Comparing E_G with a full combined analysis of its composite probes

Danielle Leonard

I.   INTRODUCTION AND MOTIVATION

The E_G statistic is a commonly-employed test of gravity. In its typical form, it combines a two-point measurement of galaxy clustering with a measurement of either galaxy-galaxy lensing or the cross-correlation of CMB lensing with galaxy positions, and with a measurement of the linear growth rate of structure from redshift-space distortions.

E_G is defined as a ratio of these probes such that on scales at which the galaxy bias is linear, it cancels from the ratio. This independence from galaxy bias is one of the main selling points of the statistic. However, E_G also has downsides. Combining the covariance of the constituent probes of E_G is not straightforward due to its nature as a ratio of noisy quantities. Furthermore, it is not clear that the E_G statistic measured in practice provides the simple, bias-independent, model-indepenent test of gravity for which it was introduced.

Given these drawbacks, one might ask the question of whether computing E_G from a given series of datasets actually provides any additional information about deviations from GR that cannot be obtained in a full, combined-probes likelihood analysis of its constituent probes. That is, in which scenarios is the modified-gravity-related *information content* of E_G greater than, less than, or comparable to, that of a full combined analysis?

II.   THEORY AND SETUP

In this section we first discuss the combined-probe analysis scenarios we will consider, and provide theoretical expressions in the fiducial GR scenario for the various relevant observables (including those which are required to compute E_G itself). We then introduce the quasistatic parameterisation of deviations from GR and detail how the required observables are modeled within this framework. We finally describe how we calculate the observable covariance matrices in each analysis scenario.

A.   Observational probes

We consider the observational definition of E_G in real space, as originally presented in [7]:

$$E_G(r_p) = \frac{\Upsilon_{gm}(r_p)}{\beta \Upsilon_{gg}(r_p)} \tag{1}$$

We will compare E_G as defined in this way to:

- A **baseline combined-probes** analysis of $\Upsilon_{gm}(r_p)$, $\Upsilon_{gg}(r_p)$, and $\beta = f/b$. This represents the minimal joint-probes case which is comparible to a measurement of E_G.

- A more **extended combined-probes** analysis of $\xi_+(r_p)$, $\xi_-(r_p)$, $w_{\rm gg}(r_p)$, $\gamma_t(r_p)$, and β. This case represents a typical '3x2pt' analysis as performed in e.g. [**?**], plus β.

We now provide theoretical expressions for the observables required in each observational scenario.

Baseline   For both $E_G(r_p)$ and the baseline joint probe analysis, we require predictions for $\Upsilon_{gm}(r_p)$, $\Upsilon_{gg}(r_p)$ and β.

$\Upsilon_{gm}(r_p)$ and $\Upsilon_{gg}(r_p)$ are annular differential surface density observables for galaxy-galaxy lensing and projected galaxy clustering, introduced in [2]. $\Upsilon_{gm}(r_p)$ is defined as:

$$\Upsilon_{gm}(r_p) = \Delta\Sigma_{gm}(r_p) - \left(\frac{r_p^0}{r_p}\right)^2 \Delta\Sigma_{gm}(r_p^0) \tag{2}$$

where r_p^0 is a projected radial scale selected to remove uncertain small scale information, and $\Delta\Sigma_{gm}(r_p)$ is the differential surface mass density at projected radius r_p. It is defined as

$$\Delta\Sigma_{gm}(r_p) = \bar{\Sigma}_{gm}(< r_p) - \Sigma_{gm}(r_p) \tag{3}$$

where $\Sigma_{gm}(r_p)$ is the projected surface mass density and $\bar{\Sigma}_{gm}(< r_p)$ is the same quantity averaged over scales less than or equal to r_p. $\Sigma_{gm}(r_p)$ can be calculated theoretically via

$$\Sigma_{gm}(r_p) = \rho_{\mathrm{m}} \int d\Pi \left(1 + \xi_{gm}(r_p, \Pi)\right) \tag{4}$$

where ρ_{m} is the matter density of the universe in comoving coordinates, ξ_{gm} is the 3D galaxy-matter correlation function, and Π is the radial separation from the lens galaxy along the line of sight.

