Mass and radius constraints for neutron stars from pulse shape modelling

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 a neutron star is emitted towards the Earth. This emission is thought be created when matter is accreted to a hot spot on a magnetic pole at the surface of the pulsar. These pulses or light curves 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 portions 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 neutron star astrophysics. This information could provide tight constraints on the equation of state of ultra-dense matter located inside the neutron star [1], [2]. This high densities of cold matter is otherwise unattainable. Studies of the light curves of pulsars can therefore help determine the properties of such matter. It is also not overstated to say that 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 burst observations of neutron stars. Models of burst oscillations waveforms (pulses during the burst) can be fit to observations of these 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 correct values for these parameters.

1 Neutron stars

Neutron stars are some of the densest and most 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 $n_0 \approx 0.16 fm^{-3}$ 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.

Neutron stars are created after the gravitational collapse of the core of a massive star (>$8M_\odot$) at the end of its life, which triggers a Type II supernova explosion. Too massive stars 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, constrains the possible mass and radius of neutron stars.

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

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

(2)

Neutron stars have also a minimum stable neutron star mass, which is about 0.1 $M_\odot$, although the neutron star's origin in a supernova gives a more realistic minimum [3].

Generally the mass-radius (M-R) relation, is determined by the equations of hydrostatic equilibrium. For a spherical object under hydrostatic equilibrium in general relativity these are the the so called Tolman-Oppenheimer-Volkov equations (Tolman, 1934 [5]; Oppenheimer & Volkoff, 1939 [6]):

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

(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. The relation we call the equation of state (EOS). For a realistic EOS the previous equations must be numerically solved to obtain M-R relation. These can be separated into three categories according to the compressibility of the matter: soft, moderate and stiff equations of state.

1.1 Accreting millisecond pulsars

Since the discovery of neutron stars as radio pulsars in 1967, many different classes have been discovered including low mass X-ray binaries (LMXBs). In LMXBs the neutron star accretes matter from a non-collapsed stellar companion (with mass M $\gtrsim M_\odot$) via an accretion disk. Accreting millisecond X-ray pulsars (AMXPs) is a subgroup of the LMXBs, in which the gas stripped from the companion is channeled out of the accretion disk onto the magnetic poles of the rotating neutron star. This gives rise to the X-ray pulsations at the spin frequency of the neutron star. By definition AMXPs are also spinning at frequencies $\nu \geq 100\,\mathrm{Hz}$ and have weak surface magnetic fields ($B \sim 10^{8-9}$) [7]. All known AMXPs have donor stars that transfer mass via Roche lobe overflow.

The first AMXP discovered was SAX J1808.4–3658. It is also the main observed target in this thesis. The source was first found in 1996 by the Italian-Dutch BeppoSAX satellite, but the first coherent pulsations (at 401 Hz) were detected during the second outburst with Rossi X-ray Timing Explorer (RXTE) in 1998. 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 slow neutron stars with high magnetic field ($B \sim 10^{12}$), into a rapidly spinning objects with a

relatively weak magnetic field ($B \sim 10^8$) [7].

SAX J1808.4–3658 went into outburst again in 2000, 2002, 2005, 2008 and 2011, with an approximate recurrence time of about 1.6-3.3 yr, and is the best sampled and studied of all AMXPs. The typical outburst can be split in 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 the disk instability model.

The X-ray spectrum of the outbursts have also been analysed and there is evidence for a two component-model: a blackbody at soft energies and a hard Comptonization component at higher energies [8] . The blackbody is interpreted as the heated hot spot on the neutron star surface and the Comptonization is produced in a accretion shock. This shock is created at the bottom of the magnetic field lines as the plasma abruptly decelerates close to the neutron star surface.

2 Methods

2.1 Pulse profile modelling

The oscillations in the flux during the outbursts (the light curves) of accreting millisecond pulsars (described in the first chapter) can be modelled with different models. The model presented here assumes that the radiation is originated from one or two polar spots locating at the polar caps of the stars. General and special relativistic effects have been taken account using so called Schwarzschild-Doppler (S+D) approximation (Miller, Lamb 1998, Poutanen, Gierlinski 2003) [9]. In the S+D approximation the gravitational light-bending effects are modelled as though the star is not rotating using the Schwarzschild metric and the formalism prescribed by Pechenick, Ftaclas and Cohen (1983) [10]. Rotational effects are added by introducing Doppler terms 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

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

between 4 and 12 per cent [11]. This amplitude depends only weakly on energy and the spectrum may be fitted with a two-component model (blackbody and Comptonization). The weak energy dependence 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 channelled 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 $\nu \geq 300\,\mathrm{Hz}$. The most important effect is purely geometrical: The directions that the light can be emitted into are different in these two cases. Thus there are certain spot locations on the star where the spot is visible if the surface is oblate but would be invisible if the surface where 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) [12]. It is also the model we use in this thesis. In that model (in geometric units where G = c = 1)

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

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 in this thesis.

2.1.2 Geometry

Besides the geometry of the star itself, we also need to know 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. [13]). The star is assumed to be rotating with frequency $\nu = P^{-1}$ (seen by a distant observer). The velocity of the spot in units of c as measured in the non-rotating frame is

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

where β_{eq} would be constant if the star was spherical, $u \equiv r_{\text{S}}/R(\theta)$, $r_{\text{S}} = 2GM/c^2$ is the Schwarzschild radius, M is mass and R 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 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 $\Gamma = (1 - \beta^2)^{-1/2}$.

Spot area measured in the corotating frame is $\mathrm{d}S'$, and its instantaneous position in the fixed lab frame is described by the unit vector

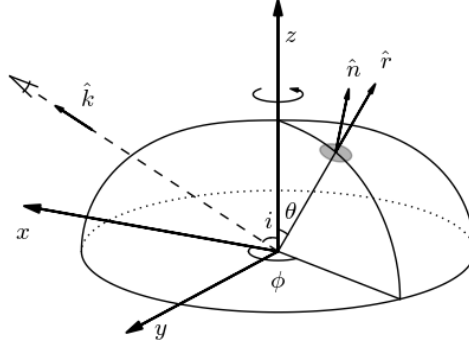$$\boldsymbol{r} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta), \tag{11}$$

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

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

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

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. The time is zero when this photon arrives at the observer.

