



Published in final edited form as:

*J Comput Graph Stat.* 2022 ; 31(2): 486–501. doi:10.1080/10618600.2021.1987254.

## Eigenvectors from Eigenvalues Sparse Principal Component Analysis (EESPCA)

H. Robert Frost

Department of Biomedical Data Science, Dartmouth College

### Abstract

We present a novel technique for sparse principal component analysis. This method, named Eigenvectors from Eigenvalues Sparse Principal Component Analysis (EESPCA), is based on the formula for computing squared eigenvector loadings of a Hermitian matrix from the eigenvalues of the full matrix and associated sub-matrices. We explore two versions of the EESPCA method: a version that uses a fixed threshold for inducing sparsity and a version that selects the threshold via cross-validation. Relative to the state-of-the-art sparse PCA methods of Witten et al., Yuan & Zhang and Tan et al., the fixed threshold EESPCA technique offers an order-of-magnitude improvement in computational speed, does not require estimation of tuning parameters via cross-validation, and can more accurately identify true zero principal component loadings across a range of data matrix sizes and covariance structures. Importantly, the EESPCA method achieves these benefits while maintaining out-of-sample reconstruction error and PC estimation error close to the lowest error generated by all evaluated approaches. EESPCA is a practical and effective technique for sparse PCA with particular relevance to computationally demanding statistical problems such as the analysis of high-dimensional data sets or application of statistical techniques like resampling that involve the repeated calculation of sparse PCs.

### Keywords

principal component analysis; sparse principal component analysis; sparse eigenvalue decomposition; eigenvector-eigenvalue identity

## 1 Introduction

### 1.1 Principal component analysis (PCA)

PCA is a widely used statistical technique that was developed independently by Karl Pearson (Pearson, 1901) and Harold Hotelling (Hotelling, 1933) in the early part of the 20th century. PCA performs a linear transformation of multivariate data into a new set

---

rob.frost@dartmouth.edu .

SUPPLEMENTARY MATERIAL

**R logic for simple running example:** Sweave generated file that contains all of the R logic needed to generate the results for the simple running example including results for SPC, SPC.lse, rifle and TPower. This example is also embedded as a vignette in the EESPCA R package.

**Supplementary results:** Simulation study results for the block covariance, high dimension and limited sparsity models and Gene Ontology enrichment results for the EESPCA.cv, SPC.lse, and rifle methods on the example scRNA-seq data.

of variables, the principal components (PCs), that are linear combinations of the original variables, are uncorrelated and have sequentially maximum variance (Jolliffe, 2002). PCA can be equivalently defined by the set of uncorrelated vectors that provide the best low-rank matrix approximation, in a least squares sense. The solution to PCA is given by the eigenvalue decomposition of the sample covariance matrix with the variance of the PCs specified by the eigenvalues and the PC directions defined by the eigenvectors.

Because PCA is defined in terms of the eigenvectors and eigenvalues of the sample covariance matrix, it is related to a wide range of matrix analysis methods and multivariate statistical techniques with an extremely large number of applications (Jolliffe, 2002; Jolliffe and Cadima, 2016). Although most commonly used for unsupervised linear dimensionality reduction and visualization, PCA has been successfully applied to statistical problems including regression (e.g., principal components regression (Hastie *et al.*, 2009)), clustering (e.g., use of PCs as input for clustering methods (Waltman and van Eck, 2013; Hafemeister and Satija, 2019)), and non-linear dimensionality reduction (e.g., seed directions for non-linear methods (McInnes *et al.*, 2018)). In the biomedical domain, PCA has been extensively employed for the analysis of genomic data including measures of DNA variation, DNA methylation, RNA expression and protein abundance (Ma and Dai, 2011). Common features of these datasets, and the motivation for eigenvalue decomposition methods, are the high dimensionality of the feature space (i.e., from thousands to over one million), comparatively low sample size (i.e.,  $p \gg n$ ) and significant collinearity between the features. The most common uses of PCA with genomic data involve dimensionality reduction for visualization or clustering (Stuart *et al.*, 2019), with population genetics an important use case (Patterson *et al.*, 2006). PCA has also been used as the basis for feature selection (Lu *et al.*, 2011), and gene clustering (Kluger *et al.*, 2003). More recent applications include gene set enrichment of bulk or single cell gene expression data (Tomfohr *et al.*, 2005; Fan *et al.*, 2016).

## 1.2 Mathematical notation for PCA

Let  $\mathbf{X}$  be an  $n \times p$  matrix that holds  $n$  independent samples drawn from a  $p$ -dimensional joint distribution with population covariance matrix  $\Sigma$ . Without loss of generality, we can assume the columns of  $\mathbf{X}$  are mean-centered. The unbiased sample covariance matrix is therefore given by  $\Sigma = 1/(n-1)\mathbf{X}^T\mathbf{X}$ . PCA can be performed via the direct eigenvalue decomposition of  $\Sigma$  with the eigenvalues  $\lambda_i$  equal to the PC variances and the unit length eigenvectors  $\mathbf{v}_i$  equal to the PC loadings. For computational reasons, PCA is more commonly performed via the singular value decomposition (SVD) of  $\mathbf{X}$ :  $\mathbf{X} = \mathbf{U}\mathbf{D}\mathbf{V}^T$ , where  $\mathbf{U}$  and  $\mathbf{V}$  are both orthonormal matrices (i.e.,  $\mathbf{U}^T\mathbf{U} = \mathbf{V}^T\mathbf{V} = \mathbf{I}$ ), the columns of  $\mathbf{V}$  represent the PC loading vectors, the entries  $d_i$  in the diagonal matrix  $\mathbf{D}$ , arranged in decreasing order, are proportional to the square roots of the PC variances ( $d_i = (n-1)\sqrt{\lambda_i}$ ) and the columns of  $\mathbf{U}\mathbf{D}$  are the principal components. Consistent with the optimal low-rank matrix approximation property of PCA, the first  $r$  components of the SVD minimize the squared Frobenius norm between the original  $\mathbf{X}$  and a rank  $r$  reconstructed version of  $\mathbf{X}$ . Specifically, if  $\mathbf{U}_r$ ,  $\mathbf{D}_r$  and  $\mathbf{V}_r$  hold the first  $r$  columns of  $\mathbf{U}$ ,  $\mathbf{D}$  and  $\mathbf{V}$ ,  $\mathbf{X}_r = \mathbf{U}_r\mathbf{D}_r\mathbf{V}_r^T$  is the SVD rank  $r$  reconstruction of  $\mathbf{X}$ , which minimizes  $\|\mathbf{X} - \mathbf{X}_r\|_F^2$  for all possible rank  $r$  reconstructions.

