DES-2025-0886
FERMILAB-PUB-25-0362-PPD

Draft version April 29, 2026
Typeset using LaTeX **default** style in AASTeX7

Dimensional reduction for sampled priors and application to photometric redshift distributions

G. M. Bernstein, W. d'Assignies, M. A. Troxel, A. Alarcon, A. Amon, G. Giannini And B. Yin
et al.
(DES Collaboration)

(Published 1 May 2026, ApJ)

ABSTRACT

A typical Bayesian inference on the values of some parameters of interest $\mathbf{q}$ from some data D involves running a Markov Chain (MC) to sample from the posterior $p(\mathbf{q},\mathbf{n}|D) \propto \mathcal{L}(D|\mathbf{q},\mathbf{n})p(\mathbf{q})p(\mathbf{n})$, where $\mathbf{n}$ are some nuisance parameters with separable prior. In some cases, the nuisance parameters are high-dimensional, and their prior $p(\mathbf{n})$ is itself defined only by a set of samples that have been drawn from some other MC. The MC for the posterior will typically require evaluation of $p(\mathbf{n})$ at arbitrary values of $\mathbf{n}$, *i.e.* one needs to provide a density estimator over the full $\mathbf{n}$ space from the provided samples. But the high dimensionality of $\mathbf{n}$ hinders both the density estimation and the efficiency of the MC for the posterior. We describe a solution to this problem: a linear compression of the $\mathbf{n}$ space into a much lower-dimensional space $\mathbf{u}$ which projects away directions in $\mathbf{n}$ space that cannot appreciably alter $\mathcal{L}$. The algorithm for doing so is a slight modification to principal components analysis, and is less restrictive on $p(\mathbf{n})$ than other proposed solutions to this issue. We demonstrate this "mode projection" technique using the analysis of 2-point correlation functions of weak lensing fields and galaxy density in the *Dark Energy Survey*, where $\mathbf{n}$ is a binned representation of the redshift distribution $n(z)$ of the galaxies.

1. MOTIVATION

Consider an inference in which we have a vector of observable summary statistics $\mathbf{c}$ that we are using to constrain a set of parameters of interest $\mathbf{q}$. There is a model $\bar{\mathbf{c}}(\mathbf{q},\mathbf{n})$ for the expectation value of the observables which involves the parameters of interest, but also a vector $\mathbf{n}$ of nuisance parameters. We wish to characterize the Bayesian posterior density

$$p(\mathbf{q}|\mathbf{c}) \propto \int dn\, \mathcal{L}(\mathbf{c}|\mathbf{q},\mathbf{n})p(\mathbf{q})p(\mathbf{n}), \tag{1}$$

where $\mathcal{L}(\mathbf{c}|\mathbf{q},\mathbf{n})$ is a known likelihood function of the data, and we assume that the priors on $\mathbf{q}$ and $\mathbf{n}$ are separable to $p(\mathbf{q})$ and $p(\mathbf{n})$.[1] This posterior is complex enough that it requires approximation by the output of a Markov Chain (MC) wandering across the space $(\mathbf{q},\mathbf{n})$. Implicit in Equation (1) is that the prior $p(\mathbf{n})$ is independent of the likelihood $\mathcal{L}$ for $\mathbf{c}$, *e.g.* the prior on $\mathbf{n}$ has been constrained using data that are distinct from those entering the likelihood.

The scenario we address here is when *the prior $p(\mathbf{n})$ is not available in evaluable form, rather we have only a set of samples of $\mathbf{n}$ known to be drawn from this distribution.* Most MC samplers require that the posterior (and hence the prior and likelihood) be an evaluable function of any value of the parameters. It is the general task of density estimators to convert the samples of $\mathbf{n}$ into an evaluable $p(\mathbf{n})$. But when $\mathbf{n}$ is of high dimension, two problems arise: first, there may be insufficient available samples to create a viable density estimator; second, sampling of the posterior in (1) becomes infeasible if the MC must traverse a high-dimensional space.

A concrete example, which motivated this paper's work, is when the observable data $\mathbf{c}$ are the binned 2-point correlation functions of cosmic fields derived from a catalog of galaxies; the parameters of interest are cosmological quantities such as the matter density Ω_m, the amplitude of density fluctuations σ_8, etc.; and the nuisance parameters

[1] Following the nomenclature for the elements of Bayes's formula in, *e.g.*, https://en.wikipedia.org/wiki/Posterior_probability.

$\mathbf{n}$ include the coefficients of some linear expansion of the redshift distributions $n(z)$ of the galaxies being observed:

$$n(z) = \sum_{k=1}^{N} n_k b_k(z). \tag{2}$$

The b_k are a set of predetermined basis functions for the redshift distribution, *e.g.* these are boxcar functions if we are modeling $n(z)$ as stepwise constant. In our case of analyzing the data from the *Dark Energy Survey* (DES), there are 10 distinct samples of galaxies—each designed to prefer galaxies in a limited range of redshift—which we can index by s. Each has its own $n_s(z)$ to be characterized by coefficients n_{sk} at ≈ 100 values of k, leading to $N = O(1000)$ parameters n_{sk} to be considered. The vector $\mathbf{n}$ of nuisance parameters would be the concatenation of all the n_{sk}. For clarity, we will still write this as $\mathbf{n} = \{n_1, n_2, \ldots, n_N\}$ and demonstrate the method with a single galaxy sample's $n(z)$.

One approach would be to run a new MC over $\mathbf{q}$ for each of the samples we have of $\mathbf{n}$, and then concatenate these to effect marginalization over $\mathbf{n}$. This is clearly infeasible if thousands or more of $\mathbf{n}$ samples are needed to characterize the prior in this space.

Facing this problem for the cosmological analyses of the 3-year data (Y3) from DES, Cordero et al. (2022) devised a scheme whereby the samples of $\mathbf{n}$, which we write as $\{\mathbf{n}_\alpha\}$ for $\alpha \in 1 \ldots N_{\rm samp}$, are assigned to equally-spaced grid points within some M-dimensional unit hypercube $\mathcal{H}$. The coordinates $\mathbf{u}$ within the hypercube are considered the compressed parameters of $n(z)$, and the decompression function $\hat{\mathbf{n}}(\mathbf{u})$ outputs the $\mathbf{n}_\alpha$ sample at the nearest grid point to any $\mathbf{u}$, *i.e.* nearest-neighbor interpolation. In the example application described in Section 3, this procedure would assign an $N = 80$-dimensional $\mathbf{n}$ to each grid point placed in an $M = 3$-dimensional hypercube. This solves the problem of creating a continuous $\mathbf{u}$ domain, and gives each sample $\mathbf{n}_\alpha$ equal probability under $p(\mathbf{u})$, maintaining the meaning of the input samples. But the *output* of the function $\hat{\mathbf{n}}(\mathbf{u})$, and the resultant likelihood function $\mathcal{L}(\mathbf{c}|\mathbf{q}, \mathbf{u})$, are discontinuous over $\mathbf{u}$. Various strategies are proposed by Cordero et al. (2022) to assign the $\mathbf{n}_\alpha$ to the grid points in $\mathcal{H}$ based on summary statistics, to reduce the discontinuities—but the function is never smooth. As a consequence, many MC samplers become quite inefficient in sampling of the cosmological posterior. In particular, samplers such as MULTINEST that assume continuity are rendered nearly non-functional. As a result, the Y3 cosmological priors could not be evaluated with this method. Instead, the $\mathbf{n}$ samples were not used, and an *ad hoc* $p(\mathbf{n})$ was adopted which allowed only shifts and dilations of the mean $n(z)$ of the $\mathbf{n}$ samples [see Equation (31)].

