Mass and radius constraints for neutron stars from pulse shape modeling

University of Turku
Department of Physics and Astronomy

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

Master's thesis, 42 s., 3 Appendices
Astronomy
September 2016

Write here!

Contents

Introduction

Rapidly rotating neutron stars, called pulsars, are the lighthouses of the universe. Pulses can be observed when an electromagnetic beam from this neutron star is emitted towards the Earth. 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 pulsars, called millisecond pulsars, have been detected in the radio, X-ray, and gamma ray bands 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 of the ultra-dense matter located inside neutron stars [1], [2]. Densities of this high of a cold matter is 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 oscillation produced by the rotation of some accreting millisecond pulsars. Models of these oscillations can be compared to observed waveforms. The mass and radius of neutron star have an effect to 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 tight 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 of a neutron star M 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 $n_0 \approx 0.16\text{fm}^{-3}$ of neutrons and protons found in laboratory nuclei [3]. Neutrons dominate the nucleonic component of neutron stars, but also some protons, electrons and muons exist. At the supernuclear densities also more exotic baryons, mesons or quarks may 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 excepted inside the star.

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

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

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, which means that the light cannot escape the object.

A more strict upper bound to the compactness (ratio between the radius and mass) 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

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

Neutron stars have also a minimum stable neutron star 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].

Generally the mass-radius (M-R) relation, is determined by the equations of hydrosta-

tic equilibrium. For a general relativistic and spherically symmetric (non-rotating) object these are the the so called Tolman-Oppenheimer-Volkov equations [6][7]:

$$\frac{dP}{dr} = -\frac{G[m(r) + 4\pi r^3 P/c^2](\rho + P/c^2)}{r[r - 2Gm(r)/c^2]},\tag{3}$$

and

$$\frac{dm(r)}{dr} = 4\pi\rho r^2,\tag{4}$$

where P and ρ are the pressure and mass-energy density, respectively, and $m(r)$ is the gravitational mass enclosed within a radius r. M-R relation can now be obtained if the relation between pressure and density $P = P(\rho)$ is known. This relation we call the equation of state (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]. The number of possible EOS can be reduced by observing which masses and radii the EOS should be able to produce.

1.1 Accreting millisecond pulsars

Rotating neutron stars were first detected as radio pulsars in 1967 [9]. After that, many different classes of pulsars have been discovered including 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 few milliseconds corresponding to the spin frequency of the neutron star. By definition, AMXPs are spin-

ning at frequencies $\nu \geq 100\,\text{Hz}$ and have weak surface magnetic fields ($B \sim 10^{8-9}$ G) [10].

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 Roche lobe. Both the neutron star and the companion have their own lobes that only join at a inner Lagrangian point L_1 [11]. In the 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 L_1. An accretion disk is formed because of the rotation of the system of two stars.

The first AMXP discovered was SAX J1808.4–3658. It is also used as a reference target in this thesis. The source was first found in 1996 by the Italian-Dutch BeppoSAX satellite [12]. The first coherent pulsations (at 401 Hz) were detected during the second outburst with NASA's Rossi X-ray Timing Explorer (RXTE) in 1998 [13]. It provided a 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) [10]. The idea is that a weakening magnetic field allows accretion to happen. The angular momentum of the accreting matter is transferred to the pulsar, which results in the increase of its spin frequency.

After the two first 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 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, they can in principle be partially explained with a so-called instability model.

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

for a two component-model, which includes a blackbody at lower energies and a hard Comptonization component at higher energies [14] . The heated hot spot on the neutron star surface is interpreted to produce the blackbody flux. The Comptonization is produced above the spot in a accretion shock. This shock is created as the plasma abruptly decelerates close to the neutron star surface at the bottom of the magnetic field lines.

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. The approach we use here 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 we use to get constraints for the parameters of the model.

2.1 Pulse profile modeling

The oscillations in the flux during the outbursts (the light curves) of accreting millisecond pulsars (described in the first chapter) can be modeled with different models. 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 account using so-called Schwarzschild-Doppler (S+D) approximation [17][18]. In the S+D approximation 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) [19]. 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 shape of the rapidly rotating neutron stars have also been taken account using an empirical formula for the oblate shape.

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

$$A = \frac{F_{\max} - F_{\min}}{F_{\max} - F_{\min}} \tag{5}$$

between 4 and 12 per cent [20], where $F_{\max}$ is the maximum observed flux and $F_{\min}$ is the minimum observed flux. Light bending tend to reduce the 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

deviate 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 polar caps (where the gas stream channeled by the neutron star 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 $\nu \geq 300\,\text{Hz}$ [21]. 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 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) [22]. It is also the model we use in this thesis. In that model (in geometric units where G = c = 1)

$$\frac{R(\theta)}{R_{\text{eq}}} = (1 + o_2(x,\bar{\Omega})\cos^2(\theta)), \tag{6}$$

where

$$o_2(x,\bar{\Omega}) = \bar{\Omega}^2(-0.788 + 1.030x), \tag{7}$$

$$x = \frac{M}{R_{\text{eq}}}, \tag{8}$$

$$\bar{\Omega} = \Omega(\frac{R_{\text{eq}}^3}{M})^{1/2}. \tag{9}$$

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. The equator radius is measured in the normal Schwarzschild coordinates as all the radii are in this thesis.

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 from the rotational pole (for a similar derivation see e.g. [23]). 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 an non-rotating observer near the star) is

$$\beta = \frac{v}{c} = \frac{2\pi R(\theta)}{c} \frac{\nu}{\sqrt{1-u}} \sin\theta = \beta_{\mathrm{eq}}(\theta)\sin\theta, \tag{10}$$

where β_{eq} would be constant if the star were spherical, $u \equiv r_{\mathrm{S}}/R(\theta)$, $r_{\mathrm{S}} = 2GM/c^2$ is the Schwarzschild radius, M is mass and $R(\theta)$ is radius of the star at spot location (given by equation 6). 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

$$\boldsymbol{r} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta), \tag{11}$$

that points to the spot from the center of the star (see Fig. 1). The rotational phase of the pulsar is $\phi = 2\pi\nu t$. The vector $\boldsymbol{r}$ is not usually directed perpendicular to the surface of the spot unless if the star is spherical. The vector that points normal to the surface is given by

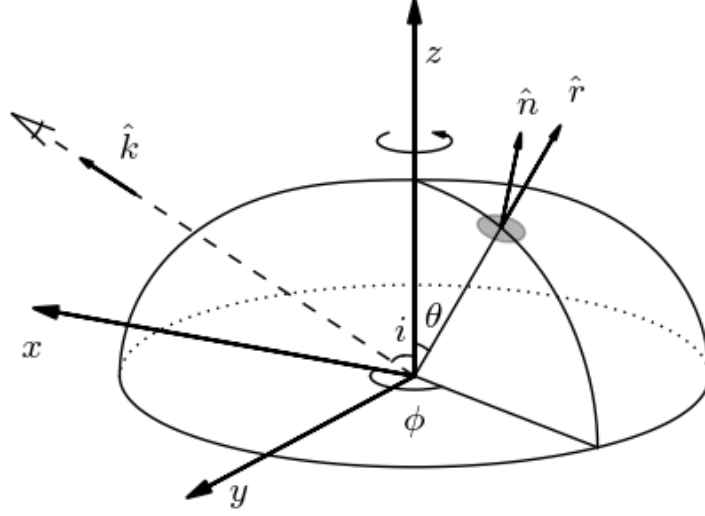$$\boldsymbol{n} = (\sin(\theta-\gamma)\cos\phi, \sin(\theta-\gamma)\sin\phi, \cos(\theta-\gamma)). \tag{12}$$

Fig. 1. Geometry of the problem. Created with https://github.com/natj/nsfig.

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

$$\cos\gamma = [1 + f^2(\theta)]^{-1/2}, \tag{13}$$

where

$$f(\theta) = \frac{1+z}{R}\frac{dR}{d\theta}. \tag{14}$$

In case of a spherical star $\frac{dR}{d\theta} = 0$ and thus $\boldsymbol{n} = \boldsymbol{r}$.

