Blinding the Hubble Constant

Blind Analysis in High-Stakes Survey Science: When, Why, and How?

KIPAC/SLAC

Bonnie Zhang (ANU)
@astrobonnie

CAASTRO
ARC CENTRE OF EXCELLENCE FOR ALL-SKY ASTROPHYSICS

case study

MNRAS 000, 1–35 (2017) Preprint 7 March 2017 Compiled using MNRAS LaTeX style file v3.0

A blinded determination of H_0 from low-redshift Type Ia supernovae, calibrated by Cepheid variables

Bonnie R. Zhang[1,2★], Michael J. Childress[3], Tamara M. Davis[2,4], Natallia V. Karpenka[3], Chris Lidman[2,5], Brian P. Schmidt[1,2], Mathew Smith[3]

[1] Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT 2611, Australia.
[2] ARC Centre of Excellence for All-sky Astrophysics (CAASTRO).
[3] School of Physics and Astronomy, University of Southampton, Southampton, SO17 1BJ, UK.
[4] School of Mathematics and Physics, University of Queensland, QLD 4072, Australia.
[5] Australian Astronomical Observatory, PO Box 915, North Ryde, NSW 1670, Australia.

Accepted XXX. Received YYY; in original form ZZZ

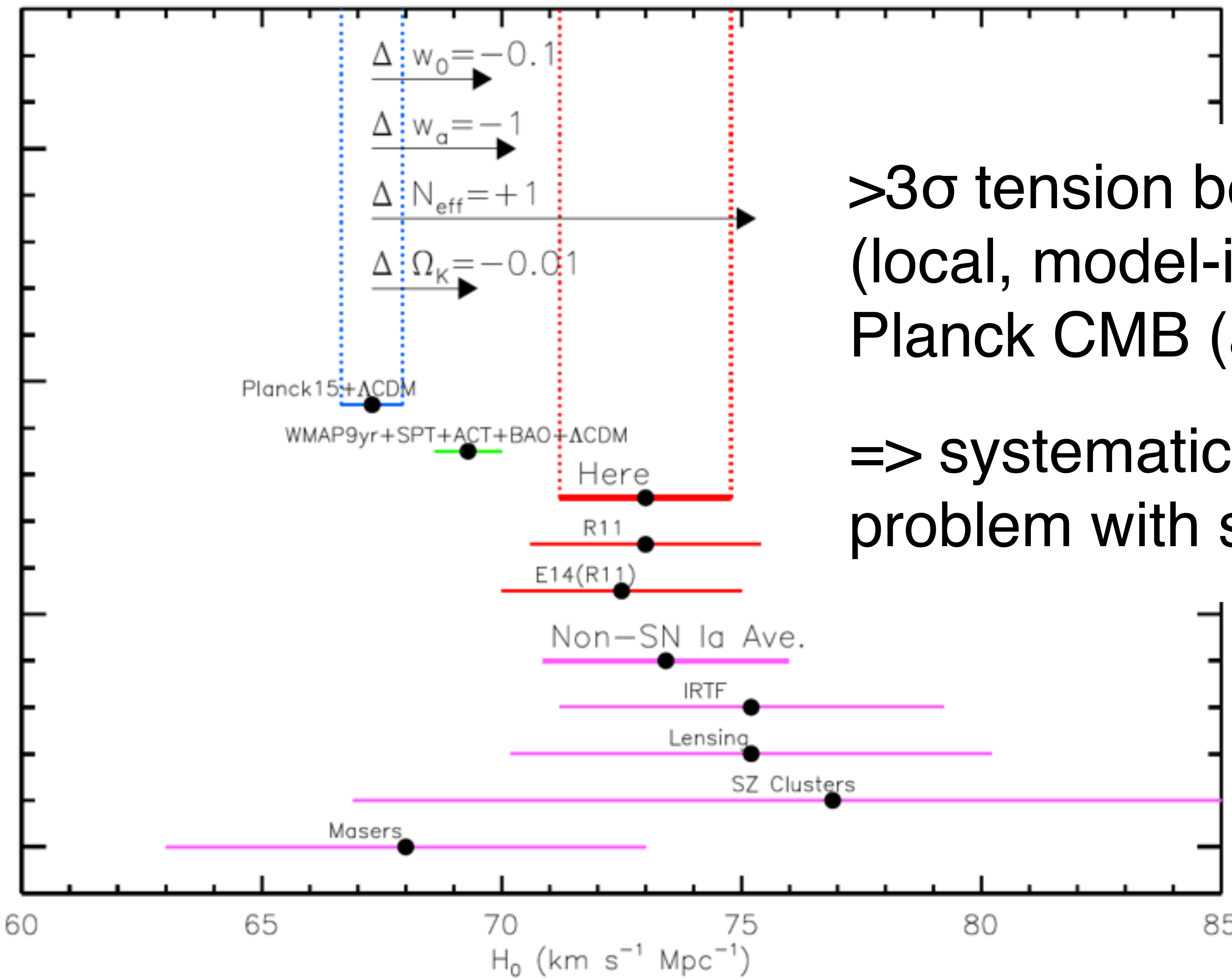
under revision (MNRAS); value of H_0 not yet revealed