A more rigorous and extremely efficient method of marginalizing over high-dimensional nuisance parameters was described by Bridle et al. (2002) and reprised by Hadzhiyska et al. (2020) for the $n(z)$ application, for the case where the following restrictions apply:

1. The likelihood of the observable $\mathbf{c}$ is normal, $\mathbf{c} \sim \mathcal{N}(\bar{\mathbf{c}}, C_c)$, with C_c fixed.
2. The prior $p(\mathbf{n})$ can also be assumed to be normal, with a mean taken to be $\bar{\mathbf{n}} = \langle \mathbf{n} \rangle$ and covariance matrix taken to be $C_n = \left\langle (\mathbf{n} - \bar{\mathbf{n}})(\mathbf{n} - \bar{\mathbf{n}})^T \right\rangle$ using the samples of $\mathbf{n}$ we are given.
3. The model $\bar{\mathbf{c}}$ can be linearized in $\mathbf{n}$ about fiducial values $\mathbf{q}_0, \mathbf{n}_0$ without loss of accuracy exceeding measurement errors, with the derivatives independent of $\mathbf{q}$.

Under these conditions, the marginalization over $\mathbf{n}$ is shown to be algebraically equivalent to adding terms to C_c, such that any MC process need not sample $\mathbf{n}$ at all.

We describe here an approach that is algebraically similar to this analytic marginalization, but does not require the 2nd condition of Gaussianity for the nuisance prior. Our approach is to seek a linear compression of $\mathbf{n}$ into a lower-dimensional set of parameters $\mathbf{u}$ that projects away variations in $\mathbf{n}$ that do not influence the likelihood $\mathcal{L}$. Standard density estimators can then be applied to the $\mathbf{u}$ values implied by the known $\mathbf{n}$ samples to yield a prior $p(\mathbf{u})$ that can be used for the MC of the cosmological posterior. The model $\bar{\mathbf{c}}$, and hence $\mathcal{L}$, will be continuous over this low-dimensional $\mathbf{u}$ space, and marginalization over $\mathbf{u}$ will yield posterior probabilities very close to marginalization over the original $\mathbf{n}$.

2. DERIVATION

We assume that we do have a multivariate normal likelihood for the observables $\mathbf{c}$ with the mean being some model $\bar{\mathbf{c}}(\mathbf{q}, \mathbf{n})$ and a fixed covariance matrix C_c, known *a priori* via analytic calculations and/or jacknifing of the data or other methods. In this section we will assume that the $\mathbf{n}$ vectors have been shifted by the mean of the samples $\mathbf{n}_0 \equiv \langle \mathbf{n}_\alpha \rangle$, generating a replacement set of $\mathbf{n}_\alpha$ that have zero mean. We seek some function $\hat{\mathbf{n}}(\mathbf{u})$ of a much lower-dimensional vector $\mathbf{u}$ which can be substituted for $\mathbf{n}$ and yield nearly the same likelihood function for any $\mathbf{n}$ in the

domain spanned by the samples $\{\mathbf{n}_\alpha\}$. This means we want maps $\mathbf{n}_\alpha \to \mathbf{u}_\alpha \to \hat{\mathbf{n}}_\alpha$, with $\dim(\mathbf{n}) = \dim(\hat{\mathbf{n}}) = N$ and $\dim(\mathbf{u}) = M \ll N$. We wish to find maps such that replacing $\mathbf{n}$ with $\hat{\mathbf{n}}$ alters the cosmological inference by much less than the other uncertainties in the model or data. Inferences use the likelihood of the observed $\mathbf{c}$ under the assumed Gaussian probability distribution:

$$-2\log\mathcal{L}(\mathbf{c}|\mathbf{q},\mathbf{n}) = |2\pi C_c| + [\mathbf{c} - \bar{\mathbf{c}}(\mathbf{q},\mathbf{n})]^T C_c^{-1} [\mathbf{c} - \bar{\mathbf{c}}(\mathbf{q},\mathbf{n})]\,. \tag{3}$$

While ideally the loss function for a compression scheme would be some direct measure of the biases induced in the inferred parameters $\mathbf{q}$, we will instead target a minimization of the expected level of mis-estimation of the likelihood function $\mathcal{L}$. The posterior credible region for $\mathbf{q}$ at fixed $\mathbf{c}$ will correspond to a range of $\Delta(2\log\mathcal{L}) \approx 1$ (unless dominated by priors.) Thus we can be confident that a compression that changes the inferred $\log\mathcal{L}$ by $|\Delta\log\mathcal{L}| \ll 1$ will change inferred $\mathbf{q}$ values by a small fraction of the measurement uncertainty. It is certainly true that as $\langle\Delta\log\mathcal{L}\rangle \to 0$, the substitution of $\hat{\mathbf{n}}$ for $\mathbf{n}$ has no effect on the calculated posterior density of $\mathbf{q}$. We do not, however, offer any formal proof or bound on the relation between $\log\mathcal{L}$ errors and $\mathbf{q}$ errors induced by compression.

If the compression scheme replaces $\mathbf{n}$ with $\hat{\mathbf{n}}$, then we will assign an incorrect $\hat{\mathcal{L}}$

$$-2\log\hat{\mathcal{L}} = |2\pi C_c| + [\mathbf{c} - \bar{\mathbf{c}}(\mathbf{c},\mathbf{n}) - \Delta_c]^T C_c^{-1} [\mathbf{c} - \bar{\mathbf{c}}(\mathbf{q},\mathbf{n}) - \Delta_c]\,, \tag{4}$$

$$\Delta_c \equiv \bar{\mathbf{c}}(\mathbf{q},\hat{\mathbf{n}}) - \bar{\mathbf{c}}(\mathbf{q},\mathbf{n}). \tag{5}$$

Expanding (4) and subtracting (3) yields the compression error in the likelihood:

$$\Delta(-2\log\mathcal{L}) = \Delta_c^T C_c^{-1}\Delta c + 2\,[\mathbf{c} - \bar{\mathbf{c}}(\mathbf{q},\mathbf{n})]^T C_c^{-1}\Delta_c. \tag{6}$$

If the model $\bar{\mathbf{c}}(\mathbf{q},\mathbf{n})$ properly describes the mean of the observations $\mathbf{c}$, then the expectation value of the rightmost term under the observational noise is zero. A useful measure of the typical mis-estimation of $\mathbf{q}$ induced by the compression of $\mathbf{n}$ is therefore the term $\Delta_c^T C_c^{-1}\Delta_c$. Marginalizing this quantity over the (sampled) prior for $\mathbf{n}$ yields the loss function that we will minimize during compression:

$$\langle\chi^2\rangle = \frac{1}{N_{\rm samp}}\sum_\alpha [\bar{\mathbf{c}}(\mathbf{q},\mathbf{n}_\alpha) - \bar{\mathbf{c}}(\mathbf{q},\hat{\mathbf{n}}_\alpha)]^T C_c^{-1} [\bar{\mathbf{c}}(\mathbf{q},\mathbf{n}_\alpha) - \bar{\mathbf{c}}(\mathbf{q},\hat{\mathbf{n}}_\alpha)]\,. \tag{7}$$