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

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

Angle ψ measures the apparent inclination of the spot to the line of sight, which is different from the true inclination because of the light bending effect. We denote the initial direction of the emitted photon by $\boldsymbol{k}_0$ and the true emission angle (relative to the radius vector) 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

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

As a photon propagates to infinity, its direction changes from $\boldsymbol{k}_0$ near the stellar surface to $\boldsymbol{k}$ at 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

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

At any moment of time, we can introduce 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 to the spot. 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 $\boldsymbol{k}_0$:

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

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

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

and $\boldsymbol{k}_0$:

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

Emission angle in the corotating frame relative to the surface normal 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 the frame comoving with the spot (with y-axis along the spot motion, 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 is the Doppler factor

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

Using equation (24), we obtain

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

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

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

where the spherical trigonometric identity $\cos i = \cos \theta \cos \psi + \sin \theta \sin \psi \cos \delta$ (where δ is the angle between spot-observer and spot-spin-axis planes) and the equation (16) are used. With small bending angles the approximation $\frac{\sin \alpha}{\sin \psi} \approx \sqrt{(1 - u)}$ is used ([14]). 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 need first 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 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 [15]

$$\psi_p(R, \alpha) = \int_R^\infty \frac{dr}{r^2} \left[\frac{1}{b^2} - \frac{1}{r^2}\left(1 - \frac{r_S}{r}\right)\right]^{-1/2}, \tag{28}$$

where b is impact parameter,

$$b = \frac{R}{\sqrt{1-u}} \sin\alpha. \tag{29}$$

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

$$\psi(R, \alpha) = 2\psi_{\max} - \psi_p(R, \pi - \alpha), \tag{30}$$

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

$$p = -\frac{2}{\sqrt{3}}b \cos([\arccos(3\sqrt{3}r_s/(2b)) + 2\pi]/3). \tag{31}$$

Numerical calculations using directly integral 28 lead to dramatic errors, but after making a variable substitution $x = \sqrt{1 - R/r}$ the integral can be computed without problems using e.g. 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 values with different α and r and then find the correct α at different phases of the rotation of the star by interpolating.

The maximum bending angle corresponds to $\sigma = \pi/2$. Otherwise the photon would be directed across the star surface. The visibility of the spot is thus defined by the condition $\cos \sigma > 0$. In addition one should also check that the photon won't hit the surface of the star in any later phase of it's trajectory (it might be possible 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 the small spot. The observed flux from the spot at photon energy E is

$$\mathrm{d}F_E = I_E \, \mathrm{d}\Omega, \tag{32}$$

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

$$\mathrm{d}\Omega = b \, \mathrm{d}b \, \mathrm{d}\varphi / D^2, \tag{33}$$

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 on ψ only, but not on φ.

The apparent area of the spot as measured by photon beams in the non-rotating frame near the stellar surface is $\mathrm{d}S = \delta \, \mathrm{d}S'$ (see Terrell 1959; Lightman et al. 1975; Ghisellini 1999) and the relation between σ and σ' is described by the relativistic aberration formula (26) (for motions parallel to the spot surface). Thus the spot area projected on to the plane perpendicular to the photon propagation direction, i.e. a photon beam cross-section, is Lorentz invariant (see e.g. Lightman et al. 1975; Lind & Blandford 1985 [16]):

$$\mathrm{d}S \, \cos \sigma = \mathrm{d}S' \, \cos \sigma'. \tag{34}$$

The solid angle measured at infinity is then (Morsink et al. 2007 [17])

$$\mathrm{d}\Omega = \frac{\mathrm{d}S' \cos \sigma'}{D^2} \frac{1}{1 - u} \frac{\mathrm{d}\cos\alpha}{\mathrm{d}\cos\psi}. \tag{35}$$

In the limit of weak gravity $u \ll 1$, this gives the usual formula $\mathrm{d}\Omega = \mathrm{d}S' \cos\sigma'/D^2$.

The combined effect of the gravitational redshift and Doppler effect results in the following relation between the monochromatic observed and local intensities [15]

$$I_E = \left(\frac{E}{E'}\right)^3 I'_{E'}(\sigma') \tag{36}$$

where $E/E' = \delta\sqrt{1-u}$. Here $I'_{E'}(\sigma')$ is the intensity computed in the frame comoving with the spot. For the bolometric intensity, one gets

$$I = \left(\delta\sqrt{1-u}\right)^4 I'(\sigma'). \tag{37}$$

The observed spectral flux (eq. 32) now reads

$$F_E = (1-u)^{1/2}\delta^4 I'_{E'}(\sigma')\cos\sigma\frac{\mathrm{d}\cos\alpha}{\mathrm{d}\cos\psi}\frac{\mathrm{d}S'}{D^2}, \tag{38}$$

where we have used the aberration formula (26). The bolometric flux is given by:

$$F = (1-u)\,\delta^5\,I'(\sigma')\cos\sigma\frac{\mathrm{d}\cos\alpha}{\mathrm{d}\cos\psi}\frac{\mathrm{d}S'}{D^2}. \tag{39}$$

The radiation spectrum can be e.g. a black body spectrum or a power-law. If it can be represented by a power-law $I'_{E'}(\sigma') \propto E'^{-(\Gamma-1)}$ with a photon spectral index Γ which does not depend on the angle σ' then

$$I'_{E'}(\sigma') = I'_E(\sigma')\left(\delta\sqrt{1-u}\right)^{\Gamma-1}. \tag{40}$$

The observed spectral flux at a distance D from the star is then given by

$$F_E = (1-u)^{\Gamma/2}\,\delta^{\Gamma+3}I'_E(\sigma')\cos\sigma\frac{\mathrm{d}\cos\alpha}{\mathrm{d}\cos\psi}\frac{\mathrm{d}S'}{D^2}. \tag{41}$$

Expression for the bolometric flux (39) may be obtained as a special case of Eq. (41) by setting $\Gamma = 2$.

