A&A 577, A83 (2015)
DOI: 10.1051/0004-6361/201424915

Astronomy & Astrophysics

Can we constrain the interior structure of rocky exoplanets from mass and radius measurements?★

Caroline Dorn[1], Amir Khan[2], Kevin Heng[1], James A. D. Connolly[3], Yann Alibert[1], Willy Benz[1], and Paul Tackley[2]

[1] Physics Institute, University of Bern, Sidlerstrasse 5, 3012 Bern, Switzerland
e-mail: caroline.dorn@space.unibe.ch
[2] Institute of Geophysics, ETH Zürich, Sonneggstrasse 5, 8092 Zürich, Switzerland
[3] Institute of Geochemistry and Petrology, ETH Zürich, Clausiusstrasse 25, 8092 Zürich, Switzerland

Received 4 September 2014 / Accepted 10 February 2015

ABSTRACT

Aims. We present an inversion method based on Bayesian analysis to constrain the interior structure of terrestrial exoplanets, in the form of chemical composition of the mantle and core size. Specifically, we identify what parts of the interior structure of terrestrial exoplanets can be determined from observations of mass, radius, and stellar elemental abundances.

Methods. We perform a full probabilistic inverse analysis to formally account for observational and model uncertainties and obtain confidence regions of interior structure models. This enables us to characterize how model variability depends on data and associated uncertainties.

Results. We test our method on terrestrial solar system planets and find that our model predictions are consistent with independent estimates. Furthermore, we apply our method to synthetic exoplanets up to 10 Earth masses and up to 1.7 Earth radii, and to exoplanet Kepler-36b. Importantly, the inversion strategy proposed here provides a framework for understanding the level of precision required to characterize the interior of exoplanets.

Conclusions. Our main conclusions are (1) observations of mass and radius are sufficient to constrain core size; (2) stellar elemental abundances (Fe, Si, Mg) are principal constraints to reduce degeneracy in interior structure models and to constrain mantle composition; (3) the inherent degeneracy in determining interior structure from mass and radius observations does not only depend on measurement accuracies, but also on the actual size and density of the exoplanet. We argue that precise observations of stellar elemental abundances are central in order to place constraints on planetary bulk composition and to reduce model degeneracy. We provide a general methodology of analyzing interior structures of exoplanets that may help to understand how interior models are distributed among star systems. The methodology we propose is sufficiently general to allow its future extension to more complex internal structures including hydrogen- and water-rich exoplanets.

Key words. planets and satellites: terrestrial planets – planets and satellites: interiors

1. Introduction

Major advances in the detection and characterization of exoplanets have been achieved over the past decade. To date, several hundred have been characterized in terms of mass and radius by space-based or ground-based observations. Continued improvements in observational techniques allow more precise inference of mass and radius. To date, even a few small-mass exoplanets (<10 Earth masses ($M_{\rm E}$)) with uncertainties in mass and radius below 20% and 10%, respectively, have been discovered (e.g., Kepler-36b, CoRot-7b) (e.g., Pepe et al. 2013). Space-based missions, like *Kepler* (Borucki et al. 2010) and CoRoT (Bordé et al. 2003), have been able to measure the transits of these objects and infer the radius. Within the first 16 months of the *Kepler* mission alone, 207 planetary candidates of radius $R < 1.25$ Earth radii ($R_{\rm E}$) and 680 super-Earth-sized ($1.25R_{\rm E} < R < 2R_{\rm E}$) planetary candidates were reported by Batalha et al. (2013).

From the knowledge of mass and radius we are able to derive constraints on the interior structure of exoplanets. Previous studies concerned with rocky exoplanets (e.g., Valencia et al. 2006; Sotin et al. 2007; Seager et al. 2007; Fortney et al. 2007; Wagner et al. 2011; Mocquet et al. 2014) have generally concentrated on computing mass-radius relations based on terrestrial-type interior structures and compositions. These studies show that different interior models are capable of explaining the observations within their uncertainties. Such "forward" approaches, however, do not quantify the inherent degeneracy of interior structure models. For example, in Valencia et al. (2006, 2007b), Sotin et al. (2007), Wagner et al. (2011) it has been shown that different core sizes and mantle compositions affect the mass-radius relationship, but to our knowledge no comprehensive study of the degeneracy of these model parameters has been performed. As a consequence, it is not fully understood which interior-structure parameters can be estimated from the data and which structure parameters tend to be strongly correlated.

In the light of the inherent ambiguity in determining interior structure from mass and radius (e.g., Howe et al. 2014; Rogers & Seager 2010), analysis is most sensibly conducted in an inverse sense (Mosegaard & Tarantola 2002). We therefore propose a complete Bayesian inverse analysis by employing a Markov chain Monte Carlo (MCMC) method to provide full probability distributions for the model parameters of interest (mantle composition and core size).

Early studies of mass-radius relations (Zapolsky & Salpeter 1969) considered only homogeneous and simple monoatomic bulk compositions (H, He, C, Fe, Mg, and H/He mixtures). More

★ Appendices are available in electronic form at http://www.aanda.org recent studies assume differentiated terrestrial-like compositions (e.g., Valencia et al. 2006; Sotin et al. 2007; Seager et al. 2007; Fortney et al. 2007). Based on solar system observations and planet formation and differentiation models, terrestrial planets are generally thought to be differentiated in an iron-rich core, a silicate mantle and a crust (see Howe et al. 2014, for a review of previous and current work). In this work, we focus on terrestrial-type rocky planets and assume that these consist of a pure iron core, a silicate mantle comprising the oxides Na_2O-CaO-FeO-MgO-Al_2O_3-SiO_2, and that volatiles have a negligible effect on mass and radius.

In order to be able to compute the radius for a planet of a given mass and composition (the "forward" model) we solve the hydrostatic equilibrium equation coupled with a thermodynamic approach based on Gibbs free-energy minimization and Equation-of-State (EoS) modeling. In this regard, our work is similar to earlier studies (Valencia et al. 2007a; Sotin et al. 2007; Seager et al. 2007; Fortney et al. 2007), except for the use of thermodynamic modeling. This addition allows us to compute the density profile of the planet for an arbitrary bulk silicate mantle composition from first principles of thermodynamics. Moreover, given that cosmochemically derived elemental abundances (e.g., Fe, Mg, and Si) have proved important as constraints on the composition of the terrestrial planets (e.g., Grasset et al. 2009; Rogers & Seager 2010), we investigate the effect of including abundance constraints in determining interior structure.

Although similarities between the present work and the Bayesian analysis of Rogers & Seager (2010) exist, our approach nonetheless differs in several respects:

- our inversion method is applicable to planets of any mantle composition and core structure;
- our inversion method is general in that it can be easily adapted to more complex planetary models that include oceans and atmospheres;
- we infer probability distributions on interior-structure parameters (core radius and mantle composition) for a wide range of masses ($0.1M_E < M < 10M_E$) and radii ($0.5R_E < R < 1.7R_E$);
- we demonstrate how observations of a few key elemental abundances (e.g., Fe, Mg, Si) of the host star's photosphere significantly reduces the degeneracy of interior structure models;
- we quantify the influence of measurement uncertainties on the determination of interior structure models.

The outline of this study is as follows: in Sect. 2 we introduce the physical model that links data and model parameters, and outline the inversion strategy. In Sect. 3, we first test our method on Earth using Earth bulk elemental abundances derived from geochemical analysis of mantle rocks. Next, we apply the method to all terrestrial solar system planets. In Sect. 4, we present results for synthetic exoplanets and Kepler-36b followed by discussion and conclusions.

2. Methodology

2.1. Abundance constraints

In summary, the importance of additional constraints in the form of chemical abundances of the elements Fe, Mg, Si are investigated in detail. Arguments from planet formation favor certain interior compositions as discussed by e.g., Grasset et al. (2009), Rogers & Seager (2010), Madhusudhan et al. (2012).

Observations and theoretical considerations suggest that exoplanet bulk composition is dictated by stellar composition: first, relative elemental abundances of Fe, Mg, and Si are similar among the Sun, Earth, Mars, the Moon as well as meteorites (Lodders 2003; Drake & Righter 2002; McDonough & Sun 1995). Meteorites are believed to be chemically similar to the building blocks of planets (Morgan & Anders 1980), both being condensates from the solar nebula that experienced the same fractionation processes. Second, it has been demonstrated that during planet formation, planetary bulk and stellar ratios of Fe/Si and Mg/Si are very similar (e.g., Bond et al. 2010; Elser et al. 2012; Johnson et al. 2012; Thiabaud et al. 2013) since the refractory elements Fe, Si, and Mg condense at similar temperatures corresponding to small distances to the host star (~1 AU).

For the Sun and our solar system planets, bulk Fe/Si and Mg/Si (henceforth $\mathrm{Fe/Si}_{\bigcirc}$ and $\mathrm{Mg/Si}_{\bigcirc}$) have been inferred from photospheric analysis and from geochemical analysis of meteoritic and cosmochemical material, respectively. For extrasolar systems, we shall assume that $\mathrm{Fe/Si}_{\bigcirc}$ and $\mathrm{Mg/Si}_{\bigcirc}$ equal to their stellar counterparts $\mathrm{Fe/Si}_{\bigstar}$ $\mathrm{Mg/Si}_{\bigstar}$.

2.2. Data

The aim of our study is to characterize the interior structure of terrestrial exoplanets based on inferred mass and radius from radial velocity and transit observations and constraints on bulk Fe, Mg, and Si abundances ($\mathrm{Fe/Si}_{\bigcirc}$ and $\mathrm{Mg/Si}_{\bigcirc}$) inferred from stellar photospheric analysis. Mass, radius and bulk abundance constraints constitute our data **d**. For the purpose of illustrating the relative importance of data on inverted model parameters, we consider the following two cases:

Case A Data are mass M and radius R.
Case B Data are mass M, radius R, and compositional constraints ($\mathrm{Fe/Si}_{\bigcirc}$ and $\mathrm{Mg/Si}_{\bigcirc}$) obtained from stellar spectroscopic observations of the host star.

$\mathrm{Fe/Si}_{\bigcirc}$ expresses the mass ratio between the mass of iron to silicate for the entire planet (core and mantle). Given that we assumed a pure iron core, this ratio is equal to the mass of iron in core and mantle divided by the mass of silicate in the mantle $M_{\mathrm{si,mantle}}$. This can be written as

$$\mathrm{Fe/Si}_{\bigcirc} = \mathrm{Fe/Si}_{\circledcirc} + M_{\mathrm{core}}/M_{\mathrm{Si,mantle}}, \qquad (1)$$

where M_{core} is mass of iron in the core, and $\mathrm{Fe/Si}_{\circledcirc}$ denotes the mass ratio of Fe/Si in the mantle. Mantle Mg/Si is referred to as $\mathrm{Mg/Si}_{\circledcirc}$. Since all Si and Mg is assumed to be in the mantle, $\mathrm{Mg/Si}_{\bigcirc} = \mathrm{Mg/Si}_{\circledcirc}$.

2.3. Model parameterization

Our exoplanet interior model consists of a layered sphere with an iron core surrounded by a silicate mantle (see Fig. 1). Model parameters include core size R_c and mantle silicate composition using the NCFMAS model chemical system that comprises the oxides Na_2O-CaO-FeO-MgO-Al_2O_3-SiO_2 and accounts for more than 98% of the mass of Earth's mantle (Irifune et al. 1994) (see Appendix C). We parameterize mantle composition by mantle Si-content defined as $M_{\mathrm{Si,mantle}}/M_{\mathrm{mantle}}$, (referred to as $\mathrm{Si}_{\circledcirc}$), $\mathrm{Fe/Si}_{\circledcirc}$ and $\mathrm{Mg/Si}_{\circledcirc}$. Na_2O, CaO, and Al_2O_3 are minor components and their relative abundance can be constrained from stellar photospheric observations (e.g., Elser et al. 2012). In this

Table 1. Prior ranges of model parameters for case A (Beirao et al. 2005; Gilli et al. 2006; Thiabaud et al. 2013) and case B (Lodders 2003).

Model parameter of **m**	Case A	Case B
$\mathrm{Si}_{\circledcirc}$	8–34 (non-uniform†)	14–30 (non-uniform†)
$\mathrm{Fe/Si}_{\circledcirc}$	0–4.15	0–2.23
$\mathrm{Mg/Si}_{\circledcirc}$	0.61–2.32	0.89 ± 0.08 (Gaussian)
Core radius R_c	0–R	0–R

Notes. Uniform distributions unless stated otherwise. Si, Fe, and Mg are in wt% (weight percentage). Minor components of Na_2O, CaO, and Al_2O_3 are fixed to chondritic abundances of 2 wt%, 2 wt% and 4.8 wt%, respectively (Lodders 2003). $^{(\dagger)}$ See Appendix A.

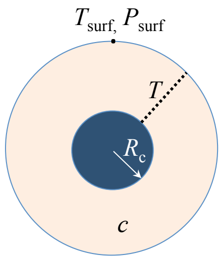

Fig. 1. Illustration of model parameterization. Parameters are core radius R_c and mantle composition c comprising the oxides Na_2O-CaO-FeO-MgO-Al_2O_3-SiO_2. Mantle temperature profile T, surface temperature T_{surf} and pressure P_{surf} are fixed input parameters. In the iron core an adiabatic temperature profile is used.

study, we assume chondritic abundances (see Table 1). The set of parameters that constitutes our model parameter **m** is thus

– core radius R_c,
– mantle Si-content $\mathrm{Si}_{\circledcirc}$,
– mantle $\mathrm{Fe/Si}_{\circledcirc}$,
– mantle $\mathrm{Mg/Si}_{\circledcirc}$.

The mantle composition parameters $\mathrm{Si}_{\circledcirc}$, $\mathrm{Fe/Si}_{\circledcirc}$, and $\mathrm{Mg/Si}_{\circledcirc}$ depend on each other, i.e., the fractions of oxides in the NCFMAS system must sum up to one. All parameters are assumed to be uniformly distributed unless stated otherwise (see Table 1).

