deformed nuclei. A number of semi-empirical functionals have been developed over the last decade [5; 31–40], and very successfully applied to studies of a diversity of structure properties, from clustering in relatively light nuclei to the stability of superheavy systems, and from bulk and spectroscopic properties of stable nuclei to the physics of exotic nuclei at the particle drip lines.

In the previous studies [28; 29], the authors have used concepts from information geometry to demonstrate that nuclear EDFs are, in general, "sloppy" [22–25; 41]. The term "sloppy" refers to the fact that the predictions of nuclear EDFs and related models are really sensitive to only a few combinations of parameters (*stiff* parameter combinations) and exhibit an exponential decrease of sensitivity to variations of the remaining combinations of parameters (*soft* parameter combinations). This means that the soft combinations of parameters are only loosely constrained by the available data, and that most nuclear EDFs in fact contain models of lower effective dimensionality associated with the stiff combinations of model parameters. In Ref. [28], by employing the MBAM [24] the authors have deduced the most effective functional form of the density-dependent coupling parameters of a representative model EDF. The data used in this calculation included a set of points on a microscopic EoS of symmetric nuclear matter and neutron matter. This choice was motivated by the necessity to calculate the derivatives of observables with respect to model parameters which is, of course, much easily accomplished for nuclear matter in comparison to finite nuclei. In Ref. [29] we have extended this calculation by employing a simple numerical approximation to calculate the derivatives of observables with respect to model parameters. Thus we were able to apply the MBAM to realistic models constrained not only by the nuclear matter EoS but also by observables measured in finite nuclei. During our analysis of parametrizations in Ref. [29] we have noticed that the numerical integration of the geodesic equation could reach the manifold boundary in a finite number of integration steps, indicating the divergence of the metric tensor determinant in a particular region of the parameter space. This surprising behavior motivated an investigation of the stability of model reductions obtained by the manifold boundary approximation method (MBAM), since the divergent region might be unintentionally missed by using too large integration steps.

In this chapter, we study the stability of the MBAM with respect to the variation of the model parameters. In Sec. 4.2 we describe the numerical implementation for finding the Dirac mass and binding energies, aided by algorithmic differentiation. The results of our investigation are given in Sec. 4.3, while further applications of information geometry to nuclear EDFs are discussed in Sec. 4.4.

4.2 Illustrative calculation

The density-dependent point-coupling (DD-PC1) interaction [4] is a semi-empirical relativistic EDF that involves the point coupling [54], and has been used in many contemporary studies of nuclear structure and dynamics. The DD-PC1 functional explicitly includes nucleon degrees of freedom and considers only second-order interaction terms. Its applicability to a wide range of atomic nuclei has been demonstrated, e.g., in Refs. [55; 66].

We use the Dirac mass and energy density data shown in Table 4.1 to constrain the density-dependent coupling constants of the DD-PC1 functional, $\alpha_s(\rho)$, $\alpha_v(\rho)$ and $\alpha_{tv}(\rho)$, modeled as [28; 29]

$$\alpha_i = a_i + \left(b_i + c_i \frac{\rho}{\rho_{\text{sat}}}\right) e^{-d_i \frac{\rho}{\rho_{\text{sat}}}} \ , \quad i \in \{s, v, tv\} \ , \tag{4.1}$$

where the indices $i = s$, v, and tv correspond to the isoscalar-scalar, isoscalar-vector, and isovector-vector channels respectively, while ρ_{sat} is the saturation density. In this paper, we take a closer look at the reduced version of the model with $\alpha_{tv} = 0$ and $c_v = 0$, which results in a seven-parameter model involving a_s, b_s, c_s, d_s, a_v, b_v, and d_v.

Table 4.1: Pseudo-data for infinite symmetric nuclear matter used to compute the best-fitting solution for the energy density functional. The adopted error for the y points is 10% for energy and 2% for the Dirac mass.

index	ρ_v [fm^{-3}]	y [1]	σ_y
1	0.152	0.58	0.055
2	0.04	−6.48	0.648
3	0.08	−12.13	1.213
4	0.12	−15.04	1.504
5	0.16	−16.	1.6
6	0.2	−15.09	1.509
7	0.24	−12.88	1.288
8	0.32	−5.03	0.503

4.2.1 Numerical implementation

We solve the equation for the Dirac mass M_D, that is given by [28]

$$M_D = m + \alpha_s \rho_s \ , \tag{4.2}$$

where m is the bare nucleon mass, and ρ_s the scalar density

$$\rho_s = \frac{2}{\pi^2} M_D \int_0^{p_F} \frac{x^2 dx}{\sqrt{x^2 + M_D^2}} \ , \tag{4.3}$$

with p_F being the Fermi momentum

$$p_F(\rho_v) = \left(\frac{3}{2} \rho_v \pi^2 \right)^{\frac{1}{3}} \ . \tag{4.4}$$

The equation (4.2) is solved numerically by using the Newton-Raphson algorithm. We have also tested the Halley's method, but found no improvement of the results in accuracy.

Upon finding M_D, we compute the binding energy of symmetric nuclear matter

$$E_a = \frac{2}{\pi^2} \int_0^{p_F} \frac{x^4 dx}{\sqrt{x^2 + M_D^2}} + m(\rho_s - \rho_v) + \frac{1}{2} \alpha_s \rho_s^2 + \frac{1}{2} \alpha_v \rho_v^2 \ . \tag{4.5}$$

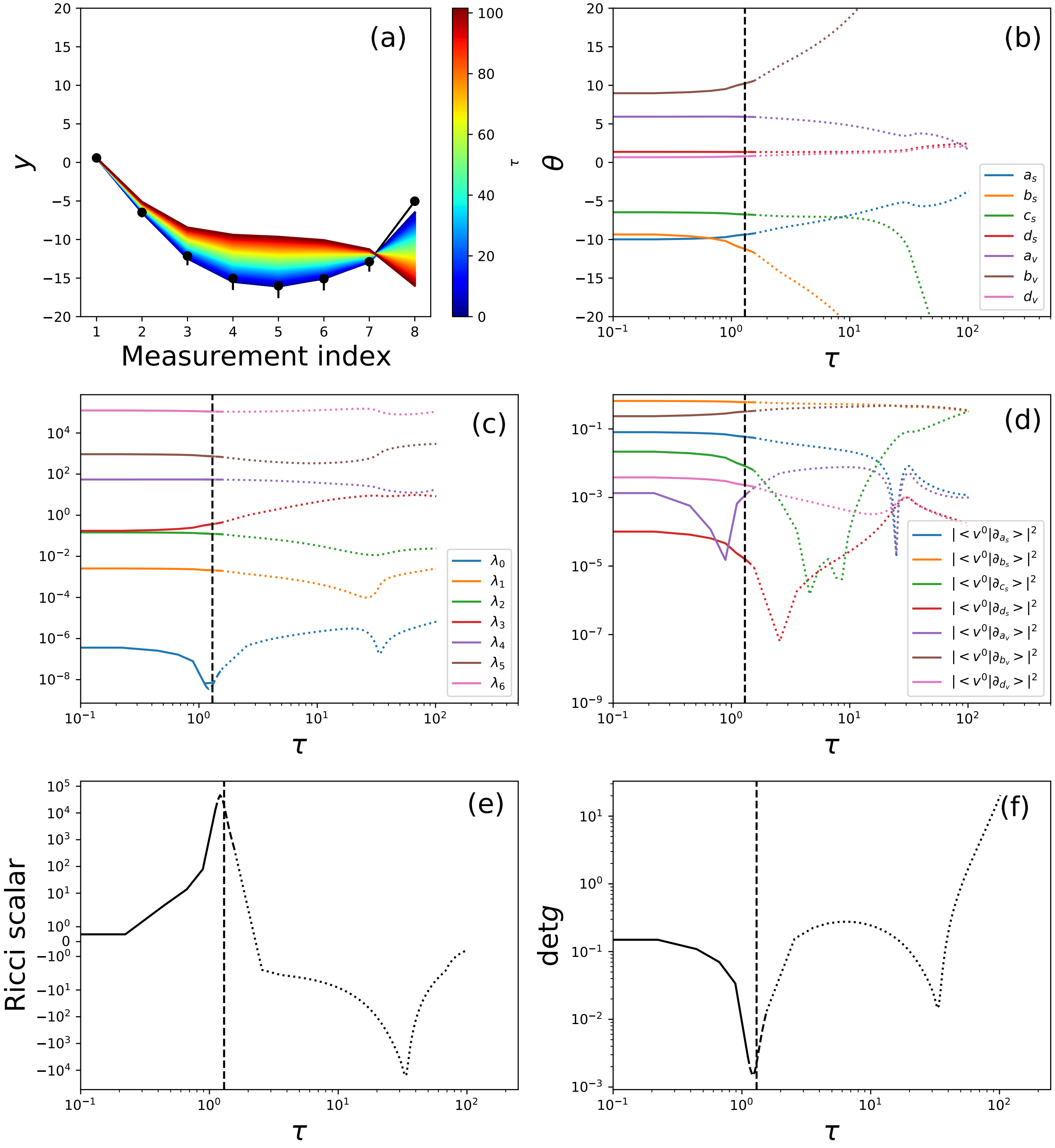

Figure 4.1: Results of extrapolating the geodesic after the $\det g = 0$ point. Shown are (a) the behavior of the evaluated model for different τ-s along the geodesic, (b) the model parameters, (c) the FIM eigenvalues as functions of τ, (d) the squares of the FIM eigenvector v^0 components, (e) the Ricci scalar, and (f) the FIM determinant along the geodesic. Solid, dashed, and dotted lines stand for, respectively, the initial odeint solutions, the linear interpolation, and the values derived using odeint starting from the endpoint of the interpolated solutions.

The best-fitting DD-PC1 parameter set is then found by computing the least-square solution to the set of measurements of M_D/m and E_a presented in Table 4.1 (see Ref. [29]). Differential equations are solved with the aid of the SciPy implementation of the ordinary differential equation integration (odeint) library [67]. These values are then used to compute the FIM and the Christoffel symbols using algorithmic differentiation implemented via the `autograd` package. We thus eliminate numerical errors due to the approximations arising from numerical differentiations.

4.3 Investigating stability of the MBAM method

In some cases, the numerical integration of the geodesic equation might slow down, or even fail. This behavior is due to the divergence of the metric tensor determinant that implicitly appears in the geodesic equation (2.7) through the metric inverse necessary for computing the Christoffel symbols [see Eq. (2.3)]. However, this divergent behavior is confined to only a small region in the parameter space, and therefore it might be easily missed by choosing too imprecise an integrator. Therefore, in Sec. 4.3.1, we investigate the impact of the size of the integration step on the MBAM procedure by artificially extrapolating the geodesic beyond the divergent region in the parameter space. Moreover, as the parameter uncertainties become larger, small perturbations to the starting point of the geodesic might influence the end result of the MBAM. In Sec. 4.3.2, we describe the impact of parameter uncertainties on the MBAM conclusions for the nuclear EDF DD-PC1 by numerical error propagation of the MBAM geodesics. Finally, in Sec. 4.3.3 we investigate the impact of using a common, physically-motivated restrictive reparametrization of the DD-PC coupling constants on the MBAM model manifold.

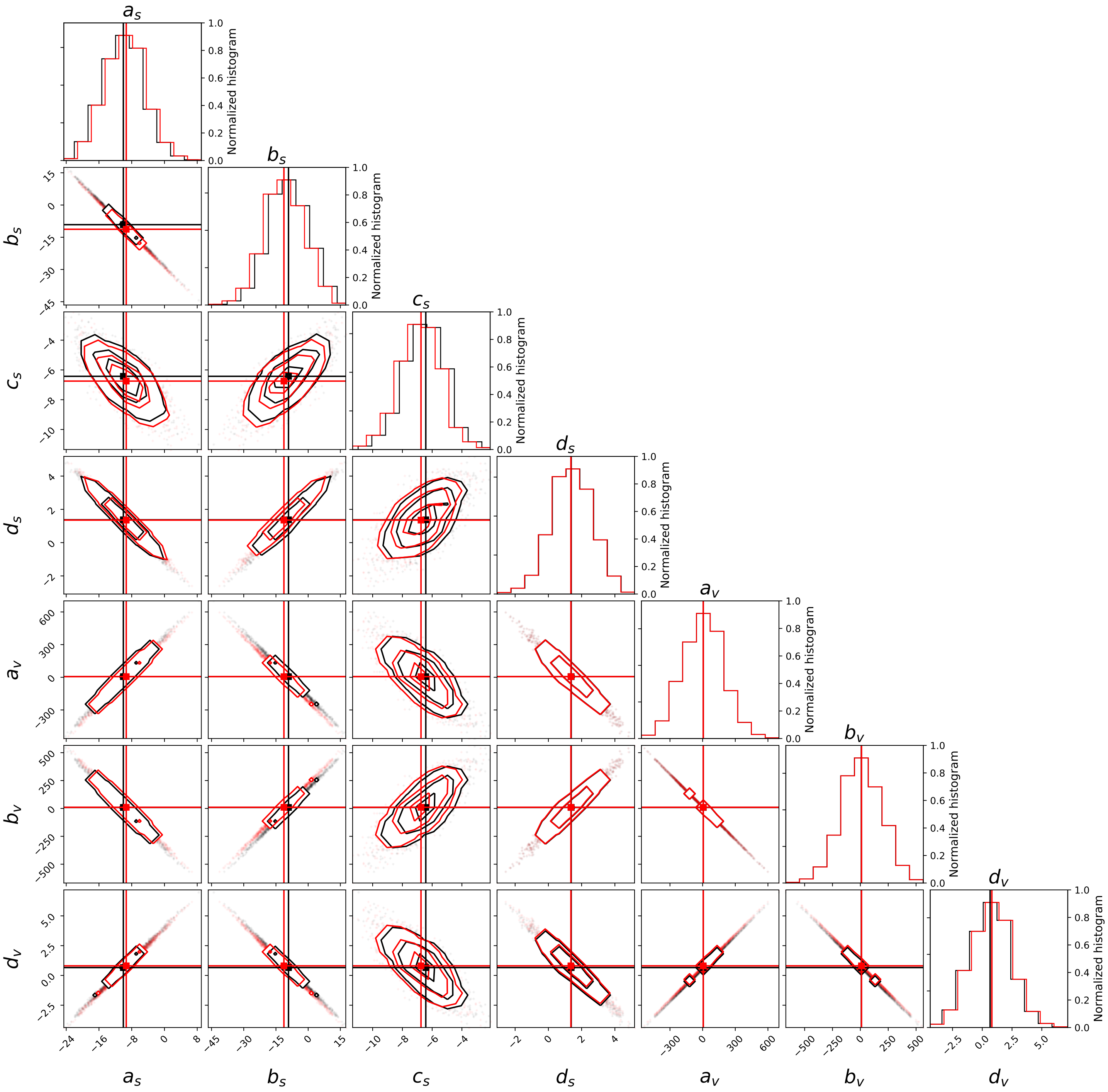

Figure 4.2: Monte Carlo simulated sample parameters using the best-fitting covariance matrix (black symbols and contours) and its propagation towards $\tau = 1.3$ along the geodesic using the Jacobi equation (4.8) (red symbols and contours).

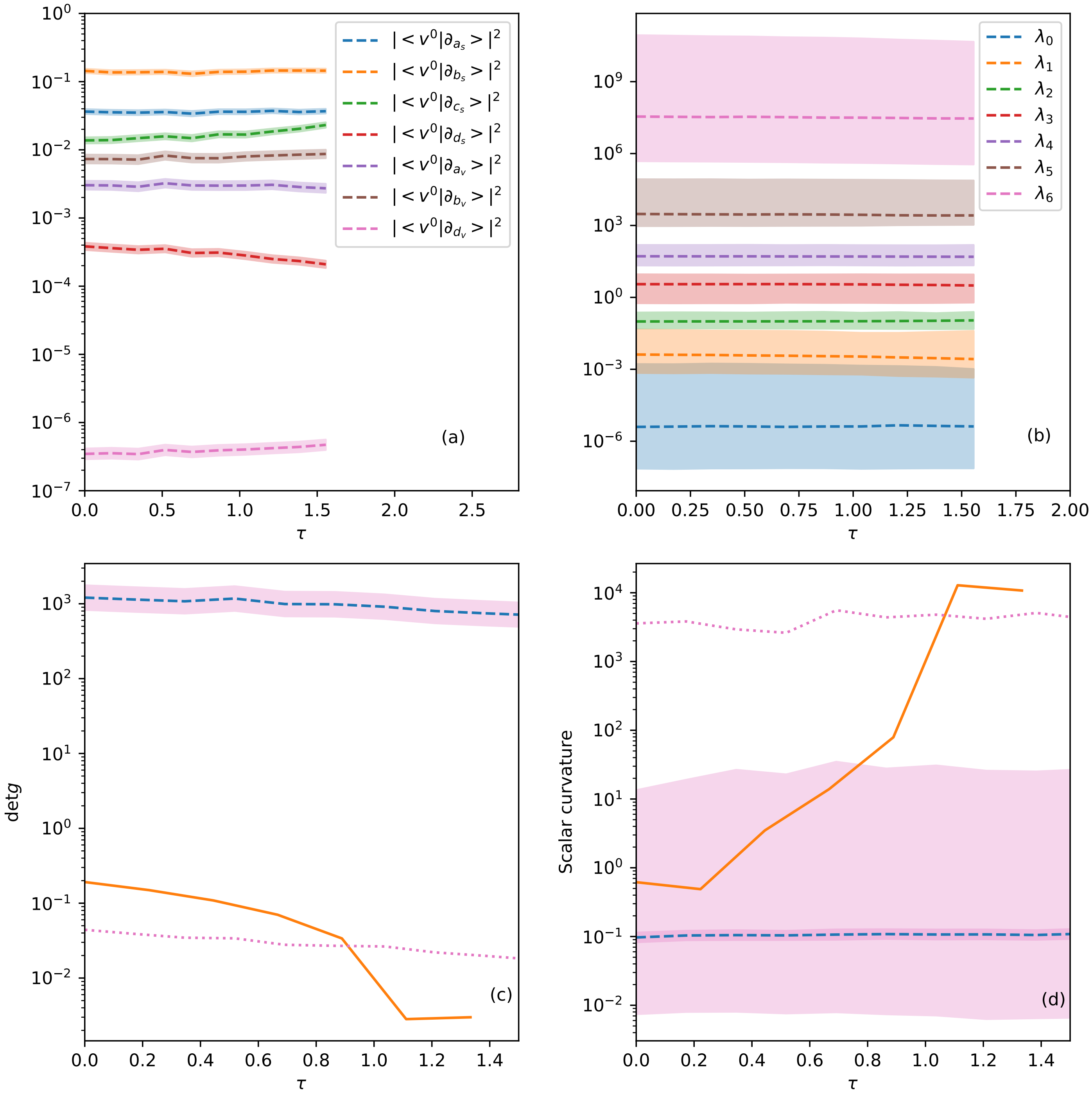

Figure 4.3: Monte Carlo simulations of uncertainty propagation using the Jacobi equation (4.8). Shown are the median and its uncertainty derived using 1300 simulated samples starting from the best fitting point. The figure shows (a) the simulated FIM v^0 eigenvector components squared, (b) FIM eigenvalues, (c) FIM determinant, and (d) scalar curvature. The shaded areas correspond to the 1σ percentile interval, while the dotted lines in panels (c) and (d) additionally show the 5-th and the 95-the percentiles, respectively. Solid orange lines in (c) and (d) stand for the respective quantities computed along the path of the MBAM geodesic.

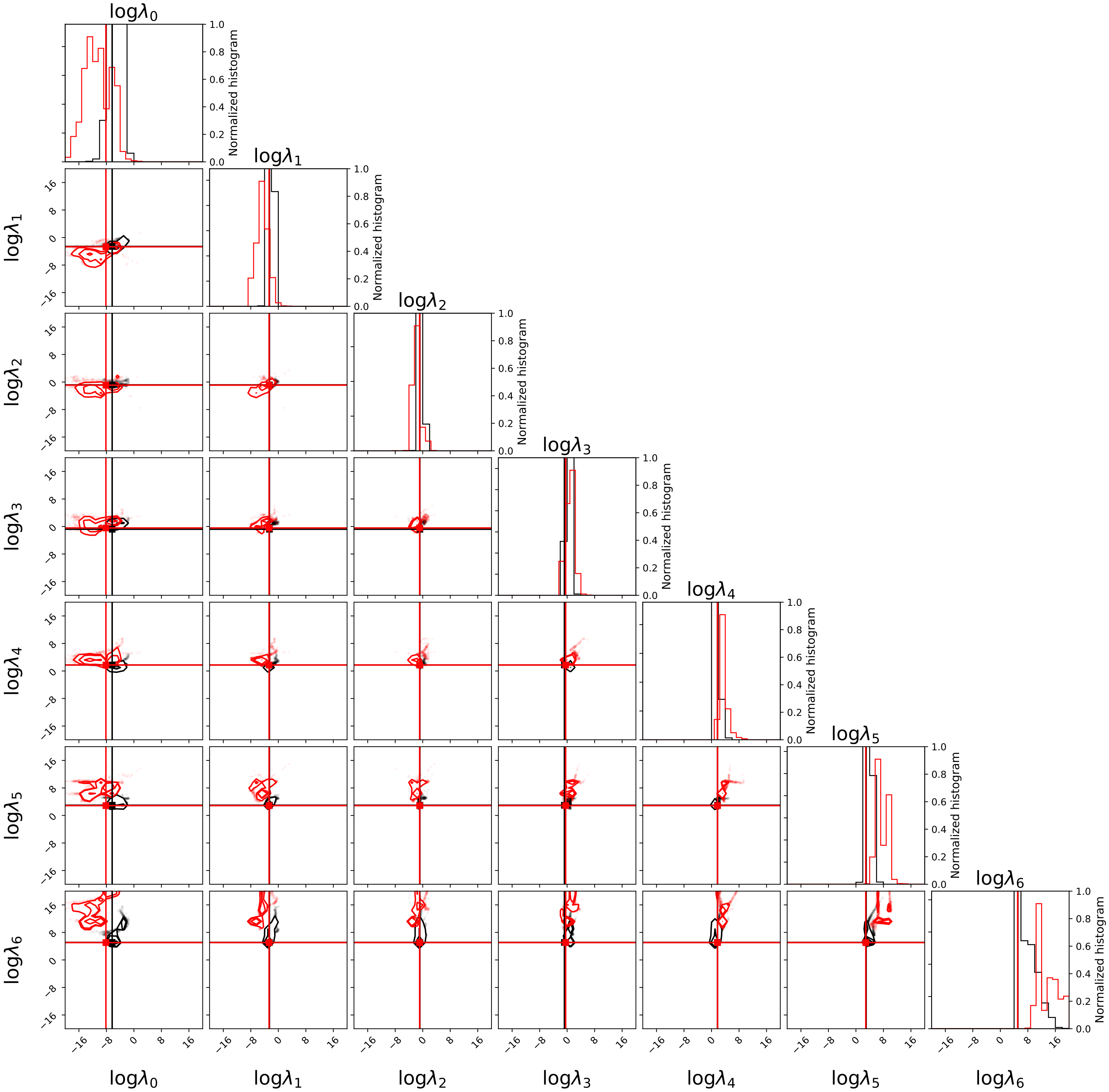

Figure 4.4: Same as Fig. 4.2, but for Monte Carlo simulated sample (base 10) logarithm of the eigenvalues of the FIM.

