Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **466,** L83–L87 (2017)
Advance Access publication 2016 November 30

doi:10.1093/mnrasl/slw240

Non-linear shrinkage estimation of large-scale structure covariance

Benjamin Joachimi[⋆]

Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK

Accepted 2016 November 28. Received 2016 November 25; in original form 2016 July 23

ABSTRACT

In many astrophysical settings, covariance matrices of large data sets have to be determined empirically from a finite number of mock realizations. The resulting noise degrades inference and precludes it completely if there are fewer realizations than data points. This work applies a recently proposed non-linear shrinkage estimator of covariance to a realistic example from large-scale structure cosmology. After optimizing its performance for the usage in likelihood expressions, the shrinkage estimator yields subdominant bias and variance comparable to that of the standard estimator with a factor of ∼50 less realizations. This is achieved without any prior information on the properties of the data or the structure of the covariance matrix, at a negligible computational cost.

Key words: methods: data analysis – methods: numerical – methods: statistical – large-scale structure of Universe.

1 INTRODUCTION

The covariance is an indispensable ingredient for inference from data, quantifying the varying levels of statistical uncertainty among the data points as well as their correlations. In many astrophysical situations, the covariance is not known a priori and has to be determined from measurements along with the data, turning the elements of the covariance matrix themselves into random variables with associated errors. Cosmological data analysis faces a particular challenge in that only a single realization of the data is available, and that treating representative subsamples of the data as quasi-independent may be inaccurate due to long-range spatial correlations of the signals under investigation (Norberg et al. 2009). Therefore, one usually resorts to estimating a standard sample covariance matrix from simulated realizations of the data.

The finite number of mock data realizations induces noise in the covariance estimate that propagates into the errors of inferred model parameters, which was first explicitly pointed out in a cosmological context by Hartlap, Simon & Schneider (2007) and subsequently investigated in detail (Dodelson & Schneider 2013; Taylor, Joachimi & Kitching 2013; Percival et al. 2014; Taylor & Joachimi 2014; Sellentin & Heavens 2016). For Gaussian distributed data, the sample covariance follows a Wishart distribution. While, for example, the fields and derived statistics probed in cosmic large-scale structure (LSS) surveys follow strongly non-Gaussian distributions, the derived properties of the covariance inferred from the Wishart distribution turn out to still be applicable to a very good approximation (Dodelson & Schneider 2013; Petri, Haiman & May 2016). A generic property of a Wishart matrix is that it becomes singular if the number of realizations used to estimate it, $N_{\rm S}$, becomes less than the size of the data vector, $N_{\rm D}$, prohibiting its use in likelihood analysis or least-squares fitting, for which the inverse covariance is required. This implies that $N_{\rm S} \gg N_{\rm D}$ often computationally expensive simulations are required to determine the covariance, where $N_{\rm D} \sim 1000$ will be readily surpassed by forthcoming cosmological surveys.

To lessen this computational bottleneck in the analysis, $N_{\rm D}$ could be reduced via data compression, but good knowledge of the covariance is necessary to achieve near-optimal compression (e.g. Tegmark, Taylor & Heavens 1997). Innovative schemes to augment a given number of simulations via resampling techniques have been proposed as well (Schneider et al. 2011; Escoffier et al. 2016). Alternatively, one can replace the sample covariance estimator with a generally biased one that has favourable noise properties. Paz & Sánchez (2015) proposed to taper correlations far from the diagonal, which, however, requires a notion of distance between all elements of the data vector. Padmanabhan et al. (2016) investigated the direct estimation of inverse covariance matrix elements, bypassing the sample covariance altogether. Linear shrinkage towards a modelled target (Pope & Szapudi 2008) or a constant correlation coefficient (Simpson et al. 2016) has seen LSS applications. These estimators relied on the accuracy of the assumed model or structure of the covariance matrix, respectively, and were limited to a single global shrinkage intensity for all matrix elements, which can be suboptimal in improving the conditioning of the covariance (Ledoit & Wolf 2012). It is therefore timely to assess the performance of a non-linear generalization of shrinkage covariance estimators, which has the added benefit of not relying on models or assumptions about the structure of the covariance matrix.

