Statistical analysis of the polarization of asteroids and comets

Antti Penttilä

Academic Dissertation for the Degree of
Licentiate of Social Science

Department of Mathematics and Statistics
Faculty of Social Sciences
University of Helsinki
2009

Statistical analysis of the polarization of asteroids and comets

Author:	Antti Penttilä
Supervisor:	Professor Emeritus Kari Lumme Observatory University of Helsinki Finland

Helsingin yliopisto - Helsingfors universitet - Helsinki University ID=2009-5

Tiedekunta-Fakultet-Faculty
Faculty of Social Sciences

Laitos-Institution-Department
Department of Statistics

Tekijä-Författare-Author
Penttilä, Antti

Työn nimi-Arbetets titel-Title
Statistical analysis of the polarization of asteroids and comets

Oppiaine - Läroämne - Subject
Statistics

Työn laji-Arbetets art-Level
Licentiate thesis

Aika-Datum-Month and year
2009-01-07

Sivumäärä-Sidantal- Number of pages
68

Tiivistelmä-Referat-Abstract
The polarization of asteroids and comets, and especially the various models for polarization observations, are studied. The goal of this study is to compare existing models for polarization and their properties. Based on the results presented here a particular model and estimation technique can be found suitable for a certain types of polarization observations and research frames.

With the polarization particularly the linear polarization ratio is referred here. The Sun radiates light or more generally electromagnetic radiation - in the framework of this study to an asteroid or to a comet - in where all the linear polarization planes are equally represented, i.e. non-polarized light. Depending on its surface structure the target might scatter the different polarization planes of the radiation with different intensities making the scattered radiation polarized. The ratio between the difference and the sum of the perpendicular and parallel polarized components is called the linear polarization ratio. The linear polarization ratio can tell the researcher something about the target properties, e.g. about the surface material composition, packing density and roughness.

Some models have been developed for the linear polarization ratio, and the models can be fitted to the observations as a function of the phase angle. These models are fully empirical and are not based on actual physical modeling of the polarization event. The estimated models are, however, useful when comparing polarization properties between different targets with often quite sparse and limited observational data.

The nonlinear regression analysis and also Bayesian nonlinear regression analysis are used in the model estimation. With the Bayesian method the a priori information about the behavior of the polarization curve can be utilized, making the estimation robust also in cases where the number of observations is small or the data is insufficient in other ways.

The polarization is also studied as a function of the wavelength of light when the interest is in the effect of the wavelength on the polarization properties. Modeling these effects reliably requires a delicate selection of the proper model and the estimation technique. The possibilities of the multiple response and the hierarchical Bayesian regression analysis in modeling are studied.

Suitable models for analyzing the polarization observations are presented and compared in the thesis, and some results concerning the wavelength effect in polarization are studied.

Avainsanat-Nyckelord-Keywords
asteroids
comets
polarization
regression analysis
bayes method

Säilytyspaikka-Förvaringsställe-Where deposited

Muita tietoja-Övriga uppgifter-Additional information

Helsingin yliopisto - Helsingfors universitet - Helsinki University ID=2009-4

Tiedekunta-Fakultet-Faculty
Valtiotieteellinen tiedekunta

Laitos-Institution-Department
Matematiikan ja tilastotieteen laitos

Tekijä-Författare-Author
Penttilä, Antti

Työn nimi-Arbetets titel-Title
Asteroidien ja komeettojen polarisaation tilastollinen mallinnus

Oppiaine - Läroämne - Subject
Tilastotiede

Työn laji-Arbetets art-Level
Lisensiaatintyö

Aika-Datum-Month and year
2009-01-07

Sivumäärä-Sidantal- Number of pages
68

Tiivistelmä-Referat-Abstract
Työssä tutkitaan asteroidien ja komeettojen polarisaatiota ja eteenkin polarisaatiohavainnoille sovitettavia malleja. Työn tarkoituksena on vertailla olemassa olevia polarisaatiomalleja ja niiden ominaisuuksia. Vertailun perusteella voidaan suositella jotain tiettyä mallia sekä sovitustekniikkaa erityyppisille polarisaatiohavaintoaineistoille ja tutkimusasetelmille.

Polarisaatiolla tarkoitetaan tässä nimenomaan lineaarista polarisaatiosuhdetta. Aurinko lähettää kohteeseen, tässä työssä joko asteroidiin tai komeettaan, valoa tai yleisemmin sähkömagneettista säteilyä jossa kaikki lineaariset polarisaatiotasot ovat tasaisesti edustettuna, ns. polarisoitumatonta valoa. Kohteen pinta voi kuitenkin sirottaa eri polarisaatiotasoissa olevaa säteilyä eri voimakkuuksilla, jolloin sironnut säteily on polarisoitunutta. Sironneen säteilyn vaaka- ja pystypolarisoitujen komponenttien erotuksen ja summan suhdetta kutsutaan lineaariseksi polarisaatiosuhteeksi. Polarisaatiosuhde voi paljastaa tutkijalle jotain kohteen pinnan ominaisuuksista, kuten pinnan materiaalikoostumuksesta, pakkaustiheydestä ja karheudesta.

Lineaariselle polarisaatiosuhteelle on kehitetty joitain malleja joita voidaan sovittaa vaihekulman funktiona havaitulle polarisaatiolle. Nämä mallit ovat perustaltaan kokeellisia eivätkä perustu varsinaisen fysiikan mallintamiseen polarisaatiotapahtumassa. Mallien sovitteet ovat kuitenkin hyödyllisiä kun halutaan verrata eri kohteiden polarisaatio-ominaisuuksia usein varsin harvan ja rajoittuneen havaintoaineiston pohjalta.

Mallien sovituksessa käytetään epälineaarista regressioanalyysia sekä bayesilaista epälineaarista regressiota. Bayesin menetelmässä voidaan hyödyntää etukäteistietoa polarisaatiokäyrän käyttäytymisestä jotta sovituksesta saataisiin vakaa myös tapauksissa joissa havaintojen lukumäärä on pieni tai aineisto on muuten puutteellinen.

Polarisaatiota tarkastellaan työssä myös valon aallonpituuden funktiona jolloin kiinnostuksen kohteena on aallonpituuden vaikutus polarisaation ominaisuuksiin. Näiden vaikutusten mallintaminen luotettavalla tavalla vaatii mallin ja sovitusmenetelmän huolellista valintaa. Tässä yhteydessä käsitellään mahdollisuutta käyttää usean vasteen regressioanalyysia sekä hierarkkista bayesilaista regressioanalyysia mallin sovituksessa.

Työssä esitellään ja vertaillaan malleja ja sovitusmenetelmiä polarisaatiohavaintojen tutkimiseen sekä johdetaan joitain tuloksia aallonpituuden vaikutuksesta polarisaatioon.

Avainsanat-Nyckelord-Keywords
pikkuplaneetat
komeetat
polarisaatio
regressioanalyysi
bayesilainen menetelmä

Säilytyspaikka-Förvaringsställe-Where deposited

Muita tietoja-Övriga uppgifter-Additional information

Preface

This licentiate thesis is partly based on the published article by A. Penttilä, K. Lumme, E. Hadamcik and A-C. Levasseur-Regourd, titled 'Statistical analysis of asteroidal and cometary polarization phase curves'. The article is published in Astronomy & Astrophysics, volume 432 (2005). As a first author of the article I was responsible of the data management, statistical analysis and conclusions made in the article.

Commonly used abbreviations in the thesis

α_0	inversion angle
λ	wavelength
GS	Gibbs sampler
IR	infrared wavelength domain
LEM	linear-exponential model
LR	likelihood ratio
LS	least squares
ML	maximum likelihood
MR	multiple response
P(α)	polarization (linear polarization ratio) as a function of α
PPC	polarization phase curve
SBN	Small Bodies Node database
sub-TRIM	restricted version of the trigonometric model
TRIM	trigonometric model
UV	ultraviolet wavelength domain
WE	wavelength effect

I will use the following font styles to indicate the different types of variables and functions: scalar variables with italic, e.g. b, and scalar-valued functions with roman slant, e.g. P. Vectors and vector-valued functions with bold, e.g. $\boldsymbol{x}$ and $\mathbf{f}$. Matrices and matrix-valued functions with capital roman bold, e.g. $\mathbf{I}$ and $\mathbf{F}$. Random variables accordingly, but with sans-serif font, e.g. y, $\mathbf{\mathsf{y}}$ and $\mathbf{\mathsf{Y}}$. Estimates are printed with an accent, $\hat{b}$, and distributions with calligraphic font, e.g. $\mathcal{N}$.

I would like to acknowledge the financial support for this work by Finnish Cultural Foundation. I thank my supervisor, prof. emeritus Kari Lumme for his support and active role in my research work, as well as the whole Planetary Systems Research group led by Karri Muinonen. Other people that have helped and influenced this work include e.g. Edith Hadamcik, Anny-Chantal Levasseur-Regourd, Imbi Traat and Pentti Saikkonen. Special thanks also to my family — Sonja, Joona and Tuomas — for patience, inspiration and support.

Contents

Chapter 1

Introduction

Photopolarimetry is one of the key methods to retrieve physical information about the Solar System objects. These include the planets, their aerosols, satellites and rings, the asteroids, the cometary comae and the interplanetary dust. Roughly speaking, they can be subdivided into objects with a surface (planets, satellites, asteroids) and clouds of particles (aerosols in planetary atmospheres, rings, coma, interplanetary dust). The Solar light scattered by such media is actually partially linearly polarized.

The linear polarization, as defined in e.g. Gehrels (1974) or Hapke (1993), is the ratio of the difference to the sum of the two polarized components of the brightness, respectively perpendicular and parallel to the scattering plane. One of the most interesting points is indeed that, since the polarization is a ratio (remaining between -1 and +1), no normalization in brightness is required to compare data obtained on different objects (or on the same object at different times). This is especially important for comets, since the brightness variations with the distances to the Sun and to the Earth depend not only upon the distances but also upon changes in the activity of the comet and upon the size of the coma.

Both types of objects have been widely modeled using the ever-increasing amount of observational data. It is fairly obvious that the models for bodies with an atmosphere have been successful and have produced unique knowledge of the physical structure. The basic reason for this is that light-scattering models are much easier to apply to atmospheres because the basic constituents are much smaller than the wavelengths used. Then the Rayleigh-type light scattering approach is quite adequate.

Seeliger's work (1887) on Saturn's rings was obviously the first attempt to gain quantitative information of the ring particles. He introduced the basic idea of the mutual shadowing concept which since then has been widely used for the regolith studies to explain the classical opposition effect, the nonlinear brightening seen near the backscattering direction. The fundamental difficulty in the shadowing

scheme is the question how to separate the shadowing contribution and that originating from a single particle because both depend on the same scattering angle or phase angle. The implicit and completely unjustified assumption normally made was that small ring or regolith particles do not have their own strong opposition brightening. It was not until much after Seelinger's work when the dynamicists realized that the packing density of Saturn's rings is about ten times as high as predicted by the shadowing mechanism (Dones et al. 1993, Mishchenko 1993).

At about the same time computer capabilities had increased so much that exact wave-optical calculations became possible for small non-spherical particles and their aggregates. This allowed a better interpretation of polarimetric data. Indeed, the question of negative polarization, sometimes wrongly called anomalous polarization, has recently been a widely studied subject in planetary research (see e.g. Muinonen, Piironen, Shkuratov, Ovcharenko & Clark 2002, Muinonen 2004, Shkuratov et al. 2004). Here the fundamental question is why the inversion angle is so insensitive to parameters like albedo or wavelength.

For very small particles in an atmosphere or for liquid (and thus spherical) aerosols, the Mie scattering theory can be used. However, such an approach is not always adequate for solid particles (see Hadamcik et al. 2003 for soot particles). Lumme et al. (2003) started a systematical study of light scattering by various aggregates of wavelength-sized constituents. These studies unequivocally show that both nonlinear backward scattering and negative polarization naturally follow without any restrictive assumptions. Unfortunately computer capabilities still limit the aggregate sizes to a few hundred constituents. Calling these aggregates a single particle allows radiative transfer calculations for large aggregates and regoliths.

An increasing amount of observational data has been compared to numerical data for modeled particles to obtain clues about the dust physical properties. These comparisons need a continuous curve to fit the data and if possible to predict the values for phase angles where data are missing. Therefore the topic of this thesis, the different models to the polarization phase curve and methods to estimate the model parameters and their errors, offers useful tools for the photopolarimetry.

1.1 Polarimetric observations on asteroids and comets

The light is electromagnetic wave-motion and the scattering of light refers to the change in the wave-motion when the incident wave confronts an interface between two media. The details of electromagnetic scattering is a vast topic and is not covered here. Instead, one can refer to e.g. van de Hulst (1957) or to more recent Mishchenko et al. (2006). It is probably sufficient just to mention that the state

of the (average) electric field vector can be expressed with the intensity and the polarization. The scattering event can change both the intensity and the polarization of the light. The polarization state can be described with the polarization ellipse in general, with the special cases of linear or circular polarization which are schematically shown in Fig. 1.1. The polarimetric observations of the light scattered from asteroids or comets measure usually the linear polarization ratio P, the ratio between the difference and the sum of the flux Z in perpendicular ($\perp$) and parallel ($\parallel$) polarized components:

$$\mathrm{P}(\alpha) = \frac{Z_\perp(\alpha) - Z_\parallel(\alpha)}{Z_\perp(\alpha) + Z_\parallel(\alpha)}, \tag{1.1}$$

usually given in percent. The phase angle α is the angle between the light source (the Sun), the target and the observer (the Earth, usually) as shown in Fig. 1.2. These three objects also define the so-called scattering plane, in which the directions parallel and perpendicular are defined. It is common that the flux is mostly perpendicularly polarized making the linear polarization ratio positive, but also negative P is regularly observed with small phase angles from e.g. asteroids and comets.

Numerous polarimetric measurements for asteroids and comets can be found in published papers. Some compilations were attempted for asteroids (available

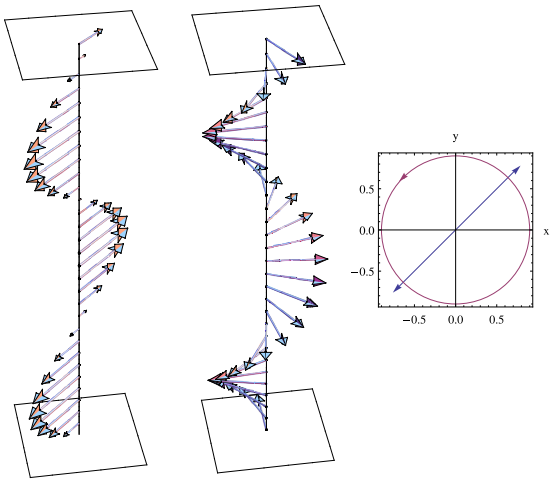

Figure 1.1: A schematic description of the electric field vector traveling upward with linear polarization state in the left and with circular polarization state in the middle. The shapes that the field vectors trace as traveling though a plane are shown in the right (blue is linear and red is circular polarization). The combination of these two states would lead to the general elliptic polarization state.

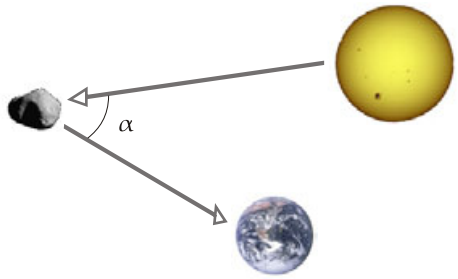

Figure 1.2: A schematic representation of the observation geometry and the definition of the phase angle α between the Sun, the target (asteroid in this case), and the Earth. The plane through these three objects is the scattering plane.

on the Internet at the SBN, the Small Bodies Node*). Except for some near-Earth objects, unfortunately, the explored phase angles are smaller than 30°. For cometary data, a general electronic database does not yet exist and it is thus necessary to use the published papers on polarimetric observations of comets.

For all these objects (asteroidal surfaces and cometary comae), phase curves at least within a given wavelength range present similar smooth shapes with a small negative branch, an inversion region where P turns from negative to positive, and a wide positive branch with a maximum near 90°. Data retrieved near opposition (phase angle equal to 0°) correspond to extremely small polarizations. Although no data have ever been obtained at 180°, with a current maximal phase angle for comets equal to 121° (Hadamcik & Levasseur-Regourd 2003*b*), polarization seems to significantly decrease near forward scattering. Maybe the best example of the polarization from atmosphereless Solar System object is available from the Moon, e.g. from Lyot (1929). The Fig. 1.3 shows the very same overall shape of the polarization phase curve as for the asteroids or comets with the maximum of α at 159.4°.

1.1.1 Asteroidal database

The asteroidal polarimetric data used in this thesis come mainly from the SBN database†. They include 1635 entries for 137 asteroids (number, date of observation, filter, phase angle, polarization, measurement error, position angle etc.). More than 85 percent of the measurements are in the U filter 362 nm, B filter (435 nm) and V filter (559 nm) and only 7 percent in the R filter (685 nm). The phase angle range is from 0.1°to 120°, but in most cases the phase angles are

*http://pdssbn.astro.umd.edu
†http://www.psi.edu/pds/archive/radarpol.html

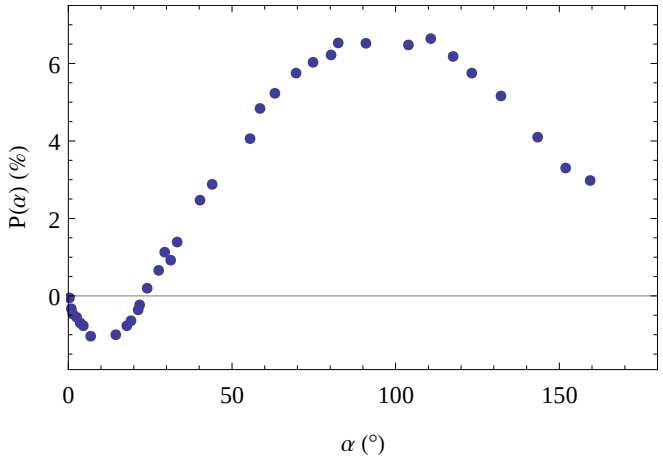

Figure 1.3: Polarization of the Moon, Lyot (1929).

limited to less than 30°. Only some near-Earth objects were explored at large phase angles. A classification based on spectra obtained in the visual and near infrared domains, and data on the physical parameters of the asteroids are also available at the same address. Further data, not included in the SBN at the time, have been published by e.g. Goidet-Devel et al. (1995), Mukai et al. (1997) and Kiselev et al. (1999, 2002).

1.1.2 Cometary database

Polarimetric observations of cometary dust are not easy. First, it is necessary to use filters to avoid the gaseous emissions. If the filters are not correctly chosen or the measured intensities are not corrected for the gaseous emissions by a study of their spectrum close to the period of observations in polarization, the measurements are not adequate for a comparison with other comets. The second difficulty is related to the variation of the polarization with the aperture due to differences of dust properties in the coma (Jockers 1997, Hadamcik & Levasseur-Regourd 2003*a*,*b*). To build up a database for cometary dust it is thus necessary to use large enough apertures (depending on the comet). The variation in the observations at a fixed phase angle is mainly due to aperture differences. With the imaging polarimetry technique it is possible to observe the coma regions and to better choose an aperture that includes the main structure (Hadamcik & Levasseur-Regourd 2003*a*,*b*). However, polarimetric observations do not require normalization and give fundamental information on the physical properties.

Levasseur-Regourd et al. (1996) used all the data already published by various groups (see references therein) to derive a first classification of comets from their polarimetric properties. The data from new observations were later added to this database (Hadamcik & Levasseur-Regourd 2003*a*,*b*, and references therein).

Finally, S. Kikuchi (personal communication with E. Hadamcik) added data, mainly for the comet 109P/Swift-Tuttle. The database includes 36 comets with more than 1000 data points between the near ultraviolet and near-infrared spectral domains, but mainly in the green and red domains. Two comets have been extensively observed, comet 1P/Halley at its 1985-1986 apparition (for which the Vegas and Giotto space probes provided some 'ground truth') and the bright comet C/1995 O1 Hale-Bopp in 1995-1997. The generally used continuum filters are from the International Halley Watch program (IHW) or later from NASA (see wavelengths and eventual contamination in Jockers 1997). The phase angle range is from 0.6°to 121°.

The cometary data that is analyzed in this article consist of observations of comets C/1995 O1 Hale-Bopp, 1P/Halley, C/1996 B2 Hyakutake, 109P/Swift-Tuttle and C/1975 V1 West. Comets will be referred hereafter in the text without the catalogue identifier number.

