Bayes-X: a Bayesian inference tool for the analysis of X-ray observations of galaxy clusters

Malak Olamaie1*, Farhan Feroz1, Keith J. B. Grainge2, Michael P. Hobson1, Jeremy S. Sanders3 and Richard D. E. Saunders1,4
1Astrophysics Group, Cavendish Laboratory, 19 J. J. Thomson Avenue, Cambridge, CB3 0HE
2Jodrell Bank Centre for Astrophysics, School of Physics and Astronomy, The University of Manchester, M13 9PL
3Max-Planck-Institut für extraterrestrische Physik, Giessenbachstrasse, 85748 Garching, Germany
4Kavli Institute for Cosmology Cambridge, Madingley Road, Cambridge, CB3 0HA

Accepted — ; Received — ; in original form 7 October 2013

ABSTRACT

We present the first public release of our Bayesian inference tool for the analysis of X-ray observations of galaxy clusters, called Bayes-X. We illustrate the approach of Bayes-X by using it to analyse a set of four simulated clusters at $z = 0.2 - 0.9$ in the X-ray band as they would be observed by a Chandra-like X-ray observatory. In both the simulations and the analysis pipeline we assume that the dark matter density follows a spherically-symmetric Navarro, Frenk and White (NFW) profile and that the gas pressure is described by a generalised NFW (GNFW) profile. We then perform four sets of analyses. These include priory-only analyses and analyses in which we adopt wide uniform prior probability distributions on $f_z(\beta)$ and on the model parameters describing the shape and slopes of the GNFW pressure profile, namely $(c_{500}, a, b, c)$. By numerically exploring the joint probability distribution of the cluster parameters given simulated Chandra-like data, we show that the model and analysis technique can robustly return the simulated cluster input quantities, constrain the cluster physical parameters and reveal the degeneracies among the model parameters and cluster physical parameters. The Bayes-X software which is implemented in Fortran 90 is available at http://www.mrao.cam.ac.uk/facilities/software/bayesx/

Key words: galaxies: clusters– cosmology: observations – methods: data analysis

1 INTRODUCTION

Clusters of galaxies, as the most massive gravitationally bound material structures, are of basic importance in the study of both baryonic and dark-matter density distributions in the Universe. In practice, measurement of line-of-sight velocity dispersions of the galaxies in a cluster (see e.g. Rines, Geller & Diaferio 2010 and Sifón et al. 2013), observation of X-ray emission from the hot gas in a cluster's gravitational potential well (see e.g. Vikhlinin et al. 2005, 2006; Allen, Evrard & Mantz 2011; Russell et al. 2012 and Sanders & Fabian 2013), observation at microwave frequencies of $2005, 2006; Allen, Evrard & Mantz 2011; Russell et al. 2012$ and in a cluster's gravitational potential well (see e.g. Vikhlinin et al. 2005, 2006; LaRoque et al. 2006; AMI Consortium; Shimwell et al. 2012, 2013) to parameterise the radial X-ray surface brightness profile and explore the constraints on both model parameters and cluster physical parameters. The Bayes-X software which is implemented in Fortran 90 is available at http://www.mrao.cam.ac.uk/facilities/software/bayesx/.

The great majority of X-ray (and indeed of SZ) measurements of cluster masses in the literature assume parameterized functional forms for the radial distribution of two independent cluster thermodynamic properties, such as electron density and temperature, to model the X-ray surface brightness (see e.g. Sarazin 1988; Vikhlinin et al. 2005, 2006; LaRoque et al. 2006; AMI Consortium; Olamaie et al. 2012 and Planck Collaboration et al. 2013). These radial profiles (e.g. $\beta$-model) have an amplitude normalisation parameter and two or more shape parameters.

In Bayes-X we use our recently developed cluster model (Olamaie et al. 2012, 2013) to parameterise the radial X-ray surface brightness profile and explore the constraints on both model param-

* email: mo323@mrao.cam.ac.uk
eters and physical parameters of simulated Chandra-like observations of four clusters. The model assumes that the dark matter density follows a Navarro, Frenk and White (NFW) profile (Navarro, Frenk & White 1996, 1997) and the ICM plasma pressure is described by the generalised NFW (GNFW) profile (Nagai, Kravtsov & Vikhlinin 2007; Arnaud et al. 2010), both in accordance with numerical simulations that take into account radiative cooling, star formation and energy feedback from supernova explosions (see e.g. Borgani 2004; Nagai, Kravtsov & Vikhlinin 2007 and Piffaretti & Valdarnini 2008).

Computational advances allow us to compare the model with the data in a fully Bayesian fashion (see e.g. Jaynes 1986 and Sivia & Valdarnini 2008).

Throughout, we assume a ΛCDM cosmology with \( \Omega_m = 0.3, \Omega_{\Lambda} = 0.7, \sigma_8 = 0.8, h = 0.7, w_0 = -1, w_a = 0. \) \( M_f(r_s), \) and \( f_g(r_s) \) represent the value of the cluster total mass, gas mass and gas mass fraction \( \text{internal} \) to the overdensity radius of \( r_s \), respectively, whereas \( T_g(r_s) \) represents the gas temperature at radius \( r_s \). The inner and outer contours in 2D marginalised posterior probability distributions indicate the areas enclosing 68% and 95% of the probability distributions.

### 2 BAYESIAN INFERENCE

Bayesian inference allows the estimation of a set of parameters \( \Theta \) within a model (or hypothesis) \( H \) using the data \( D \). Bayes' theorem states that:

\[
\Pr(\Theta|D, H) = \frac{\Pr(D|\Theta, H) \Pr(\Theta|H)}{\Pr(D|H)},
\]

where \( \Pr(D|\Theta, H) \) is the posterior probability distribution of the parameters \( \Theta \), given \( H \) and \( D \). \( \Pr(D|\Theta, H) \equiv L(\Theta) \) is the likelihood, \( \Pr(\Theta|H) \equiv \pi(\Theta) \) is a prior "prior" probability distribution of \( \Theta \) given \( H \), and \( \Pr(D|H) \equiv Z \) is the Bayesian evidence.

Bayesian inference in practice often divides into two parts: parameter estimation and model selection. In parameter estimation, the normalising evidence factor is usually ignored, since it is independent of the parameters \( \Theta \), and inferences are obtained by searching the unnormalised posterior distributions using sampling techniques. The posterior distribution can be subsequently marginalised over each parameter to give individual parameter constraints.

In contrast to parameter estimation, in model selection the evidence takes the central role and is simply the factor required to normalise the posterior over \( \Theta \):

\[
Z = \int L(\Theta) \pi(\Theta) d^D\Theta,
\]

where \( D \) is the dimensionality of the parameter space. According to Occam’s razor (see e.g. Jaynes 1986 and Sivia 2005), a simple theory with compact parameter space will have a larger evidence than a more complicated one, unless the latter is significantly better at explaining the data.

The question of the selection between two models \( H_0 \) and \( H_1 \) is then decided by comparing their respective posterior probabilities, given the observed data set \( D \), via the model selection ratio

\[
R = \frac{\Pr(H_1|D)}{\Pr(H_0|D)} = \frac{\Pr(D|H_1) \Pr(H_1)}{\Pr(D|H_0) \Pr(H_0)} = \frac{\int \Pr(D|H_1) \pi(H_1) d^D\Theta}{\int \Pr(D|H_0) \pi(H_0) d^D\Theta},
\]

where \( \Pr(H_1)/\Pr(H_0) \) is the prior probability ratio for the two models. The evaluation of the multidimensional integral for the Bayesian evidence is a challenging numerical task which can be tackled by using multiNest (Feroz & Hobson 2008; Feroz, Hobson & Bridges 2009). This Monte-Carlo method is targeted at the efficient calculation of the evidence, but also produces posterior inferences as a by-product. This method is also very efficient in sampling from posteriors that contain multiple modes or large (curving) degeneracies as is indeed the case in estimating density distributions of cluster gas from X-ray observations.

