Quantifying Suspiciousness Within Correlated Data Sets

Pablo Lemos,1⋆ Fabian Köhlinger,2 Will Handley,3,4 Benjamin Joachimi,1 Lorne Whiteway1 and Ofer Lahav1

1Department of Physics and Astronomy, University College London, Gower Street, London, WC1E 6BT, UK
2Kavli IPMU (WPI), UTIAS, The University of Tokyo, Kashiwa, Chiba 277-8583, Japan
3Astrophysics Group, Cavendish Laboratory, J.J.Thomson Avenue, Cambridge, CB3 0HE, UK
4Kavli Institute for Cosmology, Madingley Road, Cambridge, CB3 0HA, UK

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT
We propose a principled Bayesian method for quantifying tension between correlated datasets with wide uninformative parameter priors. This is achieved by extending the Suspiciousness statistic, which is insensitive to priors. Our method uses global summary statistics, and as such it can be used as a diagnostic for internal consistency. We show how our approach can be combined with methods that use parameter space and data space to identify the existing internal discrepancies. As an example, we use it to test the internal consistency of the KiDS-450 data in 4 photometric redshift bins, and to recover controlled internal discrepancies in simulated KiDS data. We propose this as a diagnostic of internal consistency for present and future cosmological surveys, and as a tension metric for data sets that have non-negligible correlation, such as LSST and Euclid.

Key words: keyword1 – keyword2 – keyword3

1 INTRODUCTION
Quantifying consistency between different data sets has become one of the main challenges in modern cosmology. With the increasing number of methods and surveys measuring different properties of the Universe, it is crucial to develop appropriate statistical tools to compare and combine these data sets. This is important for two reasons. First, differences in cosmological constraints from different data sets, i.e. ‘tensions’, could indicate unaccounted-for systematic errors in one or both data sets, or could indicate that the underlying theoretical model (for example \( \Lambda \)CDM) is not sufficient to explain both data sets. Second, because the combination of different data sets can provide powerful cosmological constraints by breaking degeneracies existing in single data sets, but data sets can only be combined meaningfully if they are in agreement.

Cosmological tensions and their correct quantification are now particularly important, given that while all existing data sets match the \( \Lambda \)CDM model of cosmology individually, there are two main disagreements between data sets which could be a hint of beyond-\( \Lambda \)CDM physics. This belief is reinforced by the fact the tensions are between high redshift measurements of the Cosmic Microwave Background (CMB) by the Planck satellite (Planck Collaboration 2018a,b), and low redshift measurements of the Hubble constant (Riess et al. 2018, 2019) and of the growth of structure measured by galaxy clustering and weak gravitational lensing by several surveys such as the Canada France-Hawaii Telescope Lensing Survey (CFHTLenS; Heymans et al. 2012, 2013; Joudaki et al. 2017), the Kilo Degree Survey (KiDS, Köhlinger et al. 2017), and the Dark Energy Survey (DES, Troxel et al. 2018a; Abbott et al. 2018). There is a recent and extensive literature on quantifying tension in the context of cosmology (for a review of some methods see Charnock et al. 2017). Perhaps the most frequently used is the Bayes ratio \( R \) introduced in Marshall et al. (2006), which has the advantages of using exclusively Bayesian quantities and of being parameterization independent. However, this method has the disadvantage of being proportional to the prior volume (which can hide existing tensions when broad priors are chosen with the goal of being uninformative). Other approaches based on differences in the best fit parameters were introduced by Lin & Ishak (2017b,a); Raveri & Hu (2019); Adhikari & Huterer (2019), Kunz et al. (2006); Seehars et al. (2014); Grandis et al. (2016b,a); Nicola et al. (2019) used methods based on the Bayesian information measured by the Kullback Leibler (KL) Divergence (Kullback & Leibler 1951). One of the latest methods suggested was the Bayesian ‘Suspiciousness’ in-

⋆ E-mail: pablo.lemos.18@ucl.ac.uk
TENSION IN CORRELATED DATASETS

The goal of this section is to extend the tension metrics and Bayesian model dimensionality introduced in \cite{Handley2019} to the case of correlated data sets. To derive the $R$ statistic, \cite{Marshall2006} propose comparing two hypotheses

- $H_0$: There exists one common set of parameters that describes both data sets.
- $H_1$: There exist two sets of parameters, one for each data set.

