Mass and radius constraints for neutron stars from pulse shape modeling

Master Thesis
University of Turku
Department of Physics and Astronomy
Astronomy
2016
Tuomo Salmi
Supervised by:
   J.P.
   J.N.

The originality of this thesis has been checked in accordance with the University of Turku

quality assurance system using the Turnitin Originality check

University of Turku
Department of Physics and Astronomy

Salmi, Tuomo Mass and radius constraints for neutron stars from pulse shape modeling

Master's thesis, 54 pages
Astronomy
November 2016

Neutron stars are the most compact directly observable objects. The matter inside a neutron star is at supranuclear densities. Thus the equation of state (EOS) of neutron stars describes the properties of such dense matter. Separation between numerous theoretical EOSs is possible if we are able to constrain the possible masses and radii of neutron stars from observations.

In this thesis we present one method that can be used to constrain masses and radii of neutron stars. The method is suitable for accreting millisecond pulsars, where a rapidly rotating neutron star, called a millisecond pulsar, accretes matter from a relatively low mass companion star, via an accretion disk, onto the magnetic poles of the neutron star. Because of the accretion, we observe radiation from two "hot spots" on the neutron star surface. This radiation is pulsating coherently at the spinning frequency of the neutron star. The exact shape of the pulses can be modeled with a theoretical model that takes into account the general and special relativistic effects via "Schwarzschild-Doppler" approximation and the oblate shape of the star caused by the fast rotation. The details of this model are discussed.

The pulse profiles carry information about the mass and radius of a neutron star since e.g., the light bending and thus pulse shape depends much on the compactness of the star. Also many other physical parameters and observing angles affect the light curves. Therefore, we use Bayesian analysis and a novel Monte Carlo sampling method, called "ensemble sampler", to obtain probability distributions for the different parameters, especially for the mass and the radius. The ensemble sampler has shown to overcome many difficulties concerning the traditional Metropolis-Hastings sampler.

We have also generated synthetic data to test our method and fitted the pulse profiles to these data. The results of our samplings, using these synthetic data, show that obtaining new constraints for radius and mass is not a very easy task. However, according to our study, prior information obtained from polarization measurements may be used the get significantly tighter constraints.

keywords : Neutron stars, Accreting millisecond pulsars, Pulse profile modeling, Bayesian analysis, Ensemble sampler, Mass and radius constraints

Acknowledgements

First of all, I would like to thank my supervisors Professor Juri Poutanen and Joonas Nättilä for a very interesting and challenging topic. The support I got was extremely helpful, especially the theoretical advices from Juri, and the tools of analysis and all the other support from Joonas. I would also like to thank all the people in the astrophysics research group for the interesting discussions and support in general. Finally, I wish to thank my family and friends for support and encouragement throughout my studies.

Contents

Introduction

Rapidly rotating neutron stars, called pulsars, are the lighthouses of the universe. Pulses can be observed when an electromagnetic beam from a neutron star is emitted towards the Earth. In some cases this emission may be created when matter is accreted onto the magnetic poles at the surface of the pulsar. Light curves from these pulses can be detected in many wavelengths. For example, the most rapidly rotating accreting pulsars, called accreting millisecond pulsars, have been detected in the X-ray band of the electromagnetic spectrum.

The exact shape of the pulses may reveal us important information about the properties of the neutron stars. The determination of the mass-radius relation of neutron stars through observations is one of the fundamental problems in astrophysics. This information could provide tight constraints on the equation of state (EOS) of the ultra-dense matter located inside neutron stars [1] [2]. Densities of this high are otherwise unattainable. Studies of the light curves of pulsars can therefore help to determine the properties of such matter. The properties of matter at extremely high densities are also among the most important questions in physics and astronomy.

One way to constrain masses and radii is to use X-ray oscillations produced by the rotation of accreting millisecond pulsars. Models of these oscillations can be compared to observed waveforms. The mass and radius of neutron star have an effect on the waveform because of the influence e.g., on the light bending due to general relativity. However, many other parameters have also an impact on the light curves making it challenging to get more strict constraints to radius and mass. Markov chain Monte Carlo sampling and high-performance computing are necessities when trying to find the correct values for these parameters.

1   Neutron stars

Neutron stars are some of the most dense and massive objects in the Universe. Typical mass M of a neutron star is on the order of 1.5 solar masses ($M_\odot$), and typical radius R on the order of 12 km. The central density n_c can be from 5 to 10 times the nuclear equilibrium density $n_0 \approx 0.16$ fm^{-3} of neutrons and protons found in laboratory nuclei [3]. Neutrons dominate the nucleonic content of neutron stars, but also some protons, electrons and muons exist. At the supernuclear densities more exotic baryons, mesons or quarks may also appear. The composition of the innermost core of the neutron star is still unknown. Superfluidity and/or superconductivity of the matter is in any case expected inside the star.

Neutron stars are created after the gravitational collapse of a core of a massive star ($> 8\ M_\odot$) at the end of its life. The collapse also triggers a Type II supernova explosion [3]. Stars more massive than about $25\ M_\odot$ collapse instead into a black hole. The general relativistic Schwarzschild condition

$$R > \frac{2GM}{c^2},$$ (1)

where G is the gravitational constant and c is the speed of light, differentiates neutron stars from black holes. More compact objects (for which $R < 2GM/c^2$) are inside the event horizon meaning that the light cannot escape the object. This constraint is presented in Figure 1 as a blue region.

A more strict upper bound to the compactness (ratio between the mass and radius) of a neutron star follows the fact [4] that the speed of sound in dense matter has to be less than the speed of light. This gives us the so-called causality condition

$$R \gtrsim \frac{3GM}{c^2}.$$ (2)

This constraint is shown in Figure 1 as a green region. Neutron stars have also a minimum stable mass, which is about $0.1\ M_\odot$, although the origin of neutron star in a supernova gives a more realistic minimum (on the order of $1.1\ M_\odot$) [5].

A lower limit for the compactness of a neutron star can be obtained from its spin frequency. The radius of the star cannot extend further than the point where the spin frequency equals to the Keplerian frequency. Assuming a rigid Newtonian sphere, this frequency can be expressed as

$$\nu_{\mathrm{K}} = 2\pi \sqrt{\frac{GM}{r^3}}, \tag{3}$$

where r is the distance from the center of the star. Thus at $r = R$, ν_{K} has to be higher than the observed spin frequency, and we get a lower limit for M/R^3. This constraint is shown in Figure 1 for the highest measured rotational frequency (641 Hz [3]) as a red region. One should also take into account the non-spherical shape and the general relativistic effects, but a lower limit is still obtained.

Generally the mass-radius ($M - R$) relation is determined by the equations of hydrostatic equilibrium. For a general relativistic and spherically symmetric (non-rotating) object these are the so-called Tolman-Oppenheimer-Volkoff equations [6][7]:

$$\frac{\mathrm{d}\mathcal{P}}{\mathrm{d}r} = -\frac{G[m(r) + 4\pi r^3 \mathcal{P}/c^2](\rho_{\mathrm{m}} + \mathcal{P}/c^2)}{r[r - 2Gm(r)/c^2]}, \tag{4}$$

and

$$\frac{\mathrm{d}m(r)}{\mathrm{d}r} = 4\pi \rho_{\mathrm{m}} r^2, \tag{5}$$

where $\mathcal{P}$ and ρ_{m} are the pressure and mass-energy density, respectively, and $m(r)$ is the gravitational mass enclosed within a radius r. The $M - R$ relation can now be obtained if the relation between the pressure and the density $\mathcal{P} = \mathcal{P}(\rho_{\mathrm{m}})$ is known. This relation is called the EOS.

For realistic EOSs the previous equations must be numerically solved to obtain $M - R$ relations. These can be separated into three categories according to the compressibility of the matter: soft, moderate, and stiff equations of state. The determination of the EOS would thus allow us to find out the structure of a neutron star and the properties of the nucleonic matter inside it [8]. Even the exteriors of neutron stars may be either normal hadronic matter or a so-called strange quark matter (SQM) depending on the EOS.
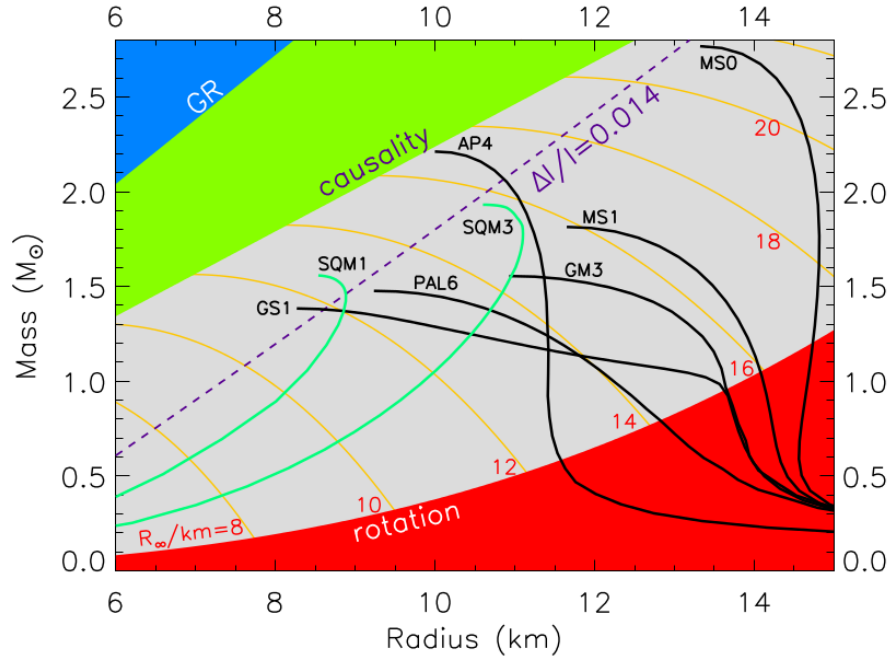
Figure 1. Mass-radius diagram for neutron stars [3]. Black curves are for normal matter and green curves for SQM equations of state. Regions excluded by general relativity (GR), causality, and rotation constraints are indicated. The rotation constraint is obtained from the highest measured spin frequency (641 Hz). The orange curves show the contours of constant radiation radii R_∞. The dashed line $\Delta I/I = 0.014$ is a limit estimated from the glitches of Vela pulsar [9]. From Lattimer and Prakash (2004) [3].

The number of possible EOSs can be reduced by observing which masses and radii the EOS should be able to produce. Different EOSs are shown in Figure 1. The highest measured mass of a neutron star is close to two solar masses [10] which rules out all the EOSs that cannot produce a mass so large. In Figure 1 this leaves only two possible EOSs, but there are also many EOSs which are not shown. The orange curves in Figure 1 show the contours of constant radiation radius R_∞, which is the apparent radius of the object at "infinity". It is larger than R, since the light bending makes the star appear larger, and is given by [9]

$$R_\infty = \frac{R}{\sqrt{1 - 2GM/Rc^2}}. \tag{6}$$

Knowing the relation between the mass and radius of any particular neutron star gives also limits to the possible EOSs. For example, suddenly increased spin rates of pulsars ("glitches") have been used to get mass-radius relations [9]. It has been a strong indication that the star's moment of inertia I acts as an angular momentum reservoir for the sudden spin-ups [11]. The measured $\Delta I/I = 0.014$ for Vela pulsar gives the mass as a function of radius, and it is also shown as a dashed line in Figure 1. The correct EOS of neutron star has to cross this line. In our work we aim to find out independent knowledge of both mass and radius of observed pulsars using pulse profile modeling.

1.1 Accreting millisecond pulsars

Rotating neutron stars were first detected as radio pulsars in 1967 [12]. After that, many different classes of pulsars have been discovered including those situated in low mass X-ray binaries (LMXBs). LMXBs are systems in which the neutron star accretes matter from a non-collapsed companion (with a relatively low mass $M \lesssim M_\odot$) via an accretion disk. Accreting millisecond X-ray pulsars (AMXPs) are a subgroup of the LMXBs in which the gas from the accretion disk (stripped from the companion) is channeled onto the magnetic poles of a rapidly rotating neutron star. The result is a pair of "hot spots" on the pulsar surface. This gives rise to the X-ray pulsations with typical periods of a
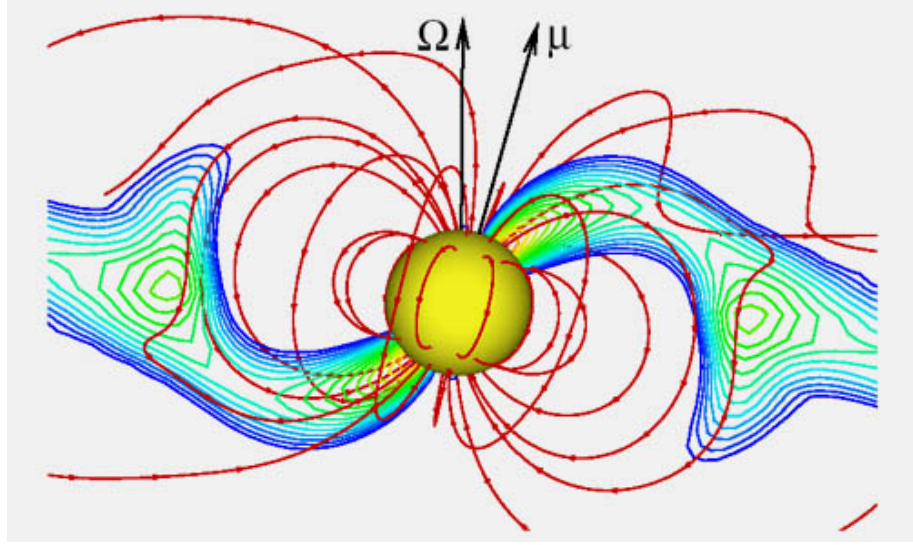
Figure 2. A view from a magnetohydrodynamical simulation for an accreting compact object with an inclined dipole field. The density of the streaming matter changes exponentially from blue to red. Red lines with arrows show selected magnetic field lines. The magnetic moment, $\boldsymbol{\mu}$, and the rotation axis, $\boldsymbol{\Omega}$, are also shown. From Romanova *et al.* (2004) [14].

few milliseconds corresponding to the spin frequency of the neutron star. By definition, AMXPs are spinning at frequencies $\nu \geq 100$ Hz and have weak surface magnetic fields ($B \sim 10^{8-9}$ G) [13]. A schematic view of the accretion onto the hot spots of the pulsar is shown in Figure 2.

The transfer of mass occurs in all known AMXPs via a Roche lobe overflow from the donor star. The equipotential surface of the binary systems gravitational and centrifugal forces (surrounding the star) is known as a Roche lobe. Both the neutron star and the companion have their own lobes that only join at the inner Lagrangian point L_1 [15]. In case when the Roche lobe of the companion star is full of matter, the gas left outside the companion's Roche Lobe starts to flow towards neutron star through the L_1 point. An accretion disk is formed because of the rotation of the system of the two stars.

