Covariance Testing

DESC Members. Decide on author list?
(Dated: August 28, 2019)

There are a number of codes that compute covariance matrices analytically; the plan is to use these to build TJPCov. In this project, we start along the path of comparing these different codes, building up a suite of tools that can be used to compare covariance matrices. We expect these tools to be useful not only for converging on a single accurate code for computing covariance matrices but also more generally for understanding which parts of the covariance matrix carry the most information (and therefore need the most attention to get right) and which are not relevant (so for example matrices that are not positive definite may still be usable if the negative eigenmodes are not relevant).

I. INTRODUCTION

II. METHODOLOGY

In this section, we explain the methodology used for each of the tests carried out for the analysis and comparison of the covariance matrices.

Our analysis is carried out using cosmic shear statistics $\xi_{\pm}(\theta)$, focusing for the most part on the Year 1 results of the Dark Energy Survey [1] (DESY1), and also on predictions for DES Year 3 (DESY3). The data is divided into four tomographic redshift bins spanning the interval $0.20-1.30$, which gives us 10 bin combinations, each one containing 20 angular bins between 2.5 and 250 arcmin. We thus have 200 data points for each $\xi_{+}(\theta)$ and $\xi_{-}(\theta)$, giving us 400 in total.

We apply the angular cuts described in [1], which removes the scales most sensitive to baryonic effects; this leaves us with 167 points for $\xi_{+}(\theta)$ and 60 for $\xi_{-}(\theta)$, resulting in 227 data, which correspond to a 227×227 covariance matrix.

The covariance matrices are obtained using two different codes: `Cosmolike` [8] and a simpler version of the code used to analyse the KiDS-450 survey [7], hereafter referred to as CL and BJ, respectively. To perform cosmological parameter inference we use the `CosmoSIS` [11] pipeline, while employing the `MultiNest` [4] sampler to explore the parameter space, with 500 `livepoints`, `efficiency` set to 0.3, `tolerance` to 0.1 and `constant efficiency` set to False.

Parameter	Prior
Cosmological	
Ω_m	$\mathcal{U}(0.1, 0.9)$
$A_s \times 10^9$	$\mathcal{U}(0.5, 5)$
$H_0(\mathrm{kms}^{-1}\mathrm{Mpc}^{-1})$	$\mathcal{U}(55, 91)$
Ω_b	$\mathcal{U}(0.03, 0.07)$
$\Omega_\nu h^2$	$\mathcal{U}(0.0005, 0.01)$
n_s	$\mathcal{U}(0.87, 1.07)$
Astrophysical	
A	$\mathcal{U}(-5, 5)$
η	$\mathcal{U}(-5, 5)$
Systematic	
m^i	$\mathcal{G}(0.012, 0.023)$
Δz^1	$\mathcal{G}(-0.001, 0.016)$
Δz^2	$\mathcal{G}(-0.019, 0.013)$
Δz^3	$\mathcal{G}(0.009, 0.011)$
Δz^4	$\mathcal{G}(-0.018, 0.022)$

TABLE I. List of the priors used in the analysis for parameter constraints. For the cosmological parameters, we fix $w = -1.0$, $\Omega_k = 0.0$ and $\tau = 0.08$. The astrophysical parameters are associated with the intrinsic alignment, they follow the relation $A(z) = A[(1+z)/1.62]^{\eta}$, we have also $z_0 = 0.62$. Lastly, for systematics we have m^i corresponding to the shear calibration, and Δz^i for the source photo-z shift, with $i = 1, 4$ in both cases. The lens photo-z shift parameters Δz^i, $i = 1, 5$ are fixed to zero.

A. One-to-one Comparison

This is a simple matter of comparing plotting the elements of two covariance matrices against each other to explain the differences in the parameter contraints. This will also be useful later on when talking about compressed covariances.

B. Eigenvalues

An intuitive approach to identifying the most contributing elements of a covariance matrix is to look at its eigenvalues, since we can associate them to the variance. The lowest eigenvalues would correspond to the least variance and therefore hold the most information, whereas the highest eigenvalues would implicate larger error and thus contribute significantly less, which implies that the most important elements of the matrix are those with the lowest eigenvalues. To verify this assumption, we need to look at the parameter constraints: remov-

ing the elements with the highest eigenvalues should not alter our results; the opposite, however, should give us broader constraints. Our procedure consists of first diagonalizing the covariance matrix in order to calculate its eigenvalues and then replacing the eigenvalues of interest with numbers of about nine orders of magnitude higher, thus removing their effective contribution; we then obtain a new covariance matrix with the modified eigenvalues. Finally, we perform a cosmological analysis with the new covariance matrix, to constrain the parameters of our model.

C. Signal-to-noise ratio

Another way of analysing covariance matrices is by obtaining the signal-to-noise ratio (SNR). A simple way of achieving this is to note that the total SNR squared is

$$\left(\frac{S}{N}\right)^2 = \sum_{ij} D_i C_{ij}^{-1} D_j \ , \tag{1}$$

where D_i and D_j are the data points and C the covariance matrix. If C were diagonal, then the eigenvectors would simply be the data points themselves, and we could estimate the SNR squared expected in each mode by simply computing T_i^2/C_{ii} where T_i is the theoretical prediction. Then we could throw out the modes with the lowest SNR. Since C is not diagonal, we have to first diagonalize it and then order the values. So, we write the expected SNR squared as

$$\begin{aligned}\left(\frac{S}{N}\right)^2 &= \sum_{ij} T_i C_{ij}^{-1} T_j \\ &= \sum_i \frac{v_i^2}{\lambda_i},\end{aligned} \tag{2}$$

where λ_i are the eigenvalues of the covariance matrix, which is diagonalized with the unitary matrix U, and the eigenvectors are

$$v_i \equiv U_{ij}^T T_j \ . \tag{3}$$

This makes it very clear which modes should be kept and which should be dropped. Modes v_i for which v_i^2/λ_i is very small can be discarded.