Bayes’ theorem is then used to calculate the ratio of degrees of belief in each hypothesis in light of the data $D$

$$
\frac{P(H_0|D)}{P(H_1|D)} = \frac{P(D|H_0)}{P(D|H_1)} \frac{P(H_0)}{P(H_1)} \propto Z_0 \equiv R,
$$

where $R$ is the Bayes ratio and $Z_0$ is the Bayesian evidence.

In the uncorrelated case, the additivity of the KL divergence implies that

$$
D = \int d\theta \log \left( \frac{P}{\Pi} \right).
$$

The Bayes ratio approach using $R$ to quantify tension for correlated data sets was used in K19, but, as they show, $R$ depends on the prior volume. This is the same prior volume dependence that is extensively discussed in H19 for uncorrelated data sets. While this would not be a problem for well-motivated priors, it means that we cannot rely on Bayesian evidence ratios obtained with priors that are purposefully chosen to be broad (with the intention of being uninformative).

To address these concerns about prior volume, we can instead use the Suspiciousness $S$ introduced in H19, which can be understood as the value of $R$ that corresponds to the narrowest possible priors that do not significantly alter the shape of the posteriors. The natural logarithm of the Suspiciousness is given by

$$
\log S = \log R - \log I.
$$

The information $I$ quantifies the $a$-priori probability that the data sets would match given the prior range. The larger the prior range relative to the posterior constraints, the lower the probability that the constraints will be consistent. For correlated data sets the natural logarithm of the information is given by

$$
\log I = D_1 - D_0,
$$

In the uncorrelated case, the additivity of the KL divergence implies that $D_1 = D(A) + D(B)$ and $D_0 = D(A, B)$, which recovers the methodology of H19.

GAUSSIAN EXAMPLE

To illustrate the formalism, we consider the example of multivariate Gaussian posterior distributions (Fig. 1). Let $A$ and $B$ be two data sets that can constrain the same set $\theta$ of $d$ parameters. Each data set $A$ or $B$ individually constrains one set of parameters $\theta_A$ or $\theta_B$ giving a posterior with parameter mean and covariance of $(\mu_A, \Sigma_A)$ or $(\mu_B, \Sigma_B)$ respectively.

Assume now that the two data sets are correlated. If we combine the two data sets and use hypothesis $H_1$, then the...
As discussed in H19, the evidence and KL divergence for a single parameter. The plot shows the two posteriors, and their combination in the uncorrelated and correlated case, for \( \sigma^2_A = \sigma^2_B = 0.1 \), and \( V_{th} = 10 \).

The likelihood for hypothesis \( H_0 \) can be found by setting \( \theta_A = \theta_B = \theta \) in Eq. (7). Notationally it is helpful to define the precision matrix \( \Lambda \) in block form via

\[
\Sigma_1 = \begin{bmatrix} \Sigma_A & \Sigma_{AB} \\ \Sigma_{AB}^T & \Sigma_B \end{bmatrix}, \quad \Sigma_1^{-1} = \Lambda = \begin{bmatrix} \Lambda_A & \Lambda_X \\ \Lambda_X^T & \Lambda_B \end{bmatrix},
\]

where the different blocks in \( \Lambda \) are related to those in \( \Sigma_1 \) by Schur’s complement. (Woodbury 1950; Zhang 2005). After some effort manipulating matrix expressions (see Appendix), we find

\[
\log L_0 = \log L_0^{\text{max}} - \frac{1}{2} (\theta - \mu_0)^T \Sigma_1^{-1} (\theta - \mu_0),
\]

(9)

\[
\log L_{\text{max}} = \log L_{\text{max}}^{\text{max}} - \frac{1}{2} (\mu_A - \mu_B)^T \Sigma_{\mu\mu}^{-1} (\mu_A - \mu_B).
\]

(10)

\[
\Sigma_{\mu\mu} = \Sigma_A - \Sigma_X - \Sigma_{AB}^T + \Sigma_B.
\]

(11)

The number of parameters constrained simultaneously by both datasets is given by subtracting the number of parameters constrained by the combination of both from the number of parameters constrained by each dataset separately, in a similar manner to Handley & Lemos (2019a) for uncorrelated datasets.