2.4. Prior information

For case A, prior bounds have been chosen based on (1) the assumption that the bulk metallic composition of a planet is similar to the abundance of the host star (Thiabaud et al. 2013) and (2) the fact that chemical abundances of observed stars with planetary companions seem to fall within certain ranges (e.g., Beirao et al. 2005; Gilli et al. 2006). Thus, the variability of observed elemental abundances in stars defines the prior bounds on $\mathrm{Fe/Si}_{\circledcirc}$ and $\mathrm{Mg/Si}_{\circledcirc}$.

For case B, prior bounds on $\mathrm{Fe/Si}_{\circledcirc}$ and $\mathrm{Mg/Si}_{\circledcirc}$ are derived from $\mathrm{Fe/Si}_{\bigcirc}$ and $\mathrm{Mg/Si}_{\bigcirc}$ inferred from the host star's photospheric abundance: (1) Since all Si and Mg is assumed to be in the mantle, $\mathrm{Mg/Si}_{\bigcirc}$ defines the prior bounds on $\mathrm{Mg/Si}_{\circledcirc}$; (2) Fe, on the other hand, is distributed between core and mantle. Thus, the bulk constraint $\mathrm{Fe/Si}_{\bigcirc}$ defines only the upper bound of the prior on $\mathrm{Fe/Si}_{\circledcirc}$. For exoplanets, $\mathrm{Fe/Si}_{\bigcirc}$ and $\mathrm{Mg/Si}_{\bigcirc}$ are defined by the host star's photospheric elemental abundance since we assume that $\mathrm{Fe/Si}_{\bigcirc} = \mathrm{Fe/Si}_{\bigstar}$ and $\mathrm{Mg/Si}_{\bigcirc} = \mathrm{Mg/Si}_{\bigstar}$. $\mathrm{Mg/Si}_{\bigstar}$ is Gaussian-distributed. We consider solar estimates for Fe, Si, and Mg (Lodders 2003) and associated uncertainties as representative of $\mathrm{Fe/Si}_{\bigcirc}$ and $\mathrm{Mg/Si}_{\bigcirc}$ throughout this study unless stated otherwise.

In both cases, the prior bounds on $\mathrm{Si}_{\circledcirc}$ depend on the prior bounds of $\mathrm{Fe/Si}_{\circledcirc}$ and $\mathrm{Mg/Si}_{\circledcirc}$, since the fractions of oxides in the NCFMAS chemical system must sum to one. This implies a non-uniform prior of $\mathrm{Si}_{\circledcirc}$ (see Appendix A).

Additional input parameters that are fixed in this study are (1) surface pressure and temperature, (2) thermal gradient in the mantle based on adiabatic gradient of Earth's mantle, and (3) absolute amount of minor components (Na_2O, CaO, and Al_2O_3) in the mantle.

Using a fixed temperature profile for the mantle is demonstrably not a limitation of our model. Many authors have shown that the variations due to temperature for rocky planets are small compared to uncertainties of material properties of silicate or iron compositions (e.g., Sotin et al. 2007; Seager et al. 2007; Grasset et al. 2009; Valencia et al. 2010). This is demonstrated in Appendix B where it is shown that varying mantle temperatures introduces little variation in computed densities. As a consequence, we fix temperature profiles by using Earth's surface temperature, lithosphere and mantle temperature gradient.

2.5. Interior structure model

Data and model parameters are linked by a physical model embodied by the forward operator $g(\cdot)$

$$\mathbf{d} = g(\mathbf{m}). \tag{2}$$

For a given model **m**, interior structure (density profile), the total mass, and $\mathrm{Fe/Si}_{\bigcirc}$ are computed for the purpose of comparing with observed data **d**. Equation (2) represents the forward problem, which is computed using thermodynamic method and Equation-of-State (EoS) modeling. The complete solution of the forward problem is summarized as

$$\{c, R_c, T\} \xrightarrow{g_1} \mathcal{M} \xrightarrow{g_2} \rho \xrightarrow{g_3} \{M, R\}, \tag{3}$$

where c denotes the NCFMAS mantle composition, T temperature, $\mathcal{M}$ equilibrium mantle mineralogy, and ρ density (all parameters are functions of radius). The forward operator g_1 embodies Gibbs free-energy minimization that computes stable mineral modes as a function of P (pressure), T, and composition, g_2 calculates density, g_3 integrates ρ over R to compute M. We assume subsolidus conditions throughout.

2.5.1. Thermodynamic modeling

Possible mantle compositions are explored within the Na_2O-CaO-FeO-MgO-Al_2O_3-SiO_2 model system. To compute the mantle density profile, we employ a self-consistent thermodynamic method. For this purpose we assume: thermodynamic equilibrium; mantle mineral phases of large exoplanets are those that potentially occur in the deep mantle of the Earth; and, when required, the thermodynamic properties of the mineralogy can

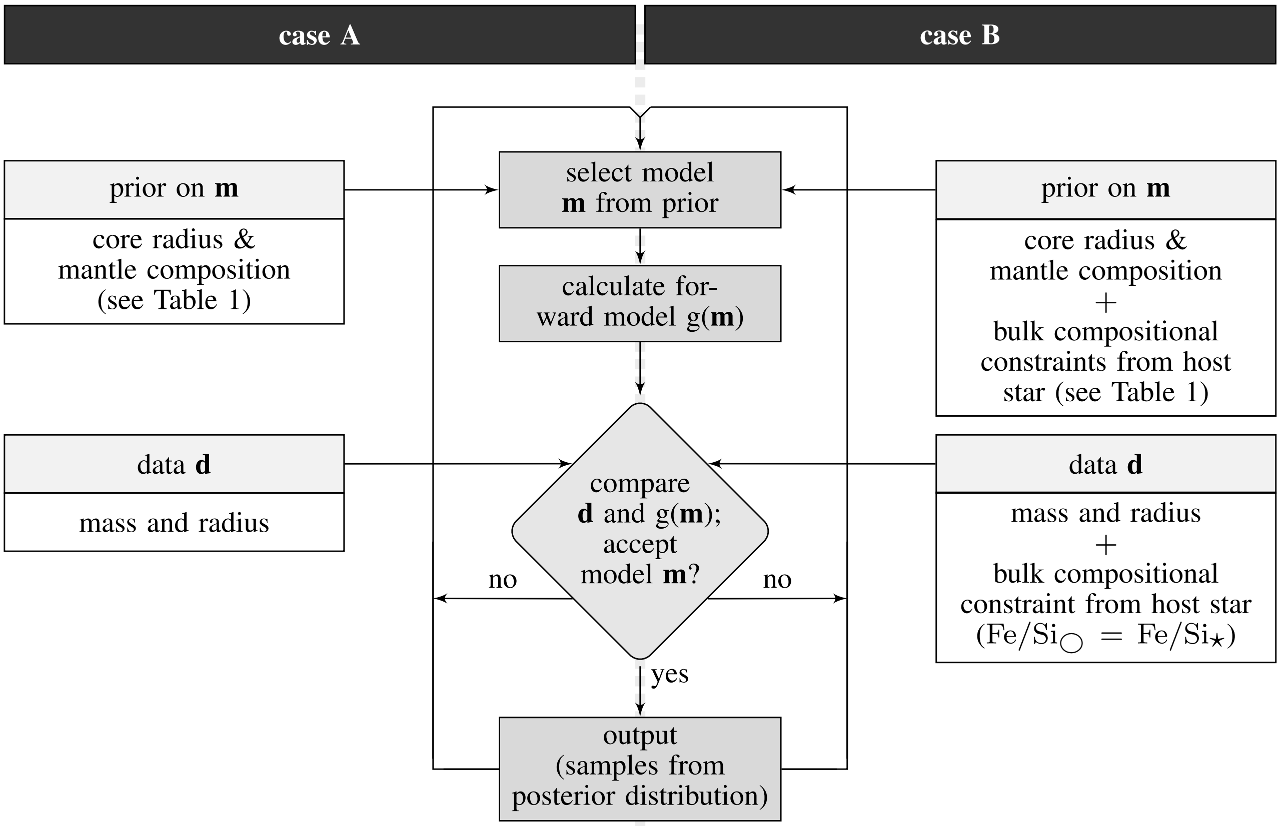

Fig. 2. Inversion scheme. The bulk compositional constraints from the host star (case B) used both in prior and data are independent. The iterative scheme is repeated until the output converges to the posterior distribution. See Sect. 2 for more details.

be extrapolated to more extreme conditions than realized on Earth. The equilibrium mineralogy and, consequently its density, is computed as a function of pressure, temperature, and bulk composition by Gibbs energy minimization (Connolly 2009). For these calculations the pressure is obtained by integrating the load from the surface boundary and temperature is obtained by integrating an Earth-like temperature gradient from an arbitrarily assumed surface temperature. Mineral equations of state and parameters are as given by Stixrude & Lithgow-Bertelloni (2005, 2011). Because the equilibrium assumption is dubious at low temperature (e.g., Wood & Holloway 1984), for models that require rock properties at temperatures below 800 K, the stable mineralogy is first calculated at 800 K and its physical properties are then computed at the temperature of interest.

2.5.2. EoS for iron core

In order to compute the core density profile, we use the EoS for pure iron derived by Belonoshko (2010), which is similar to the Mie-Grüneisen-Debye EoS (Jackson & Rigden 1996), except for a different thermal pressure term. This EoS for iron is preferred, as it allows us to closely reproduce high-precision volumetric experiments of iron under high-pressure and high-temperature conditions.

The adiabatic gradient in the core is computed as

$$\frac{\mathrm{d}T}{\mathrm{d}P} = \frac{\gamma T}{K_S}, \qquad (4)$$

where K_S is adiabatic bulk modulus, which is related to the isothermal bulk modulus K_T through $K_S = (1 + \gamma\alpha T)\, K_T$. The parameters of the Mie-Grüneisen-Debye equation $\gamma = \gamma_0 \left(\frac{\rho}{\rho_0}\right)^{-q}$ for iron are $\rho_0 = 8.334$ g/cm^3, $q = 0.489$, and $\gamma_0 = 2.434$. Density ρ and thermal expansion coefficient α are then directly computed from Belonoshko's EoS.

2.6. Inversion method

Probabilistic inversion methods are suitable when data are sparse, the physical model (Eq. (2)) is highly non-linear and/or it is expected that very different models are consistent with information and data (e.g., Mosegaard & Tarantola 1995). While stochastic sampling-based approaches for high-dimensional problems are computationally expensive, numerous global search methods exist. These include MCMC methods, simulated annealing or genetic algorithms. MCMC has the advantage that parameter uncertainties are formally assessed. Furthermore, the MCMC method generates samples whose distribution converges to a stationary distribution coinciding with the posterior probability density function (PDF) (e.g., Hastings 1970). Although the dimensionality of the presented inverse problem is low (4 parameters), offering the possibility of performing grid-search, we nonetheless employ MCMC because we intend to expand the complexity of the problem to higher dimensions, in the future. For consistency, we have also used the grid-search method to verify our results.

2.6.1. Bayesian analysis

We employ a Bayesian method to compute the posterior PDF for each model parameter (R_c, $\mathrm{Si}_{\circledcirc}$, $\mathrm{Fe/Si}_{\circledcirc}$, $\mathrm{Mg/Si}_{\circledcirc}$) from data ($M$, R and $\mathrm{Fe/Si}_{\bigcirc}$) and prior information. According to Bayes' theorem, the posterior distribution for a fixed model parameterization $\mathbf{m}$, conditional on data $\mathbf{d}$, is given by (Tarantola & Valette 1982)

$$p(\mathbf{m}|\mathbf{d}) \propto p(\mathbf{m})\, L(\mathbf{m}|\mathbf{d}), \tag{5}$$

where $p(\mathbf{m})$ represents prior information on model parameter $\mathbf{m}$. $L(\mathbf{m}|\mathbf{d})$ is the likelihood function and can be interpreted in probabilistic terms as a measure of how well a model $\mathbf{m}$ fits data $\mathbf{d}$.

Assuming uncorrelated and normally-distributed residuals (see Eq. (7)), the likelihood defines a "measure" between $\mathbf{d}$ and $\mathbf{m}$ and can be written as

$$L(\mathbf{m}|\mathbf{d}) = \frac{1}{(2\pi)^{N/2}(\prod_{i=1}^{N}\sigma_i^2)^{1/2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{N}\frac{Q_i}{\sigma_i^2}\right), \tag{6}$$

where N is total number of data points, σ_i is the estimated error in the ith datum, and Q_i is data misfit, or squared residual, given by

$$Q_i = (g_i(\mathbf{m}) - \mathbf{d}_\mathrm{i})^T (g_i(\mathbf{m}) - \mathbf{d}_\mathrm{i}), \tag{7}$$

where $g(\mathbf{m})$ is the forward model discussed in Sect. 2.5.

2.6.2. Markov chain Monte Carlo (MCMC)

For high-dimensional and non-linear inverse problems or sparse data, it is in practice impossible to derive the posterior distribution analytically. MCMC techniques offer an efficient method of performing Bayesian analysis (Eq. (5)). MCMC sampling methods iteratively search the space of feasible solutions. The iteration scheme used is based on the Metropolis algorithm, which proceeds as follows:

1. An initial starting model $\mathbf{m}_\mathrm{old}$ is drawn at random by sampling from the prior distribution.
2. The posterior density of $\mathbf{m}_\mathrm{old}$ is calculated by evaluating the product of the likelihood of the corresponding forward model and prior density.
3. The current model parameter is perturbed to obtain $\mathbf{m}_\mathrm{new}$, which is subsequently created from a proposal distribution. Here, the proposal distribution is uniformly bounded and centered around $\mathbf{m}_\mathrm{old}$. Generally, it is chosen such that the size of the model perturbation allows for a reasonable rate of accepted transitions in the MCMC procedure, typically around 30% (Gilks et al. 1996).
4. The proposal $\mathbf{m}_\mathrm{new}$ is accepted with probability (Mosegaard & Tarantola 1995):
   $P_\mathrm{accept} = \min\{1, \exp(l(\mathbf{m}_\mathrm{new}|\mathbf{d}) - l(\mathbf{m}_\mathrm{old}|\mathbf{d}))\}$.
5. If the proposal is accepted the Markov chain moves to $\mathbf{m}_\mathrm{new}$, otherwise the chain remains at $\mathbf{m}_\mathrm{old}$.
6. Steps 2 to 5 are repeated.