[⋆] E-mail: b.joachimi@ucl.ac.uk

2 SHRINKAGE ESTIMATOR

This work adopts the NERCOME[1] estimator recently proposed by Lam (2016), which in turn capitalized on earlier work by Abadir, Dinasto & Žikeš (2014) and Ledoit & Wolf (2012). Let $\mathsf{X} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_{N_\mathrm{S}})$ be a $N_\mathrm{D} \times N_\mathrm{S}$ matrix of N_S realizations (independent measurements) of the data vector, $\boldsymbol{x}_i$, each of length N_D. The data have covariance $\boldsymbol{\Sigma}$, which, however, is unknown a priori. In the following, it is assumed that the $\boldsymbol{x}_i$ are mean-subtracted and have potentially been normalized, i.e. their elements are given by $x_{\alpha i} = (x^\mathrm{raw}_{\alpha i} - \mu_\alpha)/n_\alpha$,[2] where $\boldsymbol{\mu}$ is the vector of means (also estimated from the data) and $\boldsymbol{n}$ is a normalization vector with noiseless entries. The standard sample covariance estimator reads

$$\hat{\mathsf{S}} = \frac{1}{N_\mathrm{S} - 1}\, \mathsf{X}\, \mathsf{X}^\tau\,, \tag{1}$$

which is unbiased, $\langle \hat{\mathsf{S}} \rangle = \boldsymbol{\Sigma}$. A key idea of NERCOME is to divide the data set into two subsamples, $\mathsf{X} = (\mathsf{X}_1, \mathsf{X}_2)$, with X_1 an $N_\mathrm{D} \times s$ matrix and X_2 an $N_\mathrm{D} \times (N_\mathrm{S} - s)$ matrix. The sample covariance can also be measured from each subset, denoted by $\hat{\mathsf{S}}_i$, with $i = 1, 2$. The estimator uses the diagonal decomposition of these estimates, $\hat{\mathsf{S}}_i = \mathsf{U}_i\, \mathsf{D}_i\, \mathsf{U}_i^\tau$, where U is the matrix of eigenvectors and D is a diagonal matrix with entries $d_{\alpha\beta} = \delta_{\alpha\beta}\, \lambda_\alpha$, where λ_α are the eigenvalues and δ is the Kronecker delta.

The NERCOME estimation process consists of three steps:

(1) apply the basic estimator

$$\hat{\mathsf{Z}} \equiv \mathsf{U}_1\, \mathrm{diag}\left(\mathsf{U}_1^\tau\, \hat{\mathsf{S}}_2\, \mathsf{U}_1 \right)\, \mathsf{U}_1^\tau \tag{2}$$

to a given subdivision of X;

(2) average over different compositions of $(\mathsf{X}_1, \mathsf{X}_2)$ for a given location s of the split, of which there are $\binom{N_\mathrm{S}}{s}$;

(3) find the optimal location of the data vector split by minimizing

$$Q(s) = \left|\left| \overline{\hat{\mathsf{Z}}}(s) - \overline{\hat{\mathsf{S}}}_2(s) \right|\right|_\mathrm{F}^2\,, \tag{3}$$

where the bar denotes the average of step (2), and where $||\mathsf{A}||_\mathrm{F}^2 = \mathrm{Tr}(\mathsf{A}\,\mathsf{A}^\tau)$ is the Frobenius matrix norm. An estimate for the inverse covariance is then simply provided by the inverse of the covariance estimator.

Equation (2) takes advantage of the fact that $\hat{\mathsf{S}}_1$ and $\hat{\mathsf{S}}_2$ are estimated independently of the data, so that their combination can be expected to be less adversely affected by noise when estimating the diagonal elements of the covariance. Lam (2016) showed that $\mathsf{U}_1^\tau\, \hat{\mathsf{S}}_2\, \mathsf{U}_1$ is the expression for the diagonal elements that minimizes the difference to the true covariance in the Frobenius norm. Abadir et al. (2014) demonstrated that the subsequent averaging over a moderate number of compositions of $(\mathsf{X}_1, \mathsf{X}_2)$ suppresses noise in $\hat{\mathsf{Z}}$. Here, that number is chosen to be $N_\mathrm{av} = \min\left\{ \binom{N_\mathrm{S}}{s}; 500 \right\}$, where, in the latter case, combinations are drawn at random once $\binom{N_\mathrm{S}}{s} > 3N_\mathrm{av}$. Equation (3) is minimized by evaluating Q at 20 equidistant steps in s in the range of $[0.1N_\mathrm{S}; 0.9N_\mathrm{S}]$. Since Q itself is a rather noisy quantity primarily through $\hat{\mathsf{S}}_2$, which serves as an unbiased estimate of the true covariance matrix, results for a fixed split at $s/N_\mathrm{S} = 2/3$ (meaning two-thirds of the data are used to estimate U) are also reported.