While the results of this section have been obtained for Gaussian likelihoods, they can be applied to more general posteriors. This claim is supported by the fact that the Suspicousness is invariant under coordinate transformations, such as Box-Cox transformations (Box & Cox 1964; Joachimi & Taylor 2011; Schuhmann et al. 2016) which âĂĲGaussianiseâ€‰the posterior. Furthermore, H19 showed that the Suspicousness recovers the intuitively correct answers for cosmological examples, which are a somewhat non-Gaussian, particularly in the nuisance parameters.

4 CONNECTION TO OTHER INTERNAL CONSISTENCY TESTS

K19 proposed three tiers of internal consistency between data sets. We have previously discussed how the Bayesian Suspicousness serves as an alternative to tier one (the Bayes ratio). In this section we discuss the connection between Suspicousness and tiers two and three. We argue that, while these three methods are different in their implementation, the underlying quantities being calculated are surprisingly similar.

The effective number of dimensions constrained by both data sets \( d \) is given by the Bayesian Model Dimensionality (BMD) introduced in Handley & Lemos (2019a). In particular, Handley & Lemos (2019a) show that \( d \) is exactly the same as the number of dimensions for a Gaussian likelihood. For correlated data sets, \( d \) is given by

\[
d = d_1 - d_0,
\]

(20)

where the number of parameters constrained simultaneously by both datasets is given by subtracting the number of parameters constrained by the combination of both from the number of parameters constrained by each dataset separately, in a similar manner to Handley & Lemos (2019a) for uncorrelated datasets.
after calculating the posterior distributions under hypothesis $H_1$ (the data sets are described by different sets of parameters), we can derive the posterior distribution for the differences between the parameters describing each data set. K19 then proposes finding the fraction of samples with a value of the posterior smaller than the value in the origin (which is the point of perfect agreement). This can be seen as an extension of tension metrics based on parameter shifts, such as those introduced in Raveri & Hu (2019); Adhikari & Huterer (2019), to the case of correlated data sets. From the result of Sec. 3 it is easy to see how our method is connected to this. For Gaussian posterior distributions, we show that $d - 2\log S$ is $\chi^2$ distributed, from which we can calculate the tension probability. While the methodology is different, the quantity being calculated is the same in the case of this Gaussian toy model. In fact, one could argue that the ‘more’ interpretation presented in K19 should be dependent on dimensionality, which could be calculated using the BMD (Handley & Lemos 2019a), and then the calculation of ‘more’ would become similar to Eq. (19). This is shown in Fig. 2, where we compare the results of our method and parameter differences for different one dimensional Gaussian distributions, and confirm that the results are the same.

Tier three of K19 consists of using the PPD. This technique was introduced in Gelman et al. (1996), and has been applied to cosmology in recent problems (e.g. Feeney et al. 2019; Abbott et al. 2019). When using the PPD, we calculate the probability of data $D_A$ conditional on data $D_B$ and the underlying model $M$ in the following way:

$$P(D_A|D_B, M) = \int d\theta \ P(D_A|\theta, M)P(\theta|D_B, M). \tag{21}$$

In practice this means that from samples of the posterior for $D_B$, we calculate realizations of the likelihood for $D_A$. The most challenging part of using the PPD is calibrating the distribution. $P(D_A|D_B, M)$ is a probability density which is challenging to normalize. The authors of K19, for example, built a ‘Translated Probability Distribution’ from a predicted data vector, with the goal of calibrating the PPD. However, there is an alternative way of interpreting this number, which is by taking the ratio $P(D_A|D_B, M)/P(D_A|M)$. This quantity is unitless, and can therefore be interpreted as a probability ratio. More importantly, as discussed in Handley & Lemos (2019b), this ratio is equal to the Bayes ratio $R$. In this way, we can see tiers one and three of K19 as being elements of the same calculation, but using different calibrations of the PPD. If both methods are correct, then they should produce similar results.

5 KIDS-450 DATA

In this section, we test the methods introduced in Sec. 2 on the KiDS-450 data 1 (Hildebrandt et al. 2017). KiDS uses the correlations in the shapes of galaxy images to measure weak gravitational lensing caused by large scale structure (for reviews, see Bartelmann & Schneider 2001; Kilbinger 2015). Their measurements span a redshift range $z = [0.1, 0.9]$; this range is divided into four redshift bins

---

1 See http://kids.strw.leidenuniv.nl/sciencedata.php.
bins. The data therefore consist of 130 simple statistical tests. In particular, bins 3 and 4 were found of the discrepant bins in E18. We use the code that involve redshift bin 3 to all others, since bin 3 was one and K19.

