Model-independent fitting of cosmological data

Alberto Guffanti
NBIA & Discovery Center, Niels Bohr Institute - University of Copenhagen

In collaboration with:
A. Durakovic, J. Trøst Nielsen and S. Sarkar

2nd NBIA-APCTP Workshop on Cosmology and Astroparticle Physics
Niels Bohr Institute, 18-22 August 2014

The core idea

- Can we devise **model independent** tests of **cosmic acceleration** (based on **SN Ia**)?

- A (partial) list of **recent proposals** for **model-independent studies** of cosmic acceleration, mainly based on SN Ia

  Shafieloo et al. (2006), Shafieloo (2007), Nesseris, Shafieloo (2010), …

  Schwartz and Seikel (2007, 2009)

  Benitez-Herrera et al. (2011, 2013)

  Garcia-Bellido and Nesseris (2011)

  ….

- A **crucial ingredient** of a good **model-independent method** to parametrise data is that it provides a **reliable estimation of the experimental uncertainties** so that we can tell if a given dataset can **effectively discriminate** between **competing models**

NNPDF methodology: the origins

- **Originally developed** to provide a **model independent parametrisation** of Deep Inelastic Scattering **structure functions** and **Parton Distribution Functions** (PDFs)

S. Forte et al., 2002;
L . Del Debbio et al., 2005;
A. Guffanti et al., 2006;
R. D. Ball et al., 2008;
R. D. Ball et al., 2013;
+ ……

- **Parton Distribution Functions** cannot be determined from first principles in (perturbative) QCD but need to be **extracted from (global) fits to data**

- The **NNPDF methodology** is designed to address two main **shortcomings** of **standard PDF determinations**: reliance on **linear error propagation** (Hessian method) and **parametrisation bias** (use of fixed functional form for parametrisation)

NNPDF methodology in a Nutshell

- **Generate** N_{rep} **Monte Carlo replicas** of the experimental data, taking into account all experimental correlations
- **Fit** a noncommittal functional form **to each replica** of the data providing a model-independent parametrisation of the data with a reliable uncertainty estimation
- **Expectation values** for a given observable are then given by

$$\langle O(x) \rangle = \frac{1}{N_{\rm rep}} \sum_{i=1}^{N_{\rm rep}} O_i(x)$$

.... and corresponding formulae for the estimators of Monte Carlo samples are used to compute **uncertainties**, **correlations**, etc.

NNPDF methodology: the ingredients

- **Monte Carlo** determination of **uncertainties**
  - **No** need to rely on **linear propagation** of errors
  - Possibility to test the impact of **non-gaussianly** distributed uncertainties
- Parametrisation using **Neural Networks**
  - Provide an **unbiased parametrisation**
- Determine the **best fit** functions using **Cross-Validation**
  - Ensures **proper fitting**, avoiding overlearning

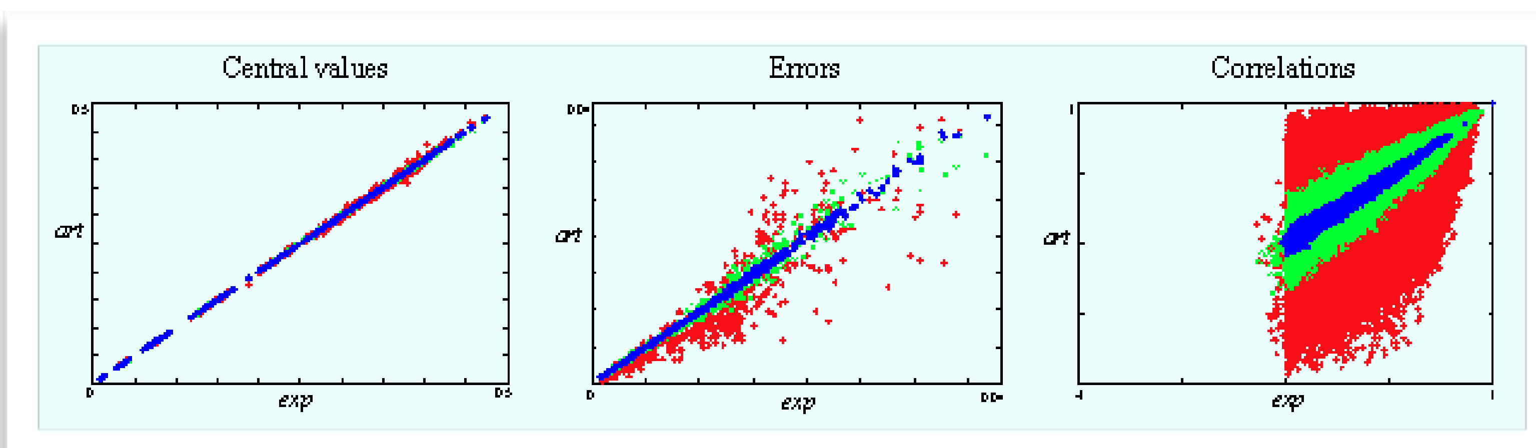
NNPDF methodology: data replicas generation