1.1.3 The wavelengths and error estimates in the observations

The properties of the wavelength filters that are used in the observations can vary between the measurement campaigns and the telescopes. For example, the width of the wavelength acceptance region can differ. In this thesis I want to compare different models for polarization for both asteroids and comets from all available data sources, and therefore the different filters need to be regrouped into more coarse but common scale. I will use the visible wavelength domains violet (380–450 nm), blue (450–490 nm), green (490–560 nm), yellow and orange (560–630 nm) and red (630–760 nm), and outside visible the ultraviolet (UV, below 380 nm) and infrared (IR, 760–2200 nm).

Estimates for the measurement errors in polarization observations are given in the asteroid and comet databases. These errors are generally instrumental errors and the rotation of asteroids whose surface may be inhomogeneous (variegation effects) or the spatial variations in a cometary coma are not taken into account. For these reasons the errors are not realistic as deviations between the observations and a unique and smooth polarization phase curve of the object. For example the asteroid Vesta has two observations $P(5.05°) = -0.81 \pm 0.01$ and $P(5.4°) = -0.45 \pm 0.01$ (in %-units). The phase angles α are so close that the observations should lie within, say two to four standard errors of each other while the difference is actually 36 standard errors. However, if the error estimates are at least in the same scale with each other, these can still be used as weights in the regression analysis. One has only to bear in mind that any quantity that is computed from the weighted residuals, e.g. the residual sum, might have an arbitrary scale.

1.1.4 The structure of this thesis

The purpose of this thesis is to study different models suitable to asteroid and comet polarization data, and their estimation techniques. First, in Sec. 2 three empirical models for polarization phase curve are compared using data from different asteroids. The models are estimated using nonlinear regression. Second, in Sec. 3 the issues encountered with the models and the regular frequentists regression are tried to tackle with a Bayesian approach, and this approach is also tested on some data from asteroid Juno and comet Halley, and on the so-called high $P_{\max}$ comets. Third, in Sec. 4 the polarization is modeled as a function of the wavelength in addition to the phase angle. The possible wavelength effect on polarization is studied with either multiple response regression model, or with a Bayesian model.

Chapter 2

Selection of the polarization phase curve model

2.1 Empirical models for the polarization

There has been two efforts at the Helsinki University Observatory to produce a suitable model function to describe the polarization phase curve (PPC). Both of these models have a fully empirical basis and there is no attempt to actually model the physics involved in the scattering event that produces the polarization state. A stable, robust and realistic empirical fit accompanied with reliable error estimates alone would already be useful in the studies about the polarization behavior of Solar System objects.

2.1.1 Trigonometric model

The first suggestion for PPC model was made by Lumme & Muinonen in a poster at ACM (Asteroid, Comets, Meteors) meeting in 1993, the so-called trigonometric fit. The idea behind the model was that, instead of the local polynomial fits, the model should be able to describe the whole phase curve from $\alpha = 0°$ to $\alpha = 180°$. This model has since been used by numerous authors to fit the observational data (e.g. Goidet-Devel et al. 1995, Levasseur-Regourd et al. 1996, Kiselev et al. 2002, Hadamcik & Levasseur-Regourd 2003*a*,*b*), but formally published as late as 2005 in Penttilä et al. The model is of the form

$$\mathrm{P}_1(\alpha) = b\,\sin^{c_1}(\alpha)\,\cos^{c_2}(\frac{\alpha}{2})\,\sin(\alpha-\alpha_0), \quad \text{where } b \in [0,1],\ c_1, c_2 > 0,\ \text{and } \alpha_0 \in [0°, 180°]. \quad (2.1)$$

There are few appealing properties in this model, which I will refer as trigonometric model (TRIM) hereafter. Firstly, the function values are by definition

always limited between -100% and 100% and the values at $\alpha = 0°$ and $\alpha = 180°$ are zero, as it should be for the polarization ratio. Secondly, the TRIM nicely produces first a negative phase that turns into a positive phase before going again to zero at $\alpha = 180°$. This is the typical behaviour of the PPC from almost any atmosphereless Solar System object. The negative phase can also be bypassed, if needed, by letting the $\alpha_0 \to 0°$.

The TRIM in Eq. 2.1 has four parameters: b, α_0, c_1 and c_2. The parameter b is mainly connected to the amplitude of the polarization. The physically reasonable range for b is $[0, 1]$. With a choice of $\alpha_0 = 0°$, $c_1 = 3$, $c_2 = 0$ and $b = 1$ we approximately get the PPC of Rayleigh scattering* with a maximum polarization of 100% at phase angle $\alpha = 90°$. Parameter b also affects the slope of the phase curve at α_0.

The parameter α_0 is the inversion angle, the phase angle where the PPC turns from negative to positive. The phenomenon of a negative branch of polarization, which will turn positive in the neighborhood of $\alpha = 20°$, is common for Solar System dust where multiple scattering and interactions between the constituent grains inside aggregates play an important role. A physically reasonable range for α_0 is obviously $[0°, 180°]$, although in observations the inversion angle seems to stay below $\sim 30°$.

The powers c_1 and c_2 have an influence on the shape of the phase curve. The parameter c_1 mainly affects the position of the minimum and the second derivative of the curve, while c_2 has influence on the maximum and on the asymmetry of the curve, moving the angle for maximum polarization away from 90°. These two parameters should have positive values.

The collection of these four parameters offers a wide variety of different, realistic shapes for phase curves. These are outlined in Fig. 2.1.

2.1.2 Restricted version of the trigonometric model

There are some parameter identification problems in the TRIM if there are only a few observations and/or they belong to a small phase angle range. The power parameter of the cosine, c_2, is usually the most challenging to estimate, since its effect is seen mostly in the positive branch with large phase angles where there is commonly a lack of observation points. Originally Lumme & Muinonen tackled this problem by setting both the power parameters c_1 and c_2 as constants which were estimated by fitting the TRIM to all available observations, separately to asteroids and comets. As a consequence, the TRIM would reduce to a linear

*Rayleigh scattering refers to scattering law named after Lord Rayleigh applicable for single particles much smaller than the wavelength.

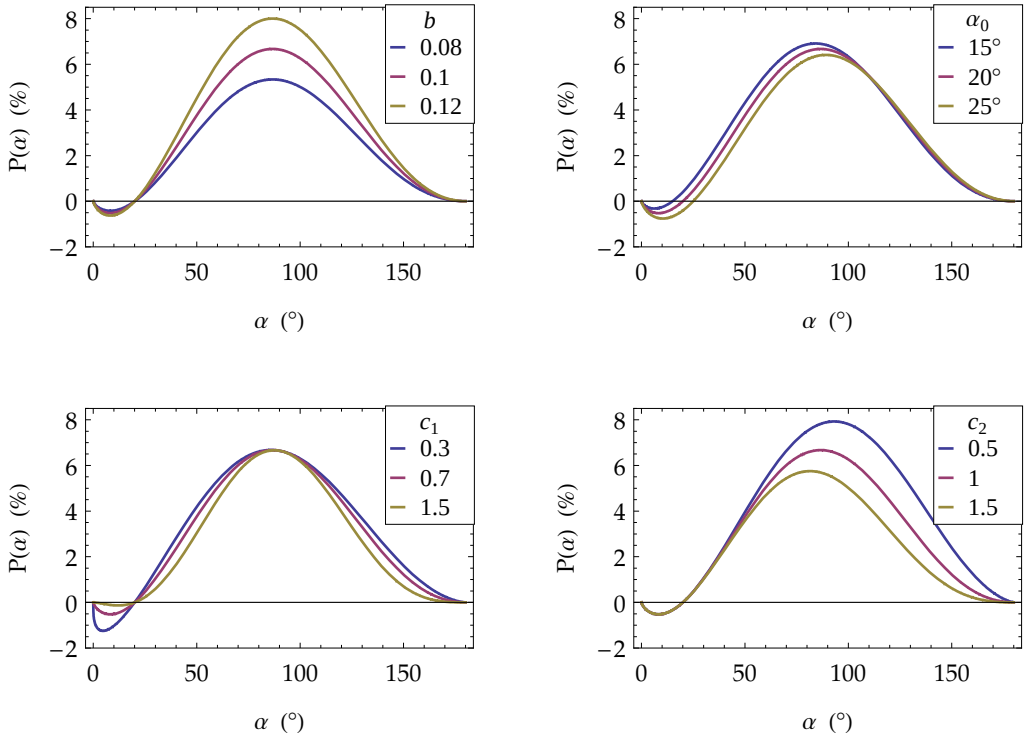

Figure 2.1: The effect of the parameters on the TRIM (given in percent). In all figures only one parameter changes, which is denoted in the corresponding legend label. The other parameters have values $b = 0.1, \alpha_0 = 20°, c_1 = 0.7, c_2 = 1$.

model:

$$\begin{aligned} P_1(\alpha; c_1 = c_1^c, c_2 = c_2^c) &= b\, \sin(\alpha)^{c_1^c}\, \cos(\frac{\alpha}{2})^{c_2^c}\, \sin(\alpha - \alpha_0) \\ &= b\, \sin(\alpha)^{c_1^c}\, \cos(\frac{\alpha}{2})^{c_2^c}\, [\sin(\alpha)\cos(\alpha_0) - \cos(\alpha)\sin(\alpha_0)] \\ &= b\, \cos(\alpha_0)[\sin(\alpha)^{c_1^c+1} \cos(\frac{\alpha}{2})^{c_2^c}] - b\, \sin(\alpha_0)[\sin(\alpha)^{c_1^c} \cos(\alpha) \cos(\frac{\alpha}{2})^{c_2^c}] \\ &= \theta_1\, x_1 - \theta_2\, x_2, \end{aligned} \quad (2.2)$$

where the θ_i are the two new parameters (e.g. $\theta_1 = b\cos(\alpha_0)$) and x_i are the new variables (e.g. $x_1 = \sin(\alpha)^{c_1^c+1}\cos(\alpha/2)^{c_2^c}$). The linearization makes the model estimation more straightforward.

To my mind the restrictions to the TRIM by setting both the c_1 and c_2 as constants are too limiting, and I propose rather a restricted version of the TRIM where only the more problematic parameter c_2 is set to a constant value. The suitable value c_2^c for the c_2 can be found by fitting the TRIM beforehand to the whole observation set in hand, e.g. to all the asteroid observations in the SBN

database. This submodel of the TRIM is now of the form

$$\mathrm{P}_2(\alpha) := \mathrm{P}_1(\alpha; c_2 = c_2^c) \tag{2.3}$$

with c_2^c as a beforehand estimated constant, and will be referred as the sub-trigonometric model (sub-TRIM) hereafter.

2.1.3 The linear-exponential model

The second model to PPC developed at the Helsinki University Observatory is the so-called linear-exponential model (LEM), which was introduced for polarimetry by Muinonen, Piironen, Kaasalainen & Cellino (2002). The main idea behind the LEM is that the same function can be used to model also the photometric observations (i.e. intensity of light) (see e.g. Piironen et al. 2000, Kaasalainen et al. 2001, 2003). With the intensity the so-called backscattering effect has been under interest in the recent years among the light scattering community (see e.g. Muinonen 2004, Mishchenko 2008, Tyynelä et al. 2008). The backscattering effect is observed from various scattering targets with small phase angles, $\alpha \lesssim 5°$. This effect was previously explained by mutual shadowing mechanisms of the target, but it has not been able to explain fully the strong nonlinear growth of the intensity. In addition to mutual shadowing, the so-called coherent backscattering or weak localization of waves has been offered to explain the observations and laboratory measurements.

The LEM for photometry is of the form $\mathrm{I}(\alpha) = a\exp(-\alpha/d) + b + k\alpha$, where $a, b, d > 0$ and $k < 0$. The exponential part is strong at small phase angles but approaches to zero with larger α, where the linear part dominates. For the polarimetry one essential demand for a model is that $\mathrm{P}(0) = 0$. For the LEM this is fulfilled if $b \equiv -a$. Thus, the LEM for polarimetry is of the form

$$\mathrm{P}_3(\alpha) = a(\exp(-\frac{\alpha}{d}) - 1) + k\alpha, \text{ where } a, d, k > 0. \tag{2.4}$$

The exponential part can model the often seen asymmetry in the negative polarization phase with the steep decrease in PPC after $\alpha = 0°$ and the linear part can model the steady increase from negative to positive around the inversion angle α_0. Obviously the LEM is not suitable for large α, since the linear part keeps the model increasing when α grows toward 180°. The model is still useful in the phase angle range where most of the observations are. The effects of the parameters on the PPC in the LEM is showed in Fig. 2.2.

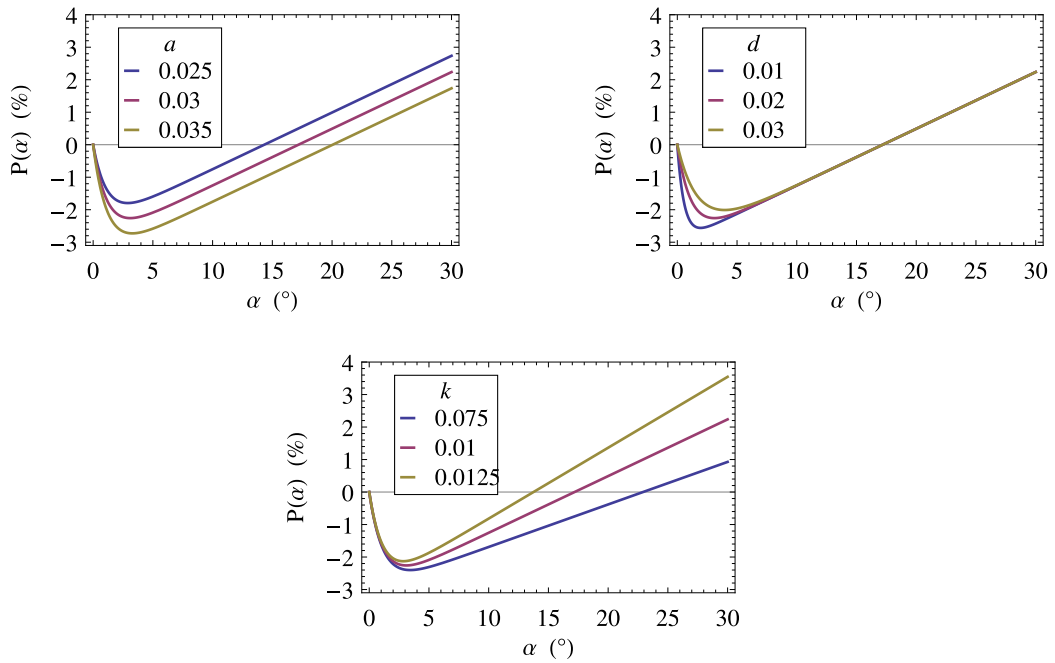

Figure 2.2: The effect of the parameters on the LEM (given in percent). In all figures only one parameter changes, which is denoted in the corresponding legend label. The other parameters have values $a = 0.03, d = 0.02, k = 0.01$.

2.2 Comparison between the empirical models for the PPC

In the following section I will compare the three aforementioned nonlinear models (TRIM, sub-TRIM and LEM) in terms of the statistical goodness and reliability of the fits. I have chosen to use a subset of the asteroid polarization data at the SBN for testing. Since the LEM is designed to model only the first negative phase and the turning from negative to positive, I will restrict the observed phase angles to corresponding range, say $\alpha \in [0°, 60°]$. Furthermore, I have required that there are enough observations per an asteroid to reliably fit the models. I have set the limit here to at least 18 observations. It is quite common that observed targets have less observations than this, and that is just the case where reliable model estimation is a challenge. Thus, for a useful model the estimation should succeed without major problems with this test dataset. The suitable objects in the asteroid dataset were, at the time of this analysis, the asteroids 324 Bamberga (CP), 1 Ceres (G), 511 Davida (C), 433 Eros (S), 8 Flora (S), 19 Fortuna (G), 40 Harmonia (S), 704 Interamnia (F), 7 Iris (S), 3 Juno (S), 39 Laetitia (S), 141 Lumen (CPF), 9 Metis (S), 2 Pallas (B), 4179 Toutatis (S) and 4 Vesta (V). The

number preceding the asteroid name is the catalogue identification number, and the letter(s) following the name indicate the asteroid class (after classification by Tholen & Barucci 1989). These asteroids are hereafter referred with the name only, without the identification number.

2.2.1 Nonlinear regression model and its estimation

All the three PPC models are nonlinear models, i.e. the outcome of the model function is not a linear combination of the parameters and functions of the explanatory variable as in linear models, but they are connected together by a nonlinear function(s). The nonlinear regression model is of the form

$$\mathbf{y} = \mathbf{f}(\boldsymbol{x}; \boldsymbol{\theta}) + \boldsymbol{\epsilon}, \tag{2.5}$$

where $\mathbf{y}$ is the dependent variable vector being modeled, $\mathbf{f}$ the vector-valued model function with argument $\boldsymbol{x}$, the vector of explanatory variables, and the parameters $\boldsymbol{\theta}$ whose values are estimated[†]. The $\boldsymbol{\epsilon}$ is the vector of random errors between the model $\mathbf{f}$ and the observed $\mathbf{y}$. By definition, the expected value $\mathbf{E}(\boldsymbol{\epsilon}) = \mathbf{0}$, and thus the expected value of the model is

$$\mathbf{E}(\mathbf{y}) = \mathbf{E}(\mathbf{f}(\boldsymbol{x}; \boldsymbol{\theta})) = \hat{\mathbf{y}}, \tag{2.6}$$

with $\hat{\mathbf{y}}$ being called the fitted, estimated or predicted value of $\mathbf{y}$. The estimation of the best-fit model parameters $\hat{\boldsymbol{\theta}}$ is done by minimizing the sum of the squared residual errors $\hat{\boldsymbol{\epsilon}}$ between the observed $\mathbf{y}$ and the modeled $\hat{\mathbf{y}}$,

$$\hat{\boldsymbol{\theta}} := \operatorname{argmin}_{\boldsymbol{\theta}} \; \| \hat{\boldsymbol{\epsilon}} \|^2 = \operatorname{argmin}_{\boldsymbol{\theta}} \; \| \mathbf{y} - \hat{\mathbf{y}} \|^2 = \operatorname{argmin}_{\boldsymbol{\theta}} \; \| \mathbf{y} - \mathbf{f}(\boldsymbol{x}; \boldsymbol{\theta}) \|^2 \, . \tag{2.7}$$

The aforementioned Eq. (2.7) defines the least squares (LS) estimator for the nonlinear model. If we add an assumption that the random errors $\boldsymbol{\epsilon}$ are distributed following the multinormal distribution with zero expected value and diagonal variance matrix $\sigma^2\mathbf{I}$, the LS estimator will coincide with the maximum likelihood (ML) estimator of the model. We can write the log-likelihood function of the model as

$$\mathrm{l}(\boldsymbol{\theta}, \sigma^2) \; = \; -\frac{n}{2} \log \sigma^2 \; - \; \frac{1}{2\sigma^2} \; \| \hat{\boldsymbol{\epsilon}} \|^2, \tag{2.8}$$

where the n is the number of observations in the model. The ML estimator for $\boldsymbol{\theta}$, the value that maximizes the log-likelihood, is clearly the same as the LS estimator in Eq. (2.7), and for the σ^2 the ML estimate is

$$\hat{\sigma}^2 = \frac{1}{n} \; \| \hat{\boldsymbol{\epsilon}} \|^2 \, . \tag{2.9}$$

[†]If not otherwise noted, all the formulae and results in this section and the following section 2.2.2 are based on Saikkonen (2004), but can also be found in e.g. Bates & Watts (1988) or Seber & Wild (1989).

The minimization in Eq. (2.7) is often impossible to calculate analytically, and needs to be solved numerically. The Levenberg-Marquardt algorithm (see e.g. Gill & Murray 1978) is often used in practice to estimate nonlinear regression models.