Equation 2, 3, and 4 offer a means to calculate $\Upsilon_{gm}(r_p)$ theoretically. However, in practice, $\Upsilon_{gm}(r_p)$ is measured observationally via galaxy-galaxy lensing: the correlation of background galaxy shapes with foreground galaxy position. In general relativity, $\Delta\Sigma(r_p)$ is related to the tangential shear of background galaxies about lens galaxies, called $\gamma_t(r_p)$, via

$$\Delta\Sigma(r_p) = \Sigma_{\mathrm{c}}\gamma_t(r_p) \tag{5}$$

for a single lens-source pair, where Σ_{c} is given by

$$\Sigma_{\mathrm{c}} = \frac{c^2}{4\pi G}\frac{\chi_s}{(\chi_s - \chi_l)\chi_l(1 + z_l)} = \frac{c^2}{4\pi G}\frac{D_s}{D_{ls}D_l(1 + z_l)^2} \tag{6}$$

where D_s and D_l are angular diameter distances to the source and lens galaxy respectively and we use comoving coordinates. Note that D_{ls} is defined such that this expression is true, i.e. $D_{ls} \neq (D_s - D_l)$.

Under modifications to gravity, however, the relationship between $\Delta\Sigma$ and γ_t given in equation 5 no longer holds. This is due to the fact that not only the clustering of matter, but also the lensing potential itself may be changed under modifications to GR. We therefore introduce a quantity $\widetilde{\Delta\Sigma}$ which we define as the measured quantity obtained via measuring tangential shear and and multiplying by averaged Σ_{c} in the typical way. Thus, in modified gravity, $\widetilde{\Delta\Sigma}$ may not correspond to the true physical projected differential surface mass density and acts instead as an analogous measured quantity. We also similarly define $\tilde{\Upsilon}_{gm}$, obtained by using $\widetilde{\Delta\Sigma}$ in equation 2.

$\tilde{\Upsilon}_{gm}(r_p)$ can be expressed as [6]:

$$\tilde{\Upsilon}_{gm}(r_p) = \frac{\rho_c\Omega_M^0}{\bar{w}}\frac{4\pi G}{c^2}\int dz_l\frac{dN}{dz_l}\int d\Pi\int_{z_l+\Pi}^{z_s^{\max}}dz_s\frac{dN}{dz_s}\Sigma_{\mathrm{c}}(z_s, z_l)\tilde{w}(z_s, z_l)\frac{(\chi_l(z_l) + \Pi)(\chi_s - \chi_l(z_l) - \Pi)(z_l + 1)}{\chi_s}$$

$$\times\left[\frac{2}{r_p^2}\int_{r_p^0}^{r_p}r_p'dr_p'\xi_{gm}\left(\sqrt{(r_p')^2 + \Pi^2}, \bar{z}_l\right) - \xi_{gm}\left(\sqrt{r_p^2 + \Pi^2}, \bar{z}_l\right) + \left(\frac{r_p^0}{r_p}\right)^2\xi_{gm}\left(\sqrt{\left(r_p^0\right)^2 + \Pi^2}, \bar{z}_l\right)\right]. \tag{7}$$

where $\frac{dN}{dz_l}$ and $\frac{dN}{dz_s}$ are respectively the lens and source galaxy distributions. (We discuss requirements on photometric redshifts of sources below in Section ??.) $\tilde{w}(z_s, z_l)$ is the minimum variance weighting typically chosen for galaxy-galaxy lensing:

$$\tilde{w}(z_s, z_l) = \frac{1}{\Sigma_{\mathrm{c}}^2(z_s, z_l)\left(\sigma_\gamma^2 + \sigma_e(z_s)^2\right)}. \tag{8}$$

Here σ_γ is the shape noise and σ_e is the per-source-galaxy measurement noise. $\bar{w}$ is the normalization over these weights:

$$\bar{w} = \int z_l\frac{dN}{dz_l}\int z_s\frac{dN}{dz_s}\tilde{w}(z_s, z_l). \tag{9}$$

Equation 7 yields the same result as equations 2, 3, and 4 in GR.