ABSTRACT

Presently a $> 3\sigma$ tension exists between values of the Hubble constant H_0 derived from analysis of fluctuations in the Cosmic Microwave Background by Planck, and local measurements of the expansion using calibrators of type Ia supernovae (SNe Ia). We perform a blinded reanalysis of Riess et al. (2011) to measure H_0 from low-redshift SNe Ia, calibrated by Cepheid variables and the geometric distance to NGC 4258. This paper is a demonstration of techniques to be applied to the Riess et al. (2016) data. Our end-to-end analysis starts from available CfA3 and LOSS photometry, providing an independent validation of Riess et al. (2011). We obscure the value of H_0 throughout our analysis and the first stage of the referee process, because calibration of SNe Ia requires a series of often subtle choices, and the potential for results to be affected by human bias is significant. Our analysis departs from that of Riess et al. (2011) by incorporating the covariance matrix method adopted in SNLS and JLA to quantify SN Ia systematics – including selection bias of the SN Ia sample, and by including a simultaneous fit of all SN Ia and Cepheid data. We find $H_0 = X \pm Y(\text{stat}) \pm Z(\text{sys})$ **(yet to be revealed)** with a three galaxy (NGC 4258+LMC+MW) anchor. The relative uncertainties are 4.3% statistical, 1.1% systematic, and 4.4% total, larger than in Riess et al. (2011) (3.3%) and the Efstathiou (2014) reanalysis (3.4%). Our error budget for H_0 is dominated by statistical errors due to the small size of the supernova sample, whilst the systematic contribution is dominated by variation in the Cepheid fits, and for the SN Ia, uncertainties in the host galaxy mass dependence and Malmquist bias.

Key words: distance scale; cosmology: observations; supernovae: general; stars: variables: Cepheids