After obtaining the SNR for the covariance matrix, we proceed to set the 50 lowest values to seven orders of magnitude lower, which is equivalent to increasing the noise (or decreasing the signal) of these modes. We then obtain a new covariance matrix with the corresponding modified SNR values.

D. Shrinkage

Our final test for comparing covariance matrices and extracting the most important elements, consists of shrinking the data vector and the covariance matrix. There have been several methods proposed in the literature to compress the data vectors, extracting as much information as possible [5, 6, 9]. Here we consider two: compression at the map level [2], where linear combinations of the tomographic maps are used to gain maximum signal-to-noise ratio, before transferring this compression scheme into real-space to compress the 2-point function. The second method directly operates in the real-space [10], where the modes used are those that maximize the Fisher information for each of the cosmological parameters.

Shrinking the data vector and its covariance matrix is equivalent to projecting the data vector onto a subspace of the original vector space. For example, if our data have a length of n and we want to compress it into a lower dimension $m \leq n$, we need to come up with a basis transformation matrix $U_{m\times n}$ that has orthonormal row vectors, or mathematically speaking, $< U_{xi} \cdot U_{yi} >= \delta_{xy}$. Once we have a basis transformation matrix $U_{m\times n}$, we can compress the data vector y and covariance matrix C into filter data $\tilde{y}$ and $\tilde{C}$

$$\tilde{y} = Uy \tag{4}$$

$$\begin{aligned}\tilde{C} &= < \tilde{y}\tilde{y}^T > \\ &= U < yy^T > U^T = UCU^T \ .\end{aligned} \tag{5}$$

For a data vector with length m, there are $m(m+1)/2$ elements in the covariance matrix to be evaluated. Once the data vector is shrunk to length n, elements in covariance matrices will shrink quadratically. Therefore, shrinking the covariance matrices is an effective way of comparing them.

Here, we will demonstrate these two methods to compress the covariance matrices. In both cases, the data vector is shrunk by 90%, which correspond to the covariance matrices being reduced by 99%.

1. *Tomographic Compression*

This compression method is based on Karhunen-Loéve decomposition for the shear power spectrum suggested by [2] and later applied to real space 2-point function in [3] for CFHTLens survey. This method generally finds the eigenmode with most of the signal-to-noise ratio contribution to the power spectrum, and then transforms the 2-point function in real space based on this eigenmode.

With `CosmoSIS`, we can generate the shear power spectrum C_ℓ of convergence $a_{\ell m}$ for a fiducial cosmology. The cosmology we choose is the best-fit value of the DES Year 1 results for cosmic shear only. With the shear power spectrum $C_\ell = S_\ell + N_\ell$ and its shape noise N_ℓ, we can calculate the Karhunen-Loéve (KL) modes matrix E_ℓ^p via a general eigenvalue problem:

$$C_\ell E_\ell^p = \lambda^p N_\ell E_\ell^p \ , \tag{6}$$

and the new observable $b_{\ell m} = E_p \cdot N^{-1} a_{\ell m}$. We should note that C_ℓ is the power spectrum of the convergence of

the weak lensing, and E_p is the transformation of basis for the convergence. We should point out that these eigenmodes are uncorrelated, so the power spectrum of the new observable D_ℓ is a diagonal matrix, with 1+SNR of the corresponding eigenmodes on the diagonal elements,

$$D_\ell = < b_{\ell m} b^T_{\ell m} > = E^p_\ell N^{-1} C_\ell N^{-1} E^{pT}_\ell \ , \qquad (7)$$

or, if we denote $E^p_\ell N^{-1}$ as R_ℓ, we can write the compression in one simple linear combination of the C_ℓ,

$$D_\ell = R^i_\ell C^{ij}_\ell R^j_\ell = U^{ij}_\ell C^{ij}_\ell \ . \qquad (8)$$

The double summation weight U^{ij}_ℓ is what we use to perform tomographic compression. Since the KL-decomposed modes of shear power spectrum are uncorrelated, we can make a compression here by only taking the first one or two modes with the highest SNR. By doing so, we compress 10 tomographic combinations to 1 or 2.

We want, however, to eventually compress the 2-point function data vector of DESY1. One possible way is to calculate the 2-point function of the KL mode of the shear power spectrum. We can calculate the 2-point function from the shear power spectrum by

$$\xi^{ij}_\pm(\theta) = \int \frac{\ell d\ell}{2\pi} J_{0/4}(\ell\theta) C^{ij}(\ell) \ . \qquad (9)$$