In the remainder of the manuscript we will use  $\|\mathbf{x}\|_0$  to refer to the  $l_0$  norm of vector  $\mathbf{x}$  (i.e., number of non-zero values),  $\|\mathbf{x}\|_2$  to refer to the Euclidean or  $l_2$  norm, and  $\mathbf{x} \odot \mathbf{y}$  to refer to the element-wise multiplication of vectors  $\mathbf{x}$  and  $\mathbf{y}$ . In general, we will use bold font for matrices and vectors and non-bold for scalars. Additional mathematical notation is defined when first used.

### 1.3 Running example

To help illustrate the concepts discussed in this paper, we introduce a simple example data set based on a 10-dimensional multivariate normal (MVN) distribution. The R logic needed to reproduce all of the results for this example can be found in the Supplementary Material. The population mean is set to the zero vector,  $\boldsymbol{\mu} = (0, 0, 0, 0, 0, 0, 0, 0, 0, 0)$ , and the population covariance matrix is given a two block covariance structure with a covariance of 0.5 among the first four variables and between the last two variables, and a covariance of 0 among all other variables. All population variances are set to one, which aligns with the common practice of performing PCA after standardization.

For this  $\boldsymbol{\Sigma}$ , the first population PC has equal non-zero loadings for just the first four variables and the second population PC has equal non-zero loadings for just the last two variables. If the matrix  $\mathbf{V}$  holds the population PC loadings, the first two columns are:

$$\mathbf{V}_{[:,(1,2)]}^T = \begin{bmatrix} -0.5 & -0.5 & -0.5 & -0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.7071068 & 0.7071068 \end{bmatrix}$$

The variances of these population PCs are  $\boldsymbol{\lambda} = (2.5, 1.5)$ .

For an example set of 100 independent samples drawn from this MVN distribution, the loadings of the first two sample PCs (rounded to three decimal places) are:

$$\widehat{\mathbf{V}}_{[:,(1,2)]}^T = \begin{bmatrix} -0.541 & -0.446 & -0.506 & -0.496 & -0.033 & 0.022 & 0.001 & -0.023 & -0.049 & -0.052 \\ 0.113 & -0.085 & -0.055 & -0.126 & 0.156 & -0.048 & 0.002 & -0.221 & 0.724 & 0.600 \end{bmatrix}$$

The variances of the first two sample PCs (again rounded to three decimal places) are  $\widehat{\boldsymbol{\lambda}} = (3.393, 1.521)$ . The minimum rank 2 reconstruction error for this case (computed as the squared Frobenius norm of the residual matrix) is 597.531 and the Euclidean distance between the first population PC and estimated PC is 0.108. Results for the SPC, SPC.1se, TPower and rifle methods detailed below on this simple running example can be found in the Supplementary Material.

### 1.4 Sparse PCA

As demonstrated by the simple example above, all variables typically make a non-zero contribution to each sample PC even when the data follows a statistical model with sparse population PCs. This property makes interpretation of the PCs challenging, especially when the underlying data is high dimensional. For example, PCA of gene expression data will generate PCs that are linear combinations of thousands of genes and attempting to ascertain the biological relevance of such a large linear combination of genes is a

very difficult task. The challenge of PC interpretation has motivated a large number of approaches for generating approximate PCs that have non-zero loadings for just a small subset of the measured variables. Such sparse PCA techniques include simple components (i.e., PC loading vectors constrained to values from  $\{-1, 0, 1\}$ ) (Vines, 2000), methods which compute approximate PCs using cardinality constraints (Moghaddam *et al.*, 2006; d'Aspremont *et al.*, 2007; Sriperumbudur *et al.*, 2011; Yuan and Zhang, 2013; Tan *et al.*, 2018), methods that use LASSO or elastic net-based penalties (Jolliffe *et al.*, 2003; Zou *et al.*, 2006; Shen and Huang, 2008; Witten *et al.*, 2009; Jung *et al.*, 2019), and methods based on iterative component thresholding (Ma, 2013). By generating approximate PCs with few non-zero loadings, all of these techniques improve interpretability by associating only a small number of variables with each PC. For the comparative evaluation of our proposed approach, we will focus on three current sparse PCA techniques: SPC (Witten *et al.*, 2009), TPower (Yuan and Zhang, 2013) and rifle (Tan *et al.*, 2018). These methods represent both the cardinality constraint and LASSO-penalization approaches and have been shown to provide state-of-the-art performance in both simulation studies and real data analyses.

The Witten *et al.* SPC method has an elegant formulation via LASSO-penalized matrix decomposition. Specifically, SPC modifies the standard SVD matrix reconstruction optimization problem to include a LASSO penalty on the PC loadings. Using the notation introduced in Section 1.2, the SPC optimization problem for the first PC can be stated as:

$$\min_{d, \mathbf{u}, \mathbf{v}} \|\mathbf{X} - d\mathbf{u}\mathbf{v}^T\|_F^2 \text{ subject to } \|\mathbf{u}\|_2^2 = 1, \|\mathbf{v}\|_2^2 = 1, \sum_{i=1}^n |\mu_i| < c, d > 0 \quad (1)$$

Due to the LASSO penalty on the components of  $\mathbf{u}$ , this optimization problem generates a sparse version of the first PC loadings vector with the level of sparsity controlled by the parameter  $c$ . Subsequent sparse PCs can be computed by iteratively applying SPC to the residual matrix generated by subtracting the rank 1 reconstruction  $d\mathbf{u}\mathbf{v}^T$  from the original  $\mathbf{X}$ . Witten *et al.* also outline a more complex approach for computing multiple PCs that ensures orthogonality. Similar to LASSO-penalized generalized linear models (Tibshirani, 2011), the penalty parameter  $c$  can be specified to achieve a desired level of sparsity in the PC loadings or can be selected via cross-validation to minimize the average out-of-sample matrix reconstruction error. Since the true level of sparsity is typically not known, the cross-validation approach is usually employed to determine  $c$  and we will use this approach in our comparative evaluation. The need to estimate the sparsity parameter via a technique like cross-validation is not unique to SPC and is a limitation shared by all existing sparse PCA methods. A common alternative to picking  $c$  to minimize the average out-of-sample reconstruction error is to set  $c$  to smallest value whose reconstruction error is within 1 standard error of the minimum error. This alternate version, which we will refer to as SPC.1se, generates a more sparse version of the PC loadings at the cost of increased reconstruction error. The SPC and SPC.1se methods are implemented in the PMA R package (Witten *et al.*, 2009).