NERCOME is close to ideal when the true covariance is a multiple of the identity, $a\,\mathsf{I}$, as asymptotically the value of U becomes irrelevant and most constraining power can be focused on estimating the single number a (Lam 2016). It is therefore advisable to whiten the covariance by an informed choice of the normalization, $\boldsymbol{n}$, if possible, in case an optimal estimate of $\boldsymbol{\Sigma}$ in a mean square error sense is the goal. However, in physical applications, one is usually more interested in controlling the uncertainty and bias of weighted sums of inverse covariance entries that enter likelihood analyses and weighted least-squares fits. This is assessed here by using the scalar quantity $F \equiv (\mathrm{S/N})^2 = \boldsymbol{m}^\tau\, \boldsymbol{\Sigma}^{-1} \boldsymbol{m}$ as the figure of merit for covariance estimation, where $\boldsymbol{m}$ is the best-guess model of the data vector. Setting $\boldsymbol{n} = \boldsymbol{m}$ for the remainder of this work, the NERCOME formalism now optimizes the performance with respect to the same combination of covariance elements as those in F.

An incorrect choice of the model $\boldsymbol{m}$ will not per se lead to biased inference on cosmological parameters but to a potentially suboptimal performance of the NERCOME algorithm (which could then imply biases on the posterior distribution). As long as signals can be predicted to within a few tens of per cent accuracy, the uncertainty should only marginally affect the NERCOME estimate.

3 A TOY EXAMPLE

The performance of NERCOME is first illustrated with a toy example, based on uncorrelated, Gaussian distributed data with standard error 5 for the first half of the data vector, and 1 for the second half. The resulting spectrum of eigenvalues is shown in Fig. 1 for the standard sample covariance and NERCOME. Since the matrix of eigenvectors is orthogonal and thus always well conditioned, all ill-conditioning due to noise shows up in the eigenvalues. This is apparent for the sample covariance with small eigenvalues decreasing and large ones

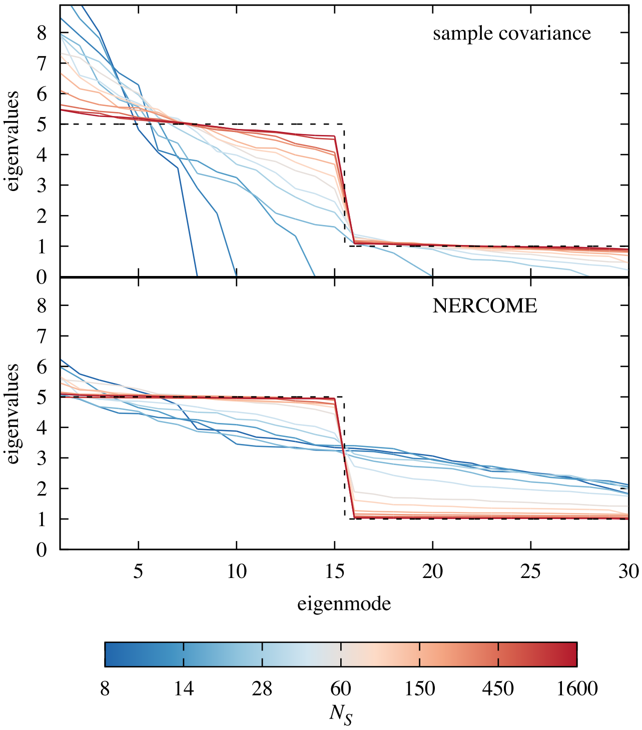

Figure 1. Spectrum of eigenvalues of an $N_\mathrm{D} = 30$-dimensional covariance estimate for the toy example of uncorrelated data with standard errors of 1 and 5 for 50 per cent of the data set, respectively (see black dashed line). Coloured curves result for different numbers of realizations of the data, N_S. Top panel: sample covariance estimates. Bottom panel: NERCOME estimates.