This quantity is the χ^2 of the difference between the original and compressed models for the data, marginalized over the prior on $\mathbf{n}$ as represented by the samples we are provided. It is also the mean-squared distance in the space $\mathbf{c}$ between the model generated by $\mathbf{n}$ and that by $\hat{\mathbf{n}}$, using the observations' covariance matrix C_c as a metric for the distance.[2]

The use of Equation (7) implicitly assumes that $N_{\rm samp}$ is sufficiently large that averaging χ^2 over the samples is equivalent to integrating over the entire $\mathbf{n}$ space. One might be concerned that a compression scheme could overtrain on Equation (7) in the sense of yielding small χ^2 at the sample points $\mathbf{n}_\alpha$, but have large χ^2 at $\mathbf{n}$ values between the samples. If, however, we use a simple linear projection for compression to M dimensions, there are NM free parameters, whereas the input samples have $NN_{\rm samp}$ values to operate on. If $N_{\rm samp} \gg M$, we assert without proof that the compression function does not have enough freedom to overtrain except in pathological cases of sampling. But our scenario already assumes that the samples are sufficient to conduct marginalization over $\mathbf{n}$ when estimating $p(\mathbf{c}|\mathbf{q},\mathbf{n})$, which means they must be numerous ($N_{\rm samp} \gg N$) and span the space of plausible $\mathbf{n}$ values under $p(\mathbf{n})$.

We next assume that the compression is linear, $\hat{\mathbf{n}} = X\mathbf{n}$, for some $N \times N$ matrix that is idempotent ($XX = X$). We will further linearize the dependence of $\bar{\mathbf{c}}$ on $\mathbf{n}$, specifically assuming that (in scalar notation)

$$\frac{\partial^2\bar{c}}{\partial n^2}n \ll \frac{\partial\bar{c}}{\partial n} \tag{8}$$

over the full range of variation of $\mathbf{n}$. With these two assumptions, Equation (7) becomes

$$\langle\chi^2\rangle = \frac{1}{N_{\rm samp}}\sum_\alpha [(I - X)\mathbf{n}_\alpha]^T F\,[(I - X)\mathbf{n}_\alpha]\,. \tag{9}$$

[2] More precisely, the inverse of the covariance matrix is the metric.

We use the Jacobian matrix of the model $\bar{\mathbf{c}}$ to define

$$F \equiv \left[\frac{\partial \bar{\mathbf{c}}}{\partial \mathbf{n}}\right]^T_{\mathbf{q}_0,\mathbf{n}_0} C_c^{-1} \left[\frac{\partial \bar{\mathbf{c}}}{\partial \mathbf{n}}\right]_{\mathbf{q}_0,\mathbf{n}_0}. \tag{10}$$

This quantity is also the Fisher matrix giving the information provided by the observations $\mathbf{c}$ about the nuisance parameters $\mathbf{n}$ under a Gaussian likelihood (Tegmark et al. 1997). In many cases this matrix will be rank-deficient and/or poorly conditioned, since the observables are not likely to be very informative on $\mathbf{n}$—if they were, we might not be concerned with establishing a prior on $\mathbf{n}$ to begin with. Fortunately, we will not require the inverse of F in our algorithm.

Our assumptions that the likelihood is normal in $\bar{\mathbf{c}}$ and that we can linearize $\bar{\mathbf{c}}$ in $\mathbf{n}$ [equations (3) and (8)] can be replaced by the slightly more general condition that the Hessian of $\log \mathcal{L}$ with respect to $\mathbf{n}$ is well approximated as constant over the domain of interest of $(\mathbf{q}, \mathbf{n})$.

The optimization implied by Equation (9) is the same as in familiar Principal Components Analysis (PCA), aside from the presence of the F matrix, which in essence defines a new metric for the variance to be captured by the principal components. Our solution will follow the typical derivation for PCA, but with an additional variable transformation to compensate for the presence of F.

Since X is idempotent, we can write

$$X = V_X P_M V_X^T, \tag{11}$$

where V_X is unitary and the projection matrix P_M is defined as

$$(P_M)_{ij} \equiv \begin{cases} 1, & i = j \leq M \\ 0, & \text{otherwise} \end{cases} \tag{12}$$

It is also useful to define

$$Y \equiv I - X = V_X P_{-M} V_X^T, \tag{13}$$

with $P_{-M} = I - P_M$.

For a chosen rank M of the transformation matrix X, our task becomes to identify the eigenvectors V_X that minimize

$$\left\langle \chi^2 \right\rangle = \frac{1}{N_{\text{samp}}} \sum_\alpha \left[Y \mathbf{n}_\alpha\right]^T F \left[Y \mathbf{n}_\alpha\right] \tag{14}$$

$$= \text{Tr}\left[C_n V_X P_{-M} V_X^T F V_X P_{-M} V_X^T\right], \tag{15}$$

$$C_n \equiv \left\langle \mathbf{n}\mathbf{n}^T \right\rangle. \tag{16}$$

This optimization is easier if we first transform the systematic variables to $\mathbf{n}' = T\mathbf{n}$ such that $C_{n'} = I$, *i.e.* make the elements of $\mathbf{n}'$ uncorrelated and unit-variance. This is accomplished by finding the eigensystem $C_n = V_n \Lambda_n V_n^T$ and setting $T = \Lambda_n^{-1/2} V_n^T$. With this transformation, we are now seeking a different unitary matrix $V_{X'}$ that minimizes

$$\left\langle \chi^2 \right\rangle = \text{Tr}\left[I V_{X'} P_{-M} V_{X'}^T \left[(T^{-1})^T F T^{-1}\right] V_{X'} P_{-M} V_{X'}^T\right] \tag{17}$$

$$= \text{Tr}\left[P_{-M} V_{X'}^T G V_{X'} P_{-M}\right], \tag{18}$$

where we have defined the transformed Fisher matrix

$$G \equiv \left(T^{-1}\right)^T F T^{-1} = \Lambda_n^{1/2} V_n^T F V_n \Lambda_n^{1/2} = V_G \Lambda_G V_G^T. \tag{19}$$

The right-hand side defines the eigensystem of G, with $\Lambda_G = \text{diag}(\lambda_1^G, \ldots, \lambda_N^G)$, and $\lambda_i^G \geq 0$. Equation (18) can now be rewritten as

$$\left\langle \chi^2 \right\rangle = \sum_{i>M} \left(V \lambda_G V^T\right)_{ii} \tag{20}$$

where $V = V_{X'}^T V_G$ is unitary. This expression must be at least as large as sum of the $N - M$ smallest λ_i^G, and that minimum is attained if $V_{X'}^T V_G = I \;\Rightarrow\; V_{X'} = V_G$, and the eigensystem of G is placed in order of decreasing eigenvalues λ_i^G. The elements surviving the projection P_{-M} yield

$$\left\langle \chi^2 \right\rangle = \sum_{i>M} \lambda_i^G. \tag{21}$$

In other words each eigenvalue of the matrix G in Equation (19) gives the contribution to $\langle\chi^2\rangle$ of one projection (mode) of $\mathbf{n}$.

Transforming the solution back into the space of $\mathbf{n}$ yields

$$X = T^{-1}V_G P_M V_G^T T \tag{22}$$
$$= \left[V_n^T \Lambda_n^{1/2} V_G P_M\right] \left[P_M V_G^T \Lambda_n^{-1/2} V_n^T\right] \tag{23}$$
$$\equiv DE. \tag{24}$$

We thus obtain our optimal encoding/compression using the nonzero rows of matrix E to give

$$\mathbf{u}_\alpha = E\mathbf{n}_\alpha \tag{25}$$

and the decoding/reconstruction of the systematic variables as

$$\hat{\mathbf{n}}_\alpha = D\mathbf{u}_\alpha. \tag{26}$$

One can confirm that this procedure yields a compressed representation $\mathbf{u}$ such that $C_u = I_M$, the M-dimensional identity matrix.