Properties of the maximum likelihood estimates of the nonlinear model

The precise statistical properties of the ML estimates $\hat{\boldsymbol{\theta}}$ and $\hat{\sigma}^2$ for nonlinear regression model are often impossible to derive. However, it can be shown that under quite general assumptions of the model behavior, the statistical properties of the linear model estimates are also asymptotically valid for nonlinear models. So, for nonlinear regression model

$$\hat{\boldsymbol{\theta}} \underset{as}{\sim} \mathcal{N}(\,\boldsymbol{\theta}, \sigma^2\,(\mathbf{F}(\boldsymbol{\theta})'\,\mathbf{F}(\boldsymbol{\theta}))^{-1}\,), \tag{2.10a}$$

$$\hat{\sigma}^2 \underset{as}{\sim} \mathcal{N}(\,\sigma^2, \frac{2\sigma^4}{n}\,), \text{ and} \tag{2.10b}$$

$$\hat{\boldsymbol{\theta}} \underset{as}{\perp\!\!\!\perp} \hat{\sigma}^2, \tag{2.10c}$$

where

$$\mathbf{F}(\boldsymbol{\theta}) = \frac{\partial \mathbf{f}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}, \tag{2.10d}$$

and where $\perp\!\!\!\perp$ marks that the variables are statistically independent. The statistical significance test about the $\hat{\boldsymbol{\theta}}$ or the $\hat{\sigma}^2$ are based on these asymptotic distributions.

Weighted nonlinear regression model

In all the aforementioned results we have assumed that the errors between the model and the observed values are all the same, i.e. σ^2. The more realistic assumption in some cases is, however, that while still being mutually independent, each y_i has its own error σ_i^2. This can easily be due to different conditions during the measurement of each y_i, for instance. For the model in Eq. (2.5) this means that the covariance matrix of the $\boldsymbol{\epsilon}$ can be expressed as

$$\mathbf{Cov}(\boldsymbol{\epsilon}) = \sigma^2\,\mathrm{diag}(\frac{1}{\delta_1}, \ldots, \frac{1}{\delta_n}) \tag{2.11}$$

with 'diag' denoting diagonal matrix. This can be further generalized to a case where the covariance matrix is of the form

$$\mathbf{Cov}(\boldsymbol{\epsilon}) = \sigma^2\,\mathbf{V}. \tag{2.12}$$

The assumption that we will make with Eq. (2.12) is that the matrix $\mathbf{V}$ is known beforehand, and only the general residual variance parameter σ^2 needs to be

estimated. The square roots of the $\delta_i = \sigma^2 / \sigma_i^2$ in the diagonal form of the $\mathbf{V}$ are called the observation weights.

The estimation and the statistical properties of the weighted nonlinear model in the case of a known $\mathbf{V}$ is quite straightforward. The matrix $\mathbf{V}$ is always possible to present as a product of a non-singular matrix $\mathbf{C}$ and its transpose, $\mathbf{V} = \mathbf{C}\mathbf{C}'$. Now, if we will multiply the model equation Eq. (2.5) from the left with the matrix $\mathbf{C}^{-1}$, we will get

$$\mathbf{C}^{-1}\mathbf{y} \,=\, \mathbf{C}^{-1}\mathbf{f}(\boldsymbol{\theta}) \,+\, \mathbf{C}^{-1}\boldsymbol{\epsilon}, \tag{2.13a}$$

which can also be denoted as

$$\mathbf{y}^* \,=\, \mathbf{f}^*(\boldsymbol{\theta}) \,+\, \boldsymbol{\epsilon}^*, \tag{2.13b}$$

a new nonlinear regression model with new, transformed variables and model function. For the new model it holds that

$$\begin{aligned}\mathbf{Cov}(\boldsymbol{\epsilon}^*) \,=\, \mathbf{Cov}(\,\mathbf{C}^{-1}\,\boldsymbol{\epsilon}\,) \,=\, \mathbf{C}^{-1}\,\mathbf{Cov}(\boldsymbol{\epsilon})\,(\mathbf{C}^{-1})' \,=\, \mathbf{C}^{-1}\,(\sigma^2\mathbf{V})\,(\mathbf{C}^{-1})' \\ =\, \sigma^2\,\mathbf{C}^{-1}\,\mathbf{C}\,\mathbf{C}'\,(\mathbf{C}')^{-1} \,=\, \sigma^2\mathbf{I}\end{aligned} \tag{2.14}$$

Thus, the weighted model can be transformed into a normal non-weighted model, and therefore all the results in Eq. (2.10) apply for weighted model, too, with only a small change in Eq. (2.10a):

$$\hat{\boldsymbol{\theta}} \underset{as}{\sim} \mathcal{N}(\,\boldsymbol{\theta}, \sigma^2\,(\mathbf{F}(\boldsymbol{\theta})'\,\mathbf{V}^{-1}\,\mathbf{F}(\boldsymbol{\theta}))^{-1}\,). \tag{2.15}$$

2.2.2 Estimated models

The weighted nonlinear regression introduced in the previous section is used to fit the three models for the test dataset of 16 asteroids. The data includes an error estimate for each polarization observation, and these are used as regression weights. The best-fit models are shown graphically, together with the data, in Figs. 2.3a and 2.3b.

Goodness-of-fit graphically

Firstly, the goodness of the model fit can be studied graphically. While being a subjective method, the human eye of the researcher is still a very good judge of the overall soundness of the result. It is quite clear that all the three models fit very well the observations — in those phase angles where the observations are. Also, if the (weighted) residuals of the model are plotted against the explanatory variable α, one can see that the residuals are, approximatively, both random and homoscedastic. When residuals are 'random', i.e. no trends can be found in

the $\hat{\epsilon}$ as a function of $\boldsymbol{x}$, it implies that the model can reproduce correctly the variability in $\mathbf{y}$, and no additional functions or explanatory variables are needed. When residuals are homoscedastic the assumption of a constant residual variance σ^2 is correct for non-weighted regression, or the weight matrix $\mathbf{V}$ is properly known for weighted regression.

Some insight of the possibilities to use the models to extrapolate the PPC beyond the observations can be gathered from the Figs. 2.3a and 2.3b. It seems that the TRIM will generally predict lower values for polarization than the two other models at larger α values. In many cases (Davida, Flora, Harmonia, Interamnia, Iris, Juno, Laetitia, Lumen, Pallas and Vesta) the P_{max}, the maximum value of polarization, is located at unrealistically low values of α, below 50°. This unwanted feature with the TRIM is due to the problems in identifying the parameter c_2 that mainly controls the location of the P_{max}. This is seen in e.g. the p-values of the estimated $\hat{c}_2$ later in this section. The sub-TRIM and LEM are quite consistent in their extrapolations.

Residual error variance

Secondly, the goodness of the model fit can be assessed by comparing the (weighted) residual error variances, $\hat{\sigma}^2$'s, of the models. The $\hat{\sigma}^2$ is the ML estimate of the model σ^2, as mentioned in Sec. 2.2.1, but it is a biased estimator. Therefore the estimator $\hat{s}^2 = \frac{1}{n-p} \parallel \hat{\epsilon} \parallel^2$, where the p is the number of the model parameters, is preferred. For linear models, the estimator $\hat{s}^2$ can be shown to be non-biased.

The $\hat{s}^2$'s for all the 16 asteroids and the three models are shown in Table 2.1, together with the $\hat{S}^2$, a 'total' residual variance over all the asteroids, calculated by

$$\hat{S}^2 = \frac{1}{N - k\,p} \sum_{i=1}^{k} \parallel \hat{\epsilon}_i \parallel^2, \tag{2.16}$$

where i goes through the asteroids in the data, from 1 to k, and the N is the total number of the observations in the data. The best model in terms of total residual variance is LEM, followed by sub-TRIM and TRIM, but the differences are not large. The worst fits are received from Ceres, Vesta and Interamnia in terms of the weighted $\hat{s}^2$, and the best fits from Lumen and Metis. However, it is highly questionable if the observation error estimates that are used as weights are really comparable between different objects and/or telescopes, so the comparison between the $\hat{s}^2$ for different asteroids can be misleading.

Parameter p-values

Thirdly, the goodness of the individual parameter estimates can be studied. Following Eq. (2.15), the ML estimates of the parameter values should follow asymptotically the multinormal distribution, and an individual parameter estimate $\hat{\theta}_i$

Table 2.1: Residual error variances $\hat{s}^2$ for the 16 asteroids and the three models, together with the total error variance $\hat{S}^2$ over all the asteroids.

Name	TRIM	sub-TRIM	LEM
Bamberga	2.20	2.34	2.30
Ceres	16.39	15.61	17.10
Davida	1.39	2.04	1.67
Eros	8.28	8.43	8.45
Flora	1.58	1.97	1.63
Fortuna	4.29	3.89	4.10
Harmonia	1.22	1.15	1.15
Interamnia	11.81	8.50	8.40
Iris	3.05	4.03	3.48
Juno	2.73	4.56	2.79
Laetitia	1.23	1.22	1.40
Lumen	0.20	0.37	0.39
Metis	0.74	0.70	0.76
Pallas	3.71	3.89	3.65
Toutatis	2.83	2.81	2.81
Vesta	16.85	16.88	16.50
$\hat{S}^2$	6.77	6.71	6.67

follows the normal distribution. From the test theory we know that using the definition of the $\mathcal{T}$-distribution and the fact that $\hat{\theta}_i \underset{as}{\perp\!\!\!\perp} \hat{s}^2$ it follows that

$$\mathtt{t} = \frac{\hat{\theta}_i - \theta_i}{\hat{s}\sqrt{(\,\mathbf{F}(\hat{\theta}_i)'\,\mathbf{V}^{-1}\,\mathbf{F}(\hat{\theta}_i)\,)^{-1}}} \underset{as}{\sim} \mathcal{T}_{n-1}. \tag{2.17}$$

The hypothesis (H_0) that we want to test is that the θ_i is zero, and would not have any effect in the model. The counter-hypothesis (H_1) is that it differs from zero. Based on the Eq. (2.17), the variable $\mathtt{t}$, under the H_0 when $\theta_i = 0$, should follow the $\mathcal{T}$-distribution with $n-1$ degrees of freedom. Unusually small or large values of $\mathtt{t}$ are suspicious if the H_0 is true. The p-value is the probability of the $\mathtt{t}$ to be as large (small) or larger (smaller) as it is. Thus the small p-values, traditionally below 5%, indicate that the H_0 is not probable and should be rejected in favor of the H_1. For a good model it is quite the necessity that all the model parameters are important for the model, and thus their p-values should be small at least in the majority of the cases. The values for the models and their parameters are shown in Table 2.2.

The fact that the H_0 cannot be rejected for some model parameter can be interpreted in two ways. The first interpretation is that the corresponding parameter is not necessary for the model and should be removed. The other is that the data

Table 2.2: The p-values from the t-test for the parameter significance in the TRIM and the LEM for the 16 asteroids. For the sub-TRIM all the p-values for the parameters b, α_0 and c_1 were below 0.01 in all the cases and are not shown here.

	TRIM				LEM		
Name	b	α_0	c_1	c_2	a	d	k
Bamberga	.00	.00	.00	.09	.09	.00	.00
Ceres	.00	.00	.00	.77	.00	.00	.00
Davida	.00	.00	.00	.00	.00	.00	.00
Eros	.00	.00	.90	.28	.00	.02	.00
Flora	.00	.00	.00	.00	.00	.00	.00
Fortuna	.00	.00	.00	.95	.26	.04	.04
Harmonia	.00	.00	.00	.28	.00	.00	.00
Interamnia	.59	.00	.00	.00	.00	.00	.00
Iris	.00	.00	.00	.00	.00	.00	.00
Juno	.00	.00	.00	.00	.00	.00	.00
Laetitia	.00	.00	.00	.03	.00	.00	.00
Lumen	.00	.00	.00	.00	.56	.28	.27
Metis	.00	.00	.00	.57	.03	.00	.00
Pallas	.03	.00	.01	.27	.00	.00	.00
Toutatis	.02	.00	.11	.62	.00	.00	.00
Vesta	.00	.00	.00	.14	.00	.00	.00

at hand does not give enough information to identify the correct parameter value. The latter interpretation is probably the case with the parameter c_2 in the TRIM model. From Table 2.2 we can see that the c_2 has p-values larger than 5% in 9 cases out of 16. The c_2 could be better estimated if some data should be available from larger phase angles showing the PPC behavior near the P_{max}. However, if the model is intented to be used regularly with data lacking the larger values of the α, then the TRIM cannot be recommended. For the sub-TRIM however, the parameter identification is not a problem since all the p-values for the parameter significance test for that model are below 1% level. The uncertainty in c_2 for the TRIM reflects also to the other parameters because they are somewhat correlated. While the sub-TRIM without the c_2 has no problems, there are large p-values for the TRIM for b with Interamnia and for c_1 with Eros and Toutatis. For the LEM there are no major problems with this data. The asteroids Fortuna and Lumen show large p-values for all a, d and k. From Figs. 2.3a and 2.3b it seems that the change in LEM from the exponential part to the linear part is harder to find with these objects. The parameter a seems to be the hardest to estimate in LEM, the p-values are large in three cases.

Indicators for ill-conditioned model

The ill-conditioned model is a model where there are structural problems with the parameter identification. The parameter $\boldsymbol{\theta}$ is said to be non-identifiable if $\mathbf{f}(\boldsymbol{x};\boldsymbol{\theta}_*) = \mathbf{f}(\boldsymbol{x};\boldsymbol{\theta}_0)$ with some $\boldsymbol{\theta}_* \neq \boldsymbol{\theta}_0$ and $\forall \boldsymbol{x}$, and in ill-conditioned cases we are close to this, i.e. $\mathbf{f}(\boldsymbol{x};\boldsymbol{\theta}_*) \approx \mathbf{f}(\boldsymbol{x};\boldsymbol{\theta}_0)$. The underlying problems in the structure of the nonlinear model might be hard to see if good data is at hand, but they will aggravate the problems in the parameter estimation in the cases where we lack either the number of observations or the coverage of the model space in the data.

The bad cases of non-identifiable parameters can be seen also in the p-values of significance tests for the individual parameter θ_i, but in general the problems show in the multicollinearity of the $\boldsymbol{\theta}$. Multicollinearity means that there are strong correlations between the parameters in the asymptotic covariance matrix $\mathbf{Cov}(\boldsymbol{\theta}) = \sigma^2 \left(\mathbf{F}(\boldsymbol{\theta})' \mathbf{F}(\boldsymbol{\theta})\right)^{-1}$ (see Eq. (2.10a) or Eq. (2.15) for the weighted model).

In fact, all the problems in an ill-conditioned model can be traced to the matrix $\mathbf{G} = \mathbf{F}(\boldsymbol{\theta})' \mathbf{F}(\boldsymbol{\theta})$ or $\mathbf{F}(\boldsymbol{\theta})' \mathbf{V}^{-1} \mathbf{F}(\boldsymbol{\theta})$ in the weighted case. The Fig. 2.4 sketches the situation with two different models in a 2-dimensional case — a case (I) where $\mathbf{G} \propto \mathbf{I}$ and a case (II) where $\mathbf{G}$ is far from identity- or diagonal matrix. The confidence ellipsoids for $\boldsymbol{\theta}$, for example, are proportional to the contours of the quadratic form of the $\mathbf{G}^{-1}$, which are shown in the figure. The area of the confidence ellipsoids are the same for both the cases, but the projections into 1-dimensional confidence intervals are not the same size. The interval (a_1, a_2) for the 'nice' case I is quite small compared to the interval (b_1, b_2) for the 'bad' case II although they have the same confidence level. For the three PPC models this multicollinearity is especially bad with the LEM. With all the 16 asteroids the correlations (not shown here) between the LEM parameters a, d and k are typically around 0.99 or more! This means that the best-fit values for $\boldsymbol{\theta}$ actually lie in a small but very elongated volume in the 3-dimensional parameter space. There are large correlations for the other models, too. For the sub-TRIM there is typically a considerable correlation between the b and the c_1. With the TRIM the b is typically correlated with both the c_1 and c_2 which are also mutually correlated.

The elongated ellipsoidal shape of the contours of the quadratic form of the $\mathbf{G}^{-1}$ imply not only multicollinearity problems for the parameters, but also difficulties in the numerical estimation of the model. The iterative numerical methods like the Gauss-Newton and similar will use the linear approximation for the function $g(\theta) = \| \epsilon(\theta) \|^2$ that is being minimized in the nonlinear regression. The approximation is that $g(\theta) \approx \| \epsilon(\theta^{(k)}) - \mathbf{F}(\theta^{(k)}) (\theta - \theta^{(k)}) \|^2$ which, in turn, leads to an update scheme $\theta^{(k)} \rightarrow \theta^{(k+1)}$ involving the matrix $\mathbf{G}^{-1} = (\mathbf{F}'\mathbf{F})^{-1}$. If the $\mathbf{G}$ is close to being a singular matrix the inversion of $\mathbf{G}$ is a numerically unstable operation.

The measure of the possible instabilities in matrix inversion is the matrix condition number κ. The condition number in a sense of the l_2 norm for the $\mathbf{G}$ is calculated by $\kappa_2(\mathbf{G}) = \sigma_{max}/\,\sigma_{min}$, where the σ_{min} and σ_{max} are the smallest and largest eigenvalues of the $\mathbf{G}$. Very large values of κ indicate that the matrix is almost singular and that the model is ill-conditioned. The condition numbers for the asteroid test data for the three PPC models are shown in Fig. 2.5. The κ's for the LEM are very large for every asteroid and several orders of magnitude larger than for TRIM or sub-TRIM. The sub-TRIM is by far the most stable model in this sense.

2.2.3 Conclusions from the PPC model comparison

The conclusions that at least I will make is that the sub-TRIM is the best choice of the models to be used for the PPC, at least for data that is similar to this test set or smaller (in number or in α-coverage). Practically this applies to almost all the current PPC data from asteroids, with the Toutatis as the only exception.

The TRIM had the worst overall residual variance, and there were cases where parameter significance tests gave large p-values for the model parameters. Furthermore, the extrapolation seems to be very unstable with this model.

The LEM performed quite nicely with only some problems with the parameter significance, but the correlations between the model parameters are huge and the model is ill-conditioned so troubles start to grow if the quality of the data should be worse. Also the extrapolation with this model is dangerous since the model is, by definition, non-physical as the α grows.

The sub-TRIM is a nice compromise between the physically realistic TRIM and the fact that the area near P_{max} cannot usually be observed. If, however, a more flexible model than the sub-TRIM is needed, the TRIM could be used together with some knowledge about the typical behavior of the c_2. This can nicely be done in the framework of Bayesian regression, which is discussed in Sec. 3.

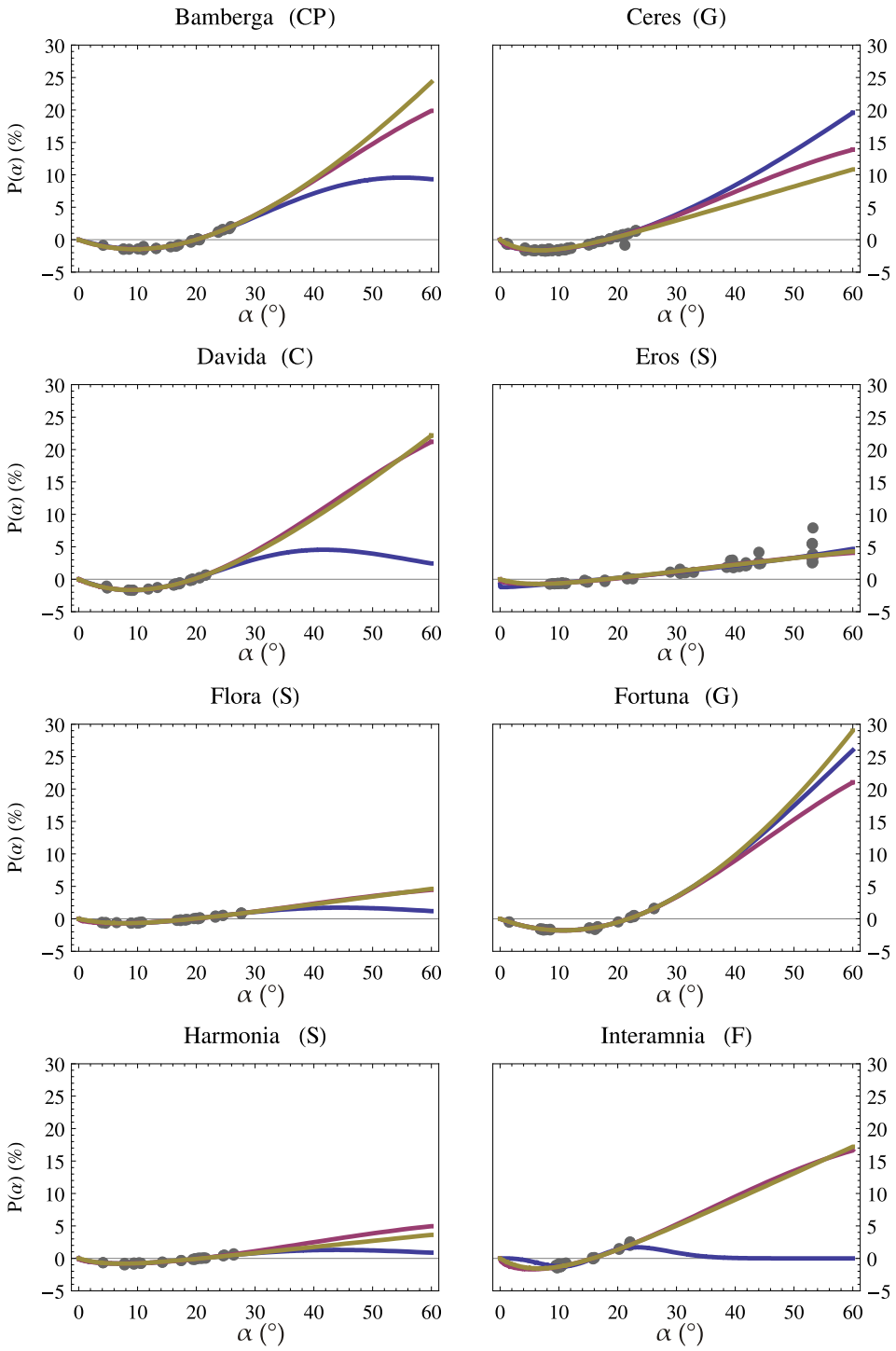

Figure 2.3a: The polarization observations and the fitted models for the first eight asteroids in the test dataset. The TRIM, sub-TRIM and LEM are shown with blue, red and yellow colors, respectively.