The data from (Hildebrandt et al. 2017) used by both E18 levels. However, in the remainder of this work, we will use reducing the discrepancies found in E18 to below significant however, we will use the data from Hildebrandt et al. (2017), proposing an accurate new approach adopted with each of the definitions of the angular tomographic model for the survey-boundary effects. This correction in- E18, Troxel et al. (2018b) improved the shot noise model sensitive to the overall goodness of fit of the data. Following this work, K19 performed their three tier consistency tests on the same data. Using these statistics, K19 found that the significance of these inconsis- tency of the data sets being consistent is \( p = 0.0674 \pm 0.0062 \), corresponding to a significance 1.83\( \sigma \). As expected, in this case the Suspiciousness provides a far more accurate assessment of tension than does \( R \), which obtains ‘very strong’ evidence for agreement according to the Jeffreys scale used in K19. However, the significance is also lower than in E18, which is a reflection on the dependence on overall goodness of fit of the statistic used in that paper. In other words, our method quantifies internal consist- ing of using the real KiDS-450 data, we use mock data vec- tors, in which the source redshift distribution is shifted by \( d_{z3} = 0.15, 0.20 \). Our results are shown in Fig. 3 and Tab. 1. Again, we see how the Bayes ratio \( R \) can hide tensions because of the width of the priors, while the Suspiciousness shows clear tension for both shifts of the redshift distribu- tion. For example, \( d_{z3} = 0.15 \) corresponds to ‘No Evidence’ in \( R \), while for the same shift \( S \) yields a more than 3\( \sigma \) discrep- any, indicating a strong tension. This highlights the value of the Suspiciousness for internal consistency tests in large data sets. The table also shows how the number of sigma is zero for \( d_{z3} = 0 \). This is because in this case we are comparing data sets whose posteriors overlap perfectly. We should generally also be suspicious of cases where the tension probability is very close to one (such as this one), as they indicate that the agreement is ‘too good’.

### 6 CONCLUSIONS

In this paper, we extended the novel tension metric intro- duced in Handley & Lemos (2019b) to the case of correlated data sets. This takes the role of “tier 1” consistency tests intro- duced in Köhlinger et al. (2019), and provides a measure of consistency for correlated data sets that uses Bayesian quantiles, is parameterization independent, and can be used for the case of wide, uninformative priors since it does not depend on the prior volume (in contrast to the Bayes ra-

| \( d_{z3} \) | \( \log R \) | Interpretation | \( \log S \) | \( d \) | \( p_t \) | \( N_o \) |
|-------|-------|-------------|-------|-----|-----|-----|
| 0     | 7.69 ± 0.15 | Decisive Agreement | 1.488 ± 0.049 | 2.68 ± 0.17 | 0.99995 ± 0.00053 | 0 |
| 0.15  | −0.04 ± 0.14 | Neutral | −5.735 ± 0.052 | 2.75 ± 0.18 | 0.00203 ± 0.00028 | 3.089 ± 0.041 |
| 0.20  | −6.18 ± 0.14 | Decisive Tension | −11.383 ± 0.051 | 2.60 ± 0.17 | 7.6 · 10^{-6} ± 1.4 · 10^{-6} | 4.478 ± 0.039 |

Table 1. Comparison of evidence \( \log R \) (interpreted with the Jeffreys scale used in K19) and Suspiciousness \( \log S \) for some redshift shifts in the sensitivity analysis. The last three columns show the Bayesian Model Dimensionality \( d \), the tension probability \( p_t \) and the corresponding number of sigma, calculated as \( N_o = \sqrt{2 \text{Erf}^{-1}(1 - p_t)} \).
We applied this formalism to the case of KiDS-450 data from Hildebrandt et al. (2017), focusing on tests of tomographic redshift bin 3 vs the rest of the data (bins 1, 2 and 4). We find that the 2.60σ tension detected in Efstathiou & Lemos (2018) reduces to 1.83σ. We interpret this as a difference in the tension metric used: the methods used in Efstathiou & Lemos (2018) depend on overall goodness of fit, while those used in this work do not. This is a desirable quality, as it allows us to identify the origin of potential problems in the data, but it stresses the need to combine the Suspiciousness with a goodness of fit test.