- **Monte Carlo replicas** are generated according to the distribution

$$O_i^{(art),(k)} = O_i^{(exp)} + \sum_{p=1}^{N_{sys}} r_p^{(k)} \sigma_p^i + r_{i,s}^{(k)} \sigma_s^i$$

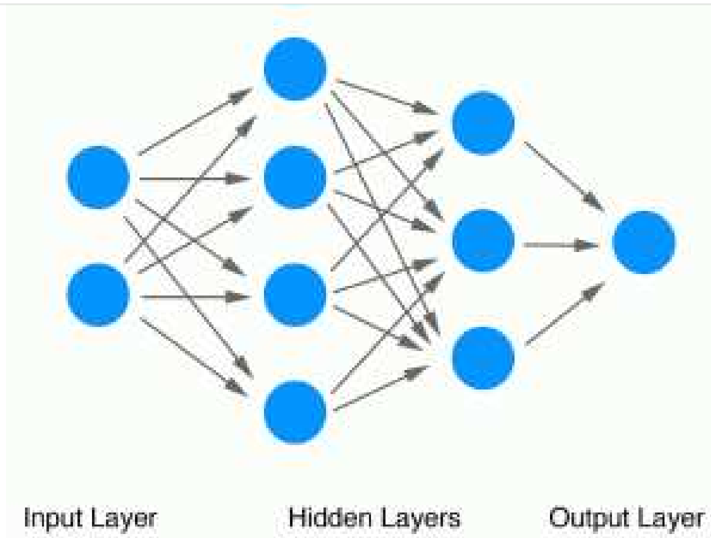
  where r_i are (gaussianly distributed) random numbers

- **Validate** Monte Carlo replicas against experimental data



- O(1000) replicas needed to **reproduce correlations** in experimental data to percent accuracy

NNPDF methodology: Neural Networks



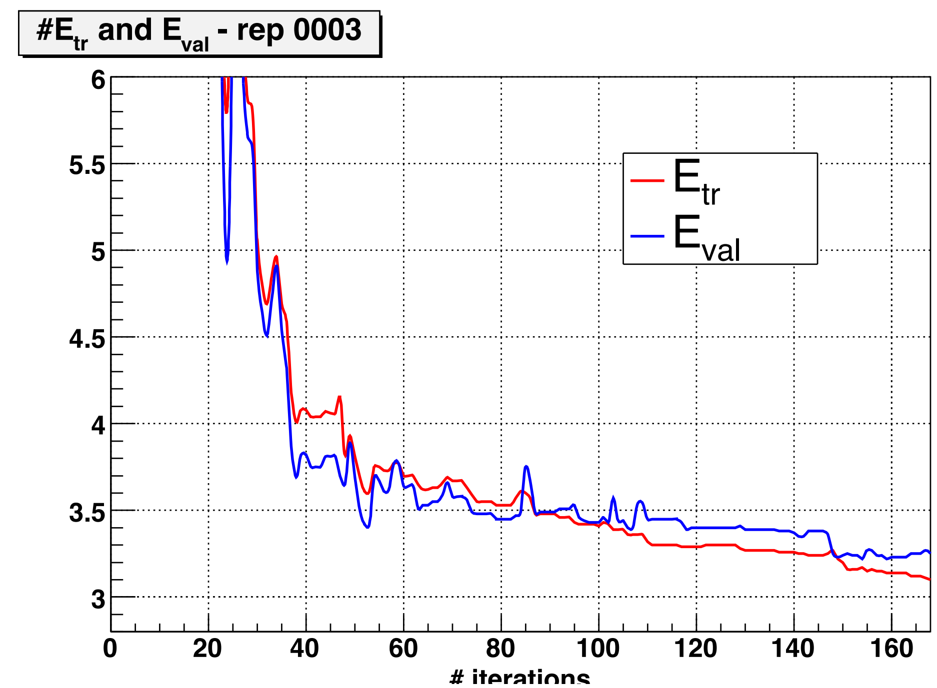
$$\xi_i^{(l)} = g\left(\sum_j^{n_l - 1} \omega_{ij}^{(l-1)} \xi_j^{(l-1)} - \theta_i^{(l)}\right)$$