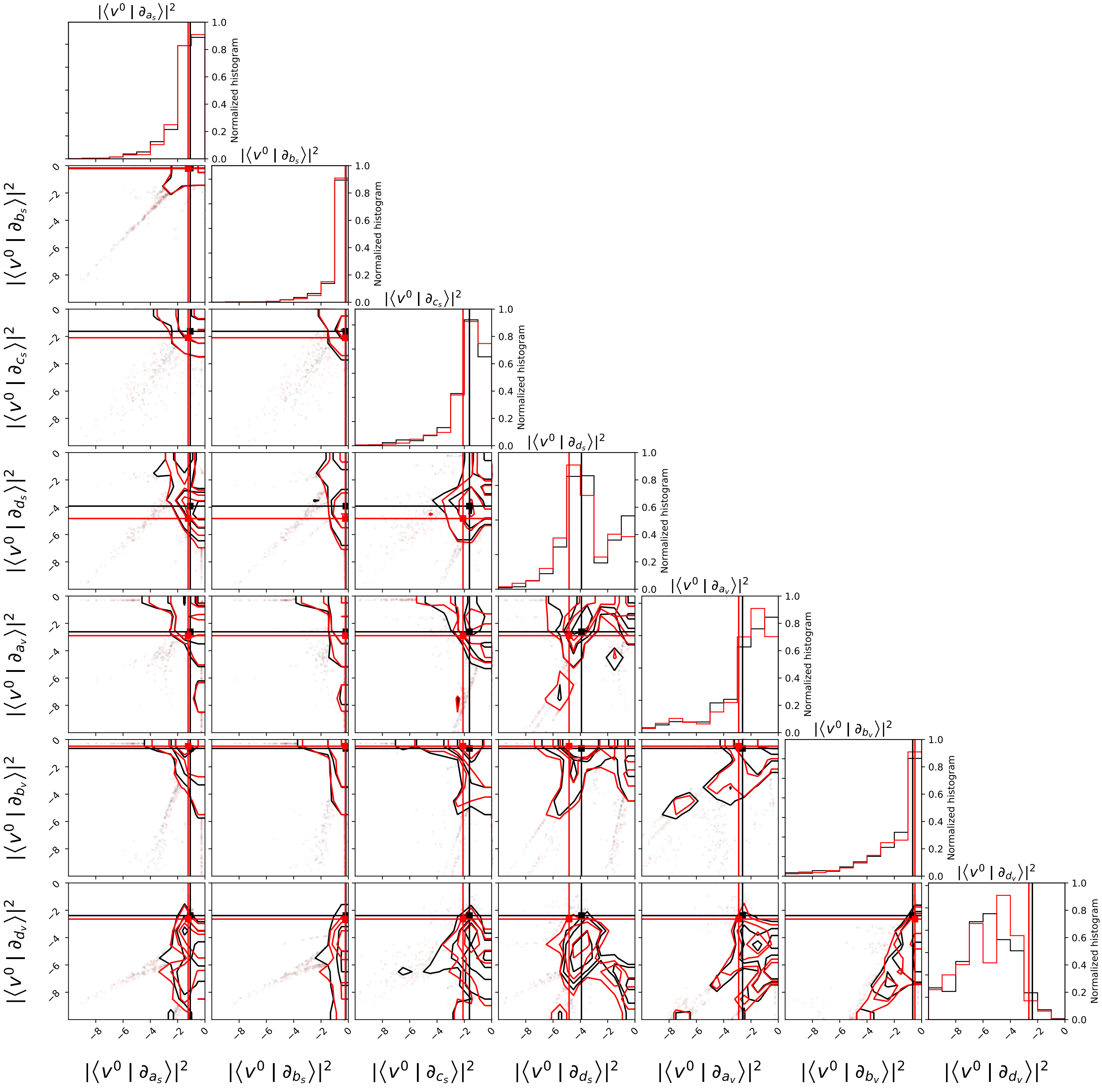

Figure 4.5: Same as Fig. 4.2, but for Monte Carlo simulated sample components of the FIM v^0 eigenvector.

4.3.1 Geodesic extrapolation

We extrapolate the geodesic by using the last point having $\det g > 0$ (labeled as τ_2) and the point before it (τ_1). We first extrapolate $\tau(t) = \tau_1(1-t) + \tau_2 t$ for $t > 0$, i.e., a straight line joining τ_1 and τ_2. We then compute the $\mathbf{p}^\mu(t)$ and $\dot{\mathbf{p}}^\mu(t)$ using their corresponding values at τ_1 and τ_2 as

$$\mathbf{p}^\mu(t) = \mathbf{p}^\mu(\tau_1)(1-t) + \mathbf{p}^\mu(\tau_2)t \tag{4.6}$$

$$\dot{\mathbf{p}}^\mu(t) = \dot{\mathbf{p}}^\mu(\tau_1)(1-t) + \dot{\mathbf{p}}^\mu(\tau_2)t \tag{4.7}$$

This procedure produces a linear extrapolation of the geodesics in the region where the geodesic equation does not hold because $\det g = 0$. The variable t is just an interpolation parameter, not connected to τ, so $\dot{\mathbf{p}}$ is not coupled as $d\mathbf{p}/d\tau$ in this region. We find that one can safely continue integrating the geodesic equation after $t = 2$, where there are no more singularities along the path.

In Fig. 4.1, the resulting model parameters along the extended geodesic (a), the corresponding model evaluation (b), the FIM eigenvalues (c), the v^0 eigenvector (d), the Ricci scalar (e), and the metric determinant (f) are shown. After the $t = 2$ point along the extrapolated geodesic, the metric tensor determinant starts to rise again. In the same figure, the linearly extrapolated geodesic, corresponding to the small region region $\tau \in [\tau(t = 1), \tau(t = 2)]$, is shown with dashed lines. The extrapolated geodesic computed using MBAM continuation starting from the point $\tau(t = 2)$ is shown with dotted lines. The initial odeint solutions (solid lines), which produce results for a few points after $\tau = 1.3$, differ significantly from the interpolated solution, indicating numerical problems due to singularity. Upon restarting the odeint procedure after the singular region, we find that the MBAM solution yields different contributions to the v^0 eigenvector, indicating an equal contribution of ∂_{b_s}, ∂_{c_s}, and ∂_{b_v} directions, while before $\tau = 1.3$, the MBAM method finds that the most significant contribution is from ∂_{b_s}. The Ricci scalar diverges at $\tau \sim 1.3$, but starts to fall and change signs at $\tau > 1.3$. Since the Ricci scalar is related to the volume element, its divergence to positive values would produce a compressed region of the parameter manifold, that begins to expand after the singularity.

The conclusion drawn from the results given in Fig. 4.1 is that one must be careful with the models where the metric tensor determinant shows significant variations, as choosing too big steps for the odeint integrator might result in “skipping” to another portion of the parameter space and continuing along it. This yields completely different contributions to the FIM eigenvector corresponding to its smallest eigenvalue and hence might lead to a completely different model reduction than expected from the simple MBAM case.

4.3.2 Parameter uncertainties

Further extension of the basic model might be the propagation of its parameter uncertainties, and this can be facilitated by looking into how the uncertainties of the best-fitting parameters propagate along the geodesics. For this purpose, we perform Monte Carlo simulations. To analyze the error propagation one would have to compute the geodesic equation many times, which is not cost-efficient. We, therefore, adopt a simplified approach that makes use of the Jacobi equation, which computes differences $\delta\mathbf{p}$ between neighboring geodesics along the already computed MBAM geodesic.

We use the covariance matrix Σ to produce Monte Carlo simulations of $\delta\mathbf{p}$ from the normal distribution, $\delta\mathbf{p} \sim \mathcal{N}(0, \Sigma)$. For each simulated $\delta\mathbf{p}$, we compute its propagation by using the Jacobi equation

$$\delta\ddot{\mathbf{p}}^{\mu} + R^{\mu}_{\alpha\nu\beta}\dot{\mathbf{p}}^{\alpha}\dot{\mathbf{p}}^{\beta}\delta\mathbf{p}^{\nu} = 0 \ . \tag{4.8}$$

We find 1300 points to sample the DD-PC1 parameter space reasonably well. Figure 4.2 shows the distributions of the parameters at the beginning (denoted by black symbols and contours) and at $\tau = 1.3$ (red symbols and contours). These two distributions are almost identical since the simulated parameters are more dispersed than the gradual changes in parameter values along the geodesic.

Even though the parameter uncertainties in the full model are large, we can estimate the error on the eigensolutions of the FIM along the geodesic. We do this by computing the FIM for every simulated point propagated along the best-fitting geodesic to various values of τ using the Jacobi equation. The results of this procedure are shown in Fig. 4.3. The top panels show the median and the corresponding 1σ confidence interval of the eigensolutions, computed using the 16th and the 84th percentile. The simulated FIM v^0 eigenvector components squared are shown in panel (a) and FIM eigenvalues in panel (b) for each τ. We see that, while the results using the simulated sample are consistently ordered when compared to the MBAM solution, there is a small offset between the median and the MBAM solution. The panels (c) and (d) of Fig. 4.3 show the median and the 1σ confidence interval for the FIM determinant and the scalar curvature, respectively. The simulated scalar curvature and the metric determinant along the geodesic show a larger variation in their values along the geodesic. In these panels we additionally show the FIM determinant and scalar curvature along the best-fitting geodesic by the solid orange lines. There is a large discrepancy between the behavior of the median of the simulated quantities and the behavior of the quantities along the best-fitting geodesic. In panel (c) [(d)] of Fig. 4.3, we see that these quantities along the best-fitting geodesic are comparable to the 5th (95th) percentile of $\det g$ (scalar curvature), shown as dotted lines. This behavior indicates that only the geodesics starting at the vicinity of the best-fitting point encounter the region corresponding to $\det g = 0$.

Furthermore, in Figs. 4.4 and 4.5 we show, respectively, the distributions of eigenvalues and components of v^0 at the beginning and at the end of the geodesic. These large differences in eigenvalues and eigenvector components propagating along the geodesic are in stark contrast to the parameter values in Fig. 4.2. The discrepancies presented in Figs. 4.3, 4.4 and 4.5 can be explained by the sensitivity of the FIM eigenproblem to small changes in DD-PC1 parameters, since diagonalization results are not expected to change linearly with inputs. We conclude that the offset is due to the non-Gaussianity of the distribution of eigenvalues and v^0 components, which arises even though the parameters were sampled using the normal distribution. Even though there is a change in the shape of these distributions, the overall qualitative MBAM conclusions remain the same along the geodesic.

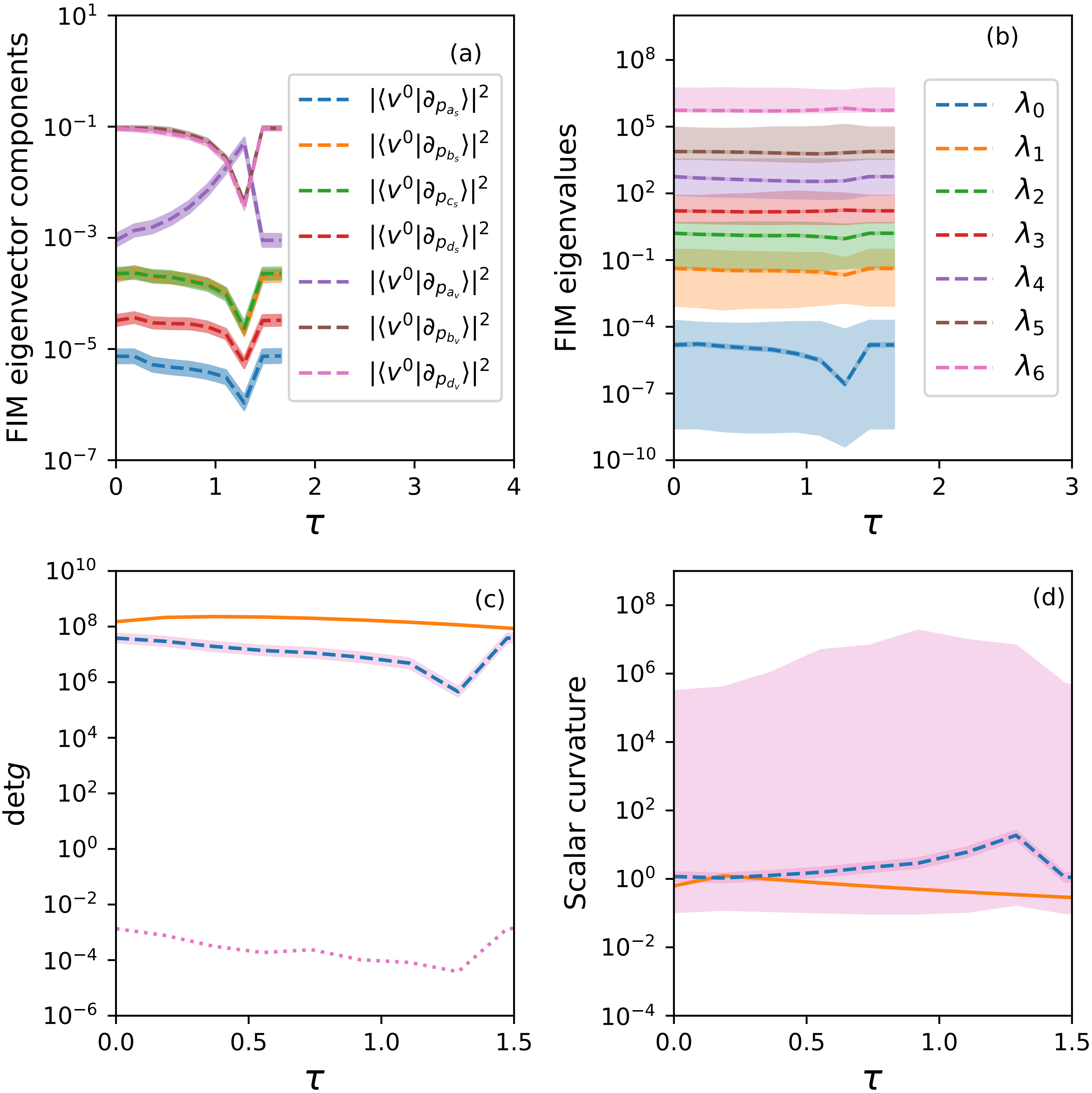

Figure 4.6: Same as Fig. 4.3, but for the reparametrized model described in Sec. 4.3.3.

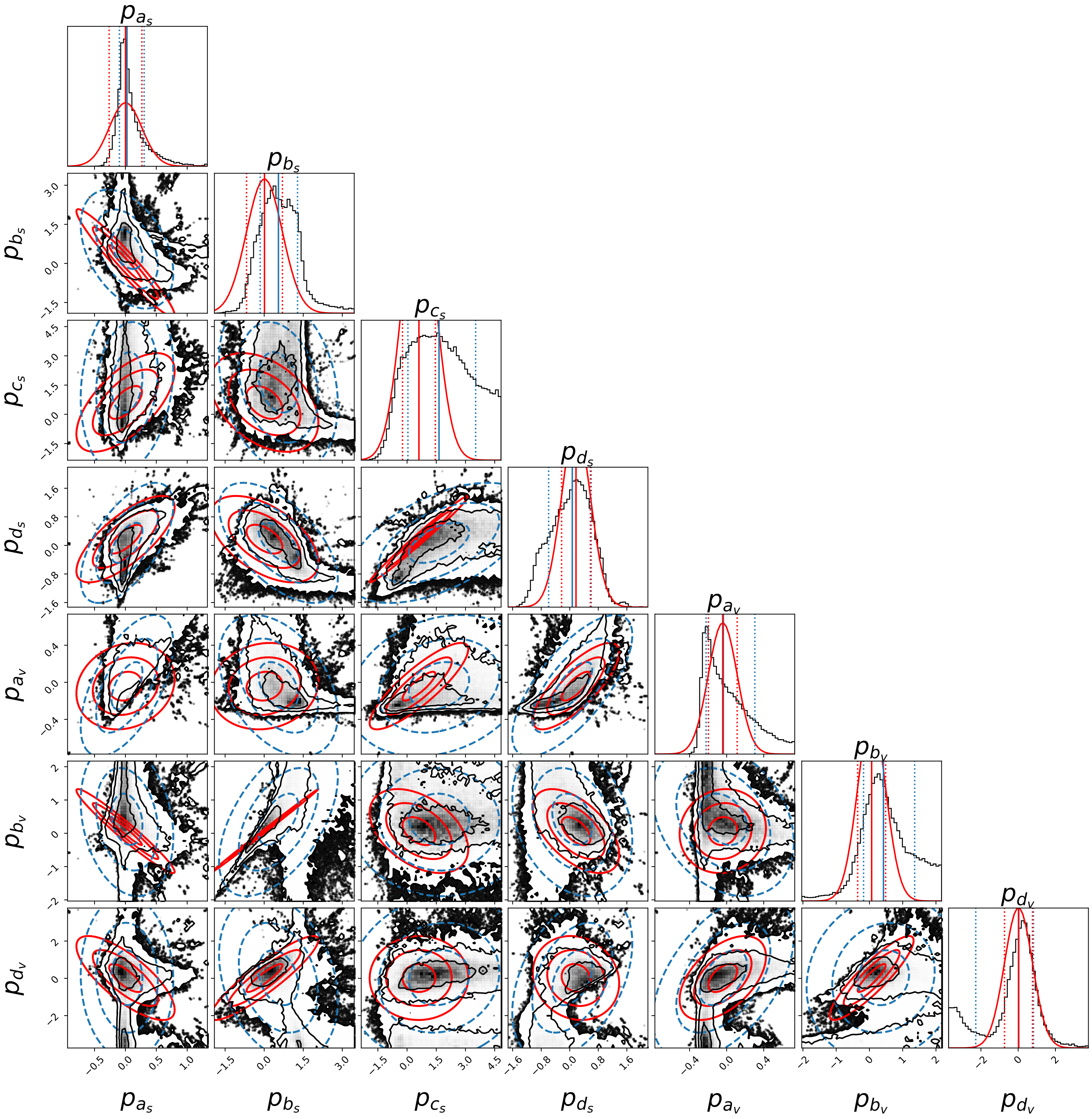

Figure 4.7: Monte Carlo simulations of posterior distributions of the error estimates for the reparametrized model, based on the MCMC algorithm. The figure shows the 1σ, 2σ and 3σ covariance ellipses in red, as estimated from the FIM inverse, and the estimates of the covariance ellipses based on the MCMC sample points in blue.

4.3.3 Model reparametrization

The authors of Ref. [28] have considered an exponential reparametrization of the seven-parameter DD-PC1 coupling constants centered at their best-fitting values [4]. This reparametrization transformation can be schematically represented as a vector

$$\mathbf{p}(\widetilde{\mathbf{p}}) = \begin{pmatrix} a_s(p_{a_s}) \\ b_s(p_{b_s}) \\ c_s(p_{c_s}) \\ d_s(p_{d_s}) \\ a_v(p_{a_v}) \\ b_v(p_{b_v}) \\ d_v(p_{d_v}) \end{pmatrix} = \begin{pmatrix} a_{s,\mathrm{bf}}\, e^{-p_{a_s}} \\ b_{s,\mathrm{bf}}\, e^{-p_{b_s}} \\ c_{s,\mathrm{bf}}\, e^{-p_{c_s}} \\ d_{s,\mathrm{bf}}\, e^{-p_{d_s}} \\ a_{v,\mathrm{bf}}\, e^{-p_{a_v}} \\ b_{v,\mathrm{bf}}\, e^{-p_{b_v}} \\ d_{v,\mathrm{bf}}\, e^{-p_{d_v}} \end{pmatrix}, \tag{4.9}$$

where $\widetilde{\mathbf{p}}$ indicates multivariate distribution of parameters $p_{a_s}, \cdots, p_{d_v}$, and the quantities such as $a_{s,\mathrm{bf}}$, $b_{s,\mathrm{bf}}$, etc. stand for the best-fitting parameter values. The exponential form of the coupling constants is chosen by the constraints (i) that the new parameters in the geodesic equation are dimensionless, and (ii) that the exponential form prevents the coupling functions a_s and a_v from changing sign along the geodesic path, thus confining them in the region described by the inequalities $\alpha_s < 0$ and $\alpha_v > 0$. Using these constraints the scalar mean-field potential remains attractive and the vector mean-field repulsive for all allowed parameter values [28].

We repeat the Monte Carlo analysis described in Sec. 4.3.2 for the reparametrized model. The resulting error estimates are shown in Fig. 4.6 in the same manner as in Fig. 4.3. By comparing the two figures panel-by-panel, we conclude that both methods produce MBAM geodesics that are stable under perturbations, even though the two FIMs do not behave in the same way along their respective geodesics. The reparametrized FIM determinant and the Ricci scalar change gradually, compared to the initial model.