[1] Non-parametric Eigenvalue-Regularized COvariance Matrix Estimator.

[2] Latin indices denote different realizations, while Greek indices cycle through the elements of the data vector.

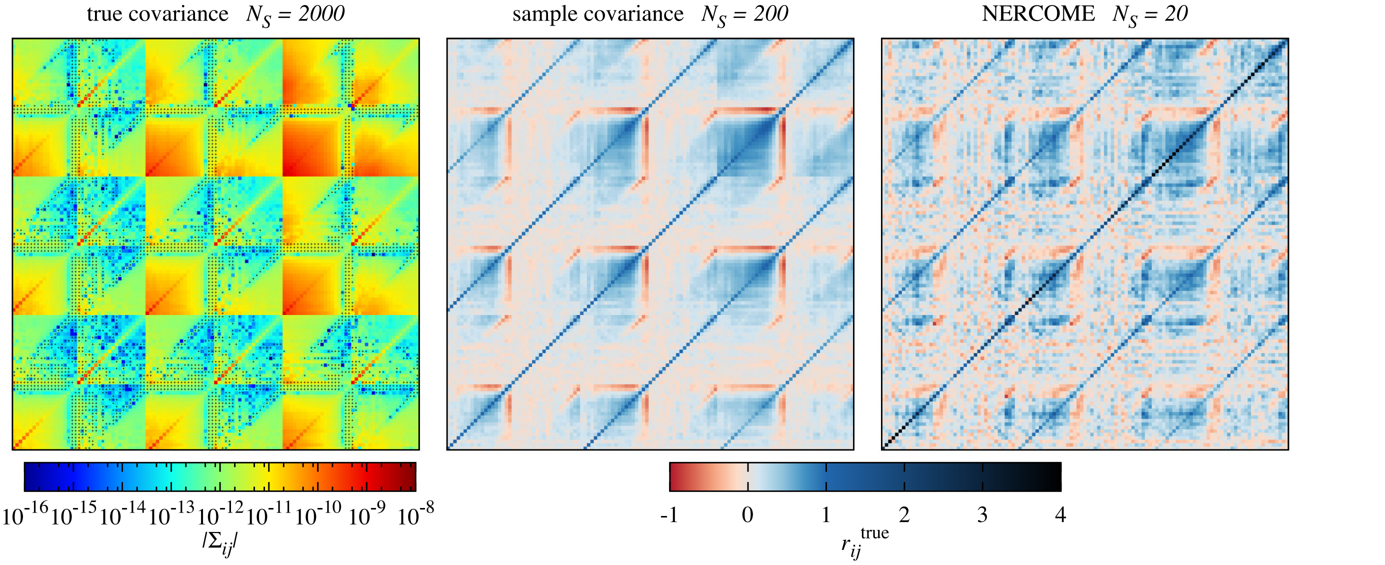

Figure 2. Left-hand panel: 'true' covariance (sample covariance determined from $N_\text{S} = 2000$ realizations). Shown are the absolute values of the covariance elements, with negative elements indicated by the additional black markers. The block structure reflects the use of three tomographic redshift bin combinations, with ξ_+ and ξ_- calculated for each combination. Middle panel: sample covariance for $N_\text{S} = 200$, normalized by the diagonal elements of the true covariance (see equation 4). Right-hand panel: same as the middle panel but for the NERCOME estimate with $N_\text{S} = 20$.

strongly increasing as the number of realizations decreases. Once $N_\text{S} < N_\text{D} + 2$, at least one eigenvalue vanishes so that the covariance becomes singular (e.g. Taylor & Joachimi 2014).

NERCOME shrinks both excessively large and small eigenvalues back towards the true values, avoiding singular values altogether. It can be shown (Lam 2016) that NERCOME estimates are positive definite with probability 1 (i.e. the exceptions constitute a set of measure zero) and consistent (approaching $\boldsymbol{\Sigma}$ for $N_\text{S} \rightarrow \infty$). The shrinkage is non-linear in that different eigenvalues are shrunk by different amounts (see Pope & Szapudi 2008 for an illustration of linear shrinkage). NERCOME consistently overestimates the smallest eigenvalues for low values of N_S, a feature that is also present in the following more realistic example.