The previous derivation ignores the possibility that C_n is singular or nearly so, such that taking $\Lambda_n^{-1/2}$ in Equation (23) is not possible. Indeed in our application, it is *required* that C_n be singular, because we have a sum normalization constraint on the initial $\mathbf{n}_\alpha$ values. Any such (nearly) zero element j of Λ_n has a corresponding eigenvector $\mathbf{v}_j$ of the $\mathbf{n}$ space which has zero amplitude in all of the input samples $\mathbf{n}_\alpha$, so that the $\mathbf{n}$ are confined to a subspace—therefore the reconstructed $\hat{\mathbf{n}}$ should also be. The compressed representations $\mathbf{u}_\alpha$ and reconstructed $\hat{\mathbf{n}}_\alpha$ should therefore be unaffected by the presence of any $\mathbf{v}_j$ component. This can be accomplished in Equation (23) by setting element j of $\Lambda_n^{-1/2}$ to zero, as is typically done during solutions of least-squares problems using singular value decompositions.

In summary, the procedure for dimensional reduction is:

1. Obtain the mean $\mathbf{n}_0$ of the input samples, the Fisher matrix F of the system defined in Equation (10) using derivatives about $\mathbf{n}_0$, plus the covariance matrix C_n of the samples.

2. From the eigensystem (Λ_n, V_n) of C_n, form the matrix G defined in Equation (19) and get its eigensystem (Λ_G, V_G). Place the eigenvalues in descending order.

3. Choose the size M of the compressed representation to be the minimum that keeps the $\langle\chi^2\rangle$ value in Equation (21) below a chosen threshold, presumably $\ll 1$. If this is achieved, the multiplicative errors in the likelihood induced by the compression are at percent levels.

4. Form the encoding matrix E and decoding matrix D as in Equation (23), taking the inverse $\Lambda_n^{-1/2}$ to be zero for any eigenvalues that are zero (or within roundoff errors).

5. Compress all incoming (mean-subtracted) samples using Equation (25). The resulting $\mathbf{u}$ values will have have unit covariance matrix and zero mean.

6. Construct a continuous density estimator in $\mathbf{u}$ space that mimics the finite sample distribution. If the distribution is normal, this becomes the multidimensional unit normal, since the construction yields $\mathrm{Cov}(u) = I_M$.. There is, however, no *a priori* reason that this must be the case, and something like a normalizing flow may be needed to approximate this lower-dimensional representation of the prior.

7. Sample over $\mathbf{u}$ space in the Markov Chain that is sampling the posterior on the parameters of interest $\mathbf{q}$, using Equation (26) to transform each sample back into a $\hat{\mathbf{n}}$ vector.

The ability to accomodate non-Gaussian distributions of the nuisance-parameter space is the principle advantage of our method over the single-step covariance-inflation method of Bridle et al. (2002) and Hadzhiyska et al. (2020), *i.e.* one can drop assumption (2) of the three listed for this method in the Introduction. But the compression method is also more robust in other ways: it defines projection of $\mathbf{n}$ values onto an M-dimensional linear subspace. We can in

fact also drop assumptions (1) of Gaussian likelihood and (3) of linearization of the model $\bar{\mathbf{c}}(\mathbf{q},\mathbf{n})$ that were used to derive the subspace, if we can instead demonstrate that: *no component of* $\mathbf{n}$ *normal to the subspace can significantly alter the likelihood of the data, when* $\mathbf{n}$ *and* $\mathbf{q}$ *range across their regions of significant probability.* If this statement is true, then marginalization over the compressed subspace will yield inferences consistent with the (usually infeasible) marginalization over the full space of $\mathbf{n}$, regardless of the nature of the model and likelihood functions.

Even in the case where all three assumptions of the covariance-inflation method hold true, there are practical advantages of compressing the nuisance variables and retaining them in the cosmological Markov chain instead of doing the analytic marginalization. One can, for example, examine the posterior distribution of $\mathbf{u}$ to see how the data have altered the prior distribution of $\mathbf{u}$. If the posterior for $\mathbf{u}$ is at the edge of the prior, this would potentially indicate an inconsistency between the data and the prior.

3. APPLICATION

As an example of the application of this straightforward dimensional reduction to a high-dimensional nuisance parameter, we examine the redshift distribution of one of the bins of "Maglim" galaxies used as a lens population and clustering tracer in the Year 6 (Y6) analysis of the DES galaxy catalogs. Each of these Maglim "lens bins" is selected with cuts on galaxy fluxes and colors in an attempt to generate a sample that is confined to a particular redshift range, as described in Weaverdyck et al. (2026). Once each bin's galaxy selection criteria are chosen, a combination of photometric techniques (Giannini et al. 2025; Yin et al. 2025) and clustering information (d'Assignies et al. 2025) is used to generate samples from the posterior distribution of $n(z)$ functions applicable to the bin members, conditioned on the photometric and clustering data, *i.e.* the $p(\mathbf{n})$ that will become the prior for the inference of $\mathbf{q}$.

In the simplest case, there are 6 cosmological parameters of interest, $\mathbf{q} = \{\Omega_m, \Omega_b, \sigma_8, h, n_s, m_\nu\}$. The nuisance vector $\mathbf{n}$ has ≈ 5 parameters in each of the following categories: galaxy biases with respect to matter; intrinsic alignments of galaxy shapes with the tidal field of the mass; magnification coefficients; and multiplicative errors in the measurement of galaxy shear, for a total of 24 non-redshift parameters. The $n(z)$ for each galaxy bin is described by Equation (2) with 80 coefficients spanning $0 < z < 4$ at intervals of $\Delta z = 0.05$. If all of these coefficients, for each of the 10 galaxy selections, were allowed to vary, the parameter space for inference would grow from 24 to 824 dimensions. The inference would become infeasible even if a reliable density estimator over the 80-dimensional space could be created from the $O(10^4)$ samples available to characterize each $n(z)$.

We hence turn to the modal projection technique herein to reduce the dimensionality to those directions in $\mathbf{n}$ space in which the samples span a large enough range to alter the $\bar{\mathbf{c}}(\mathbf{n})$ at levels comparable to the precision of the observations.

The observable quantities are the autocorrelation $w(\theta)$ of the galaxies' angular positions as a function of angular separation θ; and the cross-correlations $\gamma_t^{(s)}(\theta)$ between these galaxies' positions and the weak gravitational lensing shear observed from groups $1 \le s \le 4$ of source galaxies. The vector $\mathbf{c}$ is the concatenation of values of w and γ_t in a set of bins of θ. The model $\bar{\mathbf{c}}(\mathbf{q},\mathbf{n})$ for the observables consists of integrations over redshift z of products of the redshift distributions $n(z)$ with solutions of General Relativity differential equations for the expansion history and growth of density fluctuations in the Universe. The covariance matrices C_c are even more complex. The derivations of both are described in full detail in Sanchez-Cid et al. (2026). We use the same `Cosmosis` software (Zuntz et al. 2015) employed in that paper to calculate the partial derivatives needed in deriving the compression scheme.