### 3 THE X-RAY OBSERVABLES

The fundamental sources of X-ray emission in clusters of galaxies include both continuum and line emission processes. In a hot diffuse plasma, the X-ray continuum emission is due to three processes: free–free (\( ff \)) emission (Bremsstrahlung); free-bound (\( fb \)) emission (recombination); and two-photon (\( 2\gamma \)) emission. In addition to the continuum emission, line radiation from a diffuse plasma contributes significantly to the flux. Line radiation is in particular very important at low temperatures (< 3 keV) as it can make up most of the flux, integrated over a broad energy band. The X-ray line emission is due to collisional excitation of valence or inner shell electrons, radiative and dielectronic recombination, inner shell collisional ionisation and the subsequent emission process following any of these processes. The emissivities of these processes are proportional to the square of the electron density.

The total emissivity, \( \epsilon_C \), (the number of photons per unit volume per unit time and per unit energy interval) is the sum of contributions from both continuum and line emissions,

\[
\epsilon_C = \epsilon_C + \epsilon_L,
\]

where \( \epsilon_C \) is the total continuum emissivity and \( \epsilon_L \) is the emissivity due to the line emission.

The total continuum emissivity is described as

\[
\epsilon_C = n_e^2 \Lambda_C(E, Z, T),
\]

where \( n_e \) is the electron number density and \( \Lambda_C(E, Z, T) \) is the cooling function which is a function of photon energy, \( E \), plasma temperature, \( T \), and metallicity, \( Z \), and may be described as (Mewe 1972, 1975; Gronenschild & Mewe 1978; Mewe, Lemen & van den Oord 1986; Kastra & Verbunt 2010),

\[
\Lambda_C = 3.031 \times 10^{21} \frac{n_H}{n_e} E_{keV}^{-1} T_{keV}^{-1/2} G_C e^{-E/\eta_H^T}.
\]

in units of counts m\(^3\)s\(^{-1}\)keV\(^{-1}\). Here \( \eta_H \) is the hydrogen number density and \( G_C \) is the so-called averaged Gaunt factor which represents the contributions to the continuum emission from free–free, free-bound and two-photon processes (\( G_C = G_{ff} + G_{fb} + G_{2\gamma} \)).

The total line emissivity is proportional to the spontaneous transition probability: the probability per unit time that the ion in an excited state decays back to the ground state or any other lower energy level by emitting a photon. The line emissivity may be described as

\[
\epsilon_L = n_e^2 \sum_{Z,j} n_{Z,j} n_{Z,j} P(E, Z', T),
\]

where \( n_{Z,j} \) is the total number density of element \( Z \), \( n_{Z,j}/n_{H} \) is the abundance of element \( Z \), \( n_{Z,j} \) is the number density of the ion \( Z^j \), \( n_{Z,j}/n_{e} \) is the ionisation fraction and \( P(E, Z', T) \) is the emission rate per ion at unit electron density (see e.g. Sarazin 1988 and Kastra & Verbunt 2010). Finally the observed surface brightness, \( S_X \), in a
given direction towards a cluster of redshift $z$ is proportional to the line integral of the total emissivity through the cluster

$$ S_X = \frac{1}{4\pi(1+z)} \int_{\infty}^{\infty} e_x \, dl. \quad (8) $$

For survey observations, the primary X-ray observables are the surface brightness spectrum, luminosity and spatial extent. In general, the photon flux density incident at the telescope is related to the observed photon count rate through an integral equation involving the instrumental response $^1$

$$ C(i) = \int_0^\infty R(E, i) S_X(E) \, dE, \quad (9) $$

where $C(i)$ is the photon count within instrumental channel $i$, and $R(E, i)$ is the instrumental response which is proportional to the probability that an incoming photon of energy $E$ will be detected in channel $i$. $R(E, i)$ is a continuous function of energy $E$ and a discrete function of channel number $i$; however, since the response is never known exactly and it is not practical to perform large numbers of integrals, the energy $E$ is binned into discrete ranges, $E(j)$ to $E(j + 1)$. Hence $S_X$ is converted to $S_X(j)$ and $R(E, i)$ to $R(j, i)$ in the same energy range. The number of energy bins depends on the energy resolution of the detector, the quality of the data, and the extent that the detector response is actually known.

The values of $R(j, i)$ are elements of a 2-dimensional matrix which is calculated by Hadamard multiplication of two matrices, the Redistribution matrix (RMF) and the Ancillary Response Array (ARF):

$$ R(j, i) = RMF(j, i) \circ ARF(j), \quad (10) $$

where the $RMF$ maps photon energy $E_{j-1} < E < E_j$ to output instrument channels and in the ideal case is almost diagonal. $ARF$ accounts for the effective area of the telescope and is stored in a single one-dimensional array and has the dimension of area. We note that to perform the multiplication in equation (10) we need to expand ARF matrix to have the same dimension as RMF.

In order to determine the photon flux density $S_X(j)$ of the source, we assume a model $S_X$ that may be described in terms of a few parameters (i.e. $S_X(E, \theta_1, \theta_2, \ldots)$) and fit it to the data. For each $S_X(j)$, a predicted count spectrum $C_{pred}(j)$ is calculated as a Hadamard multiplication of two matrices and compared to the observed data:

$$ C_{pred}(j) = \sum_i R(j, i) \circ S_X(j) \Delta E_j, \quad (11) $$

where $\Delta E_j$ is the width of energy bin.

The X-ray counts follow Poisson statistics so that the X-ray likelihood function, $L_X$, is given by

$$ \ln(L_X) = \sum_i \left[ C_{obs}(j) \ln(C_{pred}(j)) - C_{pred}(j) - \ln(C_{obs}(j))! \right], \quad (12) $$

where $i$ runs over all the energy channels at each pixel, $C_{pred}(j)$ is the model prediction (including cluster and background components), and $C_{obs}(j)$ is the number of counts detected in energy channel $i$.

### 4 MODELLING THE X-RAY SURFACE BRIGHTNESS

Modelling $S_X$ in equation (8), requires: a model to calculate the emissivity of the hot plasma, $e_X$; a model to describe the radial dependencies of the electron number density, $n_e$, and temperature, $T_e$; and a model to take into account X-ray absorption by the interstellar medium.

We calculate the emissivity using the MEKAL model (after Mewe, Kaastra & Liedahl; see Mewe et al. 1995). The model is one of the most widely used in X-ray spectral fitting analyses from hot, optically-thin plasmas and is also incorporated in XSPEC$^2$. In both ionisation-balance and in the spectral calculations, it models the effects of all ions of the 15 most important elements: H, He, C, N, O, Ne, Na, Mg, Al, Si, S, Ar, Ca, Fe, and Ni; it also adopts the “standard” abundances given in Anders & Grevesse (1989). Given the ion concentrations, the code calculates the X-ray spectrum, consisting of continuum and line emission. The continuum emission is described by Gronenschild & Mewe (1978) and Mewe, Lemen & van den Oord (1986), and consists of free–free emission, free-bound emission and two-photonic emission.

The required inputs are the plasma temperature, the hydrogen density, the abundances, the energy range of interest and the required spectral resolution. The output is the emissivity and the electron density relative to that of hydrogen, $n_e/n_H$, describing the overall ionisation state of the plasma.

We also consider the photoelectric absorption of X-rays en-route from the source to us. The effect of this absorption can be written as

$$ F = F_0 \exp(-\sigma_{eff}(E) \cdot n_H), \quad (13) $$

where $F_0$ and $F$ are the pre- and post-absorption flux densities, $\sigma_{eff}(E) = \sum_{n} \sigma_{eff}^{\text{pre}}(nZ/m)$ is the effective cross-section, weighted over the abundance of elements ($n_{Z/n}$), $\sigma_{eff}(E)$ is the photoelectric absorption cross-section of element $Z$ at energy $E$ and $N_H = \int n_H \, dl$ is the hydrogen column density. The sum includes all elements in the line of sight. The dimensionless quantity $\tau = \sigma_{eff}(E)N_H$ is known as the optical depth and is typically between 0.001 and 0.01 through the centre of a rich cluster.