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 to the center of the star (e.g., the equator 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

$$\boldsymbol{k} = (\sin i, 0, \cos i), \tag{15}$$

where i is the inclination angle of the spin axis to the line of sight. Thus (see Fig. 1)

$$\cos\psi = \boldsymbol{k}\cdot\boldsymbol{r} = \cos i \ \cos\theta + \sin i \ \sin\theta \ \cos\phi, \tag{16}$$

The lensing 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

$$\cos \alpha = \boldsymbol{k}_0 \cdot \boldsymbol{r}. \tag{17}$$

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

$$\cos \sigma = \boldsymbol{k_0} \cdot \boldsymbol{n}. \tag{18}$$

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

$$\boldsymbol{k}_0 = [\sin \alpha \, \boldsymbol{k} + \sin(\psi - \alpha) \, \boldsymbol{r}] / \sin \psi. \tag{19}$$

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

$$\boldsymbol{k}_0 = \left(\cos \epsilon, \cos \xi, \cos \sigma \right), \tag{20}$$

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

$$\cos \xi = \frac{\boldsymbol{\beta}}{\beta} \cdot \boldsymbol{k}_0 = \frac{\sin \alpha}{\sin \psi} \frac{\boldsymbol{\beta}}{\beta} \cdot \boldsymbol{k} = -\frac{\sin \alpha}{\sin \psi} \sin i \, \sin \phi \,, \tag{21}$$

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

$$\boldsymbol{m} = (\cos(\theta - \gamma) \cos \phi, \cos(\theta - \gamma) \sin \phi, -\sin(\theta - \gamma)) \tag{22}$$

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

$$\cos \epsilon = \boldsymbol{m} \cdot \boldsymbol{k}_0 = \frac{\sin \alpha}{\sin \psi}(\sin i \cos(\theta - \gamma)\cos\phi - \cos i \sin(\theta - \gamma)) \\ + \frac{\sin(\psi - \alpha)}{\sin \psi}(\sin\theta \cos(\theta - \gamma) - \sin(\theta - \gamma)\cos\theta). \tag{23}$$

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:

$$\boldsymbol{k}_0' = \delta \begin{pmatrix} \cos \epsilon \\ \Gamma(\cos \xi - \beta) \\ \cos \sigma \end{pmatrix}, \tag{24}$$

where δ is the Doppler factor formulated as

$$\delta = \frac{1}{\Gamma(1 - \beta \cos \xi)}. \tag{25}$$

Using equation (24), we obtain

$$\cos \sigma' = \delta \ \cos \sigma, \tag{26}$$

Making use of equation (19) the zenith angle has the value

$$\cos \sigma = \boldsymbol{k_0} \cdot \boldsymbol{n} = \cos\alpha\cos\gamma + \sin\alpha\sin\gamma\cos\zeta = \\ \cos\alpha\cos\gamma + \frac{\sin\alpha}{\sin\psi}\sin\gamma(\cos i \sin\theta - \sin i \cos\theta\cos\phi), \tag{27}$$

where the spherical trigonometric identity $\cos i = \cos\theta\cos\psi + \sin\theta\sin\psi\cos\zeta$ (where ζ is the angle between the spot-observer and the spot-spin-axis planes) and the equation (16) are used. With small bending angles the approximation $\frac{\sin\alpha}{\sin\psi} \approx \sqrt{(1 - u)}$ may be used. This is based on the Beloborodov's approximation [24]

$$1 - \cos\alpha = (1 - \cos\psi)(1 - u). \tag{28}$$

The equation (27) is useful, since we need to know the emission angle relative to the spot normal (σ) in order to calculate fluxes and visibility conditions.

2.1.3 Light bending

Before we can get the angle σ from equation (27) we also first need to know how to get the angle α when we know the light bending angle ψ. The angles ψ we know (for every point at every phase) directly from the geometry, since we can calculate them using the equation (16). 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 [25]

$$\psi_p(R, \alpha) = \int_R^\infty \frac{dr}{r^2} \left[\frac{1}{b^2} - \frac{1}{r^2} \left(1 - \frac{r_{\mathrm{S}}}{r} \right) \right]^{-1/2}, \tag{29}$$

where b is the impact parameter,

$$b = \frac{R}{\sqrt{1-u}} \sin \alpha. \tag{30}$$

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

$$\psi(R, \alpha) = 2\psi_{\mathrm{max}} - \psi_p(R, \pi - \alpha), \tag{31}$$

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

$$p = -\frac{2}{\sqrt{3}} b \cos([\arccos(3\sqrt{3}r_{\mathrm{S}}/(2b)) + 2\pi]/3). \tag{32}$$

Numerical calculations using directly integral (29) lead to dramatic 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 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

$$\mathrm{d}F_E = I_E \, \mathrm{d}\Omega, \tag{33}$$

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 [19]

$$\mathrm{d}\Omega = b \, \mathrm{d}b \, \mathrm{d}\varphi / D^2, \tag{34}$$

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 $\mathrm{d}S = \delta \, \mathrm{d}S'$ [26][27] We also know the relation between σ and σ' from the relativistic aberration formula 26 (for motions parallel to the spot surface) $\cos\sigma' = \delta \, \cos\sigma$. Thus

$$\mathrm{d}S \, \cos\sigma = \mathrm{d}S' \, \cos\sigma', \tag{35}$$

which shows that the spot area projected on to the plane perpendicular to the photon direction is Lorentz invariant [27][28]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 [29]

$$\mathrm{d}S'(\theta) = R^2(\theta)[1 + f^2(\theta)]^{1/2} \sin\theta \mathrm{d}\theta \mathrm{d}\phi, \tag{36}$$

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 14). The solid angle measured at a distance D is then [29]

$$\mathrm{d}\Omega = \frac{\mathrm{d}S' \cos\sigma'}{D^2} \frac{1}{1-u} \frac{\mathrm{d}\cos\alpha}{\mathrm{d}\cos\psi}. \tag{37}$$

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 (33). However we still need to find out a expression for intensity I_E. Also the intensity is different in the static frame (of the distant observer) and in the local frame comoving with the spot due to both gravitational redshift and Doppler effect. The relation between monochromatic observed and local intensities is (see e.g. Misner et. al. 1973 [25] or Rybicki & Lightman 1979 [30])

$$I_E = \left(\frac{E}{E'}\right)^3 I'_{E'}(\sigma') \tag{38}$$

where $E/E' = \delta\sqrt{1-u}$. Here $I'_{E'}(\sigma')$ is the intensity computed in the corotating frame. The dependence 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 dependence on frequency goes to ν^4 and one gets

$$I = \left(\delta\sqrt{1-u}\right)^4 I'(\sigma'). \tag{39}$$

The observed spectral flux (eq. 33) is now given by

$$\mathrm{d}F_E = (1-u)^{1/2}\delta^4 I'_{E'}(\sigma') \cos\sigma \frac{\mathrm{d}\cos\alpha}{\mathrm{d}\cos\psi} \frac{\mathrm{d}S'}{D^2}, \tag{40}$$

where we have used the aberration formula (26). The bolometric flux reads:

$$\mathrm{d}F = (1 - u)\,\delta^5\,I'(\sigma')\cos\sigma \frac{\mathrm{d}\cos\alpha}{\mathrm{d}\cos\psi}\frac{\mathrm{d}S'}{D^2}. \tag{41}$$

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

$$I'_{E'}(\sigma') = I'_E(\sigma')\left(\delta\sqrt{1-u}\right)^{\Gamma-1}. \tag{42}$$

The observed spectral flux may be then expressed by

$$\mathrm{d}F_E = (1-u)^{\Gamma/2}\,\delta^{\Gamma+3}I'_E(\sigma')\cos\sigma\,\frac{\mathrm{d}\cos\alpha}{\mathrm{d}\cos\psi}\,\frac{\mathrm{d}S'}{D^2}. \tag{43}$$

We can notice that the equation for the bolometric flux (41) can be obtained as a special case of Eq. (43) by setting $\Gamma = 2$.

Let us look the expression (41) for 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 39) two comes from the solid angle transformation, one from the energy and one from the time contraction of photon arrival (Rybicki & Lightman 1979 [30]). In the case of spectral flux (equation 40) there is no Doppler factor from the energy and thus in total one power less.

As mentioned 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 multiplication of an angle-dependent ("beaming function") and an energy-dependent function. The beaming function describes what are directions where most of the intensity is emitted and the function might have e.g. a polynomial shape.

Thanks to the observed AMXPs (e.g. SAX J1808.4-3658 [18]) we know that the spectrum of the pulsar (i.e. the energy-dependence of $I'_{E'}(\sigma')$) might be fitted with a blackbody
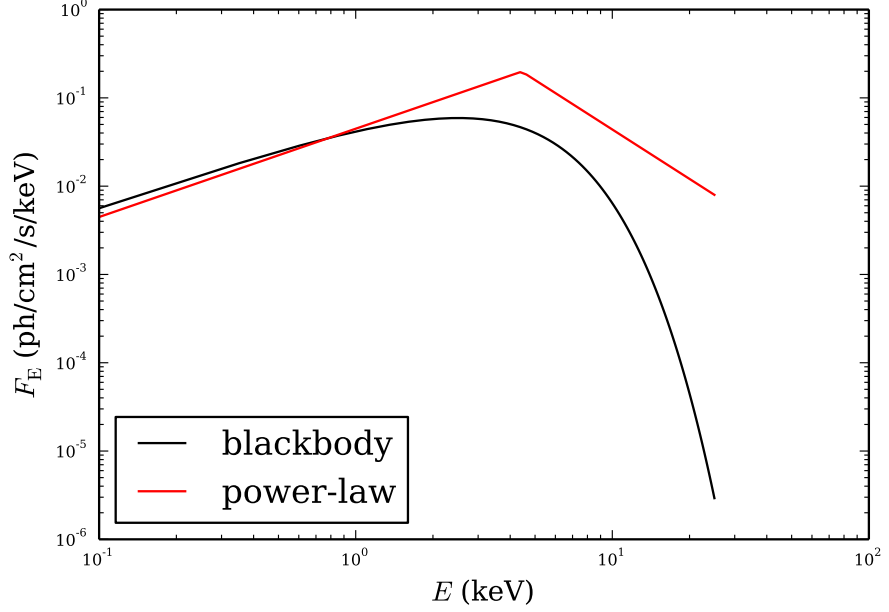
Fig. 2. The two-component power-law compared to a blackbody at 2 keV temperature. The spectra are calculated using the time-averaged fluxes at different energies of the pulse profile model with some physical parameters.

component at lower energies and with a power-law component at higher energies. However in the case of SAX J1808.4-3658 the blackbody contributes only about 30 per cent to the flux even in the 3-5 keV region (which are near the lowest observed energies) [18]. 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 a temperature of 2 keV (in the frame of the spot). Below the cut-off energy we use Rayleigh-Jeans law and above the cut-off a descending power-law. The power-law is normalized to be equal with the named blackbody at 1 keV. This model of spectrum is shown in the Figure 2.