This method can generally be used for correlated data sets where uninformative priors are used. In particular, it can be used for internal consistency tests of cosmological surveys (such as DES and KiDS), and also to quantify tension between data sets with a non-negligible covariance (such as Euclid and LSST). We conclude that due to its Bayesian nature and intuitive interpretation, this methods serves as a perfect diagnostic of internal consistency, and we encourage present and future cosmological surveys to use it.

ACKNOWLEDGEMENTS

Based on data products from observations made with ESO Telescopes at the La Silla Paranal Observatory in the programme IDs 177.A-3016, 177.A-3017 and 177.A-3018, and on data products produced by Target/OmegaCEN, INAF-OACN, INAF-OAPD and the KiDS production team, on behalf of the KiDS consortium. We would like to thank Catherine Heymans for useful comments.

PL & OL acknowledge STFC Consolidated Grant ST/R000476/1. FK acknowledges support from the World Premier International Research Center Initiative (WPI), MEXT, Japan. WJH thanks Gonville & Caius College for their continuing support via a Research Fellowship.

APPENDIX

We use Eq. (7) to derive Eq. (9) (with the associated definitions in Eq. (10) through Eq. (13)) using the general ‘complete the square’ formula

$$Q^θ - K)Σ^{-1}(θ^θ - K) = (θ - C)^T(Q^TΣ^{-1}Q)(θ - C) + K^TΣ^{-1}(K - QC) \tag{22}$$

where $C = (Q^TΣ^{-1}Q)^{-1}Q^TΣ^{-1}K$. In Eq. (22) set

$$Q = \begin{bmatrix} I & I \end{bmatrix} K = \begin{bmatrix} μ_A \\ μ_B \end{bmatrix}$$

and $Σ = Σ_1$ (so $Σ^{-1} = Λ$). \tag{23}

Also define \(P = [I \quad -I]\) (so \(PK = μ_A - μ_B\)) and set

\[
M^{-1} = Σ - Σ_X - Σ_X^T + Σ_B = PΣP^T
\]

and

\[
L^{-1} = Λ_A + Λ_X + Λ_X^T + Λ_B = Q^TΛQ.
\]

The covariance matrix in the RHS of Eq. (22) is

\[
Q^TΣ^{-1}Q = Q^TΛQ = L^{-1};
\]

this will be $Σ^{-1}$ in Eq. (9). For the constant term on the RHS of Eq. (22), begin with

\[
0 = PQ = (PΣ(ΛQ) = (Σ - Σ_X)(Λ_A + Λ_X) + (Σ_X - Σ_B)(Λ_X^T + Λ_B), \text{ so}
\]

\[
(Λ_B + Λ_X)(Λ_A + Λ_X)^{-1} = (Σ_B - Σ_X)^{-1}(Σ_A - Σ_X)
\]

\[
L^{-1}(Λ_A + Λ_X)^{-1} - I = (Σ_B - Σ_X)^{-1}M^{-1} - I
\]

\[
(Λ_A + Λ_X)L = M(Σ_B - Σ_X).
\]

This establishes the upper-right block of the identity

\[
ΛQLQ^T + P^TMΣ_{diag} = I
\]

and the other blocks can be established similarly. Thus

\[
K^TΣ^{-1}(K - QC) = K^T(1 - ΛQLQ^T)ΛK = (PK)^TMPK = (μ_A - μ_B)^TΣ^{-1}_{diag}(μ_A - μ_B);
\]

(28)

this will contribute to log $L_{max}$ in Eq. (9). Finally the mean used in the RHS of Eq. (22) is

\[
C = (Q^TΣ^{-1}Q)^{-1}Q^TΣ^{-1}K = LQΣ^{-1}K = Σ_0Q^TΛK
\]

\[
= Σ_0((Λ_A + Λ_X)μ_A + (Λ_X + Λ_B)μ_B);
\]

(29)

this will be $μ_0$ in Eq. (9).

See Fig. 4 for a diagram illustrating the behaviour of the likelihood function under the two hypotheses.