One may ask whether this discrepancy is due to using a too simplistic description of the reparametrized distributions. We then employ the Bayesian statistics to check whether the multivariate distribution $\widetilde{\mathbf{p}}$ has pronounced non-Gaussian features. To this end, we use the Markov chain Monte Carlo (MCMC) technique to sample the χ^2 posterior distribution, as implemented in the package `emcee` [68]. In Fig. 4.7 we show the behavior of the chosen 200

Markov chains as two-dimensional sections of the parameter space. The chains have been run for a long enough time to avoid the initial “burn-in” phase characteristic of the algorithm during which they follow mostly the (uniform) prior distribution instead of sampling the χ^2 posterior distribution. From the fact that the classical covariance ellipses (represented by red contours in Fig. 4.7) are well-aligned with the MCMC estimates, we conclude that one can proceed with using the simple Monte Carlo Gaussian mock sample for error propagation instead of the computationally more expensive Bayesian MCMC mock sample.

The theoretical argument for the discrepancy between the two geodesics is based on the properties of the applied transformation. Since the exponential transformations are not bijections, the geodesics on the manifold spanned by $\widetilde{\mathbf{p}}$ need not have the same behavior as the geodesics on the manifold spanned by $\mathbf{p}$. To better understand the connection between these two geodesics, we derive the FIM determinant on the $\widetilde{\mathbf{p}}$-manifold by using the transformation of Eq. (4.9),

$$\det g(\widetilde{\mathbf{p}}) = a_s^2(p_{a_s})b_s^2(p_{b_s})c_s^2(p_{c_s})d_s^2(p_{d_s})a_v^2(p_{a_v})b_v^2(p_{b_v})d_v^2(p_{d_v}) \det g(\mathbf{p}(\widetilde{\mathbf{p}})) \ . \tag{4.10}$$

The determinant of the metric is not an invariant quantity under reparametrizations, hence the additional multiplicative scaling is required. Equation (4.10) shows that, if the value of $\det g$ approaches zero for particular values of $\mathbf{p}$, both geodesics terminate. However, additional singularities appear if any of the coupling constants is allowed to change sign along a particular geodesic. In contrast to the FIM determinant, the Ricci scalar is not affected by reparametrizations. The scalar curvature distributions for different points on the geodesic in Fig. 4.6 (d) do not have the same values as those in Fig. 4.3(d). The effects of reparametrizations on the scalar curvature can be clearly seen from the comparison between these figures.

The general conclusion is, therefore, that the MBAM method is sensitive to the way the reparametrization is made, as has been shown above in the case of the reparametrization tied to domain restrictions. This is related to the fact that different reparametrizations do not lead to the same, but similar, models describing the common physical problem. Since the EDF has an arbitrarily-chosen functional form, there is no *a priori* way of identifying which parametrization is optimal. This sensitivity only emphasizes the fact that different reparametrizations may describe similar, but inherently different, physical models. Choosing a particular EDF parametrization is equivalent to choosing a particular range model parameters can take.

4.4 Chapter summary

Methods of information geometry have been applied to investigate the stability of reducing the nuclear structure models. We have constrained the error estimates of the manifold boundary approximation method (MBAM) solutions by means of the Monte Carlo simulations. In the illustrative application to the DD-PC1 model of the nuclear EDF, it has been found that the main conclusions obtained by using the MBAM method are stable under the variation of the parameters within the 1σ confidence interval of the best-fitting model. Moreover, we have found that the end of the geodesic occurs when the determinant of the FIM approaches zero, thus effectively separating the parameter space into two disconnected regions.

Further applications of information geometry to nuclear EDFs could be analyzing possible phase transitions in models of finite nuclei using scalar curvature and their impact on nuclear properties. The analysis could even be expanded to include an extended temperature-dependent model or to look for model instabilities. It would be worth investigating whether information-theoretic optimizations, could accelerate computer codes to solve nuclear many-body problems. Such 2nd-order optimization algorithms, like the natural-gradient descent, find optimal solutions by taking optimization steps in the parameter space informed by the behavior of the FIM.

Chapter 5

Classical and Bayesian error analysis of the relativistic mean-field model for doubly-magic nuclei

Even though nuclear energy density functionals (EDFs) have already enabled the description of nuclear structure phenomena for many different nuclei, statistical properties of nuclear EDF models have only recently become a topic of research. Parameter estimation of such complex models presents a difficult task due to the fact that error estimates of the model parameters cover an exponential range of values, often requiring model reduction techniques. In the previous chapter, published as our recent paper [M. Imbrišak and K. Nomura, Phys. Rev. C **107**, 034304 (2023)], we have investigated the stability of model reductions, focusing on the Manifold Boundary Approximation Method applied to the nuclear density-dependent point-coupling model of infinite nuclear matter. In this chapter, we extend our information-geometric statistical analysis to the considerably more numerically challenging case of finite nuclei.

This chapter is a new article in preparation.

5.1 Introduction

The nuclear energy density functionals (EDFs) are a widely-used framework for describing nuclear structure phenomena. Many such EDFs are based on the relativistic mean-field Lagrangian (RMF) in the finite-range meson-exchange model [53]. The density-dependent meson-nucleon couplings have been successfully applied in this framework to describe asymmetric nuclear matter [52]. Alternatively, since the exchange of heavy mesons cannot be resolved at low energies, the self-consistent relativistic mean-field framework can be formulated in terms of point-coupling (PC) nucleon interactions. This approach yields comparable results to the meson-exchange coupling approach for finite nuclei [2; 3]. For example, the successful phenomenological finite-range interaction DD-ME2 was mapped to the PC framework by relating the strength parameter of the isoscalar-scalar derivative term to different values of the mass of the phenomenological σ meson in the DD-ME2 model [4]. The resulting 'best-fit model' [such as the DD-PC1 functional, see, e.g., 42] required the fine-tuning of the density dependence of the isoscalar-scalar and isovector-vector interaction terms to nuclear matter and ground-state properties of finite nuclei.

The issue of uncertainty quantification and error propagation in nuclear EDFs has recently attracted attention, focusing on the study of error estimates by statistical analysis [10; 11], assessment of systematic errors [8; 9], and correlation analysis [9; 43]. However, the statistical analysis is more challenging for the PC models since they were found to exhibit an exponential range of sensitivity to parameter variations [42]. This behavior was found to be a feature of *sloppy* models - models that depend only on a few stiffly constrained combinations of the parameters [44].

Recent advancements in the understanding of the behavior of sloppy models [23; 25; 26; 45] yielded new approaches to analyzing sloppy models, such as the manifold boundary approximation method [MBAM, 24]. MBAM is a systematic procedure for reducing model sloppiness by constructing progressively less sloppy lower-dimensional models from an initial sloppy higher-dimensional model. This construction is based on the concepts from information geometry - an interdisciplinary field that introduces differential geometry concepts to statistical problems [20; 21].

MBAM has already been used to systematically construct effective nuclear density functionals of successively lower dimensions and smaller impact of sloppiness. This was illustrated on the DD-PC1 functional evaluated for pseudo-data for infinite symmetric nuclear matter [28]. In Ref. [29] we extended this analysis to calculate the derivatives of observables with respect to model parameters, and we were able to apply the MBAM to realistic models constrained not only by the pseudo-data related to the nuclear matter equation of state but also by observables measured in finite nuclei. In our recent paper [50], we investigated the overall stability of the MBAM procedure applied in the reduction of nuclear structure models using methods of information geometry and Monte Carlo simulations. In the illustrative application to the DD-PC1 model of the nuclear EDF, we found that the main conclusions obtained by using the MBAM method are stable under the variation of the parameters within the 1σ confidence interval of the best-fitting model.

In contrast to the simple case of infinite nuclear matter, where one would have to solve only a simple iterative procedure to obtain the Dirac mass and binding energy, finite nuclei require a careful description of the nuclear many-body problem. Broadly speaking, statistical analysis can either be performed in the Bayesian framework - by employing elaborate Monte Carlo simulations or in the 'classical' framework, found by computing the Fisher information matrix (FIM) and its inverse (the covariance matrix) from the chosen statistical model [see, e.g., 9]. The latter approach should, in principle, be less time-consuming than running a large Monte Carlo simulation. However, when computing the FIM, one has to constrain the first derivatives of the chosen statistical model, either numerically or analytically. Attempting a simple extension of existing implementations of RMF fortran codes [46–49] would introduce uncertainties due to employing numerical differentiation. We have therefore decided to implement a simple proof-of-concept version of a finite nucleus code in Python, in which well-tested libraries for algorithmic differentiation (AD) exist.

The analysis presented below is based on a procedure for determining the RMF binding energies, starting from a simple and widespread [48; 49] assumption of a Woods-Saxon potential, often used to compute the starting point for density-dependent potentials. This paper compares numerically estimating parameter errors using a chosen Bayesian statistical technique - the

Markov chain Monte Carlo (MCMC) to the faster method of directly determining the covariance matrix without sampling using the AD-determined FIM.

The chapter is organized as follows. In Sec. 5.2, we give an overview of the RMF procedure that was implemented, and in Sec. 5.3, we describe the inputs used for our Python routines. In Sec. 5.4 we present the results of our statistical analysis.

5.2 Numerical implementation of the RMF procedure

In Sec. 5.2.1 we describe the description of the matrix elements for the Dirac equation for the proton and neutron single-particle energies in the spherical system, and in Sec. 5.2.2, we give the description of the functional form of the Woods-Saxon potential that was implemented in our python codes.

5.2.1 The spherical system

The procedure is based on solving the Dirac equation for the single-particle energies for protons and neutrons in the spherical system. First, the single-particle wavefunction is decomposed into the isospin wavefunction, $\chi_{t_i}(t)$, the spin wavefunction, $\chi_{1/2}(s)$, the angular momentum wavefunction, $Y_l(\theta, \phi)$, and two spinor radial components, $f(r)$ and $g(r)$. Due to symmetry considerations, the solutions are separable in terms of the total angular momentum, j, and parity, π, yielding the following relations:

$$l(j,\pi) = j + \pi/2 \tag{5.1}$$

$$\widetilde{l}(j,\pi) = j - \pi/2 \tag{5.2}$$

$$\kappa(j,\pi) = \pi(j + 1/2)\ . \tag{5.3}$$

In practical calculations, the maximal radial quantum number needs to be truncated to obtain finite matrices. The maximum radial quantum number for the expansion of radial functions f and g (n_{max} and $\widetilde{n}_{max}$, respectively) are determined as functions of the final major shell quantum number N_F. The value of the maximal radial quantum number of the function g is greater than the maximal value for f to avoid spurious solutions. These states of a high radial quantum number close to the Fermi surface arise from the lack of coupling for the $f_{n_{max}}$ state to the g states through the $\sigma \cdot \nabla$ term when a truncation of the quantum number is applied [46–48]

$$n_{max} = \frac{N_F - l(j,\pi)}{2} \tag{5.4}$$

$$\widetilde{n}_{max} = N_F + 1 \ . \tag{5.5}$$

In this separation, a joint spin and angular momentum quantum numbers, $|ljm\rangle$, are represented with the two-dimensional spinor

$$\Phi_{ljm}(\theta,\phi,s) = [\chi_{1/2}(s) \otimes Y_l(\theta,\phi)]_{jm} \ . \tag{5.6}$$

The full wavefunction can then be written as

$$\psi(r,\theta,\phi,s,t) = \begin{pmatrix} f(r)\Phi_{ljm}(\theta,\phi,s) \\ ig(r)\Phi_{\widetilde{l}jm}(\theta,\phi,s) \end{pmatrix} . \tag{5.7}$$

After separating the isospin, spin, and angular momentum components, one can use the simplified Hamiltonian for a single (j,π) block for protons and neutrons, whose solution depends only on the radial coordinate

$$\psi_{j\pi}(r) = \begin{pmatrix} f_{j\pi}(r) \\ ig_{j\pi}(r) \end{pmatrix} . \tag{5.8}$$

Both f and g functions are expanded using the relativistic quantum harmonic oscillator basis

$$R_{n,l} = N_{n,l} L_n^{l+1/2}(\xi^2)\xi^l e^{-\xi^2/2} \ , \tag{5.9}$$

where the radial coordinate has been rescaled to a dimensionless quantity ξ using the scaling parameter $b_0 = \sqrt{1.011A^{1/3}}$. The expansion includes a finite range of radial quantum numbers that are different for f and g functions

$$\begin{pmatrix} f \\ g \end{pmatrix} = \begin{pmatrix} \sum_{n}^{n_{max}} f_n R_{n,l} \\ \sum_{\widetilde{n}}^{\widetilde{n}_{max}} g_{\widetilde{n}} R_{\widetilde{n},\widetilde{l}} \end{pmatrix} . \tag{5.10}$$

The limits n_{max} and $\widetilde{n}_{max}$ are dependent on the total quantum number N_F and angular momentum.

For each (j, π) block, the Dirac equation is solved using the effective mass M and potential V. The aforementioned ansatz, $\psi = (f(r), ig(r))$, yields the following matrix equation

$$\begin{pmatrix} V + M - m & \hbar c\left(\partial_r - \frac{\kappa-1}{r}\right) \\ -\hbar c\left(\partial_r + \frac{\kappa+1}{r}\right) & V - M - m \end{pmatrix} \begin{pmatrix} f \\ g \end{pmatrix}_{j\pi} = \epsilon \begin{pmatrix} f \\ g \end{pmatrix}_{j\pi} . \tag{5.11}$$

Using the relativistic harmonic oscillator basis introduced in Eq. 5.9, this matrix equation can be structured as

$$\begin{pmatrix} A & B^T \\ B & -C \end{pmatrix} \begin{pmatrix} f_1 \\ \vdots \\ g_{\widetilde{n}_{max}} \end{pmatrix} = \epsilon \begin{pmatrix} f_1 \\ \vdots \\ g_{\widetilde{n}_{max}} \end{pmatrix}, \tag{5.12}$$

using three matrices $A_{nn'}$, $B_{\widetilde{n},n'}$ and $C_{\widetilde{n},\widetilde{n}'}$:

$$A_{n,n'} = \int_0^\infty r^2 dr R_{n,l} R_{n',l} (V + M - m) \tag{5.13}$$

$$B_{\widetilde{n},n'} = \hbar c \int_0^\infty r^2 dr R_{\widetilde{n},\widetilde{l}} \left(-\partial_r - \frac{\kappa + 1}{r}\right) R_{n',l} \tag{5.14}$$

$$C_{\widetilde{n},\widetilde{n}'} = \int_0^\infty r^2 dr R_{\widetilde{n},\widetilde{l}} R_{\widetilde{n}',\widetilde{l}} (M + m - V). \tag{5.15}$$

Once the wavefunctions are known, the pairing is introduced as an additional weight to the density of each eigenstate, v_i^2, as outlined in Sec. 1.2 using Eq. 1.27.

5.2.2 The Woods-Saxon potential

We apply the finite nucleus procedure to the simple case of the Woods-Saxon potential. The Woods-Saxon potential is also the first step for more complex density-dependent potentials. The shape of the Woods-Saxon potential is known, and this potential does not depend on the nucleon densities. Therefore, in contrast to density-dependent potentials, the procedure need not be run iteratively, reducing computational complexity for various numerical tests.

The shape of the potential has been adapted from [51], the authors of which developed a relativistic equivalent of the simple Woods-Saxon potential. In their model, a set of twelve parameters was used to constrain the shape of the Woods-Saxon potential by describing both the potential and the effective mass. Their model accomplishes this by introducing four different potentials - the normal (U_p and U_n) and spin-orbit potentials (W_p and W_n) for protons and neutrons. These potentials were tied to the vector, V, and scalar, S, potentials in the Dirac equation by considering their non-relativistic limit as

$$U = V + S \tag{5.16}$$

$$W = V - S. \tag{5.17}$$

The strengths of all four potentials are regulated by the overall potential strength V_0 and modulating factors for different numbers of protons and neutrons, κ, and for the strength of the spin-orbit contribution, λ_p and λ_n. The shape of the potentials is regulated by four diffusivities, a_p, a_n, a_p^{ls} and a_n^{ls}, and four radii R_0^n, R_0^p, $R_{0,ls}^n$, and $R_{0,ls}^p$ [1].

The resulting potentials are as follows:

$$U_p(r) = \frac{V_0\left(1 + \kappa\frac{N-Z}{A}\right)}{1 + e^{\frac{r - R_0^p A^{1/3}}{a_p}}} + U_C(r) \tag{5.18}$$

$$U_n(r) = \frac{V_0\left(1 - \kappa\frac{N-Z}{A}\right)}{1 + e^{\frac{r - R_0^n A^{1/3}}{a_n}}} \tag{5.19}$$

$$W_p(r) = \frac{V_0\lambda_p\left(1 + \kappa\frac{N-Z}{A}\right)}{1 + e^{\frac{r - R_{0,ls}^p A^{1/3}}{a_p^{ls}}}} + W_C(r) \tag{5.20}$$

$$W_n(r) = \frac{V_0\lambda_n\left(1 - \kappa\frac{N-Z}{A}\right)}{1 + e^{\frac{r - R_{0,ls}^n A^{1/3}}{a_n^{ls}}}}. \tag{5.21}$$

An additional component describing the repulsive Coulomb potential, U_C, is added to the potential of protons using the homogeneously charged sphere potential

$$U_C(r) = \begin{cases} Ze^2\left(\frac{3}{R_0^p A^{1/3}} - \frac{r^2}{(R_0^p)^3 A}\right), & r \le R_0^p A^{1/3} \\ \frac{Ze^2}{r}, & r > R_0^p A^{1/3} \end{cases} \tag{5.22}$$

$$W_C(r) = \begin{cases} Ze^2\left(\frac{3}{R_0^p A^{1/3}} - \frac{r^2}{(R_0^p)^3 A}\right), & r \le R_0^p A^{1/3} \\ \frac{Ze^2}{r}, & r > R_0^p A^{1/3} \end{cases}. \tag{5.23}$$

---

[1] The notation of [51] has been simplified, and the signature of the spin-orbit potentials has been absorbed into λ_n and λ_p for convenience.

5.2.3 Fisher information matrix

Using indices from the beginning of the Latin alphabet for N_m measurements, and the Greek letters for N_p model parameters, here labeled as

$$\mathbf{p} = \begin{pmatrix} p^1 \\ \vdots \\ p^{N_p} \end{pmatrix}, \tag{5.24}$$

we want to compute error estimates for the problem of fitting a model $f^a(\mathbf{p})$ to measurements y^a assuming measurement errors σ^a.

In the standard maximum likelihood method, the best-fitting value of p^μ is found by minimizing the χ^2 value

$$\chi^2(\mathbf{p}) = \sum_{a=1}^{N_m} \left(\frac{y^a - f^a(\mathbf{p})}{\sigma^a} \right)^2 . \tag{5.25}$$

A useful derived quantity is the reduced χ^2 value

$$\chi^2_{red} = \frac{\chi^2}{N_m - N_p}, \tag{5.26}$$

which should be close to 1 for models that are neither over-fitted nor under-fitted.

We find parameter uncertainties using the Cramer-Rao bound on the covariance matrix Σ, which is based on the inverse of the FIM, $g_{\mu\nu}$ [21]

$$g_{\mu\nu}(\mathbf{p}) = \sum_a \frac{\partial_\mu f^a \partial_\nu f^a}{(\sigma^a)^2} . \tag{5.27}$$

We compute model derivatives using algorithmic differentiation implemented in the `autograd` package. Using AD procedures, we eliminate numerical errors related to using numerical differentiation approximations.

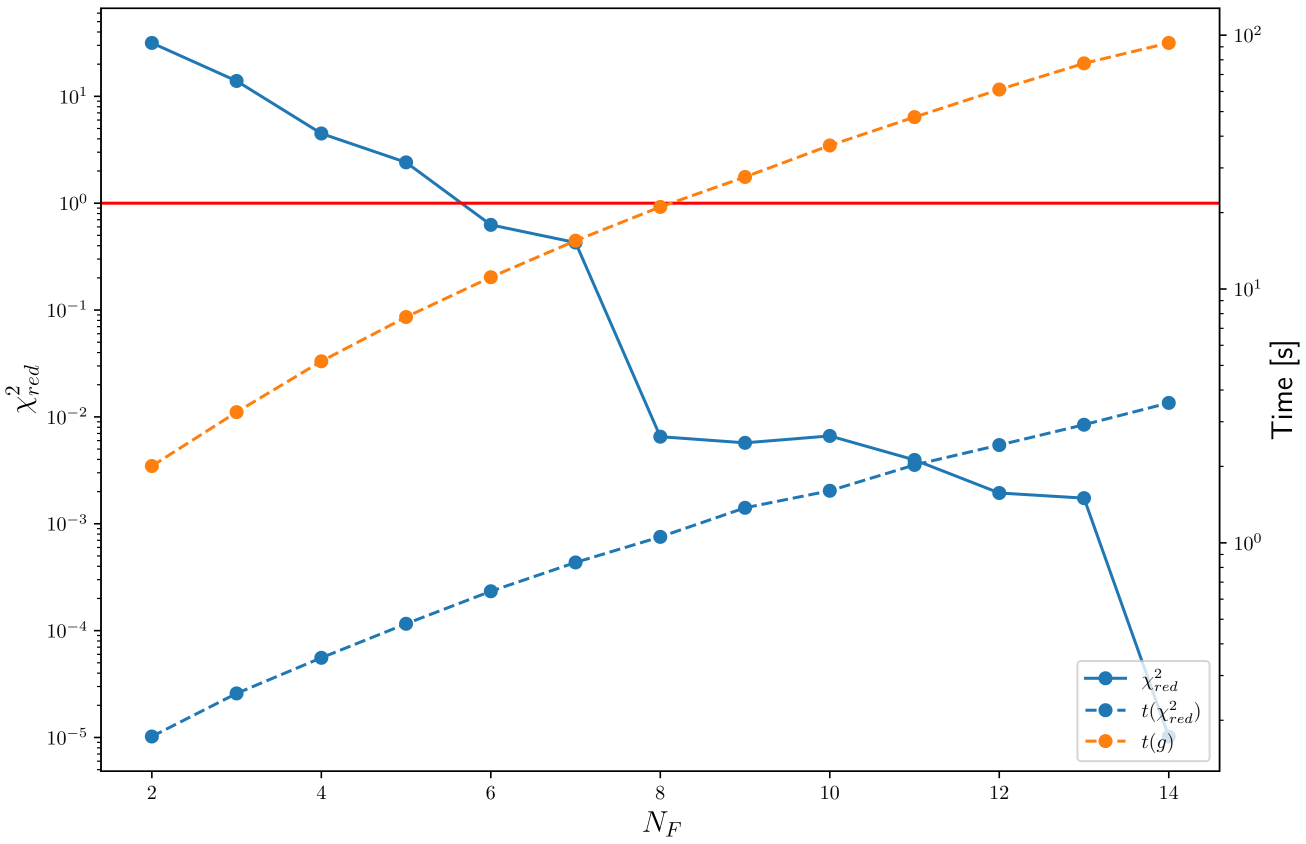

Figure 5.1: Reduced χ^2_{red} value of the finite-nucleus model as a function of N_F for the Woods-Saxon potential. The dashed lines represent the execution time of the χ^2_{red} function and the computation time of the Woods-Saxon FIM.