The Yuan & Zhang TPower method solves the largest  $k$ -sparse eigenvalue problem, which has the following formulation for sparse PCA:

$$\max_{\mathbf{v}} \mathbf{v}^T \boldsymbol{\Sigma} \mathbf{v} \text{ subject to } \|\mathbf{v}\|_2^2 = 1, \|\mathbf{v}\|_0 \leq k \quad (2)$$

The TPower method solves this cardinality constrained optimization problem using a simple truncated version of the power iteration algorithm that truncates the estimated eigenvector on each iteration by setting all loadings to 0 except for the  $k$  loadings with the largest absolute values. Similar to the SPC method, the true PC cardinality  $k$  is usually unknown and must be estimated using cross-validation. Because an R version of the TPower method is not available, we have provided an implementation in the EESPCA R package along with cross-validation logic to select the sparsity parameter  $k$ .

The Tan et al. rifle method solves the largest  $k$ -sparse generalized eigenvalue problem, which is identical to (2) in the context of PCA. For the rifle method, this optimization problem is solved using a two-stage approach that begins with a convex relaxation of (2) (Gao *et al.*, 2017) to generate an initial value of the principal eigenvector. This initialization step is followed by a non-convex optimization algorithm similar to the truncated power iteration of TPower that iteratively updates the generalized Rayleigh quotient followed by truncation of the estimated eigenvector to preserve the top  $k$  eigenvector loadings with the largest absolute values with all other loadings set to 0. For evaluation of the rifle method, we used the implementation in the rifle R package with cross-validation logic we implemented in the EESPCA package.

## 1.5 Eigenvectors from eigenvalues

Denton et al. (Denton *et al.*, 2021) recently published a survey of a fascinating association between the squared elements of the unit length eigenvectors of a Hermitian matrix and the eigenvalues of the full matrix and associated sub-matrices. This identity has a complex history in the mathematical literature with numerous rediscoveries in different disciplines since the earliest known reference in 1834 (Jacobi, 1834). Importantly in this context, the sample covariance matrix  $\boldsymbol{\Sigma}$  is Hermitian, which implies that squared PC loadings can be computed as a function of PC variances for the full data set and all of the leave-one-out variable subsets. To briefly restate the results from Denton et al., let  $\mathbf{A}$  be a  $p \times p$  Hermitian matrix with eigenvalues  $\lambda_i(\mathbf{A})$  and unit length eigenvectors  $\mathbf{v}_i$  for  $i = 1, \dots, p$ . Let the elements of each eigenvector  $\mathbf{v}_i$  be denoted  $v_{i,j}$  for  $j = 1, \dots, p$  and let the squared value of element  $j$  be denoted as  $|v_{i,j}|^2$ . Let  $\mathbf{M}_j$  be the  $(p-1) \times (p-1)$  submatrix of  $\mathbf{A}$  that results from removing the  $j^{\text{th}}$  column and the  $j^{\text{th}}$  row from  $\mathbf{A}$ . Denote the eigenvalues of  $\mathbf{M}_j$  by  $\lambda_k(\mathbf{M}_j)$  for  $k = 1, \dots, p-1$ . Given these values, Denton et al. (Denton *et al.*, 2021) state their key result in Theorem 1:

**Theorem 1 (from (Denton et al., 2021)).**—The squared elements of the unit length eigenvectors are related to the eigenvalues and the submatrix eigenvalues:

$$|v_{i,j}|^2 \prod_{k=1, k \neq i}^p (\lambda_i(\mathbf{A}) - \lambda_k(\mathbf{A})) = \prod_{k=1}^{p-1} (\lambda_i(\mathbf{A}) - \lambda_k(\mathbf{M}_j)) \quad (3)$$

Which can alternatively be represented as a ratio of the product of eigenvalue differences:

$$|v_{i,j}|^2 = \frac{\prod_{k=1}^{p-1} (\lambda_i(\mathbf{A}) - \lambda_k(\mathbf{M}_j))}{\prod_{k=1, k \neq i}^p (\lambda_i(\mathbf{A}) - \lambda_k(\mathbf{A}))} \quad (4)$$

In the context of PCA,  $\mathbf{A}$  can be replaced by the sample covariance matrix  $\mathbf{\Sigma}$  and (4) provides a formula for computing squared normed PC loadings from the PC variances of the full matrix and associated sub-matrices. It is important to note two degenerate scenarios detailed in Denton *et al.* under which both sides of (3) and (4) vanish: 1) when  $v_{i,j} = 0$ ,  $\mathbf{v}_j$  for  $\mathbf{A}$  will be an eigenvector for  $\mathbf{M}_j$  with the same eigenvalue after deleting the  $j^{\text{th}}$  coefficient, and 2) when  $\mathbf{A}$  has a repeat eigenvalue  $\lambda_i(\mathbf{A})$ ,  $\lambda_i(\mathbf{A})$  is also an eigenvalue of  $\mathbf{M}_j$ . Although the ratio (4) is undefined under the second degeneracy scenario (i.e., the denominator is 0), this ratio has the defined value of 0 under the first scenario, which is correct since the eigenvector loading  $v_{i,j}$  is 0 in this case. In the context of PCA on experimental data (i.e.,  $\mathbf{A} = \widehat{\mathbf{\Sigma}}$ ), these degenerate scenarios are extremely unlikely for the first  $k$  PCs as long as  $k$  is less than the rank of  $\mathbf{A}$ , so we can in practice safely assume that our approximation of (4) defined as (6) in Section 2.1 below is defined for  $\widehat{\mathbf{\Sigma}}$ .

In the remainder of this paper, we will detail an approximation of (4) relevant to the sparse PCA problem, our proposed EESPCA method (and EESPCA.cv variant) based on this approximation and results from a comparative evaluation of the EESPCA, EESPCA.cv, SPC, SPC.1se, TPower and rifle techniques on simulated data and real data generated using single cell RNA-sequencing (Tanay and Regev, 2017; Wagner *et al.*, 2016). The Supplementary Material associated with this paper includes a vignette showing results on the simple running example (this is also contained in the EESPCA package), and supplemental simulation and real data results. The most recent version of the EESPCA R package can be found on CRAN.

