Determination of the beam asymmetry Σ in η- and η'-photoproduction using Bayesian statistics

Jakob Michael Krause

Masterarbeit in Physik
angefertigt im Helmholtz-Institut für Strahlen- und Kernphysik

vorgelegt der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität
Bonn

Sep 2022

I hereby declare that this thesis was formulated by myself and that no sources or tools other than those cited were used.

Bonn, ...
Date Signature

1. Gutachterin: Jun. Prof. Dr. Annika Thiel
2. Gutachter: Prof. Dr. Jochen Dingfelder

Contents

CHAPTER 1

Introduction

The *Standard Model of Particle Physics* (SM) is the most successful model aiming to describe the particles and forces of the universe. It distinguishes between *fermions* and *bosons*. While all matter consists of fermions, bosons are particles that mediate the fundamental interactions.

Matter consists of (anti-)quarks and (anti-)leptons with three generations of each. Table 1.1 shows all elementary fermions including some of their most important properties. Only the first and lightest generation consists of stable particles, i.e. the up and down quark as well as the electron and its neutrino. All other particles are heavier and not stable, they will thus decay fast via the strong, electromagnetic or weak interaction.

There are in fact four interactions described by the SM: strong, electromagnetic, weak and gravitational interaction [1], where gravitation is mentioned here for the sake of completeness; on the mass scale of elementary particles gravitation is negligible. Strong and weak interaction are restricted to a finite range of the order of the nucleon radius, whereas electromagnetic interaction and gravitation have infinite range. Each interaction has its own coupling (charge). The strong interaction is mediated by gluons and couples to the color charge.

	Generation			**el. charge**	**color charge**
	1	**2**	**3**		
Quarks	u	c	t	2/3	r,g,b
	d	s	b	1/3	r,g,b
Leptons	e	μ	τ	-1	-
	ν_e	ν_μ	ν_τ	0	-

Table 1.1: Summary of the particles of the SM

Gluons and quarks carry color charge and thus interact strongly. However, an isolated quark or gluon has not been observed. Only color neutral bound systems of quarks are seen, which are called hadrons. Hadrons with integer spin are called mesons and those with half-integer spin are called baryons. Color neutrality demands mesons consist of at least one quark and one anti-quark and baryons consist of at least three quarks.

[1] they are ordered here according to their relative strength

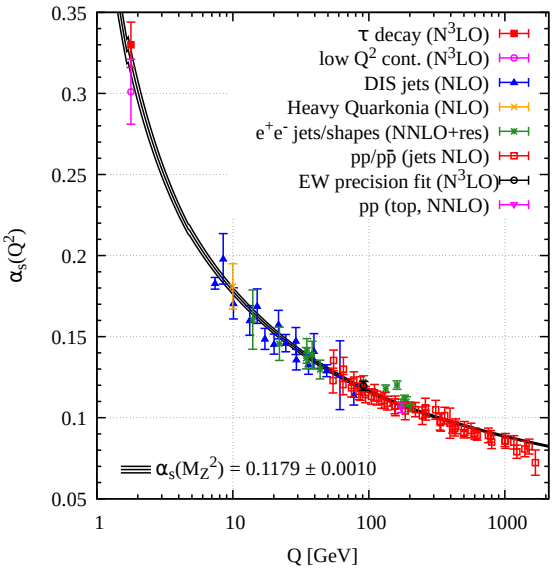

Figure 1.1: Running coupling of QCD. The colored data points represent different methods to obtain a value for α_s. For a detailed review see [Wor+22].

As already mentioned, isolated quarks are not seen. This can be understood in terms of the strong coupling constant α_s. The coupling constant is a measure of the strength of the strong interaction. Because it is highly dependent on the momentum transfer in the observed strong reaction it is also called running coupling constant, which is depicted in figure 1.1.

For low (< 1 GeV) momentum transfers or large distances the coupling constant approaches infinity whereas it decreases for high ($\gg 1$ GeV) momentum transfers or short distances. These momentum ranges are referred to as *confinement* and *asymptotic freedom*, respectively; quarks are confined to remain in a bound state since if one tried to pull them apart the color field becomes so strong it will create a new quark anti-quark pair resulting in two new bound states. On the other hand, bound quarks behave quasi-free and can be described using perturbative quantum chromodynamics (pQCD) if probed at sufficiently large momentum transfers.

It is more difficult however to describe QCD at momentum scales of ≈ 1 GeV since the coupling is too strong to justify a perturbative approach. Thus explicit modeling of QCD bound states is inevitable. One possibility is to describe baryons consisting of constituent quarks which are bound in a potential. Constituent quark models assume baryons are made up of three constituent quarks with effective masses differing from the bare quark mass. The effective mass is made up mostly from a sea of quark anti-quark pairs and gluons which surround the bare (valence) quarks. The explicit form of the binding potential is determined for each model.

The Bonn model [LMP01], for example, is formulated as a relativistically covariant constituent quark model. A potential increasing linearly with the distance is employed to adequately describe confinement. The binding potential between the constituent quarks is described by an instanton-induced interaction. Baryon resonances are then states with an orbital or angular excitation of one of the quarks. Figure 1.2 shows computed nucleon resonances with Isospin $I = 1/2$ of the Bonn model [LMP01] on the left side of each column. These are compared to measured resonances and their PDG rating [Wor+22] in the middle. Uncertainties are indicated by the colored areas. The resonances are

identified by their total angular momentum and their parity $J\pi$. In addition also the total internal angular momentum along with isospin and again the total angular momentum $L_{2T\,2J}$ is given. While

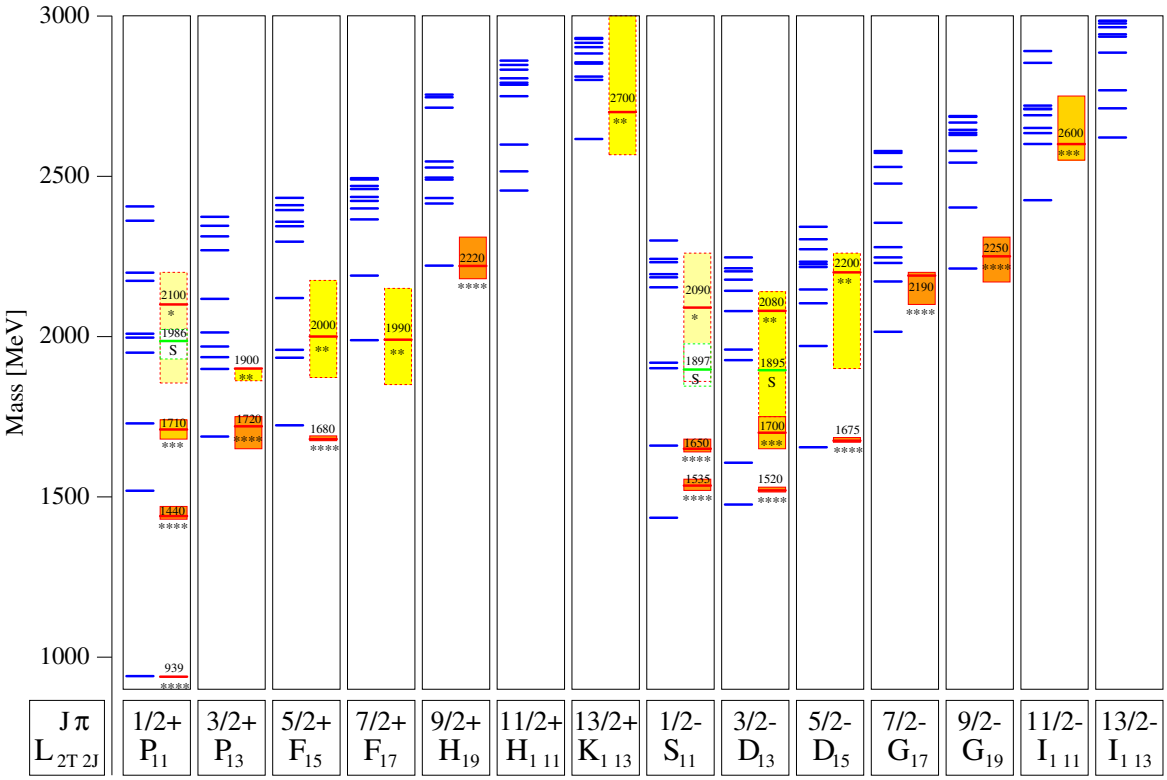

Figure 1.2: Calculated nucleon (isospin $I = 1/2$) resonances compared to measurements. Left in each column are the calculations [LMP01], the middle shows the measurements and PDG rating [Wor+22]