Table 5.1: Charge-radius and single-particle energy data set. The dataset consists of the charge radii, r_{ch}, and single particle neutron, ϵ_n, and proton, ϵ_p, energies for occupied states. The single-particle energies were computed using the Woods-Saxon potential as determined in [51].

Nucleus	r_{ch} (fm)
^{4}He	1.65
^{16}O	2.41
^{40}Ca	3.29

	ϵ_n (MeV)					
	$1s_{1/2}$	$1p_{3/2}$	$1p_{1/2}$	$2s_{1/2}$	$1d_{5/2}$	$1d_{3/2}$
^{4}He	−25.30					
^{16}O	−43.20	−24.68	−19.04			
^{40}Ca	−53.34	−39.40	−35.40	−24.95	−18.51	−17.42

	ϵ_p (MeV)					
	$1s_{1/2}$	$1p_{3/2}$	$1p_{1/2}$	$2s_{1/2}$	$1d_{5/2}$	$1d_{3/2}$
^{4}He	−24.95					
^{16}O	−40.08	−22.39	−18.36			
^{40}Ca	−45.80	−33.08	−30.32	−19.53	−14.96	−13.23

5.3 Input selection

We analyze the statistical properties of the RMF procedure on charge-radius, r_{ch}, and single-particle energy data. To this end, we chose a set of doubly-magic nuclei: ^{4}He, ^{16}O, and ^{40}Ca. Since the parameter space consists of 12 parameters and only three nuclei, the chosen data set consisted of their charge-radii and the single-particle energies of protons and neutrons for occupied states computed using the Koepf values [51]. For statistical analyses, these parameter values were taken as the best-fitting values for the Woods-Saxon potential.

Using charge-radii and the energies of the occupied single-particle states resulted in 23 data points, ensuring enough degrees of freedom for a twelve-parameter model. A further advantage of using the aforementioned doubly-magic even-even nuclei is that they have an equal number of protons and neutrons, resulting in removing the need for the parameter κ. Hence, the parameter space is reduced to 11 dimensions. We compute the charge-radius, r_{ch}, from the root-mean-square radius, $\langle r^2 \rangle$, [as in, e.g., 49] using the proton density distribution, as $r_{ch} = \sqrt{\langle r^2 \rangle + 0.64}$. A homoscedastic error of 0.1 fm and 0.1 MeV has been chosen arbitrarily since the data set consists of the model evaluation, *not* spectral measurements.

The corresponding reduced χ^2_{red} value of the finite-nucleus model as a function of N_F for the Woods-Saxon potential is shown in Fig. 5.1. The choice of a different error would only shift the χ^2_{red} curve upwards or downwards. The figure also shows the execution time as a function of the maximal total quantum number N_F, displayed as a dashed line. The simple relation $\chi^2_{red} \sim 1$ should hold to minimize the impact of over-fitting and under-fitting. The model accomplishes this near $N_F \sim 5$. Since the execution time of the χ^2_{red} function rises progressively with a larger N_F, the value of the N_F parameter was set to 5 for statistical analyses. The execution time for the FIM matrix for this model shows similar behavior. The chosen (pseudo)dataset is shown in Table 5.1 and is computed using a $N_F = 15$, which is outside the examined N_F range in Fig. 5.1 to avoid the model evaluations that would result in exactly $\chi^2 = 0$. The value of N_F was chosen to be large enough so that the values of all computed parameters differ less than 10% of the adopted value for the homoscedastic error between neighboring values of N_F.

5.4 Results

We applied the finite nucleus procedure to compute parameter uncertainties for the Woods-Saxon potential. We estimate the errors of the model parameters by computing the diagonal elements of the FIM, σ_{FIM}, as presented in Table 5.2. In Sec. 5.4.2 we compare the values of the FIM components computed using AD and the values computed using numerical differentiation and in Sec. 5.4.2 we compare the FIM-derived error estimates to the MCMC of the derived FIM to the error estimates. In Sec. 5.4.3 we present the extension of our analysis to the case of the DD-PC1 functional, which is subject of ongoing research and will be published as a separate paper.

5.4.1 Comparison with numerical differentiation

The numerical differentiation is compared to the one using a symmetric differentiation step h. The figure shows the relative error, R for the different components of the FIM computed as

$$R(g_{\mu\nu}) = \left| \frac{g_{\mu\nu}^{(A)} - g_{\mu\nu}^{(N)}}{g_{\mu\nu}^{(A)}} \right|, \tag{5.28}$$

where $g_{\mu\nu}^{(A)}$ is our AD-derived FIM estimate of the $\mu\nu$ matrix component of the FIM and $g_{\mu\nu}^{(N)}$ is the numerical estimate computed using a differentiation step h. In Figs. 5.2-5.10, we show these relative errors computed for different values of h and N_F. For very small values of $h < 10^{-7}$, the numerical errors due to floating point precision accumulate, while for $h > 10^{-2}$, the finite difference approximation tends to break down. As demonstrated by Figs. 5.2-5.10, this behavior occurs for all N_F, and the values of the relative error do not depend strongly on N_F.

Motivated by the χ^2 testing procedure, we analyze the impact of the possible overall worst-case error scenario by computing the sum of all relative errors of the FIM components in the bottom right panel of Fig. 5.10. This worst-case error estimate suggests that the optimal h is consistently $h \sim 10^{-4}$ for the entire range of N_F. One can conclude that the AD implementation provides accurate estimates of the FIM and that any discrepancy to the numerical derivative can be attributed to the inherent issues of numeric derivatives.

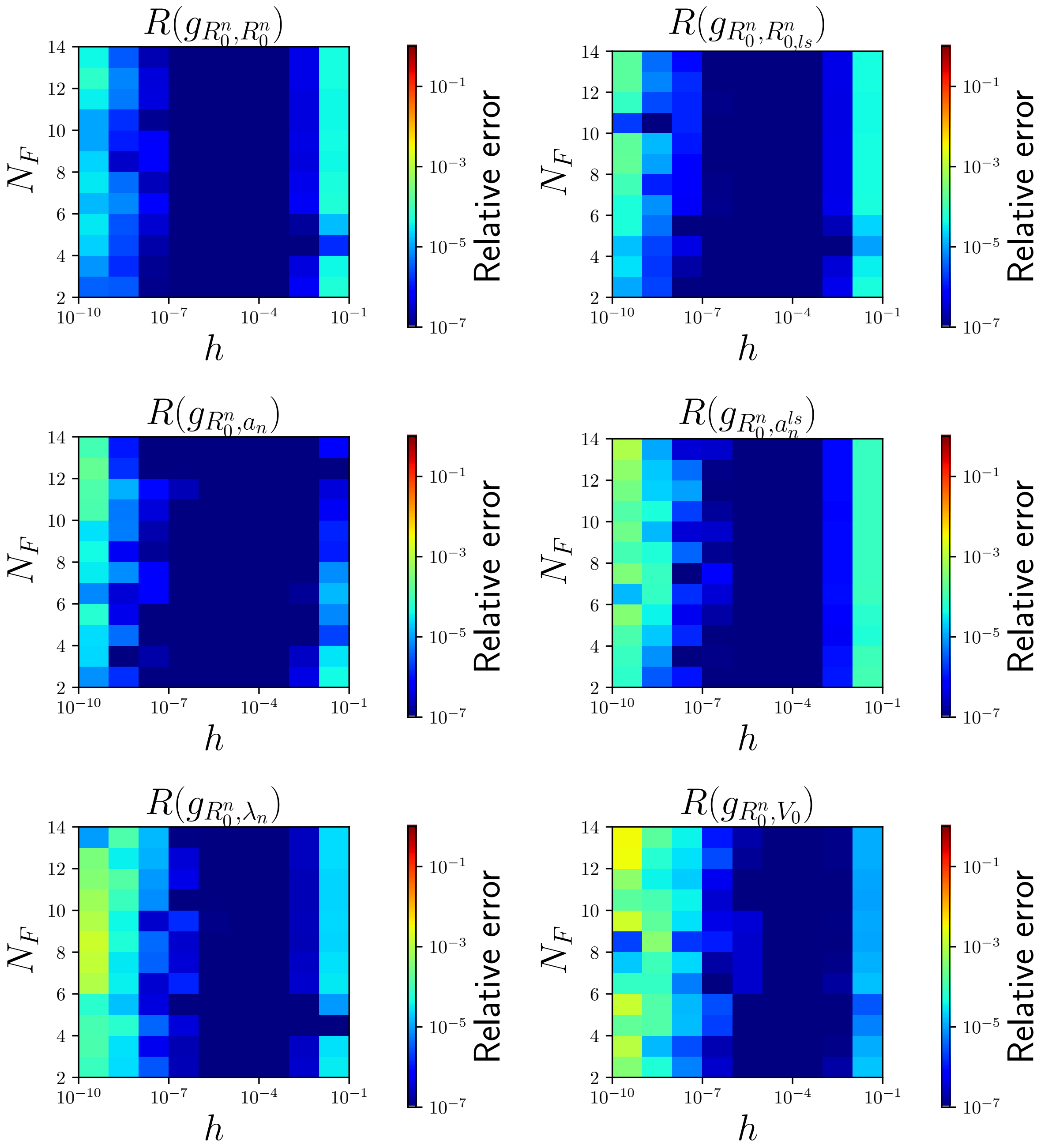

Figure 5.2: Relative error of the different FIM components $R(g_{\mu\nu})$, $\mu = R_0^n$, color-coded as a function of N_F and numerical derivative step h for the FIM components. The relative error compares the AD-derived FIM estimate, $g_{\mu\nu}^{(A)}$, to the numerical estimate $g_{\mu\nu}^{(N)}$.

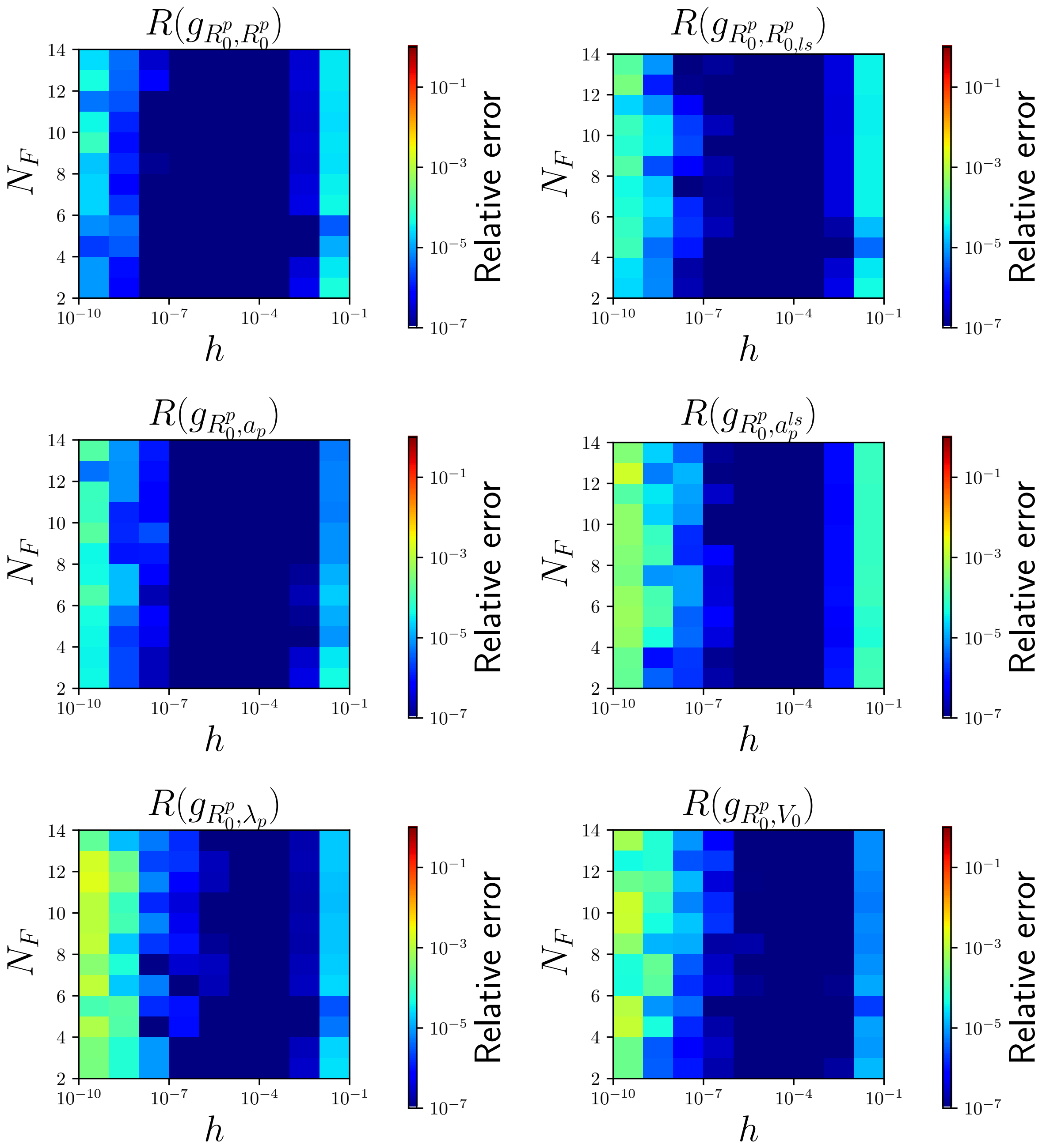

Figure 5.3: Relative error of the different FIM components $R(g_{\mu\nu})$, $\mu = R_0^p$, color-coded as a function of N_F and numerical derivative step h for the FIM components. The relative error compares the AD-derived FIM estimate, $g_{\mu\nu}^{(A)}$, to the numerical estimate $g_{\mu\nu}^{(N)}$.

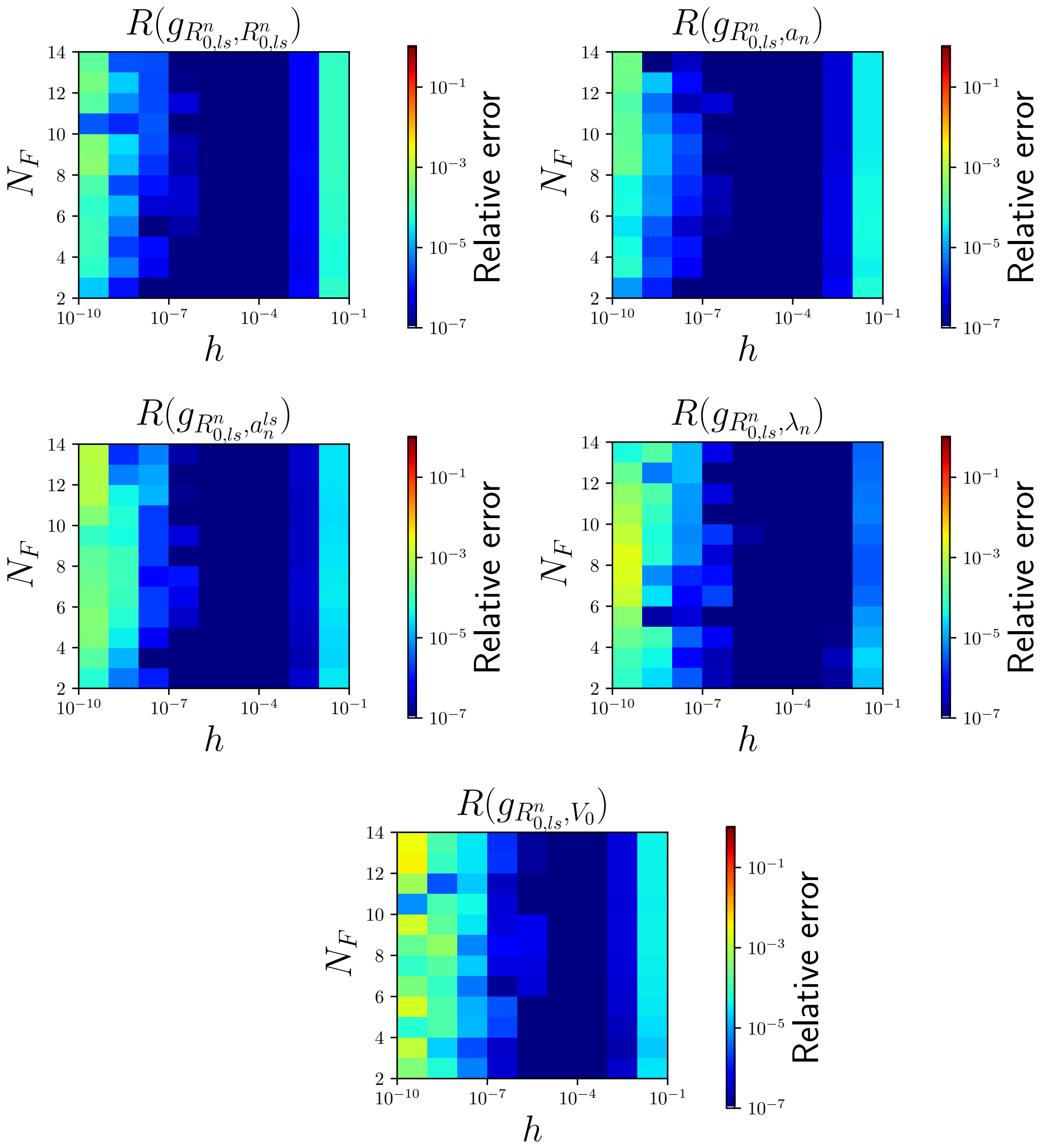

Figure 5.4: Relative error of the different FIM components $R(g_{\mu\nu})$, $\mu = R^n_{0,ls}$, color-coded as a function of N_F and numerical derivative step h for the FIM components. The relative error compares the AD-derived FIM estimate, $g^{(A)}_{\mu\nu}$, to the numerical estimate $g^{(N)}_{\mu\nu}$.

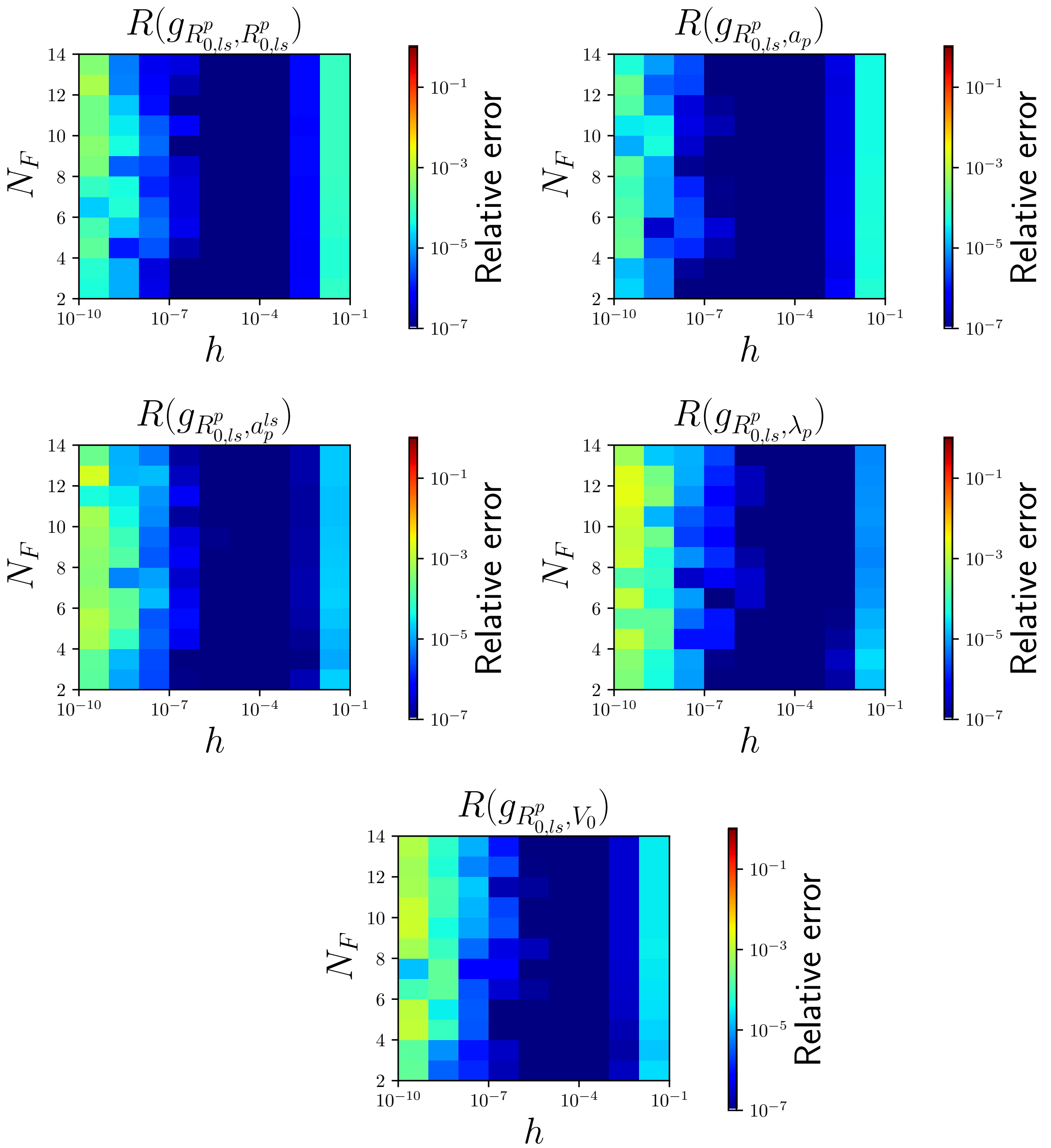

Figure 5.5: Relative error of the different FIM components $R(g_{\mu\nu})$, $\mu = R^p_{0,ls}$, color-coded as a function of N_F and numerical derivative step h for the FIM components. The relative error compares the AD-derived FIM estimate, $g^{(A)}_{\mu\nu}$, to the numerical estimate $g^{(N)}_{\mu\nu}$.