## 2 Methods

### 2.1 Approximate eigenvector from eigenvalue formulation

If  $\mathbf{A}$  is replaced by the sample covariance matrix  $\mathbf{\Sigma}$ ,  $\mathbf{\Sigma}_j$  represents the sample covariance sub-matrix with variable  $j$  removed, and computation focuses on the loadings for the first PC, (4) can be restated as:

$$|v_{1,j}|^2 = \frac{\prod_{k=1}^{p-1} (\lambda_1(\mathbf{\Sigma}) - \lambda_k(\mathbf{\Sigma}_j))}{\prod_{k=2}^p (\lambda_1(\mathbf{\Sigma}) - \lambda_k(\mathbf{\Sigma}))} \quad (5)$$

Although a theoretically interesting result, (5) does not have obvious practical utility for PCA since it only generates squared values (i.e., it fails to capture the sign of the loading) and requires computation of the eigenvalues of all of the covariance sub-matrices. However, if we make the approximation that the eigenvalues  $\lambda_i(\mathbf{\Sigma})$  for  $i = 2, \dots, p-1$  are equal to the corresponding sub-matrix eigenvalues  $\lambda_i(\mathbf{\Sigma}_j)$ , many of the eigenvalue difference terms in (5) cancel and we can simplify as follows:

$$\begin{aligned}
|v_{1,j}|^2 &= \frac{\prod_{k=1}^{p-1} (\lambda_1(\Sigma) - \lambda_k(\Sigma_j))}{\prod_{k=2}^p (\lambda_1(\Sigma) - \lambda_k(\Sigma))} \\
|v_{1,j}|^2 &= \frac{(\lambda_1(\Sigma) - \lambda_1(\Sigma_j))(\lambda_1(\Sigma) - \lambda_2(\Sigma_j)) \dots (\lambda_1(\Sigma) - \lambda_{p-1}(\Sigma_j))}{(\lambda_1(\Sigma) - \lambda_2(\Sigma)) \dots (\lambda_1(\Sigma) - \lambda_{p-1}(\Sigma))(\lambda_1(\Sigma) - \lambda_p(\Sigma))} \\
|\tilde{v}_{1,j}|^2 &= \frac{\lambda_1(\Sigma) - \lambda_1(\Sigma_j)}{\lambda_1(\Sigma) - \lambda_p(\Sigma)}
\end{aligned}$$

If we also assume that  $\lambda_p(\Sigma) \approx 0$ , we can further simplify the approximate squared loadings  $|\tilde{v}_{1,j}|^2$  to:

$$|\tilde{v}_{1,j}|^2 = 1 - \frac{\lambda_1(\Sigma_j)}{\lambda_1(\Sigma)} \quad (6)$$

This approximation has several appealing properties in the context of sparse PCA. First, it greatly reduces the computational cost since only the largest eigenvalues of the full covariance matrix and associated sub-matrices are needed, which can be efficiently computed even for very large matrices using the method of power iteration. The computational cost can be further lowered by using the estimated principal eigenvector of the full matrix to initialize the power iteration calculation for the sub-matrices. When this approximation is applied to the population covariance matrix  $\Sigma$ , it will correctly estimate zero squared loadings since  $\lambda_1(\Sigma) = \lambda_1(\Sigma_j)$  for these variables per the first degeneracy scenario outlined above; for variables that have a non-zero loading on the first population PC, the approximation will be less than or equal to the squared value of the true loading. It should also be noted that the second degeneracy case outlined above does not apply to approximation (6) since only the terms for the first eigenvalue are retained. For the example introduced in Section 1.3, the squared loadings for the first population PC are:

$$|v_1|^2 = [0.25 \ 0.25 \ 0.25 \ 0.25 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]$$

and the approximate squared loadings computed via (6) are:

$$|\tilde{v}_1|^2 = [0.2 \ 0.2 \ 0.2 \ 0.2 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]$$

When applied to the sample covariance matrix  $\Sigma$ , the approximate squared loadings will again be less than or equal to the squared PC loadings with the values for variables with a zero loading on the population PC approaching zero as  $n \rightarrow \infty$  and  $n/p \rightarrow \infty$  given the consistency of the sample PC loading vectors in this asymptotic regime (Johnstone and Lu, 2009). Let  $\mathbf{r}_j$  be a vector whose elements  $r_{i,j}$  are the ratios of the approximate-to-true absolute loadings of eigenvector  $\mathbf{v}_j$ :

$$r_{i,j} = \sqrt{|\tilde{v}_{i,j}|^2 / |v_{i,j}|^2} \quad (7)$$

Importantly, these ratios for  $\mathbf{v}_1$ , i.e.  $r_{1,j}$  will tend to be larger for variables that have a non-zero loading on the population PC than for variables with a zero population loading. Although we are not able to formally prove this statement, it is informally based on the fact that the eigenvalue difference term retained in the numerator of the approximation,  $\lambda_1(\boldsymbol{\Sigma}) - \lambda_1(\boldsymbol{\Sigma}_j)$ , tends to capture a greater proportion of  $|v_{1,j}|^2$  for variables with a true non-zero loading on the first population PC than for variables with a zero loading on the first population PC. If variable  $j$  only has a non-zero loading on the first population PC and  $\lambda_{p'}(\boldsymbol{\Sigma}) \approx 0$ , then it follows from PCA consistency (Johnstone and Lu, 2009) that  $r_{1,j} \rightarrow 1$  as  $n \rightarrow \infty$  and  $n/p \rightarrow \infty$ . In particular, the consistency of PC loadings in that asymptotic regime implies consistency of the squared loadings via (4) and, when variable  $j$  only has a non-zero value on the first population PC and  $\lambda_{p'}(\boldsymbol{\Sigma}) \approx 0$ , (4) reduces to (6), which means that both the numerator and denominator of  $r_{1,j}$  converge to the same value. On the other hand, variables that have a zero loading on the first population PC will have a sample loading that approaches 0 as  $n \rightarrow \infty$  and  $n/p \rightarrow \infty$ , and, for finite  $n$ , have a non-zero value based on random contributions spread across the spectrum of eigenvalue difference terms.