generally good agreement exists for low lying resonances, especially for high masses there are much more resonances predicted than actually found. This is also known as the problem of the "missing resonances" indicating the poor understanding of QCD in the non-perturbative region. This can be due to several reasons: most of the knowledge about nucleon resonances and their properties was obtained investigating the πN channel, biasing the data for resonances coupling weakly to this channel. Furthermore, the number of excited states with definite quantum numbers is related directly to the effective number of degrees-of-freedom accessible to the underlying theory. As a consequence, the number of degrees-of-freedom should be obtainable by comparing the measured states to the predicted states. Since nucleon resonances decay dominantly hadronic, their resonances are broad and overlapping. Thus on one hand the determination of excitation spectra proves to be a challenge on its own, demanding sophisticated methods, such as partial wave analysis (PWA). On the other hand it is not yet clear how many effective degrees-of-freedom exist for the nucleon in a constituent quark model. They could for example be decreased if the nucleon were made up of a quark and a di-quark structure. To access nucleon resonances and transitions between them for as many final states as possible the photoproduction of mesons off the proton has gained attention. Experiments dedicated to study photoproduction reactions off nucleons are located e.g. at JLAB, MAMI or ELSA. In this thesis photoproduction data taken at the CBELSA/TAPS experiment, which is located at

the accelerator ELSA in Bonn, is analyzed regarding the reactions $\gamma p \rightarrow p\eta$ and $\gamma p \rightarrow p\eta'$. The theoretical foundation underlying the photoproduction of pseudoscalar mesons and the measurement of polarization observables will be discussed in the next sessions.

1.1 Photoproduction of Pseudoscalar Mesons

From the scattering theory point of view, photoproduction of mesons is well understood [KS03]. Figure 1.3 shows schematically the process thereof off the proton:

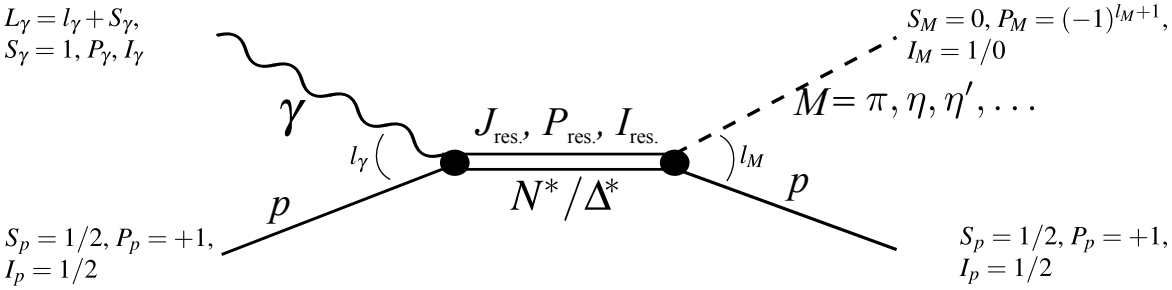

Figure 1.3: FEYNMAN diagram for the s-channel photoproduction of pseudoscalar mesons, adapted from [Afz19]

The analysis requires partial wave decomposition in both initial and final states [DT92] since the intermediate resonance N^*/Δ^* has definite angular momentum, parity and isospin $J_{\text{res.}}, P_{\text{res.}}, I_{\text{res.}}$. The resonance is excited by a photon with (iso-) spin $I_\gamma, S_\gamma = 1$ and parity P_γ coupling electromagnetically to the target proton with (iso-) spin $I_p = 1/2, S_p = 1/2$ and parity P_p. The relative momentum is l_γ, such that the total momentum of the photon is $L_\gamma = l_\gamma + S_\gamma$. The parity of the photon depends on the multipolarity of the photon and is given by $P_\gamma = (-1)^{L_\gamma}$ for electric (E) or $P_\gamma = (-1)^{L_\gamma+1}$ for magnetic (M) photon multipoles [DT92]. Subsequently the intermediate state will have the quantum numbers $J_{\text{res.}}, P_{\text{res.}}, I_{\text{res.}}$ and decay into a proton with spin $S_p = 1/2$, parity $P_p = +1$ and isospin $I_p = 1/2$ under emission of a meson. Here, only pseudoscalar mesons that have vanishing spin $S_M = 0$, isospin I_M, relative orbital angular momentum l_M und Parity $P_M = (-1)^{l_M+1}$ are considered. Note that for η and η' mesons $I_M = 0$; this exclusively limits the intermediate resonances to N^* states since the strong interaction conserves isospin [Wor+22]. The following selection rules can be derived using parity and momentum conservation [KS03; Afz19]

$$J_{\text{res.}} = L_\gamma \oplus S_p = L_\gamma \oplus 1/2, \tag{1.1}$$

$$P_{\text{res.}} = P_p \cdot P_\gamma = P_\gamma, \tag{1.2}$$

$$J_{\text{res.}} = l_M \oplus S_p = l_M \oplus 1/2 \tag{1.3}$$

$$P_{\text{res.}} = P_p \cdot P_M = (-1)^{l_M+1}, \tag{1.4}$$

where the usual rules for the coupling $\oplus$ of angular momenta [Bar+18] apply. Thus, knowledge of the photoproduction multipoles allows the identification of contributing resonances for particular mesonic final states. Table 1.2 shows a summary of allowed quantum numbers according to the shown selection rules for the lowest order of photon multipoles ($L_\gamma = 1$). The photoproduction multipoles $E_{l\pm}, M_{l\pm}$ indicate the relative momentum of the meson ($l = l_M$) and whether the total

angular momentum is obtained by adding "+" or substracting "–" the final state momenta. Resonances are identified in spectroscopic notation by the meson momentum l_M as well as by their (iso-)spin I, J and Mass M. Note that only $2I = 1$ resonances are listed since the mesons η and η' have vanishing isospin, so that only $I = 1/2$ resonances (N^* resonances) may be accessed. The determination

Photon multipole	**initial state** $\left(L_\gamma^{P_\gamma}, S_p^{P_p}\right)$	**intermed. state** $J_{\text{res}}^{P_{\text{res}}}$	**final state** $\left(S_p^{P_p}, l_M^{P_M}\right)$	**photoproduction multipole** $E_{l\pm}, M_{l\pm}$	**resonance** $\left(l_M\right)_{2I2J}\,(M)$
$E1$	$\left(1^-, \frac{1}{2}^+\right)$	$\frac{1}{2}^-$	$\left(\frac{1}{2}^+, 0^-\right)$	E_{0+}	$S_{13}(M)$
$E1$	$\left(1^-, \frac{1}{2}^+\right)$	$\frac{3}{2}^-$	$\left(\frac{1}{2}^+, 2^-\right)$	E_{2-}	$D_{13}(M)$
$M1$	$\left(1^+, \frac{1}{2}^+\right)$	$\frac{1}{2}^+$	$\left(\frac{1}{2}^+, 1^+\right)$	M_{1-}	$P_{11}(M)$
$M1$	$\left(1^+, \frac{1}{2}^+\right)$	$\frac{3}{2}^+$	$\left(\frac{1}{2}^+, 1^+\right)$	M_{1+}	$P_{13}(M)$

Table 1.2: Allowed quantum numbers for the intermediate resonance state N^* in η/η'-photoproduction. Adapted from [Afz19]

of contributing multipoles which can then be used to identify nucleon resonances is challenging and requires sophisticated (model dependent) partial wave analyses (PWA). The measurement of polarization observables helps to eliminate ambiguities in PWA calculations as will be explained in the following section.