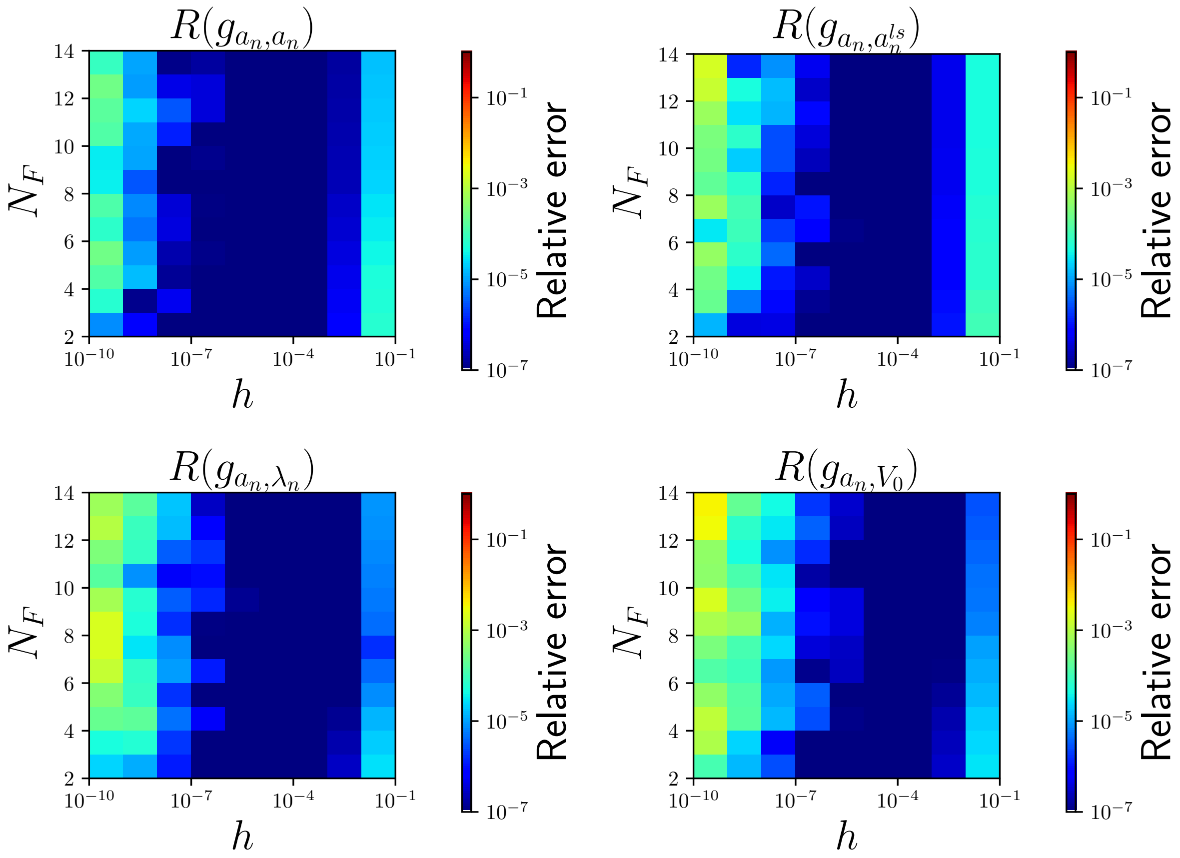

Figure 5.6: FIM components $R(g_{\mu\nu})$, $\mu = a_n$, color-coded as a function of N_F and numerical derivative step h for the FIM components. The relative error compares the AD-derived FIM estimate, $g_{\mu\nu}^{(A)}$, to the numerical estimate $g_{\mu\nu}^{(N)}$.

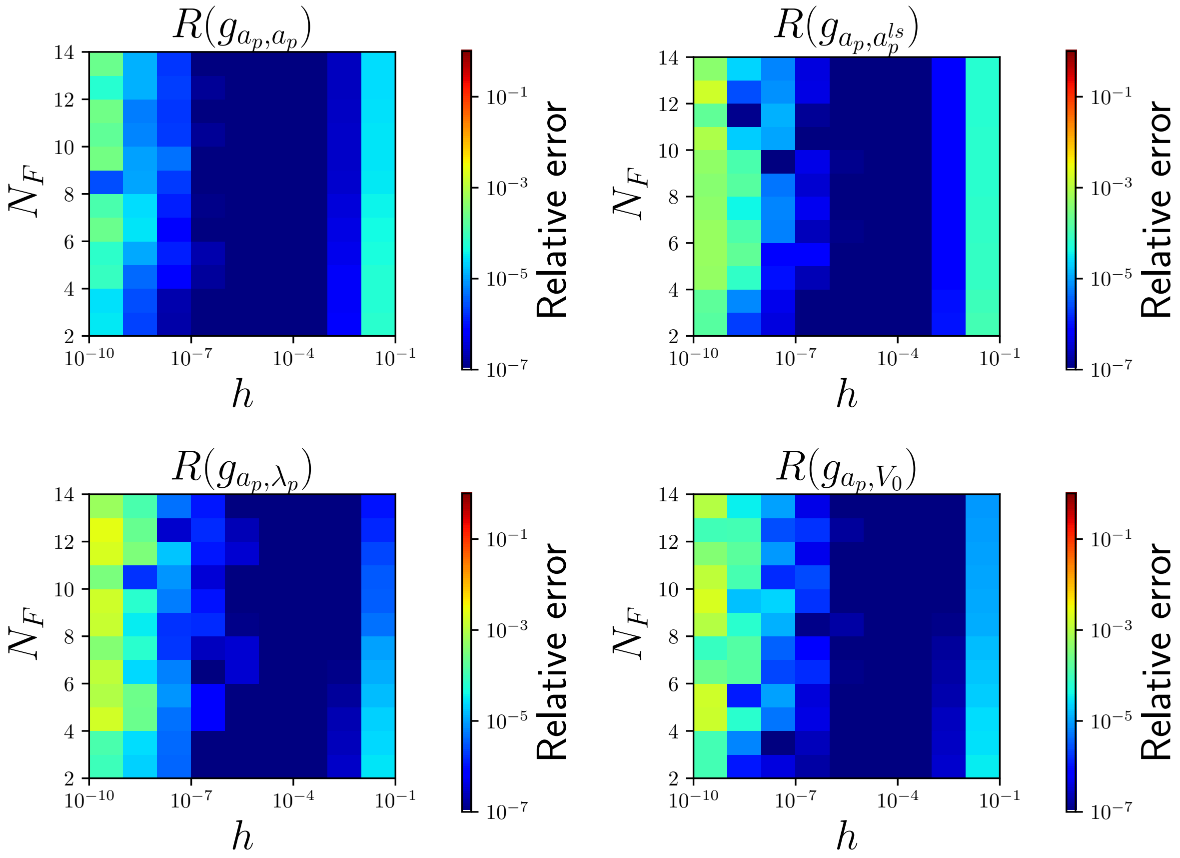

Figure 5.7: FIM components $R(g_{\mu\nu})$, $\mu = a_p$, color-coded as a function of N_F and numerical derivative step h for the FIM components. The relative error compares the AD-derived FIM estimate, $g^{(A)}_{\mu\nu}$, to the numerical estimate $g^{(N)}_{\mu\nu}$.

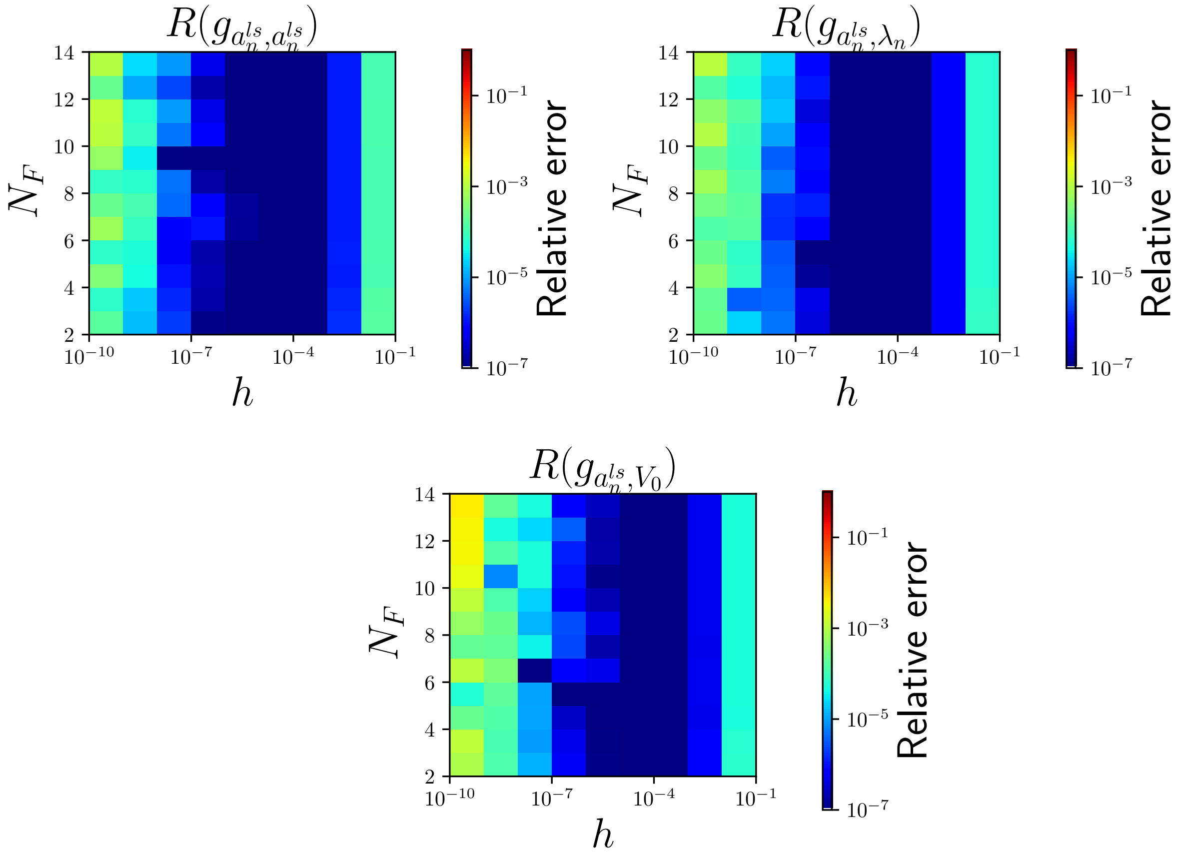

Figure 5.8: FIM components $R(g_{\mu\nu})$, $\mu = a_n^{ls}$, color-coded as a function of N_F and numerical derivative step h for the FIM components. The relative error compares the AD-derived FIM estimate, $g_{\mu\nu}^{(A)}$, to the numerical estimate $g_{\mu\nu}^{(N)}$.

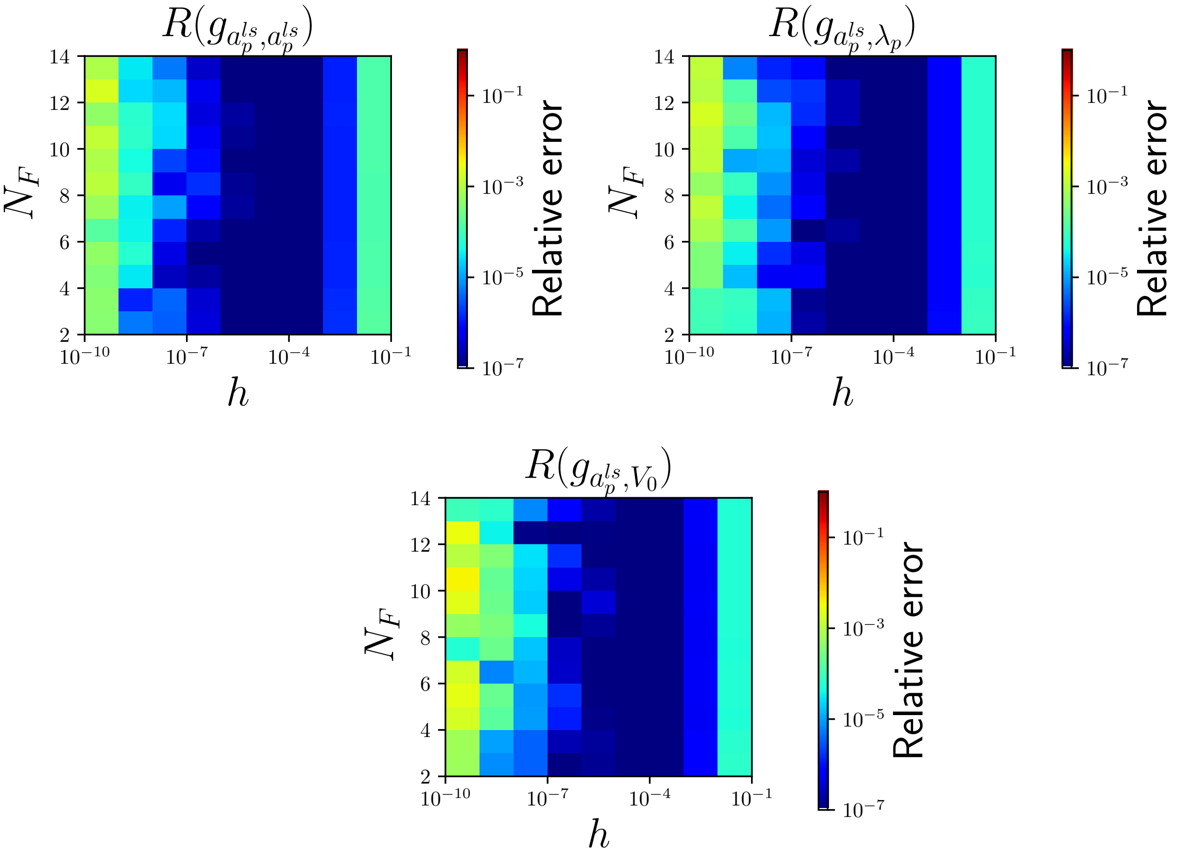

Figure 5.9: FIM components $R(g_{\mu\nu})$, $\mu = a_p^{ls}$, color-coded as a function of N_F and numerical derivative step h for the FIM components. The relative error compares the AD-derived FIM estimate, $g_{\mu\nu}^{(A)}$, to the numerical estimate $g_{\mu\nu}^{(N)}$.

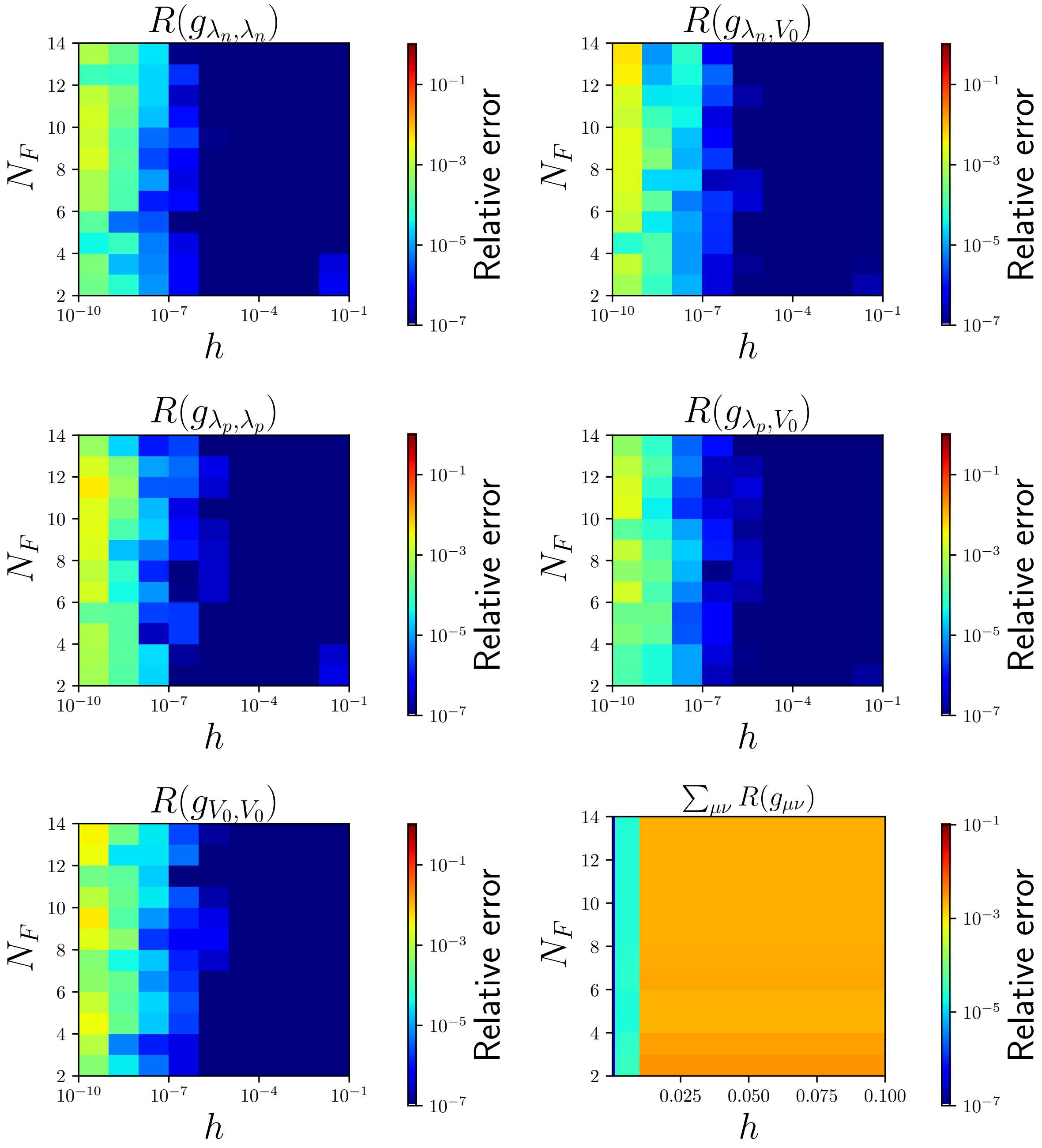

Figure 5.10: FIM components $R(g_{\mu\nu})$ for $\mu = \lambda_n$, $\mu = \lambda_p$ and $\mu = V_0$, color-coded as a function of N_F and numerical derivative step h for the FIM components. The relative error compares the AD-derived FIM estimate, $g^{(A)}_{\mu\nu}$, to the numerical estimate $g^{(N)}_{\mu\nu}$. In the bottom-right panel, the sum of all relative errors, $\sum_{\mu,\nu} R(g_{\mu\nu})$, is plotted.

5.4.2 Comparison with the Bayesian framework

We used the MCMC technique to sample the χ^2 posterior distribution, as implemented in the package `emcee` [68]. We used samples of 24 Markov chains of length 1000. The number of initialized chains has been chosen to fulfill the MCMC requirement that the number of Markov chain walkers be greater than the number of dimensions of the parameter space.

In Fig. 5.11, we plot the values of the MCMC samples of the parameter space as a function of the step in the Markov chain in which they were produced. We see that the values stabilize after $\sim$ 50 initial steps, indicating the expected *burn-in* phase for the MCMC method [68]. The sampled data points corresponding to the initial 50 steps have been excluded from further analysis. In Fig. 5.12 we show both the two-dimensional and one-dimensional marginal distributions of the MCMC samples in the parameter space. The blue lines show the value expected from the literature, which is well aligned with the distribution of the MCMC samples in all panels in Fig. 5.12.

The error estimates computed using MCMC sampling are listed alongside the FIM-based technique in Table 5.2. The medians and the 1σ confidence interval derived from the MCMC sampling align well with the Koepf estimates. To assess how significant are statistical differences between the Koepf estimates and the MCMC-based best-fitting parameter values of our dataset, in the last two columns, we compute the Z-scores,

$$Z^{\mu}(\sigma) = \frac{p^{\mu}_{\text{Koepf}} - p^{\mu}}{\sigma}. \tag{5.29}$$

We find that the differences are generally not statistically significant (i.e., they are less than 1σ) using either σ_{FIM} or σ_{MCMC}.

Reliably computing the FIM using AD enables producing error analysis without the time-consuming sampling of the parameter space by simply considering the diagonal of the FIM inverse. The resulting estimates of the sigma, σ_{FIM}, are in agreement with the MCMC estimates, σ_{MCMC}, as shown in Table 5.2.