$\Upsilon_{gg}(r_p)$ is the equivalent annular differential surface density for projected galaxy clustering, given by [2]

$$\Upsilon_{gg}(r_p) = \rho_c\left(\frac{2}{r_p^2}\int_{r_p^0}^{r_p}dr_p'r_p'w_{gg}(r_p') - w_{gg}(r_p) + \left(\frac{r_p^0}{r_p}\right)^2w_{gg}(r_p^0)\right). \tag{10}$$

For convenience, we define a function similar in character to $\Delta\Sigma_{gm}(r_p)$:

$$\Delta\Sigma_{gg}(r_p) = \rho_c\left(\bar{w}_{gg}(< r_p) - w_{gg}(r_p)\right) \tag{11}$$

such that

$$\Upsilon_{gg}(r_p) = \Delta\Sigma_{gg}(r_p) - \left(\frac{r_p^0}{r_p}\right)^2 \Delta\Sigma_{gg}(r_p^0). \tag{12}$$

The last component of the baseline joint probes analysis is $\beta(z)$, defined as

$$\beta(z) = \frac{f(z)}{b} \tag{13}$$

where b is once-again the large scale, constant galaxy bias and $f(z)$ is the linear growth rate of structure, defined as $f = \frac{d\ln D(a)}{d\ln a}$. In GR $f(z)$ can be caluculaed by solving:

$$\frac{d^2 D(a)}{d(\ln a)^2} + \left(1 + \frac{1}{aH(a)}\frac{d(aH(a))}{d\ln a}\right)\frac{dD(a)}{d\ln(a)} - \frac{3}{2}\Omega_{\rm M} D(a) = 0. \tag{14}$$

Extended:

B. μ, Σ parameterisation for modified gravity

We need to be able to compute correlation functions in modified gravity. Currently, the public MGcamb doesn't even have an option for computing power spectra with μ, Σ. Agnes has her own code for this - is this something we should get her on board for? Or ask Phil what the status of this is in CCL?

In the μ, Σ quasistatic parameterisation of modified gravity, we have

$$\frac{d^2 D(a)}{d(\ln a)^2} + \left(1 + \frac{1}{aH(a)}\frac{d(aH(a))}{d\ln a}\right)\frac{dD(a)}{d\ln(a)} - \frac{3}{2}\Omega_{\rm M}\mu(a)D(a) = 0. \tag{15}$$

C. Covariance matrices

1. For combined-probe analysis

To forecast modified-gravity information content for a combined analysis of the constituent probes of E_G, we require the full covariance matrix of a data vector consisting of $\Upsilon_{gm}(r_p)$, $\Upsilon_{gg}(r_p)$ and β, where the former two are measured in projected radial bins.

We now detail the calculation of each required type of covariance matrix term.

$\mathrm{Cov}(\Upsilon_{gm}(r_p), \Upsilon_{gm}(r'_p))$: As we have seen above $\Upsilon_{gm}(r_p)$ is given via a difference of terms involving $\Delta\Sigma_{gm}(r_p)$. We therefore begin with the covariance of $\Delta\Sigma_{gm}(r_p)$, which is given by (see, e.g., [5, 8])

$$\mathrm{Cov}\left[\Delta\Sigma_{gm}\left(r_p^i\right), \Delta\Sigma_{gm}\left(r_p^j\right)\right] = \frac{1}{4\pi f_{\rm sky}}\left(\overline{\Sigma_c^{-2}}\right)^{-1}\int\frac{ldl}{2\pi}J_2\left(l\frac{r_p}{\chi\left(z_l^{\rm eff}\right)}\right)J_2\left(l\frac{r'_p}{\chi\left(z_l^{\rm eff}\right)}\right)\left[\left(C_{g\kappa}^l\right)^2 + \left(C_{gg}^l + \frac{1}{n_l}\right)\left(C_{\kappa\kappa}^l + \frac{\sigma_\gamma^2}{n_{eff}}\right)\right],$$

where $r_p^{i,h}$ and $r_p^{i,l}$ represent the high and low edges of bin r_p^i respectively (and similarly for bin r_p^j), n_l is the surface of lens galaxies, n_{eff} is the effective surface density of sources for the source sample in question (both in galaxies per steradian), and $\overline{\Sigma_c^{-2}}$ is given by:

$$\overline{\Sigma_c^{-2}} = \int dz_l \frac{dN}{dz_l}\int dz_{ph}\frac{dN}{dz_{ph}}\Sigma_c^{-2}(z_l, z_{ph}). \tag{16}$$