The first AMXP discovered was SAX J$1808.4 - 3658$. It is also used as a reference target in this thesis. The source was first found in 1996 by the Italian-Dutch *BeppoSAX* satellite [16]. The first coherent pulsations (at 401 Hz) were detected during the second outburst with NASA's *Rossi X-ray Timing Explorer* (*RXTE*) in 1998 [17]. It provided a
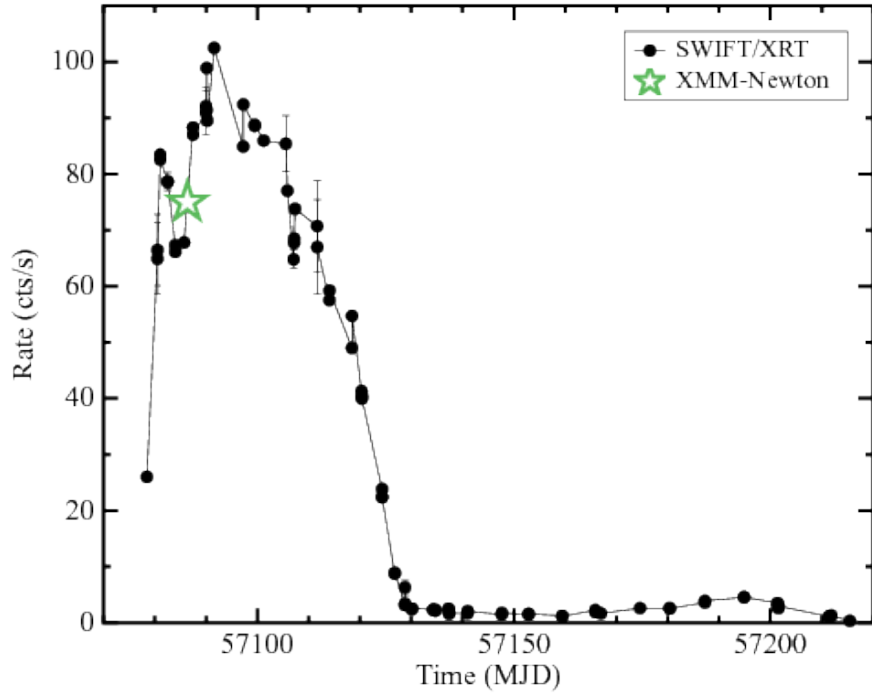
Figure 3. Light curve of the 2015 outburst of SAX J1748.9 − 2021 as observed by *Swift-XRT* (black points) and *XMM-Newton* (the green star). From Sanna *et al*. (2016) [19].

confirmation of the "recycling scenario" which states that AMXPs are the evolutionary progenitors of recycled radio millisecond pulsars. They are responsible for the conversion of slowly rotating neutron stars with high magnetic field ($B \sim 10^{12}$ G), into a rapidly spinning objects with a relatively weak magnetic field ($B \sim 10^8$ G) [13]. The idea is that a weakening magnetic field allows accretion to happen. The angular momentum of the accreting matter is transferred to the pulsar resulting in the increase of its spin frequency.

After the first two outbursts, SAX J1808.4 − 3658 has gone into outburst several times reoccurring approximately every two to three years. It is the best sampled and studied object out of all AMXPs. Typically the outburst consists of five phases: a fast rise, with a steep increase in luminosity lasting only a few days, a peak, a slow decay stage, a fast decay phase, and the flaring tail. Except the last phase, the aforementioned outburst division is typical for several X-ray binaries and can in principle be partially explained with a so-called instability model [18] (see Figure 3 for a such outburst).

The X-ray spectrum of the outbursts have also been analysed, and there is evidence for

a two component-model including a blackbody at lower energies and a hard Comptonization component at higher energies [20] . The heated hot spot on the neutron star surface is interpreted to produce the blackbody flux. The Comptonization is produced above the spot in an accretion shock. This shock is created as the plasma abruptly decelerates close to the neutron star surface.

Figure 4. Normalized photon number fluxes at different energies as a function of the phase from SAX J$1808.4 - 3658$ measured with *RXTE*.

2 Methods

There have been many different attempts to constrain the masses and radii of neutron stars and thus also the possible equations of state. Our approach is to fit a pulse profile model to the observed X-ray oscillations of accretion-powered millisecond pulsars. In this section we first present our waveform model and then the Bayesian methods which we use to get constraints for the parameters of the model.

2.1 Pulse profile modeling

The coherent flux oscillations at the rotational frequency (the light curves) of accreting millisecond pulsars (described in the first chapter) can be modeled with different models. One example of the observed pulse profiles (of SAX J$1808.4 - 3658$) is shown in Figure 4. The model presented here assumes that the radiation is originated from one or two of the polar caps of the star. General and special relativistic effects have been taken into account using a so-called Schwarzschild-Doppler (S+D) approximation [21][22]. In the S+D ap-

proximation the effects of general relativity (gravitational light bending) are modeled as though the star is not rotating using the Schwarzschild metric and the formalism specified by Pechenick, Ftaclas, and Cohen (1983) [23]. Rotational effects have been approximated by using special relativistic Doppler transformations as though the star is a rotating object with no gravitational field. The oblate shapes of the rapidly rotating neutron stars have also been taken into account using an empirical formula for the oblate shape [24] [25].

The observed pulse profiles from AMXPs appear to be rather close to sinusoidal with peak-to-peak oscillation amplitude

$$A = \frac{F_{\max} - F_{\min}}{F_{\max} + F_{\min}} \tag{7}$$

between 4 and 12 per cent [26], where $F_{\max}$ is the maximum observed flux and $F_{\min}$ is the minimum observed flux. Light bending tends to reduce this variability amplitude. The amplitude depends only weakly on energy, and the spectrum may be fitted with a two-component model (blackbody and Comptonization). At higher energies the amplitude deviates more from a sine wave because of the different emission pattern. The weak energy dependency and a fairly constant spectral shape as a function of pulse phase, support also the idea that the bulk of the observed X-ray emission originates from the polar caps (where the gas stream channeled by the neutron star's magnetic field impacts the stellar surface forming a shock). Any additional source of radiation would have to have a spectrum identical to that of the shock.

2.1.1 Oblateness

Due to the fast rotation, the millisecond pulsars have an oblate shape instead of spherical. The difference between an oblate and a spherical star is significant when the rotation frequency $\nu \geq 300$ Hz [27]. The most important effect is purely geometrical: The directions that the light can be emitted into are different in the cases of an oblate and a spherical star. Thus there are certain spot locations on the star where the spot is invisible if the surface is oblate but would be visible if the surface were purely spherical (and vice

versa).

There are different models describing the exact shape and oblateness of the neutron star (function $R(\theta)$, where θ is the colatitude measured from the spin axis). One of the most recent ones was presented by Algendy *et al.* (2014) [24]. We use primarily it in this thesis. In this model

$$\frac{R(\theta)}{R_{\text{eq}}} = (1 + o_2(x, \bar{\Omega}) \cos^2(\theta)), \tag{8}$$

where

$$o_2(x, \bar{\Omega}) = \bar{\Omega}^2(-0.788 + 1.030x), \tag{9}$$

$$x = \frac{GM}{c^2 R_{\text{eq}}}, \tag{10}$$

and

$$\bar{\Omega} = \Omega \left(\frac{R_{\text{eq}}^3}{GM}\right)^{1/2}. \tag{11}$$

In these equations R_{eq} is the radius of the rotating star measured at the equator and $\Omega = 2\pi/P$, where P is the spin period.

As another model we have also used a model presented by Morsink *et al.* (2007) [25]. In that model

$$\frac{R(\theta)}{R_{\text{eq}}} = 1 + \sum_{n=0}^{2} a_{2n}(\bar{\zeta}, \bar{\epsilon}) P_{2n}(\cos(\theta)), \tag{12}$$

where $P_n(\cos(\theta))$ is the Legendre polynomial of order n and parameters $\bar{\zeta}$ and $\bar{\epsilon}$ are obtained from

$$\bar{\zeta} = \frac{GM}{R_{\text{eq}} c^2} \tag{13}$$

and

$$\bar{\epsilon} = \frac{\Omega^2 R_{\text{eq}}^3}{GM}. \tag{14}$$

The coefficients needed are given as

$$a_0 = -0.18\bar{\epsilon} + 0.23\bar{\zeta}\bar{\epsilon} - 0.05\bar{\epsilon}^2, \tag{15}$$

$$a_2 = -0.39\bar{\epsilon} + 0.29\bar{\zeta}\bar{\epsilon} + 0.13\bar{\epsilon}^2, \tag{16}$$

and

$$a_4 = +0.04\bar{\epsilon} - 0.15\bar{\zeta}\bar{\epsilon} + 0.07\bar{\epsilon}^2. \tag{17}$$

2.1.2 Geometry

Besides the geometry of the star itself, we also need to know the relations between different angles in different frames. To derive these we consider a small spot on the stellar surface at colatitude θ_S from the rotational pole (we follow closely the derivation of Poutanen and Beloborodov (2006) [28]). The star is assumed to be rotating with a frequency $\nu = P^{-1}$ (seen by a distant observer). The velocity of the spot in units of c (as measured by a non-rotating observer near the star) is

$$\beta = \frac{v}{c} = \frac{2\pi R(\theta_S)}{c} \frac{\nu}{\sqrt{1-u}} \sin \theta_S, \tag{18}$$

where $u \equiv r_{\rm S}/R(\theta_S)$, $r_{\rm S} = 2GM/c^2$ is the Schwarzschild radius, M is mass and $R(\theta_S)$ is radius of the star at spot location (given by equation 8 or 12). The pulsar frequency has been corrected for the gravitational redshift $1/\sqrt{1-u} = 1 + z$ since the rotation frequency seen by distant observer is reduced in the gravitational field of the star due to the gravitational time dilation. The corresponding Lorentz factor for the velocity of the spot is $\Gamma = (1 - \beta^2)^{-1/2}$.

Because of the fast rotation and special relativistic effects, spot area $\mathrm{d}S'$ measured in a corotating frame is not same as the area $\mathrm{d}S$ measured in a non-rotating frame. The instantaneous position of the spot in the fixed lab frame is described by the unit vector

$$\boldsymbol{r} = (\sin \theta_S \cos \phi, \sin \theta_S \sin \phi, \cos \theta_S), \tag{19}$$

that points to the spot from the center of the star (see Figure 5). The rotational phase of the pulsar is $\phi = 2\pi\nu t$, where t denotes time. The vector $\boldsymbol{r}$ is not generally directed perpendicular to the surface of the spot unless the star is spherical. The vector that points normal to the surface is given by

$$\boldsymbol{n} = (\sin(\theta_S - \gamma) \cos \phi, \sin(\theta_S - \gamma) \sin \phi, \cos(\theta_S - \gamma)). \tag{20}$$

The angle between $\boldsymbol{n}$ and $\boldsymbol{r}$ is γ and

$$\cos \gamma = [1 + f^2(\theta_S)]^{-1/2}, \tag{21}$$
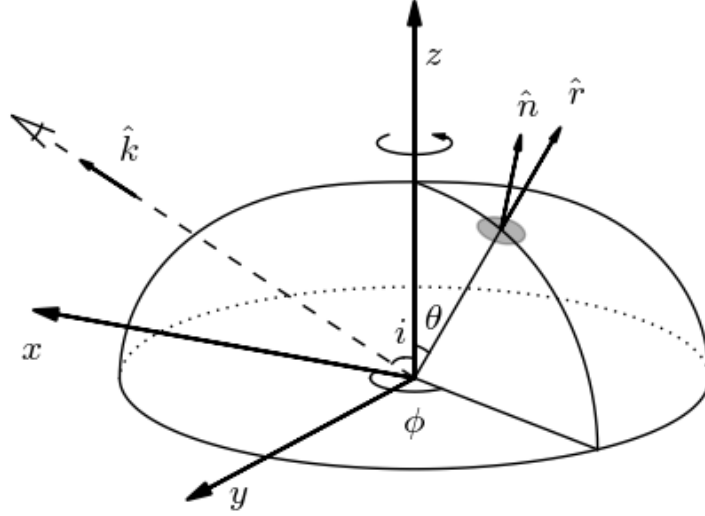
Figure 5. Geometry of the problem. Created with https://github.com/natj/nsfig.

where

$$f(\theta) = \frac{1+z}{R}\frac{\mathrm{d}R}{\mathrm{d}\theta}. \tag{22}$$

In case of a spherical star $\mathrm{d}R/\mathrm{d}\theta = 0$ and thus $\boldsymbol{n} = \boldsymbol{r}$. A normal oblate spheroid would have $f(\theta) = (\mathrm{d}R/\mathrm{d}\theta)/R$, but the Schwarzschild metric introduces an additional conformal factor $(1 + z)$ into the equation.

The computation of the waveform is started from the center of the spot in the plane defined by the spin axis and the direction to the observer (when the spot crosses the plane defined by $\boldsymbol{z}$ and $\boldsymbol{k}$). The light pulse originating at this moment at the point that is directly below the observer and with a reference distance r_{ref} to the center of the star (e.g., the equatorial radius of the star) is used to define the zero of the observer's time coordinate. Measuring of the time t is started when this photon arrives at the observer (hence fixing $t = 0$).

We denote the unit vector along the line of sight by

$$\boldsymbol{k} = (\sin i, 0, \cos i), \tag{23}$$

where i is the inclination angle of the spin axis to the line of sight. Thus (see Figure 5)

$$\cos \psi = \boldsymbol{k} \cdot \boldsymbol{r} = \cos i \ \cos \theta_S + \sin i \ \sin \theta_S \ \cos \phi. \tag{24}$$

The bending angle ψ measures the angle between the line of sight and the position vector $\boldsymbol{r}$ of the spot. The initial angle of the emitted photon with respect to $\boldsymbol{r}$ differs from ψ because of the light bending.

We denote the initial direction of the emitted photon by $\boldsymbol{k}_0$ and the true emission angle (relative to $\boldsymbol{r}$) by α so that

$$\cos \alpha = \boldsymbol{k}_0 \cdot \boldsymbol{r}. \tag{25}$$

The zenith angle σ between the $\boldsymbol{n}$ and $\boldsymbol{k_0}$ is defined by

$$\cos \sigma = \boldsymbol{k}_0 \cdot \boldsymbol{n}. \tag{26}$$

The direction of the photon changes from $\boldsymbol{k}_0$ near the stellar surface to $\boldsymbol{k}$ at infinity as it propagates to infinity so that $\cos \alpha = \boldsymbol{k}_0 \cdot \boldsymbol{r}$ changes to $\cos \psi = \boldsymbol{k} \cdot \boldsymbol{r}$. The relation between $\boldsymbol{k}_0$ and $\boldsymbol{k}$ may be written as

$$\boldsymbol{k}_0 = [\sin \alpha \ \boldsymbol{k} + \sin(\psi - \alpha) \ \boldsymbol{r}] / \sin \psi. \tag{27}$$

We would like to express this vector in the rest frame of the spot instead of the non-rotating static frame. As a first step we do the coordinate transformation to an instantaneous non-rotating frame x, y, z with the y-axis along the direction of the spot motion, x-axis along the meridian towards the equator, and z-axis along the normal of the spot. This frame can be introduced in any moment of time. In this static frame

$$\boldsymbol{k}_0 = \left(\cos \epsilon, \cos \xi, \cos \sigma \right), \tag{28}$$

where ξ is the angle between the spot velocity and the direction of the emitted photon $\boldsymbol{k}_0$. It is given as

$$\cos \xi = \frac{\boldsymbol{\beta}}{\beta} \cdot \boldsymbol{k}_0 = \frac{\sin \alpha}{\sin \psi} \frac{\boldsymbol{\beta}}{\beta} \cdot \boldsymbol{k} = -\frac{\sin \alpha}{\sin \psi} \sin i \ \sin \phi \,, \tag{29}$$

since $\boldsymbol{\beta} = \beta(-\sin\phi, \cos\phi, 0)$ in the lab frame. Angle between the meridian

$$\boldsymbol{m} = (\cos(\theta_S - \gamma)\cos\phi, \cos(\theta_S - \gamma)\sin\phi, -\sin(\theta_S - \gamma)) \tag{30}$$

and $\boldsymbol{k}_0$ is ϵ which is obtained from

$$\cos\epsilon = \boldsymbol{m} \cdot \boldsymbol{k}_0 = \frac{\sin\alpha}{\sin\psi}(\sin i \cos(\theta_S - \gamma)\cos\phi - \cos i \sin(\theta_S - \gamma)) + \frac{\sin(\psi - \alpha)}{\sin\psi}\sin\gamma. \tag{31}$$

Next we can consider a frame which have the same axes as previously but which is also corotating with the spot. We want to express the direction of the emitted light also in this frame. Emission angle relative to the surface normal in this frame is denoted by σ'. It differs from σ because of relativistic aberration. The rays of light are tilted towards the direction of the spot motion relative to the observer. In this frame comoving with the spot (with y-axis again along the spot motion and z-axis along the local normal), the unit vector along the photon momentum is obtained from the Lorentz transformation:

$$\boldsymbol{k}_0' = \delta \begin{pmatrix} \cos\epsilon \\ \Gamma(\cos\xi - \beta) \\ \cos\sigma \end{pmatrix}, \tag{32}$$

where δ is the Doppler factor formulated as

$$\delta = \frac{1}{\Gamma(1 - \beta\cos\xi)}. \tag{33}$$

Using equation (32), we obtain

$$\cos\sigma' = \delta \ \cos\sigma, \tag{34}$$

Making use of equation (27) the zenith angle has the value

$$\cos\sigma = \boldsymbol{k}_0 \cdot \boldsymbol{n} = \cos\alpha\cos\gamma + \sin\alpha\sin\gamma\cos\zeta =$$
$$\cos\alpha\cos\gamma + \frac{\sin\alpha}{\sin\psi}\sin\gamma(\cos i \sin\theta_S - \sin i \cos\theta_S \cos\phi), \tag{35}$$

where the spherical trigonometric identity $\cos i = \cos\theta_S \cos\psi + \sin\theta_S \sin\psi \cos\zeta$ (where ζ is the angle between the spot-observer and the spot-spin-axis planes) and the equation (24) are used. With small bending angles the approximation $\sin\alpha/\sin\psi \approx \sqrt{(1-u)}$ may be used. This is based on the Beloborodov's approximation [29]

$$1 - \cos\alpha \approx (1 - \cos\psi)(1 - u). \tag{36}$$

Equation (35) is useful since we need to know the emission angle σ relative to the spot normal in order to calculate fluxes and visibility conditions.