In order to compress the 2-point function based on the compression of the C_ℓ, we need to make sure that the scheme for C_ℓ is ℓ-independent, that is to say, the 2-point correlation function of D_ℓ, $\tilde{\xi}_\pm(\theta)$, can be directly calculated from other 2-point functions. We then have,

$$\begin{aligned} \tilde{\xi}_\pm(\theta) &= \int \frac{\ell d\ell}{2\pi} J_{0/4}(\ell\theta) D(\ell) \\ &= \int \frac{\ell d\ell}{2\pi} J_{0/4}(\ell\theta) U^{ij} C^{ij}(\ell) \\ &= U^{ij} \xi^{ij}_\pm(\theta) \ , \end{aligned} \qquad (10)$$

where U^{ij}, the ℓ-independent compression weight is calculated by

$$U^{ij} = \frac{\int^{\ell_{\rm max}}_{\ell_{\rm min}} (2\ell+1) U^{ij}_\ell}{\int^{\ell_{\rm max}}_{\ell_{\rm min}} (2\ell+1)} \ . \qquad (11)$$

We make a more conservative angular cut than the angular cut discussed in [ref], making sure that the cut for both $\xi_\pm$ are uniform in regard to tomographic combinations. For ξ_+, we consider an angular scale from 7.195° to 250.0°. For ξ_-, the angular scale is from 90.579° to 250.0°. Therefore, for the purpose of demonstrating KL-transform, the raw data vector has a length of 190, and by shrinking 10 tomographic combinations for each angle into 1 KL-mode, the data is shrunk to 19, and so the number of elements in the covariance matrices are reduced by 99%.

2. 2-Point Correlation Function

The compression here takes place at the 2-point level [10], with the compressed data vector containing linear combinations of the many 2-point functions. In principle, this might work with only N_p 2-point functions where N_p is the number of free parameters, and each mode, or linear combination, contains all the information necessary about the parameter of interest.

For each parameter p_α that is varied, one captures a single linear mode

$$y_\alpha = U_{\alpha i} D_i \ , \qquad (12)$$

where D_i are the data points and the coefficients are defined as

$$U_{\alpha i} \equiv \frac{\partial T_j}{\partial p_\alpha} C^{-1}{}_{ji} \ , \qquad (13)$$

with T_j being the theoretical prediction for the data point D_j. The now much smaller data set $\{y_\alpha\}$, which contains as few as N_p data points, carries with it its own covariance matrix, with which the χ^2 can be computed for each point in parameter space. Propagating through shows that this covariance matrix is related to the original C_{ij} via

$$C_{\alpha\beta} = U_{\alpha i} C_{ij} U^T_{j\beta}. \qquad (14)$$

In our case, our covariance matrix is 227×227, while the number of parameters needed to specify the model is only 16, so $C_{\alpha\beta}$ is a 16×16 matrix. We have apparently captured from the initial set of $(227 \times 228)/2 = 25,878$ independent elements of the covariance matrix a small subset (only 136 in this case) of linear combinations of these 25k elements that really matter. If two covariance matrices give the same set of $C_{\alpha\beta}$, it should not matter whether any of the other eighty thousand elements differ from one another.

3. Invertible Transformation and Tolerance Testing

In the last section, we shrink the data vector and covariance matrices and find the most cosmological-informative modes in both. However, in order to make a tolerance testing of each element in the covariance, we not only need the cosmological-informative modes but also the uninformative modes. Suppose we find the informative set of modes for cosmology, the modes that are orthogonal to them, or the complementary set of the informative one, form a set of modes that are cosmological-uninformative.

To find these modes, we start off with the compression scheme presented in Eq. 13, which will serve as the basis for a rotation matrix W_α. We then use the Gram-Schmidt decomposition to create $227 - N_p$ vectors

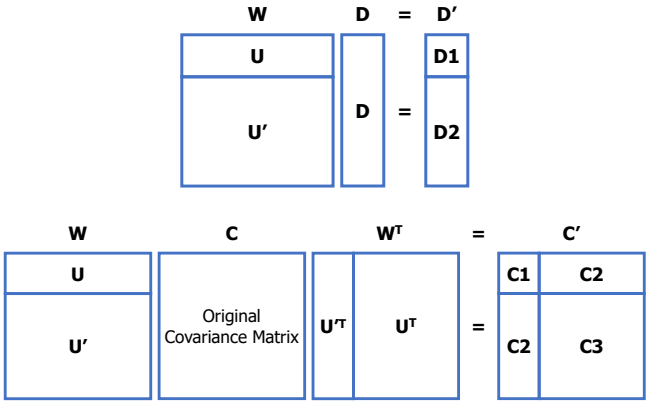

FIG. 1. Illustration of the invertible transformation, W. It's components, U and U' are the compression scheme described in Sections II D and orthogonal components obtained using the Gram-Schmidt decomposition, respectively. **Top:** Transformation of the data vector, where D1 corresponds to compressed set, and D2 are uninformative for parameter constraints. **Bottom:** Transformation of the covariance matrix. The last square on the right, C' is divided into four blocks: C1 is the most informative to cosmology, C2 are also relevant to constraints, but on a lower scale, and, finally, C3 is irrelevant to the χ^2 calculation.

orthonormal to U_α, thus obtaining a unitary 227×227 matrix. We then have,

$$C'_{\alpha\beta} = W_{\alpha i} C_{ij} W^T_{j\beta}. \tag{15}$$

With the invertible transformation W, shown in Figure 1, the data vector and covariance matrix are transformed into sectional blocks. We will describe the meaning of each block here. The data vector is split into two blocks: D1 has the transformed data points that are sensitive to changes in the cosmological parameters, that is, the cosmological-sensitive data; D2 is generated by the modes that are orthogonal to D1, they will, in principle, be unaffected by changes to the cosmological parameters, which makes them uninformative for parameter constraints.