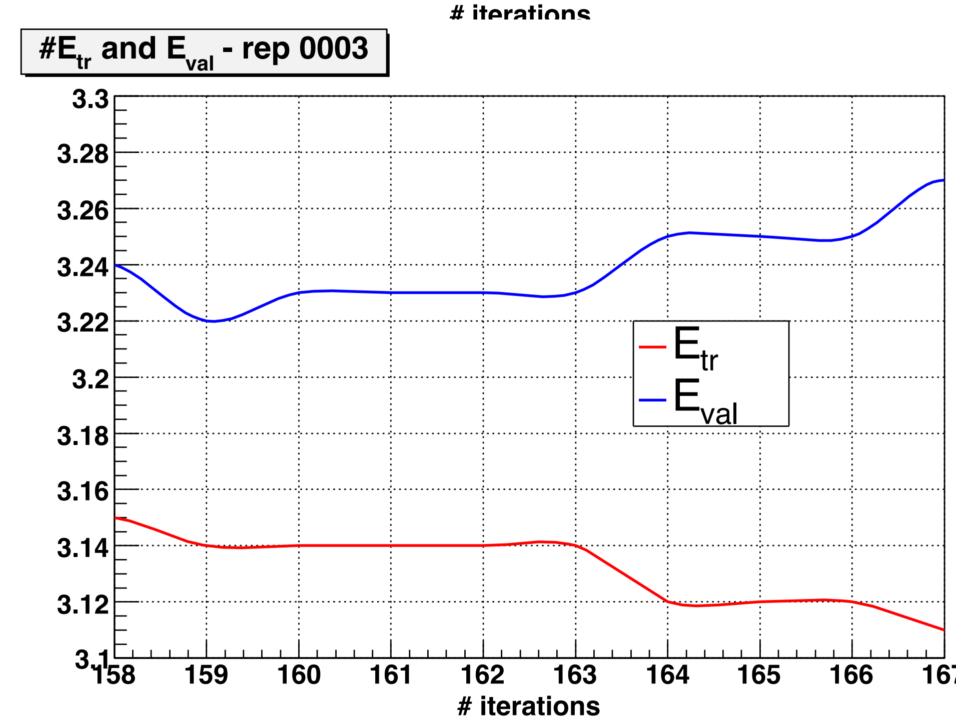
$$g(x) = \frac{1}{1 + e^{-x}}$$

- **Artificial Neural Networks** provide us with a parametrisation which is extremely **redundant** and **robust against variations**
- Very **efficient algorithms** are available which allow us to train NN (efficient fit to **large datasets** in a **very high dimensional parameter space**)
- ... but in the end they are **just another set of functions**

NNPDF methodology: Cross-validation stopping

- Separate data in each randomly in two subsets (**training** set, **validation** set)
- **Train** each neural network **minimising** an appropriate **figure of merit** on the **training set** and **monitor** the behaviour of the **figure of merit** on the **validation set**
- In our case we **train** Neural Networks **using standard back-propagation** (gradient based method), training using **Genetic Algorithms** also possibile (minimise probability of being trapped in local minima)
- Best fit defined as the point where the **validation** set f**igure of merit stops improving**