We present results for mode-compression sampling of the $n(z)$ parameters for redshift bin 4 of the DES Y6 lens galaxies. The orange violins in Figure 1 plot the distributions of the individual elements of $n(z)$ in the input 3000 samples. As per the procedure described in this paper, the mean $\mathbf{n}$ is subtracted from each sample and the covariance C_n computed. This is then combined with the derivatives and covariance matrices of the observables $\mathbf{c}$ to calculate the decomposition in Equation (19).

The left plot in Figure 2 shows the values of unmodelled χ^2 vs the dimension M of the compressed space, as per Equation (21). This modelling error induced by compression drops exponentially with the number of retained modes. We have somewhat arbitrarily chosen a threshold of $\chi^2 < 0.025$ for each galaxy sample in order to keep the total impact of compression of 10 samples to $\ll 1$. This is attained with $M = 3$ modes for this bin. The right side of Figure 2 plots the individual modes $\mathrm{U}_i(z)$, *i.e.* the rows of the decompression matrix D such that $\hat{\mathbf{n}} = \bar{\mathbf{n}} + \sum_{i \le M} u_i \mathbf{U}_i$. Recall that each mode's coefficient u_i will be a random deviate with unit variance. Each of the first three modes appears to effect some combination of a z shift of the main $n(z)$ peak, a change of the peak's shape/width, and a change in the low-z contamination. We have also plotted mode 7 in the Figure, to illustrate a mode of variation that is present in the

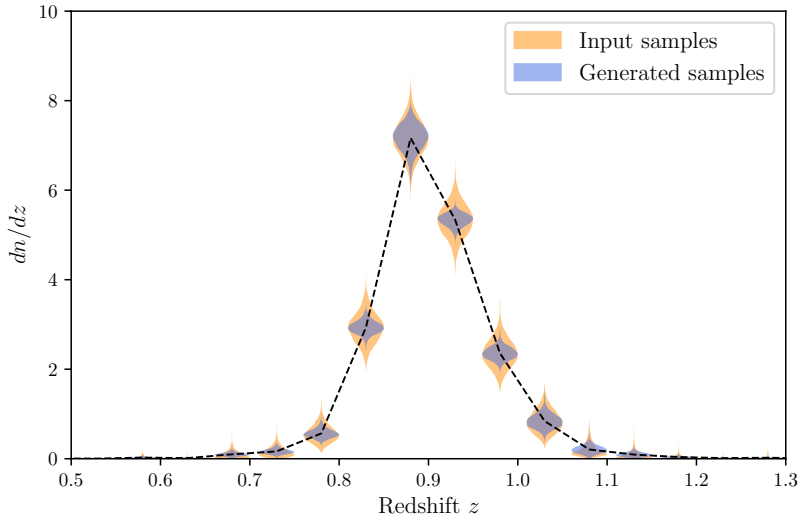

Figure 1. Violin plots for the redshift probability distribution $n(z)$ of galaxies in lens bin 4 for the DES Y6 analysis. The orange regions show the distributions for the samples of $\mathbf{n}$ derived from photometric and clustering information. The blue violins are for $\mathbf{n}$ values drawn defined by (1) subtracting the mean $\bar{\mathbf{n}}$; (2) compressing these $\mathbf{n}$ into 3 modes with coefficients $\mathbf{u}$; (2) drawing values of $\widetilde{u}_i$ from unit normal distributions; (3) transforming each component $\widetilde{u}_i$ to match the 1d distribution of the input samples' u_i; (4) decompressing the transformed u_i draws back into full-length $\mathbf{n}$ samples; finally (5) restoring the mean $\bar{\mathbf{n}}$. The dashed line connects the mean values of $n(z)$, which are the same for generated samples as for the input samples, by construction. The compression substantially lowers the variance of $n(z)$ at individual values of z without significantly altering the variation of cosmological signals that the entire $n(z)$ predicts. [Although the $n(z)$ functions are calculated for $0 < z < 4$, we truncate this plot to emphasize the redshift regime where this bin's galaxies are primarily found.]

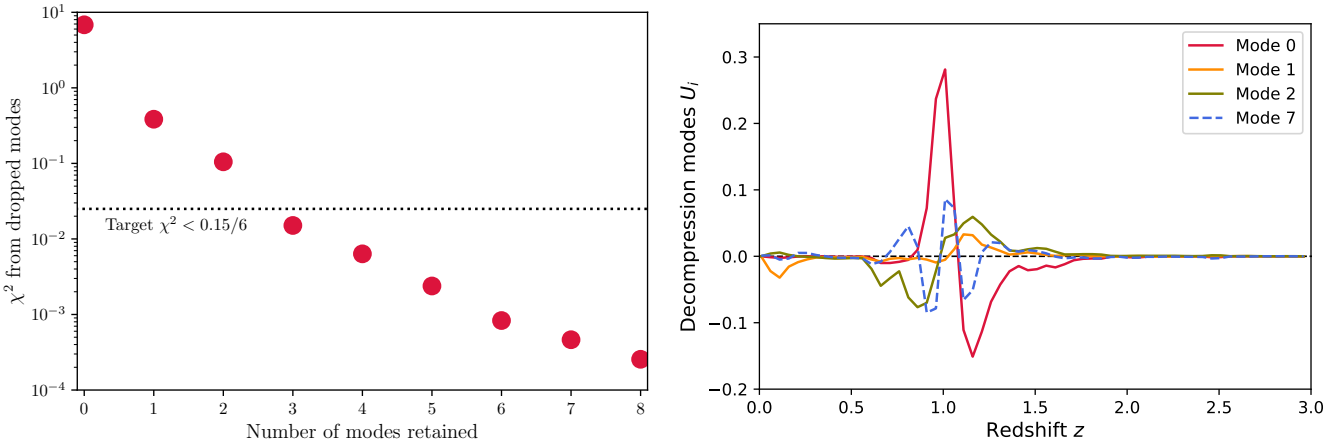

Figure 2. At left: The size of the χ^2 of modeling error attributable to compressing the $\mathbf{n}$ samples down to M modes is plotted vs M. The $M = 0$ point shows the modelling error from holding $n(z)$ fixed at its mean, and $M \geq 1$ values drop exponentially as we use more modes to reconstruct $n(z)$. Our chosen criterion of $\chi^2 < 0.025$ is attained with $M = 3$ for this bin's $n(z)$. At right: The modes of variation $U_i(z)$, *i.e.* the rows of the decompression matrix D in Equation (23), are plotted vs redshift. Each of modes 0,1,2 is multiplied by a unit-variance stochastic coefficient u_i, then they are summed with the mean $\bar{\mathbf{n}}(z)$, to form an $n(z)$ sample. Higher-numbered modes have observable consequences of decreasing statistical significance. Mode 7 is plotted as the dashed blue line as an example of what is projected out of $n(z)$; even though its typical amplitude in the input data is larger than modes 1 or 2, its oscillatory behavior does not lead to measurable changes in the cosmological statistics.