The transformed covariance matrix is split into 4 blocks, with the off-diagonal ones being the transpose of each other. Block C1 is the variance and covariance of the cosmological-sensitive data, D1, it describes how well the cosmological-sensitive data is measured, and is, therefore, the one that contributes most to the χ^2 calculation, and, consequently, to parameter constraints. Block C2 is the cross-correlation between D1 and D2, it is important for parameter constraints because it describe how D1 is affected by the uncertainty of D2. Finally, C3 relates to the uninformative data, D2; it plays a minor role to χ^2, so it affects the parameter constraints least.

The next step is then to test the tolerance of different parts of the transformed covariance matrix. We first compare the results of increasing the error of each block separately by a factor of 100 and compare the results to constraints with the unmodified covariance, then we start by introducing smaller errors to the relevant blocks, C1 and C2 and analyse the corresponding increase in the parameter constraints.

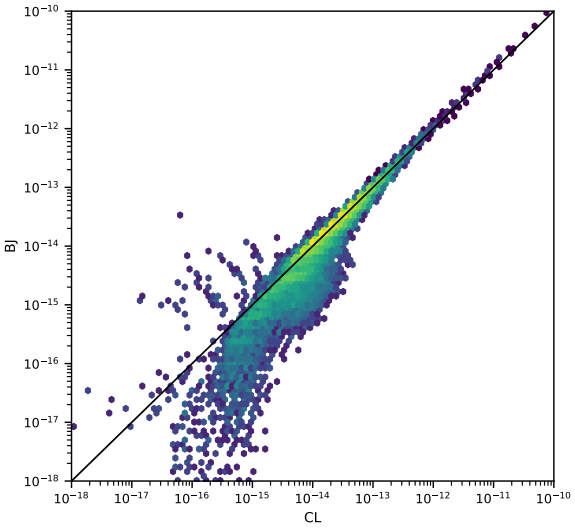

FIG. 2. Scatter plot of the elements of the covariance matrices BJ and CL.

III. RESULTS

Ultimately, what matters is how well the likelihood does at extracting parameter constraints. Since most analyses assume a Gaussian likelihood, this boils down to how well the contours in parameter space agree when computing the χ^2 using two different covariance matrices.

For brevity, we show only constraints on Ω_m and σ_8.

A. One-to-One Comparison

Using the same data vector, we compare the constraints obtained for two different covariance matrices: CL and BJ. Figure 4 shows the results for the cosmological constraints, where we can see that, while the best-fit values agree within a 2σ interval, the constraints are up to 25% broader for σ_8. This is also true for the other parameters, where, on average, constraints obtained with CL are about 18% wider.

We make a scatter plot of the elements of the two matrices in Figure 2, where we can see that the elements of CL are, in general, much larger than BJ's, differing in up to 3 orders of magnitude. This is in accordance to what we see in the parameter constraints: more variance in the covariance matrix translates to less constraining power.

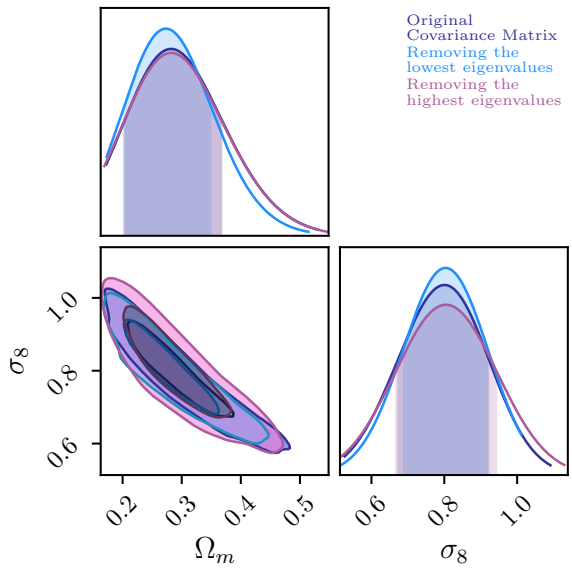

FIG. 3. Constraints on cosmological parameters Ω_m and σ_8 for the original DESY1 covariance matrix (in purple) and for two new covariance matrices obtained by setting the fifty highest eigenvalues of the original matrix to nine orders of magnitude higher (in blue), and by replacing the fifty lowest eigenvalues to nine order of magnitude lower (in magenta).

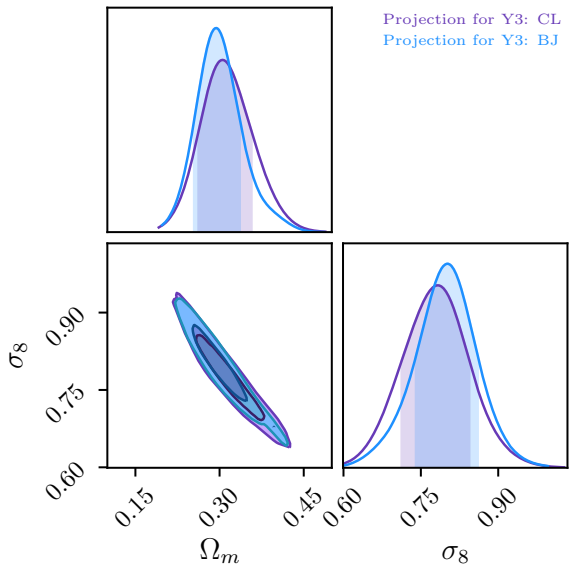

FIG. 4. Constraints on cosmological parameters Ω_m and σ_8 for two covariance matrices produced for cosmic shear for DESY3. The purple curve is for CL while the blue is for BJ. The constraints are about 18% larger for the former, indicating that the two matrices have quantifiable differences between them.