2.1.3 Light bending

Before we can get the angle σ from equation (35) we also first need to know how to get the angle α when we know the light bending angle ψ. The angles ψ (for every point at every phase) we know directly from the geometry since we can calculate them using equation (24). In case of non-infinitesimal spot (which is discussed more in the Section 2.1.6), we of course need a transformation of coordinates between the spot frame and the frame of the star (to get corresponding θ and ϕ for every point inside the spot). The angle ψ is anyway straight determined since we are going to calculate the trajectories for only those photons which are going to arrive at the observer (not in some other direction). The task is now to find out the true emission angle α using general relativity.

The exact relation between α and ψ (when $\alpha < \pi/2$) in Schwarzschild geometry (i.e., light bending) is given by [30]

$$\psi_{\mathrm{p}}(R,\alpha) = \int_R^\infty \frac{dr}{r^2} \left[\frac{1}{b^2} - \frac{1}{r^2}\left(1 - \frac{r_{\mathrm{S}}}{r}\right) \right]^{-1/2}, \tag{37}$$

where b is the impact parameter,

$$b = \frac{R}{\sqrt{1-u}}\sin\alpha. \tag{38}$$

The relation between α and ψ, when $\alpha > \pi/2$, is

$$\psi(R,\alpha) = 2\psi_{\max} - \psi_{\mathrm{p}}(R, \pi - \alpha), \tag{39}$$

where $\psi_{\mathrm{max}} = \psi_{\mathrm{p}}(\mathsf{p}, \alpha = \pi/2)$ and p is the distance of closest approach, given by

$$\mathsf{p} = -\frac{2}{\sqrt{3}} b \cos([\arccos(3\sqrt{3}r_{\mathrm{S}}/(2b)) + 2\pi]/3). \tag{40}$$

Numerical calculations using directly integral (37) lead to large errors, but after making a variable substitution $x = \sqrt{1 - R/r}$ the integral can be computed without problems using for example Simpson quadratures. Even though it is not possible to obtain α as an exact function of ψ, we can still use the previous equations to form a two-dimensional grid of ψ corresponding to different α and R. Then we find the correct α at different phases of the rotation of the star by quadratic interpolation.

The maximum bending angle corresponds to $\sigma = \pi/2$. Otherwise the photon would be directed across the star surface. The condition $\cos\sigma > 0$ is used to define the visibility of the spot. In addition one should also check that the photon won't hit the surface of the star in any later phase of its trajectory (it might happen for an oblate star). This check is not yet included in the current pulse profile code.

2.1.4 Observed flux

Next we can consider the flux originating from our small spot. The observed flux from the spot at photon energy E is

$$\mathrm{d}F_E = I_E \, \mathrm{d}\Omega, \tag{41}$$

where I_E is the specific intensity of radiation at infinity and $\mathrm{d}\Omega$ is the solid angle covered by the spot (with area $\mathrm{d}S'$ in the corotating frame) on the observer's sky. The solid angle can be expressed in terms of the impact parameter [23]

$$\mathrm{d}\Omega = b \, \mathrm{d}b \, \mathrm{d}\varphi/D^2, \tag{42}$$

where D is the distance to the source and φ is the azimuthal angle corresponding to rotation around line of sight (vector $\boldsymbol{k}$). The impact parameter b depends only on the bending angle ψ, but not on φ.

The apparent area of the infinitesimal spot measured by the observer in the non-rotating frame near the stellar surface is $\mathrm{d}S = \delta \, \mathrm{d}S'$ [31][32]. We also know the relation between σ and σ' from the relativistic aberration formula (34) (for motions parallel to the spot surface) $\cos \sigma' = \delta \, \cos \sigma$. Thus

$$\mathrm{d}S \, \cos \sigma = \mathrm{d}S' \, \cos \sigma', \tag{43}$$

which shows that the spot area projected on to the plane perpendicular to the photon direction is Lorentz invariant [32][33]. This area is also called a photon beam cross-section. The infinitesimal spot area measured in the frame comoving with the spot may be calculated as [25]

$$\mathrm{d}S'(\theta) = \Gamma R^2(\theta)[1 + f^2(\theta)]^{1/2} \sin \theta \mathrm{d}\theta \mathrm{d}\varphi, \tag{44}$$

where the factor $[1 + f^2(\theta)]^{1/2}$ takes into account the oblateness of the spot surface (function $f(\theta)$ is given in equation 22). The solid angle measured at a distance D from the star is then [25]

$$\mathrm{d}\Omega = \frac{\mathrm{d}S' \cos \sigma'}{D^2} \frac{1}{1-u} \frac{\mathrm{d}\cos \alpha}{\mathrm{d}\cos \psi}. \tag{45}$$

In case of stars with lower compactness (weak gravity) $r_\mathrm{S} \ll R$, so $u \ll 1$, and this gives the usual formula $\mathrm{d}\Omega = \mathrm{d}S' \cos \sigma'/D^2$.

We have now obtained the solid angle $\mathrm{d}\Omega$ needed to calculate the flux from our spot with equation (41). However, we still need to find out a expression for intensity I_E. The intensity is also different in the static frame of the distant observer than in the local frame comoving with the spot due to both gravitational redshift and Doppler effects. The relation between monochromatic observed and local intensities is [30][34]

$$I_E = \left(\frac{E}{E'}\right)^3 I'_{E'}(\sigma'), \tag{46}$$

where $E/E' = \delta\sqrt{1-u}$. Here $I'_{E'}(\sigma')$ is the intensity computed in the corotating frame. The dependency of redshift $(1 + z = 1/\sqrt{1-u})$ to the power of -3 can be understood by noting that $I_E \propto \nu^3$ and $\nu = \nu'(1 + z)^{-1}$. For the bolometric intensity the relation to

frequency goes as ν^4 and one gets

$$I_{\mathrm{B}} = \left(\delta\sqrt{1-u}\right)^4 I_{\mathrm{B}}'(\sigma'). \tag{47}$$

The observed spectral flux (eq. 41) is now given by

$$\mathrm{d}F_E = (1-u)^{1/2}\delta^4 I_{E'}'(\sigma')\cos\sigma\frac{\mathrm{d}\cos\alpha}{\mathrm{d}\cos\psi}\frac{\mathrm{d}S'}{D^2}, \tag{48}$$

where we have used the aberration formula (34). The bolometric flux reads

$$\mathrm{d}F_{\mathrm{B}} = (1-u)\,\delta^5\,I_{\mathrm{B}}'(\sigma')\cos\sigma\frac{\mathrm{d}\cos\alpha}{\mathrm{d}\cos\psi}\frac{\mathrm{d}S'}{D^2}. \tag{49}$$

The radiation spectrum can be e.g., a blackbody, with effective temperature T_{eff} or a power-law spectrum. We can take for example a power-law $I_{E'}'(\sigma') \propto E'^{-(\lambda-1)}$ which does not depend on the angle σ'. Here λ is the photon spectral index. With this spectrum we get

$$I_{E'}'(\sigma') = I_E'(\sigma')\left(\delta\sqrt{1-u}\right)^{\lambda-1}. \tag{50}$$

The observed spectral flux may be then expressed as

$$\mathrm{d}F_E = (1-u)^{\lambda/2}\,\delta^{\lambda+3}I_E'(\sigma')\cos\sigma\,\frac{\mathrm{d}\cos\alpha}{\mathrm{d}\cos\psi}\,\frac{\mathrm{d}S'}{D^2}. \tag{51}$$

We note that the equation for the bolometric flux (49) can be obtained as a special case of equation (51) by setting $\lambda = 2$.

Let us look the expression (49) for the bolometric flux more carefully. We see that there is a difference by a factor δ^5 between the flux from a rapidly rotating star and that from a slowly rotating star. One power of the Doppler factor came from the change in the projected area due to aberration as we already mentioned. From the four other powers of Doppler factor (appearing in the equation 47) two comes from the solid angle transformation, one from the energy and one from the time contraction of photon arrival [34]. In the case of spectral flux (equation 48) there is no Doppler factor from the energy and thus there is one power less.

As mentioned, in the most simple model we may assume that the intensity is independent of σ' ("isotropic beaming"). In a little more general case one can assume the intensity to be a product of an angle-dependent ("beaming function") and an energy-dependent function. The beaming function describes how much intensity is emitted to each direction, and the function might have e.g., a polynomial dependency on $\cos \sigma'$. An atmosphere dominated by electron scattering produces a beaming function called Hopf profile.

Thanks to the observed AMXPs (e.g., SAX J$1808.4 - 3658$ [22]) we know that the spectrum of the pulsar (i.e., the energy-dependency of $I'_{E'}(\sigma')$) might be fitted with a blackbody component at lower energies and with a power-law component at higher energies. However, in the case of SAX J$1808.4 - 3658$ the blackbody contributes only about 30 per cent to the flux even in the $3 - 5$ keV region (which is near the lowest observed energies with *RXTE*) [22]. Thus we can approximate the spectrum with two different power-laws on both sides of a cut-off energy, which is defined to be at the maximum flux of a blackbody with $T_{\text{eff}} = 2$ keV (in the frame of the spot). Below the cut-off energy we use Rayleigh-Jeans law ($I'_{E'}(\sigma') \propto E'^2$ and thus $\lambda = -1$) and above the cut-off a descending power-law with the spectral index $\lambda = 1.86$. The power-law is normalized to be equal with the reference blackbody at 1 keV. This spectrum is shown in Figure 6.

In a more realistic model we would calculate the fluxes using two overlapping spots. The other spot would produce blackbody emission with isotropic beaming and the other one the power-law emission with e.g., a linear dependency on the cosine of the beaming angle σ' (representing the accretion shock). The spots may have different shapes and sizes.

In the calculations of this thesis we use monochromatic photon number fluxes (in units of photons/cm^2/s/keV) since those are the fluxes we get from the observations. We must thus divide the fluxes we get from equation (48) with the observed energy E. When calculating the flux we make again a grid of values of the derivatives $\mathrm{d}\cos\psi/\mathrm{d}\cos\alpha$ (same as $(\mathrm{d}\cos\alpha/\mathrm{d}\cos\psi)^{-1}$) with different α and R. The derivatives for each point are found by linear interpolation.
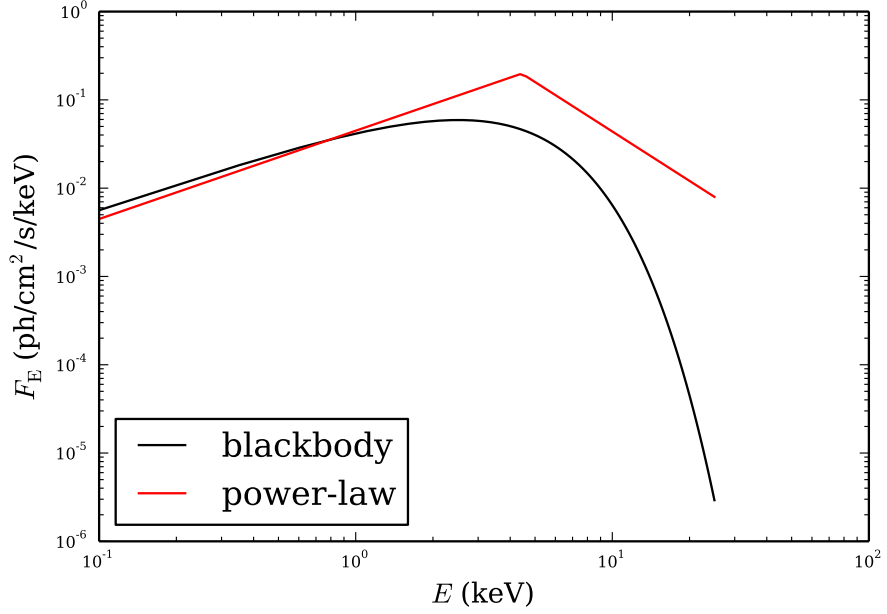
Figure 6. The two-component power-law energy spectrum compared to the reference blackbody at $2\,\mathrm{keV}$ temperature. The spectra are calculated using the time-averaged fluxes at different energies of the pulse profile model with some physical parameters.

2.1.5 Time delays

The previous expressions (48), (49), and (51) for calculating the flux of an infinitesimal spot contain the pulsar phase ϕ expressed in $\cos\sigma$. However, we should realize that photons are not necessarily observed in the same order as they are emitted due to the different light travel times when the position of the emitting spot is different. For this reason the flux actually corresponds to an observed phase ϕ_{obs}, which is different from ϕ. The delays because of light travel times become significant only for rapidly rotating pulsars.

The different light travel times in Schwarzschild metric may be calculated and subtracted to get the time delay between the photon from the point of interest and a photon from a reference point (we call this the reference photon). We choose the reference point to locate at the reference radius r_{ref} from the center of the star (it can be chosen arbitrarily as long as it exceeds r_{S}) and assume that the reference photon is emitted exactly to the radial direction (meaning that $\alpha = 0$ and thus impact parameter $b = 0$). Now a photon following the trajectory with an impact parameter b (and $\alpha < \pi/2$) is lagging the reference

photon by [23]

$$c\Delta t(R,\alpha) = c\Delta t_{\mathrm{s}}(R,\alpha) - \delta t(r_{\mathrm{ref}}, R), \tag{52}$$

where

$$c\Delta t_{\mathrm{s}}(R,\alpha) = \int_{R}^{\infty} \frac{\mathrm{d}r}{1 - r_{\mathrm{S}}/r} \left\{ \left[1 - \frac{b^2}{r^2}\left(1 - \frac{r_{\mathrm{S}}}{r}\right) \right]^{-1/2} - 1 \right\} \tag{53}$$

is the time difference between the photon in interest and a photon originating from the same radius R but with $b = 0$, and

$$\delta t(r_{\mathrm{ref}}, R) = R - r_{\mathrm{ref}} + r_{\mathrm{S}}\ln(\frac{R - r_{\mathrm{S}}}{r_{\mathrm{ref}} - r_{\mathrm{S}}}) \tag{54}$$

is the time difference between photons with $b = 0$ from R and r_{ref} [35].

In the case when $\alpha > \pi/2$ the corresponding delay is

$$\begin{aligned} c\Delta t(R,\alpha) &= 2c\Delta t_s(p, \pi/2) - c\Delta t_s(R, \pi - \alpha) \\ &+ 2\left[R - p + r_{\mathrm{S}}\ln(\frac{R - r_{\mathrm{S}}}{p - r_{\mathrm{S}}}) \right] - \delta t(r_{\mathrm{ref}}, R). \end{aligned} \tag{55}$$

The integral in equation (53) is not very easy to solve numerically. That is why we use again the variable substitution $x = \sqrt{1 - R/r}$ before using Simpson quadrature. We also make a grid of time delays corresponding to different α and R, which we later interpolate linearly in order to make the code run faster.

