MNRAS **000**, 1–3 (2016) Preprint 11 August 2016 Compiled using MNRAS LaTeX style file v3.0

Investigating the Gaussianity of Supernova SALT2 Summary Statistics

Samuel R. Hinton,[1⋆] Alex G. Kim,[3] Tamara M. Davis,[1,2] Edward Macaulay[1,2]
[1] *School of Mathematics and Physics, The University of Queensland, Brisbane, QLD 4072, Australia*
[2] *ARC Centre of Excellence for All-sky Astrophysics (CAASTRO)*
[3] *Physics Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT
Write an abstract.

1 INTRODUCTION

In the years since the discovery of the acceleration of our universe (Riess et al. 1998; Schmidt et al. 1998; Perlmutter et al. 1999), supernova cosmology continued to grow as an important cosmological probe. Recent supernova studies offer far greater statistical power than ever before, systematics are better understood, and the very phenomenon of supernova are described with improved models (Guy et al. 2007; Conley et al. 2011; Betoule et al. 2014; Rubin et al. 2015), and with increased statistical power and model fidelity, systematics are now the limiting constraint in supernova studies (Conley et al. 2011; Suzuki et al. 2012; Scolnic et al. 2014) UPDATE citations, used from Betoule. In this paper we investigate a potential source of systematics unaddressed in prior studies – the gaussianity of supernova summary statistics. We perform this investigation within the context of the Dark Energy Survey, which is introduced in Section 2. In Section 3 we summarise recent supernova models and analyses, and in Section 4 and Section 5 we investigate the effect of the assumption of gaussianity on both the supernova distributions obtained and the follow on impact on cosmology.

2 THE DARK ENERGY SURVEY

The Dark Energy Survey (DES, The Dark Energy Survey Collaboration 2005; Dark Energy Survey Collaboration et al. 2016) is a 5000 sq. deg. photometric redshift survey. The survey will image approximately 300 million galaxies in 5 broadband filters ($grizY$) using the Dark Energy Camera (DECam, Flaugher et al. 2015). Forty sq. deg. of the sky will be repeatedly images in the $griz$ bands for the supernova luminosity distance probe, and is expected to yield observations of approximately 1900 Type Ia supernova up to a redshift of $z = 1.2$. No idea where to look for more details on this. On cadence, updated estimates, cadence, difference between shallow and deep fields.

3 MOTIVATION

In order to perform a supernova analysis, an underlying Type Ia supernova model must be adopted. The most used model in recent years has been that of the SALT2 model (Guy et al. 2007, 2010; Mosher et al. 2014) and in this analysis we restrict our investigations to this model. The SALT2 model characterises the Type Ia spectral energy distribution (SED) as a function of magnitude x_0, stretch x_1, colour c, phase p and wavelength λ via

$$F(p, \lambda) = x_0 \left[M_0(p, \lambda) + x_1 M_1(p, \lambda)\right] \exp\left[cCL(\lambda)\right]. \tag{1}$$

Thus characterising an individual supernova's magnitude, colour and stretch (and associated covariance) allows recovery of the modelled SED. These summary statistics are used instead of the observed light curves in supernova analyses for computational reasons across many different fitting methodologies. In the Bayesian hierarchical analysis performed on the Unity 2.1 dataset (Suzuki et al. 2012) in Rubin et al. (2015), each supernova is described by a latent colour and stretch, which means an analysis with N supernova grows in dimensionality as $2N + x$, where x represents underlying (non-latent) parameters (such as Ω_m). If each supernova has M observations, a full hierarchical model would increase in dimensionality to $(2 + M)N + x$, providing vastly increased computational complexity in resolving the underlying posterior surface. In frequentist analyses like that of the Joint Lightcurve Analysis (Betoule et al. 2014), the correlation of all light curve points via the calibration zero points instead leads to computational difficulties in that a covariance matrix of approximate size MN is introduced into the likelihood calculation, which would be prohibitively expensive to invert. Therefore, instead of utilising light curve data directly in cosmological models, light curves are fit to the supernova data to produce summary statistics, and these are then used in cosmology analyses.

One key assumption when doing this is that these summary statistics are in fact gaussian in nature. Often this is explicitly assumed in the analysis (March et al. 2011; Rubin et al. 2015), however it is often implicitly assumed (Sullivan et al. 2011; Suzuki et al. 2012; Campbell et al. 2013; Betoule et al. 2014). Investigation into asymmetries in the summary statistics are generally limited to asymmetries in the parent colour and stretch distribution (Scolnic et al. 2014). More citations. In our paper we therefore do not investigate biases or asymmetries found in underlying population distributions, but asymmetries and biases in the SALT2 parameter probability surface, and the impact of assuming the Gaussianity of said surface.