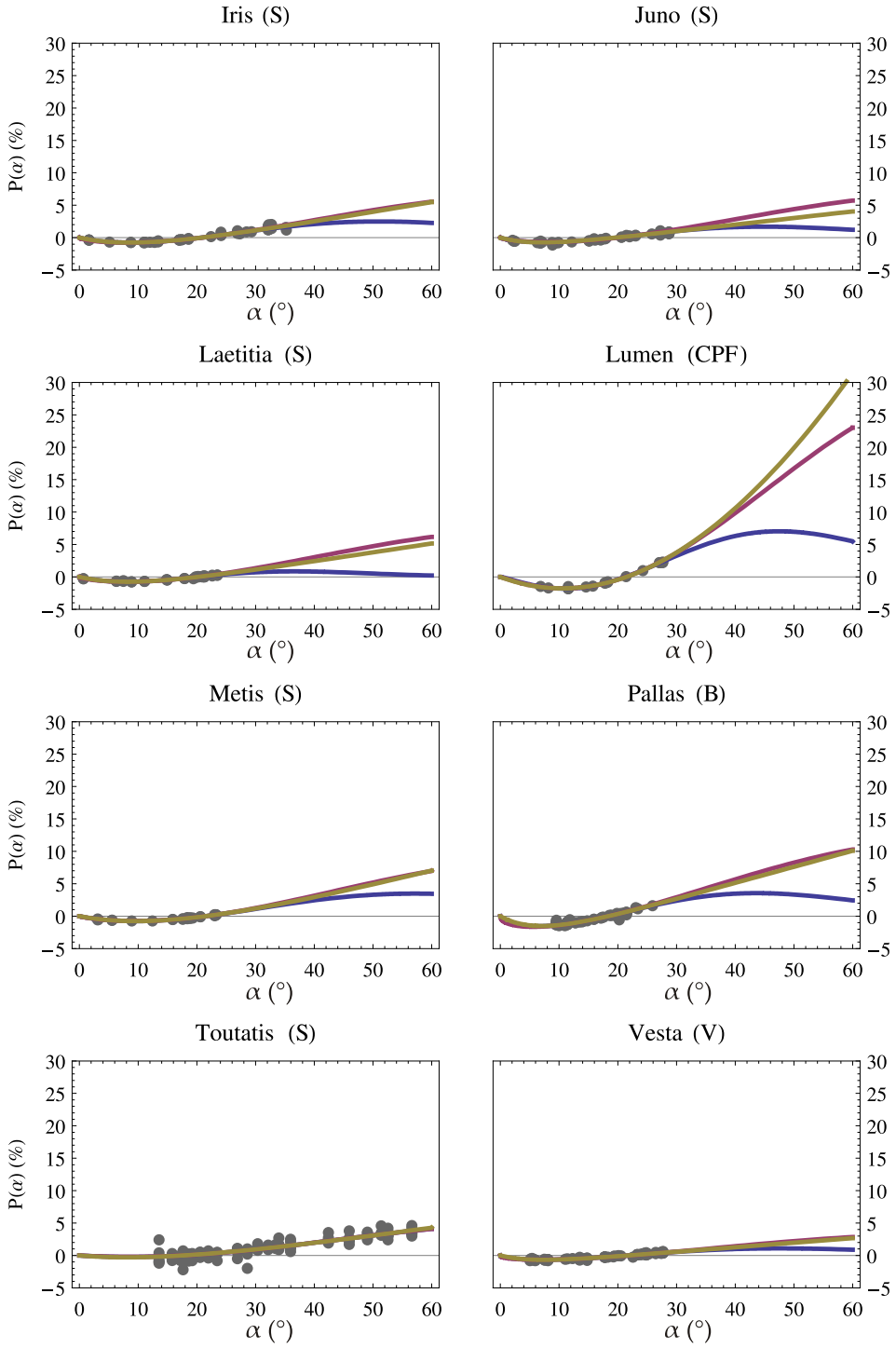

Figure 2.3b: The polarization observations and the fitted models for the last eight asteroids in the test dataset.

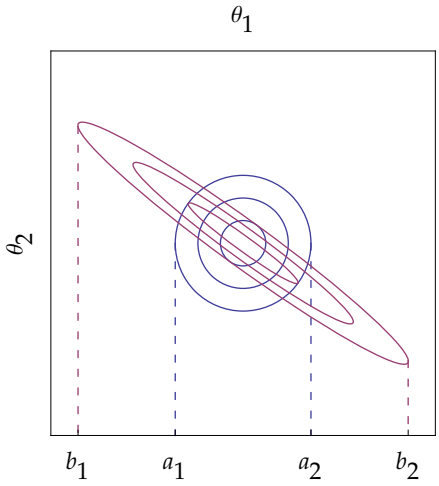

Figure 2.4: An example of the contours of the quadratic form of the $\mathbf{G}^{-1}$ for two illustrative cases I (diagonal $\mathbf{G}$) and II (non-diagonal $\mathbf{G}$).

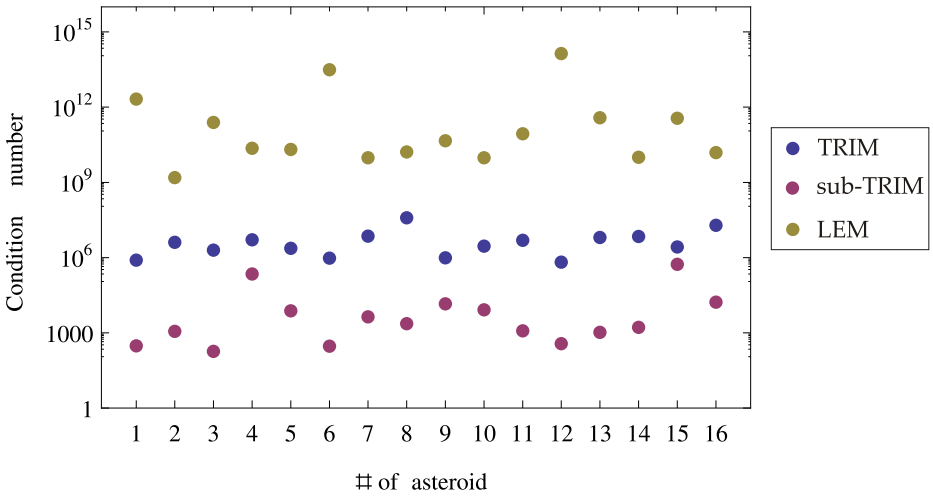

Figure 2.5: The condition number κ in logarithmic scale for the $\mathbf{G}$ for the 16 asteroids and for the three models. The number of asteroid in the x-scale refers to a rank when alphabetically ordered, as in e.g. Table 2.1.

Chapter 3

Bayesian approach in the model estimation

The common problem with empirical phase curve fitting is that, especially for asteroids, there is a lack of good datasets. For a single asteroid, there might be just a few data points from a limited phase angle range. Phase angles above $\sim 30^\circ$ are rare, and the measurement errors can be noticeable. This heavily affects the goodness and reliability of the phase curve fit, which is a nonlinear regression problem for our model. For the four parameters at least four observations is needed, but in practice only datasets starting from, say eight or ten observations, can be considered useful. Furthermore, the model parameters are not totally independent from each other, so there is multicollinearity in the model, i.e. a unique solution to the model fit is hard to find.

In many nonlinear regression procedures, the problems mentioned can result in physically unrealistic estimates, e.g. negative values for the power parameters c_1 and c_2 in the TRIM. While the fit is usually good in that small phase angle range where our observations are, any kind of extrapolation from that range is dangerous, as polarization might have values above one etc. In some applications this is not a problem, but in this study I want to examine the polarization from the whole phase angle range and consider also prediction of polarization. In some regression procedures it might be possible to introduce limits for parameters, but in practice we have noticed that this is not always reliable.

A very elegant solution to the problems of the standard nonlinear regression is the framework of Bayesian analysis and regression, where the posterior distribution of model parameters is a product of the prior distribution of the parameters and the likelihood of the parameters given by the regression model. In Bayesian regression the prior knowledge and physical limitations of the model parameters can be used in the prior distribution. Physical limitations for parameter values are handled by using distributions that have probability densities greater than zero only in the reasonable parameter range. Furthermore, since we have been working with

the asteroid and comet data for quite a while, we have acquired a good insight into the typical phase curves for different types of objects. Using this knowledge by assigning larger probabilities to the expected values of the parameters, the model fit can be guided toward a good, robust and realistic estimate. For a more detailed description of the Bayesian paradigm see e.g. Box & Tiao (1973).

3.1 Nonlinear Bayesian regression and model estimation

3.1.1 Nonlinear Bayesian regression

The Bayes formula, in short, is a simple result regarding the conditional and marginal probabilities P of the events A and B: $\mathrm{P}(A|B) = \mathrm{P}(A)\,\mathrm{P}(B|A)/\,\mathrm{P}(B)$. The same form applies also to the probability density functions p, but with p the scaling factor in the denominator can often be discarded and the proportional form $\mathrm{p}(a|b) \propto \mathrm{p}(a)\,\mathrm{L}(b|a)$ can be used. With L I denote the likelihood function, which is proportional to the density function, and can thus be used instead.

With the case of regression model the distribution $\mathcal{D}$ for $\mathbf{y}|(\boldsymbol{\theta}, \boldsymbol{\Sigma})$ is known (or at least assumed) from the specification of the model:

$$\mathbf{y}|(\boldsymbol{\theta}, \boldsymbol{\Sigma}) \sim \mathcal{N}(\mathbf{f}(\boldsymbol{x}; \boldsymbol{\theta}), \boldsymbol{\Sigma}). \tag{3.1}$$

In model estimation the interest is in knowing the $\mathcal{D}$ of the model parameters. Using the Bayes formula for probability densities it can be written as:

$$\mathrm{p}((\boldsymbol{\theta}, \boldsymbol{\Sigma}) \,|\, \mathbf{y}) \propto \mathrm{p}(\boldsymbol{\theta}, \boldsymbol{\Sigma})\,\mathrm{L}(\mathbf{y} \,|\, (\boldsymbol{\theta}, \boldsymbol{\Sigma})). \tag{3.2}$$

The Bayesian terminology is that the $\mathrm{p}((\boldsymbol{\theta}, \boldsymbol{\Sigma}) \,|\, \mathbf{y})$ is called *a posteriori* distribution $\mathcal{D}_p$ of $(\boldsymbol{\theta}, \boldsymbol{\Sigma})$ and the $\mathrm{p}(\boldsymbol{\theta}, \boldsymbol{\Sigma})$ is *a priori* distribution $\mathcal{D}_{pr}$. The philosophy of the Bayesian approach is in the use of $\mathcal{D}_{pr}$. If it would be assumed that there is no prior information about the variable under interest, a flat (i.e. non-informative) a priori could be assumed. This would, in most cases, make the inference coincide with the traditional frequentist approach. But with Bayesian inference the researcher can also use her subjective insight and decide to use a $\mathcal{D}_{pr}$ that will give information about the variable.

The Bayesian regression suits well to cases where the most flexible PPC model TRIM is preferred and where extrapolation of polarization values and estimation of some polarization features usually not well supported by the data, like the P_{max}, is needed. In Sec. 2 and especially in 2.2.2 it can be seen that the TRIM can not be well estimated as it is, but if the knowledge about the typical and probable parameter values for the PPC in the form of $\mathcal{D}_{pr}$ is used the TRIM could be guided to give more robust and sensible estimates.

3.1.2 Estimation of a nonlinear Bayesian regression model by the MCMC

The 'estimation' in Bayesian sense means that the a posteriori distribution of the model parameters in Eq. (3.2) needs to be calculated or numerically computed. The likelihood function L in the case where $\boldsymbol{\Sigma}$ is diagonal is just the L of n independent normal-distributed variables, and with a certain choice of $\mathcal{D}_{pr}$ for $(\boldsymbol{\theta}, \boldsymbol{\Sigma})$ it is possible to calculate $\mathcal{D}_p$ analytically. In a more general case, either with a more complicated model or with an unrestricted choice of $\mathcal{D}_p$, the closed analytical form is impossible to find.

In the cases where $\mathcal{D}_p$ cannot be found analytically, the Markov chain Monte Carlo (MCMC) algorithm can be used. The MCMC is a way to form a chain $\mathsf{X}_1, \mathsf{X}_2, \ldots$ where the values X_i in the chain will converge to follow the desired but unknown distribution, the $\mathcal{D}_p$ of the regression model in this case. I will not go into details of the theory of the MCMC here, instead one can turn into e.g. Gilks et al. (1995) or Robert & Casella (1999) for that. I will just mention that actually a Gibbs sampler (GS), a special case of the MCMC chain, is formed with the regression model estimation. The computations needed for the Gibbs sampling are done with the WinBUGS* software (Lunn et al. 2000).

3.1.3 The convergence of the sampler

While a modern Bayesian software packages, such as the BUGS, can form the GS for the model at hand, the researcher has to take care that the GS will converge before the sampled values are used in the analysis. The GS will, in theory, converge when i in X_i approaches infinity, but there is no guarantee that it will be converged after some large i. Therefore the convergence must be checked as the model is being sampled.

Some points about the convergence can be seen in Fig. 3.1 where two independent chains approach the same distribution. The use of two or more chains by using different starting points for the iteration is most useful. It is evident that at least the different chains should be converged to the same distribution before the chain values can be used. In the Fig. 3.1 it seems that this convergence takes place after ~ 500 iterations. The iterations before convergence form a so-called burn-in period and are not used in the final analysis.

In Fig. 3.1 the chains seem to converge to a distribution with expected value of one, and the chain variance seems constant and similar between the chains. The convergence can be followed, in addition to the raw chain values, with e.g.

*The project BUGS — Bayesian inference Using Gibbs Sampling is available on the Internet in http://www.mrc-bsu.cam.ac.uk/bugs/.

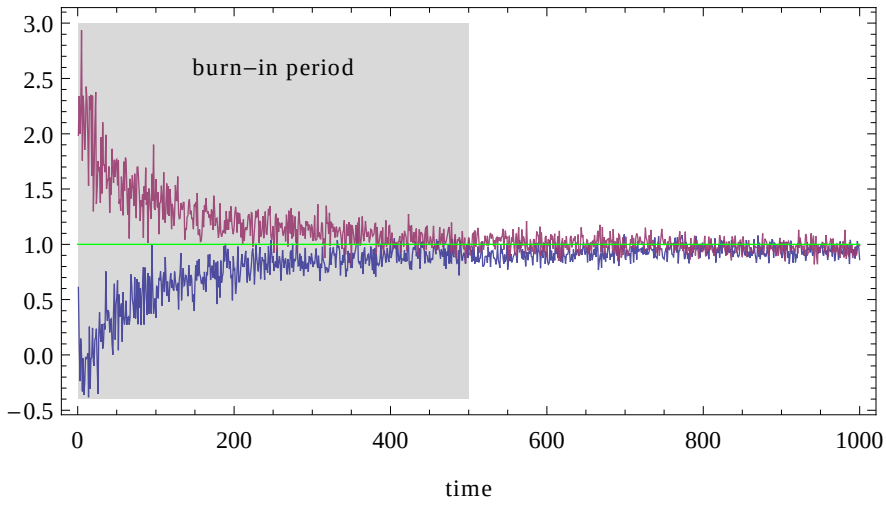

Figure 3.1: Simulated example of two MCMC chains that approach the same distribution.

running means and running confidence levels or with some dynamic estimation of the chain distribution, e.g. kernel estimation.

3.2 The specification and estimation of the TRIM in Bayesian approach

3.2.1 Model specification

The specification of the TRIM for the Bayesian regression can be presented as a graphical model, as in Fig. 3.2. First, the observed **y** together with its distribution and the functions modeling the distribution parameters are presented on a plate (large dashed square). The plate presents the data from 1 to n, and objects on the plate are dependent on the i^{th} value of the data. The distributions are marked with dashed ellipses, random variables with solid ellipses and known constants with solid squares. The objects on the plate form the same regular nonlinear model as in Sec. 2.1.1. The w_i's are the (known) weights for the corresponding observations.

The Bayesian part of the model is presented outside the plate, where the a priori distributions for the model parameters are presented. I choose to use the Beta distribution as a priori distribution for parameter b, and Weibull distributions for α_0, c_1 and c_2. In theory, the support for the a priori distribution should coincide with valid range for the model parameter. For c_1 and c_2 this is true, since they should be always positive (see Eq. (2.1)) and the support for the Weibull distribution is the positive half-space, and for b because the support for

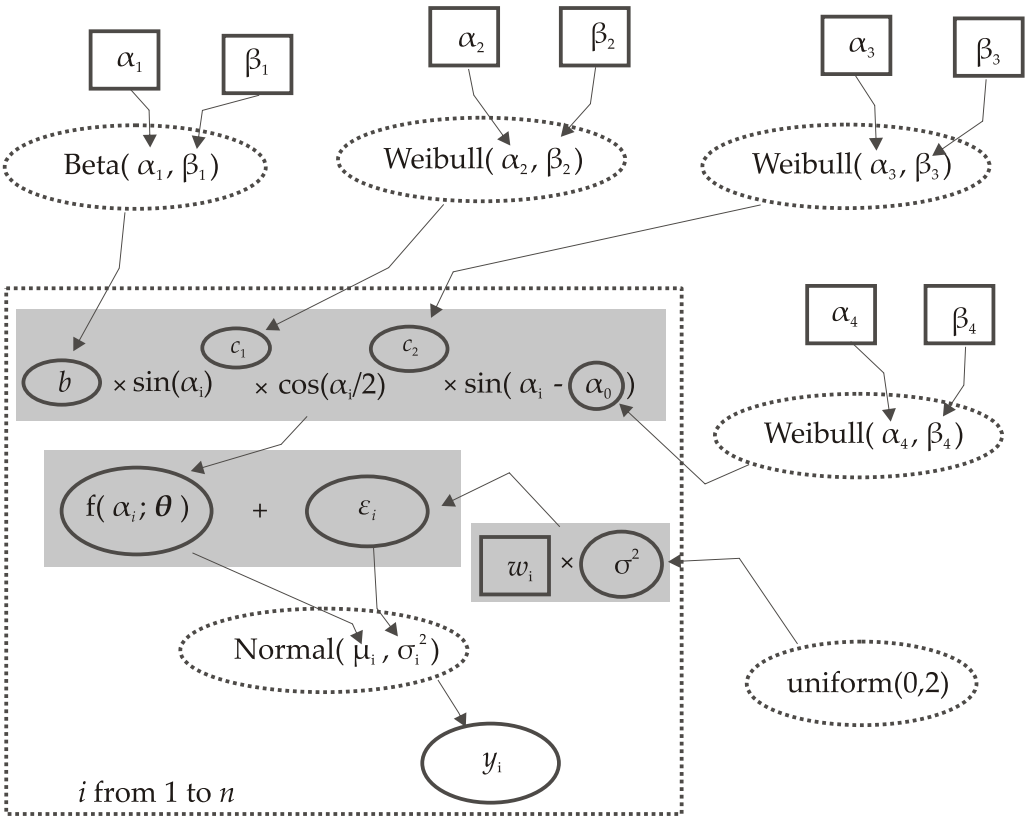

Figure 3.2: The TRIM for the Bayesian regression as a graphical model presentation.