We are now able to compute the bending angle ψ, the emission angles α and σ, and the impact parameter b using equations (37), (38), and (35) for a given pulsar phase ϕ. The corresponding delays $\Delta t(R, \alpha)$ we can find using formulae (52) and (55). When these time delays are known, we compute the delays expressed in phase with

$$\Delta\phi(\phi) = 2\pi\nu\Delta t[b(\phi)], \tag{56}$$

and the arrival phase to the observer is then $\phi_{\mathrm{obs}} = \phi + \Delta\phi$. We may thus construct a one-to-one correspondence between ϕ and ϕ_{obs}.

The flux at the observed phase ϕ_{obs} is $F_{\mathrm{obs}}(\phi_{\mathrm{obs}}) = F(\phi + \Delta\phi)$ with phase delay $\Delta\phi = 2\pi\nu\Delta t$ calculated using (52), (55), and (56). The contraction (or stretching) of the arrival times of photons is already taken account by one of the Doppler factors in the previous section, so we don't need to multiply flux again by δ.

2.1.6 Profiles from a large spot

In all the previous sections we assumed the spot to be infinitesimally small. If the spot have a finite size, we need of course to integrate over the spot surface. It may be done by splitting the spot into a number of small sub-spots and computing the fluxes to each sub-spot separately. Integration over the spot surface is then done using Gaussian quadrature in (cosine of) colatitude and trapezoidal rule for integration over the azimuth inside the spot.

One important thing is to include the time delays correctly for each sub-spot. Because of this we compute the delays relative to the photons emitted from a point directly under the observer (with $b = 0$ as said in the previous section). A more difficult problem is to find out what is the exact relation between the area of the finite spot in the rotating and non-rotating frames ($\mathrm{d}S = \delta \, \mathrm{d}S'$ was true for an infinitesimal spot). This problem has still to be studied.

In the model in this thesis we assumed a spot which have a constant angular radius ρ (angle between the center and the edge of the spot measured from the center of the star). The spot could also be modeled to have a more complicated shape.

2.1.7 Comparison of the profiles

The new versions of the waveform code have been tested by comparing the pulse profiles to profiles obtained from other theoretical models or from a similar model developed by other groups. In all comparisons we assume the neutron star to be at distance $D = 10\,\mathrm{kpc}$. In the case of spherical star we have checked that the new version gives the same results as the old one [28] (which assumed a spherical star and has been shown to give accurate results).

In the case of oblate star we have compared our light curves to those obtained by Morsink [25] and Cadeau [27] (model called oblate S+D). The waveforms are computed for one small spot (we use spot angular size $\rho = 1°$), which emits isotropic blackbody
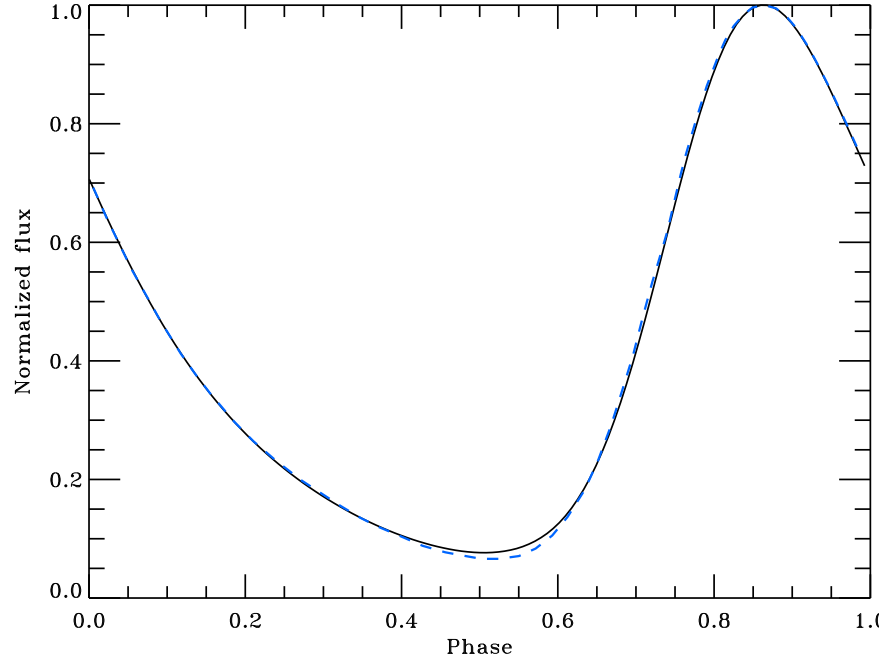
Figure 7. Light curves (bolometric flux) for light emitted from one infinitesimal spot with $R_{\mathrm{eq}} = 16.4\,\mathrm{km}$, $M = 1.4\,\mathrm{M_\odot}$, $\nu = 600\,\mathrm{Hz}$, $\theta_S = 49°$, $i = 70°$, and $T_{\mathrm{eff}} = 2\,\mathrm{keV}$. Results are compared with figure 3 of Morsink *et al.* (2007) [25]. The blue dashed curve is the result of [25] and the black solid line is our result.

radiation ($R_{\mathrm{eq}} = 16.4$ km, $M = 1.4\,\mathrm{M_\odot}$, $\nu = 600$ Hz, and $T_{\mathrm{eff}} = 2$ keV). The oblateness of the star is described with function (12). The results are shown in Figures 7, 8, and 9. They are in good agreement except in cases of some very high bending angles. In Figure 7 ($\theta_S = 49°$ and $i = 70°$) and in Figure 9 ($\theta_S = 15°$ and $i = 100°$) we see situations when the spot is on the other side of the pulsar and only photons with high bending angles can reach the observer. Instead in Figure 8 ($\theta_S = 41°$ and $i = 20°$) most of the spot is all the time observed.

The profiles of pulsars have also been compared to profiles obtained from full ray tracing method (Nättilä & Pihajoki 2017, in preparation)), which show very similar results (see Figures 10, 11, and 12). The photon number fluxes at 2, 6, and 12 keV energies, along with the bolometric photon number flux, are computed to both spherical ($R(\theta) = R_{\mathrm{eq}}$) and oblate stars ($R(\theta)$ from equation 8). Angular distribution of radiation corresponds either to a blackbody with isotropic beaming (noted as "blackbody") or to a Hopf profile.
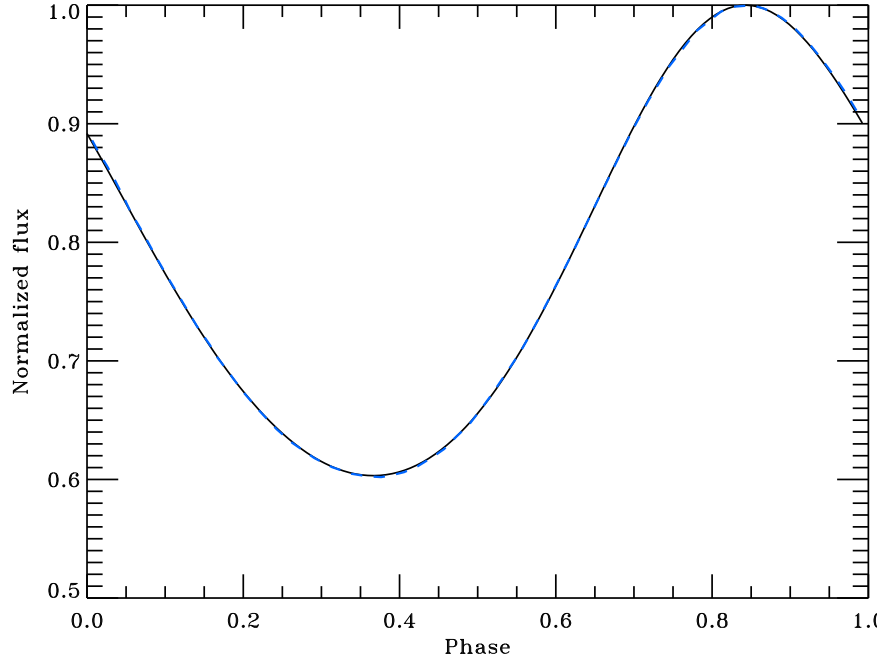
Figure 8. The light curves (bolometric flux) for light emitted from one infinitesimal spot with otherwise similar same parameters as in Figure 7 but with $\theta_S = 41°$ and $i = 20°$. The results are compared with figure 3 of Cadeau *et al.* (2006) [27]. The blue dashed curve is the result of [27] and the black solid line is our result (almost coinciding).

In the case of spherical star ($R = 12$ km, $M = 1.6$ $M_\odot$, $\theta_S = 50°$, $i = 60°$, and $T_{\text{eff}} = 2$ keV) we compute the light curves with two different spinning frequencies ($\nu = 1$ Hz and $\nu = 400$ Hz) and with the spot size $\rho = 1°$ or $30°$ (Figures 10 and 11). In the case of oblate star ($R_{\text{eq}} = 12$ km, $M = 1.4$ $M_\odot$, $\nu = 700$ Hz, $i = 45°$, $\rho = 10°$, and $T_{\text{eff}} = 2$ keV) we show the light curves with three different spot colatitudes $\theta_S = 18°$, $45°$, and $90°$ for the isotropic blackbody, along with $\theta_S = 45°$ and Hopf profile.

2.2   Bayesian inference

The model presented in the previous section can be used to constrain parameters of neutron stars by using Bayesian inference. We can fit the observed data to the model and use Markov Chain Monte Carlo (MCMC) methods to integrate over the parameter space to find the most probable values for the parameters of the model [36]. The data can also be synthetic (as in this thesis) in order to test our methods and tools. Previously the con-
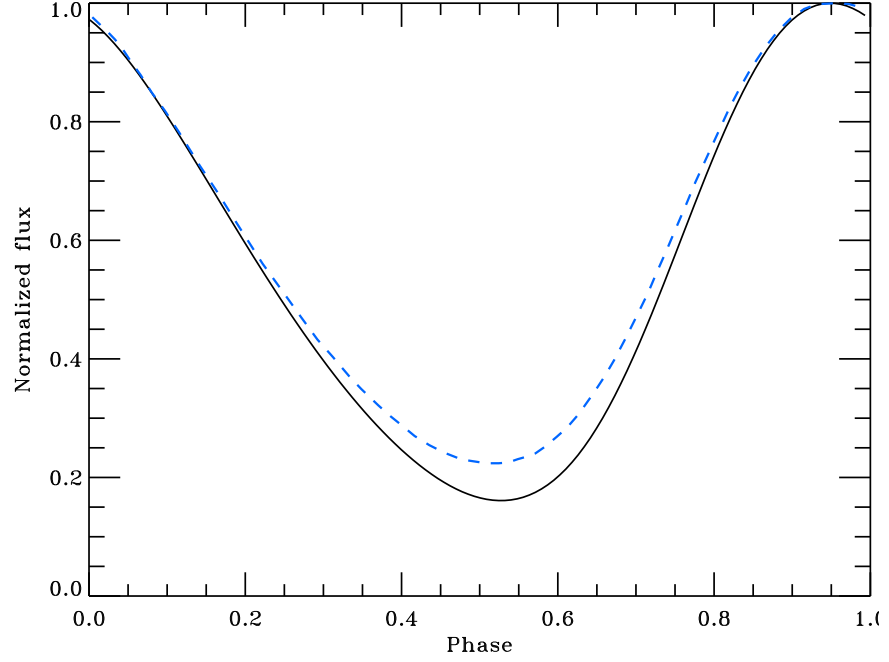
Figure 9. The light curves (bolometric flux) for light emitted from one infinitesimal spot with otherwise similar same parameters as in Figure 7 but with $\theta_S = 15°$ and $i = 100°$. The results are compared with figure 4 of Cadeau *et al.* (2006) [27]. The blue dashed curve is the result of [27] and the black solid line is our result.

straints of neutron star masses and radii with synthetic data have been studied e.g., by Ka Ho Lo *et al.*, 2013 [37] and Miller & Lamb, 2015 [38]. More about the synthetic data of this thesis is presented in the Results section.