4 METHODOLOGY

We begin our investigation by generating supernova light curves for mock supernova in both the DES deep and shallow fields. As prior studies have investigated the effects of parent colour and stretch distributions, we generate all canonical supernova ($x_1 = c = 0$). The supernova are simulated in the *griz* bands with an observational cadence of 5 days, starting from a time $t = -35$, giving 7 observations before the peak to remove the biases introduced by having little data before the luminosity peak CITE. Weather conditions and moon phase are simulated based on observed DES seeing Cite excel file?. Simulated light curves undergo a simple selection cut of having at least one observation in any band above a signal to noise of 5.

We investigate the potential asymmetries in SALT2 summary statistics by fitting the light curves and obtaining summary statistics using different algorithms in the software package `sncosmo` (Barbary 2014). Specifically, we obtain a full posterior surface using our own MCMC fitting methodology, and compare against summary statistics generated using a different MCMC approach (the `mcmc_lc` method in `sncosmo`), nestled sampling (the `nest_lc` method) and minuit (via the `fit_lc` method). The MCMC based algorithm utilises the package `emcee` (Foreman-Mackey et al. 2013), whilst the nestled sampling is based off `nestle` (Skilling 2004; Barbary 2015), and both algorithms should map out the full posterior surface, giving summary statistics from the mean and covariance of that full surface. The third algorithm, invoked via `fit_lc` uses the minuit algorithm (James & Roos 1975; Ongmongkolkul & Deil 2015), and as such is a maximum likelihood estimator and does not map out the full posterior surface. For a Gaussian surface, all three algorithms should agree, however for non-Gaussian surfaces the minuit algorithm will be expected to produce different output. As such, we test the SALT2 probability surfaces for skewness and fit with the four algorithms described previously. The summary statistics are converted into distance modulus via the Philips relation

$$\mu = m_B - M + \alpha x_1 - \beta c, \tag{2}$$

where we adopt the fiducial $\alpha = 0.14$ and $\beta = 3.15$ values from Betoule et al. (2014). This allows us to directly see the effects of any skewness or non-Gaussianity in any parameter on the cosmologically important distance modulus. We can compare the distributions in μ for the full posterior surface and for summary statistics, and also verify the skewness of the full distribution of μ. We investigate any shift in $\langle\mu\rangle$, and change in the standard deviation σ_μ, binning by redshift, with results illustrated in Figure 1. We find strong evidence that, with increasing redshift, the skewness of the SALT2 fit posterior increases, and causes the maximum likelihood estimator to not only produce biased output, but overestimate the uncertainty of the result. The likelihood surface for one of the skewed supernova fits is shown as an example in Figure 2, where the effect of the skewness is readily apparent. To confirm that the skewness is a result of redshift and not simply a result of the general trend of decreasing signal to noise, we plot the skewness as a function of both redshift and signal to noise in Figure 3. Interestingly, doing this shows a negative skewness for low redshift, low signal-to-noise light curves, and a positive skewness for high redshift light curves, regardless of signal-to-noise. Having confirmed that skewness becomes significant at higher redshift, we now seek to characterise this potential impact on cosmology.

5 COSMOLOGICAL IMPACT

Fit cosmology against simulated SN

Do for des shallow and deep, using survey area to produce z dist plus extra low-z sample (0.05<z<0.2). simulate from light curves, apply selection effect of ston>5to cull bad fits. have <x> shallow.

toy cosmology model. adopt α and β values for Phillips correction from Betoule2014, assume known gaussian intrinsic scatter, have Ω_m, w and M_B as cosmology parameters - marginalise over M_B.

Show cosmology results.

6 CONCLUSIONS

The last numbered section should briefly summarise what has been done, and describe the final conclusions which the authors draw from their work.

ACKNOWLEDGEMENTS

The Acknowledgements section is not numbered. Here you can thank helpful colleagues, acknowledge funding agencies, telescopes and facilities used etc. Try to keep it short.