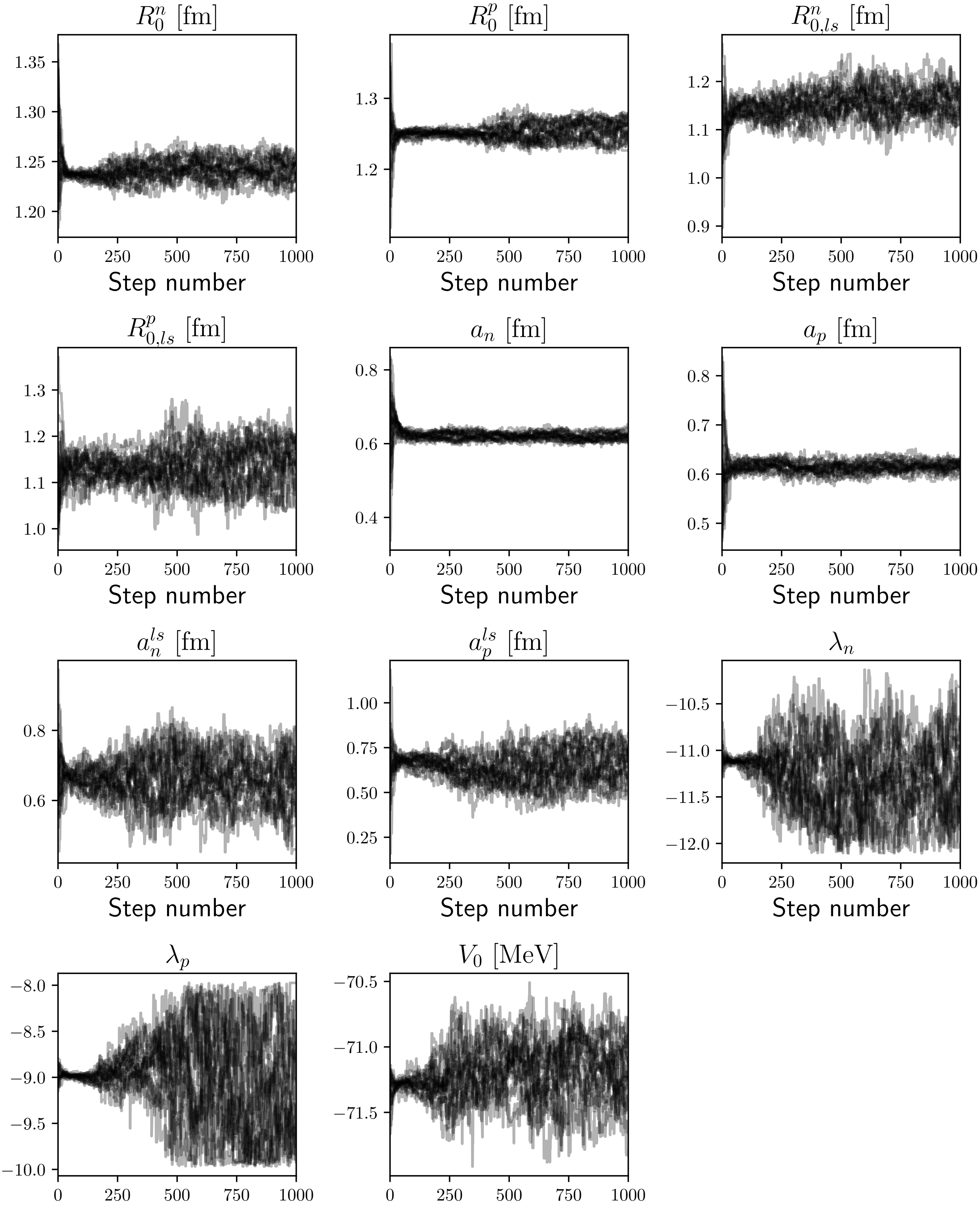

Figure 5.11: Values of the individual Markov chains of the MCMC sampling as a function of the MCMC step.

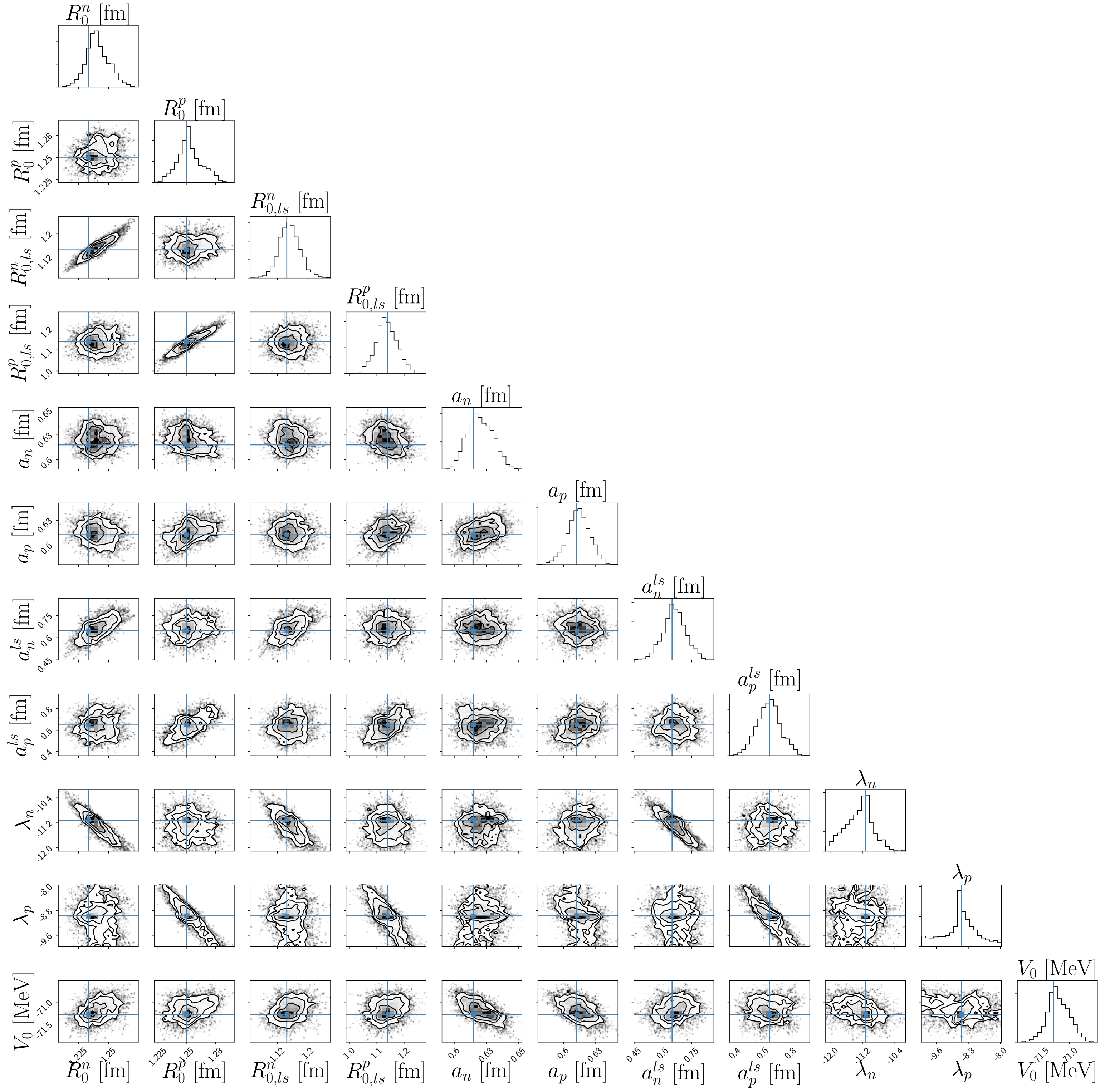

Figure 5.12: MCMC-derived sampling of the Woods-Saxon potential shown as two-dimensional sections of the parameter space.

Table 5.2: Results of Woods-Saxon potential fitting using the MCMC method for the charge radius and single-particle energy data set from Table 5.1.

Parameter	unit	Koepf estimate [51]	σ_{MCMC}	σ_{FIM}	MCMC	Confidence interval	$\lvert Z(\sigma_{MCMC})\rvert$	$\lvert Z(\sigma_{FIM})\rvert$
R_0^n	[fm]	1.2334	0.01	0.0112	1.241 ± 0.01	[1.23, 1.25]	0.79	0.66
R_0^p	[fm]	1.2496	0.0108	0.0168	1.25 ± 0.01	[1.24, 1.26]	0.28	0.18
$R_{0,ls}^n$	[fm]	1.1443	0.0273	0.0320	1.15 ± 0.03	[1.13, 1.18]	0.27	0.23
$R_{0,ls}^p$	[fm]	1.1401	0.0389	0.0563	1.14 ± 0.04	[1.10, 1.18]	0.11	0.07
a_n	[fm]	0.6150	0.0097	0.0098	0.62 ± 0.01	[0.61, 0.63]	0.56	0.55
a_p	[fm]	0.6124	0.0107	0.0108	0.61 ± 0.01	[0.60, 0.63]	0.20	0.20
a_n^{ls}	[fm]	0.6476	0.0601	0.0746	0.66 ± 0.06	[0.60, 0.72]	0.23	0.18
a_p^{ls}	[fm]	0.6469	0.0848	0.1271	0.64 ± 0.08	[0.56, 0.72]	0.05	0.03
λ_n	[1]	-11.1175	0.3391	0.4167	-11.3 ± 0.3	[-11.65, -10.98]	0.49	0.39
λ_p	[1]	-8.9698	0.4287	0.7025	-9.0 ± 0.4	[-9.47, -8.61]	0.07	0.04
V_0	[MeV]	-71.2800	0.1941	0.2228	-71.2 ± 0.2	[-71.37, -70.99]	0.45	0.39
κ	[1]	0.4616	N/A	N/A	N/A	N/A	N/A	N/A

5.4.3 The error estimates for the DD-PC1 functional

In this section, we utilize nuclear structure codes to precisely calculate error estimates for point-coupling models in finite nuclei. To initiate the DD-PC1 functional iteration, we compute densities for the Woods-Saxon potential, which serve as the basis for our subsequent calculations and analyses.

Figures 5.13, 5.14, and 5.15 showcase the charge-radius and single-particle energies behavior of the DD-PC1 functional. These figures provide an insight into the functional's behavior when applied to a group of doubly magic nuclei.Through our analysis, we have found that the iteration process shows quick stabilization within just a few steps, regardless of the N_F values observed. To ensure accuracy and consistency in our calculations, we have applied the 7-parameter constraint on the DD-PC1 model parameters, which has been previously utilized for infinite nuclear matter. Furthermore, we have incorporated homoscedastic errors of 0.1 fm and 0.1 MeV to further enhance the precision of our calculations. The behavior of the Fisher

Information Matrix (FIM) eigenvalues, eigenvector components, and parameter errors can be observed through Figures 5.16, 5.17, and 5.18. These figures provide a visual representation of the corresponding outcomes.

In like manner, the outcomes from the thorough 10-parameter model have been exhibited and are easily discernible in Figures 5.19, 5.20, and 5.21. Based on the analysis, it is evident that utilizing the 7-parameter model for estimating parameter errors is significantly more accurate than relying on the 10-parameter DD-PC1 model. This outcome was predictable as the 7-parameter model is a simplified version of the 10-parameter model, referred to as the MBAM-reduced form. We've included error estimates for both the 7 and 10-parameter models based on the FIM (Fisher Information Matrix) in Table 5.3. The information provided is crucial for assessing and examining the effectiveness of the models and can be utilized for future research.

Table 5.3: Error estimates of the DD-PC1 model parameters.

Parameter	unit	7-parameter model	10-parameter model
a_s	[fm^2]	-10.0 ± 0.5	-10 ± 1
b_s	[fm^2]	-9.2 ± 0.6	-9 ± 1
c_s	[fm^2]	-6 ± 3	-6 ± 6
d_s	[1]	1.4 ± 0.3	1.4 ± 0.8
a_v	[fm^2]	6 ± 1	6 ± 3
b_v	[fm^2]	8.9 ± 0.3	8.9 ± 0.4
d_v	[1]	0.7 ± 0.2	0.7 ± 0.6
b_{tv}	[fm^2]	N/A	2 ± 1
d_{tv}	[1]	N/A	0.6 ± 0.5
δ_s	[fm^4]	N/A	-0.81 ± 0.02

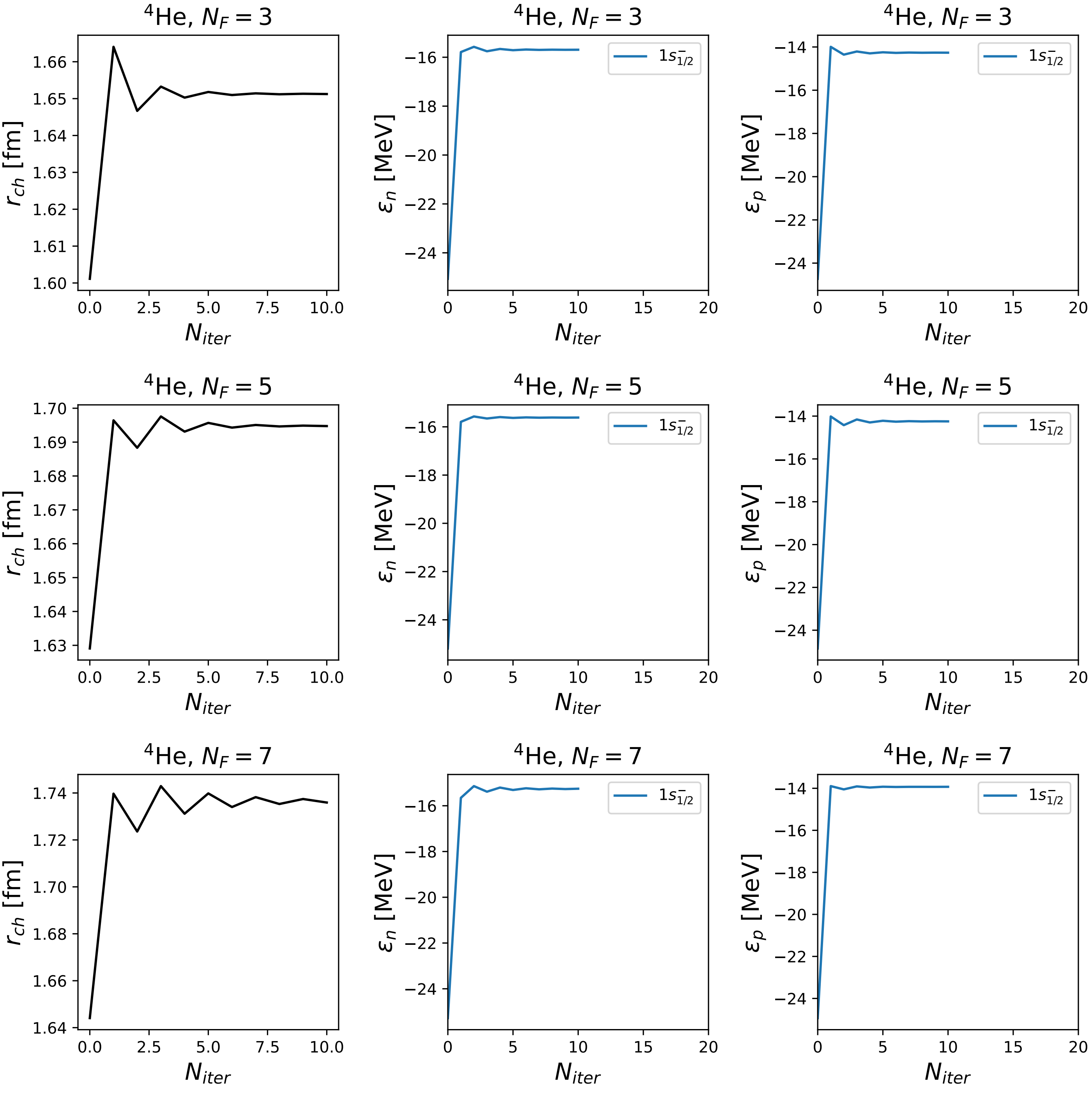

Figure 5.13: Charge-radius and single-particle energy computed using the DD-PC1 functional for the ^{4}He nucleus. The figure shows the values of the charge radii, r_{ch}, and single particle neutron, ϵ_n, and proton, ϵ_p, energies for occupied states for the different number of iterations N_{iter}.

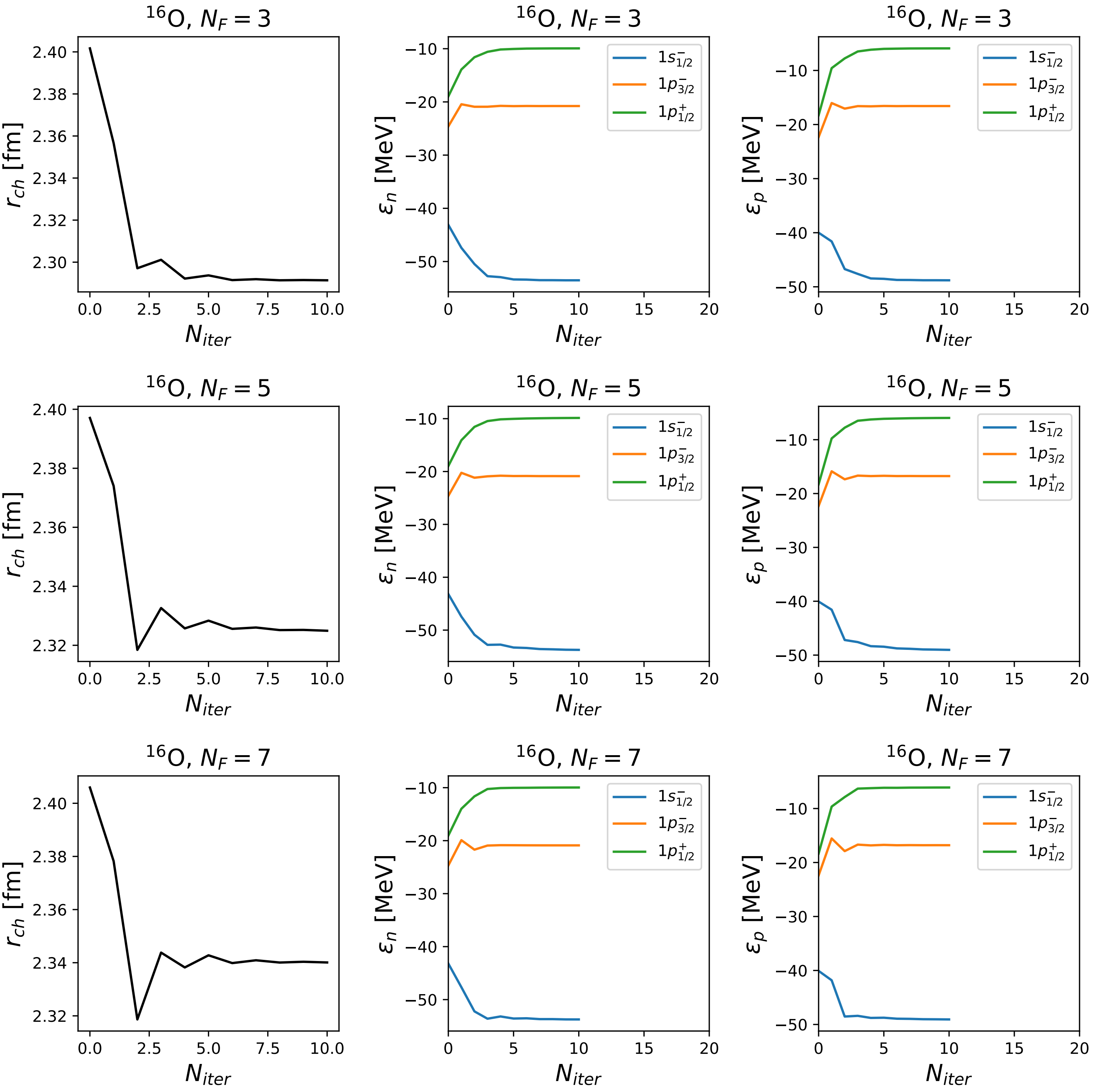

Figure 5.14: Charge-radius and single-particle energy computed using the DD-PC1 functional for the ^{16}O nucleus. The figure shows the values of the charge radii, r_{ch}, and single particle neutron, ϵ_n, and proton, ϵ_p, energies for occupied states for the different number of iterations N_{iter}.

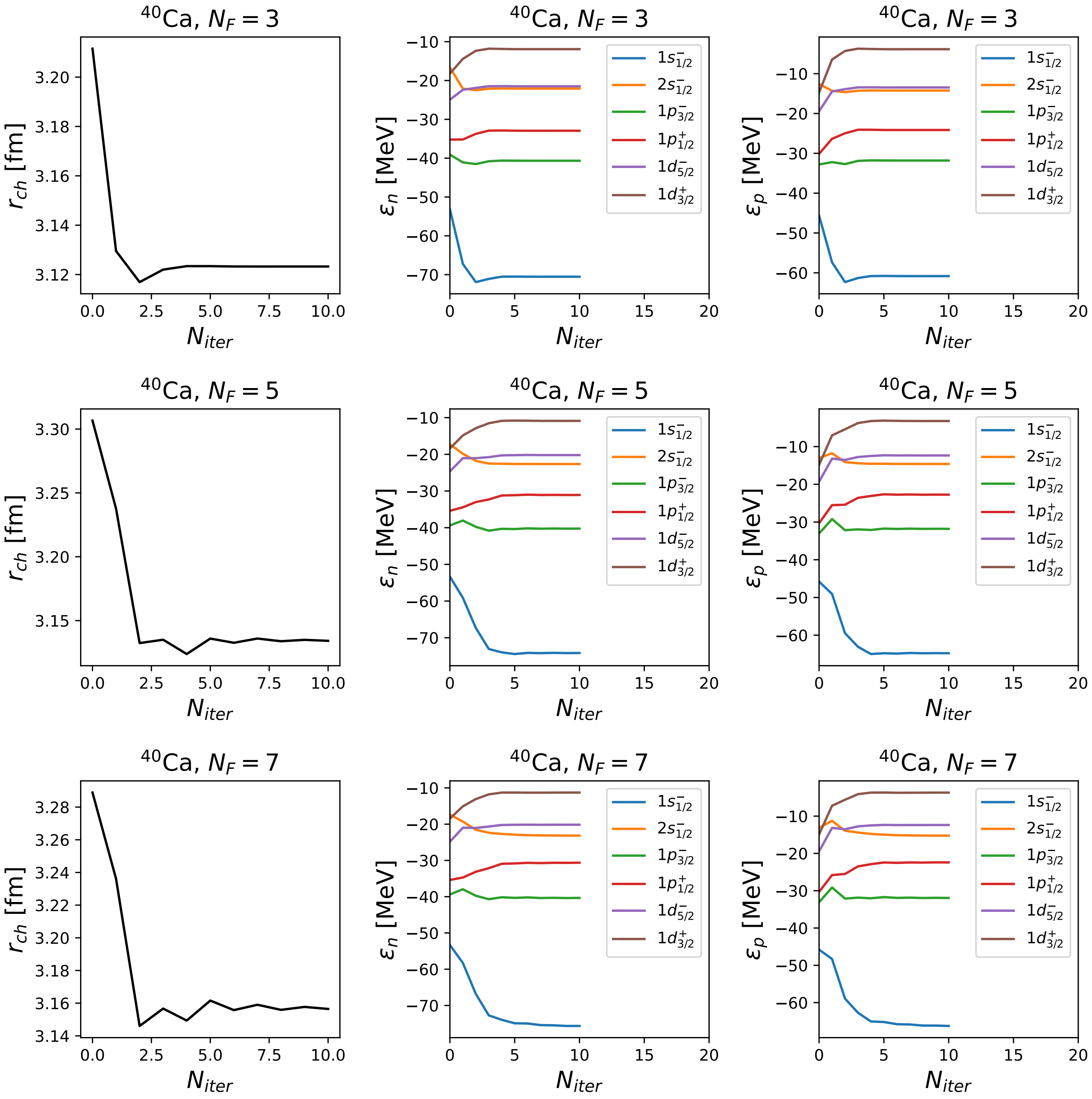

Figure 5.15: Charge-radius and single-particle energy computed using the DD-PC1 functional for the ^{40}Ca nucleus. The figure shows the values of the charge radii, r_{ch}, and single particle neutron, ϵ_n, and proton, ϵ_p, energies for occupied states for the different number of iterations N_{iter}.