Given then the covariance matrix for $\Delta\Sigma$ in r_p bins, we return to the definition of $\Upsilon_{gm}(r_p)$ given in equation 2 and find

$$\mathrm{Cov}(\Upsilon_{gm}(r_p), \Upsilon_{gm}(r'_p)) = \mathrm{Cov}(\Delta\Sigma_{gm}(r_p), \Delta\Sigma_{gm}(r'_p)) + \left(\frac{r_p^0}{r_p}\right)^2\left(\frac{r_p^0}{r'_p}\right)^2\mathrm{Cov}(\Delta\Sigma_{gm}(r_p^0), \Delta\Sigma_{gm}(r_p^0))$$

$$- \left(\frac{r_p^0}{r_p}\right)^2\mathrm{Cov}(\Delta\Sigma_{gm}(r_p^0), \Delta\Sigma_{gm}(r'_p)) - \left(\frac{r_p^0}{r'_p}\right)^2\mathrm{Cov}(\Delta\Sigma_{gm}(r_p), \Delta\Sigma_{gm}(r_p^0)). \tag{17}$$

Finally we average in bins:

$$\mathrm{Cov}(\Upsilon_{gm}(r_p^i), \Upsilon_{gm}(r_p^j)) = \frac{2}{\left(\left(r_p^{i,h}\right)^2 - \left(r_p^{i,l}\right)^2\right)} \frac{2}{\left(\left(r_p^{j,h}\right)^2 - \left(r_p^{j,l}\right)^2\right)} \int_{r_p^{i,l}}^{r_p^{i,h}} dr_p \, r_p \int_{r_p^{j,l}}^{r_p^{j,h}} dr_p' \, r_p' \mathrm{Cov}(\Upsilon_{gm}(r_p), \Upsilon_{gm}(r_p'))$$

(18)

Technically $\Delta\Sigma_{gm}(r_p^0)$ is computed via splining $\Delta\Sigma_{gm}(r_p)$, which may introduce some additional covariance contributions. We check that neglecting this is a good approximation by drawing realizations of $\Delta\Sigma_{gm}(r_p)$ from $\mathrm{Cov}(\Delta\Sigma_{gm}(r_p^i), \Delta\Sigma_{gm}(r_p^j))$, performing the typical splining procedure on these samples as described in e.g. [4], and computing the covariance of these samples empirically. We find that the two methods agree to a X percent level.

$\mathrm{Cov}(\Upsilon_{gg}(r_p), \Upsilon_{gg}(r_p'))$: Following the derivations of [8], we can obtain the covariance matrix of $\Delta\Sigma_{gg}(r_p)$ in projected radial bins:

$$\mathrm{Cov}\left[\Delta\Sigma_{gg}\left(r_p^i\right), \Delta\Sigma_{gg}\left(r_p^j\right)\right] = \frac{1}{2\pi f_{\mathrm{sky}}} \int \frac{l \, dl}{2\pi} J_2\left(l \frac{r_p}{\chi\left(z_l^{\mathrm{eff}}\right)}\right) J_2\left(l \frac{r_p'}{\chi\left(z_l^{\mathrm{eff}}\right)}\right) \left(C_{gg}^l + \frac{1}{n_l}\right)^2.$$

(19)

and in analogy with equation 17 above we find:

$$\mathrm{Cov}(\Upsilon_{gg}(r_p), \Upsilon_{gg}(r_p')) = \mathrm{Cov}(\Delta\Sigma_{gg}(r_p), \Delta\Sigma_{gg}(r_p')) + \left(\frac{r_p^0}{r_p}\right)^2 \left(\frac{r_p^0}{r_p'}\right)^2 \mathrm{Cov}(\Delta\Sigma_{gg}(r_p^0), \Delta\Sigma_{gg}(r_p^0))$$

$$- \left(\frac{r_p^0}{r_p}\right)^2 \mathrm{Cov}(\Delta\Sigma_{gg}(r_p^0), \Delta\Sigma_{gg}(r_p')) - \left(\frac{r_p^0}{r_p'}\right)^2 \mathrm{Cov}(\Delta\Sigma_{gg}(r_p), \Delta\Sigma_{gg}(r_p^0)).$$

(20)

We then obtain the binned covariance in direct analogy to equation 18.

