w.rahman17@imperial.ac.uk
SCAM - 03/10/2019

A Bayesian Approach to Selection Effects Modelling

Wahidur Rahman

Co-Authors: Roberto Trotta, David van Dyk, Vincent Chen, Evan Tey and Timothy Lucas Makinen

Imperial College London

Overview

- Intro and Analysis of SNIa within the Bayesian Hierarchical Framework BAHAMAS
- Current problems with modelling Selection Effects of SNIa
- Modelling of Selection Effects with Bayesian Methods

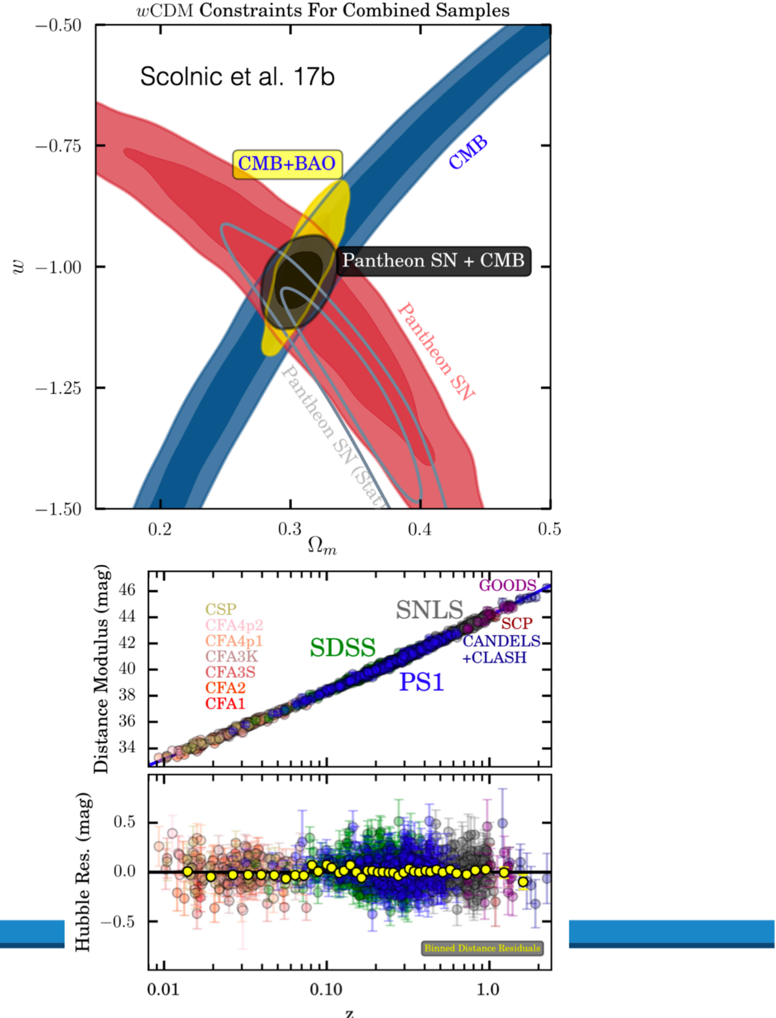
Imperial College London

Supernova Cosmology

In the era of precision cosmology we are able to place tight constraints on our cosmology which SNe contribute to.

Currently limited by systematics of SNe data.