4 SIMULATION SETUP

A realistic and challenging performance test is provided by the covariance of the two-point correlation functions $\xi_\pm$ of cosmic weak lensing, measured deeply into the non-linear regime of structure formation (see Kilbinger 2015 for a recent review). A large suite of simulated weak-lensing shear catalogues is created by producing coupled lognormal random fields from angular power spectra calculated for a vanilla flat Λ cold dark matter cosmology and assuming a minimum lensing convergence value of $\kappa_0 = -0.012$ (Hilbert, Hartlap & Schneider 2011). The redshift distribution of source galaxies is set to have a median of 0.8, and is split at the median into two tomographic bins. The mock survey is assumed to have an area of 25 deg^2 and a source galaxy number density of 10 arcmin^{-2} per tomographic bin, with a total galaxy ellipticity dispersion of 0.35. While the choice of survey area leads only to a rescaling of the covariance (note that the simulations have periodic boundary conditions and no masks), the number density determines the level of shot noise that contributes to the diagonal (and some sub-diagonals) of the covariance. Current and future surveys will choose their redshift binning such that at most a few galaxies per arcmin^2 will be in each bin, so that the choice above will lead to larger cross-correlations than expected in real-world applications.

From the 2000 mock realizations created in total, the shear correlation functions $\xi_\pm$ are measured for all three redshift–bin combinations in 20 angular bins, logarithmically spaced between 1 and 180 arcmin, with the tree code ATHENA (Kilbinger, Bonnett & Coupon 2014), constituting a total data vector of length $N_\text{D} = 120$. The resulting covariance is shown in Fig. 2 and is challenging for non-standard estimators in that it has a high condition number (~3000), a high level of correlation (numerous off-diagonal elements with correlation close to ±1; see the middle panel of Fig. 2) and a complex structure, including several discontinuities.

5 PERFORMANCE

The eigenspectra for the simulated covariance, shown in Fig. 3, are qualitatively similar to the toy case; NERCOME estimates are consistent, remain positive definite and display a positive bias across the spectrum. This is reflected in an overestimation of covariance elements, particularly the diagonal ones, which increases as N_S drops.

This is illustrated in the right-hand panel of Fig. 2, where a correlation matrix normalized with respect to the diagonal elements of the 'true' covariance (sample covariance for $N_\text{S} = 2000$),

$$r_{ij}^{\text{true}} \equiv C_{ij} / \sqrt{C_{ii}^{\text{true}}\, C_{jj}^{\text{true}}}\,, \tag{4}$$

is shown. For $N_\text{S} = 20$, diagonal elements are larger by a factor of up to 4. Otherwise, all main features in the correlation matrix have been reproduced, despite the small number of realizations (cf. the sample covariance in the centre panel). This trend persists irrespective of whether $\boldsymbol{n}$ is set to unity or to $\boldsymbol{m}$; hence, NERCOME is a poor choice of estimator if the covariance itself is the desired outcome of the estimation process.

In Fig. 4, the performance in terms of the signal-to-noise ratio, F, is shown. Under the assumption that the sample covariance is Wishart distributed, the distributions of its inverse and linear

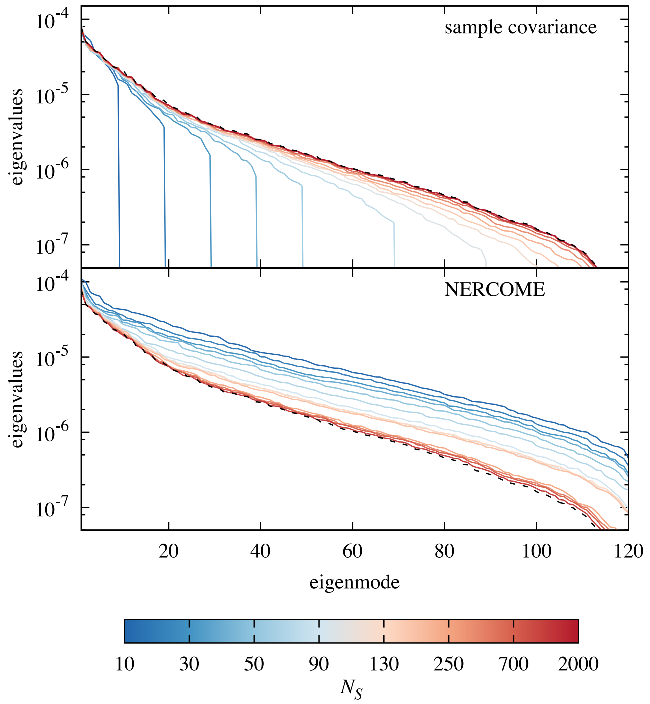

Figure 3. Same as Fig. 1 but for the cosmic weak-lensing covariance shown in Fig. 2. The 'true' spectrum as measured from the sample covariance with $N_\mathrm{S} = 2000$ realizations is given by the black dashed line. Note the logarithmic scaling of the ordinate axes.