trouble with H_0

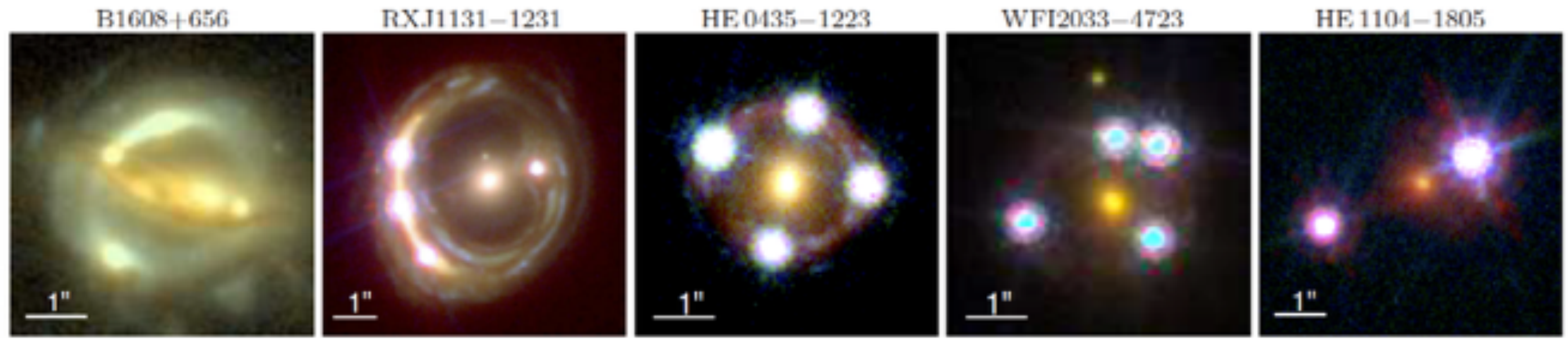


>3σ tension between SNe Ia (local, model-independent) & Planck CMB (assumes ΛCDM)

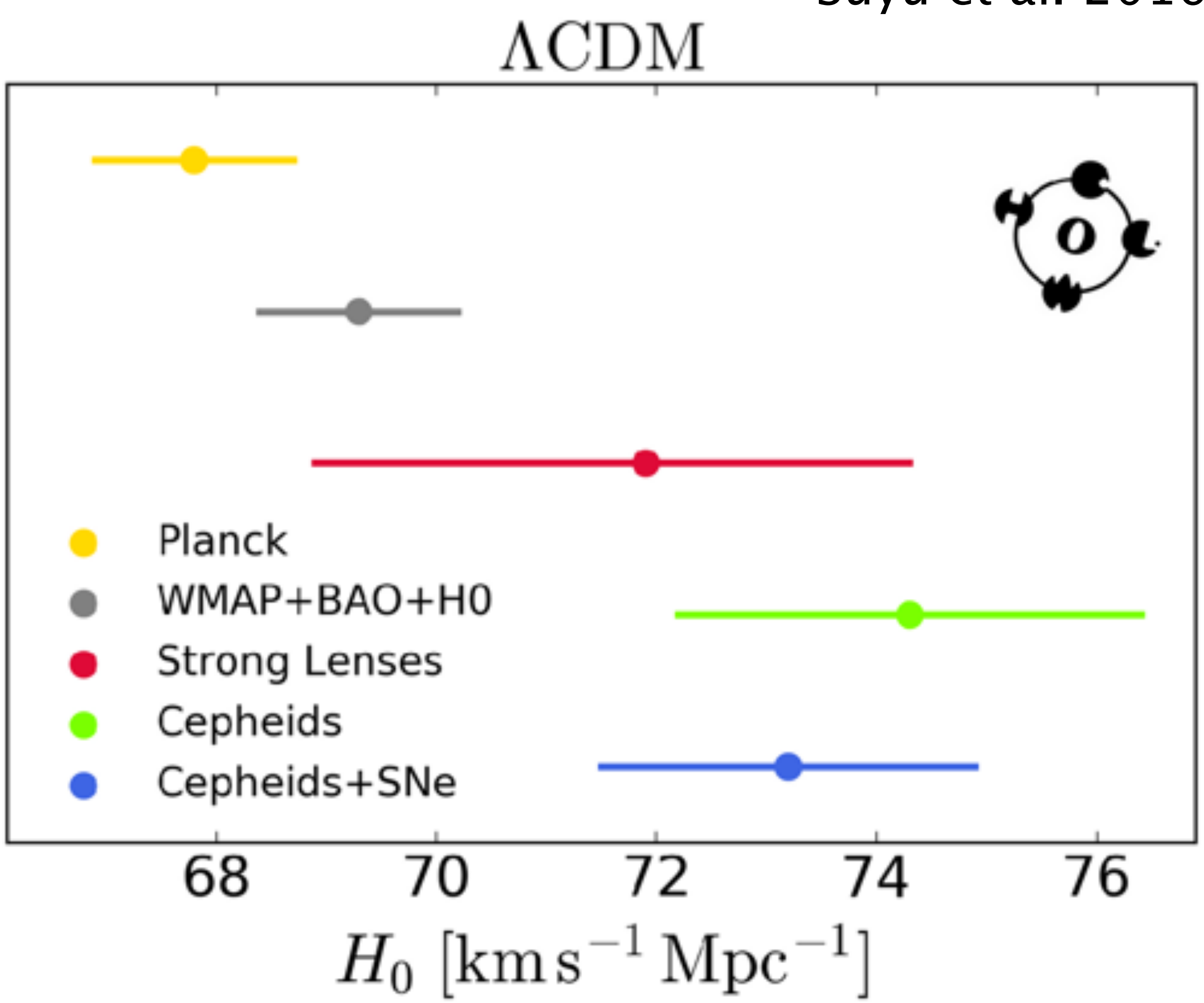
=> systematic error, or problem with standard model?