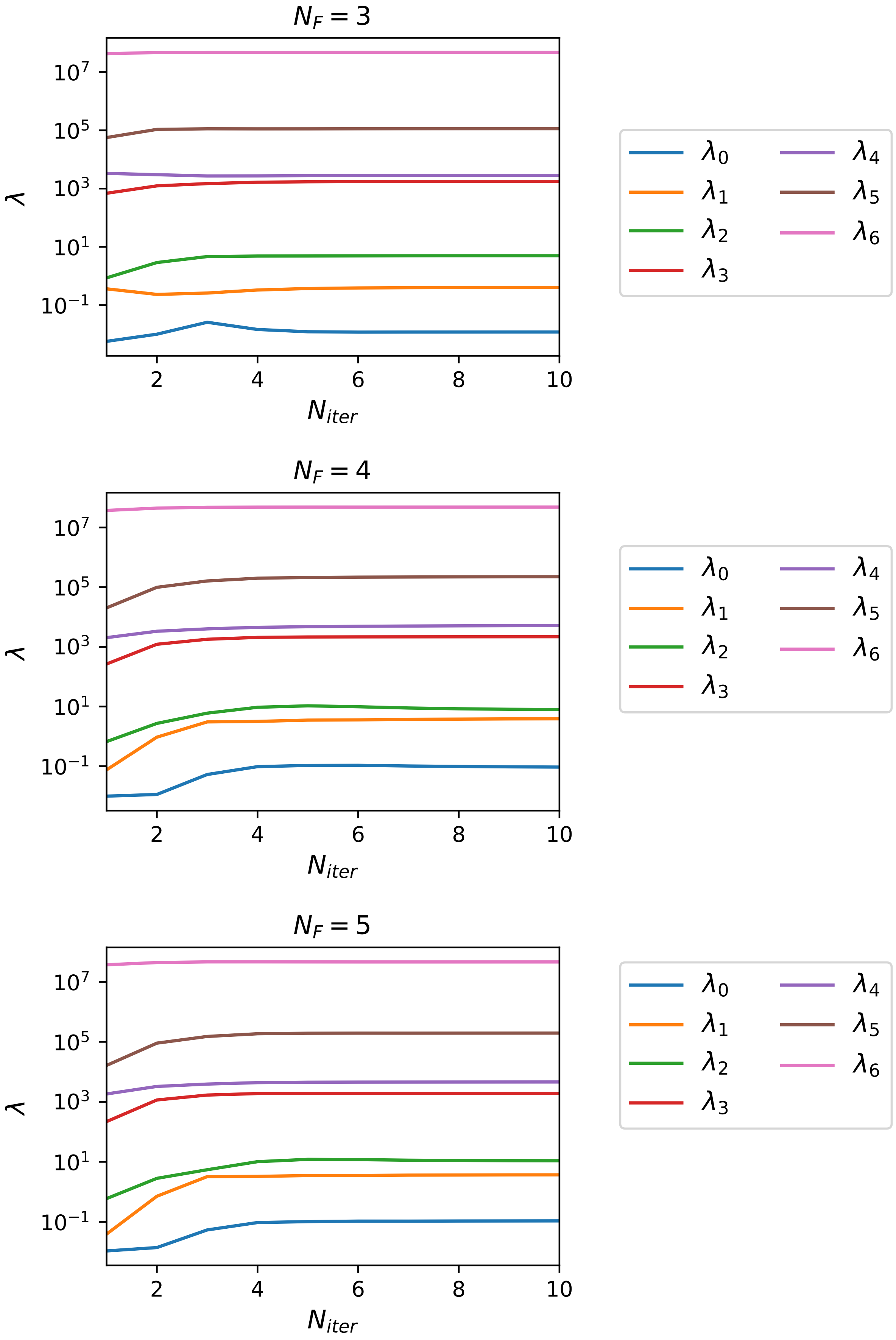

Figure 5.16: The FIM eigenvalues for the DD-PC1 functional in the 7-parameter model.

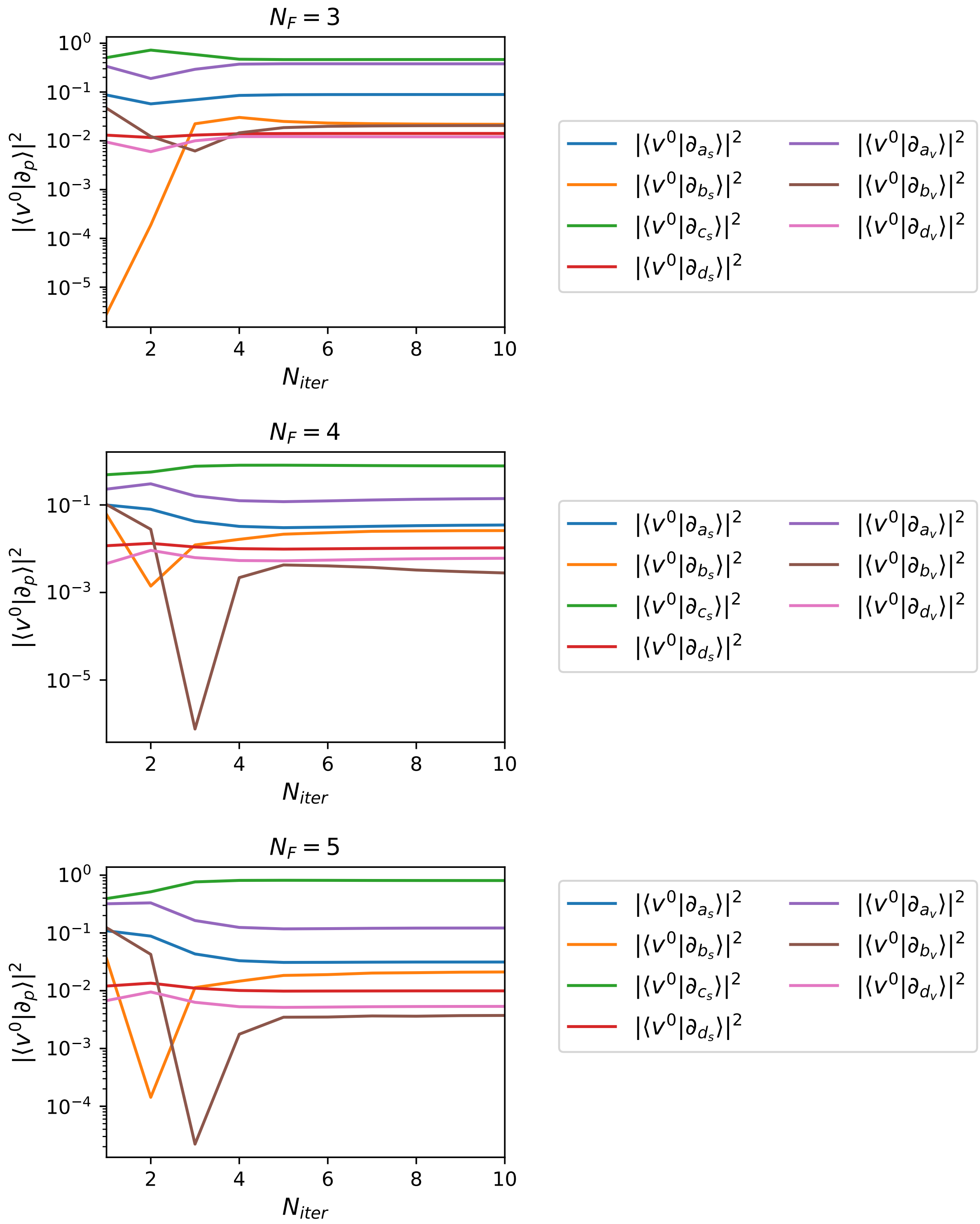

Figure 5.17: The FIM eigenvector components corresponding to the smallest FIM eigenvalue for the DD-PC1 functional in the 7-parameter model.

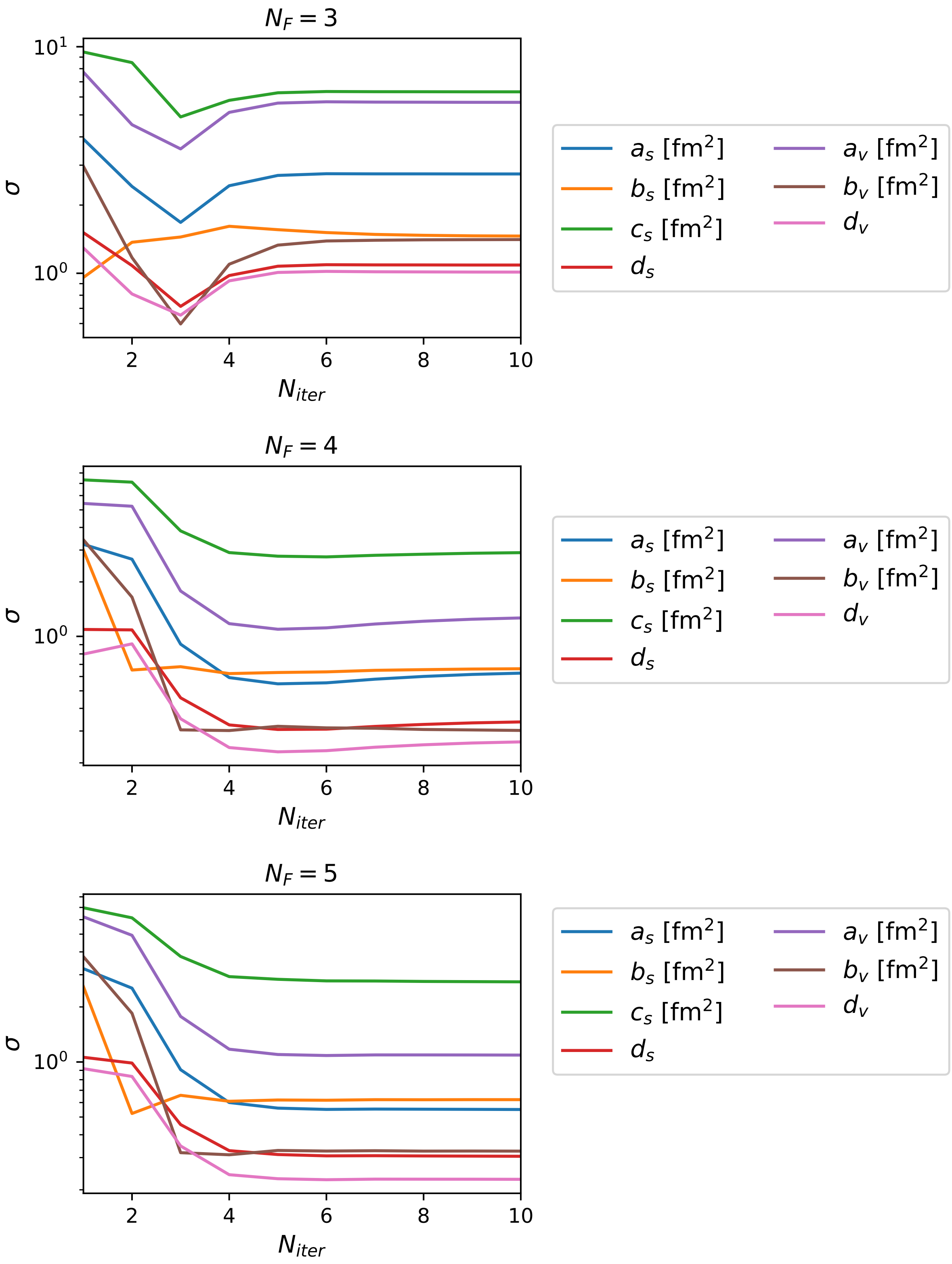

Figure 5.18: The FIM parameter error estimates for the DD-PC1 functional in the 7-parameter model.

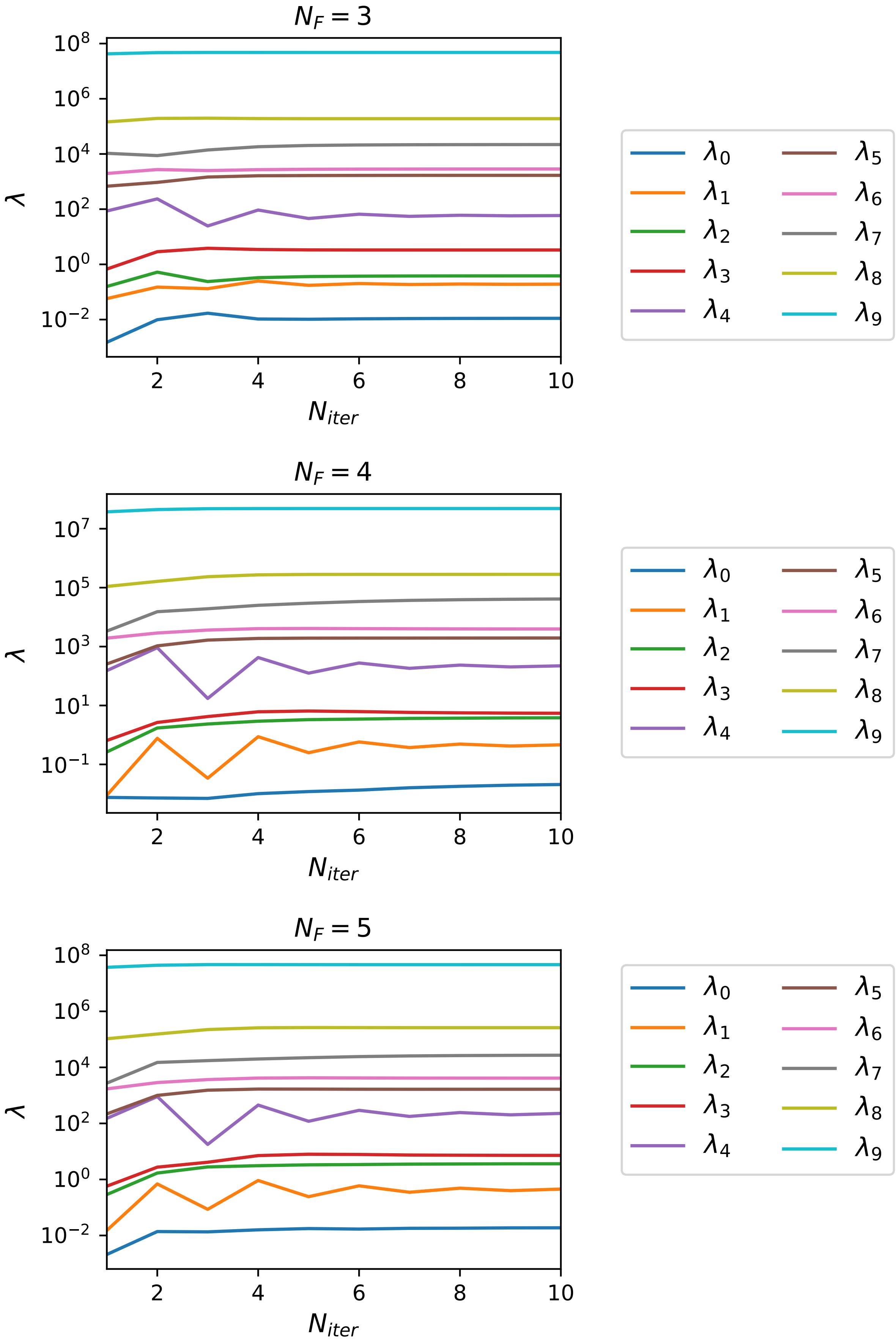

Figure 5.19: The FIM eigenvalues for the DD-PC1 functional in the 10-parameter model.

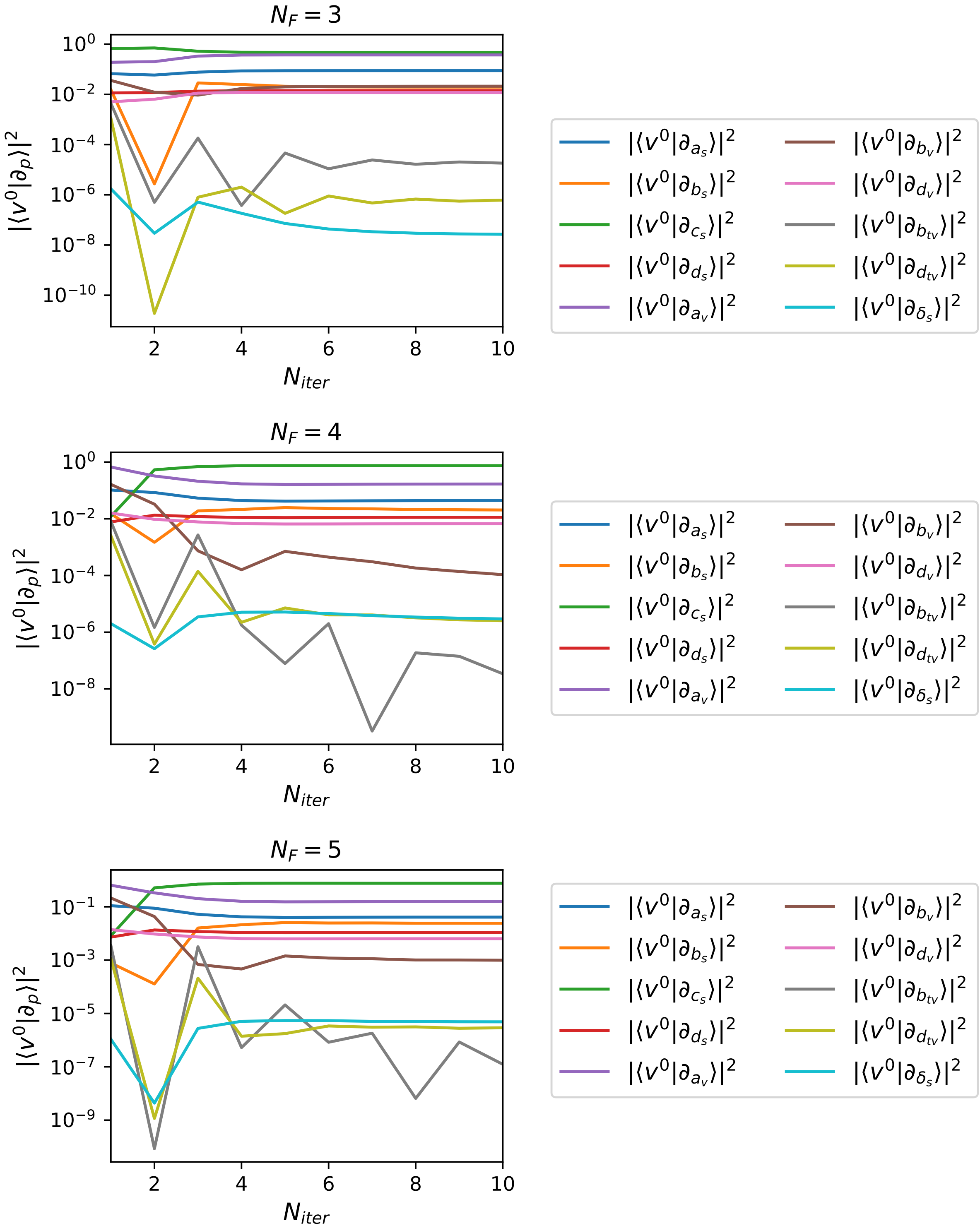

Figure 5.20: The FIM eigenvector components corresponding to the smallest FIM eigenvalue for the DD-PC1 functional in the 10-parameter model.