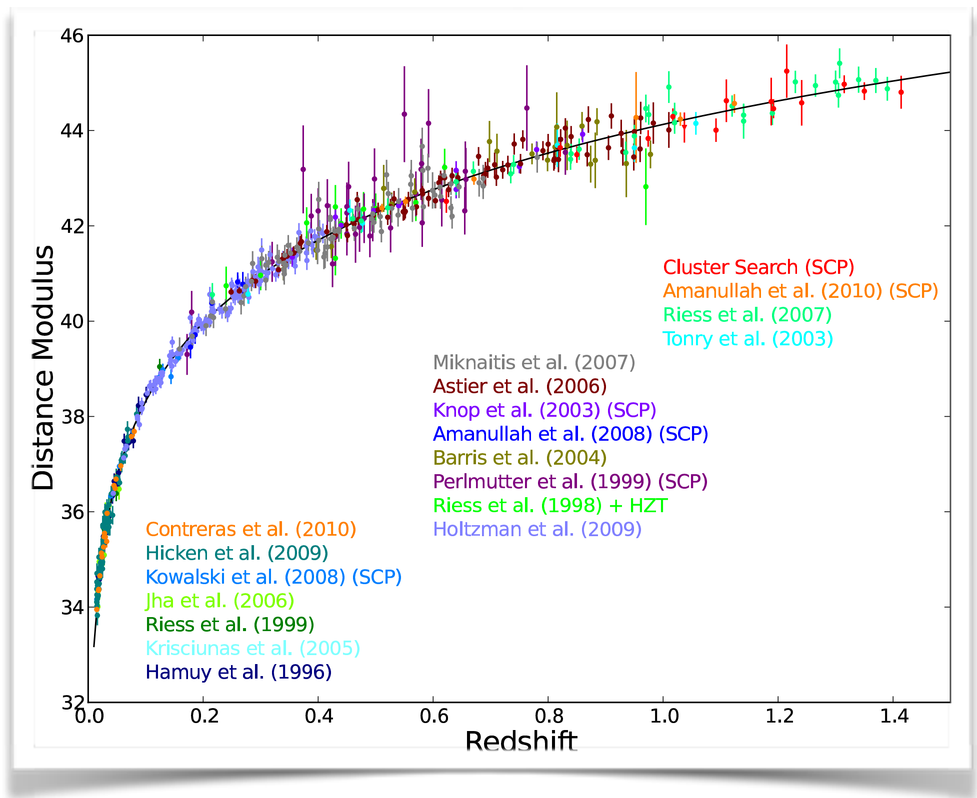


The Data - Union 2.1

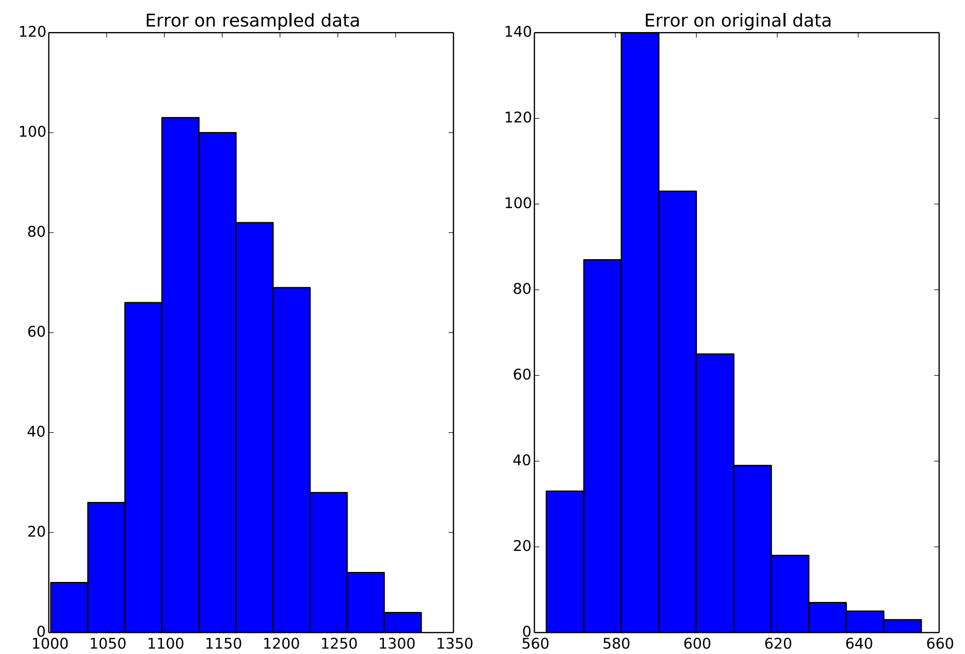
- Latest update from the Supernova Cosmology Project

  [Suzuki et al., arXiv:1105.3470]

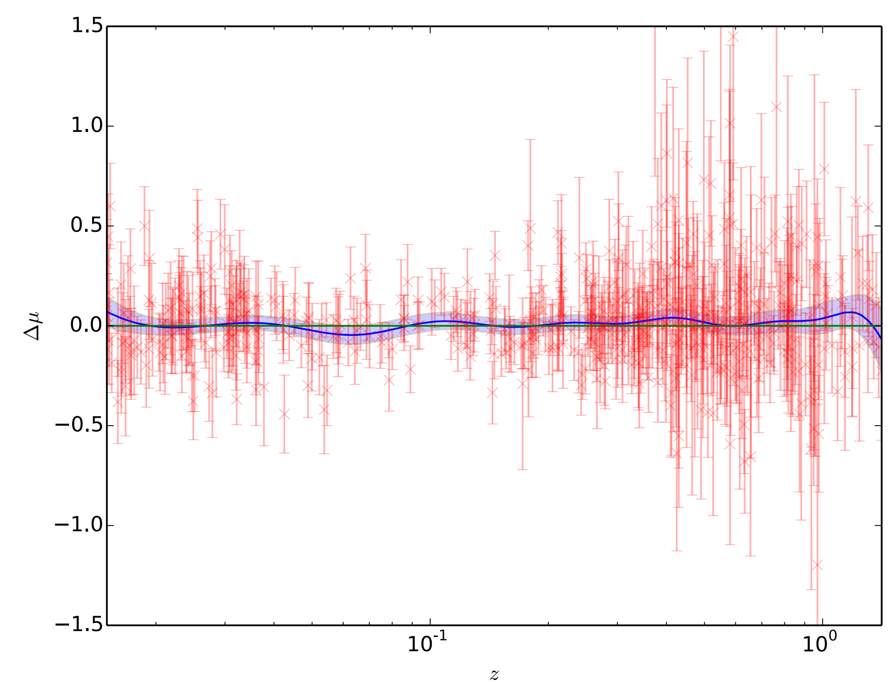
- 580 SN Ia, combined from a number of surveys
- Redshift range: **0.015 < z < 1.414**
- **Common light curve fitting** for all datasets, performed with SALT2
- **Covariance matrix** available including **correlated systematics** (not used in the preliminary results I will show)



Fitting Union2.1 using Neural Networks



Neural Network parametrisation provides a **good fit** to the data (also if we do not include correlated systematics)