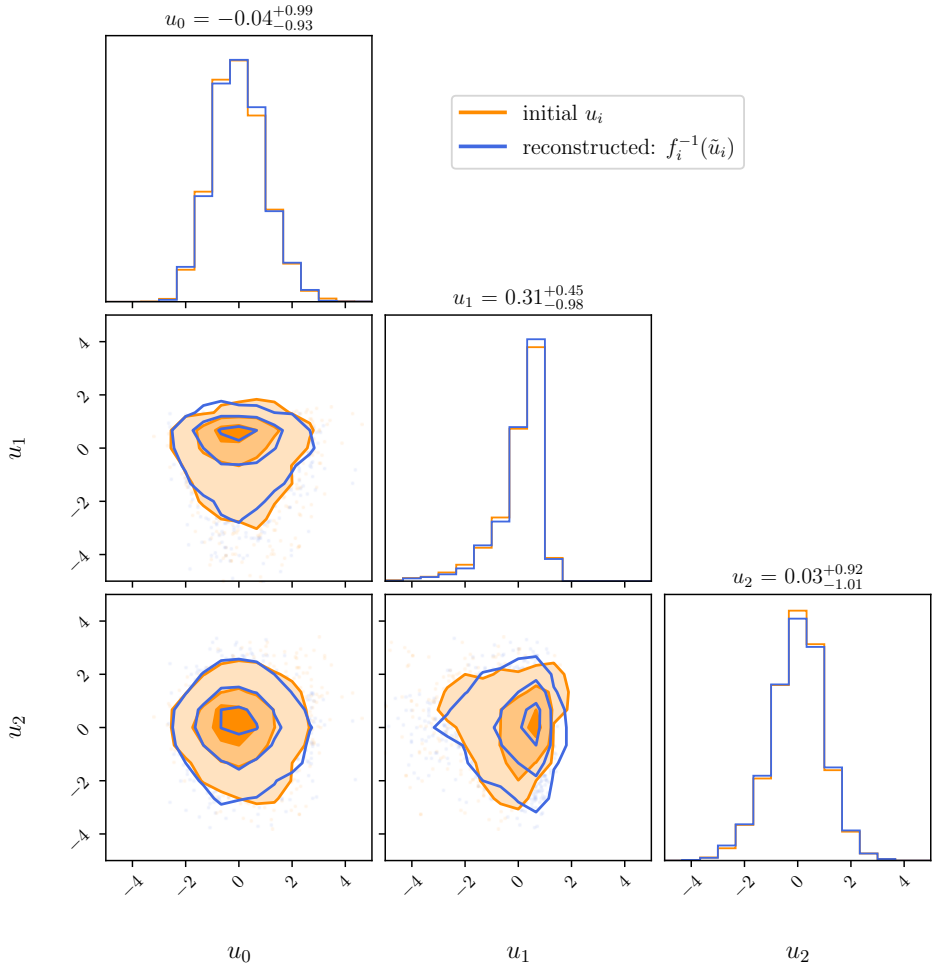

Figure 3. In orange is a corner plot of the distribution of the mode coefficients, *i.e.* the elements of $\mathbf{u} = E\mathbf{n}$. of the input samples after encoding. The coefficients, especially u_1, are significantly skewed so a normal distribution would be an inaccurate model. Instead we model each u_i as a "denormalizing" function of a unit-normal variable, as per Equation (27). The blue histograms and contours show the distributions obtained using this element-by-element transformation technique, which is seen to accurately reproduce the distribution of the input samples.

samples, but has unobservable consequences. Mode 7 is more oscillatory in redshift than the three retained modes, and does not alter the low-z tail.

Figure 3 plots the distributions of the $M = 3$ components of the compressed $\mathbf{u}$ representations of the 3000 input samples, *i.e.* the vectors $\mathbf{u}_j = E(\mathbf{n}_j - \bar{\mathbf{n}})$ obtained from each input sample $\mathbf{n}_j$. Some of the components of $\mathbf{u}$ have substantially non-Gaussian distributions, so a unit-normal prior on $\mathbf{u}$ will not faithfully represent the input samples—a better density estimator is required. We find in this case (and in all other DES cases) that it is sufficient to normalize the marginal distributions of the individual u_i components. This is done by tabulating a normalizing function f_i for each mode defined by

$$\mathrm{CDF}_n\left[f_i(u_i)\right] = \mathrm{CDF}(u_i), \tag{27}$$

where the left side is the cumulative distribution function of the unit normal, and the right-hand side is the CDF of the u_i values obtained from the input samples. The functions f_i are bijective and can be stored as a splined lookup table. Now, the cosmological Markov Chain is told that there are 3 parameters in $\tilde{\mathbf{u}}$ that have a unit-normal prior $p(\tilde{\mathbf{u}})$. This vector defines the redshift distribution via a 2-step process:

$$u_i = f_i^{-1}(\tilde{u}_i); \tag{28}$$

$$\mathbf{n} = D\mathbf{u}. \tag{29}$$

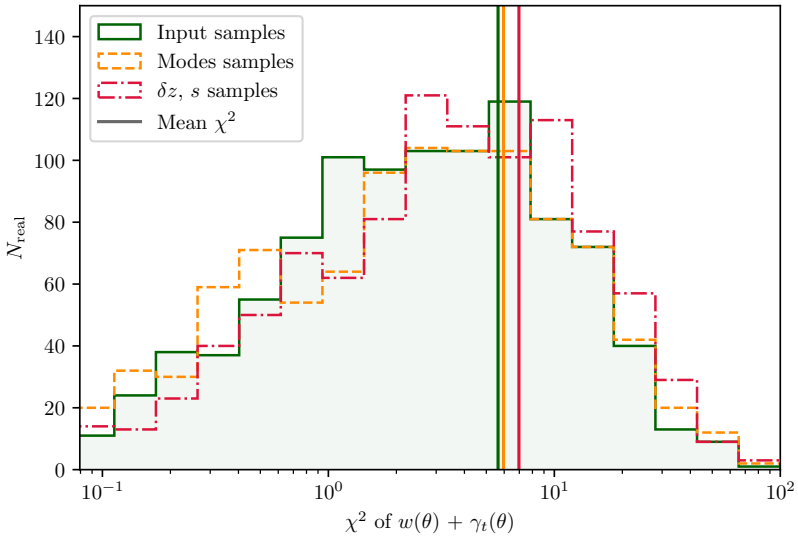

Figure 4. The histograms show the deviations of the predicted observable $w(\theta)$ quantities in DES Y6 cosmological analysis, as measured by the χ^2 in Equation (30), as we allow the $n(z)$ parameters to vary. The shaded green histogram shows the variation using the original 3000 samples of $n(z)$ produced by the photometric and clustering redshift studies. The dashed yellow histogram results from drawing 3-dimensional $\widetilde{\mathbf{u}}$ values from a unit normal and decompressing them into $n(z)$ realizations using our method. This lower-dimensional model for the prior on $n(z)$ reproduces the original samples' result very well. By comparison, the dash-dotted red histogram generates samples of $n(z)$ using the *ad hoc* method of Equation (31). This method's two parameters z and s are given priors to match the distributions of the mean of $n(z)$ over z, and the standard deviation of z, present in the input samples. The *ad hoc* method produces $n(z)$ fluctuations with $\approx 20\%$ larger χ^2 from the mean, on average, than the input distribution has. [Note the logarithmic horizontal axis.]

The blue violins in Figure 1 show the distributions of the $n(z)$ values that are generated by drawing $\widetilde{\mathbf{u}}$ values from a unit Gaussian. The means of the two distributions (centers of the violins) are, by construction, identical. The blue, compression-derived distributions are shorter (narrower distributions) than the orange input distributions, which is a result of projecting away modes of $\mathbf{n}$ that have no measurable influence on the observable $\mathbf{c}$ values. Removing modes reduces the variance of $n(z)$ at any particular z, by varying degrees at different values of z. The compression scheme acts as noise reduction on the $n(z)$ estimates, where "noise" means a fluctuation that does not detectably impact measurable quantities. Projecting away the unobservable fluctuations in $n(z)$ has substantially reduced the variance of the function at any *individual* z value, but the *collective* $n(z)$ behavior still retains the same observable influence on the summary statistics $\mathbf{c}$.