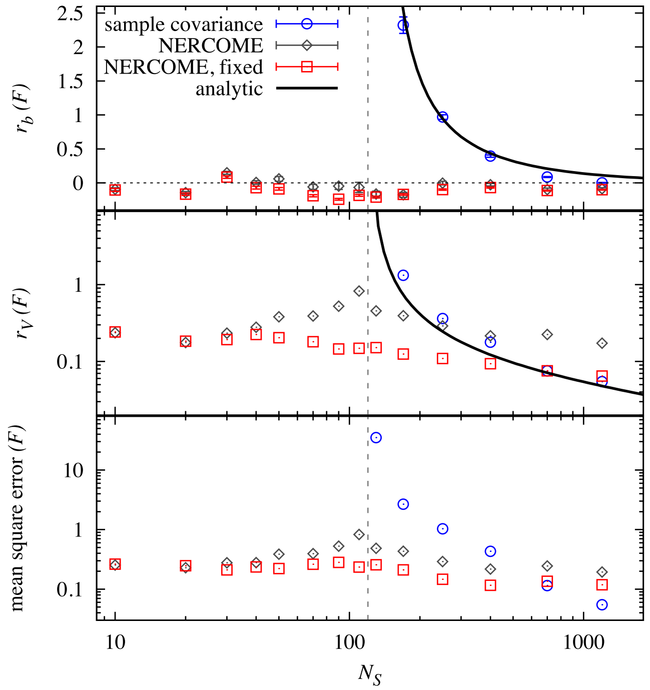

Figure 4. Relative bias (top panel), variance (centre panel) and mean square error (bottom panel) for the squared signal-to-noise ratio, F, as a function of the number of data realizations, N_S, used to estimate the covariance. Blue circles (grey diamonds, red squares) correspond to using the sample covariance (default NERCOME, NERCOME with a split fixed at $s/N_\mathrm{S} = 2/3$). The black solid curves show the analytic expectation for the sample covariance. The vertical grey dashed line indicates $N_\mathrm{S} = N_\mathrm{D} = 120$.

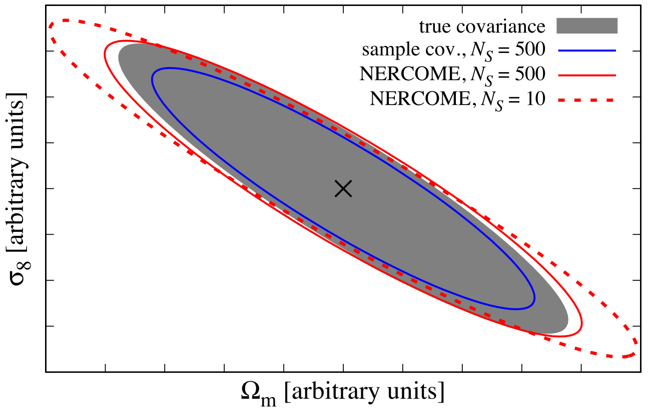

Figure 5. Comparison of parameter constraints in the Ω_m–σ_8 plane, using different covariance estimates in a Fisher matrix calculation, as indicated in the legend.

mappings thereof can be calculated analytically. It follows that an estimate of F derived from using the inverse of the sample covariance is distributed according to $\hat{F} \sim \mathrm{Inv}\text{–}\chi^2(F, N_\mathrm{S} - N_\mathrm{D})$ (Eaton 2007). This allows for the calculation of the relative bias and rms noise of F, given by

$$r_\mathrm{b} \equiv \frac{\langle \hat{F} \rangle}{F} - 1 = \frac{N_\mathrm{S} - 1}{N_\mathrm{S} - N_\mathrm{D} - 2}\;; \qquad (5)$$