In this case, the approximation based on the difference term for just the largest eigenvalue,  $\lambda_1(\boldsymbol{\Sigma}) - \lambda_1(\boldsymbol{\Sigma}_j)$ , will tend to underestimate the sample loading by a greater degree than for variables with a true non-zero population PC loading. This property of approximation (6) provides important information regarding the true sparsity of the PC loadings and plays a key role in the EESPCA method detailed below. For the example data, the ratio (rounded to three decimal places) of approximate to real normed loadings for the first sample PC are:

$$\mathbf{r}_1 = [0.919 \ 0.938 \ 0.909 \ 0.910 \ 0.837 \ 0.839 \ 0.867 \ 0.816 \ 0.815 \ 0.840]$$

Figure 1 illustrates the relationship between approximate and true squared loadings for first PC as computed on 1,000 data sets simulated according to the model in Section 1.3. As shown in this figure, the ratio for variables with a non-zero loading on the population PC is markedly larger on average than the ratio for variables with a zero population loading.

## 2.2 EESPCA

### Algorithm 1 Eigenvectors from Eigenvalues Sparse PCA (EESPCA)

**Input:**  $n \times p$  mean-centered matrix  $\mathbf{X}$  and threshold  $\alpha$

**Output:** sparse first PC loadings vector  $\mathbf{v}_1^s$  and associated eigenvalue  $\lambda_1^s$

1:  $\boldsymbol{\Sigma} = 1/(n-1)\mathbf{X}^T\mathbf{X}$  ▷ Compute sample covariance matrix

2:  $\{\mathbf{v}_1, \lambda_1\} = \text{powerIteration}(\boldsymbol{\Sigma})$  ▷ Compute first eigenvector/eigenvalue

3:  $\forall_{j \in [1:p]} \lambda_1(\boldsymbol{\Sigma}_j) = \text{powerIteration}(\boldsymbol{\Sigma}_j)$  ▷ Compute first submatrix eigenvalues

4:  $\forall_{j \in [1:p]} |\tilde{v}_{1,j}|^2 = 1 - \lambda_1(\boldsymbol{\Sigma}_j) / \lambda_1(\boldsymbol{\Sigma})$  ▷ Approximate first eigenvector using (6)

5:  $\forall_{j \in [1:p]} r_{1,j} = \sqrt{|\tilde{v}_{1,j}|^2 / |v_{1,j}|^2}$  ▷ Compute ratio of eigenvector elements

6: Compute sparse  $\mathbf{v}_1^s$  using  $\mathbf{r}_1$ :



- a:  $\mathbf{v}_1^s = \mathbf{r}_1 \odot \mathbf{v}_1 \triangleright$  Scale  $\mathbf{v}_1$  by  $\mathbf{r}_1$
- b:  $\mathbf{v}_1^s = \mathbf{v}_1^s / \|\mathbf{v}_1^s\|_2 \triangleright$  Normalize  $\mathbf{v}_1^s$  to unit length
- c: Set  $v_{1,j}^s = 0$  for all  $v_{1,j}^s < \alpha \triangleright$  Truncate  $\mathbf{v}_1^s$  using threshold  $\alpha$
- d:  $\mathbf{v}_1^s = \mathbf{v}_1^s / \|\mathbf{v}_1^s\|_2 \triangleright$  Normalize  $\mathbf{v}_1^s$  to unit length
- 7:  $\lambda_1^s = (\mathbf{v}_1^s)^T \Sigma \mathbf{v}_1^s \triangleright$  Compute sparse first eigenvalue

Our proposed EESPCA method, which is based on the eigenvector from eigenvalue approximation (6) detailed above, computes a sparse version of the first sample PC for mean-centered  $n \times p$  matrix  $\mathbf{X}$  using Algorithm 1. We explore two versions of the EESPCA method in the remainder of this paper: a version where  $\alpha = 1/\sqrt{p}$  (element of a unit length vector that has identical values) and a version where  $\alpha$  is selected via cross-validation to minimize the out-of-sample reconstruction error (i.e., the approach used by SPC to select  $\lambda$ ). In the remainder of this paper, we will refer to the version where  $\alpha = 1/\sqrt{p}$  as EESPCA and the version where  $\alpha$  is selected by cross-validation as EESPCA.cv. The rationale for using  $\alpha = 1/\sqrt{p}$  is discussed at greater length in Section 3.5 below. Additional details for each step of the EESPCA algorithm are as follows:

1. Compute the unbiased sample covariance matrix  $\Sigma = 1/(n-1)\mathbf{X}^T\mathbf{X}$ .
2. Use the power iteration method to compute the principal eigenvector  $\mathbf{v}_1$  and associated eigenvalue  $\lambda_1$  of  $\Sigma$ .
3. Use the power iteration method to compute the principal eigenvalues of the sub-matrices,  $\lambda_1(\Sigma_j)$ . For efficient calculation, the power iteration method is used with the initial value set to a subsetted version of  $\mathbf{v}_1$ , i.e.,  $\mathbf{v}_1$  with the relevant variable removed.
4. Use formula (6) to approximate the squared elements of the principal eigenvector  $|\tilde{\mathbf{v}}_1|^2$ .
5. Compute the element-wise ratio of approximate-to-real eigenvectors:
 
$$r_{1,j} = \sqrt{|\tilde{v}_{1,j}|^2 / |v_{1,j}|^2}.$$
6. Use  $\mathbf{r}_1$  to compute a sparse version of  $\mathbf{v}_1$ . As outlined in Section 2.1 above, the entries in  $\mathbf{r}_1$  will be less than one and will tend to be larger for variables with a non-zero population loading than for variables with a zero population loading. This property suggests that one could use  $\mathbf{r}_1$  to scale  $\mathbf{v}_1$  followed by renormalization to generate sample PC loadings that more closely match the population loadings, however, this would not generate sparse loadings since the elements of  $\mathbf{r}_1$  are themselves non-zero. To produce a sparse version of  $\mathbf{v}_1$ , we follow the initial scaling and renormalization with a thresholding operation that sets any adjusted loadings less than  $\alpha$  to 0 and then renormalize a final time to produce a sparse and unit length loadings vector  $\mathbf{v}_1^s$ . Specifically:
  - a. Scale  $\mathbf{v}_1$  by  $\mathbf{r}_1$ , i.e., set  $\mathbf{v}_1^s$  to the element-wise product of  $\mathbf{v}_1$  and  $\mathbf{r}_1$ .