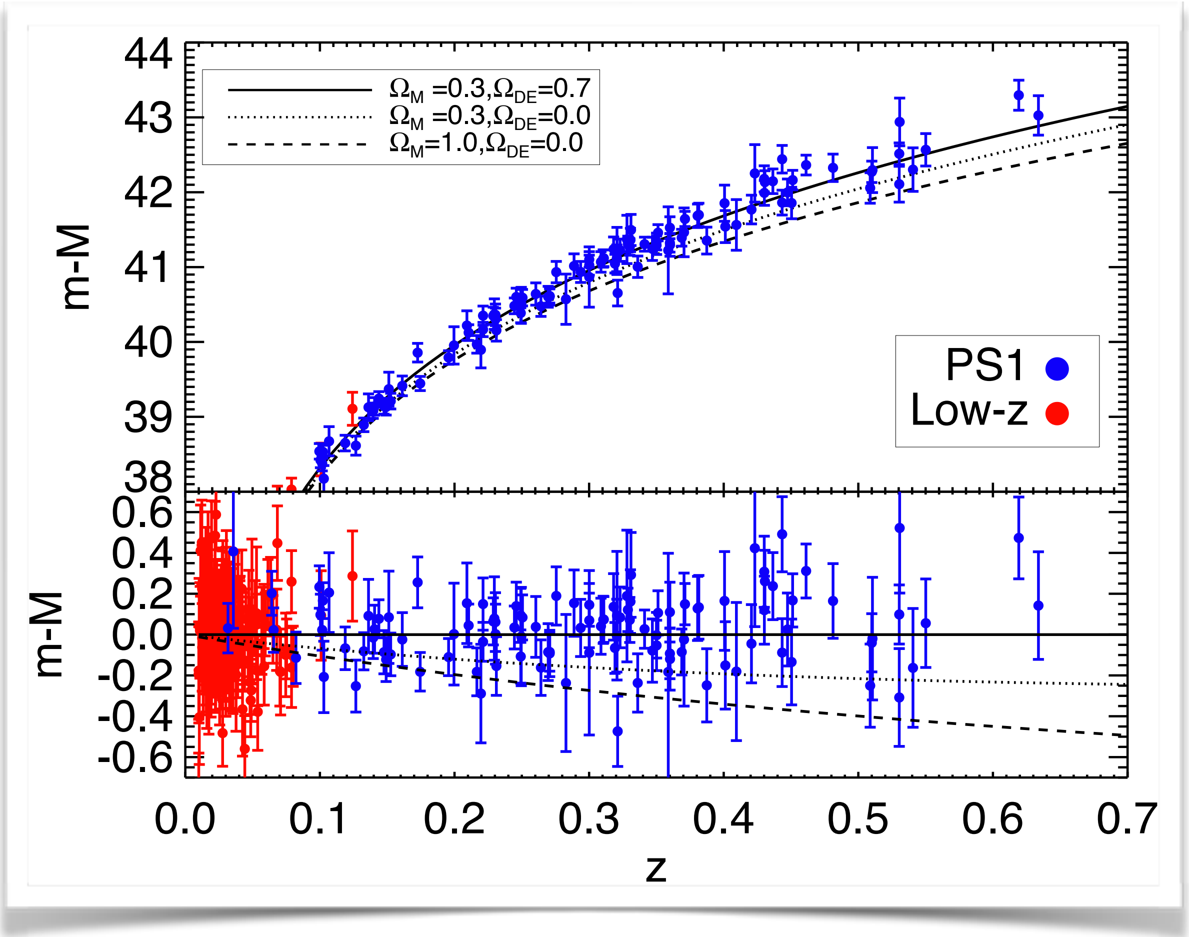
ΛCDM (Ω_m=0.3) is in **agreement** with our **Neural Network parametrisation** of the data over the **whole z-range**