After many iterations, the accepted samples that are generated with this approach are (1) independent of the starting model and (2) distributed according to the posterior distribution. The efficiency of sampling is strongly dependent on the proposal distribution. If this distribution is badly chosen, the acceptance rate of solutions might be unacceptably low, resulting in very poor efficiency regarding convergence (Hastings 1970). On the contrary, if the proposal distribution is well chosen, the MCMC sampler will rapidly explore the posterior target distribution. Here, the posterior information is gathered from a large number of sampled models ($\sim 10^4$). Only the statistical nature of sampled model features are of interest. Presently, we will concentrate on computing marginal posterior distributions in order to provide the reader with a notion of model parameter uncertainties. Information on single parameters are obtained by one-dimensional (1D) marginals; higher dimensional marginal PDFs reveal the correlation that exists among several parameters. If the data are able to constrain the model (i.e., informative), differences between prior and posterior PDFs are expected.

Given the large number of models that have to be computed, the calculation of the forward model must be computationally very efficient. In our work, generating the planet's internal structure takes on average 0.8 s of CPU time on a four quad-core AMD Opteron 8380 CPU node with 32 GB of RAM. In all, we sampled about 10^7 models and retained around 10^4 models for further analysis.

3. Method validation

3.1. Synthetic case

We apply our method to a synthetic planet that ultimately serves as a benchmark test. Structure and composition of the synthetic model are defined as:

- $\mathrm{Si}_{\circledcirc} = 23.4$ wt%;
- $\mathrm{Fe/Si}_{\circledcirc} = 0.5$; $\mathrm{Fe/Si}_{\bigcirc} = 1.69$;
- $\mathrm{Mg/Si}_{\circledcirc} = \mathrm{Mg/Si}_{\bigcirc} = 0.89$;
- $R_\mathrm{c} = 0.47\,R$.

Using the interior structure model (Sect. 2.5) we calculate the synthetic data to be $M = 1 M_\mathrm{E}$ and $R = 1 R_\mathrm{E}$. For the inversion we consider the following data uncertainties: $\sigma_R = 0$, $\sigma_M = 0.001M$, together with $\mathrm{Fe/Si}_{\bigstar} = 1.69 \pm 0.18$ and $\mathrm{Fe/Si}_{\bigstar} = 0.89 \pm 0.08$. Results are shown in Fig. 3 and demonstrate the ability of the method to retrieve the internal structure of the synthetic planet. As expected, the most likely posterior model overlaps with the true model. The inherent degeneracy in interior structure is apparent from the spread of the posterior.

3.2. Solar system planets

We test our method on the terrestrial solar system planets Earth, Moon, Mars, Venus, and Mercury, for which a series of independent model parameter estimates are available from a host of geophysical, geo-, and cosmochemical data gathered in situ as well as from orbiting spacecraft (see Table 2). These tests are important inasmuch as they allow us to benchmark our method. For simplicity, Table 2 only summarizes a single best estimate of independent interior models. Associated uncertainties are very different depending on data and model assumptions. For example, core radius R_c of the Moon, Mars and Mercury, is estimated to be 310−320 km (Kuskov & Kronrod 2001), 1680 ± 150 km (Khan & Connolly 2008), and 2020 ± 100 (Padovan et al. 2014), respectively.

In the following, we extensively test different scenarios for the Earth: case A and B. Input for cases A and B is defined in Sect. 2. These cases show that the compositional constraints $\mathrm{Fe/Si}_{\bigcirc}$ and $\mathrm{Mg/Si}_{\bigcirc}$ are crucial for determining interior structure and that $\mathrm{Fe/Si}_{\bigstar}$ and $\mathrm{Mg/Si}_{\bigstar}$ are good proxies for them. How well data for case B allow us to recover the interior structure is demonstrated for the Moon, Venus, Mars, and Mercury. For all terrestrial planets we use 1% and 0% uncertainties on observed M and R, respectively, and $\mathrm{Fe/Si}_{\bigstar} = 1.69 \pm 0.18$ (Lodders 2003). Minor components of Na_2O, CaO, and Al_2O_3

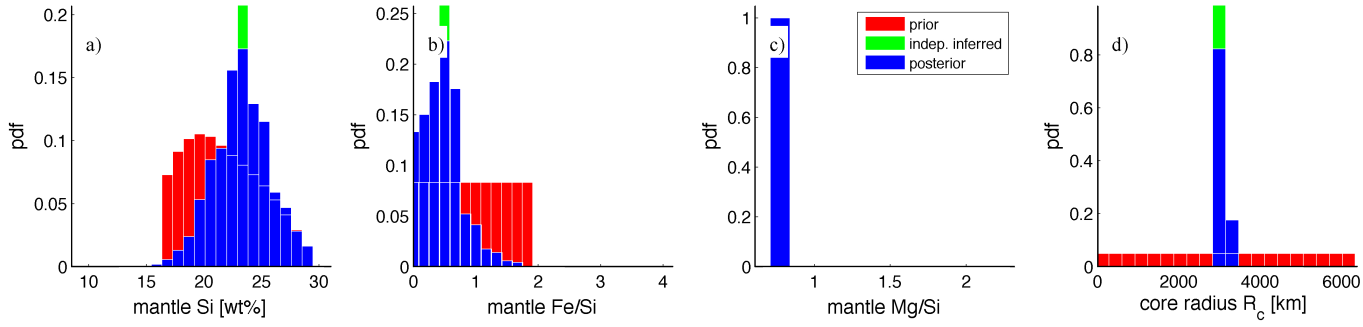

Fig. 3. Sampled 1D marginal posterior distributions for the synthetic test where model parameters are inverted for an Earth-sized planet: prior (red), posterior (blue) and independent estimates (green) of model parameters: **a)** mantle Si-content $\mathrm{Si}_{\circledcirc}$; **b)** mantle $\mathrm{Fe/Si}_{\circledcirc}$; **c)** mantle $\mathrm{Mg/Si}_{\circledcirc}$, and **d)** core radius R_c.

Table 2. Independent best estimates of model parameters for terrestrial solar system planets.

Planet	$\mathrm{Si}_{\circledcirc}$ [wt%]	$\mathrm{Fe/Si}_{\circledcirc}$	$\mathrm{Mg/Si}_{\circledcirc}$	R_c [km]	Reference
Earth	22	0.17	0.83	3400	PREM
Moon	23	0.4–0.47	0.8	310–320	Kuskov & Kronrod (2001)
Mars	21	0.64	0.97	1680	Khan & Connolly (2008)
Venus	–	–	–	3100	Aitta et al. (2012)
Mercury	–	–	–	2020	Padovan et al. (2014), Hauck et al. (2013)

Notes. The preliminary reference Earth model (PREM) refers to Dziewonski & Anderson (1981).

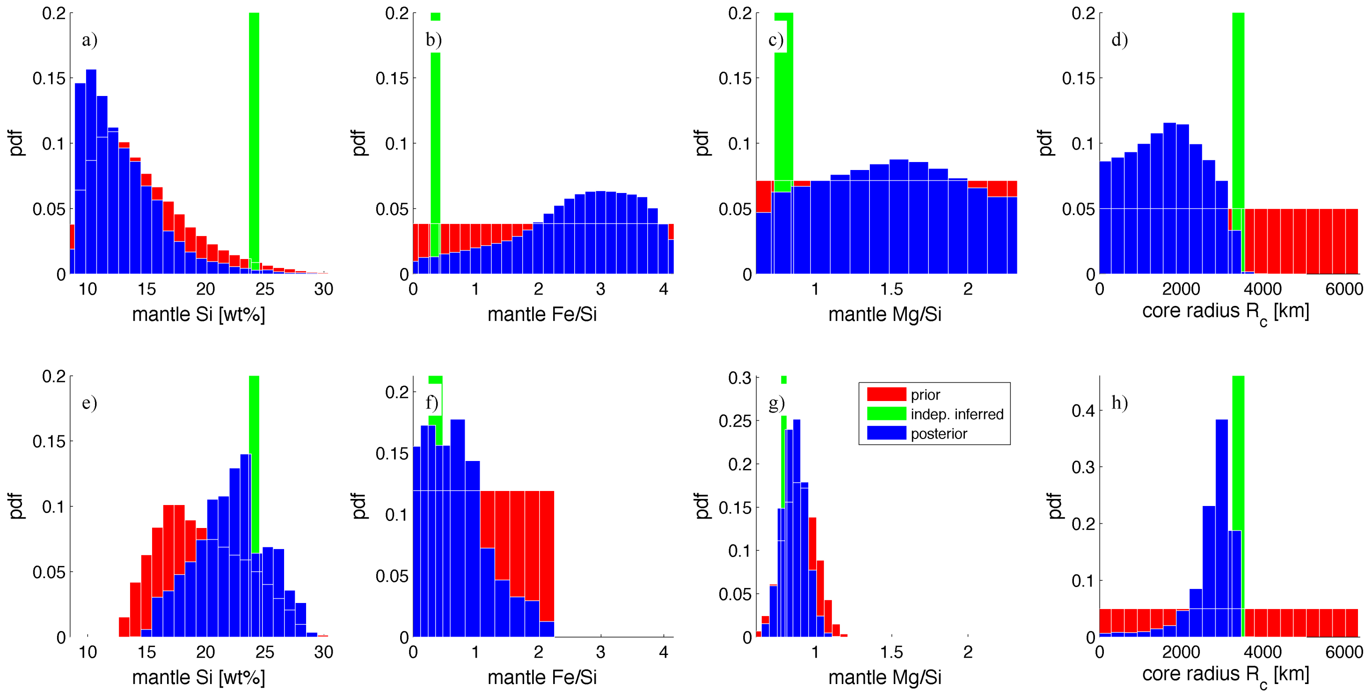

Fig. 4. Sampled 1D marginal posterior distributions for Earth, **a)–d)** case A, **e)–h)** case B: prior (red), posterior (blue) and independent estimates (green) of model parameters: **a)**, **e)** mantle Si-content $\mathrm{Si}_{\circledcirc}$, **b)**, **f)** mantle $\mathrm{Fe/Si}_{\circledcirc}$, **c)**, **g)** mantle $\mathrm{Mg/Si}_{\circledcirc}$, and **d)**, **h)** core radius R_c. Independent estimates are listed in Table 2. Prior ranges are different for the two cases, but axis ranges are kept the same to ease comparison.

are fixed to chondritic values (Lodders 2003). Prior ranges for the various parameters are listed in Table 1.

Earth. Results for the Earth for cases A and B are shown in Fig. 4. Independent estimates of mantle composition and core size (Table 2) are shown as green bars. Inversion of only mass and radius (case A) provides little information on mantle composition, but some constraints on R_c. In contrast, addition of $\mathrm{Fe/Si}_{\bigcirc}$ and $\mathrm{Mg/Si}_{\bigcirc}$ (case B) enables us, as expected, to constrain mantle composition and thereby obtain improved estimates of core size. Moreover, better agreement with the independent estimates is also observed.

There is an underestimation of the core radius because of our assumed core composition. In the PREM model (Dziewonski & Anderson 1981), the core is less dense which reflects the presence of lighter elements (e.g., S, Si, C, O) in addition to Fe-Ni. Since we consider pure iron, we systematically overestimate core density and therefore underestimate core size. For both cases A and B, core sizes are $0 \leq R_c \leq 2930$ km and $1470 \leq R_c \leq 3300$ km (95% credible intervals), respectively. Furthermore, we observe that $Mg/Si_{⊚}$ is not well constrained by data.

Figure 5 shows the correlation between individual model parameters for cases A and B. In contrast to case A, strong correlations in case B are evident between model parameters: $Si_{⊚}$, $Fe/Si_{⊚}$, and R_c. This is due to the constraint provided by $Fe/Si_{\bigcirc}$ that couples core and mantle, i.e., a larger core must be compensated by a lower $Fe/Si_{⊚}$ or a larger $Si_{⊚}$. No significant correlation of $Mg/Si_{⊚}$ to the other model parameters is observed because data are not able to constrain $Mg/Si_{⊚}$ well.

Moon. Results for the Moon are shown in Fig. 6. As for Earth, we see that $Si_{⊚}$, $Fe/Si_{⊚}$, and R_c are well-constrained, i.e., prior and posterior PDF significantly differ, except for $Mg/Si_{⊚}$. Predicted core size matches the independent estimate (Table 2) reasonably well with $0 \leq R_c \leq 380$ km (95% credible interval), which also assumes a pure iron core. The mantle composition appears not to be fully retrieved here. This is because the Moon has the lowest $Fe/Si_{\bigcirc}$ among the terrestrial planets, including chondrites and satellites of the outer solar system (Kuskov & Kronrod 2001). Therefore, the assumption that $Fe/Si_{\bigcirc} = Fe/Si_{\bigstar}$ leads us to overestimate $Fe/Si_{⊚}$ and underestimate $Si_{⊚}$. The independent estimate of mantle compositions nonetheless lies within the range of inferred posterior models.

Mars. Results for Mars are shown in Fig. 7. Apart from $Mg/Si_{⊚}$, mantle composition ($Si_{⊚}$ and $Fe/Si_{⊚}$) and core radius (R_c) appear to be well-constrained. Although independent geophysical estimates of Khan & Connolly (2008) lie within the posterior distribution, there is discrepancy that relates to the assumption of sulfur in the core model of Khan & Connolly (2008). As a consequence, we underestimate core size to $0 \leq R_c \leq 1330$ km (95% credible interval), which leads to an overestimate of $Fe/Si_{⊚}$ and an underestimate of $Si_{⊚}$. Comparing our results to models with pure iron cores (Rivoldini et al. 2011; Khan & Connolly 2008), reduces discrepancy.