B. Eigenvalues

We perform two different analysis: the first consist of modifying the fifty lowest eigenvalues, and for the second, we take the fifty highest. Figure 3 shows the results of both approaches compared to constraints for the original covariance matrix. We can see in the Posterior Density Functions (pdfs) that the mean values of both parameters are well within the 2σ interval, which is also true for all the other parameters. As for the constraining power, the sizes of the constraints are within less than 5% of each other for almost all of the parameters. Since the results are similar for both approaches, it becomes clear that ranking eigenvalues is not the most efficient way of evaluating where the most contributing elements of the covariance matrix lie.

The eigenvalues are also used to try to identify the differences in parameter constraints for CL and BJ shown in the previous section. We plot the eigenvalues for these matrices in Figure 5. At a first glance, both curves show reasonable agreement, with values differing only by an average of $\approx$ 15%, but we should keep in mind that the order of magnitude of the values vary greatly, which makes it difficult to compare them efficiently with this methodology, furthermore, it is not clear which elements are most important.

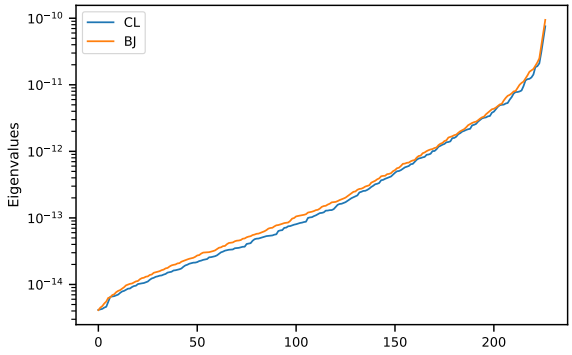

FIG. 5. A log plot showing the 227 eigenvalues of CL (blue) and BJ (orange).

C. Signal-to-noise ratio

The parameter constraints for this method are shown in Figure 6, where we note that the mean values are different about 10% and 4% for Ω_m and σ_8, respectively, with 17% broader constraints for the former, when compared to the results with the original covariance. This tells us that the modes removed are essential for obtaining better constraints on these parameters.

Another aspect that we must consider when applying this procedure is that the highest modes do not necessarily hold the most information for all of the parameters; this is evident when we analyse the SNR for each param-

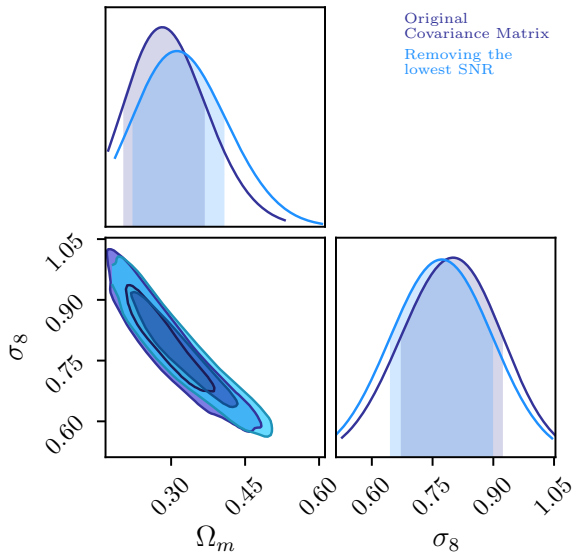

FIG. 6. Constraints on cosmological parameters Ω_m and A_s for the original DESY1 covariance matrix (in purple) and for a covariance matrix (in blue) obtained by setting fifty elements corresponding to the lowest SNR to a value seven orders of magnitude lower, in order to evaluate their contribution to parameter constraints.

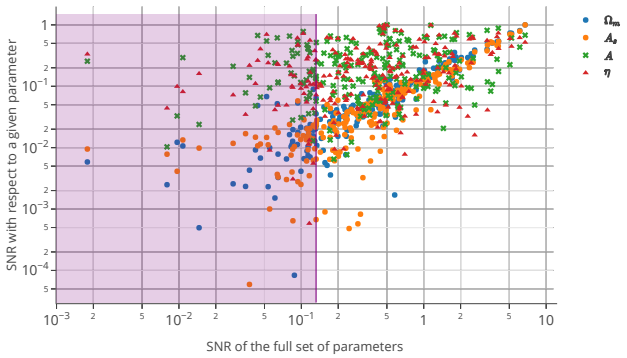

FIG. 7. Scatter plot for the relation between the signal to noise obtained with the covariance matrix for DESY1 for each parameter (x-axis) against that for the full set of parameters (y-axis). The derivatives are shown with respect to cosmological parameters Ω_m (blue) and A_s (orange), and for the intrinsic alignment parameters A (green) and η (red). The purple rectangle spreads until the fiftieth lowest value of SNR, which corresponds to the values that are modified for parameter constraints.

eter individually. To illustrate this, we take

$$\left(\frac{\partial S/\partial p_\alpha}{N}\right)^2 = \sum_i \frac{(\partial v_i/\partial p_\alpha)^2}{\lambda_i}, \tag{16}$$

with

$$v_i^\alpha \equiv U_{ij}^T \frac{\partial T_j}{\partial p_\alpha}, \tag{17}$$

where $\partial/\partial p_\alpha$ is the derivative with respect to each parameter θ. This produces the SNR for each parameter of interest. The importance of this procedure is illustrated in Figure 7, where we can see that while the highest SNR for all the parameters does indeed correspond to those that hold most information for parameters Ω_m and A_s, this is not the case, for the intrinsic alignment parameters A and η. As a result, we may end up losing constraining power over these parameters when we modify the lowest values of SNR. While this is apparent in the aforementioned figure, it is not very clear when looking at the resulting constraints on these parameters because the constraining power of the data over these parameters is not very strong (errors of about 100%). As such, these results do not encourage us to use this method for identifying the most important elements of the covariance matrix.