To check whether we have achieved our goal of leaving the observable consequences of the $n(z)$ variation unchanged, we calculate the distribution of

$$\chi^2 = [\bar{\mathbf{c}}(\mathbf{q},\mathbf{n}) - \bar{\mathbf{c}}(\mathbf{q},\mathbf{n}_0)]^T \, C_c^{-1} \, [\bar{\mathbf{c}}(\mathbf{q},\mathbf{n}) - \bar{\mathbf{c}}(\mathbf{q},\mathbf{n}_0)] \tag{30}$$

for the cases when (1) the $\mathbf{n}$ are drawn from the input samples, vs (2) are generated using the procedure defined above. This χ^2 measures the deviation of the model from that implied by the mean vector $\mathbf{n}_0$ at some chosen nominal value of cosmological parameters $\mathbf{q}$ (we also fix the other nuisance parameters of the DES model for this test).[3]

Figure 4 shows the results: the distribution of the χ^2 values defined by Equation (30) using the input samples is indistinguishable from the distribution of this χ^2 using samples of $\mathbf{n}$generated by our compressed, normalized representation. Note that this agreement indicates that the approximation of linearity of the model in Equation (8) is sufficiently accurate in this application to yield an accurate compression scheme. The Figure also shows the χ^2 distribution resulting from samples generated by the method used in the DES Y3 analysis (Myles et al. 2021). In that

[3] Please note that this χ^2 quantity is generated from nuisance parameter samples, not by any process that should reproduce the standard χ^2 distribution.

case, the $n(z)$ for each galaxy sample was given an *ad hoc* variation of the form

$$n(z) = n_0 \left(\frac{z - \Delta z}{s} \right), \tag{31}$$

with Δz and s being "shift" and "stretch" parameters. Separable Gaussian priors were assigned to Δz and s with means of 0 and 1, respectively, and standard deviations that equaled the RMS variation in the mean and width of the input $n(z)$ samples. The shift-stretch model essentially compresses the samples of $n(z)$ into two parameters, their mean and standard deviation, and executes an *ad hoc* reconstruction based on those parameters. In the Figure we can see that the shift-stretch model does not in fact reproduce the size of the deviations in the $\mathbf{c}$ observables that is implied by the original samples; the mean χ^2 induced by the shift-stretch samples is $\approx 20\%$ larger than the input samples. This mismatch indicates that some cosmologically relevant information in the original samples is being lost, or that the shift-stretch samples are imposing cosmological constraints that are not present in the original samples.

For simplicity, we have demonstrated this method for a case in which $\mathbf{n}$ specifies the $n(z)$ function for a single sample of DES galaxies. The method is fully applicable to any set of nuisance parameters for which we are given a set of samples from $p(\mathbf{n})$. For instance, in the DES Y6 analysis, we use an $\mathbf{n}$ that is a concatenation of the parameters of $n(z)$ for 4 bins of lens source galaxies. Since each output sample from the redshift-estimate process describes all 4 bins, we know the cross-correlations between distinct bins' $n(z)$ values, and the compression process described herein will properly preserve these correlations in the subsequent cosmological analysis. In the DES Y6 analysis, we have also extended the nuisance parameter vector $\mathbf{n}$ to include the multiplicative errors on the shear measurement method, and create a compressed $\mathbf{u}$ that captures correlations between the multiplicative errors and the $n(z)$ estimates.

4. SUMMARY

We present a linear dimensionality reduction technique that has the aim of making it feasible to produce continuous density estimators for nuisance parameters that have no evaluable prior, and are characterized only from a set of samples in a high-dimensional space $\mathbf{n}$. This is essentially a principle components analysis that is adjusted to separate contributions to the detectable consequences of $\mathbf{n}$ rather than contributions to the Euclidean norm of $\mathbf{n}$. This method enables a rigorous marginalization over the distribution of $\mathbf{n}$ even for non-Gaussian distributions, as long as the derivatives of the log-likelihood of the data with respect to $\mathbf{n}$ are nearly constant over the posterior domain. The compression method is also robust to any nonlinearities in the model and non-Gaussian likelihoods that lie *within* the subspace of $\mathbf{n}$ defined by the compression, as long as $\mathbf{n}$ components normal to this subspace do not have detectable effects on the data model. The mode-projection technique is essentially a method of projecting away irrelevant fluctuations in $\mathbf{n}$. The algebra to derive the compression scheme that minimizes the mean shift in computed likelihood $\langle \chi^2 \rangle$ makes the assumption that the Hessian $\partial^2 \log \mathcal{L} / \partial \mathbf{n}^2$ is constant in the parameter ranges of interest, but in fact the successful application of this compression to inference requires only the weaker assumption that fluctuations in $\mathbf{n}$ that are normal to the M-dimensional compression plane have negligible impact on $\log \mathcal{L}$.

This technique was developed for the case when $\mathbf{n}$ represents the redshift distributions of galaxies in DES. Some attempts at sampling the posterior $p(\mathbf{q}|\mathbf{c})$ while fully respecting the prior $p(\mathbf{n})$ as embodied by the samples $\{\mathbf{n}_\alpha\}$ have proven infeasible, such as concatenating MC's run at every $\mathbf{n}_\alpha$, or the nearest-neighbor hypercube embedding of Cordero et al. (2022). The method of Bridle et al. (2002) that propagates $p(\mathbf{n})$ into an inflated C_c relies on assumptions of Gaussianity for the nuisances that are not necessarily valid here, and precludes some forms of internal consistency checks. The mode-projection technique we describe is able to do a better job, with principled adherence to the prior embodied by the input samples. Extensive use of this method on the 10 different galaxy samples defined for the Year 6 analysis of DES data is reported by Yin et al. (2025); Weaverdyck et al. (2026) and d'Assignies et al. (2025).

While this new approach to marginalizing over $n(z)$ has minimal consequence for the posterior cosmological parameter estimates in the DES Y6 analysis, the mode-projection method is better motivated and will produce more accurate posteriors in future experiments where the $n(z)$ prior's uncertainty is a dominant contributor to the cosmological posterior, and can be generally applicable to other analyses where critical nuisance parameters do not have evaluable priors.

The compression technique is applicable to any inference with a high-dimensional prior known only through a set of samples. In the cosmological realm, one case of high-dimensional nuisance parameters might be a binned representation of the absolute response vs wavelength λ of some imaging bands used in constructing Type Ia supernova Hubble diagrams, as constrained by observations of a finite sample of spectrophotometric samples. Another case might be

detailed power spectrum $P_\delta(k, z)$ of density fluctuations created by gravity plus baryonic forces, as constrained by ensembles or jackknife samples of numerical simulations. This likelihood-aware PCA should be of use to inferences in a broad range of fields.

ACKNOWLEDGMENTS

G.M.B. acknowledges support from NSF grant AST-2205808 and DOE award DE-SC0007901. W. d'A. acknowledges support from the MICINN projects PID2019-111317GB-C32, PID2022-141079NB-C32 as well as predoctoral program AGAUR-FI ajuts (2024 FI-1 00692) Joan Oró. The project that gave rise to these results received the support of a fellowship to A. Alarcon from "la Caixa" Foundation (ID 100010434). The fellowship code is LCF/BQ/PI23/11970028. We also thank the anonymous referees for pointing out portions of the paper in need of improvement.

Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, the Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da Ciência, Tecnologia e Inovação, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey.