Venus. Detailed independent analyses of the mantle composition of Venus are not yet available. Aitta et al. (2012), e.g., impose an Earth-like mantle composition and find that the composition of Venus' core is likely enriched in lighter elements above that of Earth's core to fit mass and radius. Consequently, as shown in Fig. 8, we underestimate core radius to $0 \leq R_c \leq 2810$ km (95% credible interval) compared to the independent estimates of Aitta et al. (2012). Generally, R_c is well constrained, while the mantle composition ($Si_{⊚}$, $Fe/Si_{⊚}$ and $Mg/Si_{⊚}$) is only moderately to weakly constrained.

Mercury. Mercury has a high bulk density and is thought to have experienced a different accretion scenario or undergone a different post-formation process in comparison to the other terrestrial planets (Cameron 1988; Strom & Sprague 2003). An early giant impact is likely to have stripped off a large part of its silicate crust and mantle (Benz et al. 1988), leaving behind a planet with a large iron core. Hence, the constraint $Fe/Si_{\bigcirc} = Fe/Si_{\bigstar}$ is not applicable to bulk Mercury as a consequence of which it is not possible to match data. We nonetheless show the posterior model parameter PDFs in Fig. 9, but emphasize that these models have near-zero likelihoods. To improve fit to M a larger iron core is required. This, however, would imply a higher mantle Si-content ($Si_{⊚}$) in order to simultaneously match the $Fe/Si_{\bigcirc}$ constraint, but would move $Si_{⊚}$ outside the prior range.

In summary, the addition of $Fe/Si_{\bigcirc}$ as constraint (case B) is clearly the central parameter that allows us to: (1) constrain mantle composition; (2) obtain an improved estimate of core size; and (3) significantly reduce model variability. This ultimately arises because the compositional constraint ($Fe/Si_{\bigcirc}$) results in a strong correlation of the model parameters $Si_{⊚}$, $Fe/Si_{⊚}$ and R_c in the inversion as clearly demonstrated in Fig. 5. The data are not able to constrain $Mg/Si_{⊚}$, since $Mg/Si_{⊚}$ appears not to be significantly correlated to the other model parameters. We find that independent estimates of $Si_{⊚}$, $Fe/Si_{⊚}$ and R_c are relatively well predicted by our method for Earth, Moon, Mars, and Venus. This is not the case for Mercury, where no model is found that fits data given our model assumptions.

4. Results

4.1. Application to synthetic exoplanets

In this section, we apply our method to a set of synthetic planets where mass and radius uncertainties are considered to be artificially small in order to demonstrate its performance. We consider a range of synthetic exoplanets between $0.1M_E < M < 10M_E$ and $0.5R_E < R < 1.7R_E$ for both cases A (inversion of mass and radius) and B (inversion of mass, radius, and bulk abundance constraints). Prior model ranges are listed in Table 1. We employ the same standard deviations used in the case of the terrestrial solar system planets, i.e., $\sigma_M = 0.01M$, $\sigma_R = 0$, and $Fe/Si_{\bigcirc} = Fe/Si_{\bigstar} = 1.69 \pm 0.18$ (Lodders 2003), as well as fixing the minor components (Na_2O, CaO, and Al_2O_3) to chondritic values (Lodders 2003).

Figure 10 shows the range of synthetic rocky planets that fit data for cases A (left panels) and B (right panels). For every synthetic planet we solved an inverse problem, i.e., for every point in the M-R-plane in Fig. 10, a full inverse problem is solved yielding separate PDFs for all parameters. When inverting only M and R (case A), all planets are seen to fall within the extreme limits of a pure iron and a pure silicate sphere (Fig. 10) as expected. By adding the compositional constraint ($Fe/Si_{\bigcirc}$) inferred from observations of the host star (here the Sun), the extreme case of a pure iron sphere, for example, is clearly unable to satisfy this constraint. As a consequence, for case B synthetic planets plot within a smaller range than for case A. Mercury, for example, lies outside the permissible region for reasons discussed above.

With our assumption that planets are composed of a pure iron core surrounded by a rocky mantle of arbitrary size, it is not surprising that internal structures could only be derived for planets with structures lying between a pure iron and a pure rocky sphere. By adding a compositional constraint ($Fe/Si_{\bigcirc}$) the range of possible internal structures is sharply reduced, leading to exclusion of Mercury-type planets.

Between the limits of a pure iron and pure silicate sphere, bulk densities decrease toward the silicate-sphere limit. Thus, we also expect core size to decrease toward the silicate-sphere limit. To verify that this trend is indeed obtained here, we plot inverted core mass fraction M_c/M in Fig. 11 for all considered synthetic exoplanet cases shown in Fig. 10. In the limit of a pure iron (silicate) sphere $M_c/M \rightarrow 1$ (0). For case B, models with

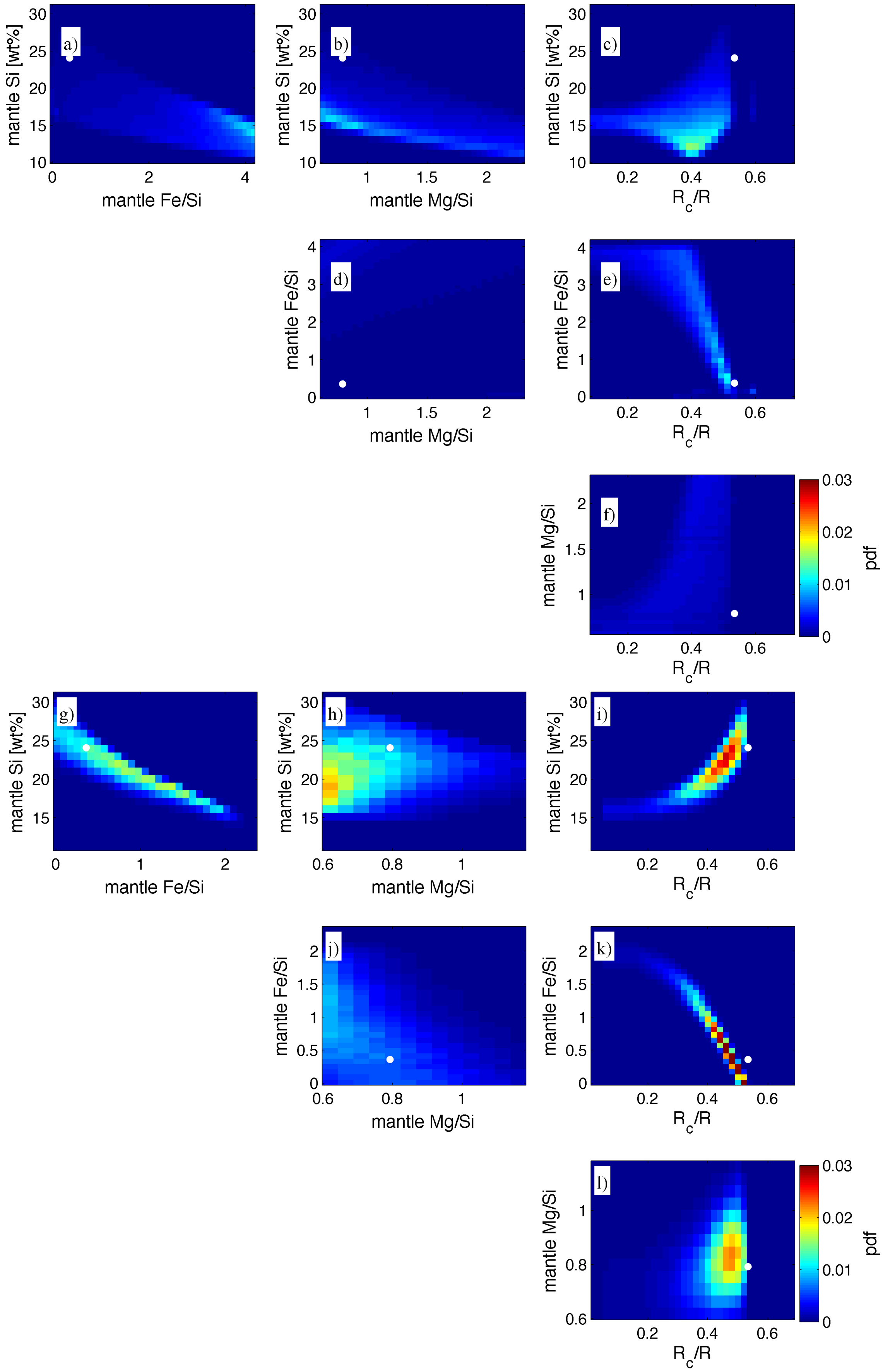

Fig. 5. Sampled 2D marginal posterior distributions for the Earth (case A: plots **a)–f)** and case B: plots **g)–l)**) showing correlation between parameters $\mathrm{Si}_{\circledcirc}$, $\mathrm{Fe/Si}_{\circledcirc}$, $\mathrm{Mg/Si}_{\circledcirc}$, and core radius R_c (corresponding to Fig. 4). Independent estimates (Table 2) are marked in white.

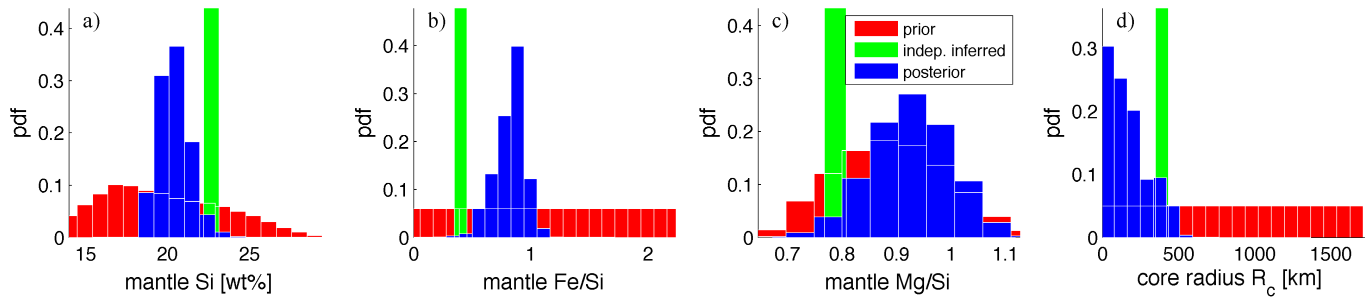

Fig. 6. Sampled 1D marginal posterior distributions for the Moon, case B: prior (red), posterior (blue) and independent estimates (green) of model parameters: **a)** mantle Si-content $\mathrm{Si}_{\circledcirc}$; **b)** mantle $\mathrm{Fe/Si}_{\circledcirc}$; **c)** mantle $\mathrm{Mg/Si}_{\circledcirc}$; and **d)** core radius R_c. Independent estimates are listed in Table 2.

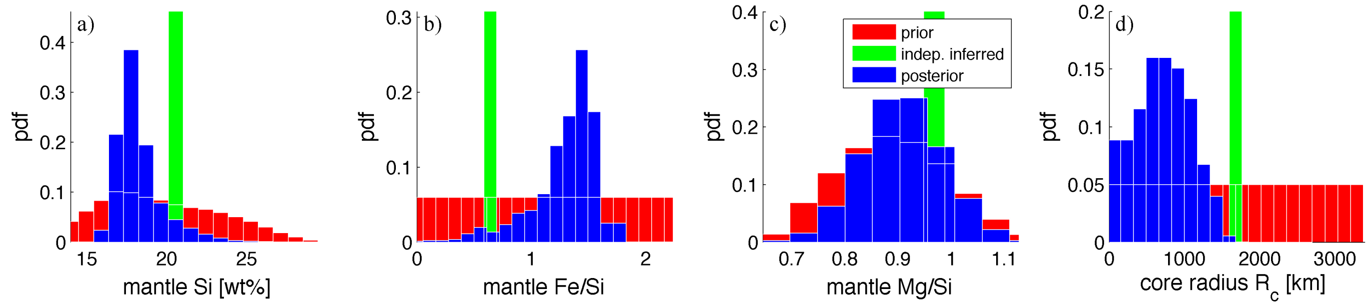

Fig. 7. Sampled 1D marginal posterior distributions for Mars, case B: prior (red), posterior (blue) and independent estimates (green) of model parameters: **a)** mantle Si-content $\mathrm{Si}_{\circledcirc}$; **b)** mantle $\mathrm{Fe/Si}_{\circledcirc}$; **c)** mantle $\mathrm{Mg/Si}_{\circledcirc}$; and **d)** core radius R_c. Independent estimates are listed in Table 2.

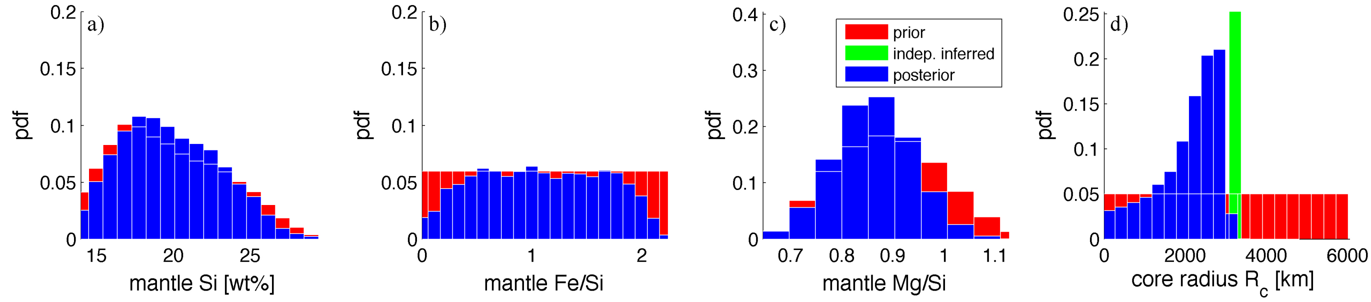

Fig. 8. Sampled 1D marginal posterior distributions for Venus, case B: prior (red), posterior (blue) and independent estimates (green) of model parameters: **a)** mantle Si-content $\mathrm{Si}_{\circledcirc}$; **b)** mantle $\mathrm{Fe/Si}_{\circledcirc}$; **c)** mantle $\mathrm{Mg/Si}_{\circledcirc}$; and **d)** core radius R_c. Independent estimates are listed in Table 2.