REFERENCES

Barbary K., 2014, sncosmo v0.4.2, doi:10.5281/zenodo.11938, `http://dx.doi.org/10.5281/zenodo.11938`
Barbary K., 2015, nestle v0.1.1, `https://github.com/kbarbary/nestle`
Betoule M., et al., 2014, A&A, 568, A22
Campbell H., et al., 2013, ApJ, 763, 88
Conley A., et al., 2011, ApJS, 192, 1
Dark Energy Survey Collaboration et al., 2016, MNRAS, 460, 1270
Flaugher B., et al., 2015, AJ, 150, 150
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Guy J., et al., 2007, A&A, 466, 11
Guy J., et al., 2010, A&A, 523, A7
James F., Roos M., 1975, Computer Physics Communications, 10, 343
March M. C., Trotta R., Berkes P., Starkman G. D., Vaudrevange P. M., 2011, MNRAS, 418, 2308
Mosher J., et al., 2014, ApJ, 793, 16
Ongmongkolkul P., Deil C., 2015, iminuit v1.2, `https://github.com/kbarbary/nestle`
Perlmutter S., et al., 1999, ApJ, 517, 565
Riess A. G., et al., 1998, AJ, 116, 1009
Rubin D., et al., 2015, ApJ, 813, 137
Schmidt B. P., et al., 1998, ApJ, 507, 46
Scolnic D., et al., 2014, ApJ, 795, 45