- b. Normalize  $\mathbf{v}_1^s$  to unit length, i.e.,  $\mathbf{v}_1^s = \mathbf{v}_1^s / \|\mathbf{v}_1^s\|_2$ .
  - c. Truncate  $\mathbf{v}_1^s$  by setting  $v_{1,j}^s = 0$  for all  $v_{1,j}^s < \alpha$
  - d. Normalize  $\mathbf{v}_1^s$  again to generate a sparse and unit length eigenvector.
7. Compute the associated eigenvalue  $\lambda_1^s$  as:

$$\lambda_1^s = (\mathbf{v}_1^s)^T \mathbf{\Sigma} \mathbf{v}_1^s \quad (8)$$

To compute multiple sparse PCs, the EESPCA method uses a similar approach to that employed by the SPC method, i.e., it repeatedly applies to the procedure outlined above for calculating the first sparse PC to the residual matrix formed by subtracting the rank 1 reconstruction of  $\mathbf{X}$  generated using  $\mathbf{v}_1^s$  from the input  $\mathbf{X}$ . Note that multiple sparse PCs generated using this recursive approach are not guaranteed to be orthogonal. When applied to the example data described in Section 1.3, the EESPCA method produces the following sparse PC loading vectors (rounded to three decimal places):

$$\widehat{\mathbf{V}}_{\text{EESPCA}} [., (\mathbf{1}, \mathbf{2})] \mathbf{T} = \begin{bmatrix} 0.543 & 0.457 & 0.503 & 0.494 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.779 & 0.627 & 0 \end{bmatrix}$$

For this example, the EESPCA.cv version generates identical results since the  $\alpha$  selected via 5-fold cross-validation out of 21 potential values equally spaced from  $0.75/\sqrt{p}$  to  $1.25/\sqrt{p}$  is the default value of  $1/\sqrt{p}$ .

### 2.3 Simulation study design

To explore the comparative performance of the EESPCA, EESPCA.cv, SPC, SPC.1se, TPower and rifle methods, we simulated MVN data for a range of sample sizes, variable dimensionality and covariance structures. Although \*.1se variants of EESPCA.cv, TPower and rifle are also possible, results are only shown for SPC.1se since the relative performance pattern is similar. Specifically, we simulated data sets containing  $n$  independent samples drawn from a  $p$ -dimensional MVN data with a zero population mean vector and population covariance matrix  $\mathbf{\Sigma}$  with all variances set to 1 and covariances set to either 0 or a non-zero  $\rho$ ; 50 data sets were simulated and analyzed for each unique combination of parameter values. The specific  $n$ ,  $p$ , and  $\rho$  values were set according to four different simulation models:

- **Basic:** A single distinguishable PC was simulated by setting the covariance between a block of  $\beta p$  variables to  $\rho$  and to 0 otherwise. According to this covariance model, the first population PC will have equal, non-zero loadings for the first  $\beta p$  variables and zero loadings for the remaining  $(1-\beta)p$  variables. Simulations were performed for  $n$  from 25 to 250,  $p$  from 20 to 200,  $\rho$  from 0.025 to 0.25, and  $\beta$  from 0.025 to 0.25.
- **Block covariance:** Three distinguishable PCs were simulated by setting the covariance to  $\rho$  for disjoint variable blocks of size  $\beta p$ ,  $0.5\beta p$ , and  $0.25\beta p$ .

According to this covariance model, the first population PC will have equal, non-zero loadings for the first  $\beta p$  variables and zero loadings for the remaining  $(1-\beta)p$  variables, the second population PC will have non-zero loadings for variables in the range  $(\beta p, 1.5\beta p]$ , and the third population PC will have non-zero loadings for variables in the range  $(1.5\beta p, 1.75\beta p]$ . Data was simulated for the same  $n$ ,  $p$ ,  $\rho$ , and  $\beta$  ranges used for the basic model.

- **High dimension:** Similar to the basic model, a single distinguishable PC was simulated but with larger values of  $n$  and  $p$  and smaller  $\rho$  values. Specifically, data was simulated for  $n$  from 50 to 500,  $p$  from 100 to 1000,  $\rho$  from 0.01 to 0.10, and  $\beta$  from 0.025 to 0.25.
- **Limited sparsity:** Similar to the basic model, a single distinguishable PC was simulated but with a larger proportion of variables with non-zero loadings on the first population PC. Specifically, simulations were performed for  $n$  from 25 to 250,  $p$  from 20 to 200,  $\rho$  from 0.025 to 0.25, and  $\beta$  from 0.55 to 1.0. The case of  $\beta = 1.0$  represents a non-sparse scenario with all loadings on the first population PC equal to  $1/\sqrt{p}$ .

The EESPCA, EESPCA.cv, SPC, SPC.1se, TPower and rifle methods were used to compute the first sparse PC of each simulated data set and method performance was quantified according to classification accuracy (i.e., ability of the method to correctly assign zero loadings to variables with a zero population loading), computational speed, out-of-sample rank 1 reconstruction error, and PC estimation error. For classification accuracy, the specificity, sensitivity and balanced accuracy of each method were computed. If  $v_i$  and  $v_i^s$  represent the  $i^{\text{th}}$  elements of the true eigenvector  $\mathbf{v}$  and sparse estimate  $\mathbf{v}^s$  and  $1(\cdot)$  is the indicator function, then the sensitivity is defined as  $sens = \sum_{i=1}^p 1(v_i^s = 0 \wedge v_i = 0) / \sum_{i=1}^p 1(v_i = 0)$  (i.e., the proportion of true zero loadings recovered by the sparse PCA method), specificity is defined as  $spec = \sum_{i=1}^p 1(v_i^s \neq 0 \wedge v_i \neq 0) / \sum_{i=1}^p 1(v_i \neq 0)$  (i.e., the proportion of true non-zero loadings recovered by the sparse PCA method), and balanced accuracy is defined as  $balacc = (sens + spec) / 2$ . To measure out-of-sample reconstruction error, the first PC loadings estimated on one simulated data set were used to generate a rank 1 reconstruction of a second data set simulated using identical parameters; reconstruction error was quantified by the squared Frobenius norm of the residual matrix formed by subtracting the rank 1 reconstruction from the simulated matrix. To measure PC estimation error, the Euclidean distance was computed between the absolute value of the estimated PC loadings and the absolute value of the population PC loadings (absolute values were used to account for potential differences in eigenvector sign). The SPC and SPC.1se methods were realized using the `SPC()` method in version 1.2.1 of the PMA R package (Witten *et al.*, 2009) with the optimal penalty parameter computed using the `SPC.cv()` method with `nfolds=5`, `niter=10` and `sumabsv=seq(1, sqrt(p), len=20)`. We created an R implementation of the TPower algorithm, which is available in the EESPCA R package via the function `tpower()`. The initial eigenvector for the TPower algorithm was computed using non-truncated power iteration. To determine the optimal cardinality parameter  $k$  for TPower, we implemented