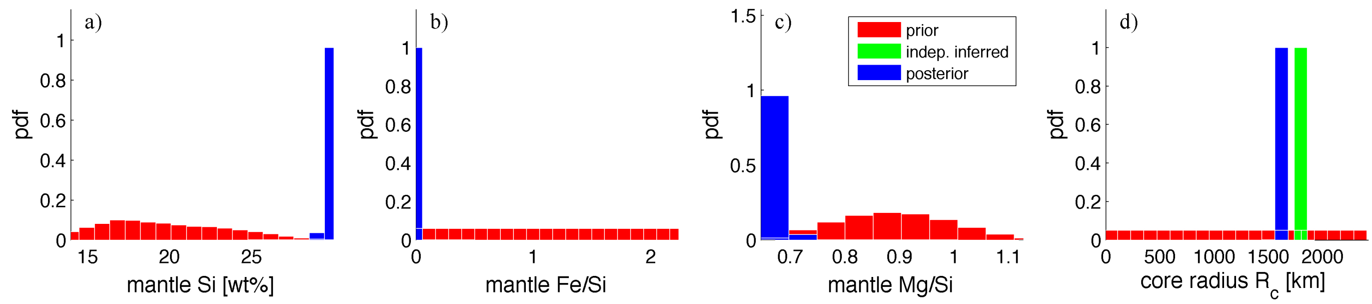

Fig. 9. Sampled 1D marginal posterior distributions for Mercury, case B: prior (red), posterior (blue) and independent estimates (green) of model parameters: **a)** mantle Si-content $\mathrm{Si}_{\circledcirc}$; **b)** mantle $\mathrm{Fe/Si}_{\circledcirc}$; **c)** mantle $\mathrm{Mg/Si}_{\circledcirc}$; and **d)** core radius R_c. No interior model fits the data, i.e., likelihoods are nearly zero. Independent estimates are listed in Table 2.

$M_c/M \geq 0.6$, i.e., large iron cores, are evidently excluded as it becomes difficult to increase mantle Si-content in order to satisfy the compositional constraint ($\mathrm{Fe/Si}_{\bigcirc}$).

In order to quantify the information content of data, i.e., the degree to which data constrain internal structure, we compute the Shannon entropy measure ($H_{\mathrm{posterior}}/H_{\mathrm{prior}}$) (Tarantola 2005), which allows the comparison of prior and posterior PDFs. If $H_{\mathrm{posterior}}/H_{\mathrm{prior}} \rightarrow 1$, data are only weakly constraining models, whereas if $H_{\mathrm{posterior}}/H_{\mathrm{prior}} \rightarrow 0$, data pose strong constraints on the models (see Appendix D for details). Figure 12 shows $H_{\mathrm{posterior}}/H_{\mathrm{prior}}$ for all inverted model parameters. For case A terrestrial-type planets, we find that core radius R_c is generally well-constrained, whereas mantle composition is weakly constrained. For $\mathrm{Si}_{\circledcirc}$, $\mathrm{Fe/Si}_{\circledcirc}$, and $\mathrm{Mg/Si}_{\circledcirc}$ the relative Shannon

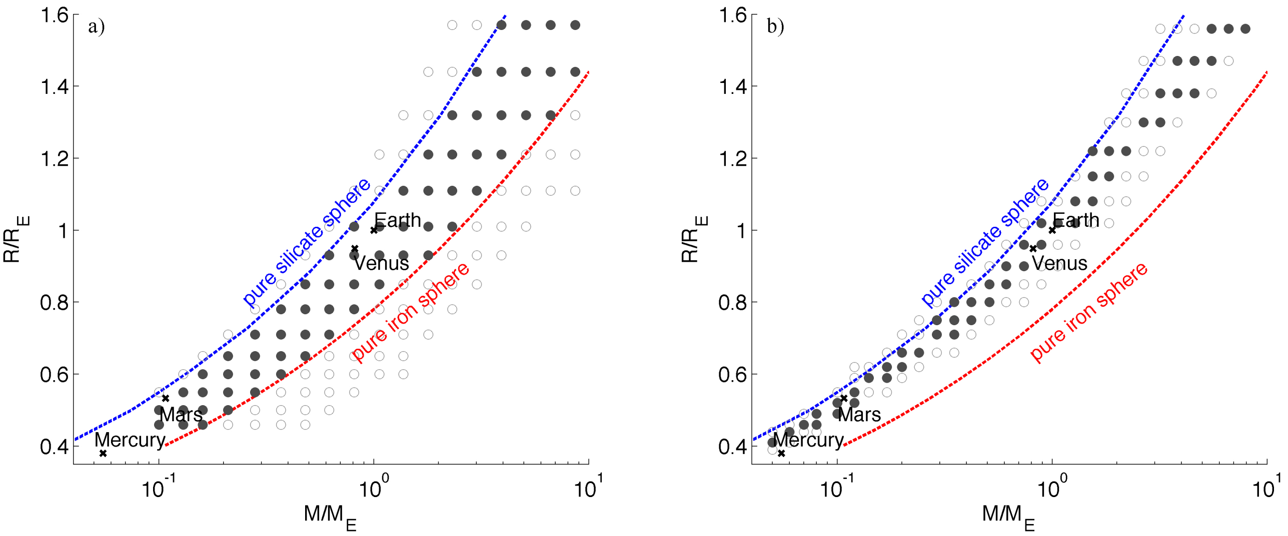

Fig. 10. Range of synthetic rocky planets that fit data (black dots) of **a)** M and R (case A), and **b)** M, R, and stellar abundance constraints (case B). Mass-radius curves for pure iron (red) and pure silicate spheres (blue, zero amount of FeO) were calculated independently. In case A, all synthetic planets that fit data (black dots) fall within the limits of a pure silicate (blue line) and pure iron sphere (red line). In case B, this region is reduced to a smaller area indicating that chemical bulk ratio ($Fe/Si_{\bigcirc}$) introduces strong additional constraints on the internal structure. Synthetic planets for which no internal structure could be derived within model assumptions (pure iron core surrounded by a rocky mantle) are indicated by gray open circles. For every circle in the plots an inverse problem has been solved. See main text for further details.

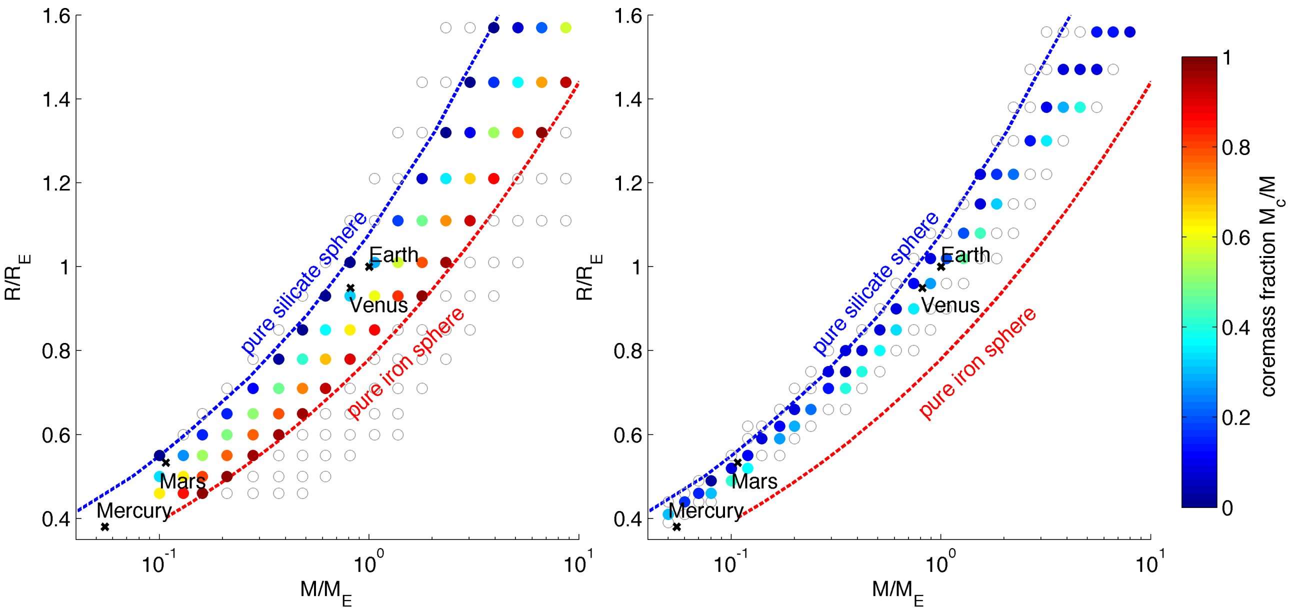

Fig. 11. Mean posterior core mass fraction M_c/M plotted for the range of synthetic rocky planets that fit **a)** M and R (case A), and **b)** M, R, and stellar abundance constraints (case B). See also Fig. 10.

entropy measures are larger than 0.8 throughout the tested range of synthetic planets. There are few planets that have bulk densities that are very close to the pure silicate sphere for which $Fe/Si_{\circledcirc}$ appears to be better constrained. Generally for case A, the inversion scheme is not sensitive to variations in mantle composition but mostly sensitive to variations in core radius R_c. Stated differently, any variations in mantle composition can be compensated by tiny variations in core radius, while variations in core radius must be compensated by large changes in mantle composition in order that data (M and R) be matched.

From Fig. 12 we observe the following trend for case A: R_c appears to be better constrained for smaller and denser planets. The denser a rocky planet is, the more it is dominated by its core, and so less mantle material is available whose compositional changes could compensate for variations in core radius. Similarly, when the planet is smaller, there is less absolute mantle mass available to compensate for a higher variability in core radius. Therefore, the precision with which core radius can be determined rests ultimately on the available mantle mass.

For case B, the addition of $Fe/Si_{\bigcirc}$ significantly reduces model variability, i.e., correlation between model parameters is increased. For this case R_c, $Si_{\circledcirc}$, and $Fe/Si_{\circledcirc}$ are better constrained compared with case A, while $Mg/Si_{\circledcirc}$ is only weakly constrained as for case A. We note, however, that the prior bounds $Mg/Si_{\circledcirc}$ are significantly narrower for case B than for case A because of the stellar abundance constraint. The relative

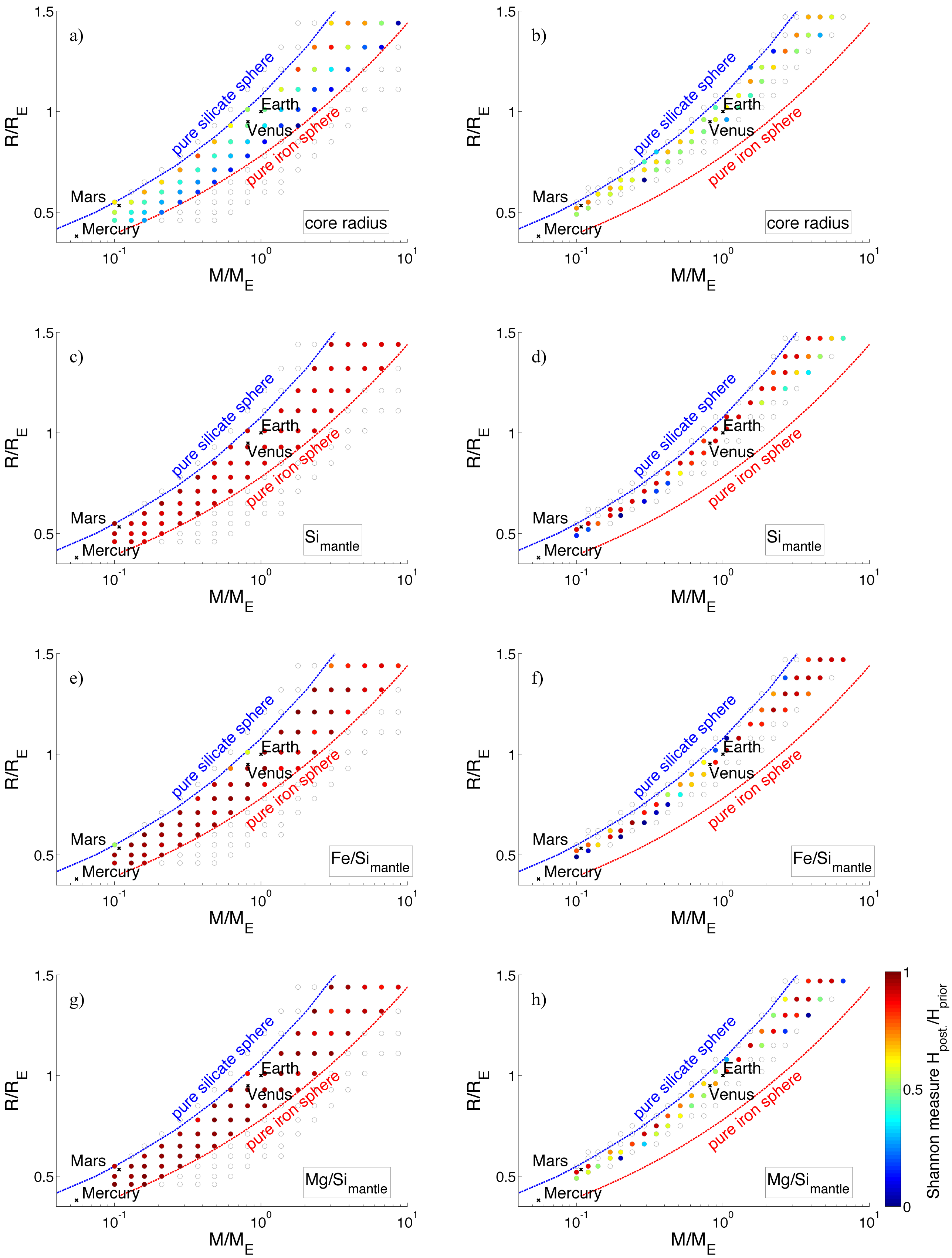

Fig. 12. Relative Shannon entropy measure $H_{\text{posterior}}/H_{\text{prior}}$ of model parameters plotted for the range of synthetic rocky planets that fit a given M and R (case A, *left panels*), and M, R, and $\mathrm{Fe/Si}_{\odot}$ (case B, *right panels*). Model parameters are well constrained for $H_{\text{posterior}}/H_{\text{prior}} \rightarrow 0$ (blue colors) and not constrained for $H_{\text{posterior}}/H_{\text{prior}} \rightarrow 1$ (red colors). For case B, variations of $H_{\text{posterior}}/H_{\text{prior}}$ along the grid appear not to be smooth which is due to the poor grid resolution. See also Fig. 10.