The absorption cross-sections are calculated using polynomial fit coefficients obtained by Balucinska-Church & McCammon (1992) for 17 elements: the 15 elements listed above plus Cl and Cr. These cross-sections are intended to be for the hydrogen-like atomic form of the elements and do not take into account the possibility of ionisation or the inclusion of material into molecules. None of these, however, has a very large effect on the total absorption (Krolik & Kallman 1984; Balucinska-Church & McCammon 1992).

To calculate the line-of-sight integral of emissivity given in equation (8) and determine a map of $S_X$, we need to take into account the radial dependencies of the electron number density, $n_e$, and temperature, $T_e$.

We use the model described in Olamaie et al. (2012), with its corresponding assumptions on the dynamical state of the cluster. As shown in Olamaie et al. (2012), the model leads to radial profiles for clusters physical properties that are consistent both with numerical simulations and multi wavelength observations of clusters (see e.g. Navarro, Frenk & White 1997; Carlberg et al. 1997; Borgani et al. 2004; Pointecouteau, Arnaud & Pratt 2005; Vikhlinin et al. 2005, 2006; Holder, McCarthy & Babul 2007; Nagai, Kravtsov & Vikhlinin 2007; McCarthy et al. 2008; Mroczkowski et al. 2009; Arnaud et al. 2010 and Plagge et al. 2010).

The model assumes that the dark matter density follows a  

---

1 see http://heasarc.gsfc.nasa.gov/xanadu/xspec/manual/manual.html

2 see http://heasarc.gsfc.nasa.gov/docs/xanadu/xspec
Table 1. The input parameters of simulated X-ray clusters describing the telescope properties, X-ray background and hydrogen column density. These parameters are the same for each simulated cluster.

| Parameter               | Value            |
|-------------------------|------------------|
| \(N_{\text{H}}\)       | \(2.2 \times 10^{24}\) m\(^{-2}\) |
| exposure time           | \(3 \times 10^5\) s       |
| energy range            | 0.7–7 keV         |
| energy bin size, \(\Delta E\) | 0.1 keV          |
| pixel solid angle, \(\Delta \Omega\) | (0.492\(^{-2}\)) \(^2\) |
| X-ray background level  | \(8.6 \times 10^{-2}\) counts m\(^{-2}\) arcmin\(^{-2}\) s\(^{-1}\) |
| \(\Lambda_{\text{eff}}\) | 2.50 \times 10^{-2}\) m\(^2\) |

NFW profile (Navarro, Frenk & White 1996, 1997) and the plasma pressure is described by the GNFW profile (Nagai, Kravtsov & Vikhlinin 2007),

\[
\rho_{\text{G}}(r) = \frac{\rho_s}{(1 + \frac{r}{R_s})^2},
\]

(14)

\[
P_s(r) = \frac{P_a}{(1 + \frac{r}{R_s})^{3\alpha}},
\]

(15)

where \(\rho_s\) is an overall normalisation coefficient, \(R_s\) is the scale radius at which the logarithmic slope of the profile \(d \ln \rho(r)/d \ln r = -2\), \(P_a\) is an overall normalisation coefficient of the pressure profile, \(r_p\) is the scale radius defined through the gas concentration parameter, \(c_{500} = r_{500}/r_p\) and the parameters \((a, b, c)\) describe the slopes of the pressure profile at \(r \approx r_p, r > r_p\) and \(r \ll r_p\) respectively. It also is a common practice to define the halo concentration parameter, \(c_{200} = r_{200}/R_s\). To calculate this we use the relation derived by Neto et al. (2007) from N-body simulations, namely

\[
c_{200} = 5.26 \left(\frac{M_{\text{halo}}(r_{200})}{10^{14} M_\odot}\right)^{-0.1}.
\]

(16)

The cluster model parameters \(\rho_s, R_s\) and \(P_a\) and hence \(\rho_{\text{G}}(r)\) and \(P_s(r)\) distributions are derived under the following assumptions: spherical symmetry; hydrostatic equilibrium; and that the local gas fraction is much less than unity (equations (3) to (11) in Olamaie et al. 2012). Thus the relevant equations are:

\[
\rho_{\text{G}}(r) = \left(\frac{\mu_e}{\mu}\right) \left(\frac{1}{4\pi G}\right) \left(\frac{P_a}{P_s}\right) \left(\frac{1}{R_p}\right) \times \frac{r}{\ln \left(1 + \frac{r}{R_p}\right) - \left(1 + \frac{r}{R_p}\right)^{-1}} \times \left[\frac{r}{r_p}\right]^{-\alpha} \left[1 + \left(\frac{r}{r_p}\right)^{-\alpha}\right] \left[\frac{b}{\left(\frac{r}{r_p}\right) + c}\right],
\]

(17)

\[
k_B T_s(r) = (4\pi G \rho_{\text{G}})(R_p^4) \times \left[\frac{\ln \left(1 + \frac{r}{R_p}\right) - \left(1 + \frac{r}{R_p}\right)^{-1}}{r}\right] \times \left[1 + \left(\frac{r}{r_p}\right)^{-\alpha}\right] \left[b\left(\frac{r}{r_p} + c\right)^{-\alpha}\right],
\]

(18)

where \(\mu_e = 1.14 m_p\) is the mean gas mass per electron, \(\mu = 0.6 m_p\) is the mean mass per gas particle and \(m_p\) is the proton mass.

Using the abundances given in Anders & Grevesse (1989), the abundances and hydrogen number density \((n_0(r))\) are determined as,

\[
n_0(r) = \frac{\rho_0(r)}{m_p \sum_i A(i) \frac{n_{iH}}{n_{H}}},
\]

(19)

where \(A(i)\) is the nucleon number, and \(n_{iH}/n_{H}\) is the ion abundance. Given the energy range, abundances, temperature and hydrogen number density distributions, the MEKAL model is used to calculate the profiles \(n_{iH}/n_{H}\) and the emissivity. The electron number density \(n_0(r)\) is estimated using this ratio and \(n_{iH}(r)\).

Combining equations (4), (8), and (13),

\[
S_X(r, E) = \left(\frac{1}{4\pi (1 + z)^3}\right) \left(\frac{\pi^2}{60^2 \times 180^2}\right) \times \int_{-\infty}^{\infty} e_X(E, Z, T(R)) \exp(-\sigma_{\text{eff}}(E) \cdot n_{\text{H}}) \, dl
\]

(20)

\[
(s) (\text{counts}) (m)^{-2} (s)^{-1} (\text{keV})^{-1} (\text{arcmin})^{-2}.
\]

Setting \(r^2 = s^2 + \ell^2\) where \(s\) is the projected distance from the centre of the cluster on the sky and \(l\) is the distance along the line of sight, we can solve the integral in equation (20) numerically. At each pixel on the sky map we calculate photon flux density in each energy bin

\[
S_X(j) \equiv S_X(l, m, j),
\]

(21)

where \(l\) and \(m\) count the spatial pixels on the sky image, \((i.e. l = 1\) to \(n_l, m = 1\) to \(n_m\)) where \(n_l\) and \(n_m\) are the number of pixels in the image, while \(j\) counts the energy bins at each \((l, m)\) spatial pixel, \((i.e. j = 1\) to \(n_{\text{bin}}\)) where \(n_{\text{bin}}\) is the number of energy bins.

Hence for a particular pixel \((l, m)\) on the sky image, the predicted count in a detector output energy channel, \([C_{\text{in}}]_{\text{pixel}}\), is calculated by Hadamard multiplication of \(S_X(l, m, j),\) with \(R(j, i),\) where \(i\) is the detector energy channel number, \(i.e. i = 1\) to \(n_{\text{bin}}\),

\[
[C_{\text{in}}]_{\text{pixel}} = \sum_j [R[j]] \circ [S_{\text{in}}]_{\text{model}} \Delta E_j.
\]

(22)

To express the predicted counts in terms of \((\text{counts}) (m)^{-2} (s)^{-1}\), we multiply by each pixel’s solid angle in arcmin\(^2\).