Scolnic et. al 2018



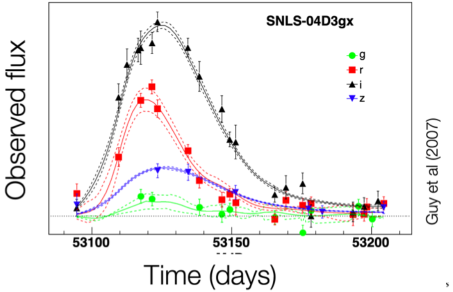
Imperial College London

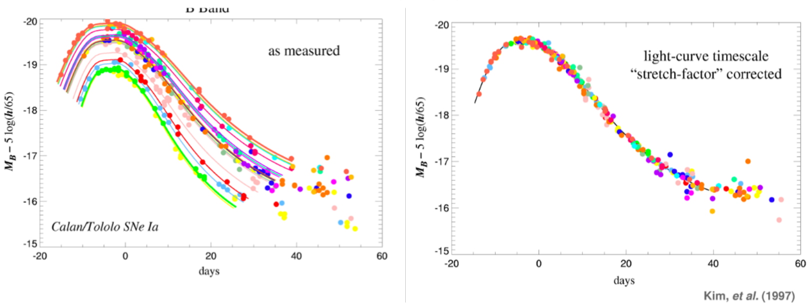
Standard Procedure

Spectroscopic Confirmation

Standardization Procedure

Chi-squared fit and minimization





$$X^2 = \sum_i \frac{(\mu(z_i, C) - [m_{B,obs,i} - M_0 + \alpha x_{1,obs,i} + \beta c_{obs,i})])\hat{}2}{\sigma_{int}^2 + \sigma_{\mu_i}^{z\;2} + \sigma_{fit}^2}$$

Imperial College London

Alternative to Chi-Squared Fit - BAHAMAS

BAyesian Hier**A**rchical **M**odelling for the **A**nalysis of **S**upernova Cosmology

- Fits into the last step and based on Bayesian Statistics and designed to allow easy to access features such as model comparison and hierarchical modelling.
- Other Bayesian models have appeared for this step which include *UNITY* by Rubin et al (2015) and Simple-*BayeSN* by Mandel et al (2017) and *Steve* by Hinton et al (2019).

Likelihood

Prior

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)}$$

Evidence

Imperial College London

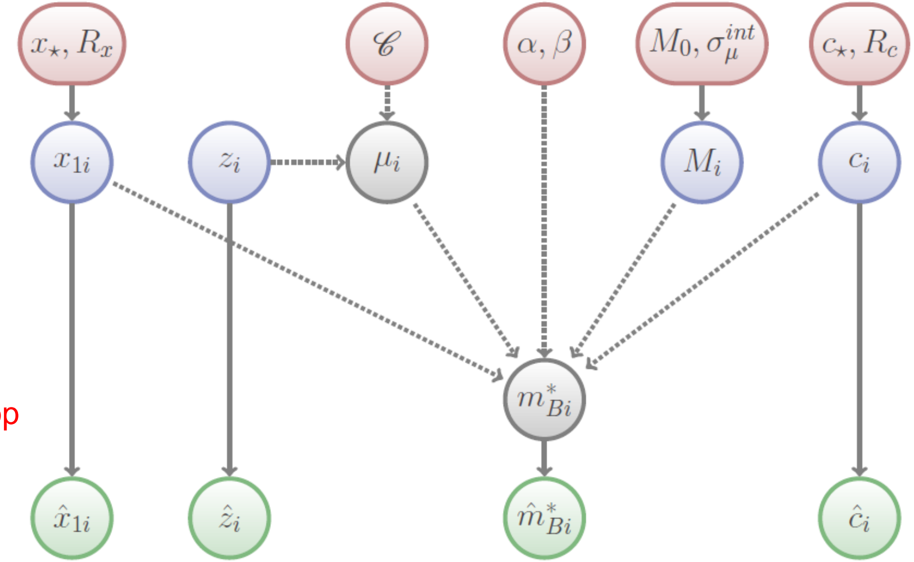
BAHAMAS

Using a Hierarchical Modelling, we can relate the observed parameters to the parameters of interest and the true latent values which can be integrated out later in the likelihood.

p(data|params) ∝ ∫ p(data, true, pop | params) dtrue dpop

= ∫ p(data|true) p(true|pop) p(pop) dtrue dpop

March et al (2011)