We again verify that using the r_p which contains r_p^0 for relevant quantities in the above expression agrees with simulating the measurement of $\Upsilon_{gg}(r_p)$ many times and computing the covariance empirically; this time we find agreement to greater than Y percent.

$\mathrm{Cov}(\beta, \beta)$: Might want something better than the fitting function we are using now, it might not be relevant for all surveys.

$\mathrm{Cov}(\Upsilon_{gm}(r_p), \Upsilon_{gg}(r_p'))$: Following again the similar derivation of [8] we find:

$$\mathrm{Cov}\left[\Delta\Sigma_{gm}\left(r_p^i\right), \Delta\Sigma_{gg}\left(r_p^j\right)\right] = \frac{1}{2\pi f_{\mathrm{sky}}} \left(\overline{\Sigma_c^{-1}}\right)^{-1} \int \frac{l \, dl}{2\pi} J_2\left(l \frac{r_p}{\chi\left(z_l^{\mathrm{eff}}\right)}\right) J_2\left(l \frac{r_p'}{\chi\left(z_l^{\mathrm{eff}}\right)}\right) C_{g\kappa} \left(C_{gg}^l + \frac{1}{n_l}\right).$$

(21)

and

$$\mathrm{Cov}(\Upsilon_{gm}(r_p), \Upsilon_{gg}(r_p')) = \mathrm{Cov}(\Delta\Sigma_{gm}(r_p), \Delta\Sigma_{gg}(r_p')) + \left(\frac{r_p^0}{r_p}\right)^2 \left(\frac{r_p^0}{r_p'}\right)^2 \mathrm{Cov}(\Delta\Sigma_{gm}(r_p^0), \Delta\Sigma_{gg}(r_p^0))$$

$$- \left(\frac{r_p^0}{r_p}\right)^2 \mathrm{Cov}(\Delta\Sigma_{gm}(r_p^0), \Delta\Sigma_{gg}(r_p')) - \left(\frac{r_p^0}{r_p'}\right)^2 \mathrm{Cov}(\Delta\Sigma_{gm}(r_p), \Delta\Sigma_{gg}(r_p^0)).$$

(22)

Again, we average over r_p to get the binned quantity.

$\mathrm{Cov}(\Upsilon_{gm}(r_p), \beta)$ and $\mathrm{Cov}(\Upsilon_{gg}(r_p), \beta)$ Sukhdeep spoke to Uros, who claimed that the sample covariance between these quantities is zero, because β is sensitive to line of sight modes while Υ_{gm} and Υ_{gg} are sensitive to projected modes. However, I think we would still have some covariance due to correlated galaxy shot noise. How do we calculate this?

Checked with Elisabeth as to whether Cosmolike provides this functionality. It does not. She recommended getting these quantities using simulations and has found it painful trying to get these types of covariances in the past.

2. For $E_G(r_p)$

To estimate the modified-gravity-related information content of $E_G(r_p)$, we need to be able to calculate a covariance matrix of a data vector consisting of E_G in projected radial bins. I already have code set up to calculate matrices for $\text{Cov}(\Upsilon_{gm}(r_p), \Upsilon_{gm}(r'_p))$, $\text{Cov}(\Upsilon_{gg}(r_p), \Upsilon_{gg}(r'_p))$, and $\text{Cov}(\beta, \beta)$ separately, as made clear above. I also have code which uses these covariance matrices to draw samples of $\Upsilon_{gm}(r_p)$, $\Upsilon_{gg}(r_p)$ and β to then get $\text{Cov}(E_G(r_p), E_G(r'_p))$ using methods to estimate the covariance from a sample. But, this is only valid under the assumption that $\Upsilon_{gm}(r_p)$, $\Upsilon_{gg}(r_p)$ and β are independent, ignoring their covariance. This is inconsistent with comparing E_G to a full combined-probe analysis. When estimating errors on $E_G(r_p)$ from data, jacknife errors are typically used, which means that covariances are taken into account.