5 SIMULATED X-RAY DATA

For our simulations, the output of a Chandra observation consists of four files, a data (“event”) file, telescope response files including the redistribution matrix (RMF) and the ancillary Response Array (ARF) files and a file containing the X-ray background emission.

We generate simulated Chandra ACIS images of four clusters with given \(z, M_T(r_{200}), f_{\text{bg}}(r_{200}), c_{500}, a, b\) and \(c\). We assume a typical Galactic neutral hydrogen column density of \(N_\text{H} = 2.2 \times 10^{24}\) m\(^{-2}\) and an exposure time equal to \(3 \times 10^5\) s to give the highest

Table 2. The physical parameters of simulated X-ray clusters. To generate the simulated clusters we assume the values of the gas concentration parameter and the slopes to be \((c_{500}, a, b, c) = (1.156, 1.0620, 5.4807, 0.3292)\) (appendix B in Arnaud et al. 2010). The value of the gas mass fraction within the overdensity radius of \(r_{200}\) was also fixed to \(f_{\text{bg}}(r_{200}) = 0.13\) in generating simulated clusters.

| Cluster         | \(z\) | \(M_T(r_{200})/(10^{14} M_\odot)\) |
|-----------------|------|-----------------------------------|
| X-ray cluster 1 | 0.2  | 6.16                              |
| X-ray cluster 2 | 0.3  | 5.80                              |
| X-ray cluster 3 | 0.5  | 5.20                              |
| X-ray cluster 4 | 0.9  | 4.10                              |
practicable signal-noise ratio. We choose parameters corresponding to Chandra ACIS detector and assume 100% optics and CCD quantum efficiencies. Our approach, however, can be applied to any X-ray telescope with its corresponding properties.

The ACIS detector provides spatially resolved X-ray spectroscopy and imaging with an angular resolution of $0.492^\prime$ and an energy resolution of $\approx 100–200$ eV. The background consists of both detector and astronomical components. At low energies there is a variable background from charge exchange (see e.g. Tawa et al. 2008, Bautz et al. 2009, Koutroumpa et al. 2009, Snowden 2009). There is also an OVIII emission line at 0.65 keV (see e.g. Koutroumpa et al. 2007) and possible contamination on the ACIS detector leading to degradation at low energies and uncertainties in calibration (see e.g. Allen et al. 2008). However, the strongest background component in ACIS on Chandra is flaring...
limit our analysis to the range. At the highest energies, the cluster emission decreases and the signal becomes background-dominated. For these reasons, we have a bigger effective area because of the spectral shape of cluster emission. When the input photon flux density is multiplied by the effective area, the result is the distribution of counts that would be seen by an ideal detector. This is the Poisson distribution with mean \(\mu\) and standard deviation \(\sigma\). Drawing from the Poisson distribution with expectation count for each energy channel of each pixel.

3 see http://cxc.harvard.edu/contrib/maxim/bgg/index.html
4 see http://cxc.harvard.edu/toolkit/pimms.jsp
5 see http://cxc.harvard.edu/caldb/prop_plan/pimms/index.html

6 Olamaie et al.

Table 3. Summary of the priors on the sampling parameters in the analyses. \(x_0\) and \(y_0\) are in arcsec and \(M_\text{T}(r_{200})\) is in M\(\odot\). Note that \(N(\mu, \sigma)\) represents a Gaussian probability distribution with mean \(\mu\) and standard deviation of \(\sigma\) and \(U(l_1, l_2)\) represents a uniform distribution between \(l_1\) and \(l_2\).

| Sampling Parameters | I          | II         | III        | IV         |
|---------------------|------------|------------|------------|------------|
| \(x_0\)             | N(0, 4)    | N(0, 4)    | N(0, 4)    | N(0, 4)    |
| \(y_0\)             | N(0, 4)    | N(0, 4)    | N(0, 4)    | N(0, 4)    |
| \(\log M_\text{T}(r_{200})\) | U(14, 15.8) | U(14, 15.8) | U(14, 15.8) | U(14, 15.8) |
| \(f_g(r_{200})\)    | N(0.13, 0.02) | U(0.01, 1.0) | N(0.13, 0.02) | N(0.13, 0.02) |
| \(a\)               | 1.0620     | 1.0620     | N(1.0620, 0.06) | U(0.3, 10) |
| \(b\)               | 5.4807     | 5.4807     | N(5.4807, 1)  | U(2.0, 15) |
| \(c\)               | 0.3292     | 0.3292     | N(0.3292, 0.02) | U(0, 1)    |
| \(c_{500}\)         | 1.156      | 1.156      | N(1.156, 0.02) | U(0.01, 6) |

\(\delta\)-function prior on redshift \(z\) at the true value for each cluster. The prior on \(M_\text{T}(r_{200})\) is taken to be uniform in \(\log M\) in the range \(M_{\text{min}} = 10^{14} M_\odot\) to \(M_{\text{max}} = 6 \times 10^{15} M_\odot\).

6 BAYESIAN ANALYSIS OF SIMULATED X-RAY CLUSTERS USING BAYES-X

We adopt the model described in Olamaie et al. (2012, 2013) and section 4 to study the sample of four simulated X-ray clusters with the input parameters as given in Tables 1 and 2.

Using BAYES-X, we perform four sets of analyses (I, II, III, IV) and also analyses with no-data in order to investigate the capability of the data, our model and the analysis to return the simulated cluster quantities and clearly reveal structure of degeneracies in the cluster parameter space.

In a general case the sampling parameter space comprises of \(\Theta_s \equiv (x_0, y_0, M_\text{T}(r_{200}), f_g(r_{200}), z, c_{500}, a, b, c)\), where \(x_0\) and \(y_0\) are cluster projected position on the sky. We further assume that the priors on sampling parameters are separable (Feroz et al. 2009b) such that

\[
\pi(\Theta_s) = \pi(x_0) \pi(y_0) \pi(M_\text{T}(r_{200})) \pi(f_g(r_{200})) \pi(z) \pi(c_{500}) \pi(a) \pi(b) \pi(c).
\]

A summary of the priors on the sampling parameters in the analyses I–IV is presented in Table 3. We use Gaussian priors on cluster position parameters, centred on the pointing centre and with standard-deviation of 4\(\arcsec\). We adopt a \(\delta\)-function prior on redshift \(z\) at the true value for each cluster. The prior on \(M_\text{T}(r_{200})\) is taken to be uniform in \(\log M\) in the range \(M_{\text{min}} = 10^{14} M_\odot\) to \(M_{\text{max}} = 6 \times 10^{15} M_\odot\). The prior on \(f_g(r_{200})\) is set to be a Gaussian centred at \(f_g = 0.13\) with a width of 0.02 in the analyses I, III, IV and in the no-data analysis and uniform with a wide range between 0.01 and 1.0 in analysis II.
Gaussian priors were chosen to ensure no singularity in the GNFW pressure profile. As well as analysing the simulated data we also fitted each individual observed cluster pressure profile with the detected at high significance in the 14-month nominal survey using a.

We fix the values of \((c_{500}, a, b, c)\) in the analyses I and II to the input values of the simulated clusters given in Table 2 and appendix B in Arnaud et al. (2010).

In analysis III we use Gaussian priors on \((c_{500}, a, b, c)\) centred on the input values of the simulated clusters with standard deviations as given in the third column of Table 3. The widths on the Gaussian priors were chosen to ensure no singularity in the GNFW pressure profile. As well as analysing the simulated data we also perform a prior-only analysis assuming this set of prior probability distributions.