The Data - Pan-STARRS1

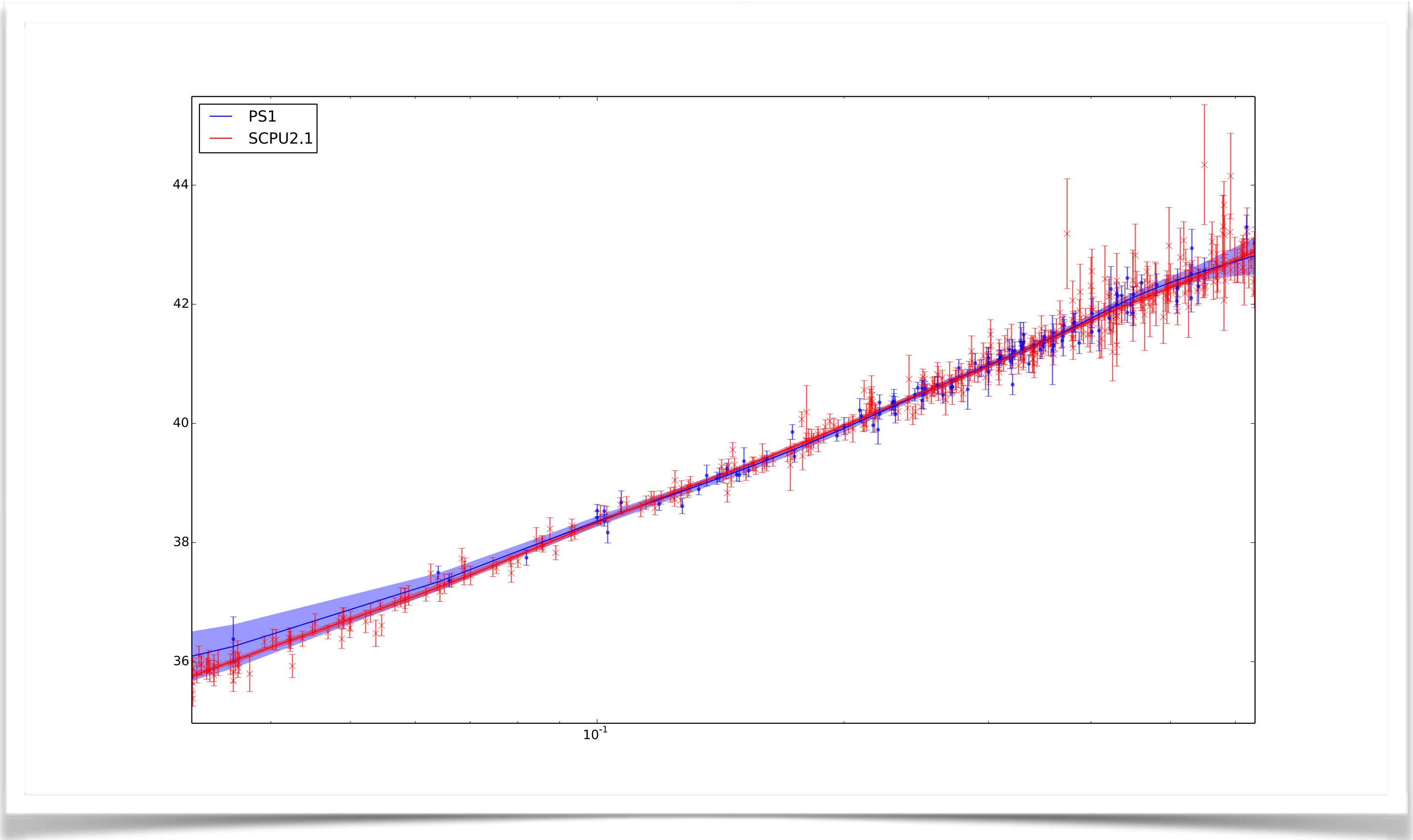
- First data release in October 2013: Pan-STARRS1

  [arXiv:1310.3828, arXiv:1310.3824]

- 146 spectroscopically confirmed SN Ia
- Redshift range: $0.03 < z < 0.65$
- Light curve fitting performed using (a modified version of) SALT2



Pan-STARRS vs. Union 2.1



- Nice **compatibility** of determination from both experiments in regions **where data overlap**
- **Uncertainties** from **Pan-STARRS1** determination grow **larger where data are sparse**

Model-independent fits to model-dependent data?

$$\mu_B = m_B^{\max} + \alpha \cdot x_1 - \beta \cdot c + \delta \cdot P(m_*^{\text{true}} < m_*^{\text{threshold}}) - M_B,$$

where M_B is the absolute B-band magnitude of a SN Ia with $x_1 = 0, c = 0$ and $P(m_\star^{\text{true}} < m_\star^{\text{threshold}}) = 0$. The parameters α, β, δ and M_B are nuisance parameters that are fitted simultaneously with the cosmological parameters. The SN Ia

4.5. *Systematic errors*

In this paper, we follow the systematics analysis we presented in Amanullah et al. (2010). Systematic errors that directly affect supernova distance measurements (calibration, and galactic extinction, for example) are treated as nuisance parameters to be fit simultaneously with the cosmology. Minimizing over these nuisance parameters gives additional terms to add to the distance modulus covariance matrix

$$U_{ij} = \sum_{\epsilon} \frac{d\mu_i(\alpha,\beta)}{d\epsilon} \frac{d\mu_j(\alpha,\beta)}{d\epsilon} \sigma_\epsilon^2 , \quad (5)$$

where the sum is over each of these distance systematic errors in the analysis. (Although the distance modulus depends on δ as well as α and β, the derivatives with respect to the zero-points do not.) In this analysis, α and β have little interaction with cosmological parameters. When computing cosmologi-

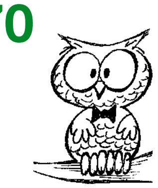
4.4. *Fitting the Cosmology*

Following Amanullah et al. (2010), the best-fit cosmology is determined by minimizing

$$\chi^2_{\text{stat}} = \sum_{\text{SNe}} \frac{[\mu_B(\alpha,\beta,\delta,M_{\text{B}}) - \mu(z;\Omega_m,\Omega_w,w)]^2}{\sigma^2_{\text{lc}} + \sigma^2_{\text{ext}} + \sigma^2_{\text{sample}}}. \quad (4)$$

A detailed discussion of the terms in this equation can be found in Amanullah et al. (2010). We only comment on the final term in the denominator, σ^2_{sample}, which is computed by setting the reduced χ^2 of each sample to unity. This term was referred to as "$\sigma^2_{\text{systematic}}$" in Kowalski et al. (2008); Amanullah et al. (2010). We note that σ^2_{sample} includes intrinsic dispersion as well as sample-dependent effects. This term effectively further deweights samples with poorer-quality data that has sources of error which have not been accounted for. As noted in Amanullah et al. (2010), this may occasionally deweight an otherwise well-measured supernova.