D. Shrinkage

1. *Tomographic Compression*

With `CosmoSIS`, we calculate the shear power spectrum C_ℓ of DES Year 1 with a fiducial cosmology at the best-fit parameters, and ℓ-range $2-5000$. The left plot in Figure 8, shows the diagonal elements of the signal part and the noise part of C_ℓ, while the right one shows the KL-transformed eigenmode D_ℓ of C_ℓ. We can see that the first KL mode contains most of the SNR contribution to the power spectrum. If we want to recover most information, we can also include the second mode.

In Figure 9, we plot the normalized KL-eigenmode E_ℓ^i of C_ℓ and its corresponding $W_\ell^{ij} = E_\ell^i E_\ell^j$. Modes with different ℓ are plotted with different depth of the color. We can see that the KL-modes do not depend a lot on scale factor ℓ by a large portion, so we take the weighted average of the eigenmodes E_ℓ^p and its quadratic form W_ℓ over ℓ's and plot them with black lines.

For different ℓ, the KL-modes do vary by a slight amount. We also observe that tomographic bins with higher redshift gains more weight than those with low redshift. This is also shown by the weight on tomographic combination that the combination of bin 3 and bin 4 gains most of the weight to maximize signal to noise ratio. This agrees with the fact that the diagonal elements C_ℓ for low redshift is much less than those with high redshift. Is my N_ℓ actually right?

We use only the first KL-mode to compress the tomographic combination for each angle θ, i.e., 190 data points

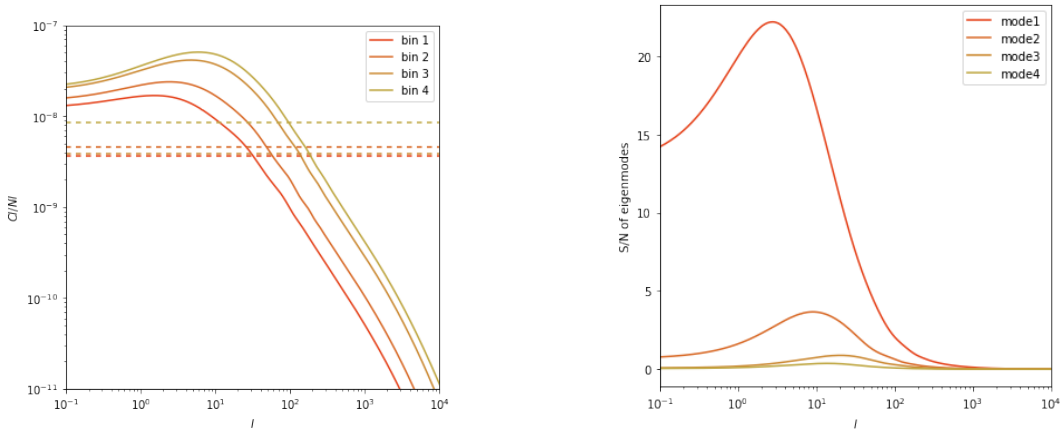

FIG. 8. **Left:** Shear power spectrum of DESY1. Solid lines are diagonal elements of the signal matrix S_ℓ, and dashed lines are the diagonal elements of noise matrix N_ℓ. **Right:** Signal to noise ratio matrix D_ℓ of KL-modes of the power spectrum on the left. Make the plots colour blind friendly.

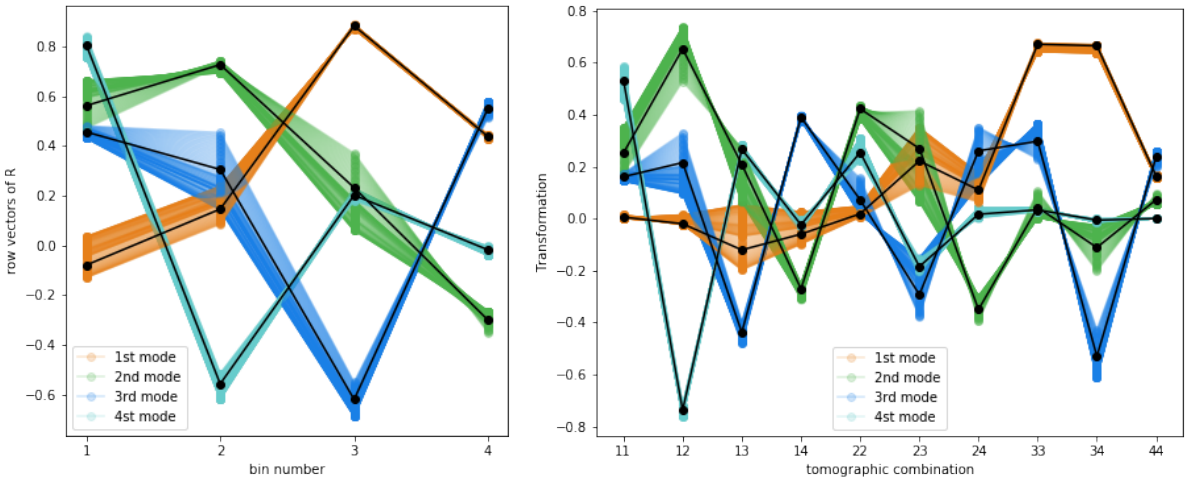

FIG. 9. **Left:** Normalized KL-eigenmodes E_ℓ^p of the shear power spectrum C_ℓ, darkness of the color is representing different ℓ. **Right:** Transformation on tomographic bin combination W_{ij} constructed by the KL-eigenmodes. Black lines are the weighted average of each mode.

for 19 angles are compressed into 19 numbers. This compression scheme is also used to compress the CL and BJ covariance matrices, and we plot them in Figure 10.