$$r_V \equiv \frac{\sqrt{\left\langle \left(\hat{F} - \langle \hat{F} \rangle \right)^2 \right\rangle}}{F} = \frac{\sqrt{2}\,(N_\mathrm{S} - 1)}{\sqrt{N_\mathrm{S} - N_\mathrm{D} - 4}\,(N_\mathrm{S} - N_\mathrm{D} - 2)}\;; \qquad (6)$$

see Taylor & Joachimi (2014) for the details regarding moment calculations. For a given N_S, the mean and variance of $\hat{F}$ entering r_b and r_V are calculated via a delete-one jackknife. The analytic predictions agree excellently with the simulation results, validating the approach. Both bias and variance diverge as $N_\mathrm{S} \to N_\mathrm{D}$, with $\hat{\mathbf{S}}$ moving ever closer to becoming singular. The standard NERCOME estimator performs very well, displaying a small, marginally significant bias over the range of N_S probed and a relative rms error that peaks around $N_\mathrm{S} \approx N_\mathrm{D}$ and reduces to ~0.2 for both large and small N_S. For $N_\mathrm{S} > 400$, the variance of $\hat{F}$ via NERCOME surpasses that using $\hat{\mathbf{S}}$, which could be beaten down by increasing N_av, but the sample covariance estimator is the more efficient choice in this regime anyway. The location of the split fluctuates typically between $s/N_\mathrm{S} = 0.5$ and 1, as N_S varies, and tends to larger values in the regime where r_V peaks. An alternative run of NERCOME with s/N_S fixed at 2/3, also shown in Fig. 4, returns a slightly smaller and almost constant mean square error, at the price of a somewhat larger bias contribution for some values of N_S. Overall, NERCOME down to $N_\mathrm{S} = 10$ (cf. Fig. 2, right-hand panel) is competitive with the sample covariance estimator at $N_\mathrm{S} \sim 500$ in terms of the mean square error.

To assess the impact on cosmological posteriors, the different covariance estimates are inserted into a Fisher matrix forecast (following Tegmark et al. 1997) for the two most strongly constrained parameters for the toy weak-lensing survey outlined above, Ω_m and σ_8, with results shown in Fig. 5. For $N_\mathrm{S} = 500$, NERCOME is very close to the result for the 'true' covariance, while the sample covariance underestimates the width of the posterior, corresponding to an overestimate of F. The latter bias can largely be removed by applying the analytic correction of the mean bias, as proposed by (Hartlap et al. (2007) (cf. Fig. 4, top panel). The $N_\mathrm{S} = 10$ NERCOME estimate accurately reproduces the degeneracy direction and the size of the minor axis of the posterior ellipse but biases the major axis moderately high. Note that the choice of normalization $\boldsymbol{n}$ in this work is optimal for F but not necessarily so for elements of the Fisher matrix.

6 CONCLUSIONS

This work marks the first application of a non-linear shrinkage estimator of covariance in an astrophysical context, using a realistic example of a tomographic cosmic weak-lensing analysis that features a high condition number, high levels of correlation and a complex structure of the covariance matrix. After rescaling with a model of the data vector, the NERCOME estimator is able to estimate a function of the inverse covariance that has the same form as a Gaussian log-likelihood with subdominant bias and with variance that scales only mildly with the number of realizations of the data vector. Well-conditioned covariance estimates in the regime of much fewer realizations than data points are readily achieved. Compared to the standard sample covariance estimator, a factor of 50 less realizations are required to achieve the same mean square error, without any assumptions on the statistical properties of the data or the form of the covariance matrix, beyond those made for the sample covariance estimator (such as independently and identically distributed data).

The NERCOME estimator is consistent and almost surely positive definite. Its algorithm is simple, mainly consisting of eigenvalue decompositions, and trivially parallelizable in the subsequent averaging and split optimization steps. On a single core of a standard UNIX work station, NERCOME takes ∼6 s of wall-clock time per split location for the 120-dimensional data vector and setup considered in this work, so it adds a negligible runtime to an analysis pipeline. A downside is the lack of a priori control over the bias of the estimator, which depends on the details of the structure of the covariance matrix, an issue which is shared with most alternative estimators of covariance (e.g. Pope & Szapudi 2008; Paz & Sánchez 2015). However, fast approximate simulations based on random fields (see e.g. Xavier, Abdalla & Joachimi 2016) are well suited to assess these biases and optimize the free parameters of the estimators. Further work is required to assess the impact of these alternative covariance estimators on the fidelity of posteriors. A first cursory test with NERCOME revealed a moderate bias for very low $N_{\rm S}$ along the least constrained direction in parameter space.