a cross-validation method based on the *SPC.cv()* function. This cross-validation method is available in the EESPCA R package via the function *tpowerPCACV()* and was executed for the simulation studies using *nfolds=5* and *k.value=round(seq(1, p, len=20))*. The rifle method was realized using the *rifle()* function in version 1.0 of the rifle R package. The initial eigenvector for the rifle method was computed using the suggested *initial.convex()* function from the rifle R package using  $K = 1$  and  $\lambda = \sqrt{\log(p) / n}$  (this is implemented in the EESPCA R package using the convenience function *rifleInit()*). Similar to the TPower method, the optimal cardinality parameter  $k$  for rifle was computed using the Witten et al. cross-validation approach as implemented by the *riflePCACV()* function in the EESPCA R package with *nfolds=5* and *k.value=round(seq(1, p, len=20))*. For the EESPCA.cv method, the optimal threshold was selected using the Witten et al. cross-validation approach as implemented by the *eespcaCV()* function the EESPCA R package with *nfolds=5* and *sparse.threshold.values=seq(from=0.75/sqrt(p), to=1.25/sqrt(p), length.out=21)*.

## 2.4 Real data analysis design

To explore the performance of the EESPCA, EESPCA.cv, SPC, SPC.1se, TPower and rifle methods on real data, we analyzed a single cell RNA-sequencing (scRNA-seq) data set generated on 2.7k human peripheral blood mononuclear (PBMC) cells. This scRNA-seq data set is freely available from 10x Genomics and is used in the Guided Clustering Tutorial (Seurat, 2020) for the Seurat single cell framework (Stuart *et al.*, 2019). Preprocessing, quality control (QC) and normalization of the PBMC data set followed the same processing steps used in the Seurat Guided Clustering Tutorial. Specifically, the Seurat log-normalization method was used followed by application of the *vst* method for decomposing technical and biological variance. Seurat log-normalization divides the unique molecular identifier (UMI) counts for each gene in a specific cell by the sum of the UMI counts for all genes measured in the cell and multiplies this ratio by the scale factor  $1 \times 10^6$ . The normalized scRNA-seq values are then generated by taking the natural log of this relative value plus 1. This technique generates normalized data whose non-zero values can be approximated by a log-normal distribution. The Seurat *vst* method fits a non-linear trend to the log scale variance/mean relationship. This estimated trend models the expected technical variance based on mean gene expression; observed variance values above this expected trend reflect biological variance. Preprocessing and QC of the PBMC data yielded normalized counts for 14,497 genes and 2,638 cells. Immune cell types were assigned using the same procedure detailed in the Seurat Guided Clustering Tutorial.

The 1,000 genes with the largest estimated biological variance according to the *vst* method were used as input for standard PCA, EESPCA, EESPCA.cv, SPC, SPC.1se, TPower and rifle. Only the first two sparse PCs were generated using the five sparse PCA methods, which were executed using the same parameter settings as detailed in Section 2.3. To capture out-of-sample reconstruction error, the scRNA-seq data set was randomly split in half 10 times and, for each split, PCs were estimated using each method on one half and the squared Frobenius norm of the residual matrix for the first two PCs was computed on the other half of the data. Gene Ontology (GO) (Gene Ontology Consortium, 2010) enrichment analysis was used to interpret the normal and sparse PCs. Specifically, the *goana()* method

(Young *et al.*, 2010) in the *limma* R package (Ritchie *et al.*, 2015) was used to determine the statistical enrichment of GO Biological Process Ontology terms among genes with either positive or negative PC loadings (GO annotations were obtained using version 3.11.4 of the GO.db Bioconductor R package (Carlson, 2020)). The *goana()* method performs this enrichment analysis using a Fisher's exact test on the  $2 \times 2$  contingency table that categorizes the 1,000 high variance genes according to whether they belong to the target GO term and whether they have a positive (or negative) loading on the target PC. False discovery rate (FDR) values were computed using the Benjamini and Hochberg method (Benjamini and Hochberg, 1995) separately for each PC and loading direction for the family of test corresponding to all GO Biological Process terms.

### 3 Results and discussion

The relative performance of the evaluated methods on data simulated according to the basic model is detailed in Sections 3.1–3.4 below. Simulation results for the block covariance, high dimension and limited sparsity models can be found in the Supplemental Material. The motivation for using  $\alpha = 1/\sqrt{p}$  for the EESPCA method is outlined in Section 3.5 and results for the scRNA-seq example are contained in Section 3.6.

#### 3.1 Classification performance

As illustrated in Figure 2, the EESPCA method has superior classification performance (as measured by balanced accuracy) relative to the other methods across almost the full range of parameter values explored for the basic model.

Performance of the EESPCA.cv and TPower methods was close to EESPCA with TPower having slightly better accuracy at large  $n$  and large  $\rho$ . As expected, the accuracy of all methods tends to increase as either  $n$ ,  $\rho$  or  $\beta$  increase; improved performance at larger  $p$  with  $n$  fixed was unexpected. Performance of the SPC and rifle methods approaches the EESPCA, EESPCA.cv and TPower methods at higher parameter values. The corresponding sensitivity (i.e., ability to assign zero loadings for variables with a zero population loading) and specificity (i.e., ability to assign a non-zero loading to variables with a non-zero population loading) values, shown in Figures 3 and 4, provide more insight into the relative performance of the methods. Looking at the sensitivity plots in Figure 3, the SPC.1se method had almost perfect sensitivity, which is consistent with the fact that the SPC.1se method generates a much sparser solution than the SPC variant. Sensitivity of the SPC, rifle and TPower methods was also very high. By contrast, the EESPCA and EESPCA.cv methods have lower sensitivity for this simulation model, although the sensitivity improves as sparsity decreases, i.e., a larger proportion of informative variables. As shown in Figure 4, EESPCA and EESPCA.cv methods have the best specificity across the range of simulated parameter values. As expected, the specificity for all evaluated methods increases as the parameter values increase. Relative to the other methods, SPC.1se has much lower specificity, which is expected given the sparser solution generated by this approach.