1.2 Measurement of Polarization Observables

Using an ansatz purely motivated by scattering theory, the differential cross section of meson photoproduction can be written as

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{q}{k} \left|\langle f|\mathcal{F}|i\rangle\right|^2, \tag{1.5}$$

where the matrix element is taken between initial and final Pauli spinors [Che+57] and q and k denote the momentum of the incident photon and final state meson, respectively. The photoproduction amplitude $\mathcal{F}$ contains all relevant information regarding the scattering process connecting the initial and final state in analogy to the S-matrix that is introduced in the general discussion of quantum-mechanical scattering [PS95]. Following the notation of reference [Che+57] it can be written as a sum of the complex Chew-Goldberger-Low-Nambu (CGLN) amplitudes F_i

$$\mathcal{F} = i(\vec{\sigma}\cdot\vec{\epsilon})F_1 + (\vec{\sigma}\cdot\hat{q})(\vec{\sigma}\cdot(\hat{k}\times\vec{\epsilon}))F_2 + i(\vec{\sigma}\cdot\hat{k})(\hat{q}\cdot\vec{\epsilon})F_3 + i(\vec{\sigma}\cdot\hat{q})(\hat{q}\cdot\vec{\epsilon})F_4, \tag{1.6}$$

where $\hat{k}$ and $\vec{\epsilon}$ are the momentum unit vector and polarization vector of the incident photon, $\hat{q}$ is the momentum unit vector of the final state meson and $\vec{\sigma}$ denote the Pauli matrices. Applying the method of partial waves onto the complex CGLN amplitudes F_i one can absorb the angular dependence into Legendre polynomials $P_l(\cos\theta)$ and derivatives thereof [2] and express the energy dependence solely

[2] Here $\cos\theta$ denotes the polar angle of the meson in the center of mass system

by the photoproduction multipoles $E_{l\pm}$, $M_{l\pm}$, e.g. for F_1 one finds[Che+57]

$$F_1 = \sum_{l=0}^{\infty} \left[l M_{l+}(W) + E_{l+}(W) \right] P'_{l+1}(\cos\theta) + \left[(l+1) M_{l-}(W) + E_{l-}(W) \right] P'_{l-1}(\cos\theta), \tag{1.7}$$

where W is the center of mass energy. Inserting the partial wave expansion of all CGLN amplitudes into Equation (1.5) directly connects the CGLN amplitudes and the photoproduction multipoles to the cross section which is a measurable quantity. However, the measurement of the differential cross section will give one real number which is not sufficient to unambigously determine four complex amplitudes. To be able to access further observables that are connected to the CGLN amplitudes and the photoproduction multipoles, a polarized target and/or a polarized photon beam can be employed. In total, one may measure 16 non-redundant polarization observables [San+11] that are grouped into four categories: single polarization observables, where either the beam photon, target proton or recoil proton are polarized, and three groups of double polarization observables where two of the mentioned particles are polarized, i.e. the three groups are beam-target (BT), beam-recoil (BR) and target-recoil (TR) observables. In Table 1.3 all single and double polarization observables in pseudoscalar meson photoproduction are listed. Eight carefully chosen observables allow an unambigous determination of all CGLN amplitudes [CT97], which are then referred to as a *complete experiment* [CT97; San+11].

category	**observables**			
non /single	$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}$ /	Σ	T	P
beam-target	G	H	E	F
beam-recoil	$O_{x'}$	$O_{z'}$	$C_{x'}$	$C_{z'}$
target-recoil	$T_{x'}$	$T_{z'}$	$L_{x'}$	$L_{z'}$

Table 1.3: All 16 non redundant polarization observables in pseudoscalar meson photoproduction [San+11]. Table adapted from [Afz12].

The CBELSA/TAPS experiment is able to produce a linearly or circularly polarized photon beam as well as a longitudinally or transersely polarized target. This introduces dependencies on the azimuthal angle into the differential cross section that couple to polarization observables Σ, T, P, E, F, G, H and the linear or circular polarization degrees of the photon beam $p_\gamma^{\mathrm{lin}}, p_\gamma^{\mathrm{circ}}$ as well as the target polarization p_x, p_y, p_z[San+11]

$$\begin{aligned} \frac{d\sigma}{d\Omega}\left(E_\gamma, \theta, \varphi\right) = \frac{d\sigma}{d\Omega}\left(E_\gamma, \theta\right) \Big[1 - p_\gamma^{\mathrm{lin}} \mathbf{\Sigma} \cos(2\varphi) + p_x \left(p_\gamma^{\mathrm{lin}} \mathbf{H} \sin(2\varphi) + p_\gamma^{\mathrm{circ}} \mathbf{F} \right) \\ - p_y \left(p_\gamma^{\mathrm{lin}} \mathbf{P} \cos(2\varphi) - \mathbf{T} \right) - p_z \left(- p_\gamma^{\mathrm{lin}} \mathbf{G} \sin(2\varphi) + p_\gamma^{\mathrm{circ}} \mathbf{E} \right) \Big] . \end{aligned} \tag{1.8}$$

In this thesis data was analyzed that was taken with a linearly polarized photon beam and an unpolarized target, so that (1.8) reduces to

$$\frac{d\sigma}{d\Omega}\left(E_\gamma, \theta, \varphi\right) = \frac{d\sigma}{d\Omega}\left(E_\gamma, \theta\right) \left[1 - p_\gamma^{\mathrm{lin}} \mathbf{\Sigma} \cos(2\varphi) \right], \tag{1.9}$$

which allows the determination of the beam asymmetry Σ, as is described in detail in chapter 4. The beam asymmetry can then be related to the CGLN amplitudes in order to enable the determination of

photoproduction multipoles if it is multiplied by the unpolarized cross section [FTS92]

$$\widehat{\Sigma} = \Sigma \cdot \left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_0 \propto \frac{\sin^2(\theta)}{2} \Re \left[\left|F_3\right|^2 + \left|F_4\right|^2 + 2\left(F_1^* F_4 + F_2^* F_3 + \cos\theta F_3^* F_4\right)\right]. \tag{1.10}$$

Unlike the unpolarized cross section, the beam asymmetry, or rather polarization observables in general, is sensitive not only to the absolute values of the photoproduction multipoles squared but is also sensitive to interference terms thereof [Afz19; Wun+17]. Ultimately the photoproduction multipoles can be linked to resonance properties like mass M and width Γ using (model dependent) partial wave analyses that consider many observables in different final states at once. Expanding the database of polarization observables helps to further eliminate ambiguities in the photoproduction multipoles, which are used to identify contributing resonances, and thus adds to the understanding of the strong interaction in the non perturbative regime.

1.3 Introduction to Bayesian statistics

To determine the beam asymmetry (see Chapter 4) Bayesian methods are applied. This section will give a short introduction of the used concepts regarding Bayesian inference and the implementation of such a Bayesian analysis.

1.3.1 Notation

First of all, a probabilistic notation is introduced that will consequently be used throughout the remainder of this thesis to ease the formulation of Bayesian models and inferences. Hereby the Bayesian approach is directly applied to the context of parameter inference.

Bayesian parameter inference aims to draw statistical conslusions about parameters $\boldsymbol{\theta} = \{\theta_1, \theta_2, \ldots, \theta_N\}$ conditioned on observed data y in the form of probability statements [Gel+14]. The probability density of the introduced parameters $\boldsymbol{\theta}$ *given* the observed data y is written as

$$p(\boldsymbol{\theta}|y). \tag{1.11}$$

Distributions that are not conditioned on other observables, i.e. marginal or prior distributions, are notated as e.g.

$$p(\boldsymbol{\theta}). \tag{1.12}$$

If a parameter θ follows a well known probability density function (PDF) like a Gaussian $\mathcal{N}$ with mean μ and standard deviation σ or Poisson distribution $\mathcal{P}$ with mean $\tilde{\mu}$ this is notated as

$$\theta \sim \mathcal{N}(\mu, \sigma) \quad \Leftrightarrow \quad p(\theta) = \mathcal{N}(\theta|\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\theta-\mu)^2}{2\sigma^2}}, \tag{1.13}$$

$$\tilde{\theta} \sim \mathcal{P}(\tilde{\mu}) \quad \Leftrightarrow \quad p(\tilde{\theta}) = \mathcal{P}(\tilde{\theta}|\tilde{\mu}) = \frac{\tilde{\mu}^{\theta}}{\tilde{\theta}!} e^{-\tilde{\mu}}. \tag{1.14}$$

1.3.2 Bayes' theorem

Bayes' theorem allows to link the conditional probabilities $p(\boldsymbol{\theta}|y)$ and $p(y|\boldsymbol{\theta})$ and can be formulated as [Gel+14]

$$p(\boldsymbol{\theta}|y) = \frac{p(y|\boldsymbol{\theta}) \cdot p(\boldsymbol{\theta})}{p(y)}. \tag{1.15}$$