The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l'Espai (IEEC/CSIC), the Institut de Física d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, NSF NOIRLab, the University of Nottingham, The Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, Texas A&M University, and the OzDES Membership Consortium.

Based in part on observations at NSF Cerro Tololo Inter-American Observatory at NSF NOIRLab (NOIRLab Prop. ID 2012B-0001; PI: J. Frieman), which is managed by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation.

The DES data management system is supported by the National Science Foundation under Grant Numbers AST-1138766 and AST-1536171. The DES participants from Spanish institutions are partially supported by MICINN under grants PID2021-123012, PID2021-128989 PID2022-141079, SEV-2016-0588, CEX2020-001058-M and CEX2020-001007-S, some of which include ERDF funds from the European Union. IFAE is partially funded by the CERCA program of the Generalitat de Catalunya.

We acknowledge support from the Brazilian Instituto Nacional de Ciência e Tecnologia (INCT) do e-Universo (CNPq grant 465376/2014-2).

This document was prepared by the DES Collaboration using the resources of the Fermi National Accelerator Laboratory (Fermilab), a U.S. Department of Energy, Office of Science, Office of High Energy Physics HEP User Facility. Fermilab is managed by Fermi Forward Discovery Group, LLC, acting under Contract No. 89243024CSC000002.

Contributions: GB devised the compression method, wrote the relevant code, and wrote the article text. MT derived the relevant derivative matrices, and Wd'A applied the method to the DES Maglim bin and produced figures. Authors Wd'A, MT, AA, AA, GG, and BY developed the redshift samples used in the demonstration, integrated the code into DES pipelines, and advised on the implementation and text. The remaining authors have made contributions to this paper that include, but are not limited to, the construction of DECam and other aspects of collecting the data; data processing and calibration; developing broadly used methods, codes, and simulations; running the pipelines and validation tests; and promoting the science analysis.

REFERENCES

Bridle, S. L., Crittenden, R., Melchiorri, A., Hobson, M. P., Kneissl, R., and Lasenby, A. N. 2002, MNRAS, 335, 1193, doi:10.1046/j.1365-8711.2002.05709.x

Cordero, J. P., Harrison, I., Rollins, R. P., et al. 2022, MNRAS, 511, 2170. doi:10.1093/mnras/stac147

d'Assignies, W. et al. 2025, arXiv:2510.23566

Giannini, G. et al. 2025, arXiv:2509.07964

Hadzhiyska, B., Alonso, D., Nicola, A., et al. 2020, JCAP, 2020, 056. doi:10.1088/1475-7516/2020/10/056

Myles, J. et al. 2021, MNRAS, 505, 4249, doi:10.1093/mnras/stab1515

Sanchez-Cid et al. 2026, arXiv:2601.14859

Tegmark, M., Taylor, A. N., & Heavens, A. F. 1997, ApJ, 480, 1, 22. doi:10.1086/303939

Weaverdyck, N. et al. 2026, arXiv:2601.14484

Yin, B. et al. 2025, arXiv:2510.23566

Zuntz, J. et al. 2015, Astronomy and Computing, 12, 45, doi:10.1016/j.ascom.2015.05.005

All Authors and Affiliations

G. M. Bernstein [1], W. d'Assignies [2], M. A. Troxel [3], A. Alarcon [4], A. Amon [5], G. Giannini [6], B. Yin [3], M. Aguena [7], S. S. Allam [8], F. Andrade-Oliveira [9], D. Brooks [10], A. Carnero Rosell [11, 7, 12], J. Carretero [2], L. N. da Costa [7], M. E. S. Pereira [13], J. De Vicente [14], S. Everett [15], J. Frieman [16, 8, 6], J. García-Bellido [17], D. Gruen [18], S. R. Hinton [19], D. L. Hollowood [20], K. Honscheid [21, 22], D. J. James [23], S. Lee [24], J. L. Marshall [25], J. Mena-Fernández [26], R. Miquel [27, 2], A. A. Plazas Malagón [28, 29], E. Sanchez [14], D. Sanchez Cid [14, 9], I. Sevilla-Noarbe [14], T. Shin [30], M. Smith [31], E. Suchyta [32], M. E. C. Swanson [33], N. Weaverdyck [34, 35], J. Weller [36, 37] And P. Wiseman [38]

(DES Collaboration)

[1] *Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA*

[2] *Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra (Barcelona) Spain*

[3] *Department of Physics, Duke University Durham, NC 27708, USA*

[4] *Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, s/n, 08193 Barcelona, Spain*

[5] *Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544, USA*

[6] *Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA*

[7] *Laboratório Interinstitucional de e-Astronomia - LIneA, Av. Pastor Martin Luther King Jr, 126 Del Castilho, Nova América Offices, Torre 3000/sala 817 CEP: 20765-000, Brazil*

[8] *Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, IL 60510, USA*

[9] *Physik-Institut, University of Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland*

[10] *Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK*

[11] *Instituto de Astrofisica de Canarias, E-38205 La Laguna, Tenerife, Spain*

[12] *Universidad de La Laguna, Dpto. Astrofísica, E-38206 La Laguna, Tenerife, Spain*

[13] *Hamburger Sternwarte, Universität Hamburg, Gojenbergsweg 112, 21029 Hamburg, Germany*

[14] *Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain*

[15] *California Institute of Technology, 1200 East California Blvd, MC 249-17, Pasadena, CA 91125, USA*

[16] *Department of Astronomy and Astrophysics, University of Chicago, Chicago, IL 60637, USA*

[17] *Instituto de Fisica Teorica UAM/CSIC, Universidad Autonoma de Madrid, 28049 Madrid, Spain*

[18] *University Observatory, Faculty of Physics, Ludwig-Maximilians-Universität, Scheinerstr. 1, 81679 Munich, Germany*

[19] *School of Mathematics and Physics, University of Queensland, Brisbane, QLD 4072, Australia*

[20] *Santa Cruz Institute for Particle Physics, Santa Cruz, CA 95064, USA*

[21] *Center for Cosmology and Astro-Particle Physics, The Ohio State University, Columbus, OH 43210, USA*

[22] *Department of Physics, The Ohio State University, Columbus, OH 43210, USA*

[23] *Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA*

[24] *Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109, USA*

[25] *George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy, and Department of Physics and Astronomy, Texas A&M University, College Station, TX 77843, USA*

[26] *Université Grenoble Alpes, CNRS, LPSC-IN2P3, 38000 Grenoble, France*

[27] *Institució Catalana de Recerca i Estudis Avançats, E-08010 Barcelona, Spain*

[28] *Kavli Institute for Particle Astrophysics & Cosmology, P. O. Box 2450, Stanford University, Stanford, CA 94305, USA*

[29] *SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA*

[30] *Department of Physics and Astronomy, Stony Brook University, Stony Brook, NY 11794, USA*

[31] *Physics Department, Lancaster University, Lancaster, LA1 4YB, UK*

[32] *Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831*

[33] *Center for Astrophysical Surveys, National Center for Supercomputing Applications, 1205 West Clark St., Urbana, IL 61801, USA*

[34] *Department of Astronomy, University of California, Berkeley, 501 Campbell Hall, Berkeley, CA 94720, USA*

[35] *Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA*

[36] *Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse, 85748 Garching, Germany*

[37] *Universitäts-Sternwarte, Fakultät für Physik, Ludwig-Maximilians Universität München, Scheinerstr. 1, 81679 München, Germany*

[38] *School of Physics and Astronomy, University of Southampton, Southampton, SO17 1BJ, UK*