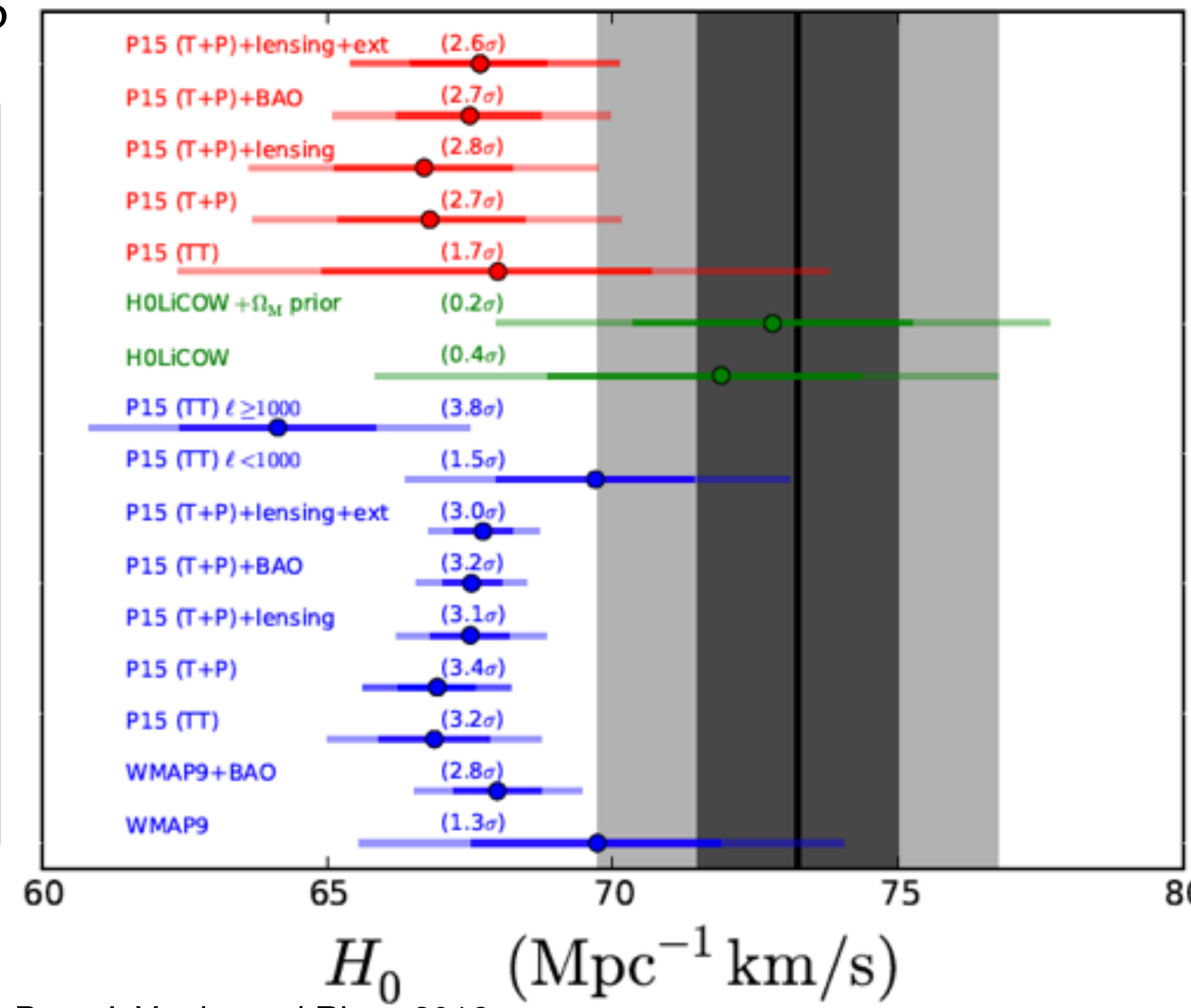
H0LiCOW

time-delay distances from strong lenses (blinded)



Suyu et al. 2016





https://shsuyu.github.io/H0LiCOW/site/

Bernal, Verde, and Riess 2016

Blinding is complex for technical and human reasons. It was made possible here by:

- applying blind at parameter/PDF level (i.e. raw data unchanged) — reduced to maths problem rather than data problem
- small team, reanalysing existing data set

CAASTRO
ARC CENTRE OF EXCELLENCE FOR ALL-SKY ASTROPHYSICS

We use data from Riess+ 2011 (as a proof of concept), with the following differences:

1. account for SN Ia systematic uncertainties with covariance matrices (SALT2 framework)
2. simultaneous fit to all 16 parameters:
   $\Theta = \{\alpha, \beta, M_B, \mathcal{H}, b_W, Z_W, M_W, \mu_{4258}, \Delta\mu_i\}$
   (where $\mathcal{H} = 5\log_{10} H_0 - 25.$)