Thus, the bolometric flux from a rapidly rotating star differs by a factor δ^5 from that for a slowly rotating star. Two powers of δ come from the solid angle transformation, one from the energy, one from the photon arrival time contraction, and the fifth from the
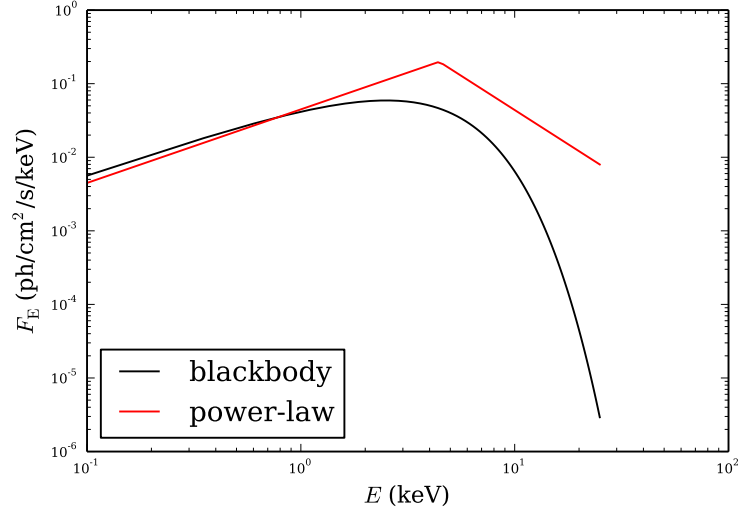
Fig. 2. The two-component power-law compared to a blackbody at 2 keV temperature. The spectra are calculated using the time-averaged fluxes at different energies of the pulse profile model with physical parameters.

change in the projected area due to aberration. Aberration may also change the specific intensity since it has to be computed for angle σ' in the comoving frame.

However in the most simple model we may assume that the intensity is independent of σ' ("isotropic beaming"). In a little more realistic case one can assume the intensity to be a multiplication of an angle-dependent ("beaming function") and an energy-dependent function.

Thanks to the observed AMXPs (e.g. SAX J1808.4-3658 [9]) we know that the spectrum of the pulsar (i.e. the energy-dependence of $I'_{E'}(\sigma')$) can be fitted with a blackbody 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) [9]. 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 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 (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 actual calculation of 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

Expressions (38), (39) and (41) contain the pulsar phase ϕ, but the flux actually corresponds to a different observed phase ϕ_{obs}, which is different from ϕ due to light travel delays. The delays become significant only for rapidly rotating pulsars. In Schwarzschild metric the maximum time delay for a neutron star of $M = 1.4\mathrm{M}_\odot$ is $\Delta t \sim 7 \times 10^{-2}$ ms (almost independent of compactness of the star M/R). This gives at most a 5 per cent correction to the arrival phase for a rotational period $P = 1.5$ ms.

The delay is caused by different travel times of emitted photons to the observer, depending on the position of the emitting spot. A photon following the trajectory with an impact parameter b (and $\alpha < \pi/2$) is lagging the photon originating from a reference radius r_{ref} (which can be chosen arbitrarily as long as it exceeds r_{S}) with $b = 0$ by [10]:

$$c\Delta t_p(R, \alpha) = c\Delta t_s(R, \alpha) - \delta t(r_{\mathrm{ref}}, R), \tag{42}$$

where

$$c\Delta t_s(R, \alpha) = \int_R^\infty \frac{\mathrm{d}r}{1 - r_{\mathrm{S}}/r} \left\{ \left[1 - \frac{b^2}{r^2}\left(1 - \frac{r_{\mathrm{S}}}{r}\right)\right]^{-1/2} - 1 \right\} \tag{43}$$

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

$$\delta t(r_{\mathrm{ref}}, R) = R - r_{\mathrm{ref}} + r_{\mathrm{S}}\ln(\frac{R - r_{\mathrm{S}}}{r_{\mathrm{ref}} - r_{\mathrm{S}}}) \tag{44}$$

is the time difference between photons with $b = 0$ from R and r_{ref} [18].

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

$$
\begin{aligned}
c\Delta t(R, \alpha) = {} & 2c\Delta t_s(p, \pi/2) - c\Delta t_s(R, \pi - \alpha) \\
& + 2\left[R - p + r_S \ln(\frac{R - r_S}{p - r_S})\right] - \delta t(r_{ref}, R).
\end{aligned}
\tag{45}
$$

The integral in equation 43 is again 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 for different α and r, which we later interpolate, to make the code run faster.

For a given pulsar phase ϕ, we compute angle ψ, then we find the corresponding emitted α, σ and the impact parameter using formulae (28), (29) and (27), and compute the corresponding delays $\Delta t(R, \alpha)$ or $\Delta t_p(R, \alpha)$ with equations (42) and (45). We then construct a one-to-one correspondence between the pulsar phase ϕ and the photon arrival phase to the observer $\phi_{obs} = \phi + \Delta\phi$, with the phase delays

$$
\Delta\phi(\phi) = 2\pi\nu\Delta t[b(\phi)].
\tag{46}
$$

The flux at observed phase ϕ_{obs} is $F_{obs}(\phi_{obs}) = F(\phi_{obs} - \Delta\phi)$ with phase delay $\Delta\phi = 2\pi\nu\Delta t$ computed using (42), (45) and (46). The effect of the photon arrival time contraction (or stretching) on the observed flux is already accounted for by one of the Doppler factors, so there is no need to multiply again flux by δ.

2.1.6 Profiles from a large spot

For a finite size spot, obviously, integration over the spot surface is required. The idea is to split the spot into a number of small sub-spots and compute the profiles from each sub-spot separately. One should be careful to include time delays correctly. For this problem it is actually good to compute the time delays relative to the photons emitted from the point directly under the observer. Integration over the spot surface is done using Gaussian quadrature in (cosine of) colatitude and trapezoidal rule for integration over the azimuth

inside the spot. The surface element of the spot is

$$\mathrm{d}S(\theta) = R^2(\theta)[1 + f^2(\theta)]^{1/2} \sin\theta \mathrm{d}\theta \mathrm{d}\phi, \tag{47}$$

where the factor $[1 + f^2(\theta)]^{1/2}$ takes care of the oblateness of the spot surface (function $f(\theta)$ is given in equation 14).

In this model we assumed a spherically symmetric spot (the angle measured from the spot center to the end of the spot is constant). The shape of the spot could also be modelled with a more complicated model. For a large spot (angular radius $\sim$30 deg) at least a few tens points in total are needed if one wants to achieve accuracy of the order of 1%.