Finally, in analysis IV we use uniform priors on \((c_{500}, a, b, c)\) with the ranges as given in fourth column of Table 3. We adopt the range according to the studies carried out by Arnaud et al. (2010) and Planck Collaboration et al. (2013). Arnaud et al. (2010) studied 31 Representative XMM-Newton Cluster Structure Survey (REXCESS) cluster sample from XMM-Newton observations (Böhringer et al. 2007; Pratt et al. 2010; Arnaud et al. 2010) within \(r_{500}\). They fitted each individual observed cluster pressure profile with the GNFW model, fixing \(b = 5.4905\). The range of the estimated best fitting parameters for this cluster sample are: \(0.01 \leq c_{500} \leq 5.51, 0.33 \leq a \leq 2.54\) and \(0.0 \leq c \leq 1\). Planck Collaboration et al. 2013 also studied the pressure profiles of 62 nearby massive clusters detected at high significance in the 14-month nominal survey using a GNFW pressure profile. Their cluster sample is a sub-sample from the ESZ catalogue (The Early release SZ sample) which comprises 189 clusters detected in SZ (Planck Collaboration et al. 2011). They fixed the value of \(c = 0.31\) and derived the best fit values for the other three parameters for each individual cluster in their sample. Their parameter values lie in \(0.01 \leq c_{500} \leq 5.51, 0.36 \leq a \leq 10\) and \(2.23 \leq b \leq 15\). We therefore selected the range of the priors on \((c_{500}, a, b, c)\) according to the minimum and maximum values of the best fitting values in these two studies rounding the numbers to the nearest integers in case of \(c_{500}\) and \(b\). Similar to analysis III we also performed a prior-only analysis assuming this set of priors.

Having established our sampling parameters in each analysis, modelling the X-ray predicted counts, \([C_{\text{count}}]_{\text{pred}}\), is performed through the calculation of the X-ray surface brightness, \(S_{\text{X}}(l, m, f)\). This requires the knowledge of (a) parameters describing the 3D plasma density and its temperature, namely \(K, \rho_s, r_p\), and \(P_\alpha\) (b) X-ray emissivity and photoelectric absorption cross-sections, and (c) background and telescope response files.

By sampling from \(M_T(r_{200})\), \(z\) and \(c_{500}\) Bayes-X calculates \(R, \rho_s, r_p\) assuming spherical geometry and using equation (16). By sampling from \(M_T(r_{200})\) and \(f_\alpha(r_{200})\) it also calculates \(M_T(r_{200}) = f_\alpha(r_{200})M_T(r_{200})\). Using \(M_T(r_{200})\) it determines the model parameter \(P_\alpha\) by sampling from \(a, b, c\), and assumption of hydrostatic equilibrium (for detailed calculation see Olamaie et al. 2012, 2013).

Following the steps described in section 4 Bayes-X then calculates the X-ray emissivity and photoelectric absorption cross-sections to determine the model map of the X-ray surface brightness on a grid of 256 by 256 and at each energy channel. 3-D predicted counts and the X-ray likelihood function are estimated using equations (22) and (12) assuming the telescope files and X-ray background level as described in section 5.

7 RESULTS

Figs. 2–7 show 2-D and 1-D marginalised posterior distributions of both sampling and derived parameters of simulated Chandra clus-
Table 4. Mean and 68%-confidence uncertainties of sampling and derived parameters of simulated cluster 1.

| Parameters | Input Values | Analysis I | Analysis II | Analysis III | Analysis IV |
|------------|--------------|------------|-------------|--------------|-------------|
| $x_0(\text{arcsec})$ | 0 | -0.03$^{+0.06}_{-0.06}$ | -0.03$^{+0.06}_{-0.11}$ | -0.030$^{+0.004}_{-0.110}$ | -0.03$^{+0.06}_{-0.05}$ |
| $y_0(\text{arcsec})$ | 0 | 0.05$^{+0.06}_{-0.06}$ | 0.05$^{+0.06}_{-0.06}$ | 0.05$^{+0.06}_{-0.06}$ | 0.05$^{+0.06}_{-0.06}$ |
| $M_T(r_{200}) \times 10^{14} M_\odot$ | 6.16 | 6.1$^{+0.1}_{-0.1}$ | 6.1$^{+0.1}_{-0.1}$ | 6.2$^{+0.2}_{-0.2}$ | 6.2$^{+0.2}_{-0.2}$ |
| $f_g(r_{200})$ | 0.13 | 0.129$^{+0.001}_{-0.000}$ | 0.129$^{+0.001}_{-0.000}$ | 0.129$^{+0.001}_{-0.000}$ | 0.129$^{+0.001}_{-0.000}$ |
| $a$ | 1.0620 | 1.0620 | 1.0620 | 1.0620 | 1.0620 |
| $b$ | 5.4807 | 5.4807 | 5.4807 | 5.4807 | 5.4807 |
| $c$ | 0.3292 | 0.3292 | 0.3292 | 0.3292 | 0.3292 |
| $r_{500}(\text{Mpc})$ | 1.36 | 1.36 | 1.36 | 1.36 | 1.36 |
| $T_g(r_{500})(\text{keV})$ | 3.45 | 3.45 | 3.45 | 3.45 | 3.45 |
| $r_{200}(\text{Mpc})$ | 1.65 | 1.65 | 1.65 | 1.65 | 1.65 |
| $M_T(r_{200}) \times 10^{13} M_\odot$ | 8.0 | 7.9$^{+0.1}_{-0.1}$ | 7.9$^{+0.1}_{-0.1}$ | 7.9$^{+0.1}_{-0.1}$ | 7.9$^{+0.1}_{-0.1}$ |
| $T_g(r_{200})(\text{keV})$ | 2.8 | 2.77$^{+0.04}_{-0.04}$ | 2.77$^{+0.04}_{-0.04}$ | 2.77$^{+0.04}_{-0.04}$ | 2.77$^{+0.04}_{-0.04}$ |
| $c_{200}$ | 3.78 | 3.79$^{+0.01}_{-0.01}$ | 3.79$^{+0.01}_{-0.01}$ | 3.79$^{+0.01}_{-0.01}$ | 3.79$^{+0.01}_{-0.01}$ |

$z = 0.5$

Figure 3. Analysis II: 2-D and 1-D marginalised posterior distributions of sampling parameters (left) and derived parameters (right) of the X-ray simulated cluster at redshift $z = 0.5$. The priors are given in the second column of Table 3. In this analysis, the prior on $f_g(r_{200})$ is assumed to be uniform between 0.01 and 1. The vertical green solid lines on the 1-D posterior distributions of the parameters show the input values used to generate the simulated cluster and the dashed magenta lines are the mean values of the distributions. $x_0$ and $y_0$ are in arcsec. $M$ and $f_g$ stand for $M_T(r_{200})$ and $f_g(r_{200})$ respectively. $M$ is in $M_\odot$. $r_{200}$ is in Mpc.
Table 5. Mean and 68%-confidence uncertainties of sampling and derived parameters of simulated cluster 2.