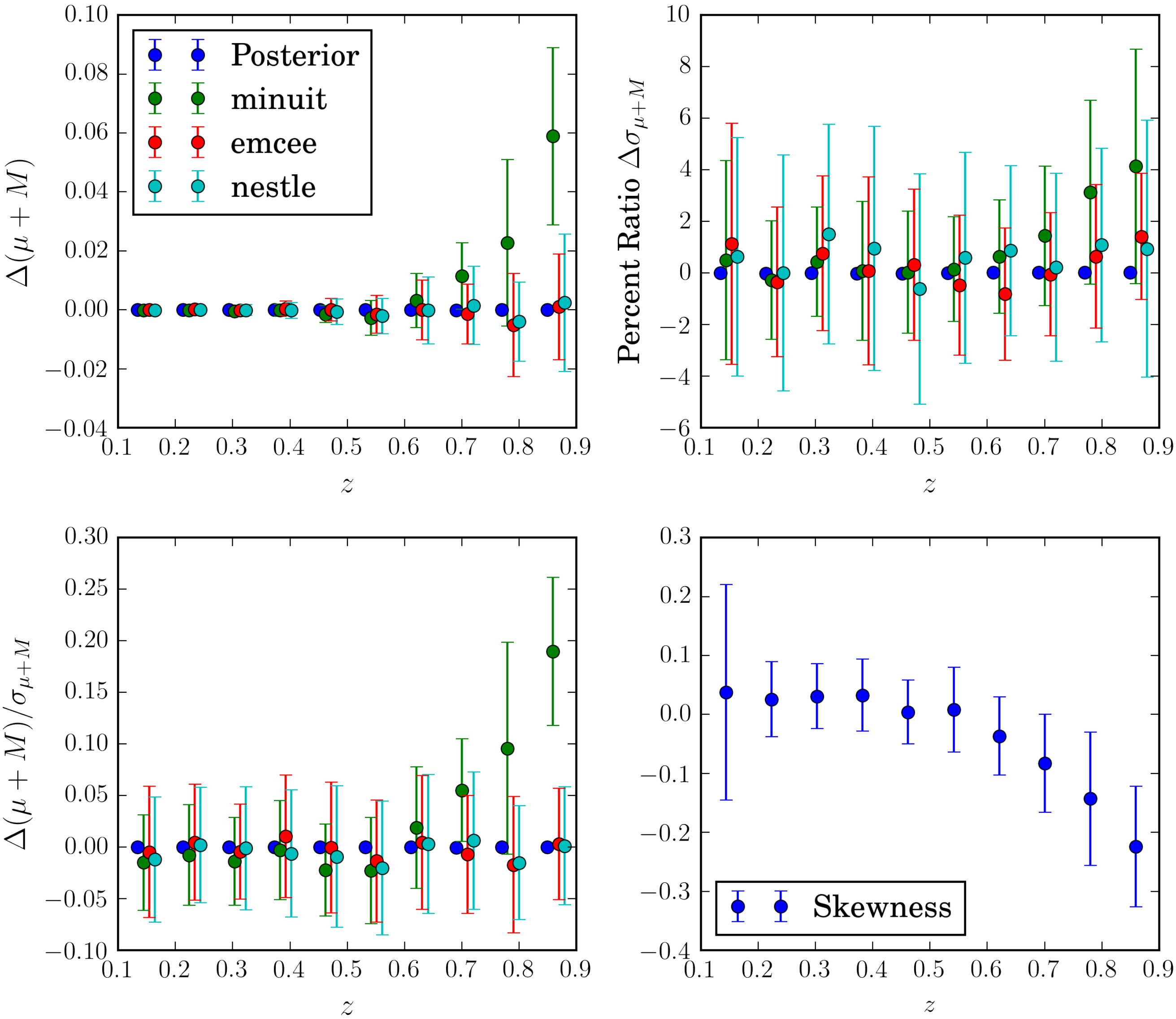

Figure 1. The mean bias in fill out caption when updated figure

Skilling J., 2004, in Fischer R., Preuss R., Toussaint U. V., eds, American Institute of Physics Conference Series Vol. 735, American Institute of Physics Conference Series. pp 395–405, doi:10.1063/1.1835238

Sullivan M., et al., 2011, ApJ, 737, 102

Suzuki N., et al., 2012, ApJ, 746, 85

The Dark Energy Survey Collaboration 2005, ArXiv Astrophysics e-prints,

This paper has been typeset from a TEX/LATEX file prepared by the author.