If I can to calculate the full covariance matrix for the joint-probe analysis, as described above, this will also enable to calculation of an $E_G(r_p)$ covariance matrix which accounts for covariances. I would simply construct a multi-variate Gaussian from this covariance matrix and jointly draw samples of $\Upsilon_{gm}(r_p)$, $\Upsilon_{gg}(r_p)$, and β, combining them to create a population of $E_G(r_p)$ samples. I would then calculate $\text{Cov}(E_G(r_p), E_G(r'_p))$ using standard formulae for estimating variances and covariances from a sample as before, but this time covariances between the probes are accounted for.

III. FORECASTING METHODOLOGY

A. Set-up

We consider two observational scenarios, one representing near-future or currently-available datasets and one representing upcoming datasets.

The near-future scenario we consider assumes Dark Energy Survey Y5 sources and SDSS LOWZ (check overlap) for lenses as well as for β.

The more futuristic scenario assumes LSST Y10 sources and DESI LRG for lenses and β.

We consider the parameter space of $\vec{p} = \{\Omega_\text{C}, \Omega_\text{B}, \sigma_8, n_s, b, \mu_0, \Sigma_0\}$ + observational systematic parameters like other bias parameters, IA....

Recall that if we assume the likelihood is a multivariate Gaussian, the Fisher matrix takes the form (see, for example, [3]):

$$\mathcal{F}_{ab} = \frac{\partial \langle \vec{d} \rangle}{\partial p_a} \mathbf{C}^{-1} \frac{\partial \langle \vec{d} \rangle}{\partial p_b} \bigg|_{\vec{p}_*} + \frac{1}{2} \text{Tr} \left(\mathbf{C}^{-1} \frac{\partial \mathbf{C}}{\partial p_a} \mathbf{C}^{-1} \frac{\partial \mathbf{C}}{\partial p_b} \right) \bigg|_{\vec{p}_*}, \tag{23}$$

B. Modified-gravity information content: defining a figure of merit

We define a 'modified gravity figure of merit' analogous to that presented in [1] for dark energy, which is equal to the area inside the forecast 1σ constrain contour in the μ_0, Σ_0 plane, marginalizing over all other parameters.

IV. RESULTS

a. Requirements on source photo-z Source galaxies for a weak lensing measurement will typically require photometric redshifts. In our forecasts, we have assumed a source galaxy redshift distribution of the form given by equation **??**. We estimate the level of uncertainty permissible on the mean and variance of this distribution by varying the parameters of **??** to adjust these quantities and comparing the level of difference in predicted observables with the forecast statistical uncertainty. We find that for the observational scenario considered, the mean of the source redshift distribution must be known to X and the variance must be known to Y.

V.   DISCUSSION AND CONCLUSIONS

[1] Albrecht A., et al., 2006, arXiv e-prints,
[2] Baldauf T., Smith R. E., Seljak U., Mandelbaum R., 2010, Physical Review D, 81
[3] Bassett B. A., Fantaye Y., Hlozek R., Kotze J., 2011, International Journal of Modern Physics D, 20, 2559
[4] Blake C., et al., 2016, Monthly Notices of the Royal Astronomical Society, 456, 2806
[5] Jeong D., Komatsu E., Jain B., 2009, Physical Review D, 80, 123527
[6] Leonard C. D., Baker T., Ferreira P. G., 2015, Physical Review D, 91, 083504
[7] Reyes R., Mandelbaum R., Seljak U., Baldauf T., Gunn J. E., Lombriser L., Smith R. E., 2010, Nature, 464, 256
[8] Singh S., Mandelbaum R., Seljak U., Slosar A., Vazquez Gonzalez J., 2016, preprint, (arXiv:1611.00752)