Further suppression of adverse noise effects from covariance estimation can be achieved by explicitly incorporating prior information, which can range from the assertion of smoothness in the elements of (sub-)matrices to physically motivated effective models of the full covariance with a small level of residual degrees of freedom (see Mandelbaum et al. 2013; O'Connell et al. 2015; Pearson & Samushia 2016, for recent applications). A combination of such covariance modelling with the principles of non-linear shrinkage estimation as employed in NERCOME is a promising avenue. Likewise, a combination of shrinkage with resampling techniques directly applied to the data could potentially obviate the need for simulated data altogether, provided the challenges of bias induced by long-range spatial correlations can be overcome (Norberg et al. 2009; Friedrich et al. 2016). Since a NERCOME-like estimator does not set any requirements on the structure of the data vector (such as smoothness or a notion of distance), it can also readily be combined with the compression of the data vector as a preprocessing step (see Taylor et al. 2013, and references therein, as well as Zablocki & Dodelson 2016 for a recent example).

An implementation of NERCOME in C is made available with this publication.[3]

ACKNOWLEDGEMENTS

The author thanks C. Lam for helpful discussions and the anonymous referee for a constructive report. BJ acknowledges support by an STFC Ernest Rutherford Fellowship, grant reference ST/J004421/1.

REFERENCES

Abadir K. M., Dinasto W., Žikeš F., 2014, J. Econometrics, 181, 165
Dodelson S., Schneider M. D., 2013, Phys. Rev. D, 88, 063537
Eaton L. M., 2007, Multivariate Statistics: A Vector Space Approach. Institute of Mathematical Statistics, Beachwood, OH:
Escoffier S. et al., 2016, preprint (arXiv:1606.00233)
Friedrich O., Seitz S., Eifler T. F., Gruen D., 2016, MNRAS, 456, 2662
Hartlap J., Simon P., Schneider P., 2007, A&A, 464, 399
Hilbert S., Hartlap J., Schneider P., 2011, A&A, 536, 85
Kilbinger M., 2015, Rep. Prog. Phys., 78, 086901
Kilbinger M., Bonnett C., Coupon J., 2014, Astrophysics Source Code Library, record ascl:1402.026
Lam C., 2016, Ann. Stat., 44, 928
Ledoit O., Wolf M., 2012, Ann. Stat., 40, 1024
Mandelbaum R., Slosar A., Baldauf T., Seljak U., Hirata C. M., Nakajima R., Reyes R., Smith R. E., 2013, MNRAS, 432, 1544
Norberg P., Baugh C. M., Gaztanaga E., Croton D. J., 2009, MNRAS, 396, 19
O'Connell R., Eisenstein D., Vargas M., Ho S., Padmanabhan N., 2015, MNRAS, 462, 2681
Padmanabhan N., White M., Zhou H. H., O'Connell R., 2016, MNRAS, 460, 1567
Paz D. J., Sánchez A. G., 2015, MNRAS, 454, 4326
Pearson D. W., Samushia L., 2016, MNRAS, 457, 993
Percival W. J. et al., 2014, MNRAS, 439, 2531
Petri A., Haiman Z., May M., 2016, Phys. Rev. D, 93, 063524
Pope A. C., Szapudi I., 2008, MNRAS, 389, 766
Schneider M. D., Cole S., Frenk C. S., Szapudi I., 2011, ApJ, 737, 11
Sellentin E., Heavens A. F., 2016, MNRAS, 456, L132
Simpson F. et al., 2016, Phys. Rev. D, 93, 023525
Taylor A., Joachimi B., 2014, MNRAS, 442, 2728
Taylor A., Joachimi B., Kitching T., 2013, MNRAS, 432, 1928
Tegmark M., Taylor A. N., Heavens A. F., 1997, ApJ, 480, 22
Xavier H. S., Abdalla F. B., Joachimi B., 2016, MNRAS, 459, 3693
Zablocki A., Dodelson S., 2016, Phys. Rev. D, 93, 083525

[3] http://www.star.ucl.ac.uk/∼joachimi/publications.html

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