3. M_B and H_0 blinded by inserting unknown but retrievable offsets into analysis — **blind not lifted until paper is accepted!**

measuring H_0

- Hubble law $\mathbf{v(z) = H_0\, D(z)}$
- LHS: redshifts + cosmology; RHS: distance ladder via standard candles: SNe Ia, Cepheids, geometrical distance to NGC 4258 (+LMC, MW Cepheids)
- simultaneously fit magnitudes of Cepheids and Type Ia supernovae

$$\Theta = \{\alpha, \beta, M_B, \mathcal{H}, b_W, Z_W, M_W, \mu_{4258}, \Delta\mu_i\}$$

Cepheids (Leavitt law)

$$m_{W\,ij} = b_W(\log_{10} P_{ij} - 1) + Z_W \Delta \log_{10}[O/H]_{ij} + M_W + \mu_{4258} + \Delta\mu_i$$

$$m_{Bi} = M_B^* - \alpha X_{1i} + \beta C_i + \mu_{4258} + \Delta\mu_i$$

SNe Ia in Cepheid hosts

$$m_{Bk} = M_B^* - \alpha X_{1k} + \beta C_k + 5\log_{10}(cz_k f(z_k)) - \mathcal{H}$$

$$\mathcal{H} = 5\log_{10} H_0 - 25.$$

low-z SNe Ia (CfA3+LOSS)