If we now again make a one-to-one comparison of these two compressed covariance matrices. We can easily notice that the elevated clump at the bottom right for scatter of original matrices, which represent elements with great difference, are gone in the compressed covariance. Instead, the two covariance matrices just have a relative constant difference because of the different model they use. This shows that the divergence between CL and BJ covariance does not affect the overall signal to noise ratio a lot.

2. *2-Point Correlation Function*

Figure 11 compares the constraints obtained for the compressed covariance and dataset with the full one. The mean values agree at the 2σ level, with the exception of η, which is not very well constrained in either analysis. The results for the compressed covariance are about 0.5% broader, which shows that the information loss is negligible.

One relevant point in this analysis is at which point to take the derivative of each parameter. When we wish to compare the results of our compression scheme with those obtained with the full covariance matrix and data set, it is important to derivate each parameter at their respective mean value (obtained by performing the analysis with the full covariance matrix). The shape and variance of the

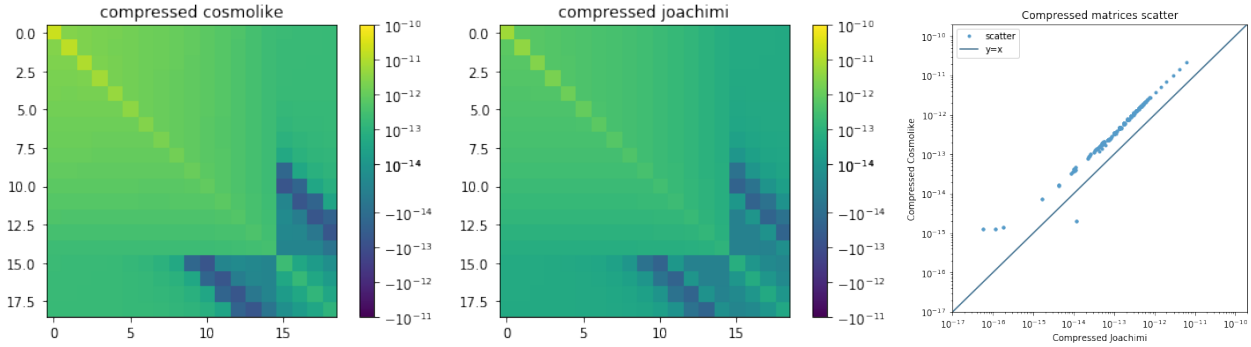

FIG. 10. **Left:** CL covariance matrix compressed by the first KL-mode. **Middle:** BJ covariance matrix compressed by the first KL-mode. **Right:** One-to-One scatter of the two compressed matrices

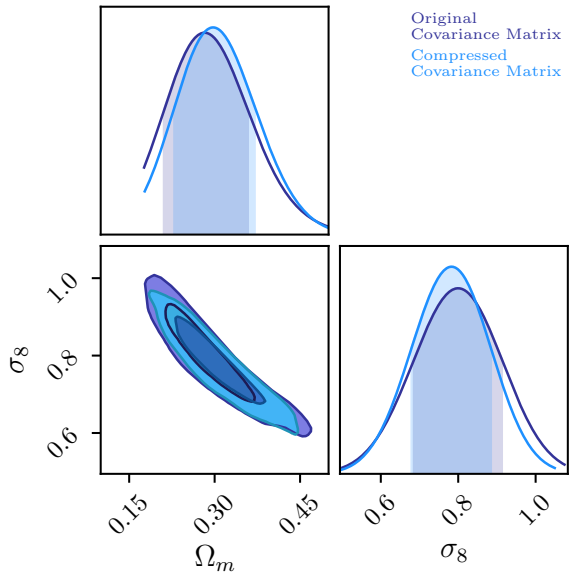

FIG. 11. Constraints on cosmological parameters Ω_m and σ_8 for the original DESY1 covariance matrix (in purple) and for the compressed one (in blue).

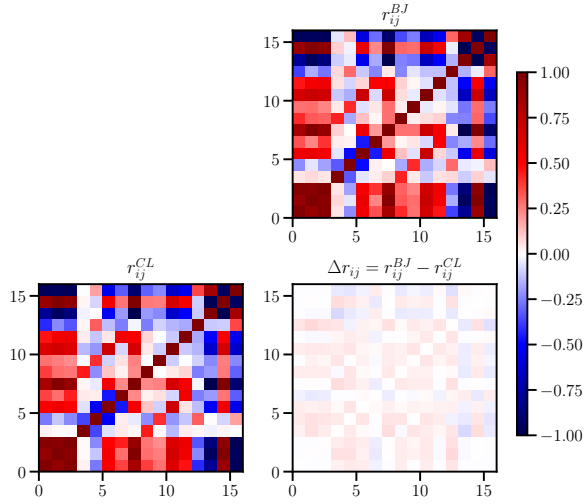

FIG. 12. The upper right and lower left plots display the correlation matrix for BJ and CL respectively, while the lower right is the difference between the two.

posterior is not dependent on the derivative, but the best-fit value shifts according to the point where the derivative is taken.