Beta is $[0, 1]$. The parameter α_0 should be in the range $[0°, 180°]$. The practical support for the Weibull distribution for α_0 with any reasonable choice for the shape of the Weibull distribution can ensure that the probabilities are practically zero long before 180°.

In the previous Sec. 2 I used the PPC data for the asteroids, but I will include also the cometary data to be used with the Bayesian regression. The parameters for $\mathcal{D}_{pr}$, called hyperparameters, can be different for asteroids and comets, thus taking into account the basic differences in their polarization behavior, mainly the greater values of polarization of comets especially on the positive branch. After an intensive analysis I ended up with a suggestion for the hyperparameters, presented in Table 3.1. The corresponding distributions are also shown in Fig. 3.3. For $\mathcal{D}_{pr}$ for the residual error variance σ^2 I choose to use an uninformative uniform distribution with a range of $(0, 2)$. Other uninformative $\mathcal{D}_{pr}$ could be used as well.

A priori distributions are always based on a subjective decision, and thus other choices for distributions and hyperparameters could be possible. Among the most important things to consider when choosing a priori is that the expected shape of

Table 3.1: Hyperparameters for a priori distributions for the TRIM for both asteroids and comets. The p(x) for Beta distribution is $x^{\alpha-1}(1-x)^{\beta-1}/\,\mathrm{B}(\alpha,\beta)$, where the B is the Euler beta function. The p(x) for Weibull distribution is $\alpha^{-\beta}\beta x^{\beta-1}\exp(-(x/\alpha)^{\beta})$.

parameter	a priori distribution	hyperparameters for asteroids	hyperparameters for comets
b	Beta	$\alpha_1 = 3.70$ $\beta_1 = 32.23$	$\alpha_1 = 1.43$ $\beta_1 = 2.39$
α_0	Weibull	$\alpha_2 = 0.42$ $\beta_2 = 2.69$	$\alpha_2 = 0.38$ $\beta_2 = 2.38$
c_1	Weibull	$\alpha_3 = 0.47$ $\beta_3 = 1.74$	$\alpha_3 = 0.64$ $\beta_3 = 1.97$
c_2	Weibull	$\alpha_4 = 1.41$ $\beta_4 = 2.20$	$\alpha_4 = 1.02$ $\beta_4 = 1.56$

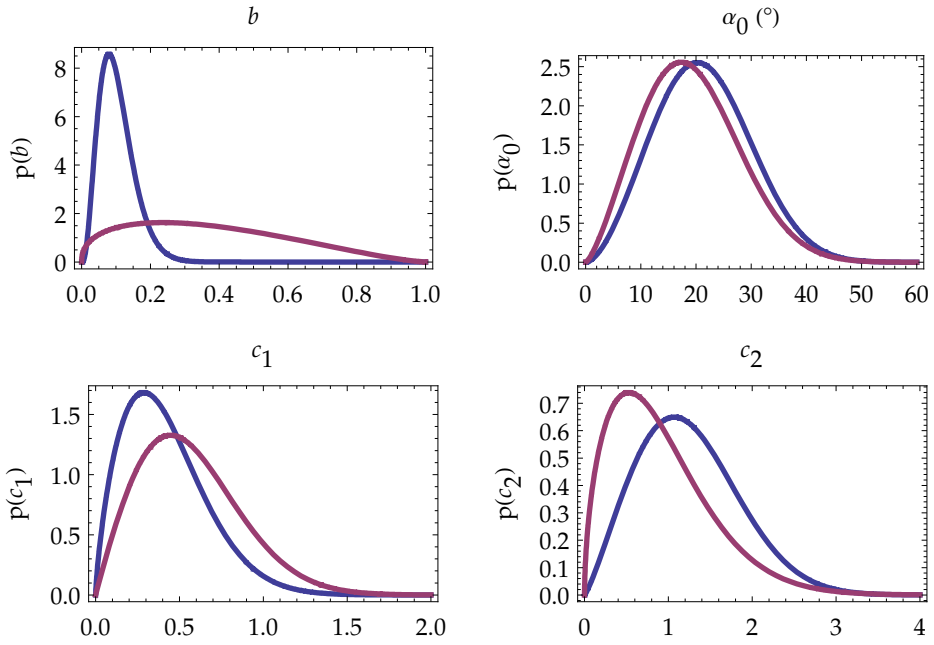

Figure 3.3: A priori distributions for the parameters in the TRIM, blue curve for asteroids and red for comets.

the PPC reflects the overall behavior of polarization observations. The expected PPC's are shown in Fig. 3.4, and they seem to cover the typical shapes of the PPC quite well. Another important subject is that $\mathcal{D}_{pr}$ and L that comes from

the data and the model are in balance. If $\mathcal{D}_{pr}$ has too much information, $\mathcal{D}_p$ will not be sensitive to the observations anymore but will only reflect the a priori information. I will study this sensitivity in Sec. 3.3.2.

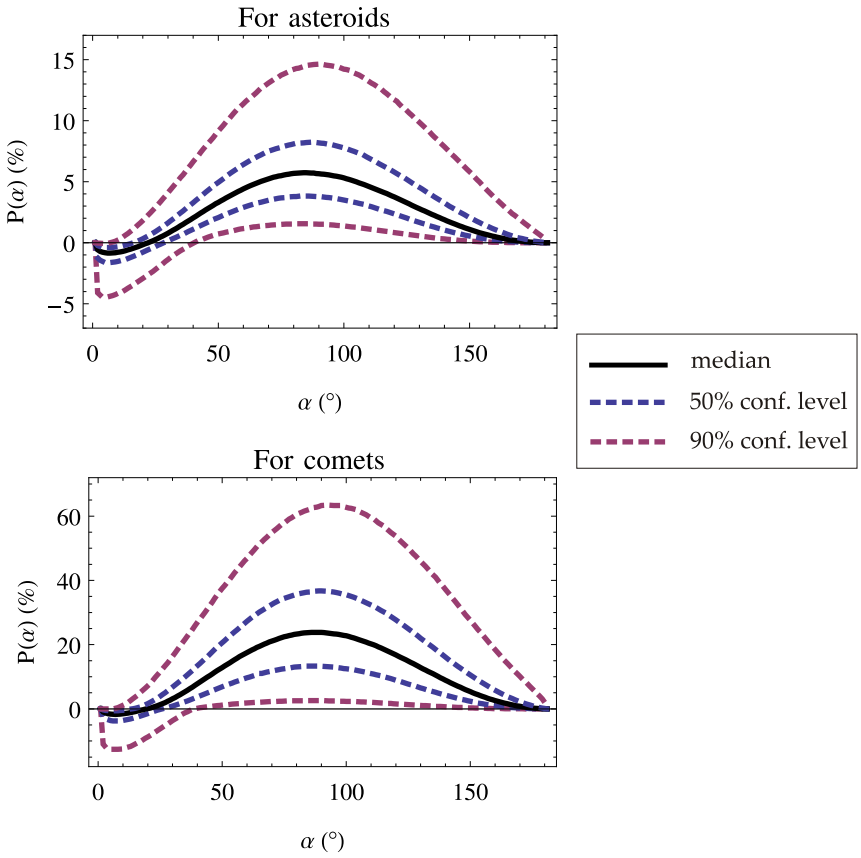

Figure 3.4: The expected shapes of the PPC (given in percents) for asteroids and comets. The median, 50% confidence level and the 90% confidence level of the shape are shown. Confidence levels are empirical, computed by simulating observations of the parameter vector (b, α_0, c_1, c_2) from the corresponding $\mathcal{D}_{pr}$ and using the simulated parameter values in the TRIM (Eq. (2.1)).

3.2.2 Derived polarization features from the PPC model

There are a few important, widely used features of polarization that are interesting and are often reported in polarization studies. These include the inversion angle α_0 where the negative branch turns to positive, the values of maximum negative and positive polarization $P_{\max}$ and $P_{\min}$, and the phase angles where these

are reached $\alpha_{\max}$ and $\alpha_{\min}$, and the slope s of the curve at α_0. The α_0 is included in the TRIM, but other features have to be calculated from the estimated parameters. For the slope s there is a simple expression, $s = b \sin^{c_1}(\alpha_0) \cos^{c_2}(\frac{1}{2}\alpha_0)$. The parameters $\alpha_{\max}$ and $\alpha_{\min}$ can also be derived analytically in closed form but the resulting formulae are too long and complicated to be useful.[†] In practice, $P_{\max}, P_{\min}, \alpha_{\max}$ and $\alpha_{\min}$ can be derived either numerically or analytically.

Within the Bayesian framework analytical results can be derived only for a limited family of parameter distributions, and I do not see any good reason to be limited by that family of conjugate a priori distributions. Instead, in the general case, the GS algorithm converges to sample from the a posteriori distribution of model parameters and after converging can produce an unlimited number of samples. The further analysis is based on these samples. The point-estimates for the model parameters are the median values of parameters sampled from their $\mathcal{D}_p$. Also the error estimates and different confidence intervals (called sometimes the credible intervals in Bayesian language) can be calculated from these samples. In the framework of standard nonlinear regression, the errors for different derived features of polarization could be estimated by the propagation of errors –method. In the MCMC case, it is more straightforward to use the a posteriori samples of the model parameters, form the model function for each and derive the feature under interest. These samples will form an estimate to the a posteriori distribution of the feature, and the confidence intervals can be found numerically.

3.3 Applications to polarization phase curve modeling

Some case studies of the possible uses of the TRIM with Bayesian regression are presented in the following sections. First, the PPC fit for one asteroid and one comet is considered in Sec. 3.3.1. Second, a possibility to predict (extrapolate) polarization values for large phase angles is studied in Sec. 3.3.2.

3.3.1 Phase curves for asteroid Juno and comet Halley

I will demonstrate the use of the TRIM with observations from the main belt asteroid 3 Juno (S) and from the comet 1P/Halley. I have a total of 46 observations of Juno's polarization with phase angles ranging from 2° to 28.7°. These observations have been made at different wavelengths from 333 nm to 952 nm, but

[†]The $\alpha_{\max}$ and $\alpha_{\min}$ can be derived by solving the equation $\frac{\partial \mathrm{P}(\alpha)}{\partial \alpha} = 0$, but the solution in closed form is quite cumbersome. With a help of some small simplifications the Mathematica®, a software package designed for symbolic calculation, will provide the answers. Without any further simplifications the so-called leaf count, the number of atomic expressions in the formula, is $\sim$ 10 000 for both the $\alpha_{\min}$ and $\alpha_{\max}$.

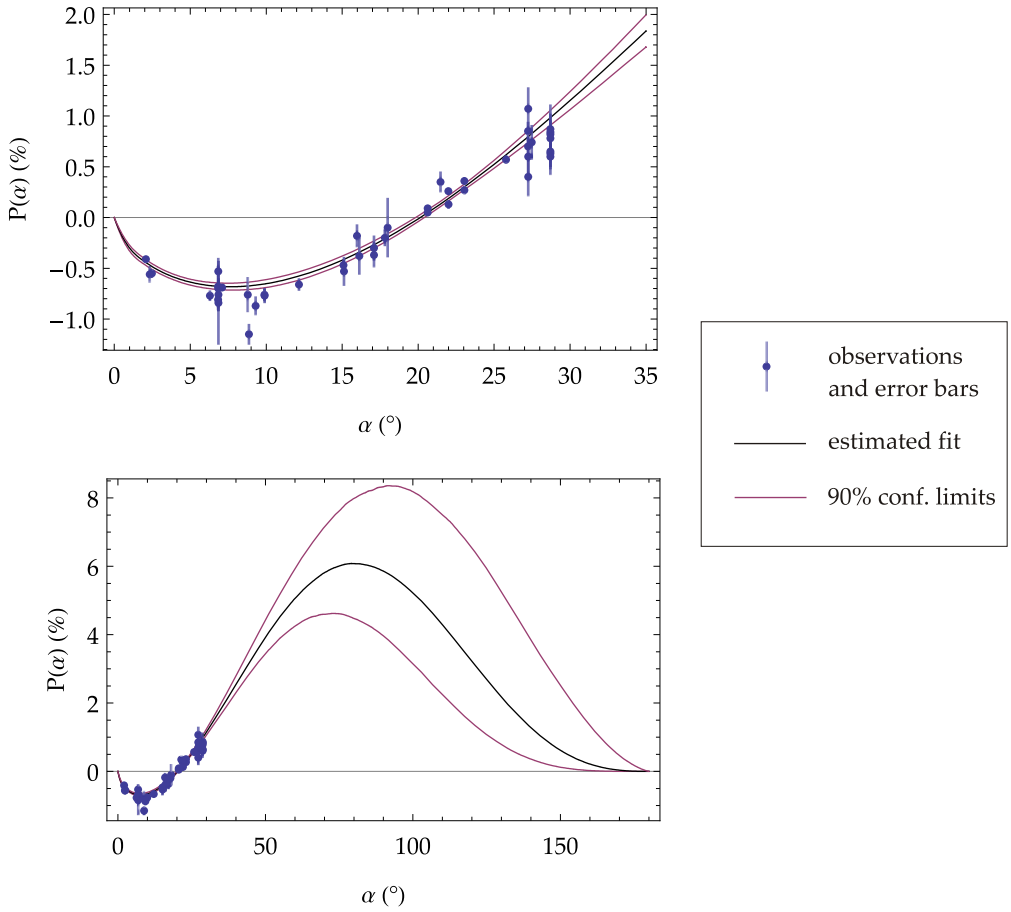

Figure 3.5: PPC fit for Juno. In the upper figure the fit is plotted for the range for which we have observations, and in the lower figure for the whole phase angle range. Estimated fit together with 90% confidence interval for the fit are plotted. Polarization is given in percent.

the (possible) wavelength effect at these phase angles is weak, and the number of observations for each wavelength is too small, so these observations are treated as one phase curve. I will discuss the wavelength effect later in Sec. 4.

Fig. 3.5 shows the estimated phase curve for Juno together with its observations. It can be seen from the figure that the fit is very good and accurate in the range where there is data. However, the confidence intervals of the fit grow noticeably at large phase angles where the data is lacking. Nevertheless, it seems that some extrapolation can be done with a reasonable accuracy, e.g. to something like $\alpha \lesssim 60°$. In Table 3.2 the key features of the polarization and their confidence intervals are presented. It can be noticed that features which are supported by the data — $\alpha_0, s, \alpha_{\min}$ and $P_{\min}$ — can be estimated quite accurately, while the extrapolated features $\alpha_{\max}$ and $P_{\max}$ have very limited accuracy.

Table 3.2: Polarization features for Juno. The 5% column is the lower limit of the 90% confidence interval, and the 95% column is its upper limit. Median is the actual estimate of the feature.

feature	5%	median	95%
α_0	19.8°	20.1°	20.4°
s	0.0895	0.0963	0.103
$\alpha_{\min}$	7.20°	7.63°	8.05°
$\alpha_{\max}$	71.6°	80.3°	88.8°
$P_{\min}$	-0.715%	-0.681%	-0.647%
$P_{\max}$	4.65%	6.08%	8.25%

In Fig. 3.6 and Table 3.3 the phase curve and its features are presented for comet Halley (see references in Levasseur-Regourd et al. 1996). For Halley there are observations for phase angles up to 66°. In that range the wavelength effect can be noticeable, so instead of all wavelengths only the red wavelengths, 630–760 nm, are included. This gives a total of 186 observations. Because of the larger phase angle range, the polarization near the maximum can be predicted with good accuracy.

Table 3.3: Polarization features for Halley. Notations as in Table 3.2.

feature	5%	median	95%
α_0	21.6°	22.2°	22.9°
s	0.234	0.246	0.259
$\alpha_{\min}$	10.1°	10.7°	11.3°
$\alpha_{\max}$	80.3°	84.9°	89.1°
$P_{\min}$	-1.59%	-1.47%	-1.35%
$P_{\max}$	26.4%	27.0%	27.4%

Estimates for the measurement errors in polarization observations are given in the asteroid and comet databases. These errors are generally instrumental errors and the rotation of asteroids whose surface may be inhomogeneous (variegation effects) or the spatial variations in a cometary coma are not taken into account. For these reasons the errors are not realistic as deviations between the observations and a unique and smooth polarization phase curve of the object. The analysis of Juno and Halley gives an estimate that the residual errors between observations and the fitted phase curve are approximately two times as large as the measurement errors for Juno, and 5.8 times as large as those for Halley.

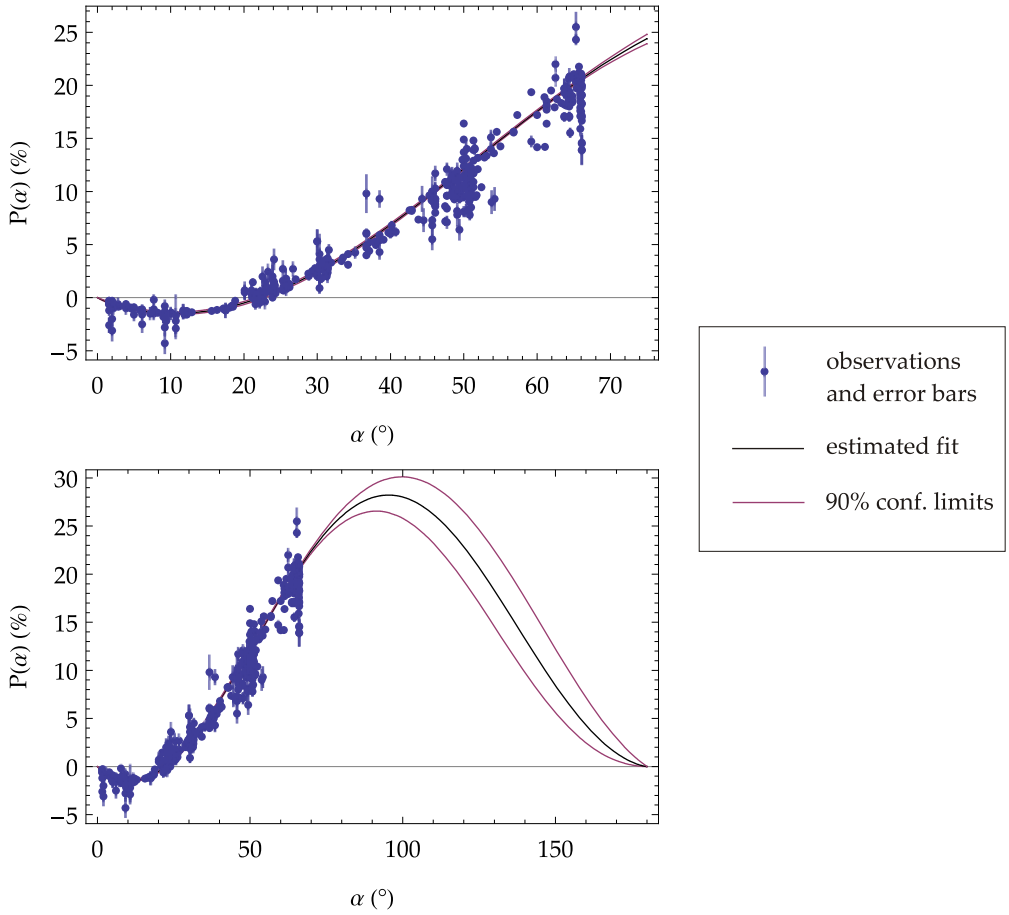

Figure 3.6: PPC fit for Halley. In the upper figure the fit is plotted for the range for which we have observations, and in the lower figure for the whole phase angle range. Estimated fit together with 90% confidence interval for the fit are plotted. Polarization is given in percents.

3.3.2 Prediction of polarization

The Bayesian approach for regression is used because the prior knowledge of the phase curve parameters is needed to obtain a good fit for the TRIM. With any kind of Bayesian analysis, though, one needs to confirm that the a priori distributions do not dominate the a posteriori's. This is especially important if the model is fitted by using just a few observations. If $\mathcal{D}_{pr}$ carry too much information, the fit is biased toward a priori. To check the possible bias in the model some simulation studies are done. I simulate a small number of observations from a phase curve that is different from the a priori phase curve (see Fig. 3.4) and fit the model. If this is repeated many times, it can be seen whether the fits are concentrated properly around the phase curve they were simulated from. In

this case it seems that $\mathcal{D}_{pr}$'s are flexible enough, and even a data set consisting of only five observations contains adequate information, and the fit is unbiased (these results are not shown here).