$$f(z) \equiv 1 + \frac{(1-q_0)z}{2} - \frac{(1-q_0-3q_0^2+j_0)z^2}{6}$$

mechanics of blinding

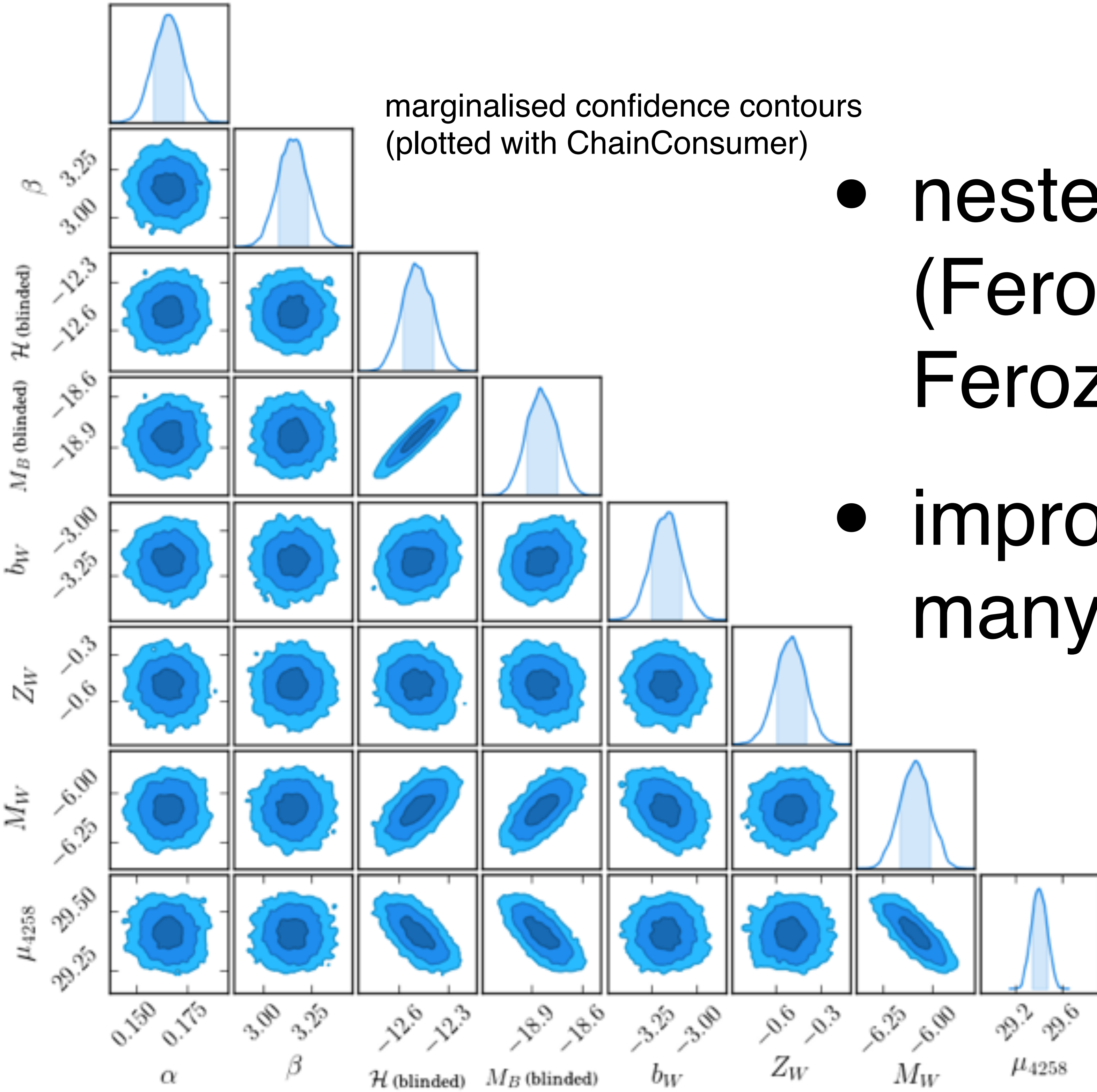
known integers (n_1, n_2) $\xrightarrow{\text{seed}}$ $N(\mu_1, \sigma_1^2)$, $N(\mu_2, \sigma_2^2)$ $\xrightarrow{\text{generate}}$ **unknown offsets (off_1, off_2)**

$$m_B \mapsto m_B + \text{off}_1 \qquad (\text{blind } \mathrm{M_B})$$

$$\mathcal{L}(\mathcal{H}) \mapsto \mathcal{L}(\mathcal{H} + \text{off}_2) \qquad (\text{blind } \mathrm{H_0})$$

Precautions: never print offsets; never see unblinded magnitudes.