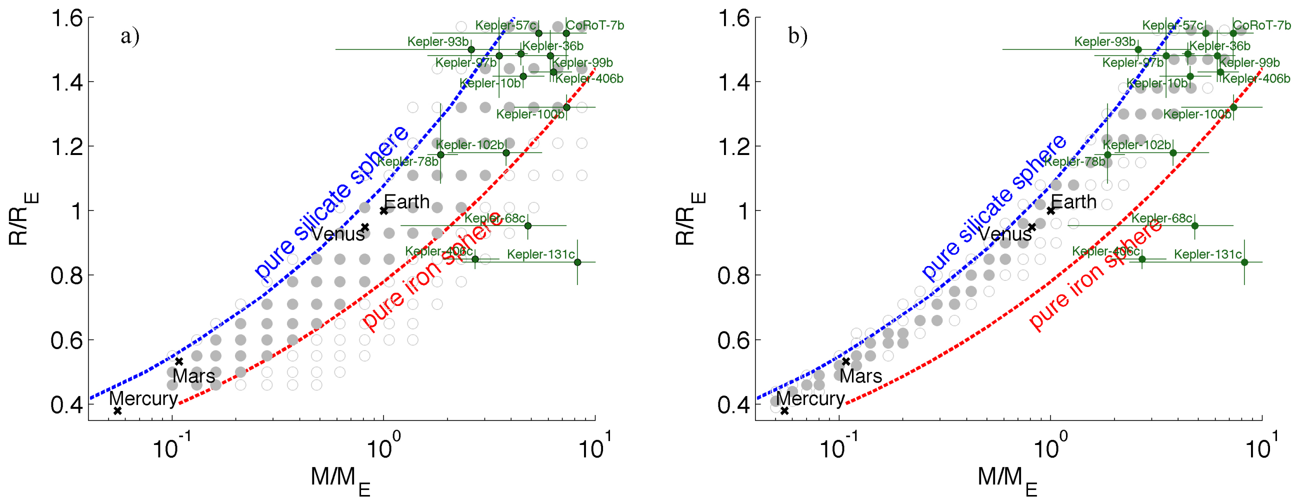

Fig. 13. Confirmed exoplanets plotted against Fig. 10: **a)** case A and **b)** case B.

Table 3. Masses and radii of observed exoplanets.

Planet	Radius R/R_E	Mass M/M_E	Reference	Mean bulk density g/cm^3
CoRoT-7b	1.55 ± 0.10	7.31 ± 1.21	Moutou et al. (2013)	10.82
Kepler-10b	1.416 ± 0.033	$4.56^{+1.17}_{-1.29}$	Batalha et al. (2013)	8.89
Kepler-36b	1.486 ± 0.035	$4.45^{+0.33}_{-0.27}$	Carter et al. (2012)	7.48
Kepler-57c	1.55 ± 0.04	5.4 ± 3.7	Steffen et al. (2013)	8.00
Kepler-68c	$0.953^{+0.037}_{-0.042}$	$4.8^{+2.5}_{-3.6}$	Gilliland et al. (2013)	30.58
Kepler-78b	$1.173^{+0.159}_{-0.089}$	$1.86^{+0.38}_{-0.25}$	Pepe et al. (2013)	6.35
Kepler-93b	1.50 ± 0.03	2.59 ± 2.00	Weiss & Marcy (2014)	4.23
Kepler-97b	1.48 ± 0.13	3.51 ± 1.90	Weiss & Marcy (2014)	5.97
Kepler-99b	1.48 ± 0.08	6.15 ± 1.30	Weiss & Marcy (2014)	10.46
Kepler-100b	1.32 ± 0.04	7.34 ± 3.20	Weiss & Marcy (2014)	17.60
Kepler-102b	1.18 ± 0.04	3.8 ± 1.8	Weiss & Marcy (2014)	12.75
Kepler-131c	0.84 ± 0.07	8.25 ± 5.90	Weiss & Marcy (2014)	76.74
Kepler-406b	1.43 ± 0.03	6.35 ± 1.40	Weiss & Marcy (2014)	11.97
Kepler-406c	0.85 ± 0.03	2.71 ± 0.80	Weiss & Marcy (2014)	24.33

Shannon entropy measures for core radius R_c in case B are similar but slightly smaller than for case A.

As discussed for the terrestrial solar system planets, only the core size can be constrained in the case of exoplanets for which only mass and radius are known (case A). In case B, the compositional constraint ($Fe/Si_{\odot}$) introduces a strong correlation between the model parameters $Si_{\circledcirc}$, $Fe/Si_{\circledcirc}$ and R_c. This enables us to constrain mantle composition.

4.2. Application to confirmed exoplanets

In this section we apply our methodology to actual exoplanet measurements in order to derive models of the internal structure. At this point in the development of our approach we are still limited by the assumptions made for the overall structure (iron core and rocky mantle with no ice and no atmosphere) and hence this application serves more as an illustration of the overall capabilities of our method. In Fig. 13 we plot all confirmed exoplanets in the mass-radius range of interest on top of our synthetic exoplanet calculations (Fig. 10). Masses, radii, and uncertainty ranges of the observed exoplanets are listed in Table 3. Kepler-68c, Kepler-131c and Kepler-406c have estimated mean densities above that of pure iron and clearly present special exotic cases. Kepler-36b has been chosen to demonstrate the application of our method to actual observations because of small uncertainties on mass and radius. To our knowledge, no stellar elemental abundances of $Fe/Si_{\star}$ and $Mg/Si_{\star}$ of its host star other than for the Sun are available. Hence, we test case B assuming solar elemental abundances. Since the *Kepler* mission has been targeting Sun-like stars (Borucki et al. 2010), this is a permissible assumption.

Figure 14 shows the results for Kepler-36b. For case A, we see that the mantle ratios $Fe/Si_{\circledcirc}$ and $Mg/Si_{\circledcirc}$ are not constrained at all, while $Si_{\circledcirc}$ and R_c are moderately and well constrained, respectively. By adding $Fe/Si_{\odot}$ to the data (case B), $Si_{\circledcirc}$ and $Fe/Si_{\circledcirc}$ are significantly better constrained. Thus, incorporating observations of stellar elemental abundances has a big impact on our ability to constrain model parameters. The posterior range of R_c is very large in both cases ranging from zero to about half the planet radius. We argue that mass and radius uncertainties of 7% and 2%, respectively, are the reason for such a large spread.

The example of Kepler-36b illustrates the necessity of having well-measured masses and radii in order to derive constraints on the interior structure. Furthermore, it exemplifies the large impact of elemental abundance ratios in order to constrain planetary interior structure. Since it will be impossible to directly

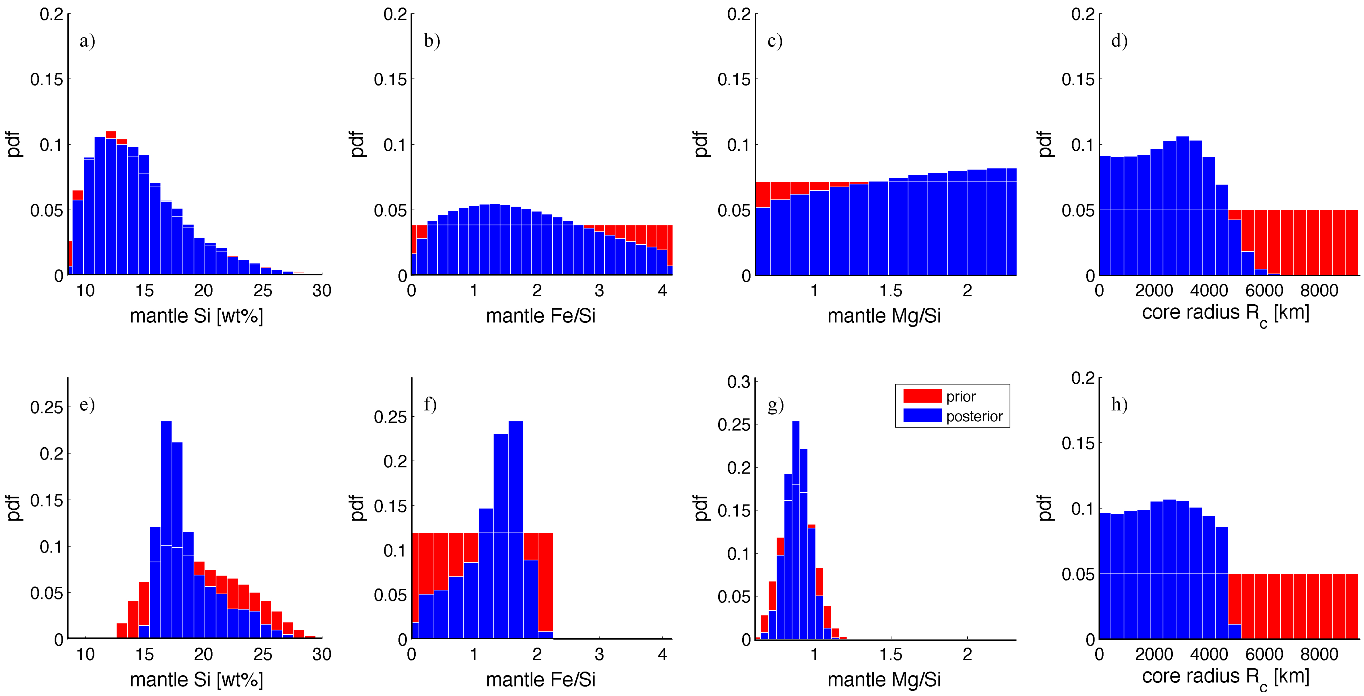

Fig. 14. Sampled 1D marginal posterior distributions for Kepler-36b, case A **a)–d)** and B **e)–h)**: prior (red) and posterior (blue) of model parameters: **a)**, **e)** mantle Si-content $Si_{\circledcirc}$, **b)**, **f)** mantle $Fe/Si_{\circledcirc}$, **c)**, **g)** mantle $Mg/Si_{\circledcirc}$, and **d)**, **h)** core radius R_c. Data uncertainties are $\sigma_{Fe/Si_{\bigcirc}} = 10\%$, $\sigma_R = 2\%$, and $\sigma_M = 7\%$.

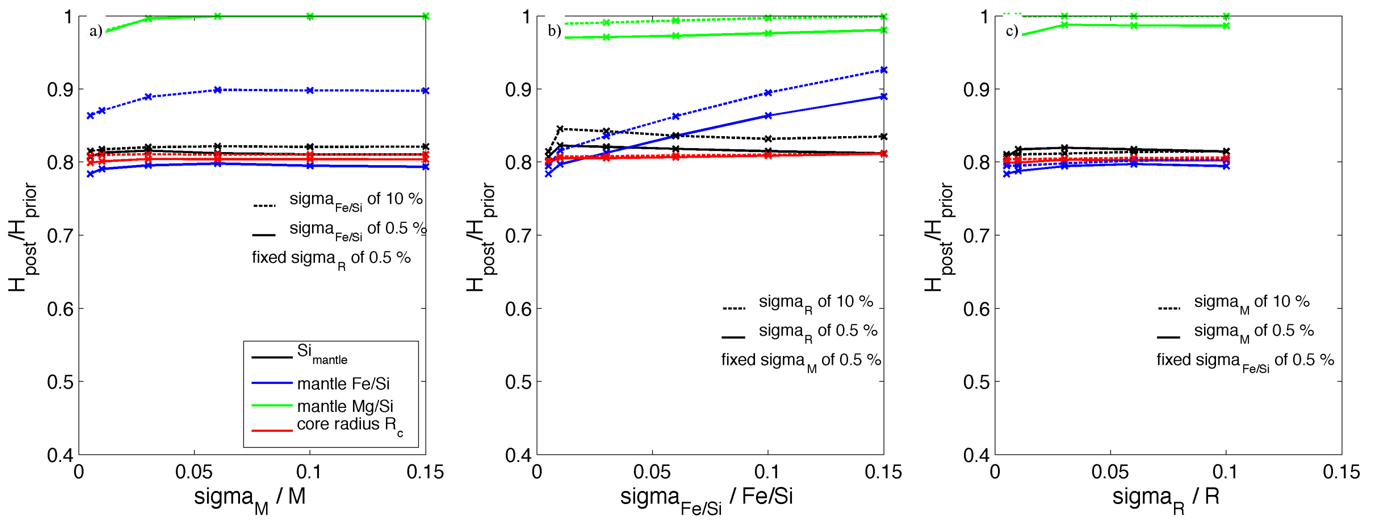

Fig. 15. Summary of information content for Kepler 36-b ($M = 4.45 M_E$, $R = 1.486 R_E$): relative Shannon measure as a function of **a)** σ_M while fixing σ_R at 0.5%, **b)** $\sigma_{Fe/Si_{\bigcirc}}$ while fixing σ_M at 0.5%, **c)** σ_R while fixing $\sigma_{Fe/Si_{\bigcirc}}$ at 0.5%.

measure these, spectroscopically-determined values measured from the photosphere of the host star should serve as proxies.