The model can be further investigated with simulation studies. From the observational point of view, it is important to have some idea on how the number of observations and the phase angles they are taken from affect the accuracy of the fit. This is extremely important if prediction or extrapolation of polarization is needed, or if a given accuracy for some polarization feature is sought. Some results of the TRIM fit with a different number of observations and phase angle ranges are shown next.

The focus is on the prediction of a typical asteroid PPC. A PPC that is slightly different from the a priori distribution is chosen. This allows to check the aforementioned bias of the fit at the same time. As observations from this phase curve are simulated, a typical measurement error calculated from the SMB asteroid data is added, and multiplied by two to obtain a realistic residual error. The multiplication by two was found realistic in the study of Juno in Sec. 3.3.1. By simulating observations and fitting the model over and over again, the confidence limits to the PPC estimate can be set. In Fig. 3.7 the results are shown in the case where 10 or 50 observations only at phase angles below 30° are available. Fig. 3.8 shows how the fit and its capabilities for accurate prediction are improved if the same 50 observations from the range $[0°, 50°]$ could be used. In that same figure there is also an example of a quite typical PPC for a comet, and the confidence intervals in the situation where observations from the range $[0°, 111°]$ are available. It can be clearly seen that if the polarization near its maximum is predicted, then the available phase angle range is the most important factor for an accurate fit.

The phase angle range of the observations has different effects on the four parameters in the model and on their accuracy. The shape parameters c_1 and c_2 are more sensitive. The c_1 affects the curvature of the PPC. If the observations can be done only with small angles, a limited information on the curvature is received which also limits the accuracy in the estimate of c_1. The estimation of c_2 is even more difficult, since the main effect of c_2 is to make the phase curve asymmetric and define the angle of maximum polarization. If the observations are from such a range that the second derivative of the curve after the inversion angle is still positive, a very poor information of the possible place of the maximum is gathered. Thus, of all four parameters, c_2 has the lowest accuracy.

High $P_{\max}$ comets

The behavior of the TRIM can also be studied with real observations instead of simulations — a subgroup of our comet data, the so-called high $P_{\max}$ comets (Levasseur-Regourd et al. 1996). In the comet data the comets Halley, Hyakutake

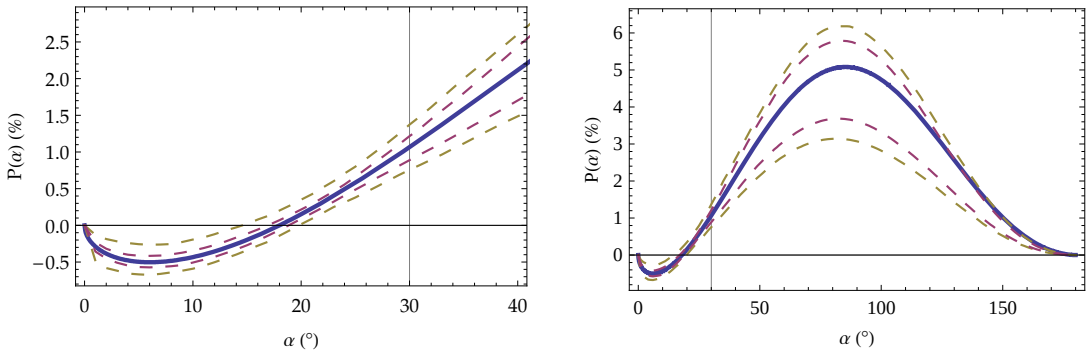

Figure 3.7: Simulated 90% confidence intervals for the PPC estimate when 10 or 50 observations are done from the phase curve marked with a solid blue line. The observations are simulated from evenly distributed angles in the range $[0°, 30°]$. On the left there is a more detailed plot from that area, and on the right the plot for the whole phase angle range $[0°, 180°]$. The larger confidence interval (dashed yellow lines) is for 10 observations, and the smaller interval (dashed red lines) is for 50 observations.

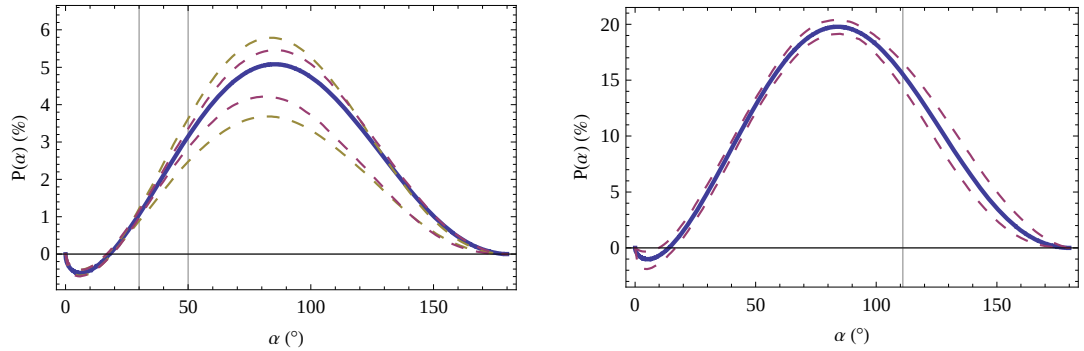

Figure 3.8: The effect of a larger phase angle range for the observations. On the left there is the 90% confidence interval for the PPC estimate when 50 observations are simulated from the range $[0°, 30°]$ (dashed yellow lines), and when the same number of observations is retrieved from the range $[0°, 50°]$ (dashed red lines). On the right, there is a confidence interval for the estimate for a typical comet phase curve (solid line) when 50 observations are retrieved from the range $[0°, 111°]$.

and West belong to this group. These comets have similar polarimetric properties at large phase angles, so it is reasonable to fit a single PPC for all the comets in this group.

The observations in the red wavelength domain are used, resulting in 229 obser-

vations in the phase angle range $[1.57°, 111.3°]$ for two experiments. In the first experiment the observations at phase angles up to 70° are used to fit the PPC. In the second experiment, only the observations at phase angles up to 30° are used. The results of these experiments are shown in Fig. 3.9, where the resulting fits and the corresponding confidence intervals are compared to the global PPC fit for the high $P_{\max}$ comets using all the observations. In both experiments the fit using a limited number of observations is very close to the global fit. The confidence interval for the polarization phase curve is significantly smaller at large phase angles for the fit using observations up to 70°, but the fit using observations up to 30° has also enough accuracy to be useful in prediction at phase angles up to, say, 60°–70°.

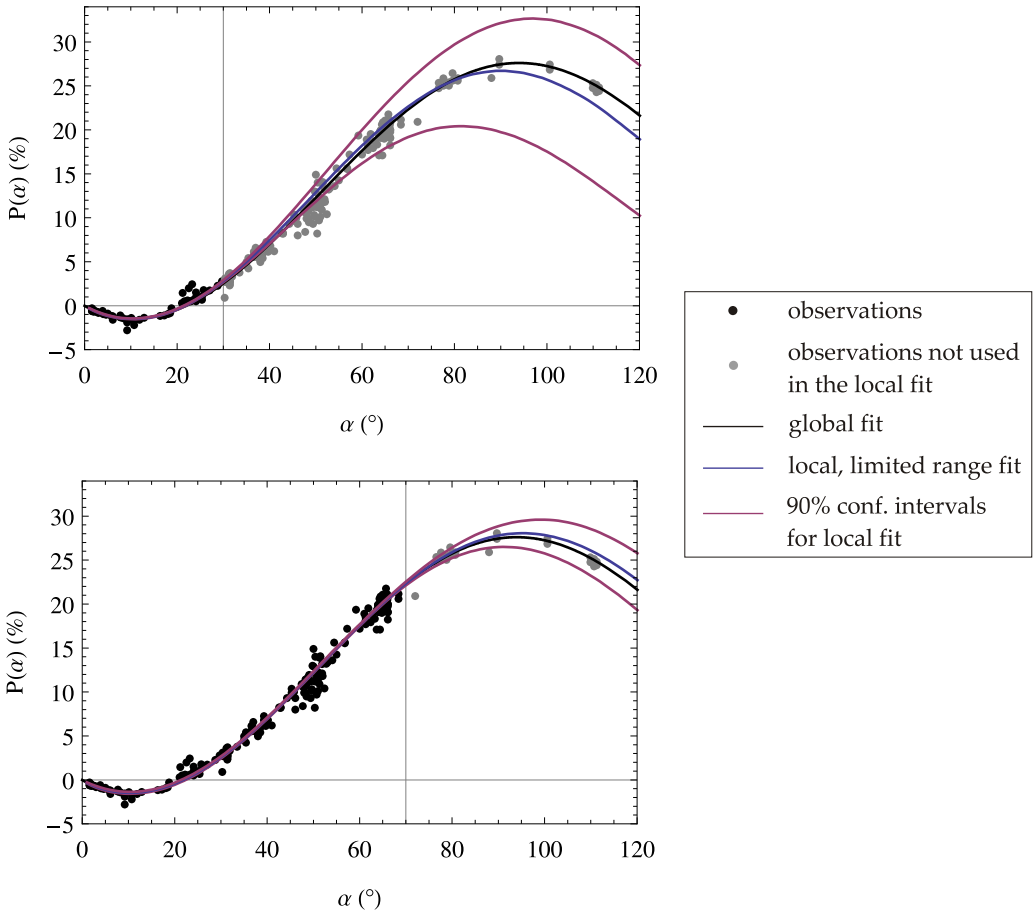

Figure 3.9: Prediction of polarization for the high $P_{\max}$ comets. On the upper figure the local PPC fit using observations up to 30° and on the lower figure the fit using observations up to 70° are shown. In both figures the global fit, using all the observation points, is shown together with the local fit using the limited range of observations and its 90% confidence intervals.

Chapter 4

Modeling multi-wavelength data

Until this point I have modeled the PPC without taking into account the wavelength λ with which the observations have been made. This approach is feasible when the observations are taken only with a single wavelength or wavelength filter, or if the effect of different wavelengths is weak in the observations. Whether the wavelength effect (WE) is weak or significant is however a question for which it is hard to answer without analyzing the data at issue.

In principle the (linear) polarization, as well as all electromagnetic scattering, is a function of the wavelength. Or, to be exact, a function of the combination of λ, the refractive index of the material with given λ, and the size of the object. For wavelengths from totally different λ domains, e.g. from radio and X-ray domains, there is surely a significant change in the polarization. For the data in the visible λ domain and its neighborhood the near-infrared and near-UV domains from where the polarimetric observations in the data are, however, the WE can be weak. Several studies (e.g. Gehrels & Teska 1963, Rosenbush 2005, Kiselev et al. 2008) have been made on the possible WE on the polarization observations from comets. Usually the comets provide more fruitful data for the study since the WE tends to be weak in the small phase angles from where the asteroid observations are.

Studying the WE can be quite tricky, since in many cases the different errors, like the measurement error of the instrument or the error due to the evolution of the comet over observation time, can be of the same magnitude as the possible WE. As there are so many error sources, the PPC model and the estimation procedure should be as robust and reliable as possible. In this section I extend the models discussed in Secs. 2 and 3 to be used with the multi-wavelength data, and use the derived models in studying the WE. First in Sec. 4.1 a (regular) nonlinear regression model with multiple response is introduced, and in Sec. 4.2 a Bayesian approach is studied together with some applications to the polarization data.

4.1 Multiple response regression model

One possibility to model the multi-wavelength data is to use a multiple response (MR) regression model, where the dependent variable $\mathbf{Y}$ is a matrix and each observed $\mathbf{y}_i$ is a vector containing multiple responses to the model. With the PPC data this means that for each observation angle α_i there are the observations $\mathsf{y}_{i1}, \ldots, \mathsf{y}_{ik}$ for the k wavelengths $\lambda_1, \ldots, \lambda_k$. To put it into a mathematical notation, the MR model is

$$\mathbf{Y} = \mathbf{F}(\boldsymbol{x}; \boldsymbol{\theta}) + \mathbf{E}, \tag{4.1}$$

where the $\mathbf{Y}$ is $n \times k$ matrix, $\mathbf{F}$ is a composite function of the $n \times 1$-valued functions $\mathbf{f}_j(\boldsymbol{x}; \boldsymbol{\theta})$ for each $\lambda_j, j = 1, \ldots, k$ and $\mathbf{E}$ is the $n \times k$ error matrix.

The difference between the simple one-response regression model and the MR model lies in the $\boldsymbol{\theta}$ and $\mathbf{E}$. The parameter vector $\boldsymbol{\theta}$ needs not to be the same as for the one-response model — this $\boldsymbol{\theta}$ can have additional parameters that are effective for only certain $\mathbf{f}_j$'s. For $\mathbf{E}$ it is usually assumed that

$$\epsilon_{ij} \not\perp\!\!\!\perp \epsilon_{ij'} \quad \text{for any } j \text{ and } j', \text{ but} \tag{4.2a}$$

$$\epsilon_{ij} \perp\!\!\!\perp \epsilon_{i'j'} \quad \text{for any } i \neq i', j \text{ and } j'. \tag{4.2b}$$

This assumption implements the underlying idea in the MR model that the multiple responses for a single x_i are correlated. If there were no correlation between the components in the response vector $\mathbf{y}_i$ the model in Eq. (4.1) could be rewritten to k regular one-response models and estimated separately.

As a result of the model formulation, the covariance structure of the error (and thus the response $\mathbf{Y}$, too) differs from the the regular model. The covariances in Eq. (4.2a) can be gathered into a matrix $\boldsymbol{\Sigma}$ so that $\mathbf{Cov}(\boldsymbol{\epsilon}_i) = \boldsymbol{\Sigma}$ for any row vector $\boldsymbol{\epsilon}_i$ in $\mathbf{E}$. The covariance matrix for the whole $\mathbf{E}$ is

$$\mathbf{Cov}(\,\mathbf{vec}(\mathbf{E})\,) = \begin{bmatrix} \boldsymbol{\Sigma} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \boldsymbol{\Sigma} \end{bmatrix} = \mathbf{I}_n \otimes \boldsymbol{\Sigma}, \tag{4.3}$$

where the function **vec** converts the matrix $\mathbf{E}$ into a long vector where the rows of $\mathbf{E}$ are stacked one after another, and $\otimes$ is the Kronecker product between matrices.

4.1.1 Likelihood and estimation in the multiple response model

Starting from the joint probability density function of the observations $\mathbf{y}_i$ one can derive the log-likelihood function of the model parameters $\boldsymbol{\theta}$ and $\boldsymbol{\Sigma}$, which can be written as

$$\mathrm{l}(\boldsymbol{\theta}, \boldsymbol{\Sigma}) = -\frac{n}{2} \log(|\boldsymbol{\Sigma}|) - \frac{1}{2} \left[\sum_{i=1}^{n} (\mathbf{y}_i - \mathbf{f}_i(\boldsymbol{\theta}))' \, \boldsymbol{\Sigma}^{-1} \, (\mathbf{y}_i - \mathbf{f}_i(\boldsymbol{\theta})) \right]. \tag{4.4}$$

If we compare this log-likelihood to the one from the one-response model in Eq. (2.8) we can see one important difference. In the one-response model it is possible to maximize the log-likelihood of $\boldsymbol{\theta}$ regardless of the σ^2 and vice versa. For MR model this is not possible in general, since the $\boldsymbol{\Sigma}$ is so profoundly entangled with the $\boldsymbol{\theta}$.

There are a few ways to find the ML estimates $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\Sigma}}$ in this case. The straightforward way is to maximize the log-likelihood numerically as a function of $(\boldsymbol{\theta}, \boldsymbol{\Sigma})$. This can be computationally challenging in some cases. The other way is to iteratively maximize the profile log-likelihood for $\boldsymbol{\theta}$ and $\boldsymbol{\Sigma}$ successively. I have chosen the first method of direct numerical optimization to be used in the following section where the MR model is used with the polarization data.

4.1.2 Wavelength effect modeled with the multiple response regression

The wavelength λ from which the observations are done is a continuous variable in principle. Continuous variable could be added to model as it is, but in the case of λ there are several reasons against it. Firstly, the observations are not actually done with a single monochromatic value of λ, but instead with a wavelength filter with a continuous absorption pattern over the spectra with the midpoint at λ. The pattern of the filter can vary from one observation set to another, so the different values of λ are not strictly speaking comparable. Secondly, the functional form of the effect of the λ should be set beforehand, which is more complicated with the nonlinear model than with a linear model. With linear model the additional variable λ, or some function of λ would be multiplied with its parameter estimate and summed up to the model. With nonlinear model the λ could be involved in the model in any imaginable way. To choose the proper functional form for the λ in the model one would have to know the WE beforehand. So, instead of a continuous variable, the different values of λ are divided into a more coarse classification of colors as mentioned in Sec. 1.1.3 making the MR model a suitable tool for the analysis.

There are some requirements to the data that can be used with the MR model. Most importantly there cannot be any missing values in the data matrix. This turns out to be quite limiting demand to the polarization data in hand. With a chosen set of filters, i.e. wavelengths $\lambda_1, \ldots, \lambda_k$ there should be polarization observation for every λ_i with every α. When observing an object, however, using all the filters simultaneously is usually not possible and the filter in use must be changed before making another observation. Due to this it is common that one set of observations are done with one filter, and another set with another filter at some later time resulting that the observation angles α are never the same. The best complete dataset can be found from the comet Halley. This dataset consists

of 22 observations from the phase angle range of $(1.6°, 64.4°)$ using three colors: blue, yellow and red domain. The data is presented in Fig. 4.1.

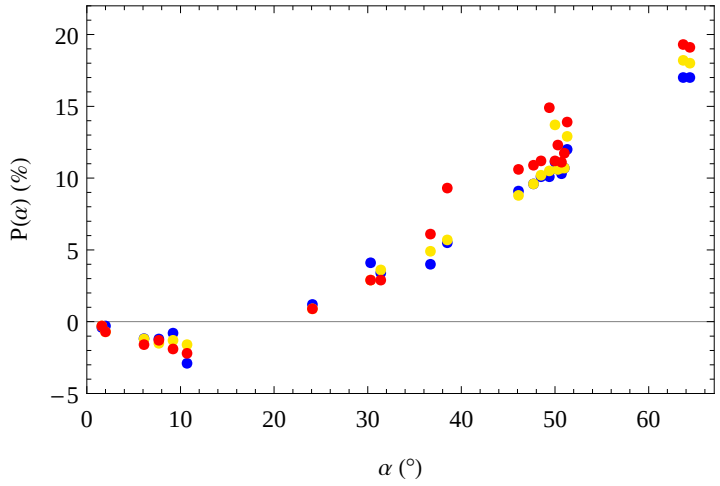

Figure 4.1: The dataset from comet Halley to be used with different MR models. Three wavelength domains, blue, yellow and red are plotted with the corresponding colors.

Different models

The MR framework can be used to specify models that will model the WE in different ways. This is done by specifying different versions of the model function $\mathbf{F}$ and its parameters $\boldsymbol{\theta}$ and of the $\boldsymbol{\Sigma}$.

First, a basic form of the model is specified to the observations. It can be seen from Fig. 4.1 that the data does not provide good information about the value of $P_{\max}$ or its location $\alpha_{\max}$. As discussed in Sec. 2.2, the LEM is ill-conditioned and perhaps not valid for this large α range, and the TRIM has parameter identification issues if the data does not support the estimation of the $P_{\max}$. Thus, I choose the sub-TRIM as the basic shape for the PPC.

The 'full model', i.e. the model where there are no shared parameters between the $\mathbf{f}_j$'s and no limitations are set to the error matrix $\boldsymbol{\Sigma}$ has a total of 15 parameters to be estimated: the b^j, α_0^j and c_1^j for the three wavelength domains $j = 1, 2, 3$ and a 3×3 matrix $\boldsymbol{\Sigma}$ with six free entries, the residual covariances $\sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{12}, \sigma_{13}$ and σ_{23}. Note that if the residual covariances between the wavelengths $(\sigma_{12}, \sigma_{13}, \sigma_{23})$ are set to zero this model can be decoupled into three separate models.