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 and with e.g. a linear dependence on the beaming angle σ' (representing the accretion shock). The spots may have different shapes and sizes. This model will be added to the pulse profile code in the near future.

In the calculations of this thesis we use spectral number fluxes, since those are the fluxes we get from the observations. We must thus divide the fluxes we get from equation (40) with the observed energy E. When calculating the flux we make again a grid of values of the derivatives $\frac{\mathrm{d}\cos\psi}{\mathrm{d}\cos\alpha}$ (same as $(\frac{\mathrm{d}\cos\alpha}{\mathrm{d}\cos\psi})^{-1}$) with different α and R. The derivatives for each point are found by interpolating.

2.1.5 Time delays

The previous expressions (40), (41) and (43) for calculating the flux of an infinitesimal spot contain the pulsar phase ϕ (e.g. 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 between 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 in interest and a photon from a reference point (we call this reference photon). We choose the reference point to locate at a reference radius r_{ref} from the center of the star (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 (Pechenick et al. [19]):

$$c\Delta t_p(R,\alpha) = c\Delta t_s(R,\alpha) - \delta t(r_{\mathrm{ref}}, R),\qquad(44)$$

where

$$c\Delta t_s(R,\alpha) = \int_R^\infty \frac{\mathrm{d}r}{1 - r_{\mathrm{S}}/r}\left\{\left[1 - \frac{b^2}{r^2}\left(1 - \frac{r_{\mathrm{S}}}{r}\right)\right]^{-1/2} - 1\right\}\qquad(45)$$

is the time difference between photons originating from the same radius (R) and

$$\delta t(r_{\mathrm{ref}}, R) = R - r_{\mathrm{ref}} + r_{\mathrm{S}}\ln(\frac{R - r_{\mathrm{S}}}{r_{\mathrm{ref}} - r_{\mathrm{S}}})\qquad(46)$$

is the time difference between photons with $b = 0$ from R and r_{ref} [31].

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

$$
\begin{aligned}
c\Delta t(R, \alpha) = {}& 2c\Delta t_s(p, \pi/2) - c\Delta t_s(R, \pi - \alpha) \\
&+ 2\left[R - p + r_{\text{S}}\ln(\frac{R - r_{\text{S}}}{p - r_{\text{S}}})\right] - \delta t(r_{\text{ref}}, R).
\end{aligned}
\tag{47}
$$

The integral in equation (45) is not very easy to solve numerically. That is why we use again a variable substitution before using Simpson quadrature. We also make a grid of time delays corresponding to different α and R, which we later interpolate, 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 (29), (30) and (27) for a given pulsar phase ϕ. The corresponding delays $\Delta t(R, \alpha)$ or $\Delta t_p(R, \alpha)$ we can find using formulae (44) and (47). When these time delays are known we compute the delays in phase with

$$
\Delta\phi(\phi) = 2\pi\nu\Delta t[b(\phi)],
\tag{48}
$$

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

The flux at observed phase ϕ_{obs} is $F_{\text{obs}}(\phi_{\text{obs}}) = F(\phi - \Delta\phi)$ with phase delay $\Delta\phi = 2\pi\nu\Delta t$ calculated using (44), (47) and (48). 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 it was a good idea to 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 tricky 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 spherically symmetric spot (the angle measured from the spot center to the end of the spot is constant). 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 the case of spherical star we have checked that the new version gives the same results as the old one (which assumed spherical star and has been shown to give accurate results). In the case of oblate star we have e.g. compared our light curves to those obtained by Morsink [29] and Cadeau [21] (model called oblate S+D). The results are shown in Figures 3, 4 and 5. They are in good agreement except in cases of some very oblate stars and high bending angles. The profiles of oblate pulsars have also been compared to profiles obtained from light-ray tracing methods (Figures?), which correspond each other well.

2.2 Bayesian inference

The model presented in the previous section can be used to constrain parameters of neutron stars by using Bayesian methods. We can fit observed data to the model and use Markov Chain Monte Carlo (MCMC) methods to sample over the parameter space and find the most probable values for the parameters of the model [32]. The data can also be synt-
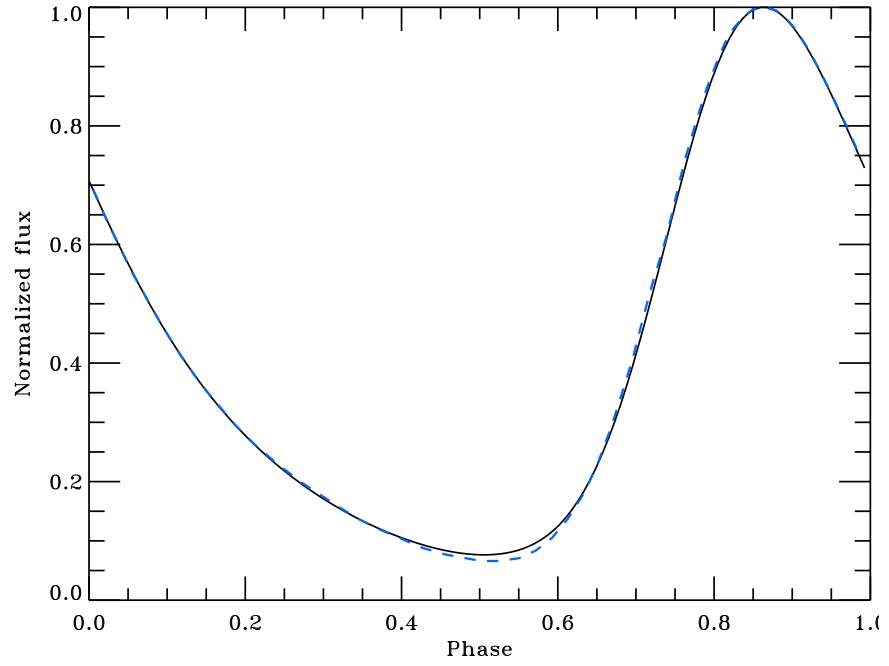
Fig. 3. The light curves (bolometric flux) for light emitted from one infinitesimal spot with parameters given in the Figure 3 of Morsink et. al. (2007) [29]. The blue dotted curve is the result of Morsink and the black solid line is our result.
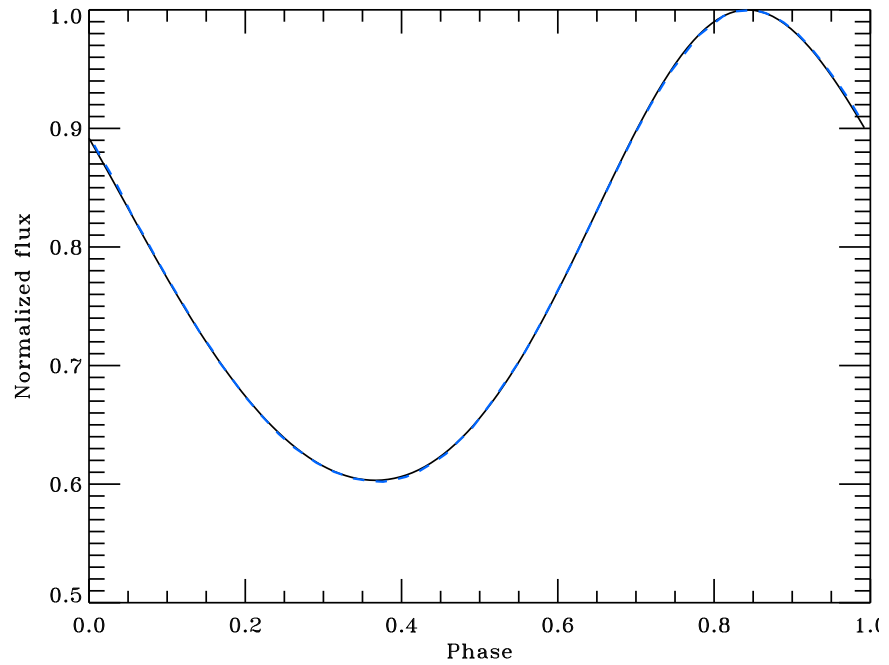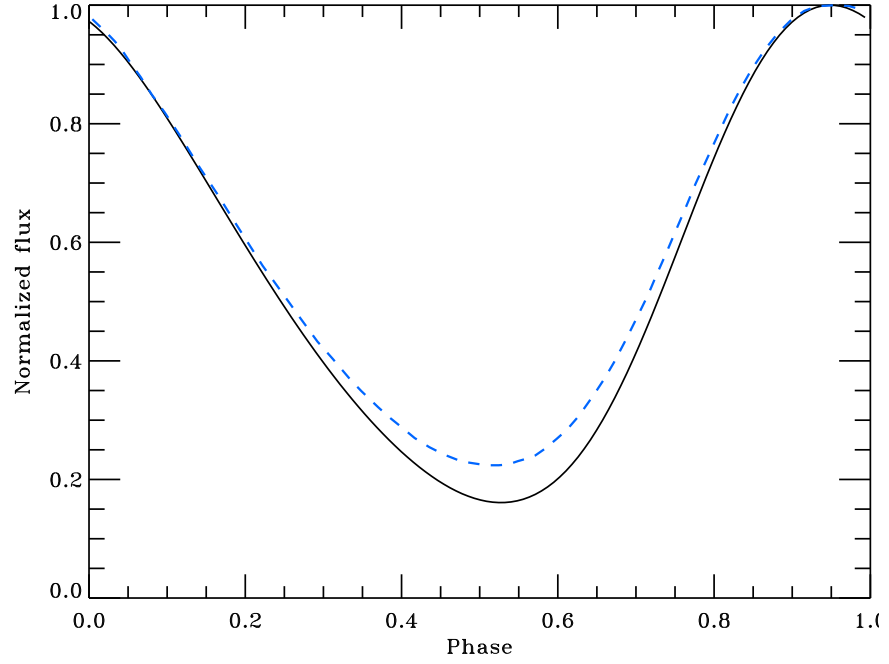