In this context, $p(y|\boldsymbol{\theta})$ is called the *likelihood* that the data y are described by the parameters $\boldsymbol{\theta}$. The second factor in the enumerator on the right hand side of Eq. (1.15) is called the *prior* of the parameters $\boldsymbol{\theta}$. It gives their probability density prior to acquiring any information from the model. The denominator $p(y)$ is a normalizing constant, for which it holds

$$p(y) = \int_{\boldsymbol{\theta}} \mathrm{d}\boldsymbol{\theta} p(y|\boldsymbol{\theta}) \cdot p(\boldsymbol{\theta}). \tag{1.16}$$

With increasing complexity of the investigated data and model, evaluating the integral (1.16) can become challenging, if not impossible. But since for a fixed dataset it is in fact only a normalizing constant, one can choose to not evaluate it to arrive at the unnormalized *posterior* on the left hand side of Eq. 1.15 and 1.17

$$p(\boldsymbol{\theta}|y) \propto p(y|\boldsymbol{\theta}) \cdot p(\boldsymbol{\theta}). \tag{1.17}$$

The posterior gives the probability density function of the parameters $\boldsymbol{\theta}$ conditioned on the observed data y. For each parameter $\theta_n \in \boldsymbol{\theta}$ a one dimensional marginal posterior can be determined by integrating out all other parameters [SS05]

$$p(\theta_n|y) = \int \mathrm{d}\theta_1 \cdots \int \mathrm{d}\theta_{n-1} \int \mathrm{d}\theta_{n-1} \cdots \int \mathrm{d}\theta_N p(\theta_1 \ldots \theta_N|y). \tag{1.18}$$

Determining marginal posteriors is the main goal of a Bayesian parameter inference. They give probability densities for each individual parameter based on the observed data and are the equivalent to point estimates with error bars that are e.g. determined from a χ^2 parameter fit but at the same time yield full distributions as a result. The determination of marginal posteriors with the ansatz formulated in Eq. 1.18 can become a highly non-trivial task using analytical methods with increasing number of parameters and complexity of the investigated model, so that the use of Monte-Carlo sampling is a suitable approach.

1.3.3 Markov-Chain-Monte-Carlo (MCMC)

Consider using Eq. (1.17) to determine unnormalized marginal posteriors for all parameters $\boldsymbol{\theta}$ without carrying out the integrations (1.18). This can be achieved by approximating the multidimensional joint unnormalized posterior $p(\boldsymbol{\theta}|y)$ using a large number of simulation draws $\boldsymbol{\theta}^{(s)}$ and projecting out each parameter θ_n while ignoring all other parameters $\theta_{k\neq n}$ [Tro08]. Thus, to perform a parameter inference, the main task is to accomplish the drawing of samples $\boldsymbol{\theta}^{(s)}$ that follow the joint posterior in order to access the marginal posteriors.

Markov chains are a finite sequence of random variables $\boldsymbol{\theta}^1, \boldsymbol{\theta}^2 \ldots \boldsymbol{\theta}^S$ where for each step t the quantity $\boldsymbol{\theta}^t$ only depends on the previous $\boldsymbol{\theta}^{t-1}$ and is independent of all other previous chain elements [Gel+14]. Such a chain is created by starting at some initial value $\boldsymbol{\theta}^0$ and generate new draws from a transition distribution $T_t(\boldsymbol{\theta}^t|\boldsymbol{\theta}^{t-1})$. The transition probabilities T_t can be constructed such that the

Markov chain reaches a stationary distribution which is the desired joint posterior (1.17) [Gel+14; Nor97]. A variety of algorithms exist to ensure the convergence of Markov chains.

Metropolis-Hastings algorithm

A first approach is the Metropolis-Hastings algorithm [Met+53; Has70]. Here, at step t in the chain random proposals $\boldsymbol{\theta}^*$ are drawn from the transition distribution $T_t(\boldsymbol{\theta}^*|\boldsymbol{\theta})$. To proceed in the algorithm, the ratio of posteriors

$$r = \frac{p\left(\boldsymbol{\theta}^*|y\right) \cdot T_t\left(\boldsymbol{\theta}^{t-1}|\boldsymbol{\theta}^*\right)}{p\left(\boldsymbol{\theta}^{t-1}|y\right) \cdot T_t\left(\boldsymbol{\theta}^*|\boldsymbol{\theta}^{t-1}\right)} \tag{1.19}$$

is considered. If the new proposal $\boldsymbol{\theta}^*$ increases the posterior density, i.e. $r > 0$, it is accepted with probability 1, such that $\boldsymbol{\theta}^t = \boldsymbol{\theta}^*$. In the case $r < 0$ the update $\boldsymbol{\theta}^t = \boldsymbol{\theta}^*$ is only accepted with probability r and otherwise discarded so that $\boldsymbol{\theta}^t = \boldsymbol{\theta}^{t-1}$ [Gel+14]. The Metropolis-Hastings algorithm exhibits the behavior of a random walk in the parameter space [Gel+14]. Figure 1.4 shows an example of five independent Markov-chains generated with the described algorithm for a two dimensional parameter $\boldsymbol{\theta}$ with a bivariate normal distribution as posterior $\boldsymbol{\theta} \sim \mathcal{N}(0, \mathbb{1}_2)$ at different stages in the simulation [Gel+14]. It is evident that Markov chains require a "burn-in" period before the simulation draws are a valid approximation of the target distribution that has to be chosen in the context of the applied model [Gel+14].

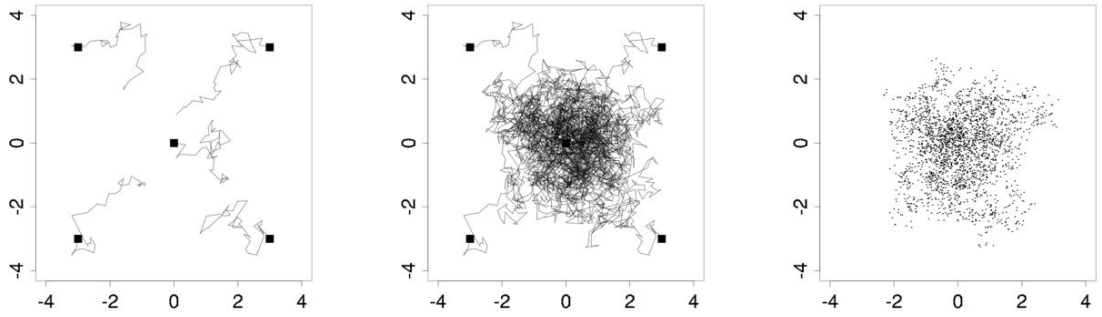

Figure 1.4: Five different MCMC simulations for a bivariate normal distribution, with starting points marked by black squares. Left: After 50 iterations no convergence is yet to be seen. Middle: After 1000 iterations the chains are closer to convergence. Right: Removing the first halve of simulation draws leaves a set of draws from the target distribution. Taken from [Gel+14]

Hamiltonian Monte-Carlo (HMC) and No-U-Turn-Sampling (NUTS)

The random walk behavior of the Metropolis-Hastings algorithm can make computation of complex models very inefficient [Gel+14], therefore more sophisticated methods have been developed. Hamiltonian Monte-Carlo (HMC) [Dua+87] is an elaborate algorithm that allows to explore the parameter space much more efficient than a simple random walk. It will be described qualitatively in the following, for a detailed discussion see e.g. reference [Gel+14].