4.2.1. Information content of the data

The limiting factor on our ability to constrain interior structure is measurement precision. In this section, we study in more detail the link between measurement precision and resulting constraints on internal structure.

For this, we compute the relative Shannon entropy measure $H_{posterior}/H_{prior}$ for each model parameter for a range of measurement uncertainty. In these calculations, inversions were performed using a grid-search method. The results are shown in Fig. 15 for the example of Kepler-36b and in Fig. 16 for an Earth-like planet. The results can be summarized as follows:

- Overall, a decrease in data uncertainty leads to a smaller $H_{posterior}/H_{prior}$ and thus better constrained parameters. However, for large uncertainties $H_{posterior}/H_{prior}$ is nearly constant and relatively large.
- Uncertainties on M, R, and $Fe/Si_{\bigcirc}$ appear equally important.
- The inherent degeneracy of the problem is limiting our ability to constrain interior structure even in the case of very small data uncertainties. For the cases shown, $H_{posterior}/H_{prior}$ does not fall below 0.5.
- Our ability to constrain interior structure is significantly different for larger and denser exoplanets such as Kepler-36b compared to Earth-like bodies. As shown for Kepler-36b,

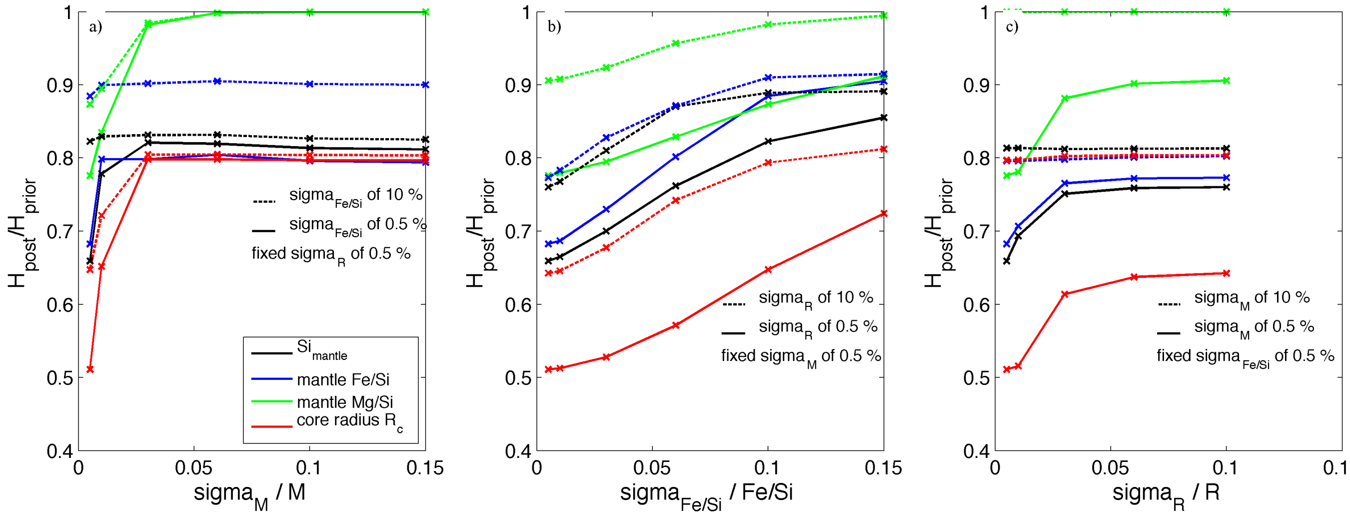

Fig. 16. Summary of information content for an Earth-sized body ($M = 1M_{\rm E}$, $R = 1R_{\rm E}$): relative Shannon measure as a function of **a)** σ_M while fixing σ_R at 0.5%, **b)** $\sigma_{\rm Fe/Si_{\bigcirc}}$ while fixing σ_M at 0.5%, **c)** σ_R while fixing $\sigma_{\rm Fe/Si_{\bigcirc}}$ at 0.5%.

$H_{\rm posterior}/H_{\rm prior}$ decreases only at extremely small data uncertainties whereas for an Earth-like body, $H_{\rm posterior}/H_{\rm prior}$ decreases significantly earlier. Given their large difference in mass, this indicates that measurements with a given precision are able to better constrain the internal structure of small-mass planets than large-mass planets. Large planets have more absolute mantle mass available whose compositional changes are able to compensate for a higher variability in core radius, ultimately reducing our ability to constrain interior models.

– Over the entire tested range of uncertainties on $\rm Fe/Si_{\bigcirc}$ (0−15%), significant improvements in our ability to constrain interior structure can be observed, whereas for mass and radius this is only evident in the case of small uncertainty ranges (<5%).

5. Discussion

The results presented here have to be interpreted in the light of the assumptions made. Firstly, we consider terrestrial-like planets that consist only of silicate mantles and pure iron cores. Sotin et al. (2007) showed that rocky planets can be described with great accuracy by only using the four elements Si, Mg, Fe, O. Secondly, we assume a pure iron composition for the core. Volatiles (H, He, H_2O, etc.) are currently neglected. Compared to Earth, where the core include lighter elements (e.g., S, Si, C, O), our method therefore systematically overestimates core density and underestimates core radius. This introduces biases in the estimates of mantle composition due to the strong correlations between mantle composition and core radius. In the future, lighter elements (e.g., Si) will be included in the core. Thirdly, our approach assumes subsolidus conditions and perfectly known Equation-of-State parameters for all considered compositions. For exoplanets that are larger than Earth, extrapolations of parameters to pressures and temperatures beyond those measured in the laboratory are required. This introduces uncertainty in the EoS modeling that we will address in the future.

While the mean density of a terrestrial-like planets can be derived from measurements of mass and radius, its internal structure is only relatively poorly constrained by these two quantities as also shown by Rogers & Seager (e.g., 2010). We find that adding $\rm Fe/Si_{\bigcirc}$ and $\rm Mg/Si_{\bigcirc}$ as compositional constraints significantly improves this situation. These additional constraints significantly narrow the prior range of mantle compositions (for details on use of different priors see, e.g., Rogers & Seager 2010). In particular, $\rm Fe/Si_{\bigcirc}$ introduces a strong correlation between model parameters $\rm Si_{\circledcirc}$, $\rm Fe/Si_{\circledcirc}$, and $R_{\rm c}$, which significantly reduces the range of possible models. Hence, $\rm Fe/Si_{\bigcirc}$ plays a central role in constraining mantle composition and internal structure.

Our ability to constrain interior structure is directly limited by data uncertainties. We have demonstrated that uncertainties on mass, radius, and elemental abundances appear to be equally important. While it is worth striving to increase the measurement precision on all input data, the knowledge of $\rm Fe/Si_{\bigstar}$ and $\rm Mg/Si_{\bigstar}$ is particularly important as it represents a relatively powerful means of reducing the degeneracy of possible interior models. The level of inherent degeneracy in model solutions actually depends on the specific values of M and R (as discussed for Fig. 12) (Rogers & Seager 2010): for given measurement accuracy of mass and radius, interior structure is better constrained for an Earth-sized planet than for a larger body (e.g., Kepler-36b).

In an earlier study, Rogers & Seager (2010) employed Bayesian analysis to study exoplanet GJ 581d in an attempt to constrain core mass and silicate mantle mass fractions of $MgSiO_3$ and $FeSiO_3$. Based on mass and radius alone, Rogers & Seager (2010) also concluded that an exact interior solution cannot be inferred since the problem is highly underdetermined. However, Rogers & Seager (2010) argue against the use of bulk abundance constraints inferred from $\rm Fe/Si_{\bigstar}$ and $\rm Mg/Si_{\bigstar}$ in order to remain as independent as possible of planet formation models. It should be noted that if Rogers & Seager (2010) were to use bulk abundance constraints core size and mantle composition would be fixed since they use a parameterized phase-diagram approach and restrict the mantle mineralogy to $\rm Mg_{1-\chi}Fe_{\chi}SiO_3$. Consequently, the problem would no longer be degenerate. Clearly, our approach is more general because it allows for arbitrary elemental abundance ratios while using stellar abundance constraints as a means of further reducing model variability.

With regard to planet formation, these studies suggest that some interior compositions are more likely than others. Recent planet formation studies accounting for equilibrium condensation of material (e.g., Thiabaud et al. 2013) have shown that most planets have a bulk refractory (Fe, Si, and Mg) composition that appears to be indistinguishable from that of the host star because species of Fe, Si and Mg condense at similar temperatures ($\sim$1000 K). Hence, most planets will have similar refractory elemental ratios as those in the original disk which is assumed to be well-represented by stellar ratios. Furthermore, O, Fe, Si, and Mg are predicted to be among the most frequent solid species in circumstellar disks (e.g., Elser et al. 2012; Gilli et al. 2006) as a consequence of which models with large iron cores and small silicate mantles are generally unlikely to form. Based on these observations, Valencia et al. (2007a) also proposed the general use of a minimum bulk value for Fe/Si. Here, we have followed Grasset et al. (2009), who considered bulk Mg/Si and Fe/Si to be dictated by their abundance in the host star's photosphere and adopt these as proxies for planet bulk composition.

For the terrestrial solar system planets, we showed that using solar photospheric abundances ($\mathrm{Fe/Si}_\bigstar$ and $\mathrm{Mg/Si}_\bigstar$) yields tighter constraints on the internal structure, except for Mercury. Mercury's interior is unique among the terrestrial planets and is usually explained by the removal of a large fraction of its silicate mantle by either a giant impact or by evaporation due to the proximity of the planet to the Sun (Benz et al. 1988). Consequently, our scheme is not able to find interior models that fit data for Mercury. In spite of this, we argue for the use of stellar abundance constraints because of the ability of our scheme to predict for which planets alternative accretion scenarios or post-formation processes must be envisaged. We also note that meteorites, which are assumed to be the building blocks of the terrestrial planets, have abundances of refractory elements within 10% of solar values (Lodders 2003).

6. Conclusions and outlook

We have presented a Bayesian inversion method in order to constrain the interior structure of rocky exoplanets from measurements of mass and radius. Since stellar refractory elemental abundances are likely to be generally a good proxy for planetary bulk compositions, they represent powerful additional constraints allowing us to significantly improve determination of internal structure. Our proposed scheme is more general than previous works on mass-radius relationship of rocky exoplanets, in that we obtain confidence regions of interior structures of general mantel composition and core structure. Furthermore, we investigated how interior structure models depend on various parameters in the form of prior information, data, and data uncertainties.

We have applied the method to the terrestrial planets of our solar system and compared the results to independent estimates and, except for Mercury, find generally good agreement. Following this, we have applied our method to synthetic exoplanets in a large mass and radius range as well as to exoplanet Kepler-36b. We have investigated two different cases: case A where we invert mass and radius, and case B where we additionally consider stellar elemental abundances obtained from photospheric observations of the host star. We summarize our findings in the following:

- Photospheric elemental abundances of the host star ($\mathrm{Fe/Si}_\bigstar$ and $\mathrm{Mg/Si}_\bigstar$) are main parameters in reducing model degeneracy through the introduction of correlations between mantle composition and core size.
- How well mantle composition and core radius can be generally constrained depends on data and data uncertainties. There is an inherent degeneracy that limits our ability to constrain the interior structure even in the case of small data uncertainties. Independently of these, it seems that model variability depends on mass and density of a planet. Thus, it is our contention that a case-by-case probabilistic inversion that provides model parameter uncertainties is indispensable in order to rigorously characterize interior structure.
- Measurement uncertainties on mass, radius, and stellar abundance constraints appear to be equally important.
- For improved characterization of interior structure better estimates of $\mathrm{Fe/Si}_\bigstar$ and $\mathrm{Mg/Si}_\bigstar$ are required. We have shown for Kepler-36b and an Earth-sized planet that significant improvements on model parameter estimation can be achieved by reducing the uncertainty of $\mathrm{Fe/Si}_\bigstar$ and $\mathrm{Mg/Si}_\bigstar$.

Space missions that aim at characterizing exoplanets by means of the transit method seek the most precise measurement of R and hope for follow-up missions to provide precise measurements for M. However, since data precision also depends on characteristics of target star and observation time (integration time, number of transits observed), a careful weighing of costs and benefits is crucial for the success of a mission. We have demonstrated how the relative Shannon entropy can be used in future missions (e.g., CHEOPS, TESS, PLATO) as a means of optimizing the scientific return. In particular, we are able to quantitatively assess the improvement in interior structure-determination of a specific exoplanet target as a result of an increase in the measurement precision of mass and radius. To illustrate this we consider the case (discussed in Fig. 15), where the predominantly flat curves imply that smaller uncertainties on mass and radius do not provide significantly better insights on the interior structure. For an Earth-sized exoplanet, however, smaller uncertainties on mass and radius can lead to an improved understanding of interior structure (Fig. 16).

An increasing number of Super-Earths are being observed. Super-Earths lie in the intermediate mass-range between terrestrial planets and the gas/ice giants in the solar system with different scenarios for their interiors. Their interior structures (e.g., purely rocky composition or predominantly water/carbon compounds) and formation histories are, however, a matter of debate, as is the question of whether Super-Earths can harbor life (e.g., Valencia et al. 2007c; Howe et al. 2014). A potential way of answering these questions is to invert for the physico-chemical structure in order to narrow down the possible types of interior models.