Imperial College London

Application of BAHAMAS Real Data - JLA

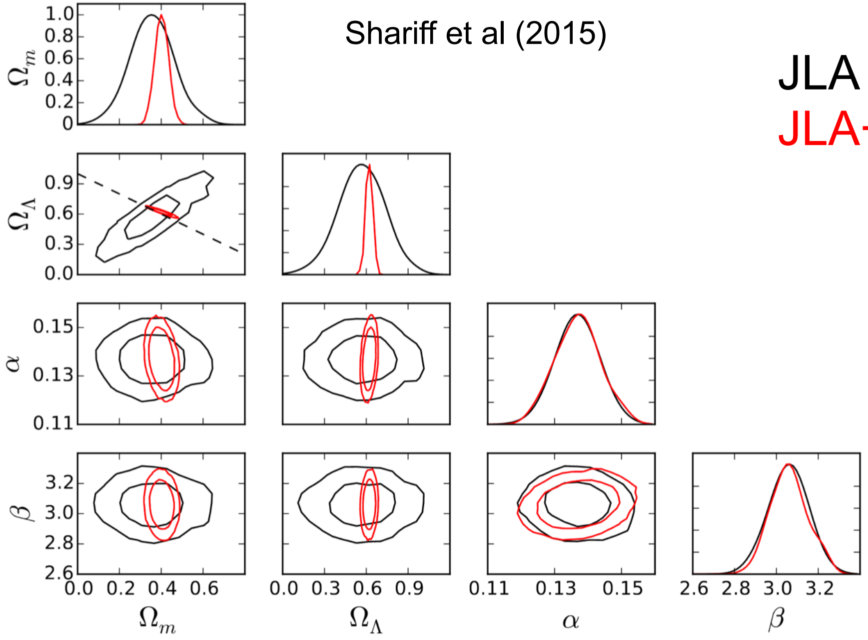
Set w = - 1

JLA Alone

$$\Omega_M = 0.340 \pm 0.101$$
$$\Omega_K = 0.119 \pm 0.249$$

+Planck

$$\Omega_M = 0.399 \pm 0.027$$
$$\Omega_K = -0.024 \pm 0.008$$



Shariff et al (2015)

JLA (Betoule et al, 2014)

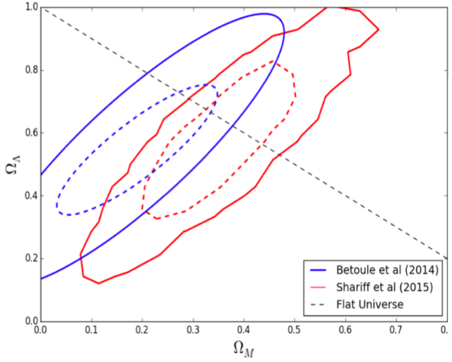
JLA+PLANCK 2015

Shariff et al (2015)