Fig. 4. The light curves (bolometric flux) for light emitted from one infinitesimal spot with parameters given in the Figure 3 of Cadeau et. al. (2006) [21]. The blue dotted curve is the result of Cadeau and the black solid line is our result.

Fig. 5. The light curves (bolometric flux) for light emitted from one infinitesimal spot with parameters given in the Figure 4 of Cadeau et. al. (2006) [21]. The blue dotted curve is the result of Cadeau and the black solid line is our result.

hetic (as in this thesis) in order to test if the sampler will actually find the same values that were used when generating the data. Previously the constraints of neutron star masses and radii with synthetic data have been studied e.g. by Ka Ho Lo et al, 2013 [33] (for future space missions). 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's theorem, which is presented in the next section. The standard Metropolis sampling method of the parameter space and a quite new ensemble sampler (used in this thesis) are discussed in the two following sections.

2.2.1 Bayesian analysis

We are interested in the probability distributions of the parameters $\mathbf{y}$ of the waveform model when the observed waveform is known. The probability is marked $p(\mathbf{y}|D)$, where D is the energy- and oscillation phase-resolved waveform data (synthetic in our case). According to the Bayes's theorem this (posterior) probability distribution can be obtained from the likelihood of the data, given the parameter values [15]:

$$p(\mathbf{y}|D) \propto p(D|\mathbf{y})p(\mathbf{y}). \tag{49}$$

In the previous equation $p(D|\mathbf{y})$ is the likelihood or the probability distribution of the data given the parameters. The next 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 account the information of polarization measurements and make the priors of inclination and spot colatitude non-uniform to see how or if it is improving the fits. 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 (Metropolis et al., 1953 [34]; Hastings, 1970 [35]). It is a MCMC method for obtaining a sequence of random samples from a probability distribution. It generates random samples by moving in a random walk; that is some sequence $X_1...X_t$. Metropolis algorithm, like all the other MCMC sampling methods, satisfy the so called Markov property. It means that the conditional probability distribution of X_{t+1}, given all past elements, is independent of all the other states but not the previous state [36]:

$$P(X_{t+1} = x|X_t...X_1) = P(X_{t+1} = x|X_t). \tag{50}$$

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 the 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 any case compared to that of the previous sample. The likelihood ratio α between the to states is calculated with the following equation [37]:

$$\alpha = \frac{p(Y|D)q(X_t|Y)}{p(X_t|D)q(Y|X_t)}. \tag{51}$$

The conditional probabilities given the data are calculated using the Bayesian formula (49). The ratio of these probabilities is multiplied with the ratio of proposal distributions (sometimes called transition kernels) so that the algorithm satisfy condition called detailed balance. It is an important property for proving the convergence of the chain. It states that a step from X_t must have the same probability as a step from Y to X_t. The proposal density $q(X_t|Y)$ describes the probability of a transition from Y j 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 $\alpha \geq 1$, but otherwise with a probability α (if α is greater than a random number 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 perhaps tens of thousands of the first iterations ("burn-in"), since the probability distribution has not yet converged in those. There

is no general theory for how long burn-on is required.

2.2.3 Autocorrelation and difficulties with Metropolis

Metropolis(-Hastings) algorithm has also some limitations. The samples are not independent (what we would them like 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) [38]:

$$\tau = \sum_{t=-\infty}^{\infty} \frac{C(t)}{C(0)}, \tag{52}$$

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

$$C(t) = \lim_{t' \to \infty} \text{Cov}(V(X_{t+t'}), V(X_t)). \tag{53}$$

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 a function sample as $V(X_t)$. If the weights with different lags are correlated strongly τ will be large. Of course most of the correlation is originated from samples close to each other. This can be seen by plotting the autocovariance as function of lag (see Figure 11 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/τ, where N is the total number of steps [36]. However especially with Metropolis τ is typically high (unless a 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 example of such is a two dimensional skewed probability distribution is given by (shown
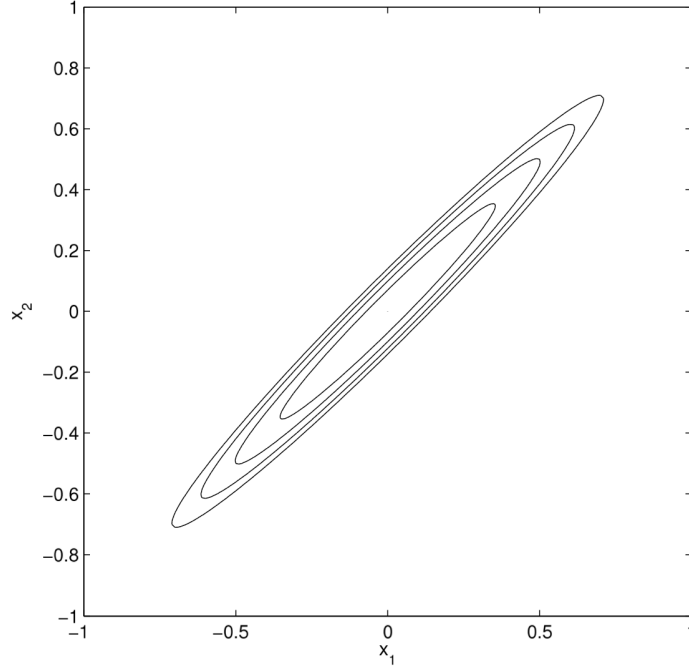
Fig. 6. Contours of the Gaussian density defined in expression 54. Figure 1 from Goodman and Weare (2010) [38].

in figure 6)

$$f(\vec{x}) = \exp(-\frac{(x_1 - x_2)^2}{2\epsilon} - \frac{(x_1 + x_2)^2}{2}). \tag{54}$$