The estimation of the full model is very challenging, at least it was with the Halley dataset of only 22 observations, so I tried assumptions that would reduce the

number of free parameters. I restricted myself to models where a single parameter c_1 is shared between the wavelengths instead of the three separate c_1^j's, reducing the number of parameters from 15 to 13. This did not help, and the optimization of the log-likelihood in Eq. (4.4) did not succeed.* While the optimization did not converge well enough, it seemed however that the residual covariances between the wavelengths could be quite small. This can imply that as the modeled part is subtracted from the observations the residuals are not anymore correlated, which would be a very desirable feature and would tell that the model is adequate. Therefore I made a further assumption that the between-λ covariances are set to zero and the error matrix is of diagonal form, $\mathbf{\Sigma} = \mathrm{diag}(\sigma_{11}, \sigma_{22}, \sigma_{33})$.

With the aforementioned restriction the following types of models M were finally estimated and compared:

short name	parameters to be estimated in $\boldsymbol{\theta}$	note
M_0	$b^1, \alpha_0^1, c_1^1,\ b^2, \alpha_0^2, c_1^2,\ b^3, \alpha_0^3, c_1^3$	full model, separable into three models
M_1	$b^1, \alpha_0^1,\ b^2, \alpha_0^2,\ b^3, \alpha_0^3,\ c_1$	only joint c_1
M_2	$b^1,\ b^2,\ b^3,\ \alpha_0, c_1$	joint c_1 and α_0
M_3	$\alpha_0^1,\ \alpha_0^2,\ \alpha_0^3,\ b, c_1$	joint c_1 and b
M_4	b, α_0, c_1	same model for all, except for σ_{jj}

For all the models, in addition to $\boldsymbol{\theta}$, the residual covariances $(\sigma_{11}, \sigma_{22}, \sigma_{33})$ need also to be estimated, resulting the number of estimated parameters to vary from 12 for M_0 to 6 for M_4. All the models are nested, i.e. the model M_0 is the full model and all the other models M_1–M_4 are derived from the M_0 by setting restrictions for the parameters, such as $c_1^1 = c_1^2 = c_1^3$ when moving from M_0 to M_1.

Model comparison

The fact that the models are nested is important for the comparison of the models and for choosing the proper statistics for the comparison. The likelihood ratio (LR) is a suitable test statistic to compare nested models and the (asymptotic) distribution of the LR can be derived under the hypothesis (H_0) that the full model and the restricted model are equally good. The LR statistics is

$$LR = -2\left[\mathrm{l}(\hat{\boldsymbol{\theta}}, \hat{\mathbf{\Sigma}}) - \mathrm{l}(\tilde{\boldsymbol{\theta}}, \tilde{\mathbf{\Sigma}})\right], \tag{4.5}$$

where the $(\tilde{\boldsymbol{\theta}}, \tilde{\mathbf{\Sigma}})$ are estimated parameters for the full model, and $(\hat{\boldsymbol{\theta}}, \hat{\mathbf{\Sigma}})$ for the restricted model. By definition the maximum of the log-likelihood for the full

*The built-in Nelder-Mead simplex algorithm in Mathematica® with multiple random seeds is used in the optimization.

model is always non-positive but as large or larger than for the restricted model. Thus, for equally good models the LR is 0 and large values of the LR indicate that the restricted version is worse. Under H_0 the asymptotic distribution of the LR is

$$LR \underset{as}{\sim} \mathcal{X}^2_q, \tag{4.6}$$

the chi-squared distribution with q degrees of freedom, the q being the difference in the number of parameters for the models. The LR values and the p-values for the hypothesis H_0 that the models are equally good are presented in Table 4.1. From the p-values we can see that only the model M_4 seems to be inferior, statistically speaking, to the full model in modeling the observations. The M_4 was a model where the different wavelengths shared all the model parameters b, α_0, c_1, so this implies that there actually is some kind of wavelength effect that needs to be modeled.

Table 4.1: Comparison of the nested models M_0–M_4. The last column is the p-value for the LR test.

Model	$l(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\Sigma}})$	LR	q	p
M_0	-19.4			
M_1	-20.0	0.965	2	0.62
M_2	-20.2	1.51	4	0.82
M_3	-22.0	5.02	4	0.29
M_4	-29.8	20.7	6	0.002

The models M_2 and M_3 both have the same number of parameters. The wavelength effect is modeled with varying b in M_2 and with varying α_0 in M_3. From these two models the M_2 is better in terms of the LR statistics. The conclusion is that the wavelength effect is not so strongly visible in the inversion angle α_0 as it is in the overall amplitude b.

When comparing the more complicated M_1 to M_2 we see that the LR is only slightly better for the M_1 and the larger degrees of freedom makes the p-value of the M_1 smaller. Thus, it is perhaps not wise to include both the wavelength-dependent b and α_0 to the model since including only wavelength-dependent b seems to be enough.

In addition to the statistics about the model goodness we should also try to compare the models and their soundness visually, and by studying the possible wavelength effect they propose for some interesting polarization features. The estimated models, together with the data, are shown in Fig. 4.2. There seems to be no large differences between the models. Estimates for the range $\alpha \in [0^\circ, 30^\circ]$ or so are quite similar between all the models, except for M_3 where the λ-dependence is modeled with varying α_0.

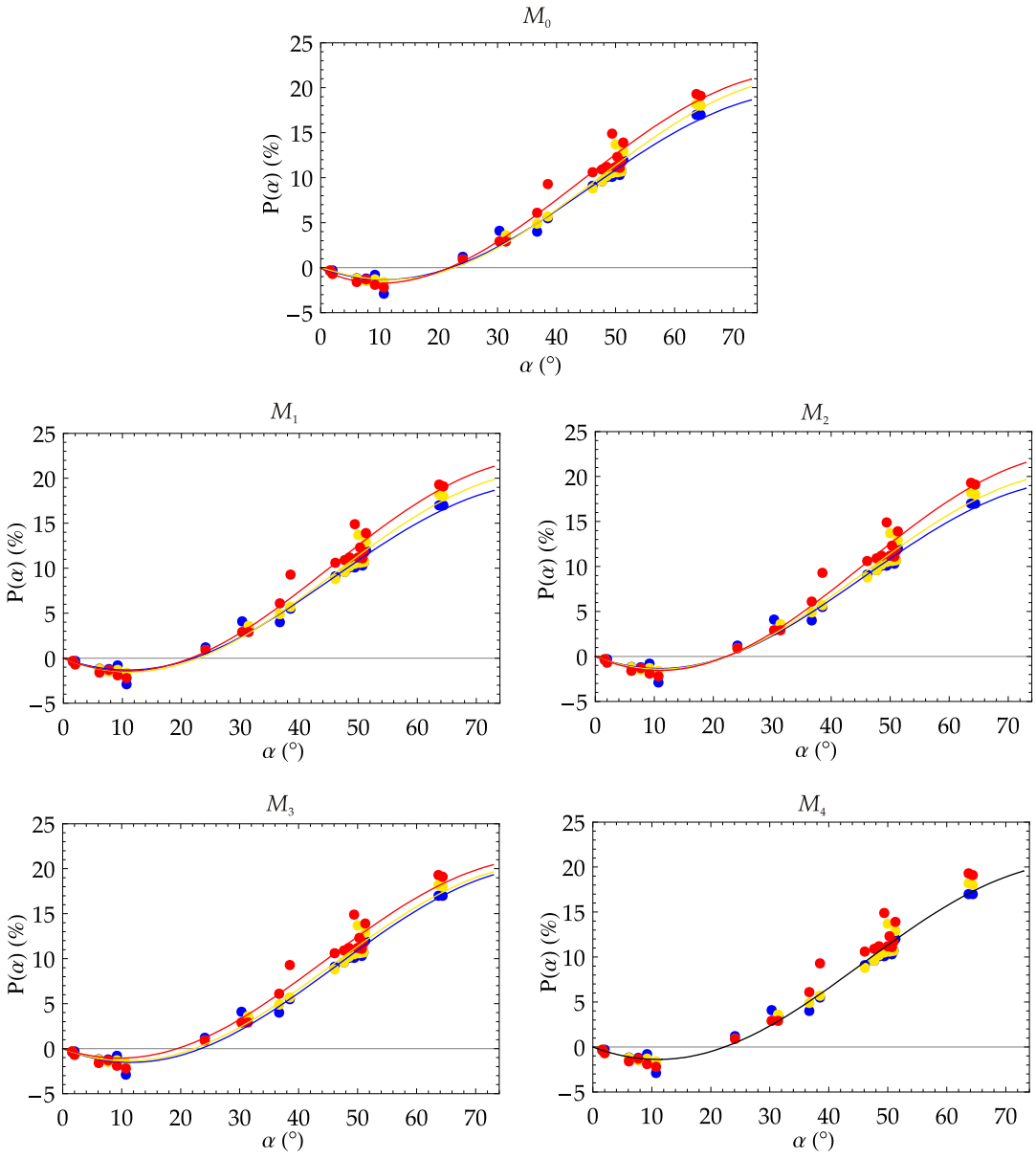

Figure 4.2: The different MR models M_0–M_4 for the data from comet Halley. The line and dot colors indicate the three wavelength domains, blue, yellow and red.

All the models in the aforementioned figure seem to be quite similar globally. However, there are some fine differences if we study some important features of polarization and the modeled wavelength effect in those. Based on the conclusions made from the LR statistics in Table 4.1 the models M_1–M_3 are the most prominent. The estimates for the $P_{\min}$, $P_{\max}$, and slope at α_0, the s, are plotted in Fig. 4.3. The λ-dependence is not modeled similarly in the models. If we assume

that WE is a monotonic in these colors the estimates should be ordered as blue, yellow and red, or vice versa. For M_2 and M_3 this is true, but not for M_1 for $P_{\min}$ or s. The estimation of both the λ-dependent b and α_0 simultaneously with M_1 might be too delicate task and I might judge the result to be highly questionable. With M_2 and M_3 we have the problem that their prediction for the WE with $P_{\min}$ and s are the opposite. The M_2 predicts that the $P_{\min}$ decreases with increasing λ and that the s increases with increasing λ, but the M_3 predicts otherwise. So, the nature of the modeled WE depends on which model we choose.

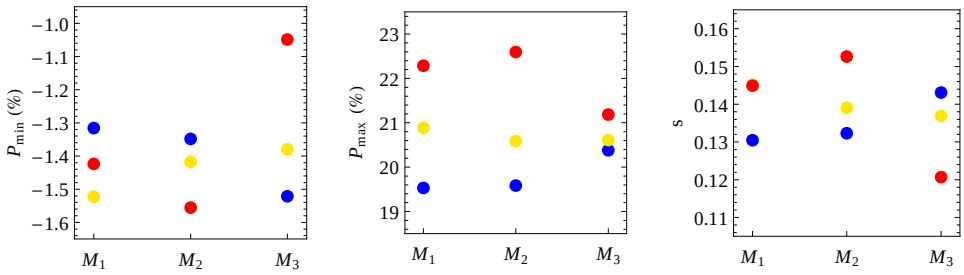

Figure 4.3: The polarization features $P_{\min}, P_{\max}$ and s for the MR models M_1–M_3. The estimates are presented as dots with color corresponding the wavelength domain blue, yellow or red.

I tend to favor the M_2, because the WE with M_1 supports more the effect with M_2, and because the LR statistics is best with M_2. It is however clear that we need either more suitable data for the MR model analysis or other approaches to draw any final conclusion about the WE in cometary data. The Bayesian approach that is dealt in the next section can give a more confident answer to the question about the nature of the WE.

4.2 Bayesian approach for multi-wavelength observations

The downside with the classical frequentist MR modeling in previous section is the lack of datasets that are of proper form to be used in the analysis. Because the continuous wavelength variable was divided into more coarse color classes, one common approach is to include the class information into the model as a set of so-called dummy variables, i.e. k classes are coded into $k-1$ binary and mutually independent explanatory variables. The evident change to the MR model is that the multiple responses are exchanged into one response per one explanatory variable, but the dimension of the explanatory variable vector increases to include

the wavelength information as the dummy variables. There are no requirements anymore that for a given α there should be observations for every color class, and this makes it possible to include all the observed data in the analysis.

While the dummy variable approach is certainly superior to the MR modeling in that sense that it can take advantage of all the available data, the classical (nonlinear and multiple) regression analysis has one weakness. The idea in the MR modeling was to include the correlation between the observations with similar λ and with the same α value in the model and use it to increase the information about the WE. The simple regression does not offer tools to gain from the fact that some observations should be correlated.

The approach that I will use is to introduce a sort of hierarchy into the model. There is a branch in classical statistics called hierarchical modeling or multi-level modeling, but I have used a Bayesian approach to the model. The approach is described in the following section.

4.2.1 Bayesian model for multi-wavelength observations

The shape of the model function is basically the TRIM, the same as in the case of one-wavelength data in Sec. 3.2. The addition is that the wavelength variable will be added to the explanatory variables in a form of a dummy variable set $\boldsymbol{\lambda}$. With k different wavelength (color) classes or domains an arbitrary color will be the 'reference group' with $\boldsymbol{\lambda} = \mathbf{0}$ and the other $k-1$ colors will be coded with $\boldsymbol{\lambda}$ having components $\lambda_{i=j} = 1$ and $\lambda_{i\neq j} = 0$ for a color j.

The possible WE is set to manifest itself in the model via the parameters b and α_0, but the parameters c_1 and c_2 are left untouched. This is because the shape parameters c_1 and c_2 are the most difficult to estimate, and in most cases the shapes of different wavelength phase curves are very similar. Thus, the wavelength effect is only studied through the inversion angle α_0 and the magnitude parameter b. The model is of the form

$$\mathrm{P}(\alpha, \boldsymbol{\lambda}) \;=\; (b+\boldsymbol{\lambda}'\,\mathbf{b}^*)\,\sin^{c_1}(\alpha)\,\cos^{c_2}(\frac{\alpha}{2})\,\sin(\alpha-(\alpha_0+\boldsymbol{\lambda}'\,\boldsymbol{\alpha}_\mathbf{0}^*)), \qquad (4.7)$$

where the parameters for the WE, $\mathbf{b}^*$ and $\boldsymbol{\alpha}_\mathbf{0}^*$ are vectors of the length $k-1$ as the $\boldsymbol{\lambda}$, and components b_j^* and α_{0j}^* model the difference between the reference group and the corresponding color domain j. The $\boldsymbol{\lambda}'$ is the transpose of the vector $\boldsymbol{\lambda}$.

The one arbitrary reference color will use the same parameters and a priori distributions as in the one-wavelength Bayesian model described graphically in Fig. 3.2. For other colors the parameters b_j^* and α_{0j}^* introduce deviations from that 'main' curve. By using suitable a priori distributions for the parameters $\mathbf{b}^*$ and $\boldsymbol{\alpha}_\mathbf{0}^*$ we can force the phase curves for different colors to have similar shapes. This is a very important feature because it is known from practice that the PPC at different wavelengths for the same object correlate strongly. If a completely separate

modeling for each wavelength would be used, also quite a lot of information would be lost making the already difficult analysis of the WE even more cumbersome.

Both the vectors $\mathbf{b}^*$ and $\boldsymbol{\alpha}_\mathbf{0}^*$ are chosen to have a multinormal distribution $\mathcal{N}_{k-1}$ as a priori distribution. The expected value in the a priori will be zero vector, thus assuming no WE a priori. The covariance matrices are set to diagonal matrices $\sigma_b^2\mathbf{I}$ and $\sigma_{\alpha_0}^2\mathbf{I}$. If the values for the hyperparameters σ_b^2 and $\sigma_{\alpha_0}^2$ are kept small the a priori belief is that the PPC's for all the colors are highly correlated, unless the observations strongly indicate otherwise. The asteroids and the comets can have somewhat different WE, and I will use values $\sigma_b^2 = 0.015, \sigma_{\alpha_0}^2 = 2^\circ$ for asteroids and $\sigma_b^2 = 0.03, \sigma_{\alpha_0}^2 = 3^\circ$ for comets.

4.2.2 Bayesian modeled wavelength effect

The comets Hale-Bopp, Halley, Hyakutake, Swift-Tuttle and West are included in the analysis because these objects can be studied at different wavelengths with data at phase angles larger than 40°. For asteroids, the only good dataset for this purpose is available from Toutatis. When asteroid datasets are divided into different color domains, there are usually just a few observations per domain. Furthermore, these observations tend to be at small phase angles, where the WE is nearly inexistent between the visible color domains compared to the measurement errors. There are 174 observations of Toutatis ranging from 13.55° to 111°, which makes it possible to use it in the analysis.

The wavelength filters used in polarimetric measurements can be slightly different between instruments, and also it seems possible that the practice of reporting the filters can vary inside the database from one observer to another. The filters for cometary measurements have to be adapted to avoid the gaseous emission lines and, when possible, narrow band filters are used and the results eventually corrected by the knowledge of the emission spectrum. Therefore, as mentioned in Sec. 1.1.3, the wavelengths of the observations are divided into UV, violet, blue, green, yellow and orange, red and IR.

Figs. 4.4a and 4.4b present the observations and the fitted multi-wavelength models for the data. The model seems to behave reasonably, the PPC for the different wavelengths tend to be roughly the same, and the differences are quite logical, e.g. the curves are more or less ordered according to the wavelength.

From the fitted models in Figs. 4.4a and 4.4b the estimates and confidence intervals for polarization features as a function of the wavelength can be derived. There are some notable WE, but in most cases these effects are not similar for all the comets, which implies that these are not global effects, but only applicable to that particular comet. An example of this is in Fig. 4.5, where the wavelength effect for Hale-Bopp is presented for polarization features α_0, $\alpha_{\min}$ and $P_{\min}$. As seen in Fig. 4.5, Hale-Bopp shows a clear WE in the inversion angle. On the

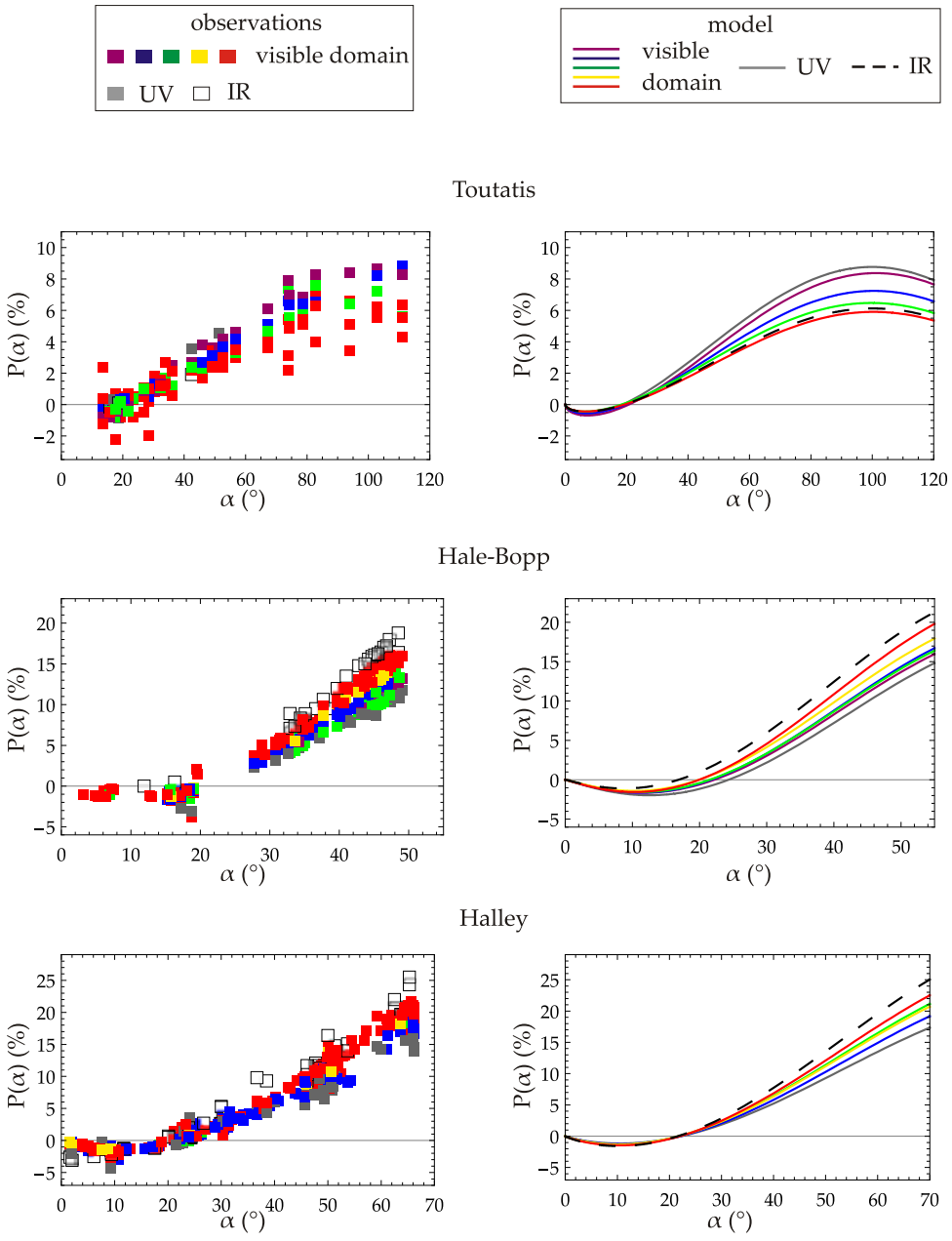

Figure 4.4a: The polarization observations in the left panel and the estimated PPC models in the right panel for one asteroid and two comets.