At the heart of the HMC algorithm lies an accept-reject step that is similar to the Metropolis-Hastings algorithm. To arrive at new proposals $\boldsymbol{\theta}^*$ at step t in the chain however an additional auxiliary variable $\boldsymbol{\phi}$ is introduced that has the same dimension as the parameters $\boldsymbol{\theta}$ but is independent of $\boldsymbol{\theta}$ and the data y. Introducing the new parameters leads to the joint posterior [Sta22]

$$\begin{aligned} & p(\boldsymbol{\theta}, \boldsymbol{\phi}|y) = p(\boldsymbol{\theta}|y) \cdot p(\boldsymbol{\phi}|\boldsymbol{\theta}, y) = p(\boldsymbol{\theta}|y) \cdot p(\boldsymbol{\phi}) \\ \Leftrightarrow & -\log p(\boldsymbol{\theta}, \boldsymbol{\phi}|y) = -\log p(\boldsymbol{\theta}|y) - \log p(\boldsymbol{\phi}) \\ & \overset{\text{Def.}}{\Leftrightarrow} H(\boldsymbol{\theta}, \boldsymbol{\phi}) := V(\boldsymbol{\theta}) + T(\boldsymbol{\phi}), \end{aligned} \tag{1.20}$$

where in the last step the artificial Hamiltonian H has been introduced as a sum of kinetic energy T and potential V. Usually, one chooses $p(\boldsymbol{\phi})$ to be a multivariate normal centered at 0 with covariance matrix M [Gel+14]. Applying the well known equations of motion to this Hamiltonian one finds

$$\frac{\mathrm{d}\theta_i}{\mathrm{d}t} = \frac{\partial H}{\partial \phi_i} = \frac{\partial T}{\partial \phi_i} \qquad \frac{\mathrm{d}\phi_i}{\mathrm{d}t} = -\frac{\partial H}{\partial \theta_i} = -\frac{\partial V}{\partial \theta_i}. \tag{1.21}$$

This allows the simultaneous direct numerical integration to evolve "position" and "momentum" in $L \cdot \epsilon$ discrete time steps with the so-called *Leap-Frog* algorithm [KN98]. Consequently L is called the number of Leap-Frog steps and ϵ is the scale on which the time is discretized. After performing L Leap-Frog steps a proposal $\left(\boldsymbol{\theta}^*, \boldsymbol{\phi}^*\right)$ is fed into a Metropolis-Hastings accept-reject step that is then used to update the current parameter value in the Markov chain at position t. The auxiliary momentum is discarded because it is drawn anew at the next iteration from its posterior distribution $p(\boldsymbol{\phi})$. When tuned right, the HMC algorithm allows for a rapid exploration of the parameter space [Gel+14]. An extension of HMC that adaptively determines the number of Leap-Frog steps, the step size ϵ and covariance matrix M is given by the No-U-Turn-Sampler (NUTS) [HG14]. The number of Leap-Frog steps L is determined individually for each chain iteration t; as soon as the distance between the original set $\left(\boldsymbol{\theta}^{(t-1)}, \boldsymbol{\phi}^{(t-1)}\right)$ and the proposals $\left(\boldsymbol{\theta}^*, \boldsymbol{\phi}^*\right)$ decreases (i.e. the trajectory makes a U-Turn in parameter space) the Leap-Frog integrator is stopped. The step size ϵ and covariance matrix M are determined during the warm-up phase and then kept fixed for all iterations that are saved for posterior inference [Gel+14].

1.3.4 Diagnosing convergence of Markov-chains

In order to quantitatively monitor the convergence of a Markov chain, there exist several diagnostic tools. Two of them will be introduced in the following.

Potential scale reduction statistic $\widehat{R}$

One way to monitor the convergence of a Markov chain towards its equilibrium distribution is to compare its performance to several other independent Markov chains [Sta22]. Motivated by this Gelman and Rubin introduced the potential scale reduction statistic, $\widehat{R}$ [GR92]. It measures the ratio of average variance of the samples within each chain to the variance of the pooled samples across all chains [Sta22]. If all chains have reached their stationary distribution this ratio will be one and otherwise greater than one [Sta22]. In practice one should tune the number of samples so that $R \lesssim 1.01$ [Veh+19].

The $\widehat{R}$-statistic is defined for a set of M Markov chains $\boldsymbol{\theta}_m$ with N samples $\boldsymbol{\theta}_m^{(n)}$. The *between-chain-variance* B is estimated as

$$B = \frac{N}{M-1} \sum_{m=1}^{M} \left(\bar{\boldsymbol{\theta}}_m^{(\bullet)} - \bar{\boldsymbol{\theta}}_\bullet^{(\bullet)} \right)^2, \tag{1.22}$$

where

$$\bar{\boldsymbol{\theta}}_m^{(\bullet)} = \frac{1}{N} \sum_{n=1}^{N} \boldsymbol{\theta}_m^{(n)}, \qquad \bar{\boldsymbol{\theta}}_\bullet^{(\bullet)} = \frac{1}{M} \sum_{m=1}^{M} \bar{\boldsymbol{\theta}}_m^{(\bullet)}. \tag{1.23}$$

Further the *within-chain variance* W is averaged over all chains

$$W = \frac{1}{M} \sum_{m=1}^{M} s_m^2, \tag{1.24}$$

where

$$s_m^2 = \frac{1}{N-1} \sum_{n=1}^{N} \left(\boldsymbol{\theta}_m^{(n)} - \bar{\boldsymbol{\theta}}_m^{(\bullet)} \right)^2. \tag{1.25}$$

Combining the between and within chain variances into the *variance estimator*

$$\widehat{\mathrm{var}}^+(\boldsymbol{\theta}|y) = \frac{N-1}{N} W + \frac{1}{N} B \tag{1.26}$$

one finally arrives at the *potential scale reduction statistic*

$$\widehat{R} = \sqrt{\frac{\widehat{\mathrm{var}}^+(\boldsymbol{\theta}|y)}{W}}. \tag{1.27}$$

The previous definitions and equations can be found in a similar fashion in [Sta22]. Advanced optimizations to the $\widehat{R}$ statistic include the splitting of all M chains into two parts (split-$\widehat{R}$) to provide an additional means for detecting non-stationarity [Sta22] and rank normalization [Veh+19].

Monte-Carlo Standard Error (MCSE)

Given S independent simulation draws from a given Markov chain $\boldsymbol{\theta}^{(s)}$ one can estimate the accuracy of the mean obtained from the posterior distribution for each parameter θ_n as the so called *Monte-Carlo standard error* (MCSE)

$$\mathrm{MCSE}_{\theta_n} = \frac{\mathrm{std}(p(\theta_n|y))}{S} = \frac{\mathrm{std}(\theta_n^{(s)})}{S}, \tag{1.28}$$

where $\mathrm{std}(x)$ is the standard deviation of the set x. As the mean can be sensitive to few extreme outliers of the posterior distribution, the median of the posterior may provide a better statistic to summarize a parameter inference. Vetahri et al. [Veh+19] propose the following method to determine a MCSE estimate for any α-quantile for a single parameter θ_n:

1. Compute a and b as the 16% and 84% quantile of

$$\mathrm{B}(S\alpha + 1, S(\alpha - 1) + 1), \tag{1.29}$$

where $\mathrm{B}(v, w) = \int_0^1 \mathrm{d}t\, t^{v-1}(1-t)^{w-1}$ is the Beta-function [AS72].

2. Determine A and B as the corresponding quantiles on the original scale of draws

$$A = \theta_n^{(s')} \text{ where } s' \leq Sa < s' + 1 \tag{1.30}$$

$$B = \theta_n^{(s'')} \text{ where } s'' - 1 < Sb \leq s'', \tag{1.31}$$

where the samples $\theta_n^{(s)}$ have been sorted in ascending order.

3. Compute the MCSE estimate of the α-quantile as

$$\mathrm{MCSE}_\alpha = \frac{B - A}{2} \tag{1.32}$$

For the remainder of this thesis the MCSE will refer to $\mathrm{MCSE}_{\alpha=0.5}$. Note that the MCSE is only then a sensible quantity when it is compared to the magnitude of e.g. the mean or the median of the posterior draws. Therefore one can define the *relative MCSE* of any posterior $p(\theta_n|y)$ as

$$\text{rel. MCSE} = \frac{\mathrm{MCSE}}{\mathrm{median}(p(\theta_n|y))}. \tag{1.33}$$

Each Markov chain element $\boldsymbol{\theta}^t$ depends on its predecessor $\boldsymbol{\theta}^{t-1}$. However the element $\boldsymbol{\theta}^{t-1}$ is derived from $\boldsymbol{\theta}^{t-2}$ and so on. This means that the elements of the constructed chain are not completely uncorrelated, which is called *autocorrelation* [Gel+14]. In all above equations S was always specified as the number of *independent* draws from a single chain which is referred to as the *effective sample size* (ESS) in the literature [Gel+14; Veh+19]. Because of autocorrelation the effective sample size does not necessarily equal the total number of posterior draws. Determining the autocorrelation allows in turn to determine the ESS from the posterior samples using fast Fourier transformations [Gey92; Gey11]. A detailed description thereof can be found in references [Sta22; Gel+14].