Union2.1 - arXiv:1105.3470



Model-independent fits to model-dependent data?

$$\mu_B = m_B^{\max} + \alpha \cdot x_1 - \beta \cdot c + \delta \cdot P(m_*^{\text{true}} < m_*^{\text{threshold}}) - M_B,$$

C. EMPIRICAL ADJUSTMENT OF UNCERTAINTIES

The propagated uncertainties are underestimates, as they do not account for the pixel-pixel covariance introduced by warping, sub-sampling, stacking, and convolution of the images. In order to empirically determine by how much the uncertainties are underestimated, we measure the flux f_r and its uncertainty σ_r at random positions in a given difference image in exactly the same way we measure the SN flux. We calculate the weighted mean $\bar{f}_r$ of these flux measurements. In order to guard against reduction and image artifacts, we apply a 3σ cut to the normalized flux distribution $(f_r - \bar{f}_r)/\sigma_r$, rather than cutting on the underestimated errors, σ_r, for the following reason: let's assume that all uncertainties are underestimated by the same factor s_r. If we nominally apply a N-sigma cut using these underestimated uncertainties, we effectively apply an N/s_r-cut, e.g. for a nominal 3-sigma cut and $s_r = 1.5$, the real cut-off is at 2-sigma. In order to avoid this, we determine the normalized flux distribution $(f_r - \bar{f}_r)/\sigma_r$, which has a standard deviation of s_r. The true 3-sigma outliers can then be identified and removed by doing an 3-sigma cut on the normalized flux distribution. Note that the standard deviation s_r is equivalent to the square-root of the chi-square distribution

$$s_r = \sqrt{\chi_r^2} = \frac{1}{N-2}\sum^{N}\left(\frac{f_r}{-}\bar{f}_r\sigma_r\right)^2 \qquad \text{(C1)}$$

We multiply all uncertainties by the factor s_r in order to empirically correct the uncertainties. We find that it is imperative to employ this robust way of determining s_r for the method to work correctly. The fact that the reduced chi-square of the baseline flux of the SN light curves peaks at 1.0 validates our method (see §5.3.3).

~~In addition, for a given difference image, $\bar{f}_r$ is an estimate of the bias in the flux measurements.~~ The values of $\bar{f}_r$ are in general very small, much smaller than the typical uncertainties. Nevertheless, we adjust all fluxes by this value.

Pan-STARRS1 - arXiv:1310.3828

Model-independent fits to model-dependent data?

Table 6

Constraints on standardization and cosmological parameters for subsets. M_B is the B-band corrected absolute magnitude; α, β, and δ are the lightcurve shape, color, and host mass correction coefficients, respectively. The outlier rejection is redone each time, so the totals may not add up to the whole sample. The constraints are computed including BAO, CMB, and H_0 constraints and supernova systematic errors.