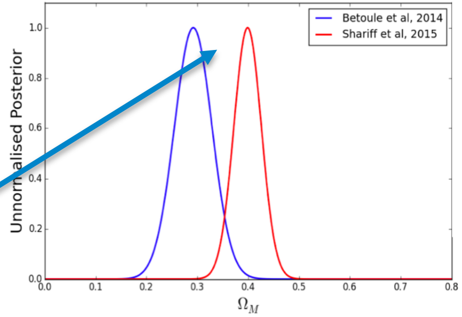
Issues

Sensitive to incorrect model specification

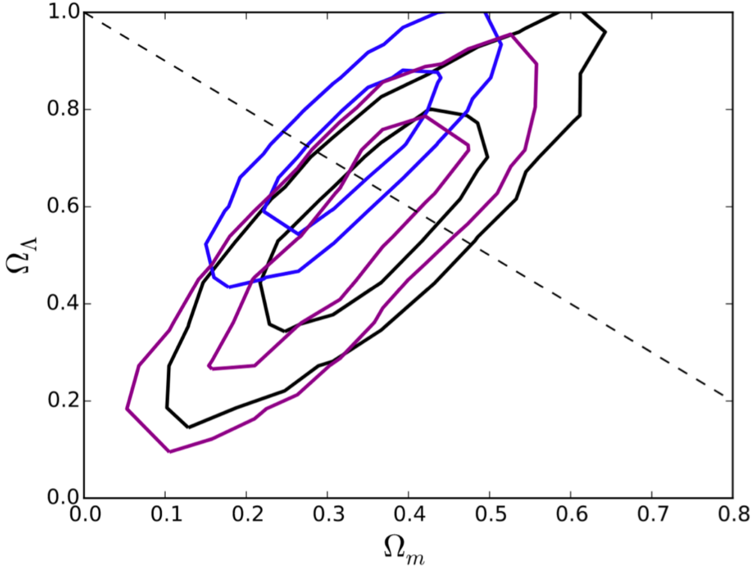
SNIa alone



SNIa + Planck



2.8 sigma tension



Blue:Statistical

Purple:Stat + m_B syst

Black:Combined

Imperial College London

Selection Effects – The Problem

Problem: There is a Malmquist bias present in SNIa, which is both magnitude and colour dependent, as intrinsically fainter and redder SNIa are harder to detect.

Current Method: solution is to simulate data and apply adhoc corrections to the original data such as the BBC method (Kessler and Scolnic, 2017).

Downsides: Dependent on the cosmological parameters used for simulation.

Uncertainty in correction is difficult to incorporate.

Imperial College London

Solution – A Bayesian Way

SNIa are indicated as selected or missed using an indicator function on some data Y= $\{Y_1, \ldots, Y_n\}$ of which N objects are observed with N<n

$$I = \begin{cases} 1 & \text{if SNe } Y_i \text{ is observed} \\ 0 & \text{If SNe } Y_i \text{ is unobserved} \end{cases}$$

$$y_{obs} = P(Y_i | I = 1, i \,\epsilon\, [1, \ldots, n]),$$
$$y_{miss} = P(Y_i | I = 0, i \,\epsilon\, [1, .., n])$$

$$p(y_{obs}, I | \theta, \phi) = \int p(Y, I | \theta, \phi) \mathrm{d}y_{miss}, \;\; \theta = \text{cosmo params}, \phi = \text{selection params}$$

We can the construct our posterior conditional on the observed data using Bayes Theorem and combine with our original likelihood. Many ways to model the Selection Function $P(Y_i, I = 1 | \theta, \phi)$. We use a Normal CDF which is linear in the cut parameters with smoothly varying selection efficiency. Cut parameters often inferred from simulations from tools such as SNANA.

Imperial College London

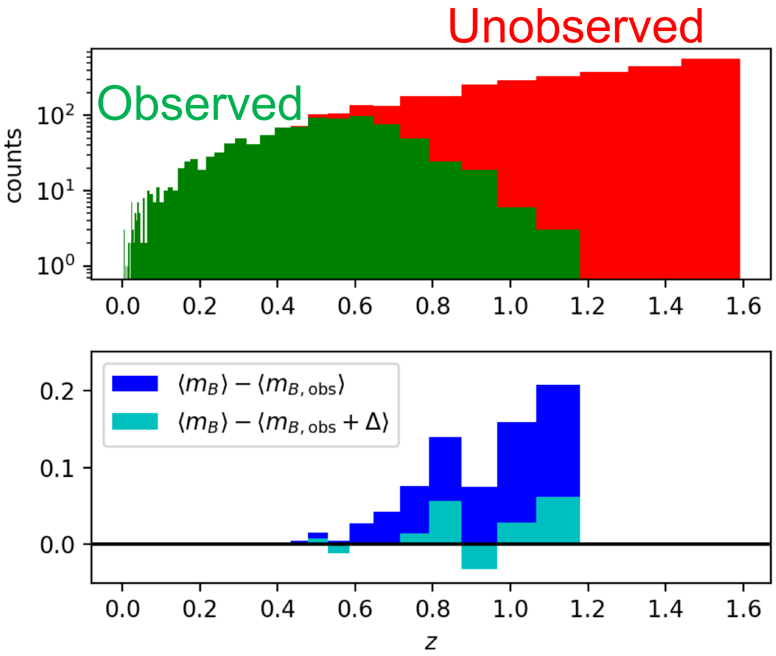
Solution – A Bayesian Way

Likelihood is similar to that used by Rubin et al (2015), but differs due to some fundamental assumptions about how we model the redshift distribution of unobserved SNe.