1.3.5 Combining inferences

If there are several independent parameter inferences, i.e. posteriors $p_1\left(\boldsymbol{\theta}|y^{(1)}\right), p_2\left(\boldsymbol{\theta}|y^{(2)}\right), \ldots,$ there are several ways to combine the information from all posteriors into a single posterior distribution. The simplest approach is to combine the draws from all posteriors, effectively just increasing the sample size, and histogram them at once. This is the mathematical equivalent of summing all posteriors with equal weight [Sto61]

$$p(\boldsymbol{\theta}|y) = \sum_{i=1}^{N} p\left(\boldsymbol{\theta}|y^{(i)}\right). \tag{1.34}$$

The resulting distribution will represent the information from all posteriors and allows to evaluate the performance of a probabilistic model that is applied to several datasets when compared with according

expectations. Suppose the combined posterior $p(\boldsymbol{\theta}|y)$ results in an unimodal Gaussian, then the mean and standard deviation will represent a combined estimate of all posteriors that can be used to summarize the results of all posteriors.

An alternative way of combining probabilistic information is the so called *independent likelihood pool* [MD94] which is a fully Bayesian approach. Consider the combined posterior

$$p\left(\boldsymbol{\theta}\middle|y^{(1)}, y^{(2)}, \ldots, y^{(N)}\right) = \frac{p\left(y^{(1)}, y^{(2)}, \ldots, y^{(N)}\middle|\boldsymbol{\theta}\right) \cdot p(\boldsymbol{\theta})}{p\left(y^{(1)}, y^{(2)}, \ldots, y^{(N)}\right)}, \tag{1.35}$$

where Bayes' theorem has been applied to write the posterior as product of likelihood and prior. Since the datasets $y^{(i)}$ are assumed independent we can write

$$p\left(\boldsymbol{\theta}\middle|y^{(1)}, y^{(2)}, \ldots, y^{(N)}\right) = \frac{\prod_{i=1}^{N} p\left(y^{(i)}|\boldsymbol{\theta}\right) \cdot p(\boldsymbol{\theta})}{p\left(y^{(1)}, y^{(2)}, \ldots, y^{(N)}\right)}, \tag{1.36}$$

and, applying Bayes' theorem once more

$$p\left(\boldsymbol{\theta}\middle|y^{(1)}, y^{(2)}, \ldots, y^{(N)}\right) = \frac{\prod_{i=1}^{N} p\left(y^{(i)}\right)}{p\left(y^{(1)}, y^{(2)}, \ldots, y^{(N)}\right)} \cdot \frac{\prod_{i=1}^{N} p\left(\boldsymbol{\theta}|y^{(i)}\right)}{p(\boldsymbol{\theta})^{N-1}} \tag{1.37}$$

$$\Rightarrow p\left(\boldsymbol{\theta}\middle|y^{(1)}, y^{(2)}, \ldots, y^{(N)}\right) \propto \frac{\prod_{i=1}^{N} p\left(\boldsymbol{\theta}|y^{(i)}\right)}{p(\boldsymbol{\theta})^{N-1}}. \tag{1.38}$$

The above calculations can be found in a similar fashion in [**probcomb**]. To summarize, the combined posterior can be written as a product of the individual posteriors normalized by the chosen prior. This method of combination is particularly useful when the datasets $y^{(i)}$ are generated data that follow a particular function with known parameters. The combined posterior in a pooled likelihood model allows to evaluate at one glance if the expectations have been met. This includes the widths of the individual posteriors because only when they are a valid approximation of estimated parameter error the combined posterior will be centered around the expected value within its own width. The simple adding of posteriors does not inherently provide this information. However, the application of this method is challenging if the number of posteriors $N \gg 1$ since the multiplication has to be carried out bin-wise; the product of posteriors is thus very sensitive to statistical fluctuations since a single empty bin will annihilate all contributions for this particular bin. The multiplication of posteriors has been implemented as a logarithmic sum.

Both presented methods are used in the context of toy Monte Carlo fits (see Chapter 4).

1.3.6 Posterior predictive checks

The quality of a Bayesian inference may be quantitatively investigated using so called *posterior predictive checks*. If the chosen model fits the data, then replicated data generated using the same model should look similar to observed data [Gel+14]. Samples $y^{\mathrm{rep},(s)}$ from the posterior predictive

distribution

$$p\left(y^{\text{rep}}|y\right) = \int \mathrm{d}\boldsymbol{\theta} p\left(y^{\text{rep}}|\boldsymbol{\theta}\right) p(\boldsymbol{\theta}|y) \tag{1.39}$$

can be drawn given posterior draws $\boldsymbol{\theta}^{(s)} \sim p(\boldsymbol{\theta}|y)$ by randomly generating from the sampling distribution with the parameter draw plugged in [Sta22]

$$y^{\text{rep},(s)} \sim p\left(y|\boldsymbol{\theta}^{(s)}\right). \tag{1.40}$$

In a regression analysis with predictors x the posterior predictive distribution (1.40) can be generated for each predictor. Introducing the test-statistic $T\left(y^{\text{rep}} > y\right)$ allows to draw conclusions about the accuracy of the prediction and thus of the chosen model. The test statistic serves as a Bayesian p-value p_B [Gel+14] and will be defined as

$$p_B = T\left(y^{\text{rep}} > y\right) = p(y^{\text{rep}} > y), \tag{1.41}$$

i.e. the fraction of replicated samples that exceed the original data points. Values near 0 or 1 will indicate a faulty measurement or model, while p-values near 0.5 are expected if the measurement and model agree well [Gel+14]. The p-value is a quantitative means to identify discrepancies between data and model and to assess whether those discrepancies may have arisen by chance [Gel+14].

1.3.7 Example: Fitting a $\cos(x)$ distribution

In order to summarize the previous discussion of Bayesian methods and also to illustrate the application thereof, the simple example of fitting a $\cos(x)$ distribution will be discussed in the following. A Frequentist approach, as well as a Bayesian approach will be applied. In the end, key differences between both approaches are investigated.

Assume there are N precise predictors $\{x_i\}$ and corresponding measurements $\{y_i\}$ with measurement errors $\{\sigma_i\}$. Additionally, the data y is expected to follow a functional $y = f(x;\boldsymbol{\theta})$ with parameters $\boldsymbol{\theta}$. Figure 1.5 shows $N = 15$ datapoints that follow the function $f(x;\boldsymbol{\theta}) := \theta_1 \cdot \cos(x) + \theta_2$ with Gaussian noise ϵ

$$y = f(x;(\theta_1 = 1, \theta_2 = 1)) + \epsilon \qquad \epsilon \sim \mathcal{N}(0,\sigma) \qquad \sigma \sim \mathcal{N}(0.25, 0.05). \tag{1.42}$$

From the noisy data, the parameters $\boldsymbol{\theta}$ shall be recovered in the following.

Frequentist Approach

The traditional approach to this fitting problem is a χ^2 fit. Then the test statistic

$$\chi^2 = \sum_{i=1}^{N} \left[\frac{y_i - f(x_i;\boldsymbol{\theta})}{\sigma_i}\right]^2 \tag{1.43}$$

can be minimized with respect to the parameters $\boldsymbol{\theta}$ giving according estimates with error bars which are calculated using error propagation from the original data points $\{y_i\}$. The minimization can be

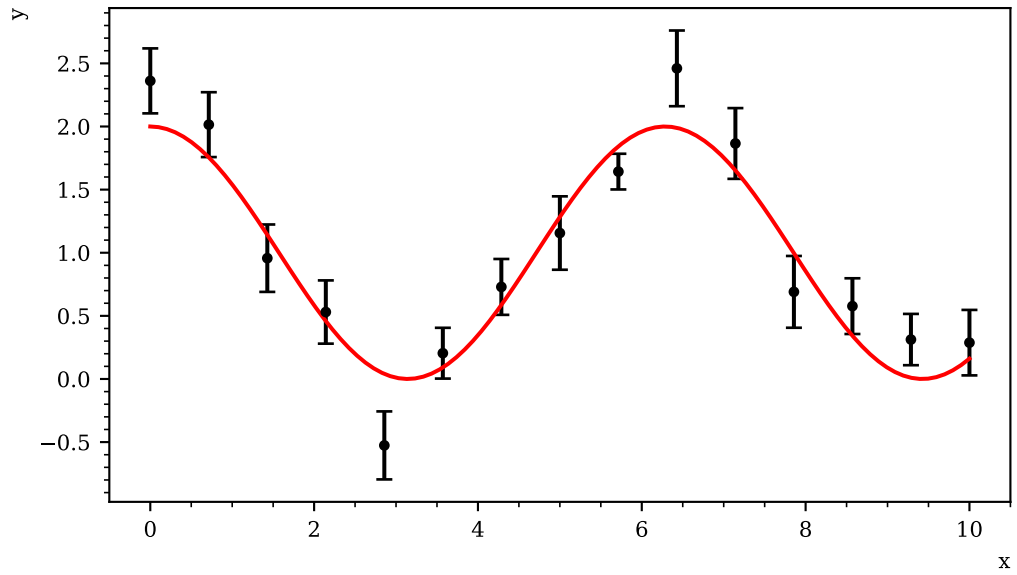

Figure 1.5: Randomly generated data with Gaussian noise (errorbars) following the function $y(x) = 1 \cdot \cos(x) + 1$ (red line).