In future studies, we will extend the dimensionality of our problem to include more elements, e.g., hydrogen and water, so as to model also exoplanets with atmospheres and/or oceans. This will eventually allow us to study how interior types are distributed among stars and what can be learned from such distributions about planet formation. However, for a given data set, increased model complexity is generally accompanied with a larger degeneracy in model solutions. Additional data would help to restrict model variability. For example, dynamic observations can constrain a planet's tidal dissipation and thereby provide constraints on the distribution of rigidity and density (through the Love number, e.g., Delisle et al. 2014).

Acknowledgements. We would like to thank a very enthusiastic referee for comments that led to a much improved manuscript. This work was supported by the Swiss National Foundation under grant 15-144 and the ERC grant 239605. It was in part carried out within the frame of the National Centre for Competence in Research PlanetS. We would like to thank Lauren Cooper for assistance in the early stages of this project.

References

Aitta, A. 2012, Icarus, 218, 967
Batalha, N. M., Rowe, J. F., Bryson, S. T., et al. 2013, ApJSS, 204, 24
Beirao, P., Santos, N. C., Israelian, G., & Mayor, M. 2005, A&A, 438, 251
Belonoshko, A. B. 2010, Condens. Matter. Phys., 13, 23605
Benz, W., Slattery, W. L., & Cameron, A. G. W. 1988, Icarus, 74, 516
Bond, J. C., O'Brien, D. P., & Lauretta, D. S. 2010, ApJ, 715, 1050
Bordè, P., Rouan, D., & Léger, A. 2003, A&A, 405, 1137
Borucki, W. J., Koch, D., Basri, G., et al. 2010, Science, 327, 977
Cameron, A. G. W., Fegley Jr, B., Benz, W., & Slattery, W. L. 1988, in The strange density of Mercury: Theoretical considerations, eds. F. Vilas, C. Chapman, & M. Matthews (Tucson: University of Arizona Press), Mercury, 692
Carter, J. A., Agol, E., Chaplin, W. J., et al. 2012, Science, 337, 556
Connolly, J. A. D. 2005, Earth Planet. Sci. Lett., 236, 524
Connolly, J. A. D. 2009, Geochem. Geophys. Geosyst., 10, Q10014
Delisle, J. B., Laskar, J., & Correia, A. C. M. 2014, A&A, 566, A137
Drake, M. J., & Righter, K. 2002, Nature, 416, 39
Dziewonski, A. M., & Anderson, D. L. 1981, Phys. Earth Plan. Interiors, 25, 297
Elser, S., Meyer, M. R., & Moore, B. 2012, Icarus, 221, 859
Fortney, J. J., Marley, M. S., & Barnes, J. W. 2007, ApJ, 659, 1661
Gilks, W. R., Richardson, S., & Spiegelhalter, D. 1996, in Markov chain Monte Carlo in practice (London: Chapman&Hall/CRC), 1
Gilli, G., Israelian, G., Ecuvillon, A., Santos, N. C., & Mayor, M. 2006, A&A, 449, 723
Gilliland, R. L., Marcy, G. W., Rowe, J. F., et al. 2013, ApJ, 766, 40
Grasset, O., Schneider, J., & Sotin, C. 2009, ApJ, 693, 722
Hastings, W. K. 1970, Biometrika, 57, 97
Hauck, S. A., Margot, J. L., Solomon, S. C., et al. 2013, J. Geophys. Res. Planets, 118, 1204
Howe, A. R., Burrows, A. S., & Verne, W. 2014, ApJ, 787, 173
Irifune, T. 1994, Nature, 370, 131
Jackson, I., & Rigden, S. M. 1996, Phys. Earth Planet. Interiors, 96, 85
Javoy, M., Kaminski, E., Guyot, F., et al. 2010, Earth Planet. Sci. Lett., 293, 259
Jephcoat, A., & Olson, P. 1987, Nature, 325, 332
Johnson, T. V., Mousis, O., Lunine, J. I., & Madhusudhan, N. 2012, ApJ, 757, 192
Khan, A., & Connolly, J. A. D. 2008, J. Geophys. Res. Planets, 113, E07003
Khan, A., Connolly, J. A. D., Maclennan, J., & Mosegaard, K. 2007, Geophys. J. Int., 168, 243
Kuskov, O. L., & Kronrod, V. A. 2001, Icarus, 151, 204
Lodders, K. 2003, ApJ, 591, 1220
Longhi, J., Knittle, E., & Holloway, J. R. 1992, in Mars, Vol. 20 (Tucson: AZ Univ. Arizona Press), 184
Madhusudhan, N., Lee, K. K., & Mousis, O. 2012, ApJ, 759, L40
McDonough, W. F., & Sun, S. S. 1995, Chemical Geology, 120, 223
Mocquet, A., Grasset, O., & Sotin, C. 2014, Phil. Trans. Roy. Soc. A, 372, 20130164
Morgan, J. W., & Anders, E. 1980, Proc. Natl. Acad. Sci., 77, 6973
Mosegaard, K., & Tarantola, A. 1995, J. Geophys. Res., 100, 12431
Mosegaard, K., & Tarantola, A. 2002, Int. Geophys. Ser., 81, 237
Moutou, C., Deleuil, M., Guillot, T., et al. 2013, Icarus, 226, 1625
Padovan, S., Margot, J. L., Hauck, S. A., Moore, W. B., & Solomon, S. C. 2014, J. Geophys. Res. Planets, 119, 850
Pepe, F., Cameron, A. C., Latham, D. W., et al. 2013, Nature, 503, 377
Rivoldini, A., Van Hoolst, T., Verhoeven, O., Mocquet, A., & Dehant, V. 2011, Icarus, 213, 451
Rogers, L. A., & Seager, S. 2010, ApJ, 712, 974
Rogers, L. A., Bodenheimer, P., Lissauer, J. J., & Seager, S. 2011, ApJ, 738, 59
Seager, S., Kuchner, M., Hier-Majumder, C. A., & Militzer, B. 2007, ApJ, 669, 1279
Sotin, C., Grasset, O., & Mocquet, A. 2007, Icarus, 191, 337
Steffen, J. H., Fabrycky, D. C., Agol, E., et al. 2013, MNRAS, 428, 1077
Stixrude, L., & Lithgow-Bertelloni, C. 2005, Geophys. J. Int., 162, 610
Stixrude, L., & Lithgow-Bertelloni, C. 2011, Geophys. J. Int., 184, 1180
Strom, R. G., & Sprague, A. L. 2003, Exploring Mercury: the iron planet (Springer Praxis Books/Space Exploration)
Tarantola, A. 2005, Inverse problem theory and methods for model parameter estimation (Philadelphia: SIAM)
Tarantola, A., & Valette, B. 1982, J. Geophys., 50, 159
Thiabaud, A., Marboeuf, U., Alibert, Y., et al. 2013, A&A, 562, A27
Valencia, D., O'Connell, R. J., & Sasselov, D. D. 2006, Icarus, 181, 545
Valencia, D., Sasselov, D. D., & O'Connell, R. J. 2007a, ApJ, 665, 1413
Valencia, D., Sasselov, D. D., & O'Connell, R. J. 2007b, ApJ, 656, 545
Valencia, D., O'Connell, R. J., & Sasselov, D. D. 2007c, ApJ, 670, L45
Valencia, D., Ikoma, M., Guillot, T., & Nettelmann, N. 2010, A&A, 516, A20
Wagner, F. W., Sohl, F., Hussmann, H., Grott, M., & Rauer, H. 2011, Icarus, 214, 366
Weiss, L. M., & Marcy, G. W. 2014, ApJ, 783, L6
Wood, B. J., & Holloway, J. R. 1984, Geochim. Cosmochim. Acta, 48, 159
Zapolsky, H. S., & Salpeter, E. E. 1969, ApJ, 158, 809

Pages 17 to 18 are available in the electronic edition of the journal at http://www.aanda.org

Appendix A: Prior of mantle Si-content

The prior bounds on $Si_{\circledcirc}$ depend on the prior bounds of $Fe/Si_{\circledcirc}$ and $Mg/Si_{\circledcirc}$, since the mass fractions of all oxides in the NCFMAS chemical system must sum to one. The mass fraction of Si can be expressed as

$$Si_{\circledcirc} = \frac{(1-\chi_{Na_2O,CaO,Al_2O_3})}{\left(\frac{\mu_{SiO_2}}{\mu_{Si}} + Fe/Si_{\circledcirc}\frac{\mu_{FeO}}{\mu_{Fe}} + Mg/Si_{\circledcirc}\frac{\mu_{MgO}}{\mu_{Mg}}\right)},$$

where μ are respective molar weights and the mass fraction of minor oxides are combined in χ_{Na_2O,CaO,Al_2O_3}. Transforming from a uniform prior in $Fe/Si_{\circledcirc}$ and $Mg/Si_{\circledcirc}$ to the variable $Si_{\circledcirc}$ we find the prior of $Si_{\circledcirc}$ being

$$p(Si_{\circledcirc}) \propto \left|\frac{\partial Fe/Si_{\circledcirc}}{\partial Si_{\circledcirc}}\right| = \frac{(1-\chi_{Na_2O,CaO,Al_2O_3})}{Si_{\circledcirc}^2\frac{\mu_{FeO}}{\mu_{Fe}}}.$$

This implies a non-uniform prior for $Si_{\circledcirc}$. Outside the given interval in Table 1 the prior distribution of $Si_{\circledcirc}$ is essentially zero.

Appendix B: Temperature profile

We assume a fixed mantle temperature profile that is based on the Earth model. A variable temperature profile for the mantle only introduces negligible variations in the density profile (<1%). This is demonstrated in Fig. B.1. Surface temperature is fixed at 1000 K and although temperature varies between 2100 K to 3500 K at the core mantle boundary, resultant changes in density are relatively small (Fig. B.1).

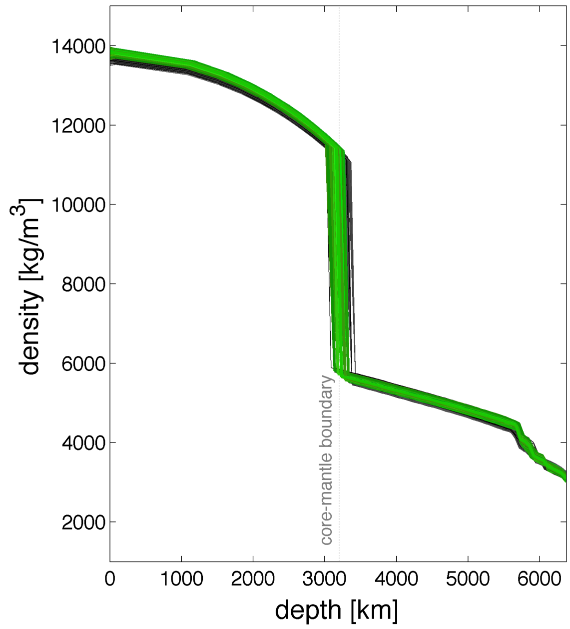

Fig. B.1. Sampled profiles of density profiles for Earth. A variable mantle temperature (gray profiles) only introduces little additional variation in density compared to profiles with fixed mantle temperature (green colors).

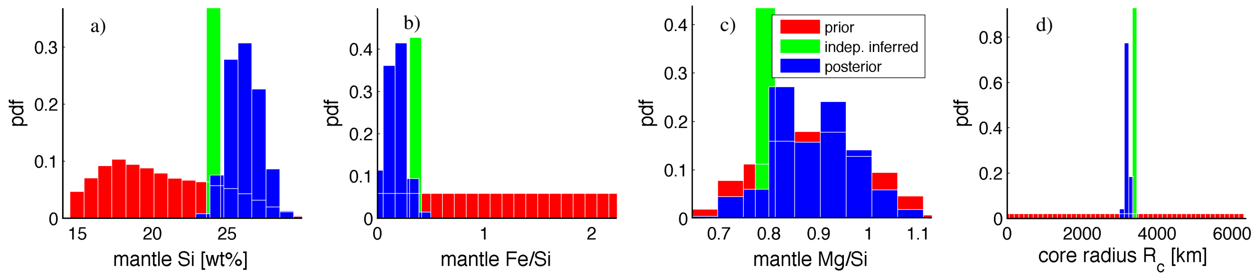

Fig. C.1. Sampled 1D marginal posterior distributions for Earth using the FMS system: prior (red), posterior (blue) and independent estimates (green) of model parameters: **a)** mantle Si-content $\mathrm{Si}_{\circledcirc}$; **b)** mantle $\mathrm{Fe/Si}_{\circledcirc}$; **c)** mantle $\mathrm{Mg/Si}_{\circledcirc}$; and **d)** core radius R_c. Independent estimates are listed in Table 2.

Appendix C: Silicate model chemical system

To illustrate the influence of the chosen model chemical system, we solve the inverse problem for the case of the Earth as in Sect. 3.2, but using the simpler FMS system. Sampled model parameter PDFs are shown in Fig. C.1 and relative to NCFMAS (Fig. 4), we observe that with the FMS system mantle composition and core radius are less well predicted, although posterior model parameter variability is, as expected, smaller (fewer chemical components). In addition, future use of stellar determinations of Ca, Na, and Al abundances support the use of the NCFMAS system.

Appendix D: Shannon entropy measure

The Shannon entropy measure of a discrete variable X of possible states $\{x_1, x_2, .., x_N\}$, with N being the number of bins, can be written as

$$H = -\sum_{i=1}^{N} p(x_i) * \log_2 p(x_i),$$

where the x_i states represent the distribution histogram bins in which the probability of occurrence is given by

$$p(x_i) = \frac{\text{number of realizations in the bin}}{\text{total number of realizations}}.$$

The Shannon entropy H has a maximum $H_{\max} = \log_2(N)$ when all N states are equally likely (e.g., a uniform distribution) and a minimum $H_{\min} = 0$ when all realizations x_i fall into a single histogram bin. Clearly, H is bin-size dependent. Here, the bin size used is defined such that the prior range of each model parameter is divided in ≈20 different bins.