$$p(n_k) \propto 1/n_k$$

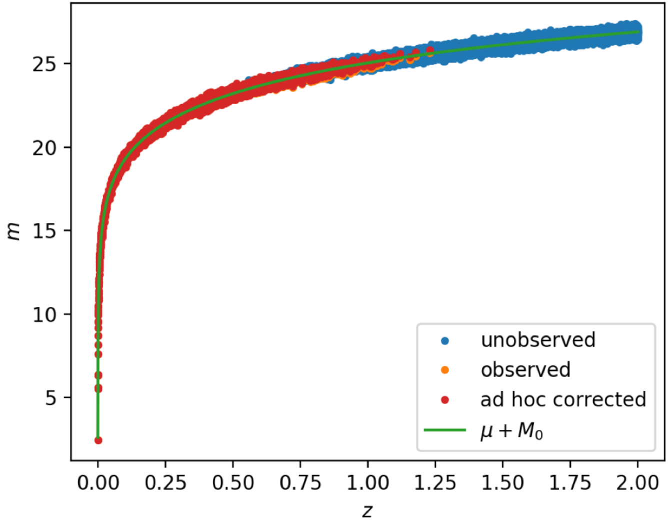
Prior is on the ***total*** unobserved supernova and marginalize out the redshift. R15 model implicitly assumes the same prior on observed and unobserved SNe. Leads to a different Likelihood, but it can be shown that it is related to our likelihood via:

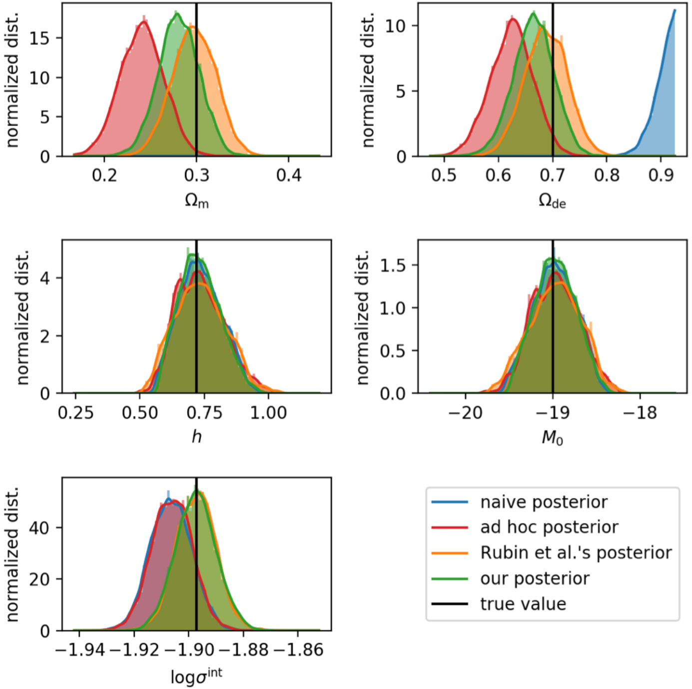
$$P(y_{obs}|\theta,\phi) = \int L_{r15}\, P(z|\theta,\phi)\mathrm{dz}$$