The aim of this work is to determine the most probable ("best-fit") values of the parameters in the pulse profile model, given the "observed" waveform, and the confidence regions for the values of these parameters. The goal is also to study the effect of polarization measurements on to the posterior probability distributions. Probabilities are calculated using Bayes' theorem which is presented in the next section. The standard Metropolis sampling method of the parameter space and a modern ensemble sampler (used in this thesis) are discussed in the following three sections.
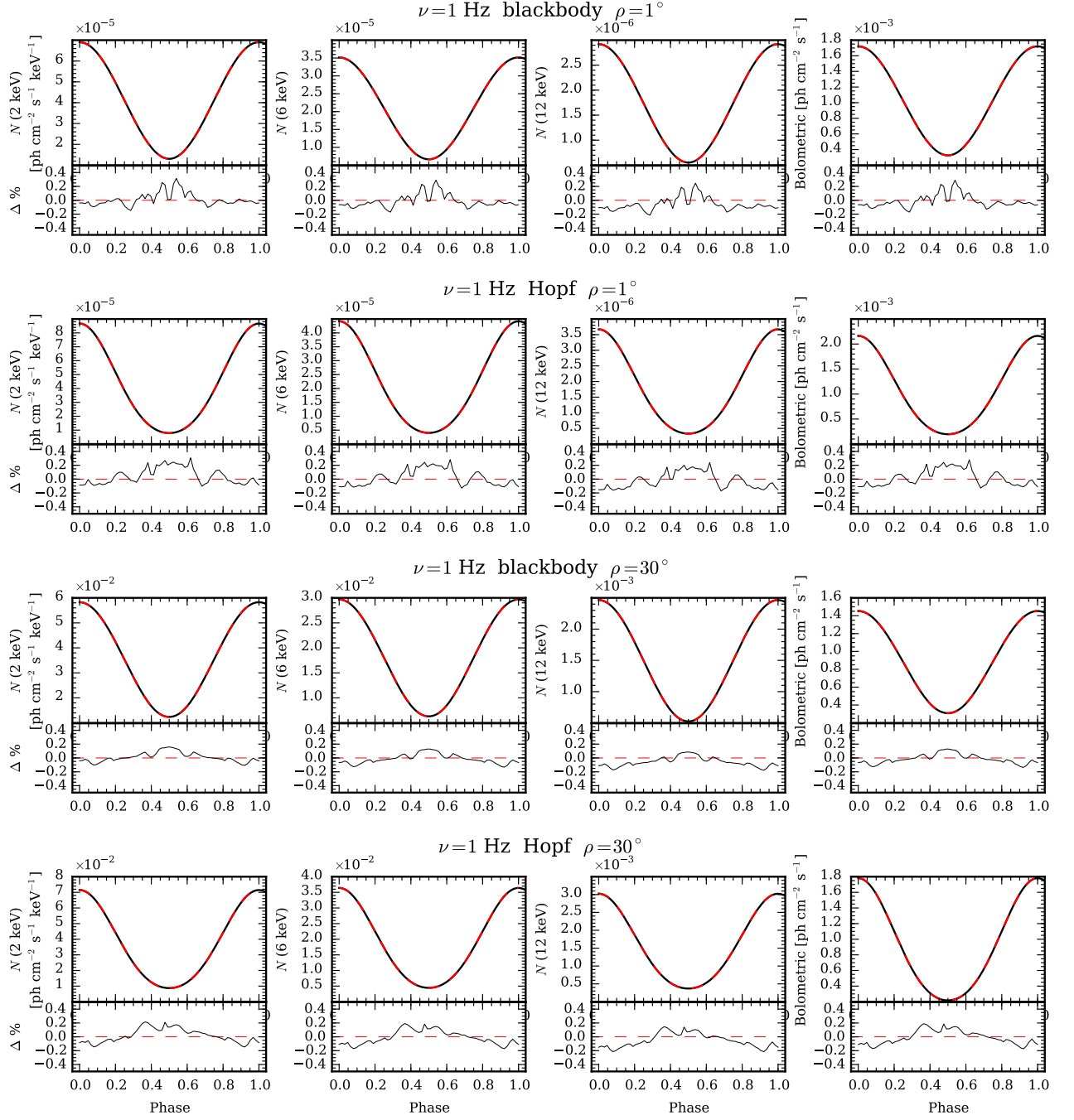
Figure 10. The light curve (monochromatic photon number fluxes at 2, 6, and 12 keV along with bolometric photon number flux) comparisons with slowly rotating spherical star ($R = 12$ km, $M = 1.6$ M$_\odot$, $\nu = 1$ Hz, $i = 60°$, $\theta_S = 50°$, and $T_{\text{eff}} = 2$ keV) emitting according to isotropic blackbody law or Hopf profile with spot size either 1 or 30 degrees. Black solid line shows the pulse profiles computed using the method described in this thesis (S+D), and red dashed line is a profile computed with the full ray tracing method (FRTM). The residuals are shown in the lower panel as $\Delta = (\text{model}_{\text{S+D}}/\text{model}_{\text{FRTM}}) * 100\%$. From Nättilä & Pihajoki 2017 (in preparation).
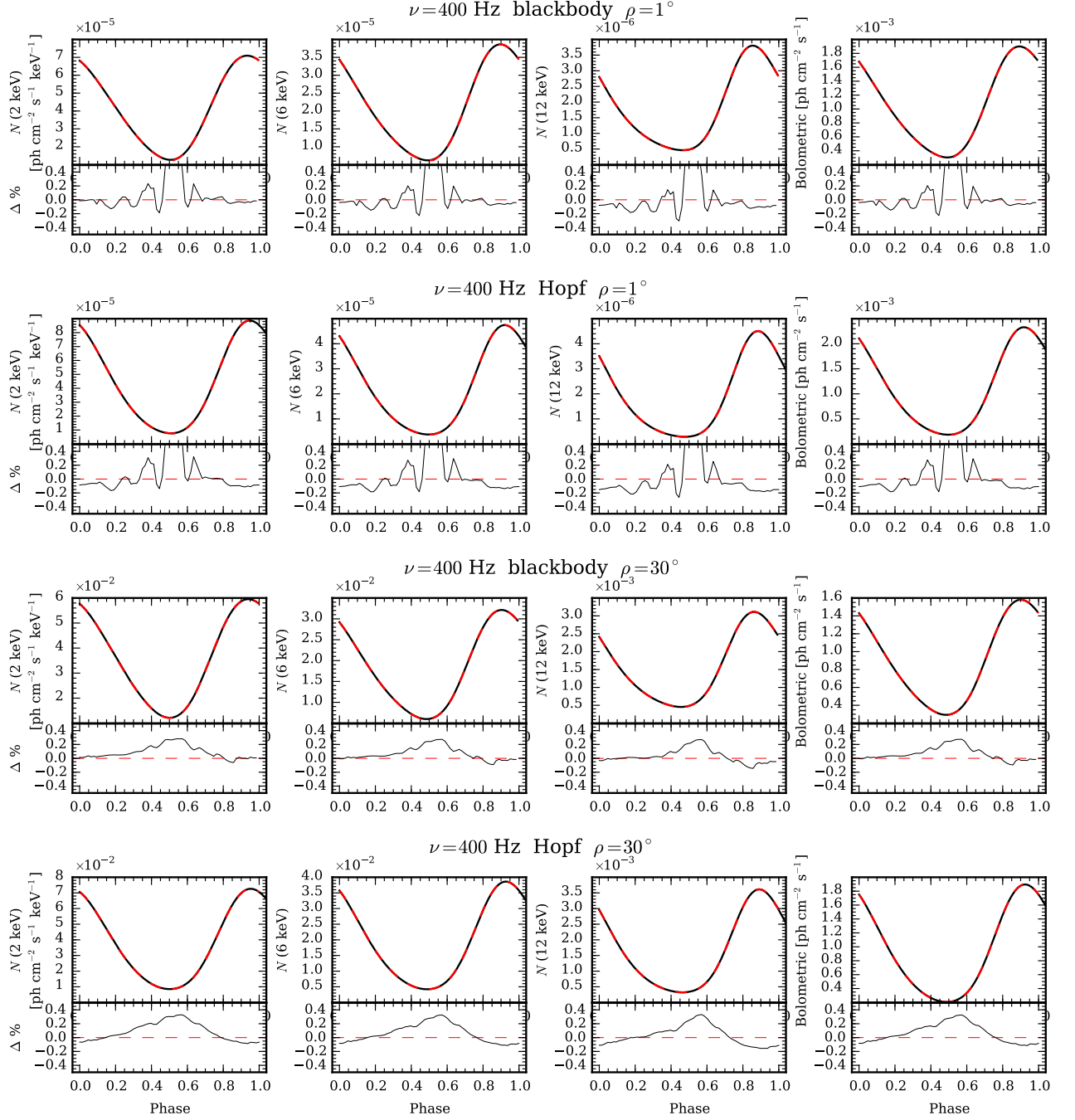
Figure 11. The light curve (monochromatic photon number fluxes at 2, 6, and 12 keV along with bolometric photon number flux) comparisons with rapidly rotating spherical star ($\nu = 400$ Hz). Otherwise the parameters and symbols are the same as in Figure 10. From Nättilä & Pihajoki 2017 (in preparation).
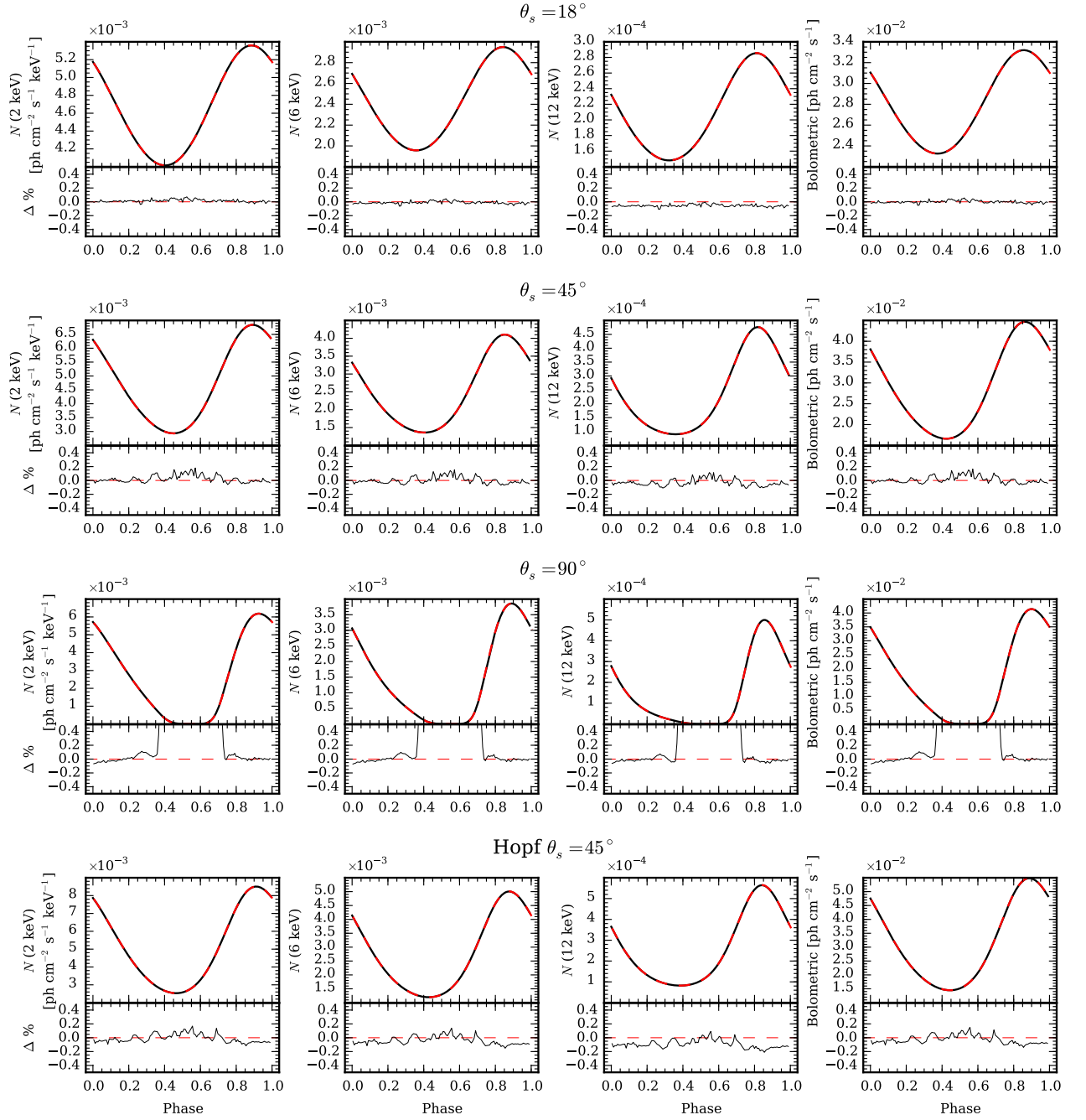
Figure 12. The light curve (monochromatic photon number fluxes at 2, 6, and 12 keV along with bolometric photon number flux) comparisons with an oblate star ($R_{eq} = 12$ km, $M = 1.4\,M_\odot$, $\nu = 700$ Hz, $i = 45°$, $\rho = 10°$, and $T_{eff} = 2$ keV) emitting according to isotropic blackbody law or Hopf profile with different spot colatitudes: $\theta_S = 18°$, $45°$ and $90°$. The symbols are the same as in Figures 10 and 11. From Nättilä & Pihajoki 2017 (in preparation).

2.2.1 Bayesian analysis

We are interested in the probability of the parameters $\mathbf{y}$ of the waveform model when the observed waveform is known. The probability distribution of the parameters given the data is $p(\mathbf{y}|\mathcal{D})$, where $\mathcal{D}$ is the energy- and oscillation phase-resolved waveform data (synthetic in our case). According to the Bayes' theorem this (posterior) probability distribution can be obtained from the likelihood of the data, given the parameter values as [39]

$$p(\mathbf{y}|\mathcal{D}) \propto p(\mathcal{D}|\mathbf{y})p(\mathbf{y}). \tag{57}$$

In the previous equation $p(\mathcal{D}|\mathbf{y})$ is the likelihood or the probability distribution of the data given the parameters. The factor $p(\mathbf{y})$ is the prior probability distribution of the parameter values. As a first approximation we use uniform prior which is the most uninformative prior. Later we take into account the information of polarization measurements and make the priors of inclination and spot colatitude non-uniform to study their impact on the fit. The constant of proportionality is the inverse of the normalization factor, but it is irrelevant when estimating the values of the parameters in a given model.

2.2.2 Metropolis-Hastings

The original standard MCMC algorithm is called the Metropolis algorithm [40][41]. It is a MCMC method for obtaining a sequence of random samples from a probability distribution. It generates these random samples by moving in a random walk; that is in some sequence $X_1...X_\mathrm{t}$. Metropolis algorithm, like all the other MCMC sampling methods, satisfies the so-called Markov property. It means that the conditional probability distribution of $X_\mathrm{t+1}$, given all the past elements, is independent of all the other states except the previous state [42]:

$$p(X_\mathrm{t+1} = x|X_\mathrm{t}...X_1) = p(X_\mathrm{t+1} = x|X_\mathrm{t}). \tag{58}$$

So generally we can conclude that a Markov chain is a random walk with Markov property.

The Metropolis algorithm works by taking an arbitrary move near the current point in the sequence. If X_t is the sample at time t, then the new sample Y is proposed from a proposal distribution $q(Y|X_t)$. In the original Metropolis algorithm this distribution must be symmetric, meaning that $q(X_t|Y) = q(Y|X_t)$. However, in Metropolis-Hastings algorithm it may be non-symmetric. The likelihood of the new sample is then in both cases compared to that of the previous sample. The likelihood ratio $\mathcal{R}$ between the two states is calculated from the following equation [43]:

$$\mathcal{R} = \frac{p(Y|\mathcal{D})q(X_t|Y)}{p(X_t|\mathcal{D})q(Y|X_t)}. \tag{59}$$

The conditional probabilities given the data are calculated using the Bayes formula (57). The ratio of these probabilities is multiplied with the ratio of proposal distributions q (sometimes called transition kernels) so that the algorithm satisfies a detailed balance condition. It is an important property for proving the convergence of the chain. It states that the number of transitions from X_t to Y must be the same as from Y to X_t. The proposal density $q(X_t|Y)$ describes the probability of a transition from Y to X_t in the parameter space.

The next step is to decide whether the likelihood ratio is high enough in order to accept the new step. In Metropolis(-Hastings) algorithm we automatically accept a step if $\mathcal{R} \geq 1$. Otherwise we accept only if $\mathcal{R}$ is greater than a random number drawn between 0 and 1. If the new state is accepted, we repeat the first steps using now Y as the current point. Otherwise we propose a new sample using again the same X_t.

It may be noticed that the acceptance rule simplifies significantly when the proposal density is indeed symmetric ($q(X_t|Y) = q(Y|X_t)$). The algorithm is typically repeated until the obtained chains are long enough in the sense that their statistics do not change considerably when adding new members to the chain. The statistical representativeness of the sample can be verified if multiple chains with different initial states result in the

same posterior density. It is also necessary to remove the so-called initial warm-up period ("burn-in") since the probability distribution has not yet converged in that case. There is no general theory to decide the optimal burn-in period.

2.2.3 Autocorrelation and difficulties with Metropolis

Metropolis(-Hastings) algorithm has also some limitations. The samples are not independent (as we would like them to be) since they are generated in a chain (concerns though also other MCMC methods). The correlation between the samples can be measured by the autocorrelation time (also called integrated autocorrelation time) [44]:

$$\tau = \sum_{t=-\infty}^{\infty} \frac{C(t)}{C(0)},$$ (60)

where the function $C(t)$ is autocovariance with lag t:

$$C(t) = \lim_{t' \to \infty} \mathrm{Cov}(V(X_{t+t'}), V(X_t)).$$ (61)

Here $V(X_t)$ is an observable. Basically this measures how related the series of numbers $V(X_t)$ is with itself.

One way to compute the autocorrelation time is to use the weights (log-likelihoods) of the fits as the observable $V(X_t)$. If the weights are strongly correlated with many different lags, τ will be large. Of course most of the correlation is originated from samples close to each other (small lags). This can be seen by inspecting the autocovariance as function of the lag (see Figure 17 in Results). There is no way to get completely rid of autocorrelation, but at least some of the problem may be solved by making the sequence of samples thinner. In principle the number of samples which should be independent is N_{step}/τ, where N_{step} is the total number of steps [42]. However, especially with Metropolis algorithm τ is typically high (unless with very carefully designed proposal densities) meaning that very high fraction of the samples should be removed.

One other major disadvantage is that the Metropolis algorithm may work poorly for skewed or anisotropic distributions depending on the proposal distribution. One simple

Figure 13. Contours of the Gaussian density defined in expression (62). From Goodman and Weare (2010) [44].

example of such is a two dimensional skewed probability distribution (shown in Figure 13)

$$H(\vec{x}) = \exp\left(-\frac{(x_1 - x_2)^2}{2\epsilon_c} - \frac{(x_1 + x_2)^2}{2}\right),$$ (62)

where ϵ_c is a constant. Metropolis (like many other MCMC strategies too) would be forced to make perturbations of order $\sqrt{\epsilon_c}$ and would have slow equilibration. A better MCMC sampler would use larger perturbations in (-1,1) than in (1,1) direction. However, the problem gets worse in high dimensions where it is difficult to tune all the step sizes correctly. As a next step to address this issue we can use the ensemble sampler which is presented in the next section.