solved analytically in the case of linear functions by solving the equation system

$$\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{\theta}}\chi^2 = 0$$

and otherwise numerically. The minimization of χ^2 to get best-fit estimates for the desired parameters can be motivated if one considers the likelihood $\mathcal{L}$ that the data follow the function f. It is given by

$$\ln \mathcal{L} = -\frac{1}{2}\sum_{i=1}^{N}\left[\frac{y_i - f(x_i;\boldsymbol{\theta})}{\sigma_i}\right]^2 - \sum_{i=1}^{N}\ln \sigma_i\sqrt{2\pi}, \tag{1.44}$$

if one assumes Gaussian errors at each data point. To maximize the (log-) likelihood with respect to all parameters is then equivalent to minimizing χ^2. As a byproduct, the χ^2 fit also gives a goodness of fit estimate which is given by $\chi^2/\mathrm{NDF} \approx 1$, where NDF are the number of degrees of freedom. Significantly smaller or larger values indicate too small error estimates or a bad fit, respectively. [Bar89]

Bayesian approach

To gather samples from the marginal posteriors $p(\theta_1|y)$ and $p(\theta_2|y)$ likelihood and priors have to be specified. In order to introduce intentionally vague priors that do not give significant information [3]

[3] This, of course, requires knowledge of the order of magnitude the fit parameters will take.

both parameters are assigned a standard normal distribution as a prior

$$\theta_1 \sim \mathcal{N}(0,1) \qquad\qquad \theta_2 \sim \mathcal{N}(0,1). \tag{1.45}$$

The likelihood is determined using

$$\epsilon_n = y_i - f(x_i;\boldsymbol{\theta}) \sim \mathcal{N}(0,\sigma_i), \tag{1.46}$$

see Eq. (1.42), and the independence of all datapoints such that the likelihood evaluates to the product

$$p(y|\boldsymbol{\theta}) = \prod_{i=1}^{N} \mathcal{N}(f(x_i;\boldsymbol{\theta})|y_i,\sigma_i). \tag{1.47}$$

With prior and likelihood specified, samples from the joint posterior $p(\boldsymbol{\theta}|y)$ can now be generated using MCMC. [Sta22]

Figure 1.6 shows the fit results from a Bayesian fit using MCMC sampling and also from a χ^2 fit for the parameters $\boldsymbol{\theta}$. The point estimates with error bars correspond to a 1σ interval of the unimodal posterior distributions. The true values underlying the simulated data are recovered by both fits.

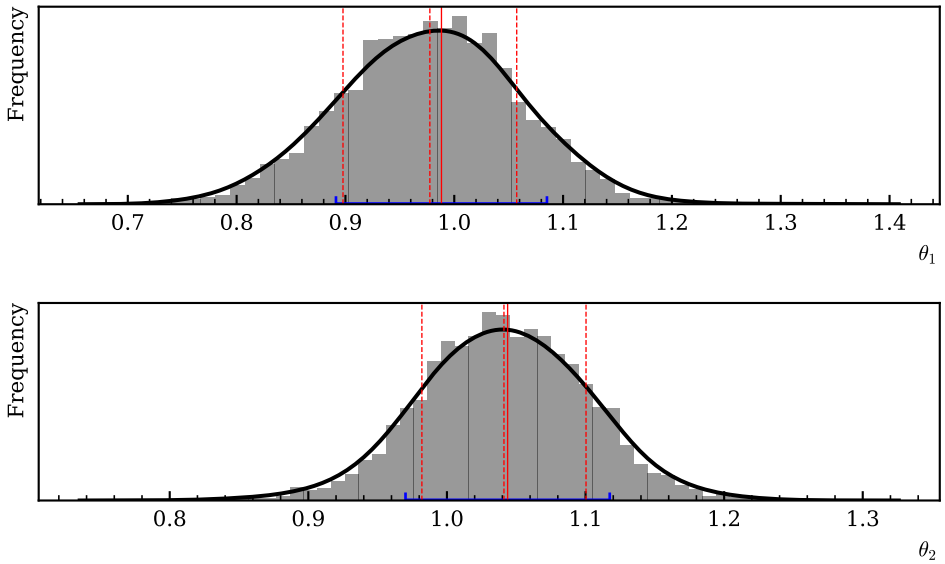

Figure 1.6: Marginal posteriors of the parameters θ_1 and θ_2 as obtained from MCMC sampling. The histograms represent the posterior draws, the black solid line shows a Kernel-Density-Estimate (KDE) to smooth the distribution. The solid vertical red line shows the most probable value determined from the KDE, while the dashed red vertical lines show a 1σ interval around the mean μ. The blue error bars show the results obtained from a χ^2 fit.

Concerning the goodness of fit, the χ^2 exhibits a value of $\chi^2/\text{NDF} = 1.3$ which is close to 1 but also significantly greater than one, indicating only moderate success of the fit due to the few number of datapoints. This issue is discussed further in chapter 4 and the appendix **??**. For the Bayesian fit, a

posterior predictive check is made, see Figure 1.7. The p values are scattered around 0.5 indicating a successful and unbiased fit. The MCMC diagnostics have to be monitored for each parameter

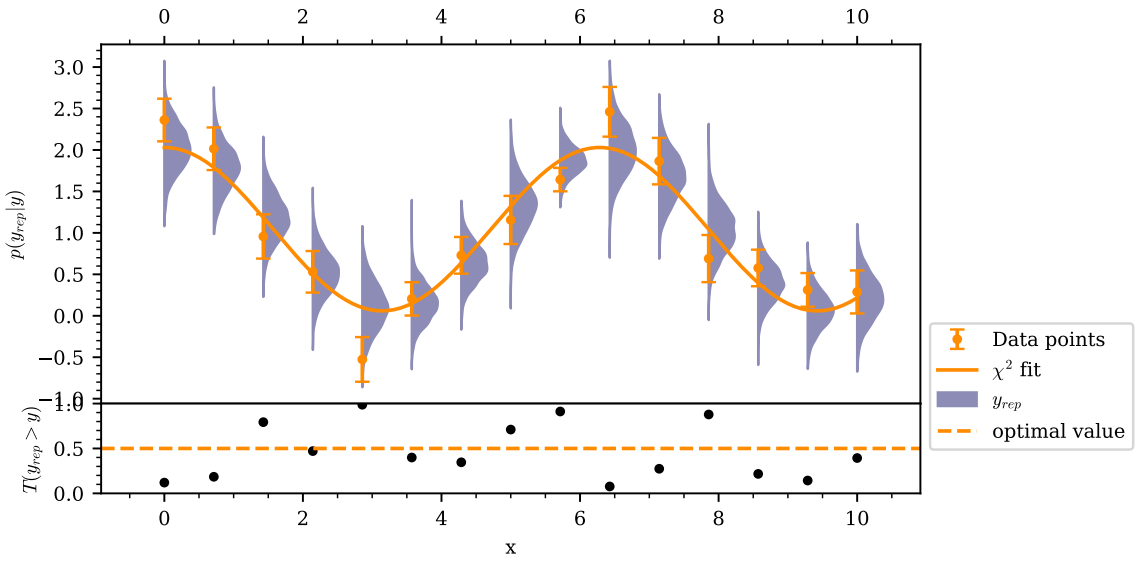

Figure 1.7: Posterior predictive check of the fitted model. For each drawn sample of $\boldsymbol{\theta}$ a random number according to the sampling distribution $p(y|\boldsymbol{\theta})$ is drawn for every predictor x_i. The p value is then determined from the posterior predictive distribution.

separately. In this particular example 1000 samples from the posterior have been drawn in four independent chains. This leads to

$$\widehat{R}_1 = 1.001 \qquad \widehat{R}_2 = 1.0004 \tag{1.48}$$

$$\frac{\text{MCSE}_1}{\text{median}(p(\theta_1|y))} = 0.002 \qquad \frac{\text{MCSE}_2}{\text{median}(p(\theta_2|y))} = 0.001, \tag{1.49}$$

indicating good within and between chain convergence as well as good accuracy regarding the median of the marginal distributions.