Imperial College London

Solution – when the selection function is known and correct

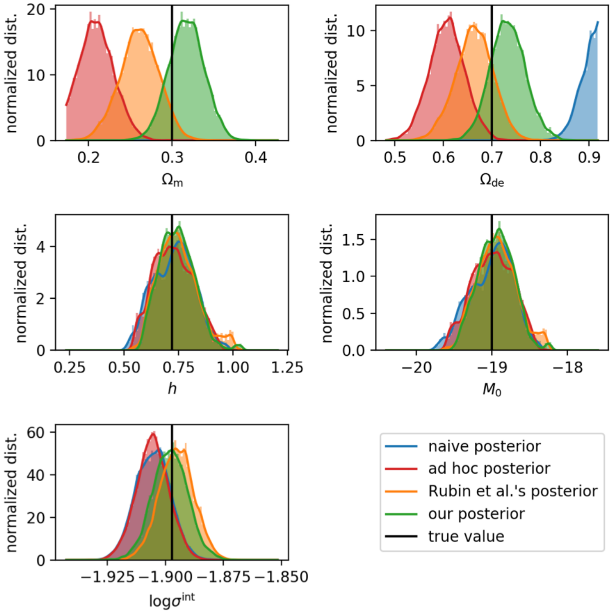




Imperial College London

Solution – when the selection function is known and incorrect

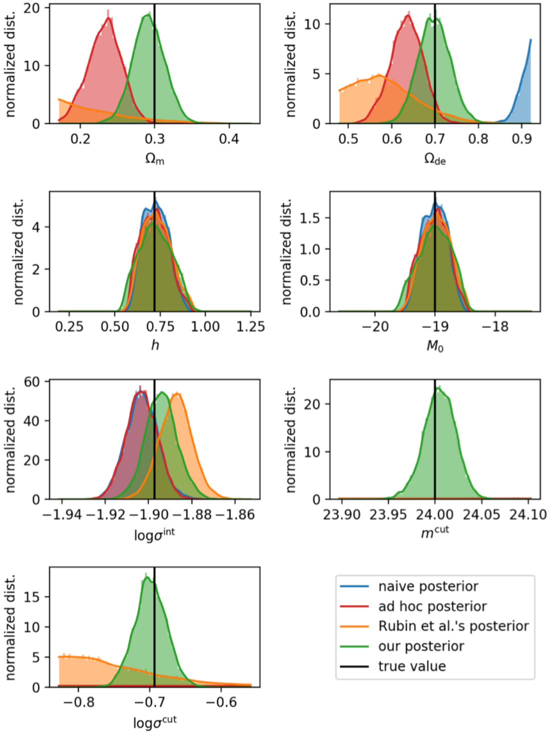
Cut on m_B biased low ~ 0.1Mag



Imperial College London

Solution – when the selection function is unknown

Perform inference over the selection parameters to solve this issue

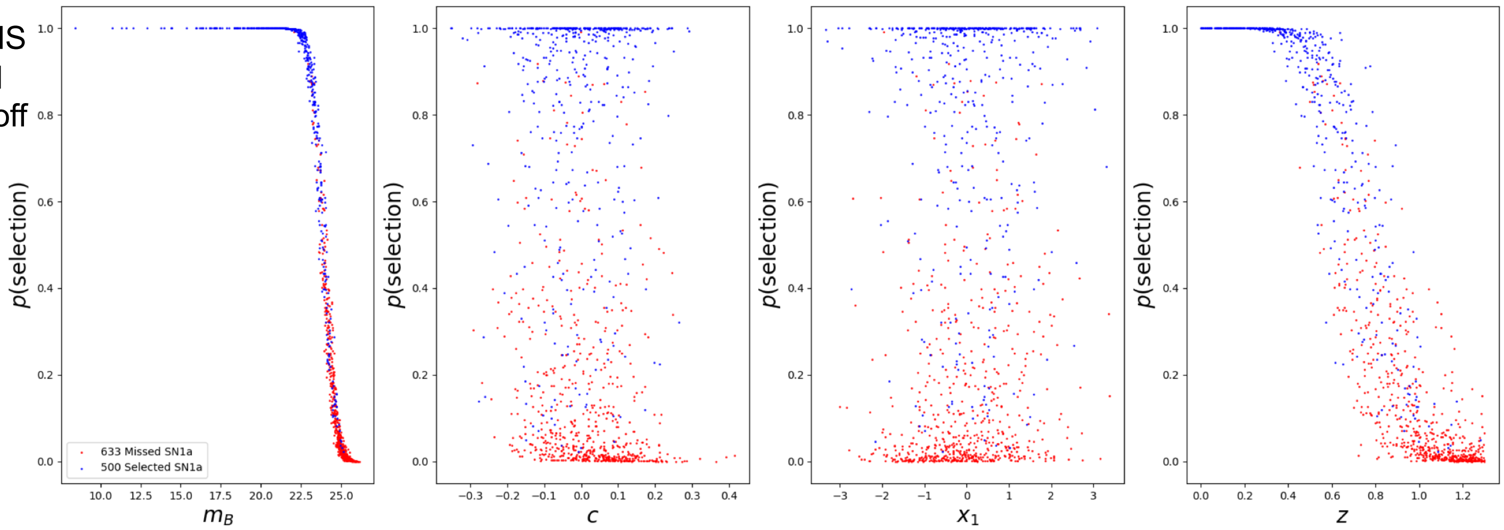


Imperial College London

Solution – A Realistic Case

Simulated SNLS-like selection efficiency based off SNANA-SIMS and applied to forward modelled data based off BAHAMAS

Simulated Selection Efficiency Based on SNLS-like SNANA simulations.