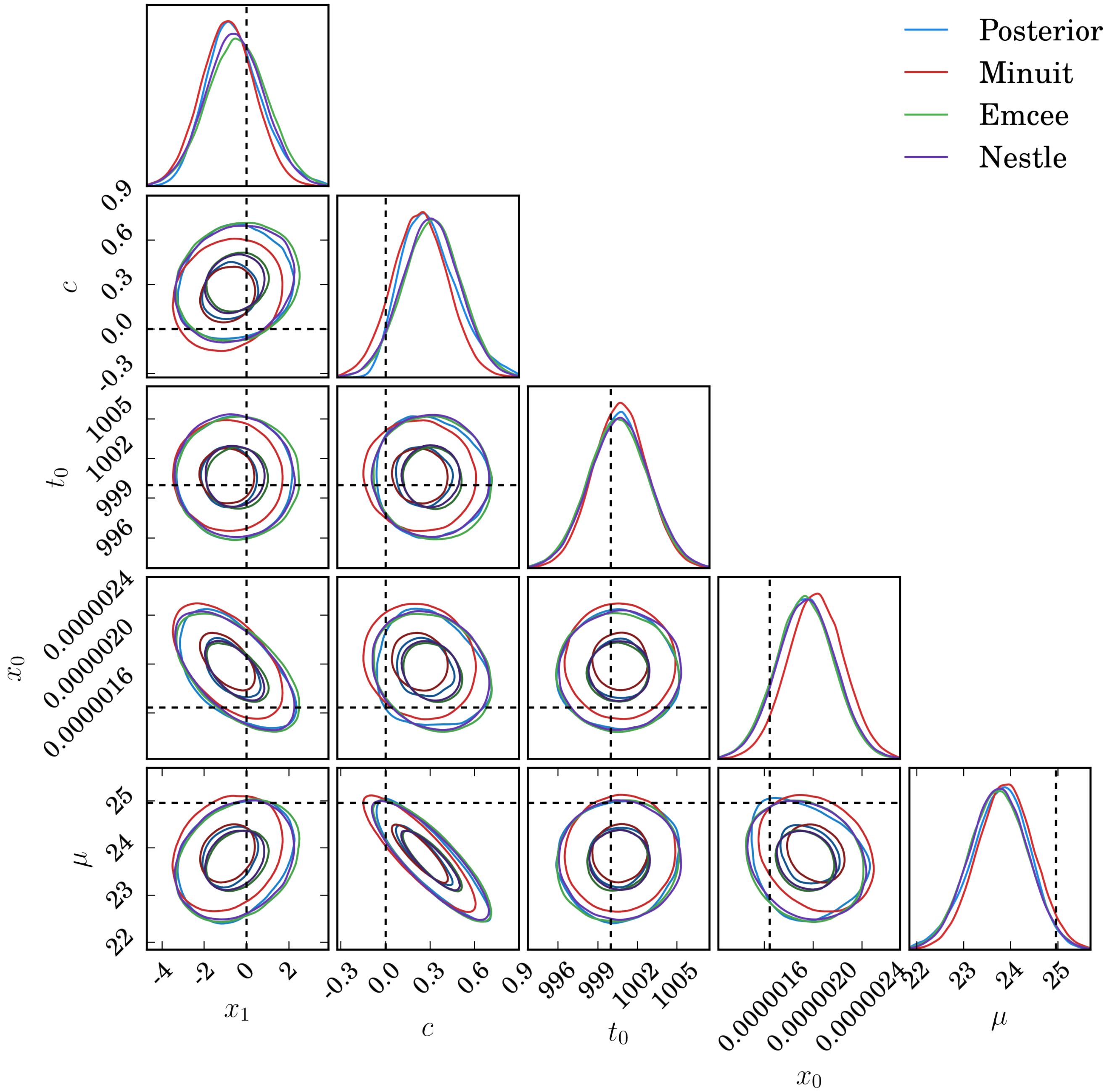

Figure 2. Fill out caption in updated figure

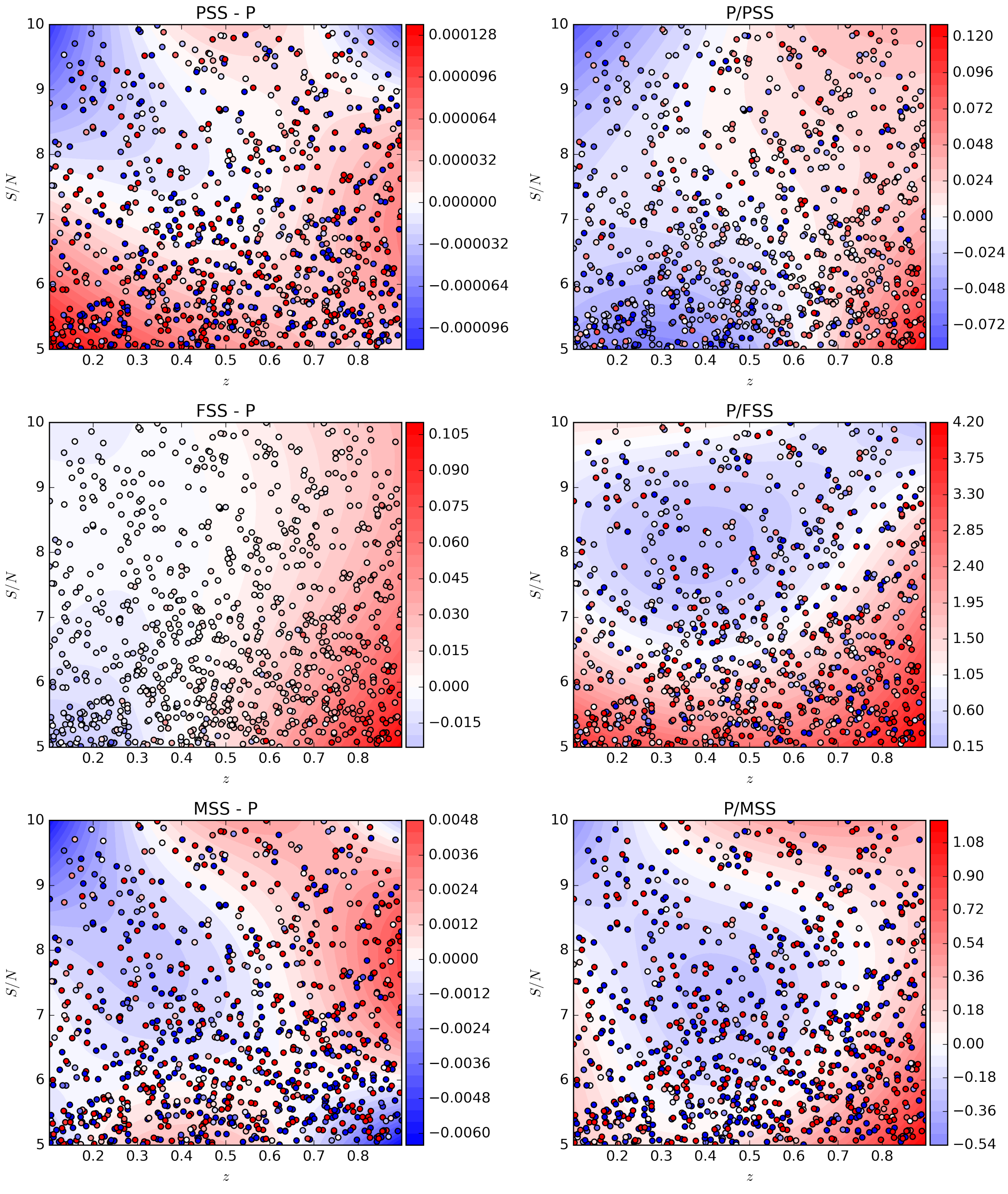

Figure 3. Fill out caption in updated figure