1.4 Current data situation

1.5 Motivation and Structure of this Thesis

bla

Bibliography

[Wor+22] R. L. Workman et al., *Review of Particle Physics*, PTEP **2022** (2022) 083C01 (cit. on pp. 2–4).

[LMP01] U. Löring, B. Metsch and H. Petry, *The light-baryon spectrum in a relativistic quark model with instanton-induced quark forces*, The European Physical Journal A **10** (2001) 395, ISSN: 1434-601X, URL: http://dx.doi.org/10.1007/s100500170105 (cit. on pp. 2, 3).

[KS03] B. Krusche and S. Schadmand, *Study of nonstrange baryon resonances with meson photoproduction*, Prog. Part. Nucl. Phys. **51** (2003) 399, arXiv: nucl-ex/0306023 (cit. on p. 4).

[Afz19] F. N. Afzal, *Measurement of the beam and helicity asymmetries in the reactions $\gamma p \to p\pi^0$ and $\gamma p \to p\eta$*, PhD thesis: Rheinische Friedrich-Wilhelms-Universität Bonn, 2019, URL: https://hdl.handle.net/20.500.11811/8064 (cit. on pp. 4, 5, 7).

[DT92] D. Drechsel and L. Tiator, *Threshold pion photoproduction on nucleons*, J. Phys. G **18** (1992) 449 (cit. on p. 4).

[Bar+18] M. Bartelmann et al., *Theoretische Physik 3 | Quantenmechanik*, 2018, ISBN: 978-3-662-56071-6 (cit. on p. 4).

[Che+57] G. F. Chew, M. L. Goldberger, F. E. Low and Y. Nambu, *Relativistic Dispersion Relation Approach to Photomeson Production*, Phys. Rev. **106** (6 1957) 1345, URL: https://link.aps.org/doi/10.1103/PhysRev.106.1345 (cit. on pp. 5, 6).

[PS95] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*, Reading, USA: Addison-Wesley, 1995 (cit. on p. 5).

[San+11] A. M. Sandorfi, S. Hoblit, H. Kamano and T.-S. H. Lee, *Determining pseudoscalar meson photoproduction amplitudes from complete experiments*, Journal of Physics G: Nuclear and Particle Physics **38** (2011) 053001, ISSN: 1361-6471, URL: http://dx.doi.org/10.1088/0954-3899/38/5/053001 (cit. on p. 6).

[CT97] W.-T. Chiang and F. Tabakin, *Completeness rules for spin observables in pseudoscalar meson photoproduction*, Physical Review C **55** (1997) 2054, ISSN: 1089-490X, URL: http://dx.doi.org/10.1103/PhysRevC.55.2054 (cit. on p. 6).

[Afz12] F. N. Afzal, *Analysis of Crystal Barrel Data - Measurement of the double polarization observable E in the reaction $\vec{\gamma}\vec{p} \to \eta' p$*, Master thesis: Rheinische Friedrich-Wilhelms-Universität Bonn, 2012 (cit. on p. 6).

[FTS92] C. G. Fasano, F. Tabakin and B. Saghai,
Spin observables at threshold for meson photoproduction, Phys. Rev. C **46** (1992) 2430 (cit. on p. 7).

[Wun+17] Y. Wunderlich, F. Afzal, A. Thiel and R. Beck,
Determining the dominant partial wave contributions from angular distributions of single- and double-polarization observables in pseudoscalar meson photoproduction,
The European Physical Journal A **53** (2017),
URL: https://doi.org/10.1140%2Fepja%2Fi2017-12255-0 (cit. on p. 7).

[Gel+14] A. Gelman et al., *Bayesian Data Analysis*, vol. 3, Chapman & Hall/CRC, 2014 (cit. on pp. 7–10, 12–14).

[SS05] D. Sivia and J. Skilling, *Data Analysis – A Bayesian Tutorial*, vol. 2,
Oxford University Press, 2005 (cit. on p. 8).

[Tro08] R. Trotta, *Bayes in the sky: Bayesian inference and model selection in cosmology*,
Contemporary Physics **49** (2008) 71,
URL: https://doi.org/10.1080%2F00107510802066753 (cit. on p. 8).

[Nor97] J. R. Norris, *Markov Chains*,
Cambridge Series in Statistical and Probabilistic Mathematics,
Cambridge University Press, 1997 (cit. on p. 9).

[Met+53] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller,
Equation of State Calculations by Fast Computing Machines,
The Journal of Chemical Physics **21** (1953) 1087,
eprint: https://doi.org/10.1063/1.1699114,
URL: https://doi.org/10.1063/1.1699114 (cit. on p. 9).

[Has70] W. K. Hastings,
Monte Carlo sampling methods using Markov chains and their applications,
Biometrika **57** (1970) 97, ISSN: 0006-3444,
eprint: https://academic.oup.com/biomet/article-pdf/57/1/97/23940249/57-1-97.pdf,
URL: https://doi.org/10.1093/biomet/57.1.97 (cit. on p. 9).

[Dua+87] S. Duane, A. Kennedy, B. J. Pendleton and D. Roweth, *Hybrid Monte Carlo*,
Physics Letters B **195** (1987) 216, ISSN: 0370-2693, URL:
https://www.sciencedirect.com/science/article/pii/037026938791197X (cit. on p. 9).

[Sta22] Stan development team, *Stan Modeling Language Users Guide and Reference Manual*, vol. 2.29, 2022, URL: https://mc-stan.org (cit. on pp. 10–12, 14, 16).

[KN98] C. Kaya and L. Noakes, "Geodesics and an optimal control algorithm", vol. 5, 1998 4918 (cit. on p. 10).

[HG14] M. D. Hoffman and A. Gelman,
The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo,
Journal of Machine Learning Research **15** (2014) 1593,
URL: http://jmlr.org/papers/v15/hoffman14a.html (cit. on p. 10).

[GR92] A. Gelman and D. B. Rubin,
Inference from Iterative Simulation Using Multiple Sequences,
Statistical Science **7** (1992) 457, URL: https://doi.org/10.1214/ss/1177011136
(cit. on p. 10).

[Veh+19] A. Vehtari, A. Gelman, D. Simpson, B. Carpenter and P.-C. Bürkner,
Rank-normalization, folding, and localization: An improved $\widehat{R}$ for assessing convergence of MCMC, arXiv e-prints, arXiv:1903.08008 (2019) arXiv:1903.08008,
arXiv: 1903.08008 [stat.CO] (cit. on pp. 10–12).

[AS72] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*,
Dover Publishing, 1972 (cit. on p. 12).

[Gey92] C. J. Geyer, *Practical Markov Chain Monte Carlo*, Statistical Science **7** (1992) 473,
URL: https://doi.org/10.1214/ss/1177011137 (cit. on p. 12).

[Gey11] C. J. Geyer, *Handbook of Markov Chain Monte Carlo*,
ed. by S. Brooks, A. Gelman, G. L. Jones and X.-L. Meng,
Chapman and Hall / CRC, 2011 (cit. on p. 12).

[Sto61] M. Stone, *The Opinion Pool*, The Annals of Mathematical Statistics **32** (1961) 1339,
URL: https://doi.org/10.1214/aoms/1177704873 (cit. on p. 12).

[MD94] J. Manyika and H. Durrant-Whyte,
Data Fusion and Sensor Management: A Decentralized Information-Theoretic Approach,
Ellis Horwood, 1994 (cit. on p. 13).

[Bar89] R. J. Barlow,
Statistics, A Guide to the Use of Statistical Methods in the Physical Sciences,
Wiley, 1989 (cit. on p. 15).

List of Figures

List of Tables