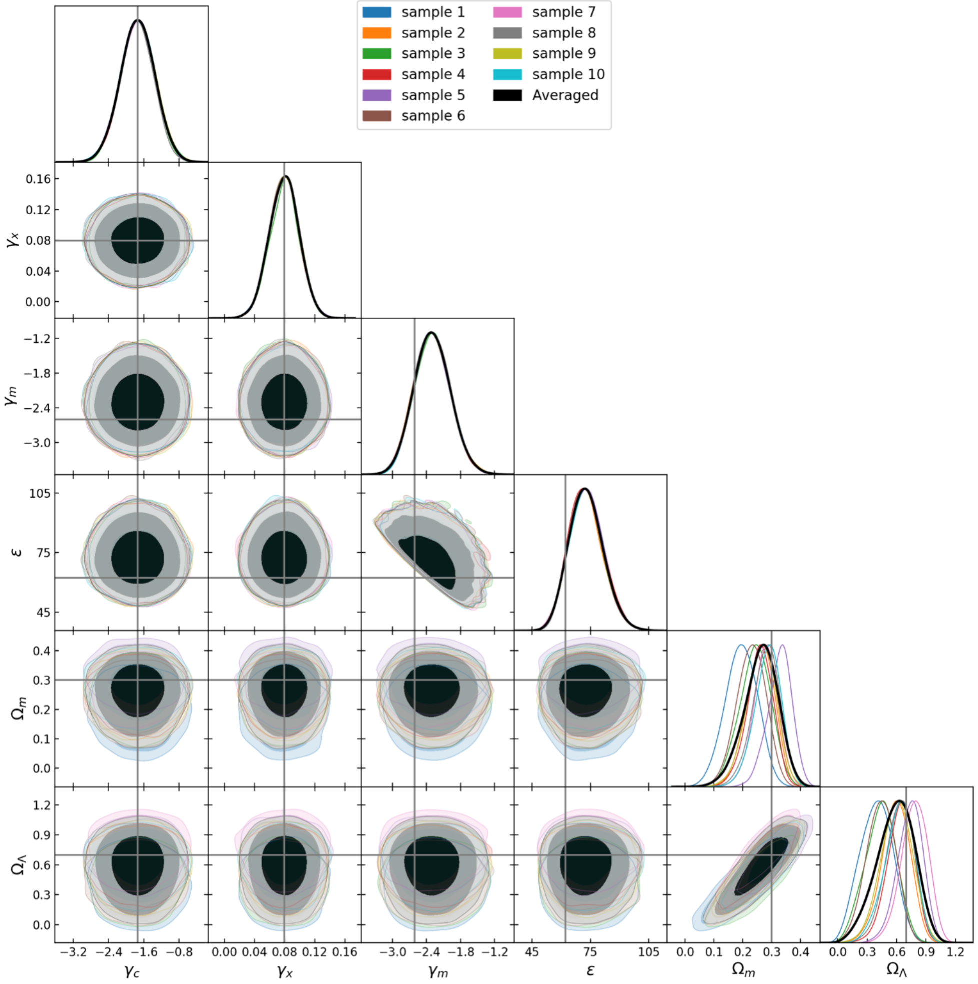


Ongoing Work

Imperial College London

Solution – A Realistic Case

Using SNLS-like cuts on forward-generated data from the BAHAMAS model, we can assess the method on more realistic SNIa data



Ongoing Work

Imperial College London

Conclusion:

Our method performs demonstrably well in the (realistic) scenario where the selection function is unknown.

BAHAMAS model still incomplete – work needed to incorporate other sources of systematic effects in model such host-metallicity and star formation rates.

Imperial College London

Thank you