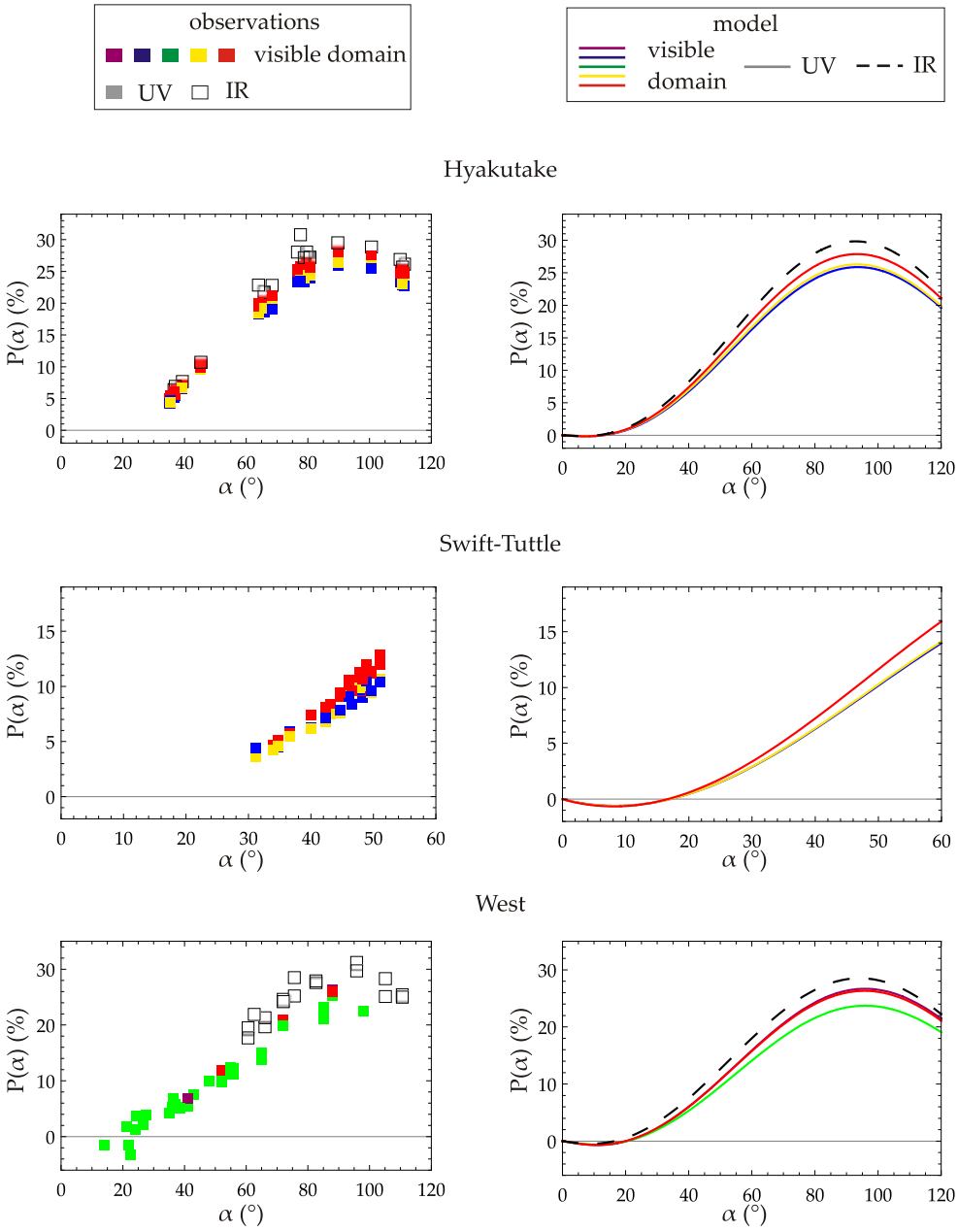

Figure 4.4b: The polarization observations and the estimated PPC models for three comets.

other hand, the inversion angle for comet Halley is practically constant at all the wavelengths. Comets Hyakutake and Swift-Tuttle have no observations near the inversion angle and therefore it is impossible to draw conclusions about the WE in the inversion angle, and the same applies for comet West, for which there is a lack of data at some wavelengths.

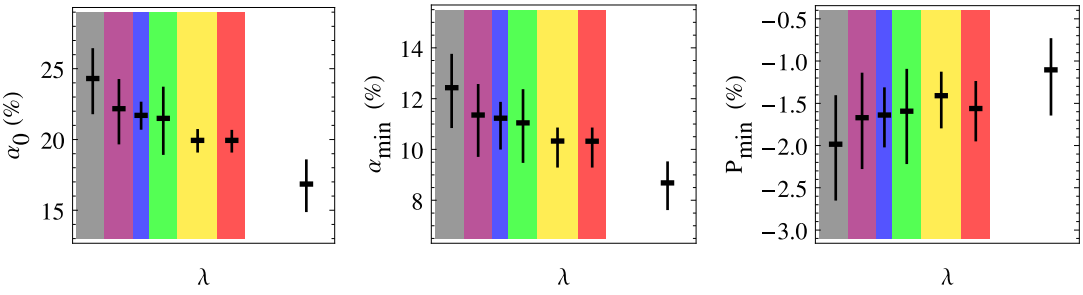

Figure 4.5: Wavelength effects for comet Hale-Bopp in α_0, $\alpha_{\min}$ and $P_{\min}$. The vertical lines are the 90% confidence intervals, and the horizontal bars are the point-estimates. The y-scale is the wavelength with background color showing either the visible spectra color, or UV with gray and IR with white.

The only clear WE is seen either in the minimum value of polarization, or much clearer in the maximum value of polarization $P_{\max}$ in Fig. 4.6. For all the comets the $P_{\max}$ increases with wavelength, and an inverse effect is observed for the S-type asteroid Toutatis.

4.3 Condluding remarks

If there is a need to fit a continuous curve to a sufficient number of polarization observations from small phase angles and perhaps use the model to interpolate polarization values between the observations, the task is quite trivial. Any of the three models introduced here, TRIM, sub-TRIM or LEM, can be used in a standard (nonlinear) regression analysis, or even just a polynomial fit could be adequate. The problem becomes more cumbersome in the cases where there is a lack of observations, some extrapolation needs to be done, or some detailed analysis of e.g. the wavelength effect is required. In these complex cases more effort needs to be given to the selection of proper model and suitable regression procedure. In Sec. 2 I compared the three PPC models and found severe problems in LEM and also in TRIM to some extent. The models are ill-conditioned by construction which manifests in high correlations between the parameters for the LEM or problems in parameter identification for the TRIM. The sub-TRIM is the

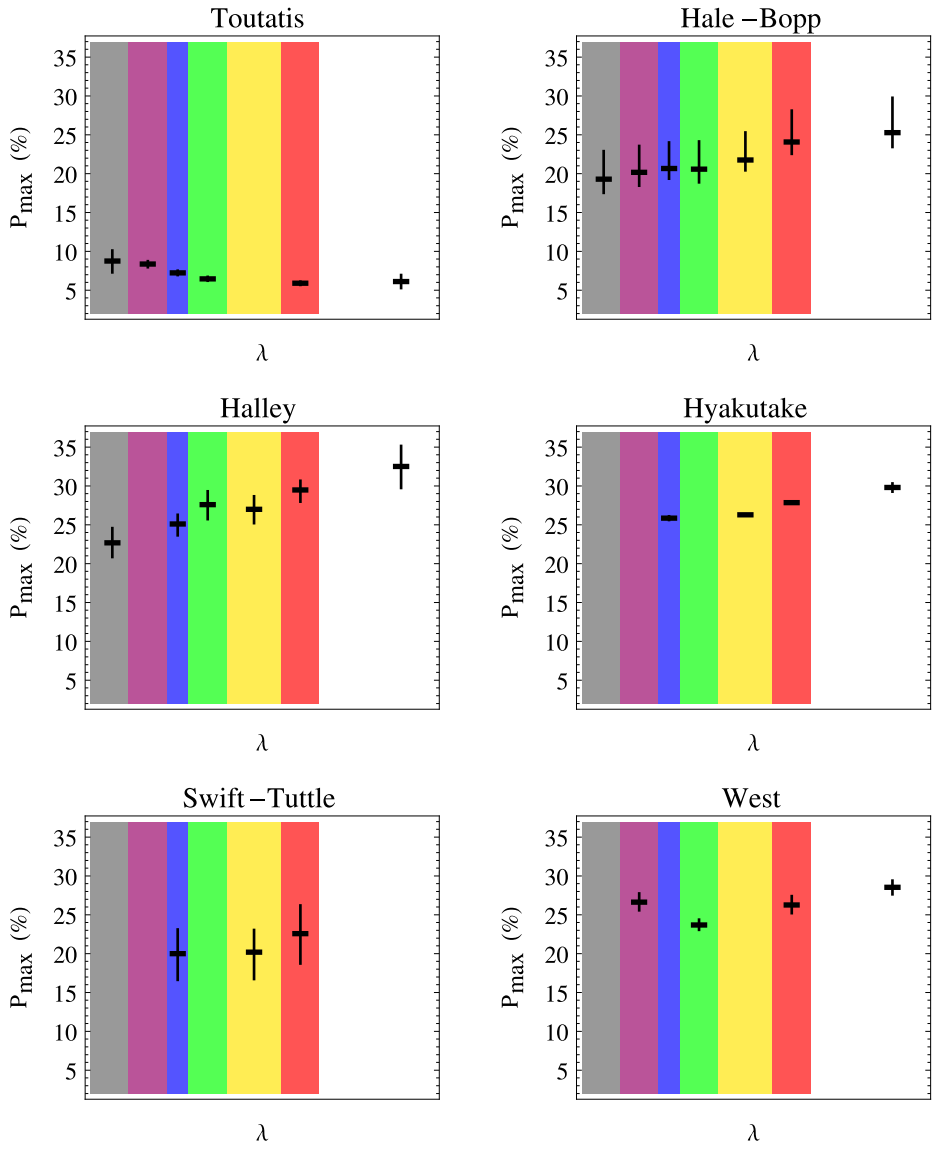

Figure 4.6: Wavelength effects for all objects in maximum polarization $P_{\max}$. The vertical lines are the 90% confidence intervals, and the horizontal bars are the point-estimates. The y-scale is the wavelength with background color showing either the visible spectra color, or UV with gray and IR with white.

most stable model and recommended to be used in regular (nonlinear) regression analysis. The robustness and physically realistic structure of the model are the advantageous features of the sub-TRIM.

If the more flexible form of the full TRIM model is needed, it is mandatory in many cases to guide the regression procedure to overcome the parameter identification issues. This can be done in the Bayesian regression framework, as shown

in Sec. 3. Informative a priori distributions are a necessity. The result of the Bayesian regression is robust to be used in e.g. extrapolation and the MCMC sampling scheme offers a straightforward sample-based method to empirically derive estimates together with error estimates to virtually any property of the PPC.

The most sophisticated methods must be used if we want to gain from the correlation between the observations at different wavelengths. The multiple response method could be suitable in observation campaigns where practically simultaneous observations at multiple wavelengths can be made, as discussed in Sec. 4.1. If we want to study the existing data catalogues a more suitable method is the multiwavelength Bayesian model developed in Sec. 4.2. This model enables e.g. the analysis of the possible wavelength effect in the polarization. The conclusion about the WE is, that within the wavelength range of our observations, from UV to near-IR, there are only a few cases where the hypothesis of the WE is strongly enough supported by the data. These include the WE in comet Hale-Bopp observations and the WE in $P_{\max}$ for comets and the asteroid Toutatis. The possibility of the WE is not debarred in other cases, but more supporting data should be observed.

Bibliography

Bates, D. & Watts, D. (1988), *Nonlinear regression analysis and its applications*, Wiley.

Box, G. & Tiao, G. (1973), *Bayesian Inference in Statistical Analysis*, Addison-Wesley.

Dones, L., Cuzzi, J. & Showalter, M. (1993), 'Voyager photometry of Saturn's A ring', *Icarus* **105**, 184–215.

Gehrels, T., ed. (1974), *Planets, Stars & Nebulae studied with Photopolarimetry*, University of Arizona Press.

Gehrels, T. & Teska, T. (1963), 'The wavelength dependence of polarization', *Applied Optics* **2**, 67–77.

Gilks, W., Richardson, S. & Spiegelhalter, D., eds (1995), *Markov Chain Monte Carlo in Practice*, CRC Press.

Gill, P. & Murray, W. (1978), 'Algorithms for the solution of the nonlinear least-squares problem', *SIAM Journal on Numerical Analysis* **15**, 977–992.

Goidet-Devel, B., Renard, J.-B. & Levasseur-Regourd, A.-C. (1995), 'Polarization of asteroids: Synthetic curves and characteristic parameters', *Planet. Space Sci.* **43**, 779–786.

Hadamcik, E. & Levasseur-Regourd, A.-C. (2003*a*), 'Dust coma of comet C/1999 S4 (LINEAR): Imaging polarimetry during the nucleus disruption', *Icarus* **166**, 188–194.

Hadamcik, E. & Levasseur-Regourd, A.-C. (2003*b*), 'Imaging polarimetry of cometary dust: Different comets and phase angles', *J. of Quant. Spec. and Rad. Transf.* **79–80**, 661–678.

Hadamcik, E., Renard, J.-B., Levasseur-Regourd, A.-C. & Worms, J.-C. (2003), 'Laboratory light scattering measurements on 'natural' particles with the PROGRA^2 experiment: An overview of the results', *J. of Quant. Spec. and Rad. Transf.* **79–80**, 679–693.

Hapke, B. (1993), *Theory of reflectance and emittance spectroscopy*, Cambridge University Press.

Jockers, K. (1997), 'Observations of scattered light from cometary dust and their interpretation', *Earth, Moon, and Planets* **79**, 221–245.

Kaasalainen, S., Muinonen, K. & Piironen, J. (2001), 'Comparative study on opposition effect of icy solar system objects', *J. of Quant. Spec. and Rad. Transf.* **70**, 529–543.

Kaasalainen, S., Piironen, J., Kaasalainen, M., Harris, A., Muinonen, K. & Cellino, A. (2003), 'Asteroid photometric and polarimetric phase curves: Empirical interpretation', *Icarus* **161**, 34–46.

Kiselev, N., Jockers, K. & Rosenbush, V. (2002), 'Comparative study of the dust polarimetric properties in split and normal comets', *Earth, Moon, and Planets* **90**, 167–176.

Kiselev, N., Rosenbush, V. & Jockers, K. (1999), 'NOTE: Polarimetry of Asteroid 2100 Ra-Shalom at Large Phase Angle', *Icarus* **140**, 464–466.

Kiselev, N., Rosenbush, V., Kolokolova, L. & Antonyuk, K. (2008), 'The anomalous spectral dependence of polarization in comets', *J. of Quant. Spec. and Rad. Transf.* **109**, 1384–1391.

Levasseur-Regourd, A.-C., Hadamcik, E. & Renard, J.-B. (1996), 'Evidence for two classes of comets from their polarimetric properties at large phase angles', *Astronomy & Astrophysics* **313**, 327–333.

Lumme, K., Muinonen, K., Green, K. & Penttilä, A. (2003), Light scattering by large densely-packed aggregates: A wave-optical approach, *in* 'Proceedings of the 7th Conference on Electromagnetic and Light Scattering by Nonspherical Particles: Theory, Measurements and Applications', Bremen, Germany, pp. 220–223.

Lunn, D., Thomas, A., Best, N. & Spiegelhalter, D. (2000), 'WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility', *Statistics and Computing* **10**, 325–337.

Lyot, B. (1929), 'Recherches sur la polarisation de la lumière des planètes et de quelques substances terrestres', *Ann. Obs. Meudon* **VIII**.

Mishchenko, M. (1993), 'On the nature of the polarization opposition effect exhibited by Saturn's rings', *Astrophysical Journal* **411**, 351–361.

Mishchenko, M. (2008), 'Multiple scattering, radiative transfer, and weak localization in discrete random media: Unified microphysical approach', *Reviews of Geophysics* **46**, RG2003.

Mishchenko, M., Travis, L. & Lacis, A. (2006), *Multiple scattering of light by particles. Radiative transfer and coherent backscattering*, Cambridge university press.

Muinonen, K. (2004), 'Coherent backscattering of light by complex random media of spherical scatterers: numerical solution', *Waves in Random Media* **14**, 365–388.

Muinonen, K., Piironen, J., Kaasalainen, S. & Cellino, A. (2002), 'Asteroid photometric and polarimetric phase curves: Empirical modeling', *Memorie della Societa Astronomica Italiana* **73**, 716–721.

Muinonen, K., Piironen, J., Shkuratov, Y., Ovcharenko, A. & Clark, B. (2002), Asteroid photometric and polarimetric phase effects, *in* W. Bottke, A. Cellino, P. Paolicchi & R. Binzel, eds, 'Asteroids III', University of Arizona Press, pp. 123–138.

Mukai, T., Iwata, T., Kikuchi, S., Hirata, R., Matsumura, M., Nakamura, Y., Narusawa, S.-Y., Okazaki, A., Seki, M. & Hayashi, K. (1997), 'Polarimetric observations of 4179 Toutatis in 1992/1993', *Icarus* **127**, 452–460.

Penttilä, A., Lumme, K., Hadamcik, E. & Levasseur-Regourd, A.-C. (2005), 'Statistical analysis of asteroidal and cometary polarization phase curves', *Astronomy & Astrophysics* **432**, 1081–1090.

Piironen, J., Muinonen, K., Keränen, S., Karttunen, H. & Peltoniemi, J. (2000), Backscattering of light by snow: field measurements, *in* M. Verstraete, M. Menenti & J. Peltoniemi, eds, 'Observing land from space: Science, customers and technology', Vol. 4 of *Advances in Global Change Research*, Springer, pp. 219–228.

Robert, C. & Casella, G. (1999), *Monte Carlo statistical methods*, Springer, New York.

Rosenbush, V. (2005), 'Opposition effects in brightness, color, and polarization of comet 1P/Halley: Comparison with atmosphereless Solar System bodies', *Solar System research* **39**, 312–321.

Saikkonen, P. (2004), Epälineaarinen regressioanalyysi. Lecture notes (in Finnish) for the course given at the University of Helsinki, Finland.

Seber, G. & Wild, G. (1989), *Nonlinear regression*, Wiley.

Seeliger, H. (1887), 'Zur theorie der beleuchtung der grossen planeten, insbesondere des Saturn', *Abh. Bayer. Akad. Wiss. Math. Naturwiss. Kl. II* **16**, 405–516.

Shkuratov, Y., Ovcharenko, A., Zubko, E., Miloslavskaya, O., Muinonen, K., Piironen, J., Nelson, R., Smythe, W., Rosenbush, V. & Helfenstein, P. (2004), 'The opposition effect and negative polarization of structural analogs for planetary regoliths', *Icarus* **159**, 396–416.

Tholen, D. & Barucci, A. (1989), Asteroid taxonomy, *in* R. Binzel, T. Gehrels & M. Matthews, eds, 'Asteroids II', The University of Arizona Press, pp. 298–315.

Tyynelä, J., Muinonen, K., Zubko, E. & Videen, G. (2008), 'Interrelating scattering characteristics to internal electric fields for gaussian-random-sphere particles', *J. of Quant. Spec. and Rad. Transf.* **109**, 2207–2218.

van de Hulst, H. (1957), *Light scattering by small particles*, Wiley, New York.