Subset	Number	$M_B(h=0.7)$	α	β	δ	Ω_m	w
Whole Sample							
$z \geq 0.015$	580	$-19.321^{+0.030}_{-0.030}$	$0.121^{+0.007}_{-0.007}$	$2.47^{+0.06}_{-0.06}$	$-0.032^{+0.031}_{-0.031}$	$0.271^{+0.015}_{-0.014}$	$-1.013^{+0.068}_{-0.074}$
Correction Coefficients, Split by Redshift							
$0.015 \leq z \leq 0.10$	175	$-19.328^{+0.037}_{-0.038}$	$0.118^{+0.011}_{-0.011}$	$2.57^{+0.08}_{-0.08}$	$-0.027^{+0.054}_{-0.054}$	0.270 (fixed)	−1.000 (fixed)
$0.100 \leq z \leq 0.25$	75	$-19.371^{+0.054}_{-0.054}$	$0.146^{+0.019}_{-0.019}$	$2.56^{+0.18}_{-0.17}$	$-0.087^{+0.060}_{-0.060}$	0.270 (fixed)	−1.000 (fixed)
$0.250 \leq z \leq 0.50$	152	$-19.317^{+0.046}_{-0.046}$	$0.116^{+0.014}_{-0.013}$	$2.46^{+0.12}_{-0.12}$	$-0.042^{+0.066}_{-0.066}$	0.270 (fixed)	−1.000 (fixed)
$0.500 \leq z \leq 1.00$	137	$-19.307^{+0.048}_{-0.049}$	$0.124^{+0.019}_{-0.019}$	$1.46^{+0.19}_{-0.19}$	$0.023^{+0.060}_{-0.060}$	0.270 (fixed)	−1.000 (fixed)
$z \geq 1.000$	25	$-19.289^{+0.217}_{-0.254}$	$-0.019^{+0.072}_{-0.076}$	$3.48^{+1.13}_{-0.89}$	$-0.151^{+0.384}_{-0.446}$	0.270 (fixed)	−1.000 (fixed)
Effect of δ on w							
$z \geq 0.015$	580	$-19.340^{+0.026}_{-0.026}$	$0.123^{+0.007}_{-0.007}$	$2.47^{+0.06}_{-0.06}$	−0.080 (fixed)	$0.272^{+0.015}_{-0.014}$	$-1.004^{+0.067}_{-0.072}$
$z \geq 0.015$	580	$-19.303^{+0.031}_{-0.031}$	$0.120^{+0.007}_{-0.007}$	$2.47^{+0.06}_{-0.06}$	0.000 (fixed)	$0.271^{+0.015}_{-0.014}$	$-1.013^{+0.069}_{-0.075}$
Cosmological Results, Split by Lightcurve Color and Shape							
$c \geq 0.05$	256	$-19.387^{+0.037}_{-0.038}$	$0.118^{+0.011}_{-0.011}$	$2.77^{+0.09}_{-0.09}$	$-0.057^{+0.052}_{-0.052}$	$0.269^{+0.015}_{-0.014}$	$-1.028^{+0.077}_{-0.084}$
$c \leq 0.05$	321	$-19.323^{+0.030}_{-0.030}$	$0.125^{+0.011}_{-0.010}$	$1.29^{+0.32}_{-0.33}$	$-0.057^{+0.038}_{-0.038}$	$0.275^{+0.015}_{-0.014}$	$-0.982^{+0.069}_{-0.075}$
$x_1 \geq -0.25$	311	$-19.366^{+0.041}_{-0.041}$	$0.020^{+0.026}_{-0.025}$	$2.58^{+0.10}_{-0.10}$	$-0.004^{+0.047}_{-0.047}$	$0.269^{+0.015}_{-0.014}$	$-1.037^{+0.077}_{-0.085}$
$x_1 \leq -0.25$	269	$-19.386^{+0.044}_{-0.045}$	$0.152^{+0.021}_{-0.020}$	$2.43^{+0.08}_{-0.08}$	$-0.087^{+0.050}_{-0.050}$	$0.267^{+0.015}_{-0.014}$	$-1.045^{+0.077}_{-0.084}$
Correction Coefficients and M_B for the Large Datasets							
Hicken et al. (2009)	94	$-19.314^{+0.055}_{-0.055}$	$0.115^{+0.015}_{-0.015}$	$2.74^{+0.11}_{-0.11}$	$-0.053^{+0.098}_{-0.099}$	0.270 (fixed)	−1.000 (fixed)
Holtzman et al. (2009)	129	$-19.336^{+0.051}_{-0.051}$	$0.149^{+0.014}_{-0.013}$	$2.40^{+0.15}_{-0.14}$	$-0.061^{+0.050}_{-0.050}$	0.270 (fixed)	−1.000 (fixed)
Miknaitis et al. (2007)	74	$-19.325^{+0.078}_{-0.080}$	$0.113^{+0.037}_{-0.035}$	$2.49^{+0.17}_{-0.16}$	0.000 (fixed)	0.270 (fixed)	−1.000 (fixed)
Astier et al. (2006)	71	$-19.292^{+0.047}_{-0.048}$	$0.145^{+0.019}_{-0.018}$	$1.70^{+0.18}_{-0.18}$	$-0.023^{+0.040}_{-0.040}$	0.270 (fixed)	−1.000 (fixed)
$z > 0.9$, Split by Galaxy Host							
Early Type $z > 0.9$	13	$-19.388^{+0.139}_{-0.186}$	$0.112^{+0.139}_{-0.151}$	$3.16^{+1.84}_{-1.26}$	0.000 (fixed)	0.270 (fixed)	−1.000 (fixed)
Late Type $z > 0.9$	15	$-19.141^{+0.067}_{-0.067}$	$0.094^{+0.049}_{-0.041}$	$0.49^{+0.85}_{-0.69}$	0.000 (fixed)	0.270 (fixed)	−1.000 (fixed)

Union2.1 - arXiv:1105.3470

What is the **dependence** on the "**nuisance parameters**" (α,β)?

Table 3
Effects of Choices for Intrinsic Scatter

Intrinsic Scatter	α	β
$\sigma_{\text{int},m_B} = 0.122$	0.147 ± 0.010	3.10 ± 0.12
$\sigma_{\text{int},c} = 0.025$	0.141 ± 0.010	3.86 ± 0.15

Note. — Intrinsic scatter σ_{int,m_B} and $\sigma_{\text{int},c}$ in the PS1+lz sample, and how α and β vary for each method. The magnitudes of each scatter given above is such that the total reduced χ^2 of the sample is $\sim$1.0.

Pan-STARRS1 - arXiv:1310.3828

Can we address the question by taking a step back and parametrising the data on distance moduli as a function of (z, x_1, c)

Instead of conclusions …

- Is there space for **model-independent** tests of **cosmic acceleration**?
- If so, there is a **wide range of methods** that could be employed and that could provide **parametrisations of the data** which could be used for **reliable and fast testing** of **different models**
- We propose such a **“new” technique** based on **Monte Carlo methods** for uncertainty estimation and **Neural Networks** to parametrise the data and applied it to **SN Ia** measurements

… but in the end

it is of no use to use a **model-independent method** to look at **model-dependent data**