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 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 [17] and Cadeau [19] (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 by 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 [20]. The data can also be synthetic (as in this Thesis) in order to test if the sampler will actually find the same values
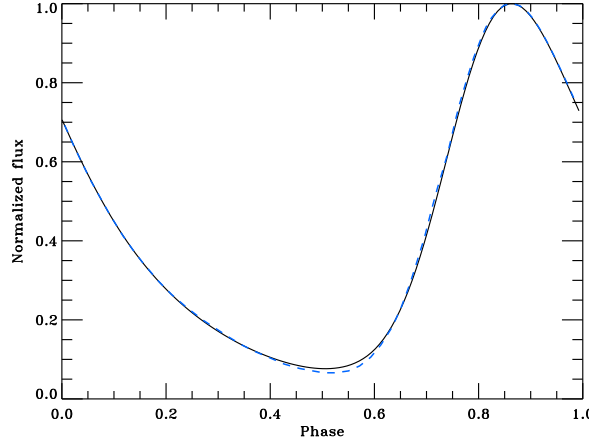
Fig. 3. The light curves for light emitted from one infinitesimal spot with parameters given in the Figure 3 of Morsink et. al. (2007) [17]. The blue dotted curve is the result of Morsink and the black solid line is our result.
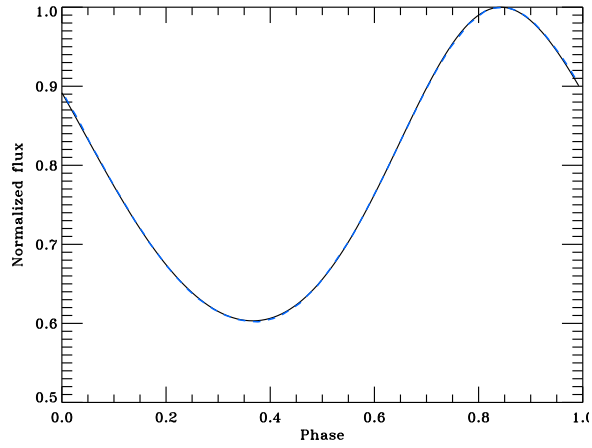

Fig. 4. The light curves for light emitted from one infinitesimal spot with parameters given in the Figure 3 of Cadeau et. al. (2006) [19]. The blue dotted curve is the result of Cadeau and the black solid line is our result.
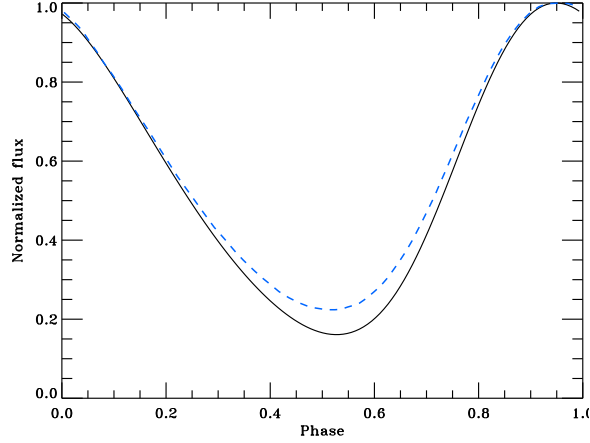
Fig. 5. The light curves for light emitted from one infinitesimal spot with parameters given in the Figure 4 of Cadeau et. al. (2006) [19]. The blue dotted curve is the result of Cadeau and the black solid line is our result.

that were used when creating 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 [21] (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 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 this Thesis). According to the Bayes's theorem this (posterior) probability distribution can be obtained

from the likelihood of the data, given the parameter values [22]:

$$p(\mathbf{y}|D) \propto p(D|\mathbf{y})p(\mathbf{y}). \tag{48}$$

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 (Metropo- lis et al., 1953 [23]; Hastings, 1970 [24]). 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 distribution of X_{t+1} given all past elements is independent of all but the previous state [25]:

$$P(X_{t+1} = x|X_t...X_1) = P(X_{t+1} = x|X_t). \tag{49}$$

A Markov chain is a random walk with Markov property.

The Metropolis algorithm works by taking an arbitrary move near the current point. 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 [26]:

$$\alpha = \frac{p(Y|D)q(X_t|Y)}{p(X_t|D)q(Y|X_t)}.$$ (50)

The conditional probabilities given the data are calculated using the Bayesian formula (48). 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 the new sample using again the same X_t.

It can be seen that the acceptance rule simplifies considerable 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 significantly when adding new members to the chain. Whether the obtained samples are statistically representative of the posterior can be verified if several 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 length of burn-in required.

Metropolis algorithm has also some limitations. The generated samples are not independent, since they are generated in a chain (concerns also other MCMC methods). The correlation between the samples can be measured by the autocorrelation time (ADD DEFINITION...and more).
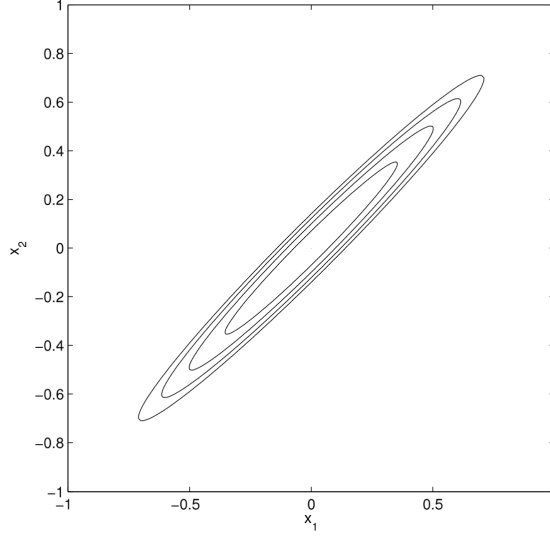
Fig. 6. Contours of the Gaussian density defined in expression 51. Figure 1 from Goodman and Weare (2010) [27].

One major disadvantage is that the Metropolis algorithm may work poorly for skewed, or anisotropic, distributions, depending on the proposal distribution. One example of such a two dimensional skewed probability distribution is given by (shown in Figure 6)

$$f(\vec{x}) = \exp(-\frac{(x_1 - x_2)^2}{2\epsilon} - \frac{(x_1 + x_2)^2}{2}). \tag{51}$$

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.