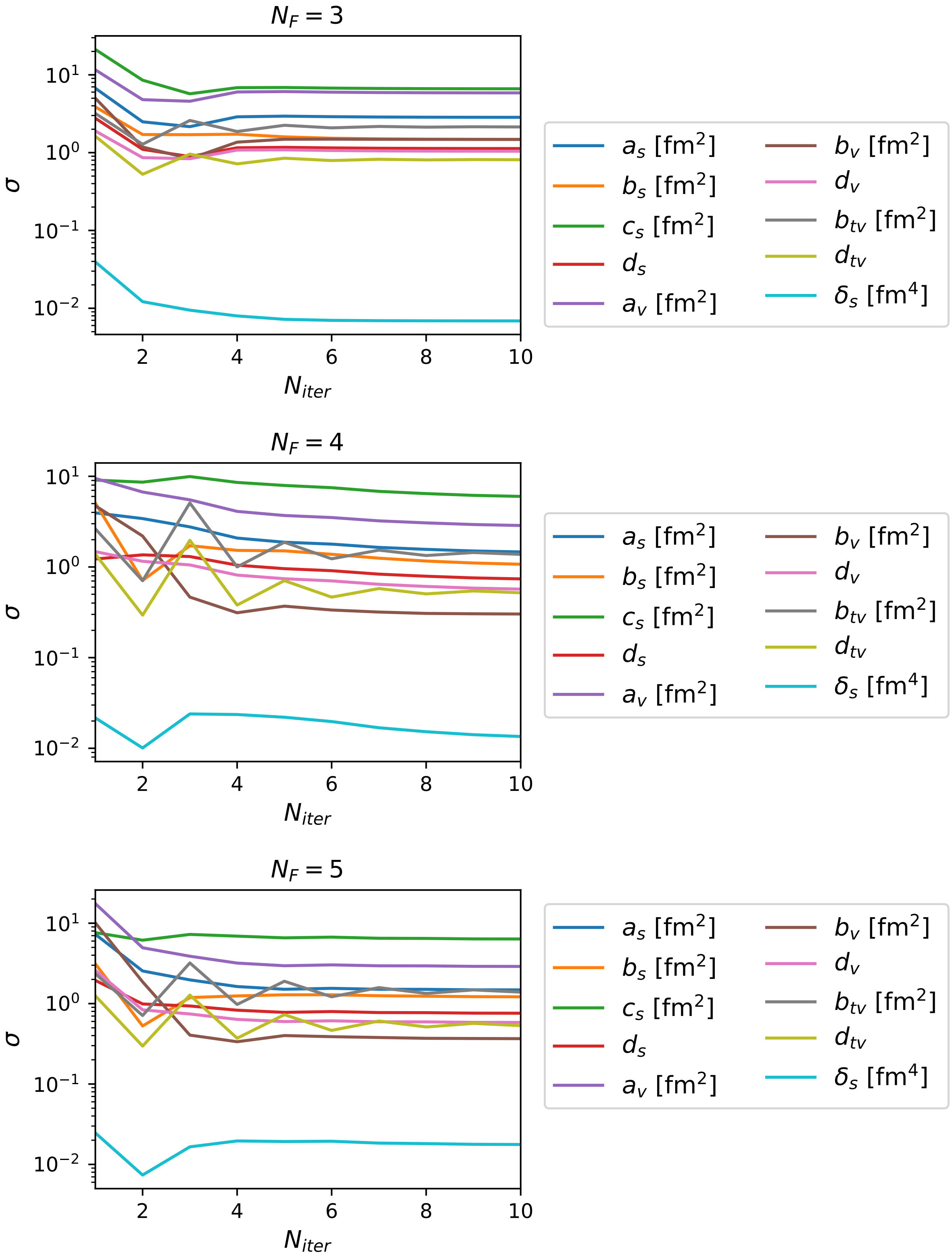

Figure 5.21: The FIM parameter error estimates for the DD-PC1 functional in the 10-parameter model.

5.5 Chapter summary

Exploring parameter estimation uncertainties in nuclear Effective Field Theories (EDFs) has recently emerged as a compelling research area. Our team has continued to build upon our past research [50], which delved into applying information geometry to EDFs in nuclear matter. Our current study presents a comprehensive statistical analysis of a straightforward method to establish the charge radius and single-particle energies in a series of doubly-magic nuclei utilizing the Woods-Saxon potential. Our findings offer valuable insights into the nature of nuclear matter and pave the way for future research in this field. We have evaluated the accuracy of error estimates using two different methods. The first method is a faster procedure employing the Fisher Information Matrix (FIM), while the second is a more intricate Bayesian Markov Chain Monte Carlo (MCMC) approach. We compared the results obtained from both methods to determine their effectiveness in estimating errors. When dealing with the complex situation of finite nuclei, it can be challenging to accurately determine the uncertainties of EDF parameters. However, by utilizing FIM in combination with algorithmic differentiation, it becomes possible to measure these uncertainties more precisely. This method of error analysis also eliminates the need for time-consuming sampling of the parameter space, which is usually required when using Bayesian statistical methods.

Chapter 6

Conclusion

The EDF framework serves as an extensive and all-encompassing theoretical methodology that offers a global perspective on nuclear occurrences across the nuclide chart. A scientifically sound approach to modelling nuclear interactions is the implementation of a relativistic mean-field Lagrangian that utilizes the point-coupling model. This model is specifically tailored to address the ground-state properties of nuclear matter and finite nuclei. Despite relying on point-coupling interactions, the outcomes generated by this model are comparable to those obtained from finite-range models.

The use of statistical analysis is of utmost importance in identifying errors in theoretical models, assessing risks, and detecting instabilities in the model. When it comes to nuclear energy density functionals, it has only recently been recognized that examining the unreliability of individual model parameters is crucial to gaining a better understanding of nuclear structure phenomena. However, analyzing error parameters can be challenging due to the possibility of dealing with imprecise parameters or their linear combinations. To overcome this obstacle, information geometry methods have been recently employed to obtain a better understanding of parameter uncertainties and model behavior. One notable instance of sophisticated procedures utilized in information geometry is the manifold boundary approximation method (MBAM). This technique has been effective in reducing the number of model parameters in EDF models. To accomplish this, MBAM relies on the principles of differential geometry to tackle statistical problems. This approach has been successfully implemented in various scientific fields.

In the fourth chapter, we made use of the MBAM technique to investigate the unreliability of the parameter estimates of the point-coupling EDF functional in nuclear matter. This allowed us to gain a deeper understanding of the behavior and properties of this functional in the context of nuclear physics. To this end, the manifold boundary approximation method (MBAM) solutions' error estimates were constrained in the framework of classical statistics, by means of the Monte Carlo simulations, as well as in the framework of Bayesian statistics, by means of the Markov Chain Monte Carlo method. To improve the reliability of parameters and streamline the execution time of code, we took the initiative of implementing the EDF and MBAM codes using Python, and incorporating automatic differentiation. This implementation proved to be highly advantageous, as it allowed for a much more accurate evaluation of the parameters, and completely eliminated any possibility of numerical imprecision. By utilizing this approach, we were able to ensure that the results obtained were of the highest possible quality and that the entire process was as efficient and effective as possible. After conducting a thorough analysis of the DD-PC1 model within the nuclear EDF, it has been confirmed that the conclusions drawn from utilizing the MBAM method remain steadfast and unwavering, even when the parameters are altered within the 1σ confidence interval of the optimal-fit model. This indicates a strong level of reliability and consistency in the results obtained from this methodology, lending further credibility to its use in future research and experimentation. In order to accurately identify the geodesic boundary, it is essential to carefully analyze the point at which the determinant of the Fisher Information Matrix (FIM) begins to approach zero. This particular condition plays a critical role in the formation of distinct and separate areas within the parameter space. Therefore, it is imperative to pay close attention to this critical point during the identification process.

In the fifth section of this study, an extensive examination was conducted on finite nuclei. The charge radius and single-particle energies of a specific set of doubly-magic nuclei were determined through a simple yet efficient technique. The method employed the Woods-Saxon potential and underwent a meticulous statistical analysis.

The research focused on comparing error estimates of two methods. One involved utilizing the FIM for a faster process, while the other required the more demanding Bayesian MCMC method. After comparing the results of both methods, it was discovered that they varied by less than 1σ. This implies that both techniques are dependable and precise.

Based on the findings, it can be deduced that the FIM-based approach combined with algorithmic differentiation is a dependable method for accurately estimating EDF parameter uncertainties even in the intricate scenario of finite nuclei. The proposed approach for error analysis offers a significant advantage in that it eliminates the requirement for an exhaustive sampling of the parameter space. This is particularly beneficial when using Bayesian statistical techniques, which can be a tedious and time-consuming process.

Bibliography

[1] G. Colò. Nuclear density functional theory. *Advances in Physics X*, 5(1):1740061, January 2020.

[2] B. A. Nikolaus, T. Hoch, and D. G. Madland. Nuclear ground state properties in a relativistic point coupling model. *Physical Review C*, 46(5):1757–1781, 11 1992.

[3] T. Bürvenich, D. G. Madland, J. A. Maruhn, and P.-G. Reinhard. Nuclear ground state observables and QCD scaling in a refined relativistic point coupling model. *Physical Review C*, 65(4), 3 2002.

[4] T. Nikšić, D. Vretenar, G. A. Lalazissis, and P. Ring. Finite- to zero-range relativistic mean-field interactions. *Phys. Rev. C*, 77(3):034302, March 2008.

[5] M. Kortelainen, T. Lesinski, J. Moré, W. Nazarewicz, J. Sarich, N. Schunck, M. V. Stoitsov, and S. Wild. Nuclear energy density optimization. *Phys. Rev. C*, 82(2):024313, August 2010.

[6] S. Bogner, A. Bulgac, J. Carlson, J. Engel, G. Fann, R.J. Furnstahl, S. Gandolfi, G. Hagen, M. Horoi, C. Johnson, M. Kortelainen, E. Lusk, P. Maris, H. Nam, P. Navratil, W. Nazarewicz, E. Ng, G.P.A. Nobre, E. Ormand, T. Papenbrock, J. Pei, S.C. Pieper, S. Quaglioni, K.J. Roche, J. Sarich, N. Schunck, M. Sosonkina, J. Terasaki, I. Thompson, J.P. Vary, and S.M. Wild. Computational nuclear quantum many-body problem: The UNEDF project. *Computer Physics Communications*, 184(10):2235–2250, 10 2013.

[7] P. G. Reinhard. Estimating the relevance of predictions from the Skyrme-Hartree-Fock model. *Phys. Scr*, 91(2):023002, February 2016.

[8] J. Dobaczewski, W. Nazarewicz, and P. G. Reinhard. Error estimates of theoretical models: a guide. *Journal of Physics G Nuclear Physics*, 41(7):074001, July 2014.

[9] Nicolas Schunck, Jordan D. McDonnell, Jason Sarich, Stefan M. Wild, and Dave Higdon. Error analysis in nuclear density functional theory. *Journal of Physics G Nuclear Physics*, 42(3):034024, March 2015.

[10] N. Schunck, J. D. McDonnell, D. Higdon, J. Sarich, and S. M. Wild. Uncertainty quantification and propagation in nuclear density functional theory. *European Physical Journal A*, 51:169, December 2015.

[11] J. Piekarewicz, Wei-Chia Chen, and F. J. Fattoyev. Information and statistics: a new paradigm in theoretical nuclear physics. *Journal of Physics G Nuclear Physics*, 42(3):034018, March 2015.

[12] X. Roca-Maza, N. Paar, and G. Colò. Covariance analysis for energy density functionals and instabilities. *Journal of Physics G: Nuclear and Particle Physics*, 42(3), 2015.

[13] S. E. Agbemava, A. V. Afanasjev, and A. Taninah. Propagation of statistical uncertainties in covariant density functional theory: Ground state observables and single-particle properties. *Phys. Rev. C*, 99(1):014318, January 2019.

[14] I. Dedes and J. Dudek. Predictive power of theoretical modelling of the nuclear mean field: examples of improving predictive capacities. *Phys. Scr*, 93(4):044003, March 2018.

[15] I. Dedes and J. Dudek. Propagation of the nuclear mean-field uncertainties with increasing distance from the parameter adjustment zone: Applications to superheavy nuclei. *Phys. Rev. C*, 99(5):054310, May 2019.

[16] V Kejzlar, L Neufcourt, W Nazarewicz, and P G. Reinhard. Statistical aspects of nuclear mass models. *Journal of Physics G Nuclear Physics*, 47(9):94001, 9 2020.

[17] A Taninah, S. E. Agbemava, A. V. Afanasjev, and P Ring. Parametric correlations in energy density functionals. *Physics Letters B*, 800:135065, 1 2020.

[18] Shun-ichi Amari. Natural Gradient Works Efficiently in Learning. *Neural Computation*, 10(2):251–276, 2 1998.

[19] Sheng Ma, Chuanyi Ji, and James Farmer. An Efficient EM-based Training Algorithm for Feedforward Neural Networks. *Neural Networks*, 10(2):243–256, 3 1997.

[20] Shun-Ichi Amari. Differential Geometry of Curved Exponential Families-Curvatures and Information Loss. *The Annals of Statistics*, 10(2):357–385, 6 1982.

[21] Shun-ichi Amari. *Information Geometry and Its Applications*. Applied Mathematical Sciences. Springer Japan, Tokyo, 2016.

[22] Mark K. Transtrum, Benjamin B. MacHta, and James P. Sethna. Why are nonlinear fits to data so challenging? *Physical Review Letters*, 2010.

[23] Mark K. Transtrum, Benjamin B. MacHta, and James P. Sethna. Geometry of nonlinear least squares with applications to sloppy models and optimization. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 2011.

[24] Mark K. Transtrum and Peng Qiu. Model reduction by manifold boundaries. *Physical Review Letters*, 2014.

[25] Mark K. Transtrum, Benjamin B. Machta, Kevin S. Brown, Bryan C. Daniels, Christopher R. Myers, and James P. Sethna. Perspective: Sloppiness and emergent theories in physics, biology, and beyond. *Journal of Chemical Physics*, 2015.

[26] Mark K Transtrum. Manifold boundaries give “gray-box” approximations of complex models. *arXiv e-prints*, page arXiv:1605.08705, 5 2016.

[27] K. Tisanić, V. Smolčić, M. Imbrišak, M. Bondi, G. Zamorani, L. Ceraj, E. Vardoulaki, and J. Delhaize. The VLA-COSMOS 3 GHz Large Project: Average radio spectral energy distribution of active galactic nuclei. *Astronomy & Astrophysics*, 643, 11 2020.

[28] Tamara Nikšić and Dario Vretenar. "sloppy" nuclear energy density functionals: Effective model reduction. *Physical Review C*, 94(2), 2016.

[29] T. Nikšić, M. Imbrišak, and D. Vretenar. Sloppy nuclear energy density functionals. II. Finite nuclei. *Physical Review C*, 95(5), 5 2017.

[30] Benjamin L. Francis and Mark K. Transtrum. Unwinding the model manifold: Choosing similarity measures to remove local minima in sloppy dynamical systems. *Phys. Rev. E*, 100(1):012206, July 2019.

[31] S. Typel and H.H. Wolter. Relativistic mean field calculations with density-dependent meson-nucleon coupling. *Nuclear Physics A*, 656(3-4), 9 1999.

[32] P. Finelli, N. Kaiser, D. Vretenar, and W. Weise. Relativistic nuclear model with point-couplings constrained by QCD and chiral symmetry. *Nuclear Physics A*, 735(3-4), 5 2004.

[33] P. Finelli, N. Kaiser, D. Vretenar, and W. Weise. Relativistic nuclear energy density functional constrained by low-energy QCD. *Nuclear Physics A*, 770(1-2), 5 2006.

[34] M. Baldo, P. Schuck, and X. Viñas. Kohn–Sham density functional inspired approach to nuclear binding. *Physics Letters B*, 663(5), 6 2008.

[35] M. Baldo, L. M. Robledo, P. Schuck, and X. Viñas. New Kohn-Sham density functional based on microscopic nuclear and neutron matter equations of state. *Physical Review C*, 87(6), 6 2013.

[36] T. Nikšić, D. Vretenar, and P. Ring. Relativistic nuclear energy density functionals: Adjusting parameters to binding energies. *Physical Review C - Nuclear Physics*, 78(3), 9 2008.

[37] T. Nikšić, D. Vretenar, and P. Ring. Relativistic nuclear energy density functionals: Mean-field and beyond, 7 2011.

[38] M. Kortelainen, J. McDonnell, W. Nazarewicz, P.-G. Reinhard, J. Sarich, N. Schunck, M. V. Stoitsov, and S. M. Wild. Nuclear energy density optimization: Large deformations. *Physical Review C*, 85(2), 2 2012.

[39] M. Kortelainen, J. McDonnell, W. Nazarewicz, E. Olsen, P.-G. Reinhard, J. Sarich, N. Schunck, S. M. Wild, D. Davesne, J. Erler, and A. Pastore. Nuclear energy density optimization: Shell structure. *Physical Review C*, 89(5), 5 2014.

[40] Aurel Bulgac, Michael McNeil Forbes, and Shi Jin. Nuclear Energy Density Functionals: What do we really know? *arXiv e-prints*, page arXiv:1506.09195, June 2015.

[41] Ryan N Gutenkunst, Joshua J Waterfall, Fergal P Casey, Kevin S Brown, Christopher R Myers, and James P Sethna. Universally Sloppy Parameter Sensitivities in Systems Biology Models. *PLoS Computational Biology*, 3(10), 10 2007.

[42] T. Nikšić, N. Paar, P. G. Reinhard, and D. Vretenar. Optimizing relativistic energy density functionals: covariance analysis. *Journal of Physics G Nuclear Physics*, 42(3):034008, March 2015.

[43] P. G. Reinhard and W. Nazarewicz. Information content of a new observable: The case of the nuclear neutron skin. *Phys. Rev. C*, 81(5):051303, May 2010.

[44] Joshua J. Waterfall, Fergal P. Casey, Ryan N. Gutenkunst, Kevin S. Brown, Christopher R. Myers, Piet W. Brouwer, Veit Elser, and James P. Sethna. Sloppy-Model Universality Class and the Vandermonde Matrix. *Phys. Rev. Lett.*, 97(15):150601, October 2006.

[45] B. B. Machta, R. Chachra, M. K. Transtrum, and J. P. Sethna. Parameter Space Compression Underlies Emergent Theories and Predictive Models. *Science*, 342(6158):604–607, 11 2013.

[46] Y. K. Gambhir, P. Ring, and A. Thimet. Relativistic mean field theory for finite nuclei. *Annals of Physics*, 198(1):132–179, February 1990.

[47] Y. K. Gambhir and P. Ring. Recipe for the Solution of the Dirac Equation by the Basis Set Expansion Method. *Modern Physics Letters A*, 8(9):787–795, January 1993.

[48] P. Ring, Y. K. Gambhir, and G. A. Lalazissis. Computer program for the relativistic mean field description of the ground state properties of even-even axially deformed nuclei. *Computer Physics Communications*, 105(1):77–97, September 1997.

[49] T. Nikšić, N. Paar, D. Vretenar, and P. Ring. DIRHB-A relativistic self-consistent mean-field framework for atomic nuclei. *Computer Physics Communications*, 185(6):1808–1821, June 2014.

[50] M. Imbrišak and K. Nomura. Stability of the manifold boundary approximation method for reductions of nuclear structure models. *Phys. Rev. C*, 107:034304, Mar 2023.

[51] W. Koepf and P. Ring. The spin-orbit field in superdeformed nuclei: a relativistic investigation. *Zeitschrift fur Physik A Hadrons and Nuclei*, 339(1):81–90, March 1991.

[52] D Vretenar, A Afanasjev, G Lalazissis, and P Ring. Relativistic Hartree-Bogoliubov theory: static and dynamic aspects of exotic nuclear structure. *Physics Reports*, 409(3-4):101–259, 4 2005.

[53] Michael Bender, Paul-Henri Heenen, and Paul-Gerhard Reinhard. Self-consistent mean-field models for nuclear structure. *Reviews of Modern Physics*, 75(1):121–180, 1 2003.

[54] P. Finelli, N. Kaiser, D. Vretenar, and W. Weise. Relativistic nuclear model with point-couplings constrained by QCD and chiral symmetry. *Nuclear Physics A*, 735(3-4):449–481, 5 2004.

[55] S. E. Agbemava, A. V. Afanasjev, D. Ray, and P. Ring. Global performance of covariant energy density functionals: Ground state observables of even-even nuclei and the estimate of theoretical uncertainties. *Physical Review C*, 89(5):054320, 5 2014.

[56] P. Ring. Relativistic mean field theory in finite nuclei. *Progress in Particle and Nuclear Physics*, 37:193–263, January 1996.

[57] D. Vautherin. Hartree-Fock Calculations with Skyrme's Interaction. II. Axially Deformed Nuclei. *Phys. Rev. C*, 7(1):296–316, January 1973.

[58] F. de Jong and H. Lenske. Asymmetric nuclear matter in the relativistic Brueckner-Hartree-Fock approach. *Phys. Rev. C*, 57(6):3099–3107, June 1998.

[59] S. Typel and H. H. Wolter. Relativistic mean field calculations with density-dependent meson-nucleon coupling. *Nucl. Phys. A*, 656(3):331–364, September 1999.

[60] John M. Lee. *Introduction to Riemannian Manifolds*. Graduate Texts in Mathematics. Springer International Publishing, Cham, 2018.

[61] J. F. Nolan. Analytical differentiation on a digital computer. Master's thesis, Massachusetts Institute of Technology. Dept. of Mathematics, 1953.

[62] Atilim Gunes Baydin, Barak A. Pearlmutter, Alexey Andreyevich Radul, and Jeffrey Mark Siskind. Automatic differentiation in machine learning: a survey. *Journal of Machine Learning Research*, 18(153):1–43, 2018.

[63] Arun Verma. An introduction to automatic differentiation. *Current Science*, 78(7):804–807, 2000.

[64] Andreas Griewank and Andrea Walther. *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*. Society for Industrial and Applied Mathematics, USA, second edition, 2008.

[65] Dougal Maclaurin, David Duvenaud, and Ryan P Adams. Autograd: Effortless gradients in numpy. In *ICML 2015 AutoML Workshop*, volume 238, page 5, 2015.

[66] S. E. Agbemava, A. V. Afanasjev, T. Nakatsukasa, and P. Ring. Covariant density functional theory: Reexamining the structure of superheavy nuclei. *Physical Review C*, 92(5):054310, 11 2015.

[67] Pauli Virtanen, Ralf Gommers, Travis E Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J van der Walt, Matthew Brett, Joshua Wilson, K Jarrod Millman, Nikolay Mayorov, Andrew R J Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E A Quintero, Charles R Harris, Anne M Archibald, Antônio H Ribeiro, Fabian

Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. {SciPy} 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020.

[68] Daniel Foreman-Mackey, David W. Hogg, Dustin Lang, and Jonathan Goodman. emcee: The MCMC Hammer. *Publications of the ASP*, 125(925):306, March 2013.