2.2.4 Ensemble sampler

The affine invariant ensemble sampler is a novel MCMC method developed a few years ago [44]. The algorithm has a similar structure to Metropolis scheme and still uses a proposal and accept/reject step. But instead of only one sequence $X_1...X_t$, we use now a

group or ensemble of sequences. Each member of the group is called a walker. On each iteration the algorithm generates a new sample for every walker using the current positions of all the other walkers in the ensemble. However, each walker is independent of the other walkers. They are correlated only with the previous states of the other walkers and their own previous state. After burn-in the distribution of each walker should be converged to the invariant posterior probability distribution.

The ensemble sampler is called affine invariant since its performance is unaffected by affine transformations of space. This means that one can sample a probability distribution function $\mathcal{G}(x) = \mathcal{A}\mathcal{F}(x) + \mathcal{B}$ either directly or sample first $\mathcal{F}$ and then apply the constants $\mathcal{A}$ and $\mathcal{B}$ afterwords. For this reason the algorithm is particularly useful for sampling badly scaled distributions. Computational tests show that the affine invariant methods can be significantly faster than standard MCMC methods on highly skewed distributions.

In the ensemble sampler the Markov chain is evolved by moving one walker at time. Each walker $X(k)$ is updated using the current positions of all of the other walkers in the ensemble. We call them a complementary ensemble. The motivation is that the distribution of the walkers in the complementary ensemble carries useful information about the probability density of the parameters. This way the trial move can be adapted to the target density.

One of the simplest affine invariant trial moves, and the move we use here, is the so-called stretch move. In the stretch move algorithm each walker is moved using only one randomly selected complementary walker (see Figure 14). The move of current walker k is proposed along a line between the walker itself and the complementary walker j according to the distribution [45]

$$Y(k) = X(j) + Z(X(k) - X(j)),\tag{63}$$

where the scaling variable Z is a random variable drawn from a probability distribution $g(\mathcal{Z})$ satisfying the symmetry condition

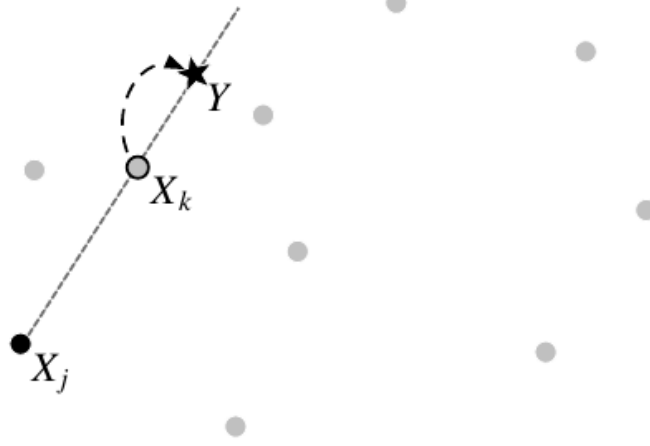$$g(1/\mathcal{Z}) = \mathcal{Z}g(\mathcal{Z}).\tag{64}$$

Figure 14. A stretch move for updating the position of X_k based on the position of another random walker, X_j (subscript identifies here the walker instead of the step count). The light-gray walkers are other members of the ensemble. From Goodman and Weare (2010) [44].

This ensures that the move (63) is symmetric in the sense that the transition probability from $X_t(k)$ to $Y(k)$ is the same as from $Y(k)$ to $X_t(k)$.

When deciding whether the new sample is accepted or not, the Metropolis-Hastings selection rules are again applied. However, instead of equation (59) the likelihood ratio is now calculated (for a given walker) as

$$\mathcal{R} = Z^{N-1} \frac{p(Y|\mathcal{D})}{p(X_t|\mathcal{D})}, \qquad (65)$$

where N is the number of parameters sampled [45]. The transition probabilities q are not needed since in the case of symmetric move (63) the acceptance ratio (65) can be shown to satisfy the detailed balance condition.

3 Results

As mentioned earlier the aim of this work is to determine the probability distributions of the parameters in the pulse profile model (especially the mass and radius), given the synthetic pulse profile, and study the effects of polarization measurements to the results. Some details of the data, sampling, and constraints obtained are presented in the three following sections.

3.1 Synthetic data

The synthetic data in this work is made to resemble the real observations as closely as possible. The X-ray count data from SAX J$1808.4 - 3658$ is used as a reference. The amplitude of the modulation in the normalized number fluxes have been set close to the same value in the synthetic data as in the observations. We also use the same energy intervals and same the number of phase points as in the real data. The original data is shown in Figure 4.

The parameters of the pulse profile model are chosen such that their values are physically reasonable and they produce a light curve similar to the SAX-data. The physical parameters are given in Table 1. The observed variability amplitude (see equation 7) is then obtained by choosing inclination i and the colatitude of the spot θ_S appropriately since the amplitude can be approximated (neglecting the Doppler effect and using the Beloborodov's approximation [46])

$$A \approx \frac{(1 - r_\mathrm{S}/R)\sin i \sin \theta_S}{r_\mathrm{S}/R + (1 - r_\mathrm{S}/R)\cos i \cos \theta_S}. \tag{66}$$

We see that i and θ_S can be changed in equation (66) without affecting the amplitude. The exact calculation shows also only a small difference between these two parameters. Thus we have created two identical synthetic datasets which differ only in i and θ_S. We call them datasets for polar and equatorial spots even though the spots are not exactly at the pole or equator.

Table 1. Parameters of the synthetic datasets.

Parameter	Value
Radius R_{eq}	12.0 km
Mass M	1.5 $\mathrm{M}_\odot$
Inclination i	5° or 75°
Spot colatitude θ_S	75° or 5°
Spot angular size ρ	10.0°
Distance D	2.5 kpc
Temperature T_{eff}	2.0 keV

The datasets have been generated using our pulse profile model. Since the observed data always includes noise, we want to have it also in our synthetic data. For this reason a different random number, from a normal distribution with mean equal to zero, is added to the original flux at each phase point. The average size of this random number have been chosen to match closely the known errors of the SAX J$1808.4 - 3658$ data. Thus a relatively larger noise has been added to the highest energy channels where the count rates are lowest. The normalized synthetic spectral fluxes and their errors (amount of noise added) are shown in Figures 15 and 16.

In our simulations we see how the two datasets with different $i - \theta_S$ solutions can be separated from each other at least with help of the polarization information giving constraints to i and θ_S. We also show their effect on the constraints of mass and radius which are our parameters of interest.

3.2   Sampling methods

All the parameters presented in Table 1, except temperature, are sampled using the ensemble sampler. There are also other parameters in the model which are not currently
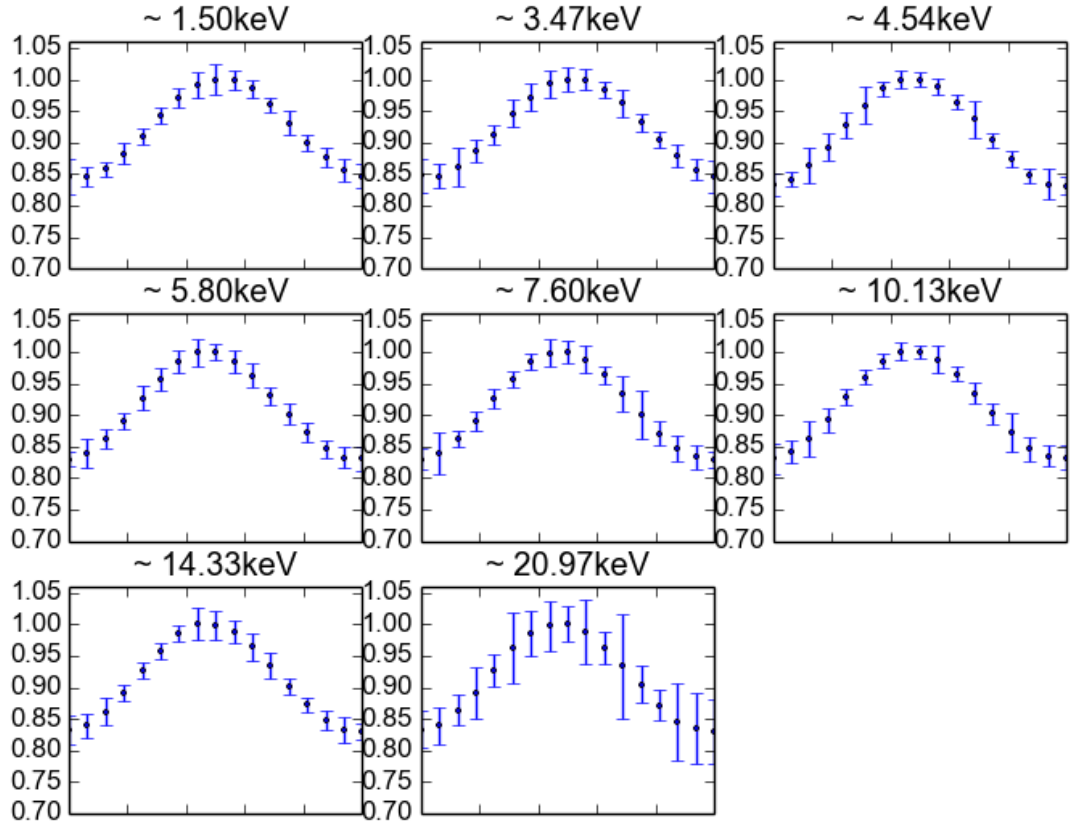
Figure 15. Synthetic data: Normalized number fluxes as a function of the phase at different energies for $i = 75°$ and $\theta_S = 5°$. The phase is from 0 to 1. Other parameters are shown in Table 1.

Figure 16. Synthetic data: Normalized number fluxes as a function of the phase at different energies for $i = 5°$ and $\theta_S = 75°$. The phase is from 0 to 1. Other parameters are shown in Table 1.

sampled. "Beaming parameter" is set to zero meaning that the intensity from the spot does not depend on the emission angle (isotropic beaming).

We also use only one circular hot spot. The oblate shape of the star is described by equation (8). Some parameters affecting the accuracy of the pulse profile calculation have been changed after creating the synthetic data in order to get faster sampling. For example the quadratic interpolation of light bending angles have been replaced by linear interpolation. The accuracy is still more than enough for our purposes.

To exploit all possible computational power we have also made the ensemble sampler code to run parallel. In this thesis we did simultaneously six independent simulations on six different processors and then combined the results of each individual chain. Thus each processor had its own ensemble (or also called the chain). The number of walkers was set to 20 in each ensemble.

The choice of the reference radius r_{ref} needed in time delay calculation affects only the phase shift of the waveform. Since we want to simulate a realistic situation where the phase shift is unknown, we use different r_{ref} in the generation of the data than in the fitting procedure of the ensemble sampler.

The likelihood of the data for each sample ($p(\mathcal{D}|\mathbf{y})$ in equation 57) is calculated by fitting the data to the pulse profile model. To be more specific, we assume that the probability density of the data fluxes F_{data} given the modeled fluxes F_{model} is normally distributed around the F_{model} with the known errors as standard deviation (σ_{E} in the following equation):

$$p(F_{\mathrm{data}}|F_{\mathrm{model}}) \propto \exp\left(\frac{-(F_{\mathrm{data}} - F_{\mathrm{model}})^2}{2\sigma_{\mathrm{E}}^2}\right). \tag{67}$$

Normalization is not needed since it will cancel out when calculating likelihood ratios. The total probability density of the data given the sample is the product of $p(F_{\mathrm{data}}|F_{\mathrm{model}})$ from each phase point and at each energy (or the sum of log-probabilities).

Of course we also need to take into account that we have different phase shift in the data and in the model. For that reason we calculate the model with five times higher phase

resolution than the resolution of the data. We calculate the probability densities using equation (67) with different phase shifts in order to marginalize over the uncertainty of phase shift. The total probability $p(\mathcal{D}|\mathbf{y})$ (which will be used to calculate the likelihood ratio given by equation 65) is the marginalized probability over all phase shifts.

As a first step we assume that all the prior probabilities of the parameters ($p(\mathbf{y})$) are uniform so that equations (57) and (65) give

$$\mathcal{R} = Z^{N-1} \frac{p(D|Y)}{p(D|X_\mathrm{t})}. \tag{68}$$

As a second step we make the prior probabilities of inclination and spot colatitude non-uniform so that the previous likelihood ratio is multiplied with the ratio of the prior probabilities, to yield

$$\mathcal{R} = Z^{N-1} \frac{p(D|Y)p(Y)}{p(D|X_t)p(X_\mathrm{t})}. \tag{69}$$

It is not obvious what kind of prior probabilities one should use. It depends on how informative constraints we are able to get from the observations to the given parameter. Unlike in the case of the usual flux, the measured polarization differ significantly if i and θ_S are switched [46]. Thus from modeling the polarization we can get the prior probability distributions for these parameters. In this thesis we choose to use normally distributed prior probabilities for i and θ_S around 0° or 90° to study the effect on the posterior probabilities of mass and radius. The standard deviations for these priors are chosen to be 15°.

Other prior probabilities of the parameters are kept always uniform inside their sampling intervals. The sizes of these intervals are the same as the sizes of the corresponding axes in the posterior probability figures in the next section.

3.3   Parameter constraints

Before obtaining the final posterior probability distributions we have reduced the burn-in phase from each ensemble by removing the first $20,000$ samples from each chain.
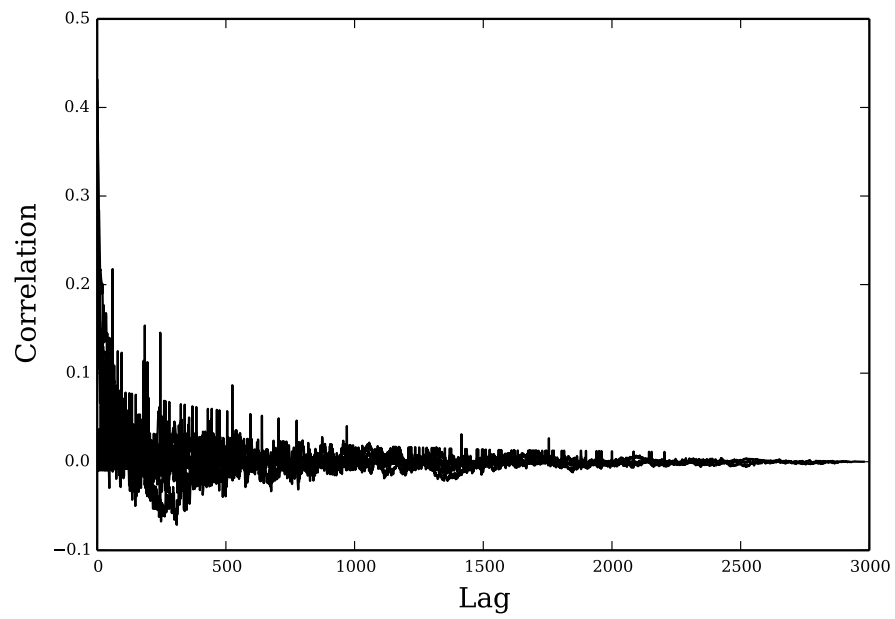
Figure 17. Autocorrelations of all the ensembles of the polar spot with only uniform priors.
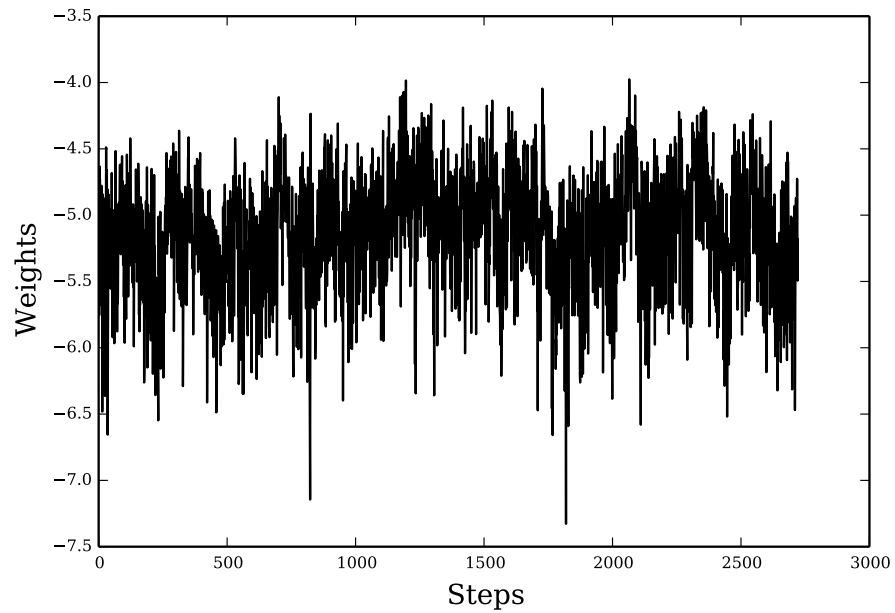

Figure 18. Log-likelihoods of one ensemble as a function of steps after the burn-in has been removed.