2.2.3 Ensemble sampler

The affine invariant ensemble sampler is a novel MCMC method developed a few years ago (Goodman and Weare, 2010 [27]). The algorithm has a similar structure to Metropolis, and still uses a proposal and 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 ensemble 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 and is not correlated with the previous states of other walkers (it is correlated only to its own previous state). After burn-in each walker is also distributed according 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 sampling a PDF g(x) = Af(x) + b is equivalent to sampling f then applying A and b after. This is the reason why 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 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 is stretch-move (we use it). In a stretch-move each walker is moved using only one randomly selected complementary walker (see Fig. 7). The move proposed alone a line between the current walker k and the complementary walker j according to the distribution [28]

$$Y = X(j) + Z(X(k) - X(j)),$$
(52)

where the scaling variable Z satisfy the symmetry condition

$$g(1/z) = zg(z).$$
(53)

This ensures that the move 52 is symmetric in the usual informal way Metropolis is discussed (detailed balance is satisfied).

When deciding whether the new sample is good enough, the Metropolis-Hastings selection rules are again applied. However instead of equation 50 the likelihood ratio is now
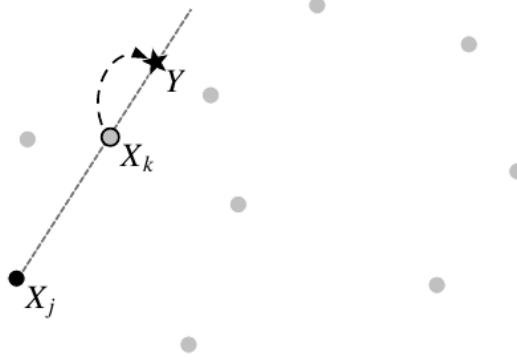
Fig. 7. A stretch move for updating the position of X_k based on the position of another random walker, X_j (subscript identifies here the walker instead of the step count). The light-gray walkers are other members of the ensemble. Figure 2 from Goodman and Weare (2010) [27].

calculated (for a given walker)

$$\alpha = Z^{N-1} \frac{p(Y|D)}{p(X_t|D)},$$ (54)

where N is the number of parameters sampled [27].

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 effect of polarization measurements to these. Some details of the data and constraints obtained are showed in the two following sections.

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 in Figure 8.
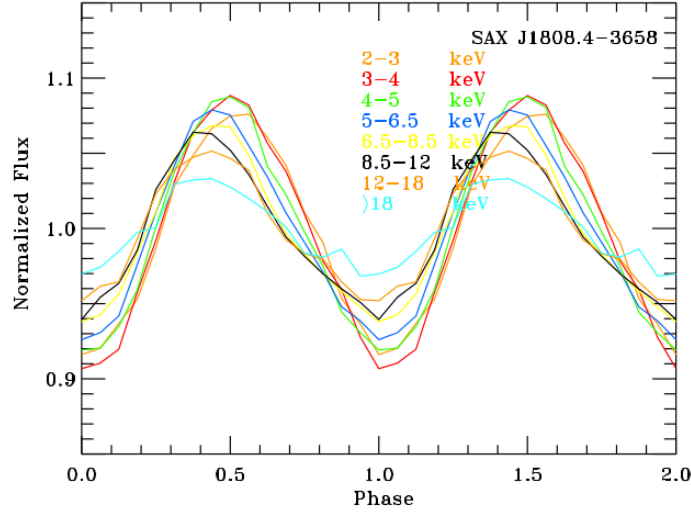
Fig. 8. Sax-data

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 [29])

$$A \approx \frac{(1 - r_{\mathrm{S}}/R) \sin i \sin \theta}{r_{\mathrm{S}}/R + (1 - r_{\mathrm{S}}/R) \cos i \cos \theta}. \tag{55}$$

We see that i and θ can be changed in the equation 55 without affecting the result. 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 (we may call them datasets for polar and equatorial spots even though they are not exactly that). These are shown in the Figures 9 and 10.They have been created using the pulse profile model and adding random Gaussian noise to fluxes of every phase point. In our simulation we see how these two solutions can be separated from each other at least with help of the polarization information giving constraints to i to θ, and we show their effect on the constraints of mass and radius.

Table 1. Paramters of the synthetic datasets.

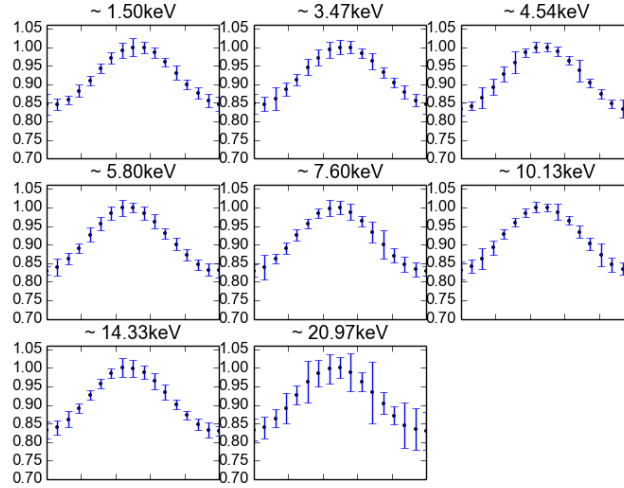| Parameter	Value
Radius R	12.0 km
Mass M	1.5 $M_\odot$
Inclination i	5 ° or 75 °
Spot colatitude θ	75 ° or 5 °
Spot angular size ρ	10.0 °
Distance D	2.5 kpc
Temperature t_{planck}	2.0 keV

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.
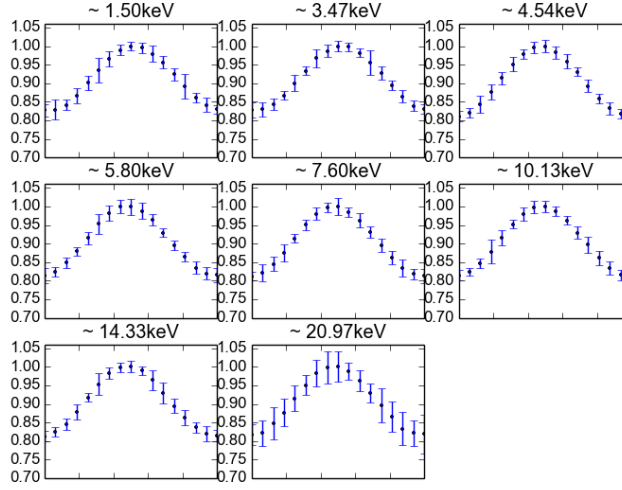
Fig. 10. Synthetic data: Normalized number fluxes as function of phase at different energies for $i = 5 \deg$ and $\theta = 75 \deg$. Other parameters are shown in the Table 1.