Metropolis (like many other MCMC strategies too) would be forced to make perturbations of order $\sqrt{\epsilon}$ and would have slow equilibration. A better MCMC sampler would use larger perturbations in (-1,1) than in (1,1) direction. However the problem is worse in high dimensions, where it would be difficult to tune all the step sizes correctly. As we encountered this problem, we moved to 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 (Goodman and Weare, 2010 [38]). The algorithm has a similar structure to Metropolis, 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 of the other walkers in the ensemble. However each walker is independent of the other walkers. They are correlated only with the previous states of other walkers and their own previous state. After burn-in the distribution of each walker has 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 PDF $g(x) = Af(x) + b$ either directly or sample first f and then apply A and b after. 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 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 so called stretch-move. In a stretch-move each walker is moved using only one randomly selected complementary walker (see fig. 7). The move of current walker k is proposed along a line between the walker itself and the complementary walker j according to the distribution [39]

$$Y = X(j) + Z(X(k) - X(j)), \tag{55}$$

where the scaling variable Z satisfy the symmetry condition

$$g(1/z) = zg(z). \tag{56}$$

This ensures that the move (55) is symmetric in the usual informal way Metropolis is discussed (detailed balance is satisfied).
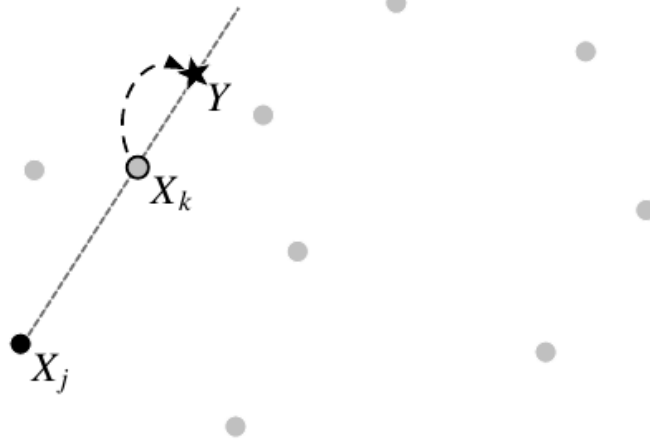
Fig. 7. A stretch move for updating the position of X_k based on the position of another random walker, X_j (subscript identifies here the walker instead of the step count). The light-gray walkers are other members of the ensemble. Figure 2 from Goodman and Weare (2010) [38].

When deciding whether the new sample is good enough, the Metropolis-Hastings selection rules are again applied. However instead of equation (51) the likelihood ratio is now calculated (for a given walker)

$$\alpha = Z^{N-1} \frac{p(Y|D)}{p(X_t|D)},$$ (57)

where N is now the number of parameters sampled [38].

3   Results

As mentioned already the aim of this work is to determine the probability distributions of the parameters in the pulse profile model (especially mass and radius), given the synthetic pulse profile, and study the effects of polarization measurements to these. Some details of the data, sampling and constraints obtained are presented in the three following sections.
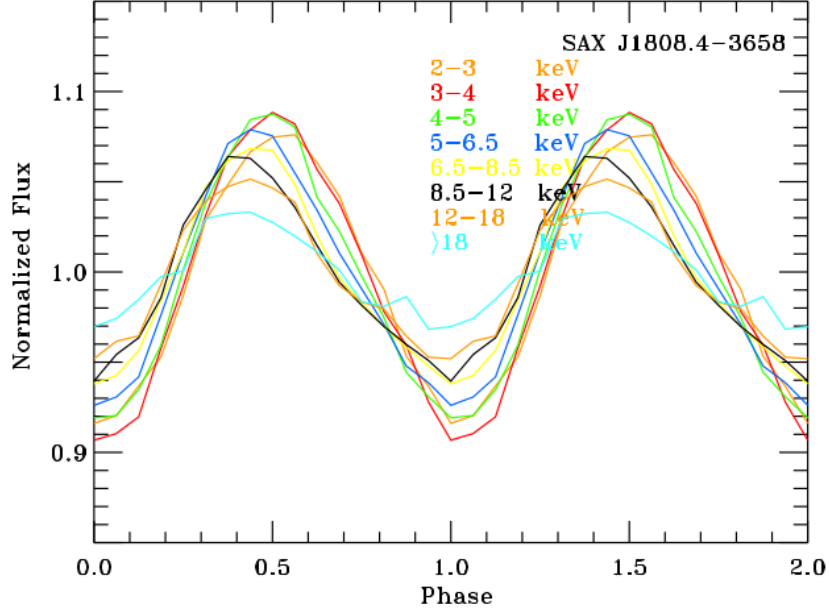
Fig. 8. Sax-data (Source?)

3.1 Synthetic data

The synthetic data in this work is made to resemble to real observations as closely as possible. The X-ray count data from SAX J1808.4-3658 is used as a reference. The amplitude of the modulation in the normalized number fluxes have the same order of magnitude. We use also the same energy intervals and same number of phase points as in the real data. The original data is shown as plotted in Figure 8.

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 the Table 1. The observed variability amplitude (see equation 5) is then obtained by choosing inclination i and the colatitude of the spot θ appropriately, since the amplitude can be approximated (neglecting the Doppler effect and using the Beloborodov's approximation [40])

$$A \approx \frac{(1 - r_{\mathrm{S}}/R)\sin i \sin \theta}{r_{\mathrm{S}}/R + (1 - r_{\mathrm{S}}/R)\cos i \cos \theta}. \tag{58}$$

We see that i and θ can be changed in the equation (58) without affecting the result. The exact calculation shows also only a small difference between these two parameters.

Table 1. Paramters of the synthetic datasets.

Parameter	Value
Radius R	12.0 km
Mass M	1.5 $\mathrm{M}_\odot$
Inclination i	5 $^\circ$ or 75 $^\circ$
Spot colatitude θ	75 $^\circ$ or 5 $^\circ$
Spot angular size ρ	10.0 $^\circ$
Distance D	2.5 kpc
Temperature t_{planck}	2.0 keV

Thus we have created two identical synthetic datasets which differ only in i and θ (we may call them datasets for polar and equatorial spots even though they are not exactly those). The normalized synthetic spectral fluxes are shown in the Figures 9 and 10. They have been created using the pulse profile model and adding random Gaussian noise to fluxes for every phase point. A larger noise was added to the highest energy channel. In our simulation we see how these two $i - \theta$ solutions can be separated from each other at least with help of the polarization information giving constraints to i to θ, and we also show their effect on the constraints of mass and radius.

3.2 Sampling methods

All the parameters presented in the Table 1 except temperature, are sampled using the ensemble sampler. There are also other parameters in the model, which are not sampled (at least not in this work yet). "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 spherically symmetric hot spot. The oblate shape of the star is described by the function in equation (6). Some parameters affecting on the accuracy of the pulse profile calculation have been changed after creating the synthetic data in order
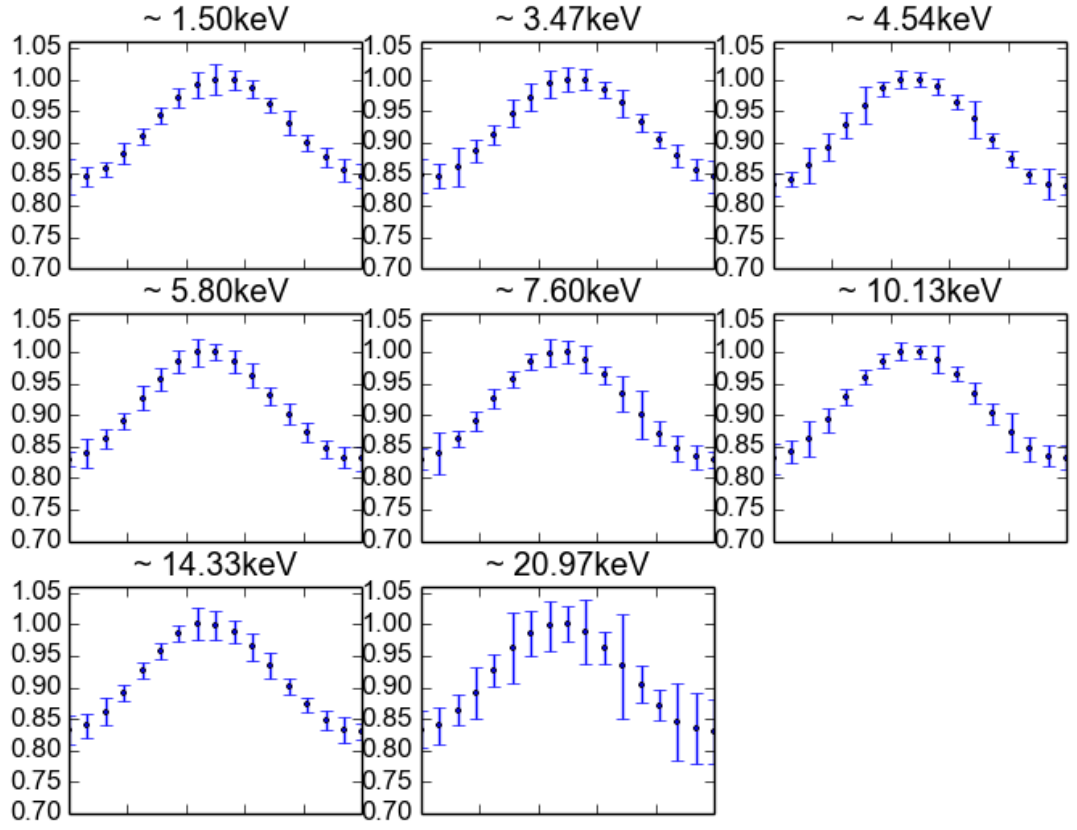
Fig. 9. Synthetic data: Normalized number fluxes as function of phase at different energies for $i = 75\deg$ and $\theta = 5\deg$. Other parameters are shown in the Table 1.
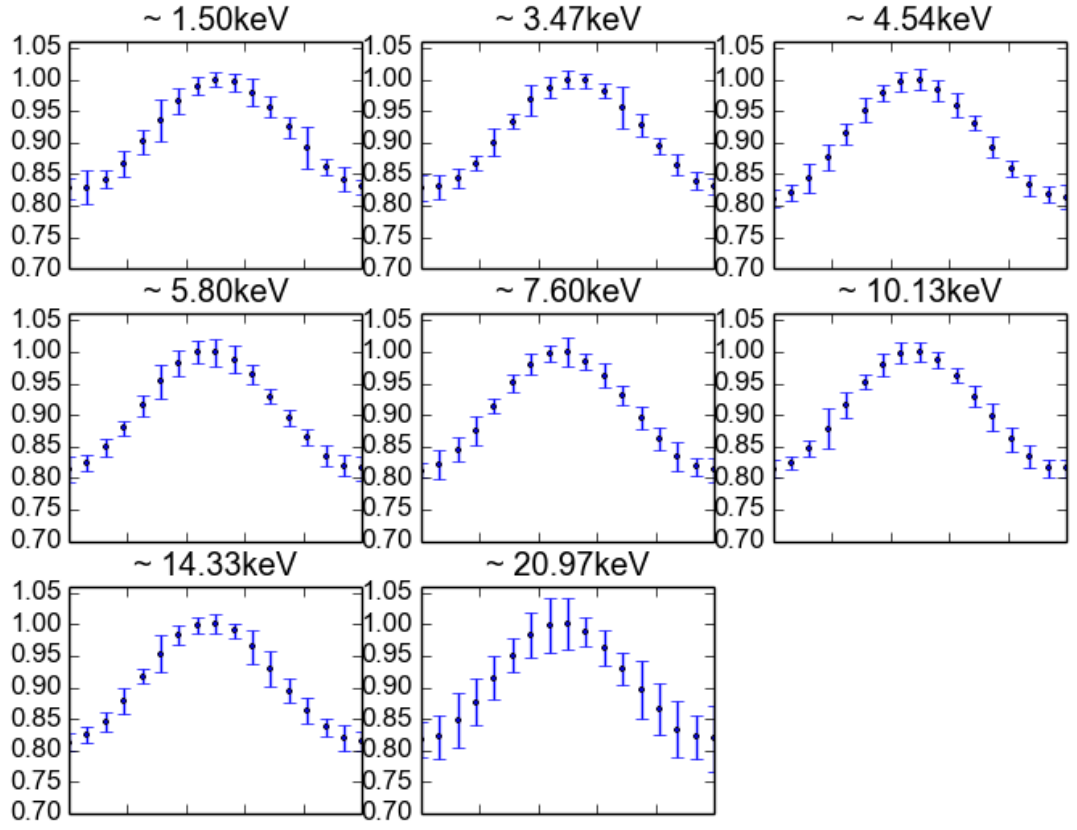
Fig. 10. Synthetic data: Normalized number fluxes as function of phase at different energies for $i = 5\,\deg$ and $\theta = 75\,\deg$. Other parameters are shown in the Table 1.