MultiNest



marginalised confidence contours
(plotted with ChainConsumer)

- nested sampling technique (Feroz and Hobson 2008, Feroz et al. 2009, 2013)
- improves on MCMC for many parameters

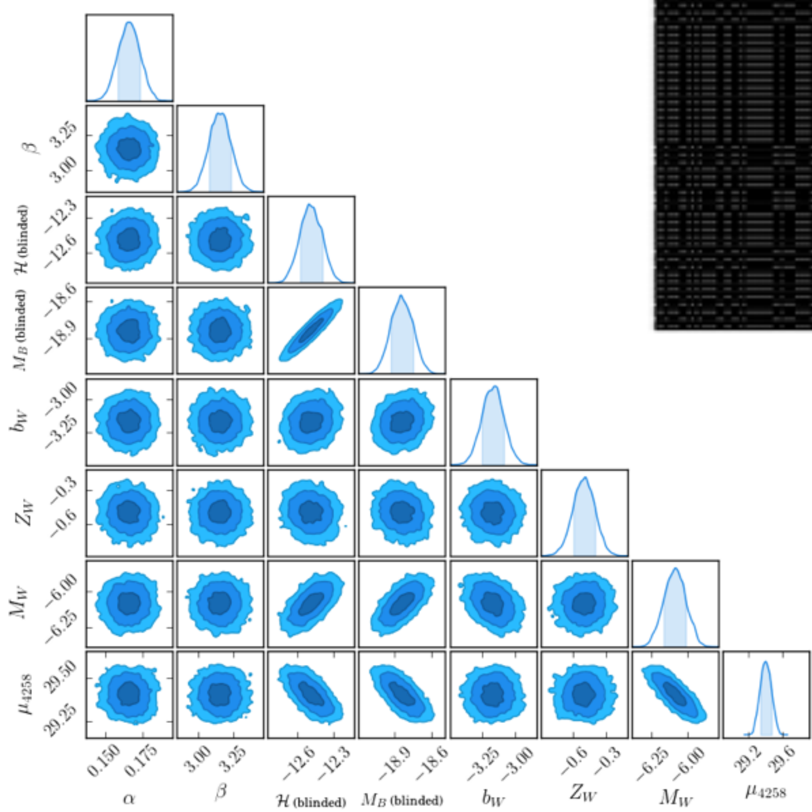
final uncertainties

Anchor	all	NGC 4258 only
$\mathcal{H}$	-12.381	-12.431
$\sigma_{\mathcal{H}}$		
Statistical	0.093	0.107
Systematic	0.023	0.023
Relative H_0 error (%)		
Statistical	4.3	4.9
Systematic	1.1	1.1
Total	4.4	5.0

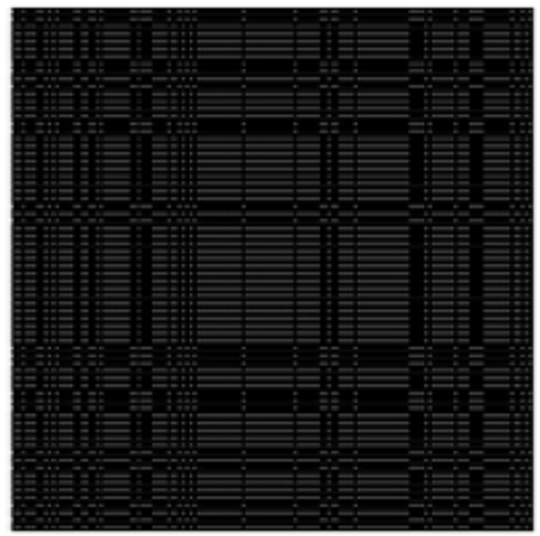
- ~1% larger than Riess et al. 2011, and Efstathiou 2014 reanalysis:
  - R11: 3.3% (4.1% with NGC 4258 only)
  - E14: 3.4% (4.7% with NGC 4258 only)