We also apply this methodology to comparing the covariance matrices of interest, i. e. CL and BJ. In order to do this, we take two different approaches: first, we assume that $U_{\alpha,i}$ is the same for both covariance matrices and we calculate it with BJ. The second approach is that each compression scheme must use the original covariance matrix that will be compressed, so that $U_{\alpha,i}$ will be different for each covariance matrix. We find that the mean values of the parameter contraints for the two methods agree to 1σ, which shows that they are equivalent to each other. Figure 12 is obtained for the first method, which will be the one adopted from here on, it shows the correlation matrix for BJ and CL, and that of the difference between them; we find this figure important because we can clearly see the difference between the two matrices by simply looking at only $(16 \times 17)/2$ elements, as opposed to having to analyse the larger correlation matrix for the full covariance matrices. It is also crucial that the matrices used for comparison here are those obtained via the same compression scheme, so that we can be sure that their differences are indeed only related to the differences in the original matrices.

3. *Tolerance Testing*

Given that the Block C1 contains the most relevant elements, it was expected that any changes to it would also modify the parameter constraints. It is clear in Figure 13 that this was not the case.

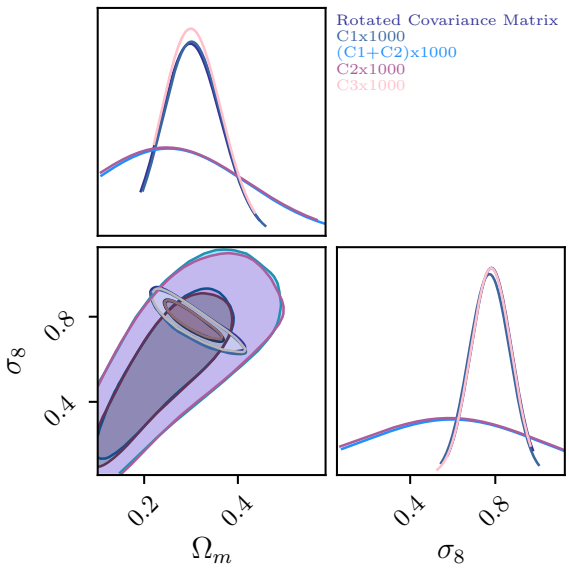

FIG. 13. The upper right and lower left plots display the correlation matrix for BJ and CL respectively, while the lower right is the difference between the two.

IV. DISCUSSION

By making the transformation described above, we will be able to reorganize the covariance matrices into cosmological informative blocks and uninformative blocks. Intrinsically, we can tolerate more errors in the uninformative blocks and have a stronger requirement on the informative blocks. This will put a lot of simulation time into more valuable work.

The gold of the invertibility of the transformation is that we can first assign higher tolerance to C3 and assign lower tolerance to C1 and C2, then use the transformation matrix to recover the covariance matrices in the original basis. Then, we can compare the recovered matrix with the original one and check out how much tolerance is allowed on each element. Without invertibility, we cannot accomplish this task, which allows us to precisely quantify the tolerance on each element in the covariance matrices.

V. CONCLUSION

Acknowledgments

The DESC acknowledges ongoing support from the Institut National de Physique Nucléaire et de Physique des Particules in France; the Science & Technology Facilities Council in the United Kingdom; and the Department of Energy, the National Science Foundation, and the LSST Corporation in the United States. DESC uses resources of the IN2P3 Computing Center (CC-IN2P3–Lyon/Villeurbanne - France) funded by the Centre National de la Recherche Scientifique; the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231; STFC DiRAC HPC Facilities, funded by UK BIS National E-infrastructure capital grants; and the UK particle physics grid, supported by the GridPP Collaboration. This work was performed in part under DOE Contract DE-AC02-76SF00515.

Need to thank CAPES as well.

[1] Abbott, T. M. C., et al. 2018, Physical Review D, 98, 043526
[2] Alonso, D. 2018, Monthly Notices of the Royal Astronomical Society, 473, 4306
[3] Bellini, E., Alonso, D., Joudaki, S., & Waerbeke, L. V. 2019, arXiv:1903.04957
[4] Feroz, F., Hobson, M. P., & Bridges, M. 2009, Monthly Notices of the Royal Astronomical Society, 398, 1601
[5] Gualdi, D., Manera, M., Joachimi, B., & Lahav, O. 2018, Monthly Notices of the Royal Astronomical Society, 4070, 4045
[6] Joachimi, B. 2017, Monthly Notices of the Royal Astronomical Society: Letters, 466, L83
[7] Köhlinger, F., et al. 2017, Monthly Notices of the Royal Astronomical Society, 471, 4412
[8] Krause, E., & Eifler, T. 2017, Monthly Notices of the Royal Astronomical Society, 470, 2100
[9] Tegmark, M., Taylor, A. N., & Heavens, A. 1997, The Astrophysical Journal, 480, 22
[10] Zablocki, A., & Dodelson, S. 2016, Physical Review D, 93, 083525
[11] Zuntz, J., Paterno, M., Jennings, E., et al. 2015, Astronomy and Computing, 12, 45