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 those. Thus each processor had its own ensemble (or also called 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 effects only on 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 and in the fitting procedure of the ensemble sampler.

The likelihood of the data for each sample ($p(D|\mathbf{y})$ in equation 49) 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 in the following equation):

$$p(F_{\text{data}}|F_{\text{model}}) = \exp(\frac{-(F_{\text{data}} - F_{\text{model}})^2}{2\sigma^2}). \tag{59}$$

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 multiplication of $p(F_{\text{data}}|F_{\text{model}})$ from each phase point and at each energy (or the sum of log-probabilities).

Of course we need also take to account that we have different phase shift in the data and in the model. For that reason we calculate the model with e.g. 5 times higher phase resolution and then calculate the probability densities using all the possible different phase shifts by shifting the values of fluxes in a correct way. The total probability $p(D|\mathbf{y})$ (which will be used to calculate likelihood ratio 57) is then 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 flat, so that equations (49) and (57) give

$$\alpha = Z^{N-1} \frac{p(D|Y)}{p(D|X_t)}.$$

(60)

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

$$\alpha = Z^{N-1} \frac{p(D|Y)p(Y)}{p(D|X_t)p(X_t)}.$$

(61)

It is not obvious what kind of prior probabilities one should use. It depends on how tight constraints we can get to the given parameter. Unlike in case of the usual flux, the measured polarization differ significantly if i and θ are switched [40]. Thus from modeling the polarization we can get the prior probability distributions to these parameters. In this thesis we choose to use normally distributed prior probabilities of i and θ around $0°$ or $90°$ to study the effect on the posterior probabilities of mass and radius. The standard deviation of that distribution is 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 than the sizes of the corresponding axes in the posterior probability plots 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. After that we have still left approximately 50 000 samples belonging to each ensemble. Figure (12) shows one example of 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).
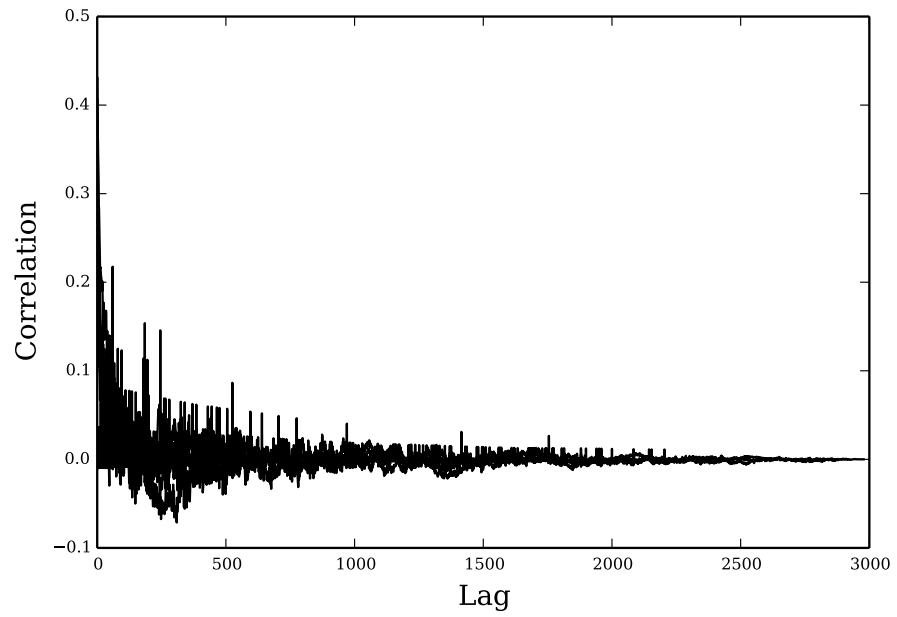
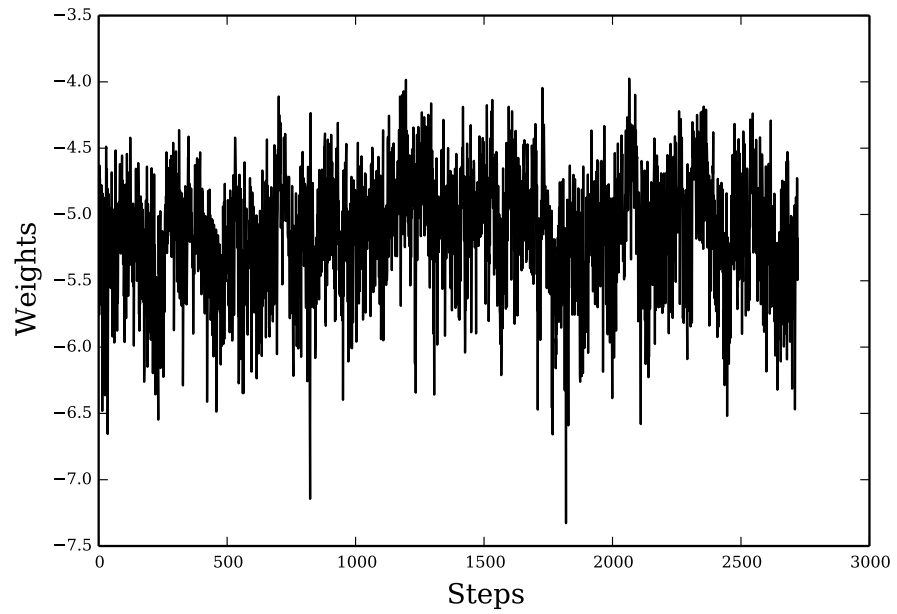Fig. 11. Autocorrelations of the all the ensembles of the polar spot with only flat priors.

Fig. 12. Log-likelihoods after burn-in removed in one ensemble.