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 spherical 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 to get the sampling faster. 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 the possible computational power we have also made it possible to run the ensemble sampler code parallel. In this thesis we did simultaneously six independent simulations on six different processors and then combined the results of those. Thus each processor has its own ensemble (or 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 creation 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 48) 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 modelled fluxes F_{model} is normally distributed around the F_{model} with the known errors as standard deviation (σ in the following equation):

$$p(F_{\text{data}}|F_{\text{model}}) = \exp(\frac{-(F_{\text{data}} - F_{\text{model}})^2}{2\sigma^2}). \tag{56}$$

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 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 different phase shifts. The total probability $p(D|\mathbf{y})$ (which will be used to calculate likelihood ratio 54) 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 48 and 54 give

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

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:

$$\alpha = Z^{N-1}\frac{p(D|Y)p(Y)}{p(D|X_t)p(X_t)}. \tag{58}$$

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 normal flux, the

measured polarization differ significantly if i and θ are switched [29]. Thus from model-ling 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

The results of our samplings are presented in the Figures 11-14. 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 12 and 11 or Figures 14 and 13).

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 as excepted (Figures 11 and 13). 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 dominate 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 12 and 14) 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.

..Show also weights after burn-in removal and autocorrelations...TBD
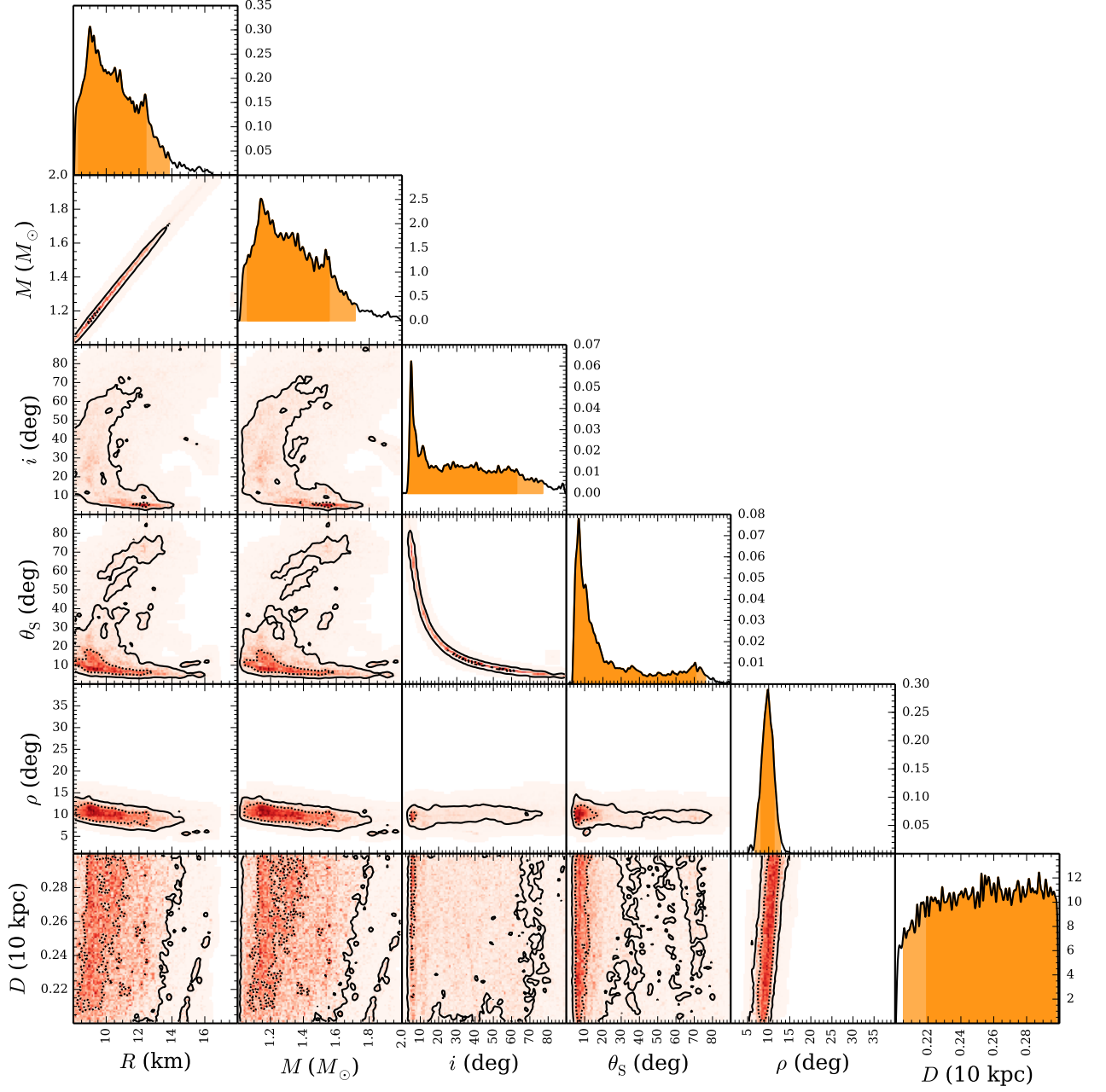
4 Summary and Conclusions