| Parameters | Input values | Analysis |
|------------|--------------|----------|
| $x_0$ (arcsec) | 0 | -0.07$^{+0.06}_{-0.06}$ | -0.07$^{+0.06}_{-0.06}$ | -0.07$^{+0.06}_{-0.06}$ |
| $y_0$ (arcsec) | 0 | 0.08$^{+0.06}_{-0.06}$ | 0.08$^{+0.06}_{-0.06}$ | 0.07$^{+0.06}_{-0.06}$ |
| $M_f(r_{200}) \times 10^{14} M_\odot$ | 5.80 | 5.9$^{+0.2}_{-0.2}$ | 5.9$^{+0.2}_{-0.2}$ | 6.0$^{+0.2}_{-0.2}$ |
| $f_g(r_{200})$ | 0.13 | 0.12$^{+0.02}_{-0.02}$ | 0.12$^{+0.02}_{-0.02}$ | 0.12$^{+0.02}_{-0.02}$ |
| $a$ | 1.0620 | 1.0620 | 1.0620 | 1.14$^{+0.06}_{-0.06}$ |
| $b$ | 5.4807 | 5.4807 | 5.4807 | 5.4$^{+1.8}_{-1.6}$ |
| $c$ | 0.3292 | 0.3292 | 0.3292 | 0.34$^{+0.01}_{-0.01}$ |
| $c_{500}$ | 1.156 | 1.156 | 1.156 | 1.3$^{+0.5}_{-0.6}$ |
| $M_f(r_{500}) \times 10^{14} M_\odot$ | 4.3 | 4.4$^{+0.1}_{-0.1}$ | 4.4$^{+0.1}_{-0.1}$ | 4.5$^{+0.2}_{-0.2}$ |
| $f_g(r_{500})$ | 0.1208 | 0.117$^{+0.001}_{-0.001}$ | 0.117$^{+0.001}_{-0.001}$ | 0.115$^{+0.005}_{-0.005}$ |
| $r_{500}$ (Mpc) | 1.038 | 1.04$^{+0.01}_{-0.01}$ | 1.04$^{+0.01}_{-0.01}$ | 1.05$^{+0.01}_{-0.01}$ |
| $M_f(r_{500}) \times 10^{15} M_\odot$ | 5.19 | 5.1$^{+0.08}_{-0.08}$ | 5.1$^{+0.08}_{-0.08}$ | 5.1$^{+0.1}_{-0.1}$ |
| $T_{500}$ (keV) | 3.40 | 3.43$^{+0.06}_{-0.06}$ | 3.43$^{+0.06}_{-0.06}$ | 3.4$^{+0.1}_{-0.1}$ |
| $r_{200}$ (Mpc) | 1.557 | 1.56$^{+0.01}_{-0.01}$ | 1.56$^{+0.01}_{-0.01}$ | 1.57$^{+0.02}_{-0.02}$ |
| $M_f(r_{200}) \times 10^{15} M_\odot$ | 7.54 | 7.5$^{+0.1}_{-0.1}$ | 7.5$^{+0.1}_{-0.1}$ | 7.4$^{+0.3}_{-0.3}$ |
| $T_{200}$ (keV) | 2.79 | 2.81$^{+0.05}_{-0.05}$ | 2.81$^{+0.05}_{-0.05}$ | 2.8$^{+0.1}_{-0.1}$ |
| $c_{200}$ | 3.52 | 3.51$^{+0.01}_{-0.01}$ | 3.51$^{+0.01}_{-0.01}$ | 3.5$^{+0.01}_{-0.01}$ |

Figure 4. Analysis III: 2-D and 1-D marginalised posterior distributions of sampling parameters (left) and derived parameters (right) of the X-ray simulated cluster at redshift $z = 0.5$. The priors used for the analysis are given in third column of Table 3. The vertical green solid lines on the 1-D posterior distributions of the parameters show the input values used to generate the simulated cluster and the dashed magenta lines are the mean values of the distributions. $x_0$ and $y_0$ are in arcsec. $M$ and $f_g$ stand for $M_f(r_{200})$ and $f_g(r_{200})$ respectively. $M$ is in $M_\odot$. 

$z = 0.5$
Figure 5. 1-D marginalised posterior distributions of sampling parameters (left) and derived parameters (right) of the no-data run (black solid lines) and X-ray simulated cluster 3 at redshift $z = 0.5$ (blue solid lines). The priors used for the analysis are given in the third column of Table 3.

$z = 0.5$

Figure 6. Analysis IV: 2-D and 1-D marginalised posterior distributions of sampling parameters (left) and derived parameters (right) of the X-ray simulated clusters at redshift $z = 0.5$. The priors used for the analysis are given in fourth column of Table 3. The vertical green solid lines on the 1-D posterior distributions of the parameters show the input values used to generate the simulated cluster and the dashed magenta lines are the mean values of the distributions. $x_0$ and $y_0$ are in arcsec. $M$ and $f_{g}$ stand for $M_T(r_{200})$ and $f_{g}(r_{200})$ respectively. $M$ is in $M_{\odot}$. 
Figure 7. 1-D marginalised posterior distributions of sampling parameters (left) and derived parameters (right) of the no-data run (black solid lines) and simulated Chandra cluster 3 at redshift $z = 0.5$ (blue solid lines). The priors used for the analysis are given in fourth column of Table 3.

Table 6. Mean and 68%-confidence uncertainties of sampling and derived parameters of simulated cluster 3.

| Cluster 3 Parameters | Input values | Analysis I | Analysis II | Analysis III | Analysis IV |
|----------------------|-------------|------------|-------------|--------------|-------------|
| $x_0$(arcsec)        | 0           | 0.04$^{+0.07}_{-0.07}$ | 0.04$^{+0.07}_{-0.07}$ | 0.04$^{+0.06}_{-0.06}$ | 0.04$^{+0.06}_{-0.06}$ |
| $y_0$(arcsec)        | 0           | 0.06$^{+0.07}_{-0.07}$ | 0.06$^{+0.07}_{-0.07}$ | 0.06$^{+0.07}_{-0.07}$ | 0.06$^{+0.07}_{-0.07}$ |
| $M_\gamma(r_{200})\times10^{14}M_\odot$ | 5.2         | 5.4$^{+0.2}_{-0.2}$ | 5.4$^{+0.2}_{-0.2}$ | 5.4$^{+0.3}_{-0.3}$ | 5.4$^{+0.3}_{-0.3}$ |
| $f_\gamma(r_{200})$ | 0.13        | 0.126$^{+0.002}_{-0.002}$ | 0.126$^{+0.002}_{-0.002}$ | 0.128$^{+0.007}_{-0.008}$ | 0.13$^{+0.01}_{-0.01}$ |
| $a$                  | 1.062       | 1.062      | 1.062      | 1.05$^{+0.03}_{-0.02}$ | 1.05$^{+0.08}_{-0.08}$ |
| $b$                  | 5.4807      | 5.4807     | 5.4807     | 5.4$^{+0.2}_{-0.2}$ | 6.2$^{+2.6}_{-2.3}$ |
| $c$                  | 0.3292      | 0.3292     | 0.3292     | 0.32$^{+0.01}_{-0.01}$ | 0.32$^{+0.01}_{-0.01}$ |
| $c_{500}$            | 1.156       | 1.156      | 1.156      | 1.15$^{+0.02}_{-0.02}$ | 1.2$^{+0.6}_{-0.69}$ |
| $M_\gamma(r_{500})\times10^{14}M_\odot$ | 3.85        | 4.0$^{+1.0}_{-1.0}$ | 4.0$^{+1.0}_{-1.0}$ | 3.9$^{+0.2}_{-0.2}$ | 3.9$^{+0.2}_{-0.2}$ |
| $f_\gamma(r_{500})$ | 0.122       | 0.119$^{+0.002}_{-0.002}$ | 0.118$^{+0.002}_{-0.002}$ | 0.120$^{+0.005}_{-0.005}$ | 0.120$^{+0.006}_{-0.006}$ |
| $r_{500}$(Mpc)       | 0.927       | 0.93$^{+0.01}_{-0.01}$ | 0.93$^{+0.01}_{-0.01}$ | 0.93$^{+0.01}_{-0.01}$ | 0.93$^{+0.01}_{-0.01}$ |
| $M_\gamma(r_{500})\times10^{13}M_\odot$ | 4.72        | 4.7$^{+0.1}_{-0.1}$ | 4.7$^{+0.1}_{-0.1}$ | 4.7$^{+0.1}_{-0.1}$ | 4.7$^{+0.1}_{-0.1}$ |
| $T_\gamma(r_{500})$(keV) | 3.35       | 3.42$^{+0.08}_{-0.08}$ | 3.42$^{+0.08}_{-0.08}$ | 3.4$^{+0.1}_{-0.1}$ | 3.4$^{+0.1}_{-0.1}$ |
| $r_{200}$(Mpc)       | 1.39        | 1.40$^{+0.02}_{-0.02}$ | 1.41$^{+0.02}_{-0.02}$ | 1.40$^{+0.02}_{-0.02}$ | 1.40$^{+0.02}_{-0.02}$ |
| $M_\gamma(r_{200})\times10^{13}M_\odot$ | 6.76        | 6.8$^{+0.1}_{-0.1}$ | 6.8$^{+0.1}_{-0.1}$ | 6.8$^{+0.2}_{-0.2}$ | 6.8$^{+0.4}_{-0.4}$ |
| $T_\gamma(r_{200})$(keV) | 2.8        | 2.86$^{+0.07}_{-0.07}$ | 2.86$^{+0.07}_{-0.07}$ | 2.8$^{+0.1}_{-0.1}$ | 2.8$^{+0.1}_{-0.1}$ |
| $c_{200}$            | 3.08        | 3.07$^{+0.01}_{-0.01}$ | 3.07$^{+0.01}_{-0.01}$ | 3.07$^{+0.01}_{-0.01}$ | 3.07$^{+0.02}_{-0.02}$ |
ter at $z = 0.5$. Fig. 5 shows the results when we adopted Gaussian priors on $(c_{500}, a, b, c)$ according to column three in Table 3. Fig. 7 shows the results when we adopted uniform priors on these parameters as given in column four of Table 3. In both analyses, we fixed the redshift to $z = 0.5$ corresponding to the redshift of cluster 3. We have also plotted the 1-D marginalised posterior distributions of both sampling and derived parameters (blue solid lines) of cluster 3 in both figures.