breakdown of uncertainties

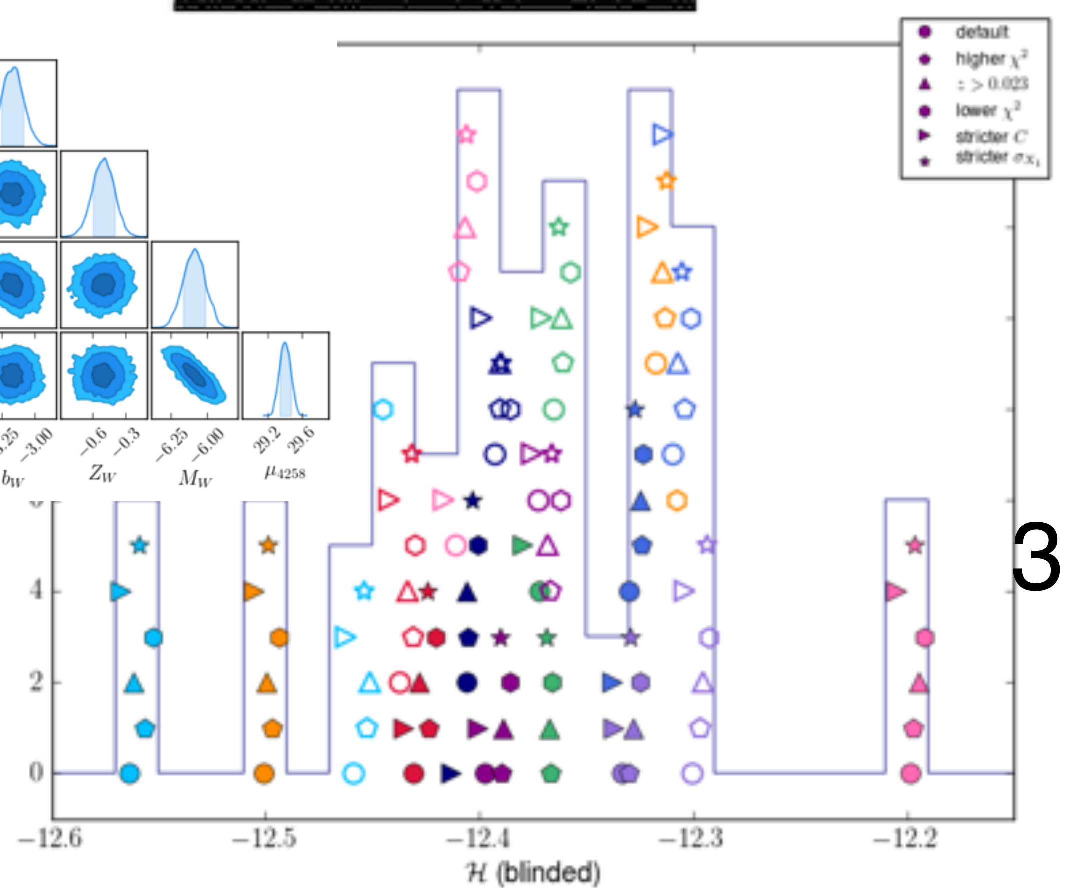
1&2



2



3



1. pure statistical (from number of SNe, Cepheids)
2. contribution to ‘statistical’ uncertainty from covariance matrices (correlated SN errors)
3. systematic, from varying choices in SN and Cepheid fits

criticisms & reactions

1. "[Blinding the referee process] hasn't been done before, is it because you don't trust the editor/referee?"
2. "Why should we trust you?"
3. "What if you find $H_0 = 20$ [km/s/Mpc]?"
4. "But we're careful and thorough in the first place and have accounted for all systematics so frankly it's unnecessary."
5. More time and effort than it's worth
6. At what point do we learn from the process?
7. "I found myself checking through everything more carefully than I would have if I had known the answer."

summary & lessons

Reanalysis of Riess et al. 2011 data finds larger uncertainty in H_0 (3.3% —> 4.4%); value will be unblinded once paper is accepted. Blinding H_0 is relatively simple, and necessary given the current tension. Some lessons:

- One size does not fit all (this analysis is small in scale and relatively simple).
- Blinding is a means to reduce bias, not an end in itself. Other mechanisms: salting, public code/data.
- Potential biases/choices made should go into the error budget. The more unbiased we think we are, the more necessary it is to actively preclude bias.
- Sociological: implement first, deal with critics second.