After that we still have approximately $50,000$ samples left belonging to each ensemble. Figure 18 shows one example of the log-likelihoods of a converged chain in one of the simulations. In that Figure we use an averaged step of ensemble which is an average of 20 (number of walkers in the ensemble) consecutive single accepted steps. However, for many ensembles we found few walkers which didn't seem to achieve convergence even after a long time (they appeared as sharp peaks in the log-likelihood plots).

The same averaged steps are also used in Figure 17, where we show an example of the autocorrelation. The autocorrelations of all the six ensembles from one simulation are presented using the weights as observables (see equation 61). The Figure shows the autocorrelation only approximatively since one should make sure that the ensemble averaged weights are calculated for walkers being at a same state of their sequence. We may, however, see that in general the correlation is not very strong, and we have reduced it by taking every fifth sample from the chain.

The results of our samplings are presented in Figures 19-22. From these posterior probability distributions we can easily see how the prior information of angles θ_S and i improves the fits of the mass and radius of the neutron star (by comparing Figures 20 and 19 or Figures 22 and 21). However, the correct ratio between mass and radius (and thus the gravitational redshift) seems to be in every case well found.

Without the prior information we get the highest probabilities for the lowest masses and radii, since those correspond to more possible solutions in θ_S and i than the highest masses and radii. In this case we are only able to get upper limits to the mass and radius. All calculated confidence limits and most probable values are shown in Table 2. The limits are calculated using the highest posterior probability credible intervals. In case of the polar spot, the radius is below 14 km, and the mass is below $1.7\,\mathrm{M_\odot}$, with 95% probabilities. The most probable value for radius is 9 km and for mass $1.2\,\mathrm{M_\odot}$. In case of the equatorial spot, the radius is below 16 km with a 95% probability, but for the mass the 95% credible interval extends until the upper limit of the sampling interval ($2.0\,\mathrm{M_\odot}$). The most probable

value for radius is 9 km and for mass 1.1 $M_\odot$.

The polar and equatorial spots give quite similar results compared to each other as expected (Figures 19 and 21). However, they are not identical which shows that switching θ_S and i has a small but significant effect on the light curves of the pulsar. It seems also slightly easier to find out the correct $\theta_S - i$ solution in the case of the equatorial spot. Although, in the one-dimensional histograms the small angles give always a higher peak in the probability density because of the marginalization and the shape of two-dimensional probability distribution of $\theta_S - i$. For the polar spot, the 95% credible interval for i is between $3°$ and $77°$, when the correct value is $75°$. However, the 68% confidence limits are $3°$ and $63°$. The 95% interval for θ_S is between $3°$ and $76°$, when the correct value is $5°$. In this case the 68% confidence limits are from $3°$ to $71°$. For the equatorial spot, the 95% limits are almost the same (and the correct values are switched), but we see that the probability density of θ_S is distributed more towards the higher angles compared to results of the polar spot. Inclination is now also better constrained. It is between $3°$ and $18°$ with a 68% probability.

In case of non-uniform priors for θ_S and i (Figures 20 and 22), we find the correct solutions for θ_S and i even though our prior distributions are not centered around the correct values. In case of the polar spot, θ_S is now between $3°$ and $9°$ with a 95% probability, and i is larger than $73°$ with a 68% probability. In case of the equatorial spot, i is between $4°$ and $6°$ with a 68% probability, and θ_S is larger than $67°$ with a 68% probability. In both cases the most probable values are close to the correct values (see the Table 2), especially for the smaller one of the angles. The posterior probability distribution for the larger angle is almost entirely determined by the assumed prior probability distribution.

The non-uniform priors for θ_S and i have also a strong effect on the posterior probability distributions for the mass and the radius. They are now also distributed on both sides of the correct values having the maximum near the correct value. The prior information shifts the probability densities of mass and radius towards higher masses and radii, since

the incorrect $\theta_S - i$ solutions are in this case not contributing to the small masses and radii so much. Most importantly both the lower and upper limits can now be determined. In case of the polar spot, the radius (correct value 12 km) is between 11 km and 15 km, and the mass (correct value 1.5 $M_\odot$) is between 1.4 $M_\odot$ and 1.8 $M_\odot$, with 68% probabilities. The most probable values are 13 km and 1.6 $M_\odot$. In case of the equatorial spot and with same the confidence limits, the radius is between 10 km and 14 km, and the mass is between 1.2 $M_\odot$ and 1.7 $M_\odot$. The most probable values are 11 km and 1.5 $M_\odot$.

The angular size of the spot ρ is found to be the closest to the correct value from all the parameters in our sampling. The highest probability density is very close to $10°$ (ρ is between $6°$ and $14°$ at least with a 95% probability in every simulation) which is the value used when creating the datasets. Partly this is due to the choice to sample the distance in a quite small interval (between 2 and 3 kpc). The distance is not constrained, and that causes problems with a larger interval. A uniform posterior probability distribution for the distance is obtained using the narrow limits. For this reason the size of the spot cannot change very much in order to keep the observed count rate at a correct level.

Table 2. Most probable values and 68% and 95% confidence limits for 4 different simulations.

| Quantity	95% lower limit	68% lower limit	Most probable value / median	68% upper limit	95% upper limit
Polar spot with only uniform priors					
Radius R_{eq} (km)	8.1	8.3	9.0	12.4	13.8
Mass M (M$_\odot$)	1.0	1.1	1.2	1.6	1.7
Inclination i (deg)	3.2	3.4	6.2	62.7	77.1
Spot colatitude θ_S (deg)	3.3	3.4	6.8	70.8	76.2
Spot angular size ρ (deg)	6.8	8.3	9.8	11.2	12.7
Distance D (kpc)	2.1	2.2	2.5	3.0	3.0
Polar spot with non-uniform i and θ_S priors					
Radius R_{eq} (km)	8.7	10.8	12.8	14.7	16.1
Mass M (M$_\odot$)	1.1	1.4	1.6	1.8	2.0
Inclination i (deg)	11.6	72.8	80.8	88.6	89.5
Spot colatitude θ_S (deg)	2.8	3.6	5.4	5.5	9.4
Spot angular size ρ (deg)	7.4	8.4	10.0	11.4	13.0
Distance D (kpc)	2.1	2.2	2.5	2.9	3.0
Equatorial spot with only uniform priors					
Radius R_{eq} (km)	8.1	8.2	9.1	10.8	16.1
Mass M (M$_\odot$)	1.0	1.0	1.1	1.4	2.0
Inclination i (deg)	3.3	3.7	7.9	18.1	79.1
Spot colatitude θ_S (deg)	4.6	4.8	8.2	73.5	77.2
Spot angular size ρ (deg)	6.5	8.7	10.3	11.6	13.1
Distance D (kpc)	2.0	2.1	2.5	2.9	3.0
Equatorial spot with non-uniform i and θ_S priors					
Radius R_{eq} (km)	8.2	9.8	10.7	14.1	16.1
Mass M (M$_\odot$)	1.1	1.2	1.5	1.7	2.0
Inclination i (deg)	2.3	3.5	5.5	6.1	67.1
Spot colatitude θ_S (deg)	4.0	67.4	77.4	85.4	89.3
Spot angular size ρ (deg)	7.2	8.2	9.9	11.3	13.5
Distance D (kpc)	2.1	2.1	2.5	2.9	3.0
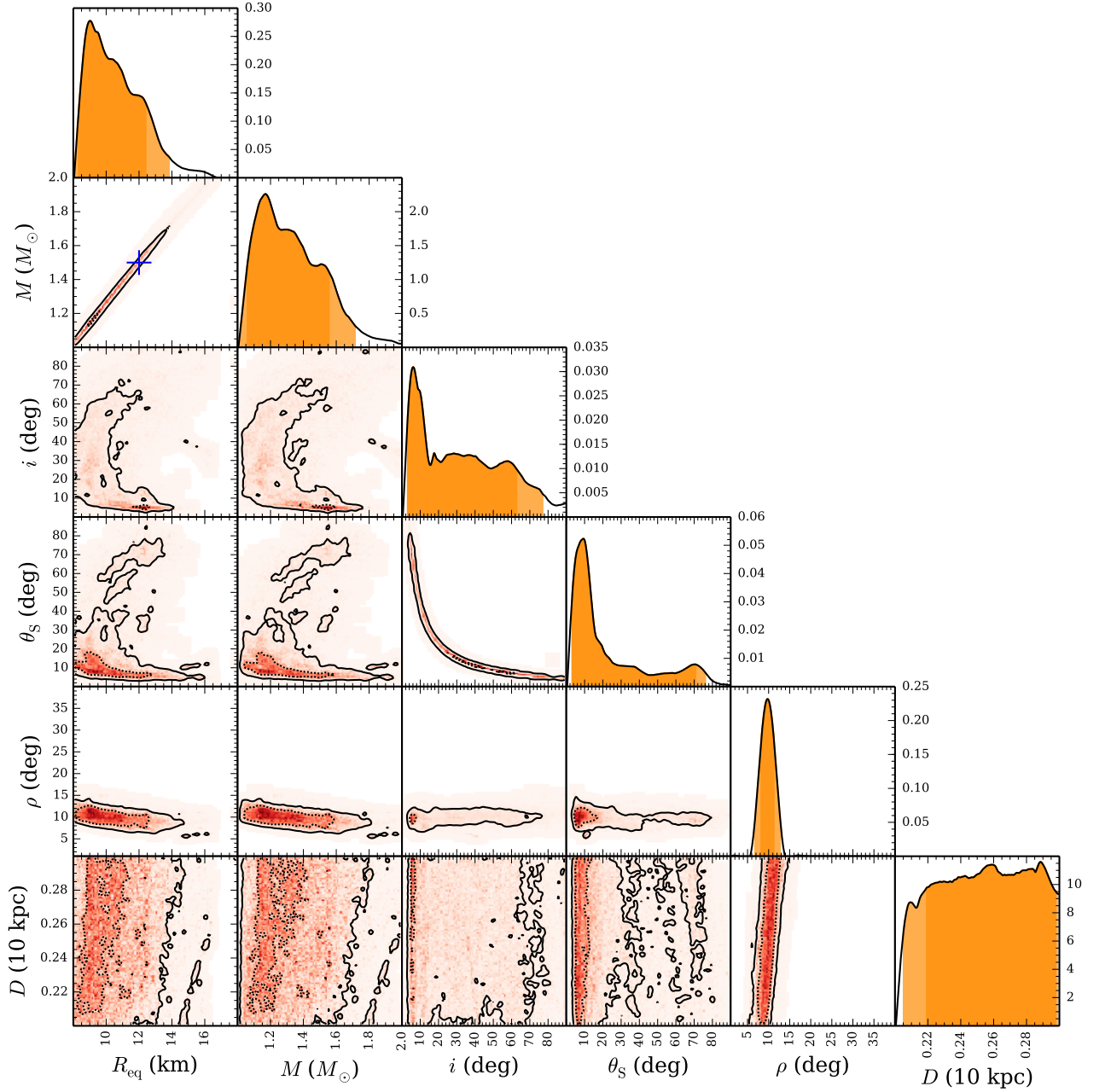
Figure 19. Posterior probability distributions for radius $R_{\rm eq}$, mass M, observer inclination i, spot colatitude θ_S, spot angular size ρ, and distance D for the polar spot with only uniform priors. The one-dimensional distributions are normalized to give 1 as the total probability. The dark orange color shows a 68% and the light orange color a 95% highest posterior density credible interval. In the 2D posterior distributions the solid contour shows a 95% and the dashed contour a 68% highest posterior density credible region. The blue cross shows the correct solution for mass and radius.
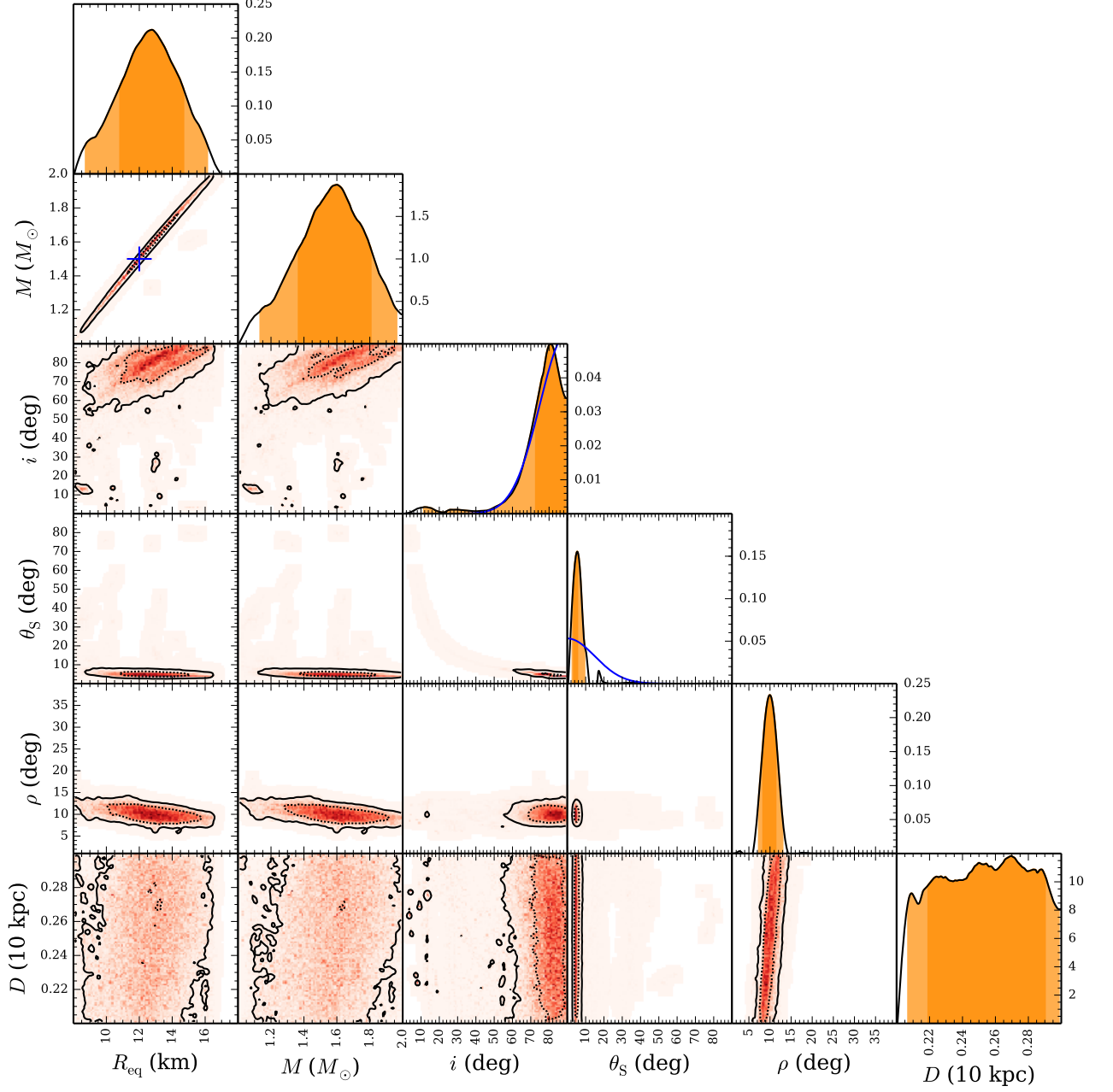
48