The same averaged steps are also used in Figure 11, where we show an example of autocorrelation. In that Figure we plot autocorrelations of all six ensembles from one simulation using the weights as observables (see equation 53). The plot shows autocorrelation only approximatively, since one should make sure that the ensamble averaged weights are calculated for walkers being at a same state of their sequence. We may however see that in general correlation is not very bad, and we may reduce it using a thinning of...(?)

The results of our samplings are presented in the Figures 13-16. From these posterior probability distributions we can easily see how the prior information of angles θ and i improves the fits of the mass and radius of the neutron star (by comparing Figures 14 and 13 or Figures 16 and 15).

Without prior information we get the highest probabilities for the lowest masses and radii, since those correspond to more possible solutions in θ and i than the highest masses and radii. In this case we are only able to get upper limits to the mass and radius. The polar and equatorial spots give very similar results compared to each other as excepted (Figures 13 and 15). However they are not identical, which shows that switching θ and i has a significant (but not very large) effect on the light curves of the pulsar. It seems also slightly easier to find the correct θ-i solution in the case of equatorial spot (although in the one-dimensional histograms the small angles give always a higher peak in probability density because of the marginalization and the shape of two-dimensional distribution of θ-i).

In case of non-uniform priors for θ and i (Figures 14 and 16) we find correct solutions for θ and i even though our prior distributions are not centered around the correct values. The posterior probability distributions for the mass and radius are now also distributed on both sides of the correct values having the maximum near the correct value. Both lower and upper limits can now be determined.

The angular size of the spot ρ is the best-fitted parameter in our sampling. The highest probability is found to be very close to $10°$, 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). For that reason the size of the spot cannot change very much in order to keep to the observed count rate at a correct level.
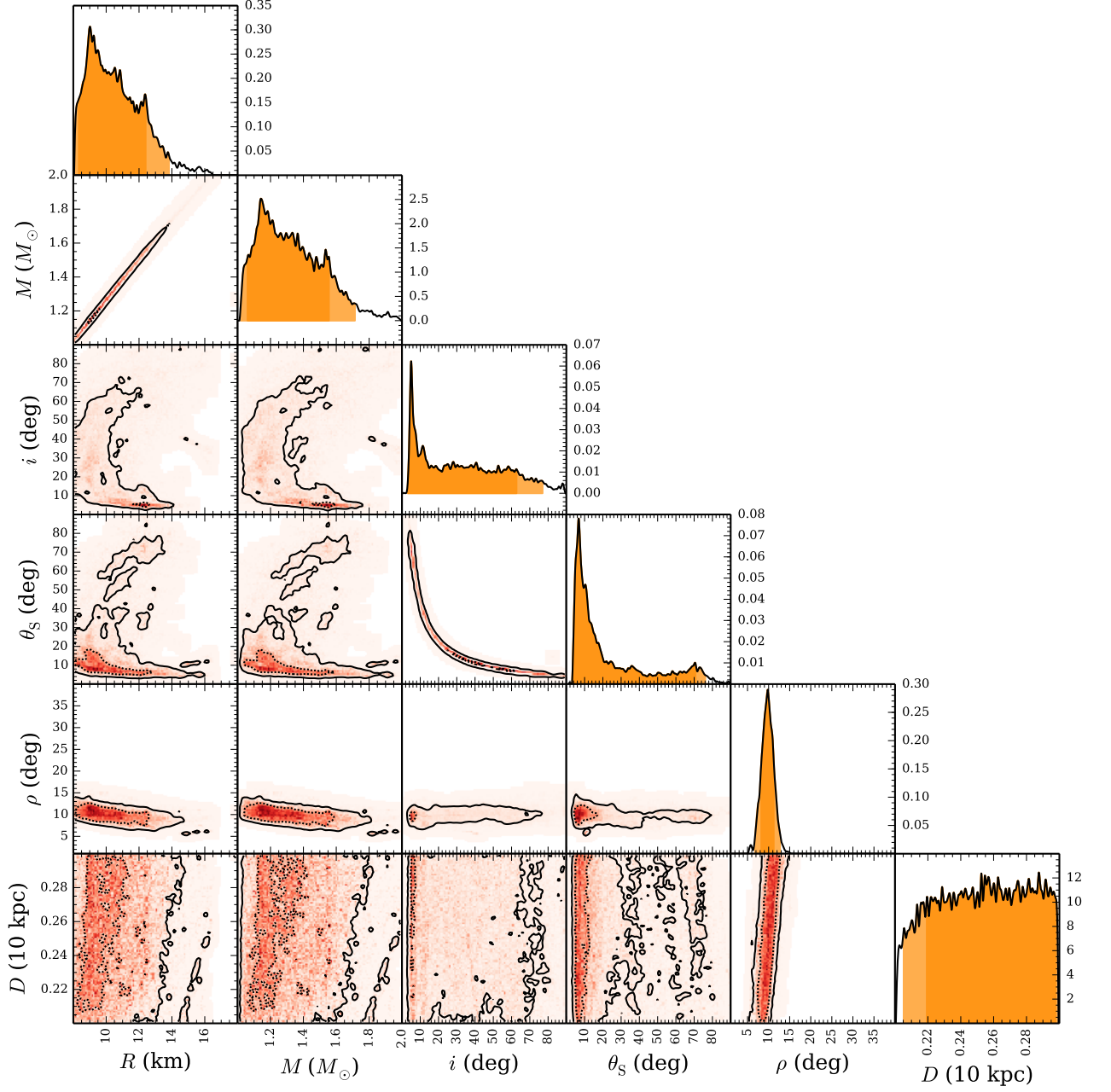
4 Summary and Conclusions