8 DISCUSSION

From the plots described above we note that the cluster position ($x_0$ and $y_0$) on the sky is firmly constrained in all cases and the true values all lie within 1σ of the means of the posterior probability distributions.

Throughout the analyses, the tight constraints on $M_\text{f}$, $T_\text{f}$ and $M_\text{T}$ and their insensitivity to the choice of priors are also clear, which shows the strong correlation between the X-ray luminosity and the cluster mass.

Similarly, from the 1-D marginalised posterior distributions of both $f_g(r_{200})$ and $f_g(r_{500})$, it is clear that we are able to constrain $f_g$ even in the analysis II where we assume a wide prior on $f_g(r_{200})$. The negative degeneracy between $f_g$ and $M_f$ is also apparent in the corresponding 2-D marginalised probability distributions in all the analyses, as one would expect ($f_g = M_g/M_f$).

In order to investigate the capability of our simulated X-ray data and analysis pipeline to constrain parameters describing the shape and slopes of the GNFW model, namely $(c_{500}, a, b, c)$, and to return the simulated cluster input values we let these parameters vary in the analyses III and IV. As was mentioned in section 6, we first assumed Gaussian prior probability distributions on these parameters centred on the input values used to generate the simulated clusters and with narrow widths (see third column in Table 3). Fig.4 shows the results of the analysis. Then in order to make sure that the results are not biased by the narrow Gaussian prior distributions, and to reveal the degeneracy between the parameters more clearly, we assume uniform priors on the pressure profile shape and slope parameters in analysis IV. We selected the range of priors based on the studies in Arnaud et al. (2010) and Planck Collaboration et al. (2013) (see fourth column in Table 3). Fig.6 shows the results of the analysis. From the plots we note that the simulated X-ray data can constrain the cluster model parameters and recover the input true values. This confirms that X-ray data can probe the cluster core and constrain the shape and slope parameters of the plasma pressure profile.

The 2-D marginalised posterior probability distributions of $(c_{500}, a, b, c)$ also show clear degeneracies among these parameters. This implies that obtaining an unbiased estimate of cluster parameters requires taking into account the degeneracies among these parameters in the analysis which can only be achieved by letting all four parameters vary.

There is also no correlation between the values of the physical parameter, $M_\text{T}$, and the shape parameters, $a$, $b$ and $c$. Further, the mean values of the cluster parameters do not change significantly upon changing the prior probability distributions in different analyses.

We have also investigated our methodology in the absence of data when we sample from the whole set of parameters, namely $\Theta = (x_0, y_0, M_\text{f}(r_{200}), f_g(r_{200}), z, c_{500}, a, b, c)$. This is carried out by setting the likelihood to a constant value and hence the algorithm explores the prior space. This analysis is crucial for understanding the underlying biases and constraints imposed by the priors and the model assumptions. The comparison of this analysis with the analysis using the simulated Chandra data reveals the constraints that measurements of the X-ray signal place on the cluster physical parameters and the robustness of the assumptions made.

Figs 5 and 7 represent the results of the prior-only analysis showing 1-D marginalised posterior distributions of both sampling and derived parameters (black solid lines). The results not only show that we are able to recover the assumed prior probability distributions of cluster parameters but also demonstrate the tight constraints on the cluster parameters arising from the simulated X-ray data. We note that the constraint on $c_{500}$ in Fig.5 from the simulated X-ray data overlaps the one from the no-data run but as Fig.7 shows this effect is a direct result of imposing a tight and very informative prior on $c_{500}$.

We also note the tight constraints that these simulated X-ray data sets can place upon the $c$ and $a$ parameters, which describe the slopes of the GNFW pressure profile at $r < r_p$ and $r \approx r_p$ and the fairly wide constraint on the $b$ parameter that describes the slope where $r > r_p$ (Fig.7).

Finally, we note that our analysis has not included effects such as finite energy resolution of the detector, PSF, varying background, point sources and substructure (e.g. related to AGN feedback, oblateness, mergers, etc). The first two effects are readily included in Bayes-X, with minimal modification. The possibility of a varying background can also, in principle, be included in our approach by making the appropriate modifications to the likelihoods. For example, one might model the background in each pixel as being drawn from a Poisson distribution with a mean that varies smoothly across the image. The final two issues of point sources and substructure may be addressed by extending our model for data to include them. Indeed our Bayesian approach is well-suited, through the evaluation of the evidence, to determining the number of additional model components required to accommodate point sources and substructures. We will address the complications in a future work.

9 CONCLUSIONS

By performing a Bayesian analysis of simulated Chandra data we have investigated the capability of our model (where we assume that the dark matter density follows a NFW-profile and that the gas pressure is described by a GNFW profile) and Bayes-X to return the cluster input quantities and constrain the cluster physical parameters.

We simulated Chandra-like observations of four clusters in a redshift range of 0.2–0.9 all with the same $f_g(r_{200}) = 0.13$. We have performed four sets of analyses including prior-only analysis and assuming different types of priors on $f_g(r_{200})$ and model parameters $(c_{500}, a, b, c)$.

We have demonstrated that Bayes-X faithfully recovers the input values of the model parameters used in the simulations and can constrain clusters positions on the sky and clusters physical parameters including $M_\text{f}$, $T_\text{f}$ and $M_\text{T}$.

We find that we can still constrain $f_g$ as well as other cluster parameters even when we assume a wide uniform prior on $f_g(r_{200})$.

By letting $(c_{200}, a, b, c)$ vary in the analysis we have shown that Bayes-X is able to reveal the degeneracy among these parameters which must be taken into account for an unbiased estimate of cluster parameters. We did this by assuming Gaussian and uniform prior probability distributions on $(c_{500}, a, b, c)$ respectively. The re-
Table 7. Mean and 68%-confidence uncertainties sampling and derived parameters of simulated cluster 4.