Figure 20. Posterior probability distributions for radius R_{eq}, mass M, observer inclination i, spot colatitude θ_S, spot angular size ρ, and distance D for the polar spot with non-uniform i and θ_S priors (shown with a blue line). The one-dimensional distributions are normalized to give 1 as the total probability. The colors, contours, and other symbols are the same as in Figure 19.
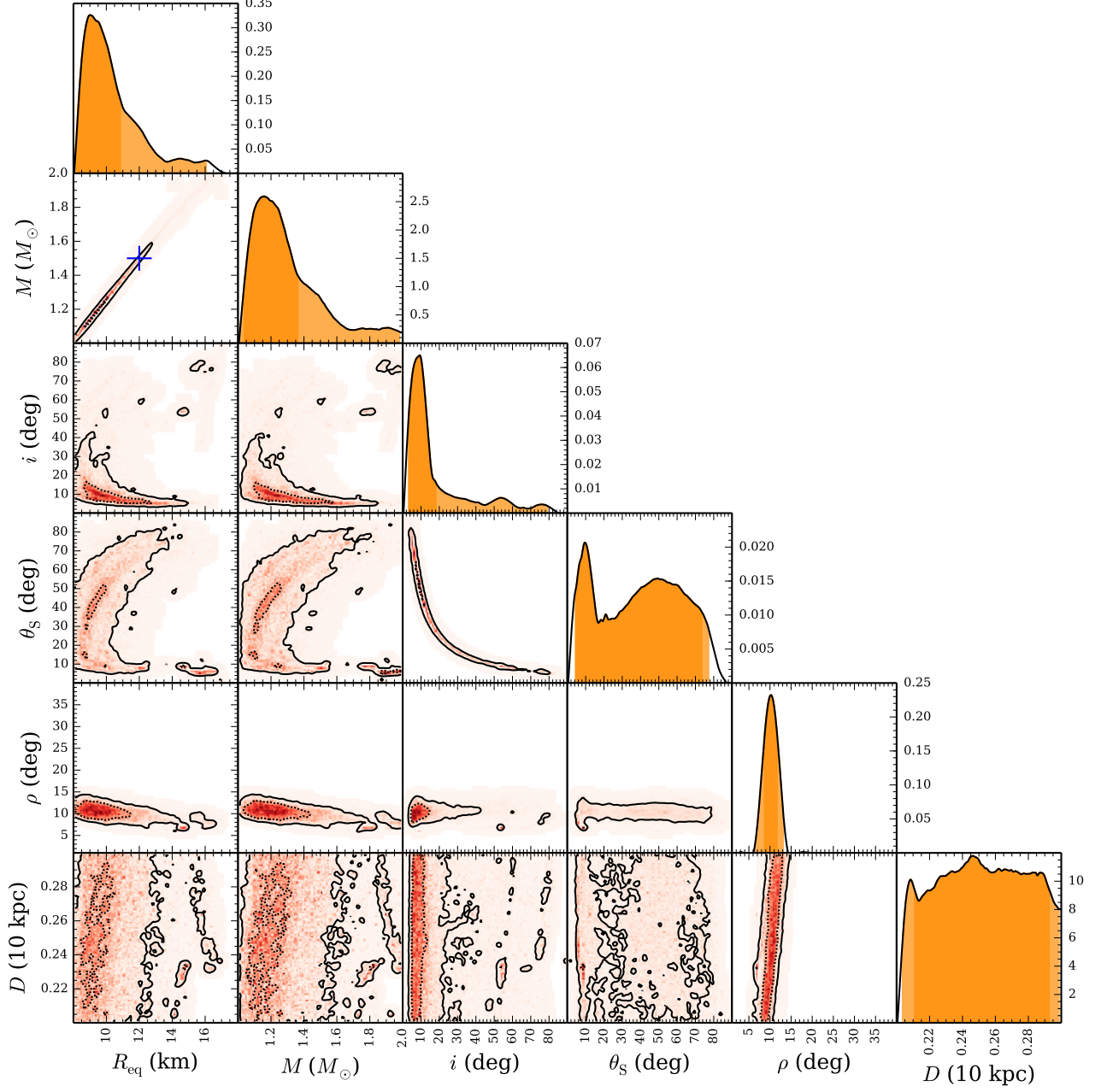
Figure 21. Posterior probability distributions for radius R_{eq}, mass M, observer inclination i, spot colatitude θ_S, spot angular size ρ, and distance D for the equatorial spot with only uniform priors. The one-dimensional distributions are normalized to give 1 as the total probability. The colors, contours, and other symbols are the same as in Figure 19.
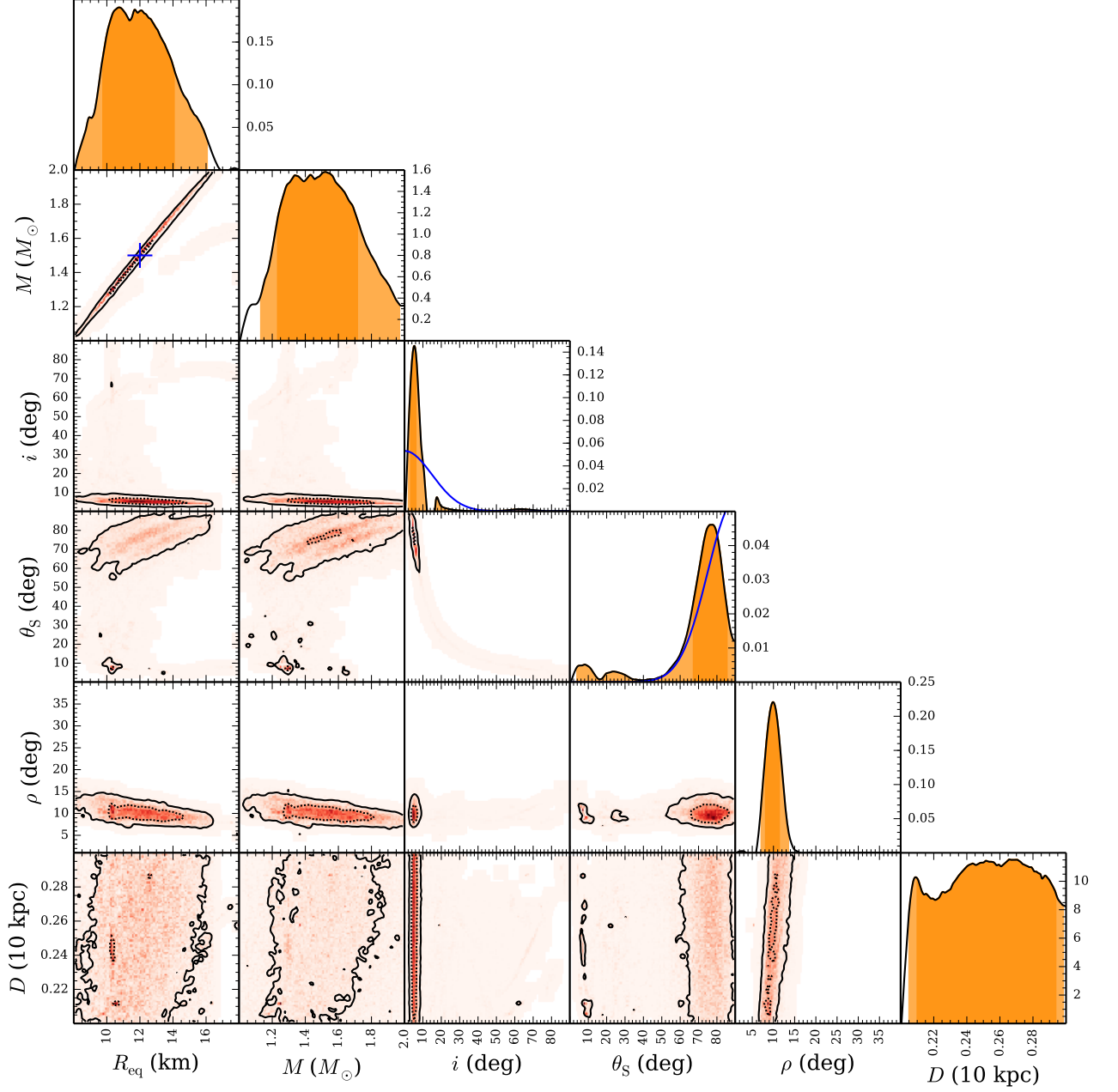
Figure 22. Posterior probability distributions for radius R_{eq}, mass M, observer inclination i, spot colatitude θ_s, spot angular size ρ, and distance D for the equatorial spot with non-uniform i and θ_S priors (shown with a blue line). The one-dimensional distributions are normalized to give 1 as the total probability. The colors, contours, and other symbols are the same as in Figure 19.

4 Summary and Conclusions

We began this thesis by introducing the basic concept of neutron stars. The mass-radius relation of a neutron star was explained to be connected to the EOS of neutron stars and thus to the properties of the nucleonic matter inside them. We introduced also AMXPs, a subgroup of LMXBs, in which gas from a companion star is channeled onto the magnetic poles of a rapidly rotating neutron star. We examined few examples of AMXPs and their outbursts during which we have observed coherent oscillations at the spinning frequency of the pulsar.

The methods used to obtain constraints on mass-radius relation were shown next. We presented our pulse profile model that explains the observed coherent pulsations of AMXPs (whether or not they occur during an outburst). The special and general relativistic effects, including light bending, time delays, and the aberration of light, have been taken into account using the Schwarzschild-Doppler approximation. The oblate shape of the pulsar have been described with an empirical function, and the geometrical effects on the pulse profiles due to the oblateness have been taken care of.

We also presented our method to fit observed data to the profiles obtained from the waveform calculation. We used Bayesian inference and the ensemble sampler which is in our case the most efficient way to obtain the probability distributions for pulse profile parameters given the data. Instead of real observations we used synthetic data (discussed in the Results section) to inspect and calibrate our methods and tools. The synthetic data was generated using our pulse profile model, and it was made as similar as possible with the real data of SAX J$1808.4-3658$. The autocorrelation and burn-in phase of the samples were also discussed.

Finally, we obtained posterior probability distributions for the parameters with promising results. We were most interested in the probability distributions for mass and radius and how to get tighter constraints for those. We found that the ratio between mass and radius is found accurately, but to get independent knowledge of the mass and the radius

is remarkably easier if we have independent knowledge of the spot-colatitude θ_S and the observer inclination i. Without prior information, we were only able to get upper limits. Using non-uniform priors for θ_S and i, we got both lower and upper limits. Prior information to these angles is expected to be obtained from the X-ray polarization measurements of AMXPs.

The next step to the work of this thesis is to add a few more free parameters to the ensemble sampler and see if the results still remain the same. The pulse profile model may also be developed further e.g., to take into account both the blackbody and power-law parts of the spectrum. Finally, we are able to fit the real observations and hopefully get new constraints to the possible EOSs of neutron star.

References

[1] J. M. Lattimer and M. Prakash, Phys.Rept., 442, 109-165 (2007).

[2] K. Hebeler, J. M. Lattimer, C. J. Pethick and A. Schwenk, Phys.Rev.Lett., 105, 161102 (2010).

[3] J. M. Lattimer and M. Prakash, Science, 304, 536-542 (2004).

[4] C. Rhoades and R. Ruffini, Phys. Rev. Lett., 32, 324 (1974).

[5] J. M. Lattimer, Annu. Rev. Nucl. Part. Sci., 62, 485 (2013).

[6] R. C. Tolman, Proceedings of the National Academy of Science, 20, 169 (1934).

[7] J. R. Oppenheimer and G. M. Volkoff, Physical Review, 55, 374 (1939).

[8] A. Akmal, V. R. Pandharipande and D. G. Ravenhall, Phys.Rev., C58, 1804-1828 (1998).

[9] J. M. Lattimer and M. Prakash, Astrophys.J., 550, 426-442 (2001).

[10] P. Demorest *et al.*, Nature, 467, 7319, 1081-1083 (2010).

[11] B. Link, R. I. Epstein and J. M. Lattimer, Phys.Rev.Lett., 83, 17, 3362-3365 (1999).

[12] T. Gold, Nature, 218, 5143, 731-732 (1968).

[13] A. Patruno and A. L. Watts, ArXiv e-prints (2012).

[14] M. M. Romanova, G. V. Ustyugova, A. V. Koldoba and R. V. E. Lovelace, Astrophys.J., 610, 920-932 (2004).

[15] J. Frank, A. R. King and D. J. Raine, *Accretion Power in Astrophysics* (Cambridge University Press, 1985).

[16] J. J. M. in 't Zand *et al.*, Astron.Astrophys., 331, L25-L28 (1998).

[17] R. Wijnands and M. van der Klis, Nature, 394, 6691, 344-346 (1998).

[18] J.-P. Lasota, New Astronomy Reviews, 45, 7, 449-508 (2001).

[19] A. Sanna *et al.*, MNRAS, 459, 2, 1340-1349 (2016).

[20] M. Gierli´nski, C. Done and D. Barret, MNRAS, 331, 141 (2002).

[21] M. C. Miller and F. K. Lamb, Astrophys.J., 499, L37 (1998).

[22] J. Poutanen and M. Gierlinski, MNRAS, 343, 1301 (2003).

[23] K. R. Pechenick, C. Ftaclas and J. M. Cohen, APJ, 274, 846-857 (1983).

[24] M. AlGendy and S. M. Morsink, The Astrophysical Journal, 791, 2, 78, 11 (2014).

[25] S. M. Morsink, D. A. Leahy, C. Cadeau and J. Braga, Astrophys.J., 663, 1244-1251 (2007).

[26] J. Poutanen, Adv.SpaceRes., 38, 2697-2703 (2006).

[27] C. Cadeau, S. M. Morsink, D. Leahy and S. S. Campbell, Astrophys.J., 654, 458-469 (2006).

[28] J. Poutanen and A. M. Beloborodov, MNRAS, 373, 836-844 (2006).

[29] A. M. Beloborodov, Astrophys.J., 566, L85-L88 (2002).

[30] C. Misner, K. Thorne and J. Wheeler, *Gravitation* (W. Freeman, 1973).

[31] J. Terrell, Phys. Rev., 116, 1041 (1959).

[32] A. P. Lightman, W. H. Press, R. H. Price and S. A. Teukolsky, *Problem book in relativity and gravitation* (Princeton University Press, 1975).

[33] K. R. Lind and R. D. Blandford, APJ, 295, 358-367 (1985).

[34] G. B. Rybicki and A. P. Lightman, *Radiative Processes in Astrophysics* (Wiley-Interscience, 1979).

[35] S. Falkner, Master's thesis (Friedrich-Alexander University of Erlangen-Nuremberg) (2002).

[36] M. H. Kalos and P. A. Whitlock, *Monte Carlo methods* (WILEY-VCHWeinheim, 2008).

[37] K.-H. Lo, M. C. Miller, S. Bhattacharyya and F. K. Lamb, (2013).

[38] M. C. Miller and F. K. Lamb, Astrophys.J., 808, 1, 31, 20 (2015).

[39] J. Nättilä *et al.*, A&A, 591, A25 (2016).

[40] N. Metropolis *et al.*, J. Chem. Phys. **21**, 1087 (1953).

[41] W. K. Hastings, Biometrika **57**, 97 (1970).

[42] A. D. Kaiser, Master's thesis (New York University) (2013).

[43] M. Tuomi, Doctoral thesis (University of Turku) (2013).

[44] J. Goodman and J. Weare, Commun. Appl. Math. Comput. Sci., 5, 65-80 (2010).

[45] D. Foreman-Mackey, D. W. Hogg, D. Lang and J. Goodman, Publications of the Astronomical Society of Pacific, 125, 925, 306-312 (2013).

[46] K. Viironen and J. Poutanen, Astron.Astrophys., 426, 985-997 (2004).