Fig. 13. Posterior probability distributions of the polar spot with only flat priors. The one-dimensional distributions are normalized to give 1 as the total probability.
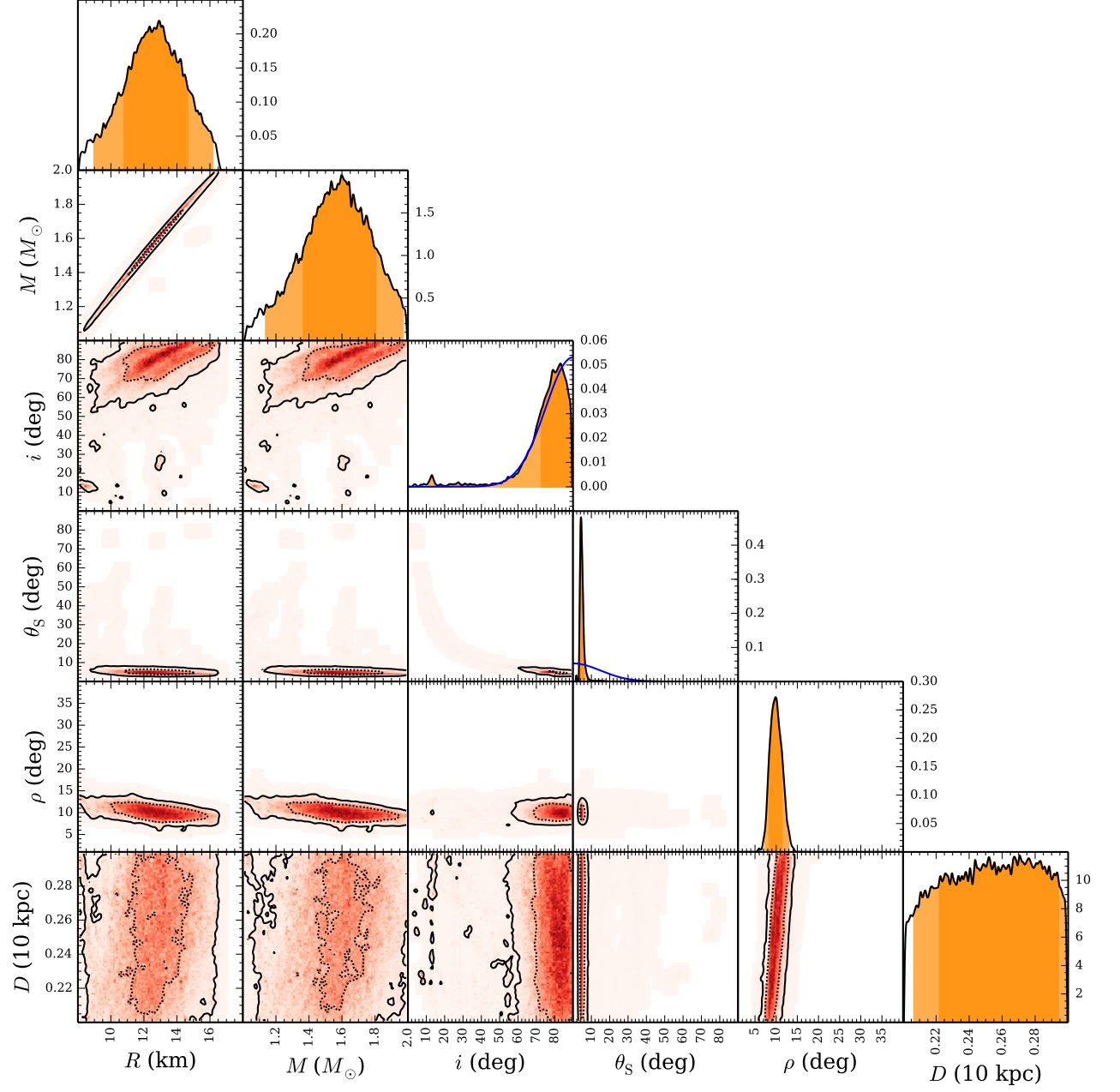
Fig. 14. Posterior probability distributions of the polar spot with non-uniform i and θ priors (shown with a blue line). The one-dimensional distributions are normalized to give 1 as the total probability.
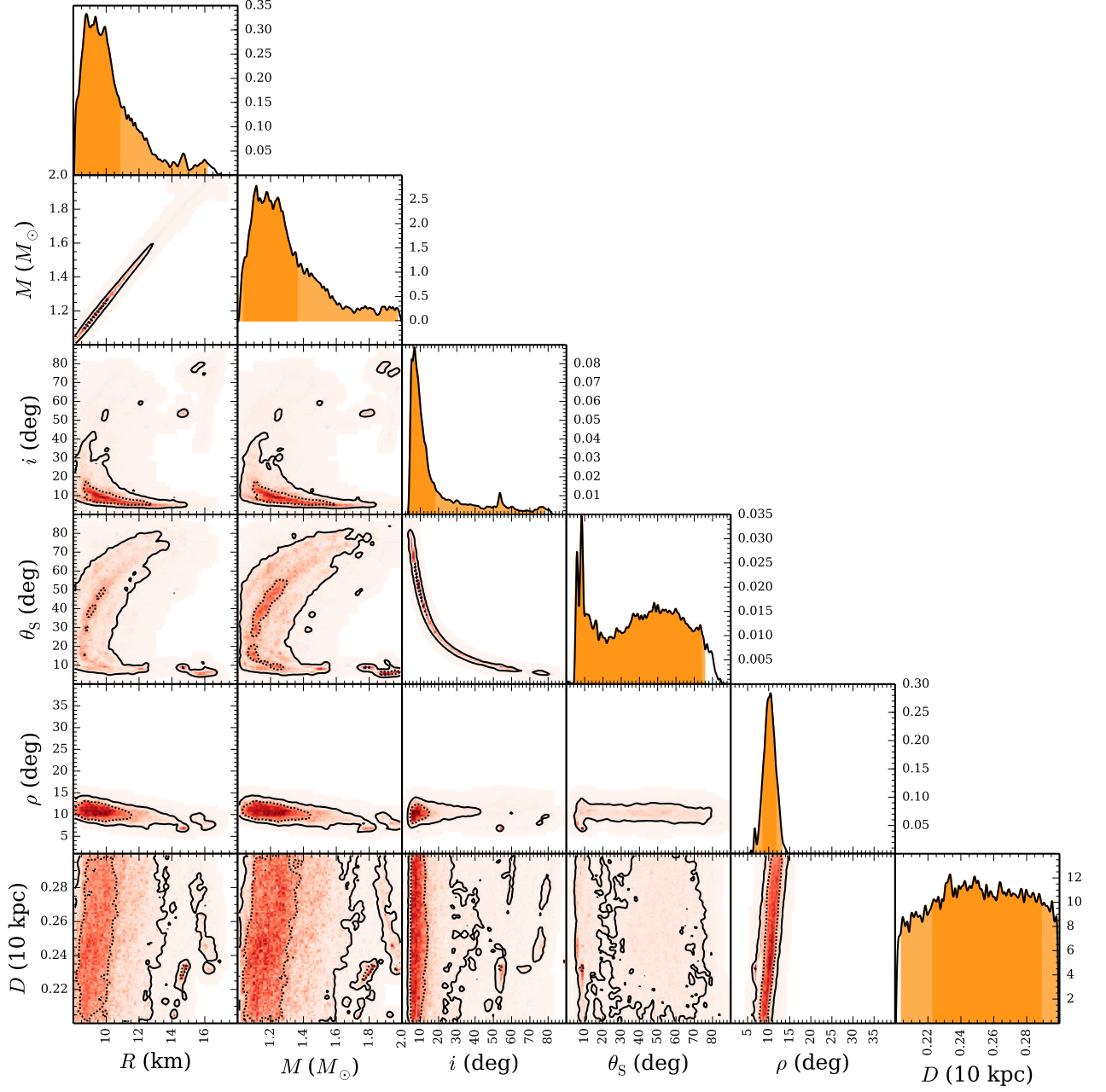
Fig. 15. Posterior probability distributions of the equatorial spot with only flat priors.The one-dimensional distributions are normalized to give 1 as the total probability.
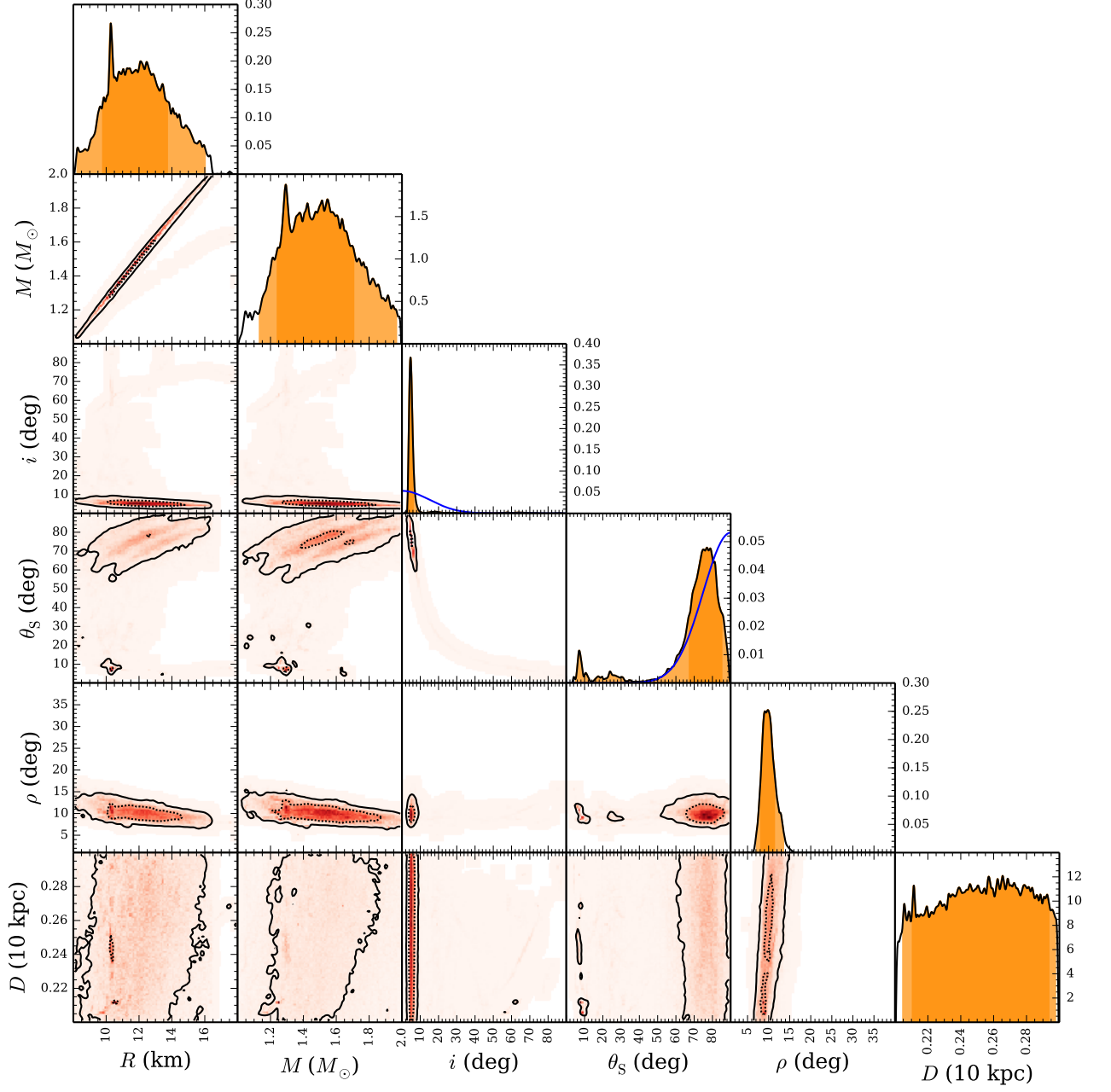
Fig. 16. Posterior probability distributions of the equatorial spot with non-uniform i and θ priors (shown with a blue line). The one-dimensional distributions are normalized to give 1 as the total probability.

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

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

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

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

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

[14] M. Gierliński, C. Done and D. Barret, MNRAS, 331, 141 (2002).

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

[16] A. Majczyna and J. Madej, Acta Astron., 55, 1-16 (2006).

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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