| Cluster 4 Parameters | Input values | Analysis |
|----------------------|--------------|----------|
| $x_0$ (arcsec)       | 0            | -0.02$^{+0.07}_{-0.07}$ | -0.02$^{+0.07}_{-0.07}$ | -0.02$^{+0.07}_{-0.07}$ | -0.02$^{+0.06}_{-0.06}$ |
| $y_0$ (arcsec)       | 0            | 0.05$^{+0.07}_{-0.07}$ | 0.05$^{+0.07}_{-0.07}$ | 0.04$^{+0.07}_{-0.07}$ | 0.04$^{+0.06}_{-0.06}$ |
| $M_{\gamma}(r_{200}) \times 10^{14} M_{\odot}$ | 4.10 | 4.1$^{+0.02}_{-0.02}$ | 4.1$^{+0.02}_{-0.02}$ | 4.3$^{+0.03}_{-0.03}$ | 4.3$^{+0.03}_{-0.03}$ |
| $f_{\gamma}(r_{200})$ | 0.13$^{+0.003}_{-0.003}$ | 0.13$^{+0.003}_{-0.003}$ | 0.125$^{+0.008}_{-0.008}$ | 0.125$^{+0.008}_{-0.008}$ |
| $a$                  | 1.0620       | 1.0620 | 1.0620 | 1.08$^{+0.03}_{-0.03}$ | 1.1$^{+0.1}_{-0.1}$ |
| $b$                  | 5.4807       | 5.4807 | 5.4807 | 5.6$^{+0.2}_{-0.2}$ | 6.2$^{+0.9}_{-0.9}$ |
| $c$                  | 0.329        | 0.3292 | 0.3292 | 0.34$^{+0.01}_{-0.01}$ | 0.35$^{+0.02}_{-0.02}$ |
| $r_{500}$ (Mpc)      | 1.156        | 1.156  | 1.156  | 1.15$^{+0.02}_{-0.02}$ | 1.3$^{+0.7}_{-0.7}$ |
| $M_{\gamma}(r_{500}) \times 10^{14} M_{\odot}$ | 3.04 | 3.0$^{+0.1}_{-0.1}$ | 3.0$^{+0.1}_{-0.1}$ | 3.2$^{+0.2}_{-0.2}$ | 3.2$^{+0.2}_{-0.2}$ |
| $f_{\gamma}(r_{500})$ | 0.1254         | 0.125$^{+0.003}_{-0.003}$ | 0.125$^{+0.003}_{-0.003}$ | 0.119$^{+0.006}_{-0.006}$ | 0.117$^{+0.006}_{-0.006}$ |
| $r_{200}$ (Mpc)      | 0.73         | 0.73$^{+0.01}_{-0.01}$ | 0.73$^{+0.01}_{-0.01}$ | 0.74$^{+0.02}_{-0.02}$ | 0.74$^{+0.02}_{-0.02}$ |
| $M_{\gamma}(r_{200}) \times 10^{14} M_{\odot}$ | 3.8 | 3.7$^{+0.1}_{-0.1}$ | 3.7$^{+0.1}_{-0.1}$ | 3.7$^{+0.1}_{-0.1}$ | 3.7$^{+0.1}_{-0.1}$ |
| $T_{\gamma}(r_{200})$ (keV) | 3.25 | 3.2$^{+0.1}_{-0.1}$ | 3.2$^{+0.1}_{-0.1}$ | 3.2$^{+0.1}_{-0.1}$ | 3.2$^{+0.1}_{-0.1}$ |
| $c_{500}$            | 2.80         | 2.76$^{+0.08}_{-0.08}$ | 2.76$^{+0.09}_{-0.09}$ | 2.7$^{+0.1}_{-0.1}$ | 2.7$^{+0.5}_{-0.5}$ |

The results also show no correlation between $M_{\gamma}$ and $a$, $b$, or $c$ as one would expect.

The results of prior-only analyses show that we recover the assumed prior probability distributions for cluster positions, model parameters and physical parameters.

We find that the results of the analyses do not depend on the choice of prior probability distributions on the sampling parameters for these high signal-noise simulations and in all cases we were able to recover the input values of the simulated clusters and expected degeneracies among the cluster parameters. We therefore conclude that BAYES-X is robust and can be used to analyse X-ray data and in future multi-wavelength analysis of clusters of galaxies.

Acknowledgments

The analysis work was conducted on the Darwin Supercomputer of the University of Cambridge High Performance Computing Service supported by HEFCE and COSMOS UK National Supercomputer at DAMTP, University of Cambridge. The authors thank Stuart Rankin and Andrey Kalizian for their computing assistance and Dave Green for his invaluable help with EfHx MO acknowledges a Research Fellowship from Sidney Sussex College, Cambridge.

References

Allen S. W., Rapetti D. A., Schmidt R. W., Ebeling H., Morris R. G., Fabian A. C., 2008, MNRAS, 383, 879
Allen S. W., Evrard A. E., Mantz A. B., 2011, ARA&A, 49, 409
AMI Consortium: Hurley-Walker et al., 2012, MNRAS, 419, 2921
AMI Consortium: Olamaie et al., 2012, MNRAS, 421, 1136
Anders E., Grevesse N., 1989, GeCoA, 53, 197
Arnaud M., Pratt G. W., Piffaretti R., Böhringer H., Croston J. H., Pointecouteau E., 2010, A&A, 517, A92
Balucinska-Church M., McCammon D., 1992, ApJ, 400, 699
Bautz M. W., et al., 2009, PASJ, 61, 1117
Böhringer H. et al., 2007, A&A, 469, 363
Borgani S., et al., 2004, MNRAS, 348, 1078
Borgani S., 2004, Ap&SS, 294, 51
Carlberg R. G., et al., 1997, ApJ, 485, L13
Corless V. L., King L. J., Clowe D., 2009, MNRAS, 393, 1235
Feroz F., Hobson M. P., 2008, MNRAS, 384, 449
Feroz F., Hobson M. P., Bridges M., 2009, MNRAS, 398, 1601
Feroz F., Hobson M. P., Zwart J. T. L., Saunders R. D. E., Grainge K. J. B., 2009, MNRAS, 398, 2049
Green D. A., 2011, BASI, 39, 289
Gronenschild E. H. B. M., Mewe R., 1978, A&AS, 32, 283
Holder G. P., McCarthy I. G., Babul A., 2007, MNRAS, 382, 1697
Jaynes E. T., 1986, Bayesian methods: an introductory tutorial, Cambridge University Press
Kaastra J. S., Verbunt F., 2010, High Energy Astrophysics.
Koutrompata D., Acero F., Lallement R., Ballet J., Kharchenko V., 2007, A&A, 475, 901
Koutrompata D., Lallement R., Kharchenko V., Dalgarao A., 2009, SSRv, 143, 217
Krolik J. H., Kallman T. R., 1984, ApJ, 286, 366
LaRoque S. J., Bonamente M., Carlstrom J. E., Joy M. K., Nagai D., Reese E. D., Dawson K. S., 2006, Ap, 652, 917
McCarthy I. G., Babul A., Bower R. G., Balogh M. L., 2008, MNRAS, 386, 1309
Mewe R., 1972, SoPh, 22, 459
Mewe R., 1975, SoPh, 44, 383
Mewe R., Lemen J. R., van den Oord G. H. J., 1986, A&AS, 65, 511
Mewe R., Kaastra J.S., Liedahl D.A., 1995, Legacy 6, 16
Mroczkowski T., et al., 2009, ApJ, 694, 1034
Nagai D., Kravtsov A. V., Vikhlinin A., 2007, ApJ, 668, 1
Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Neto A. F., et al., 2007, MNRAS, 381, 1450
Olamaie M., Hobson M. P., Grainge K. J. B., 2012, MNRAS, 423, 1534
Olamaie M., Hobson M. P., Grainge K. J. B., 2013, MNRAS, 430, 1344
Piffaretti R., Valdarnini R., 2008, A&A, 491, 71
Plagge T., et al., 2010, ApJ, 716, 1118
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2011, A&A, 536, A8
Planck Collaboration, et al., 2013, A&A, 550, A128
Planck Collaboration, et al., 2013, A&A, 550, A131
Pointecouteau E., Arnaud M., Pratt G. W., 2005, A&A, 435, 1
Pratt G. W., et al., 2010, A&A, 511, A85
Rines K., Geller M. J., Diaferio A., 2010, ApJ, 715, L180
Russell H. R., et al., 2012, MNRAS, 423, 236
Sanders J. S., Fabian A. C., 2013, MNRAS, 429, 2727
Sarazin C. L., 1988, X-ray Emission from Clusters of Galaxies, Cambridge University Press
Shimwell A. C. T. W., et al., 2013, MNRAS, 433, 2036
Sifón C., et al., 2013, ApJ, 772, 25
Sivia D. S., Skilling J., 2005, Data Analysis A Bayesian tutorial, Oxford University Press
Snowden S. L., 2009, SSRv, 143, 253
Sunyaev R. A., Zeldovich Y. B., 1970, CoASP, 2, 66
Tawa N., et al., 2008, PASJ, 60, 11
Vikhlinin A., Markevitch M., Murray S. S., Jones C., Forman W., Van Speybroeck L., 2005, ApJ, 628, 655
Vikhlinin A., Kravtsov A., Forman W., Jones C., Markevitch M., Murray S. S., Van Speybroeck L., 2006, ApJ, 640, 691