Fig. 11. Posterior probability distributions of the polar spot with only flat priors. The one-dimensional distributions are normalized to give 1 as the total probability.
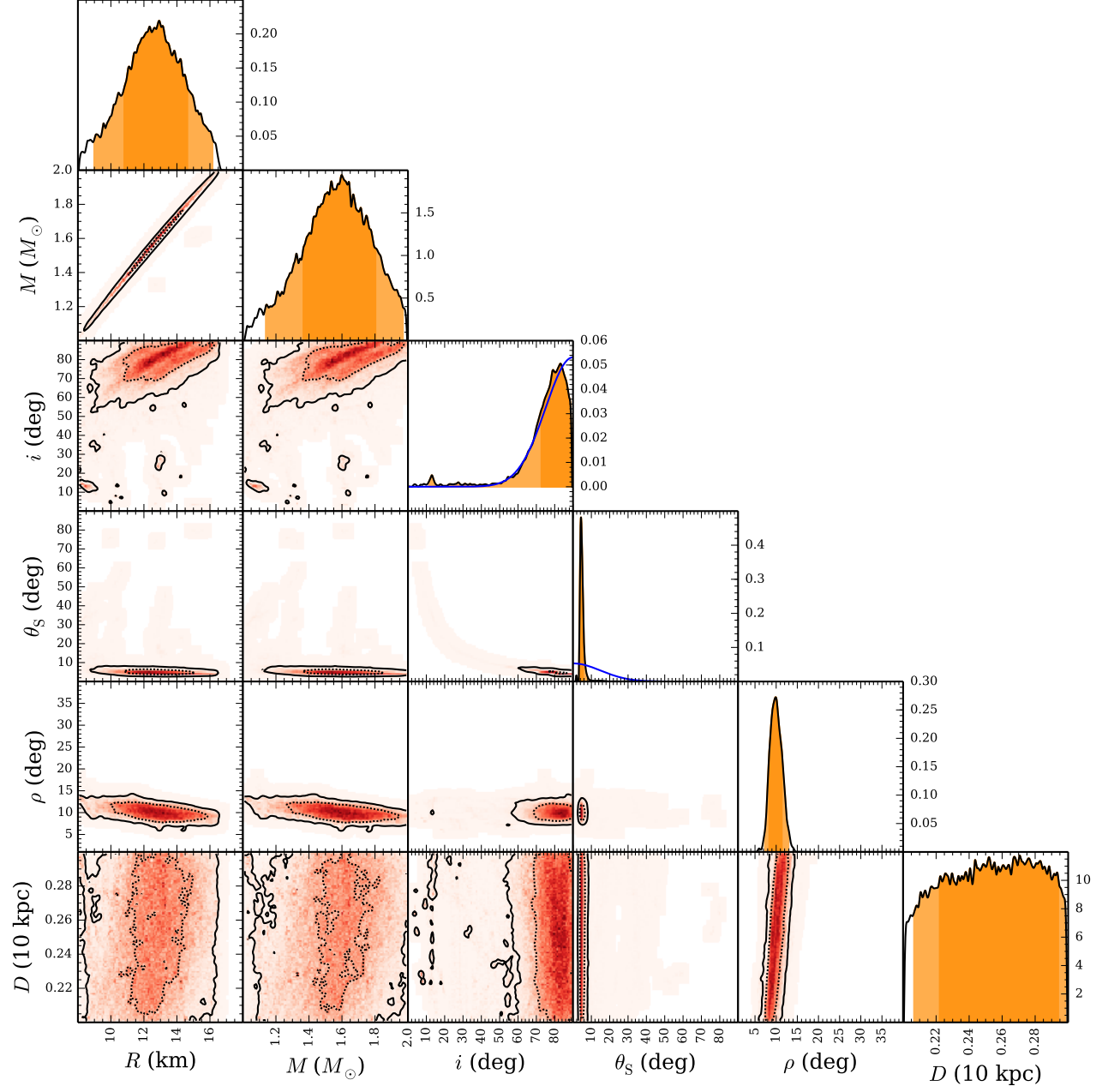
Fig. 12. Posterior probability distributions of the polar spot with non-uniform i and θ priors (shown with a blue line). The one-dimensional distributions are normalized to give 1 as the total probability.
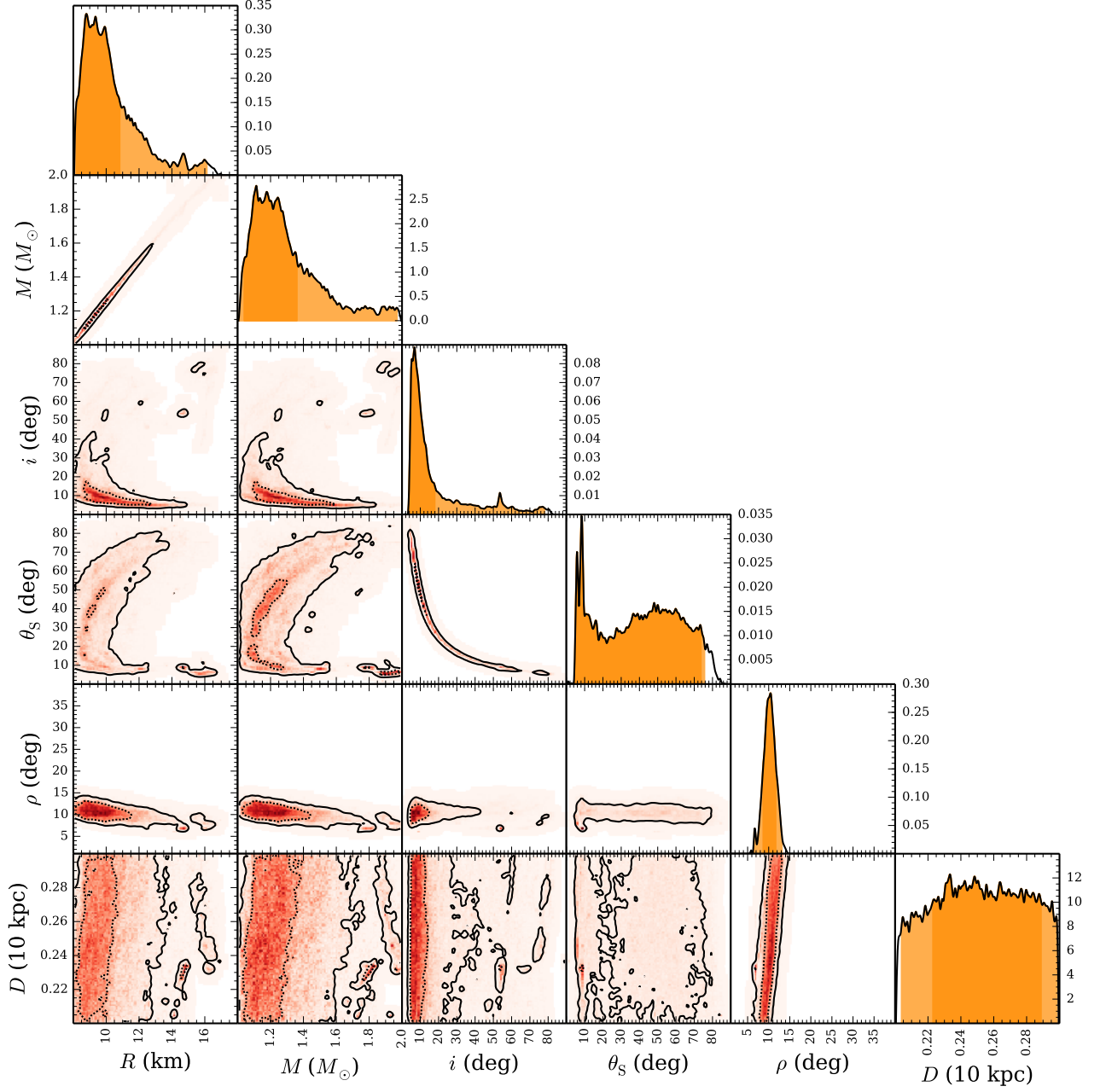
Fig. 13. Posterior probability distributions of the equatorial spot with only flat priors.The one-dimensional distributions are normalized to give 1 as the total probability.
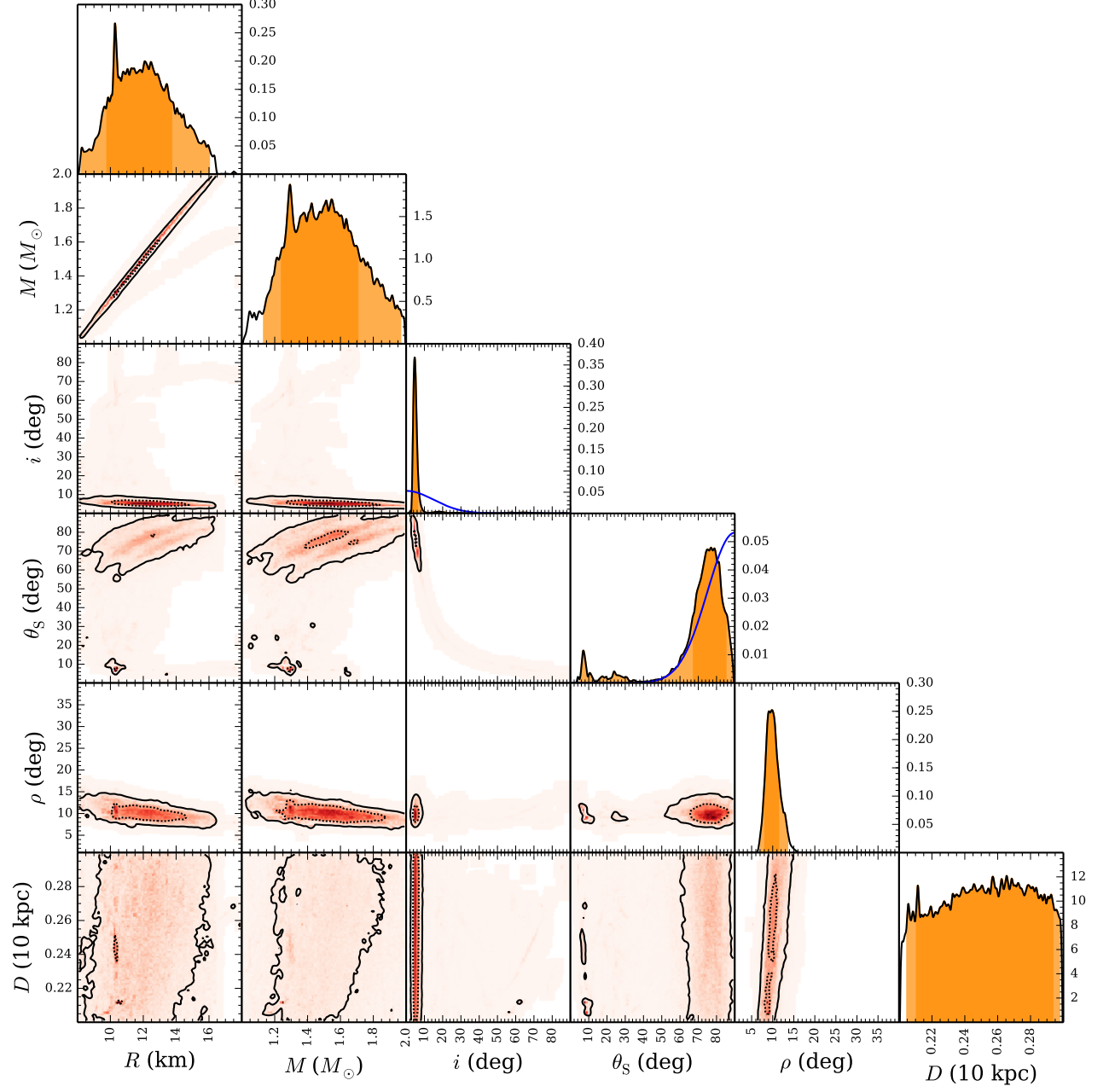
Fig. 14. Posterior probability distributions of the equatorial spot with non-uniform i and θ priors (shown with a blue line). The one-dimensional distributions are normalized to give 1 as the total probability.

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

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

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

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

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

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

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

[12] M. AlGendy and S. M. Morsink, Universality of the Acceleration Due to Gravity on the Surface of a Rapidly Rotating Neutron Star, 2014.

[13] J. Poutanen and A. M. Beloborodov, Mon.Not.Roy.Astron.Soc.373:836-844,2006 (2006).

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

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

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

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

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

[19] C. Cadeau, S. M. Morsink, D. Leahy and S. S. Campbell, Astrophys.J.654:458-469,2007 (2006).

[20] M. H. Kalos and P. A. Whitlock, *Monte Carlo methods* (WILEY-VCHWeinheim, 2008), pp. XII, 203 S.

[21] K.-H. Lo, M. C. Miller, S. Bhattacharyya and F. K. Lamb, Determining neutron star masses and radii using energy-resolved waveforms of X-ray burst oscillations, 2013.

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

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

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

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

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

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

[28] D. Foreman-Mackey, D. W. Hogg, D. Lang and J. Goodman, emcee: The MCMC Hammer, 2013.

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