REFERENCES

Abbott T. M. C., et al., 2018, Phys. Rev. D, 98, 043526

Abbott T. M. C., et al., 2019, Phys. Rev. D, 100, 023541

Adhikari S., Huterer D., 2019, JCAP, 2019, 036

Asgari M., et al., 2019, A&A, 624, A134

Audren B., Lesgourgues J., Benabed K., Prunet S., 2013, JCAP, 2013, 122

Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291

Box G. E. P., Cox D. R., 1964, Journal of the Royal Statistical Society. Series B (Methodological), 26, 211

Charnock T., Battye R. A., Moss A., 2017, Phys. Rev. D, 95, 123535

Efstathiou G., Lehoucq R., 2018, MNRAS, 476, 151

Feeney S. M., Peiris H. V., Williamson A. R., Nissimke S. M., Mortlock D. J., Asling J., Scolnic D., 2019, Phys. Rev. Lett., 122, 061105

Gelman A., li Meng X., Stern H., 1996, Statistica Sinica, pp 733–807

Grancini S., Papetti D., Saro A., Mohr J. J., Dietrich J. P., 2016a, MNRAS, 463, 1416

Grancini S., Seehars S., Refregier A., Amara A., Nicola A., 2016b, JCAP, 2016, 034

Handley W., 2019, JOSS, 4

Handley W., Lemos P., 2019a, Phys. Rev. D, 100, 023512

MNRAS 000, 000–000 (2019)
Figure 4. Inference of mean(s) from two correlated data observations $d_A$ and $d_B$. The blue contours show (under hypothesis $H_1$ i.e. separate means) the likelihood (assumed Gaussian with covariance $\Sigma$) of the mean values $(\mu_A, \mu_B)$; the maximum likelihood (and with flat priors the mode of the posterior) for these means is the black circle at $(d_A, d_B)$. Under hypothesis $H_0$ (i.e. one common mean) the likelihood is restricted to the red dashed line $\mu_A = \mu_B$ where it will be Gaussian with covariance $\Sigma_0$; here the maximum likelihood (and hence posterior mode) is at the red diamond $(\mu_0, \mu_0)$.

Handley W., Lemos P., 2019b, Phys. Rev. D, 100, 043504
Heymans C., et al., 2012, MNRAS, 427, 146
Heymans C., et al., 2013, MNRAS, 432, 2433
Hildebrandt H., et al., 2017, MNRAS, 465, 1454
Hildebrandt H., et al., 2018, arXiv e-prints, p. arXiv:1812.06076
Ivezić Z., et al., 2019, ApJ, 873, 111
Jenkins C. R., Peacock J. A., 2011, MNRAS, 413, 2895
Joachimi B., Taylor A. N., 2011, MNRAS, 416, 1010
Joachimi B., Lemos P., Handley W., Köhlinger F., 2019, In preparation
Joudaki S., et al., 2017, MNRAS, 465, 2033
Kilbinger M., 2015, Reports on Progress in Physics, 78, 086901
Köhlinger F., et al., 2017, MNRAS, 471, 4412
Köhlinger F., Joachimi B., Asgari M., Viola M., Joudaki S., Tröster T., 2019, MNRAS, 484, 3126
Kullback S., Leibler R. A., 1951, Ann. Math. Statist., 22, 79
Kunz M., Trotta R., Parkinson D. R., 2006, Phys. Rev. D, 74, 023503
Laureijs R., et al., 2011, arXiv e-prints, p. arXiv:1110.3193
Lin W., Ishak M., 2017b, Phys. Rev. D, 96, 083532
Lin W., Ishak M., 2017a, Phys. Rev. D, 95, 023522
Marshall P., Rajguru N., Slosar A., 2006, Phys. Rev. D, 73, 067302
Nicola A., Amara A., Refregier A., 2019, JCAP, 2019, 011
Planck Collaboration 2018a, arXiv e-prints, p. arXiv:1802.01854
Planck Collaboration 2018b, arXiv e-prints, p. arXiv:1807.06209
Raveri M., Hu W., 2019, Phys. Rev. D, 99, 043506
Rhodes J., et al., 2017, ApJS, 233, 21
Riess A. G., et al., 2018, ApJ, 855, 136
Riess A. G., Casertano S., Yuan W., Macri L. M., Scolnic D., 2019, ApJ, 876, 85
Schuhmann R. L., Joachimi B., Peiris H. V., 2016, MNRAS, 459, 1916
Schuhmann R. L., Heymans C., Zuntz J., 2019, arXiv e-prints, p. arXiv:1901.08586

MNRAS 000, 000–000 (2019)