Classification performance for the block covariance, high dimension and limited sparsity models is contained in Figures S1–S3, S7–S9, and S13–S15 in the Supplementary Results. Relative classification performance of the evaluated methods for both the block covariance

model and high dimension model is very similar to the performance for the basic model, i.e., EESPCA achieves the best balanced accuracy across almost all evaluated parameter values with EESPCA.cv and TPower just slightly worse. For the limited sparsity model, however, EESPCA.cv achieves the best balanced accuracy with SPC.1se a close second, EESPCA and TPower significantly lower, and rifle and SPC at the bottom. In this scenario, the poor relative performance of the EESPCA model is due to an overly sparse model (i.e., the  $1/\sqrt{p}$  threshold is too high resulting in poor specificity).

### 3.2 Computational speed

Figure 5 illustrates the relative computational cost of the evaluated methods for the basic simulation model. Because the SPC.1se and SPC methods have equivalent computational cost, the lines overlap and only the SPC.1se values are visible. As shown in this figure, the execution times of the EESPCA.cv, SPC, SPC.1se, rifle and TPower methods are between one and two-orders-of-magnitude larger than the EESPCA method across all tested parameter values. The dramatic differences in computational cost are primarily driven by the required cross-validation-based selection of sparsity constraints for these methods. As expected, computational speed is fairly insensitive to changes in  $\rho$  or  $\beta$  for all methods and tends to increase as either  $n$  or  $p$  are increased.

Computational speed for the block covariance, high dimension and limited sparsity models is contained in Figures S4, S10, S16 in the Supplementary Results. The relative computational cost of the evaluated methods for both the block covariance model and limited sparsity model is very similar to the cost measured on the basic model, i.e., the computational cost of EESPCA is between one and two-orders-of-magnitude better than other methods. For the high dimension model, however, the SPC and SPC.1se methods have the best computational cost at large  $p$  with fixed  $n$ . This is due to the fact that the penalized matrix decomposition technique used SPC and SPC.1se can in effect solve the eigenvalue problem for the smaller  $n \times n$  matrix  $\mathbf{X}\mathbf{X}^T$  rather than for the  $p \times p$  matrix  $\mathbf{X}^T\mathbf{X}$ .

### 3.3 Reconstruction error

Figure 6 shows the relative rank 1 out-of-sample reconstruction error of the evaluated sparse methods as compared to standard PCA for the basic simulation model. The reconstruction error for the EESPCA, EESPCA.cv and TPower methods was similar and, for higher parameter values, lower than standard PCA. Compared to these methods, the reconstruction error for rifle, SPC and SPC.1se was markedly higher. As expected given the more sparse solution, the SPC.1se had the worst reconstruction error of all evaluated methods.

Reconstruction error for the block covariance, high dimension and limited sparsity models is contained in Figures S5, S11, S17 in the Supplementary Results. The relative out-of-sample reconstruction error of the evaluated methods for both the block covariance model and high dimension model is very similar to the reconstruction error measured on the basic model. For the limited sparsity model, however, the EESPCA method has a dramatically larger reconstruction error than the other methods with SPC and TPower demonstrating the best performance. Although the reconstruction error for the EESPCA.cv method in this case is much lower than the error for EESPCA, it is larger than SPC, rifle and TPower. This pattern

is in contrast to classification accuracy for the limited sparsity model, where EESPCA.cv had the best performance.

### 3.4 PC estimation error

Figure 7 shows the relative PC estimation error of the evaluated sparse methods as compared standard PCA for the basic model. The plotted estimation error is quantified by the Euclidean distance between the estimated PC loadings and the population loadings of the first PC. As expected, the relative PC estimation error follows a similar trend to the relative out-of-sample reconstruction error.

PC estimation error for the block covariance, high dimension and limited sparsity models is contained in Figures S6, S12, S18 in the Supplementary Results. The relative PC estimation error of the evaluated methods for both the block covariance model and high dimension model is very similar to the PC estimation error measured on the basic model. For the limited sparsity model, however, the EESPCA method has a dramatically larger estimation error than the other methods with SPC and TPower demonstrating the best performance. As expected, the results for PC estimation error on these models are similar to those found for out-of-sample reconstruction error.

### 3.5 Rationale for $\alpha = 1/\sqrt{p}$

Although we are unable to formally prove the optimality of  $\alpha = 1/\sqrt{p}$ , this choice for the default threshold can be motivated by both heuristic arguments and simulation results. For a unit length eigenvector of length  $p$  with identical elements, all elements have the same absolute value of  $1/\sqrt{p}$ . Importantly, this form of eigenvector occurs as the loadings for the first population PC in the case where all population variances are equal and all covariances are equal to a non-zero  $\rho$ . In the context of sparse PCA, such a population covariance matrix can be considered as a null case, i.e., a covariance matrix with no internal substructure and first population PC loadings that diverge, in a qualitative sense, as far as possible from a sparse eigenvector. When sparsity does exist, some of the population PC loadings will equal 0, forcing the non-zero elements to be greater than  $1/\sqrt{p}$  for a unit length eigenvector. For estimation of sample PCs when consistency holds (i.e.,  $p/n \rightarrow 0$  (Johnstone and Lu, 2009)), the expected absolute value of the loadings will converge in probability to  $1/\sqrt{p}$  in the null scenario and to 0 or values above  $1/\sqrt{p}$  in the sparse scenario. These properties of the population and sample PC loadings in the null and sparse cases motivate our use of  $1/\sqrt{p}$  as a default threshold for the EESPCA method. Informally, when a sparse population PC is assumed, the non-sparse elements can be expected to have sample values above  $1/\sqrt{p}$  and the sparse elements can be expected to have sample values below  $1/\sqrt{p}$ . The EESPCA method attempts to magnify these deviations away from  $1/\sqrt{p}$  by scaling the sample PC loadings by the ratio  $r_1$  (7) of approximate-to-real loadings. Because this ratio tends to be larger for non-sparse elements than for sparse elements, the non-sparse elements will be shifted further above  $1/\sqrt{p}$  and the sparse elements further below  $1/\sqrt{p}$ .

Simulation studies also provide empirical support for using  $1/\sqrt{p}$  as a default threshold when significant sparsity can be assumed for the population PC loadings. As shown in Sections

3.1, 3.3 and 3.4 and the Supplementary Results, the EESPCA variant performs as well as or better than the EESPCA.cv variant across a wide range of simulation models (i.e., the basic, block covariance and high dimension models detailed in Section 2.3) in terms of sparsity estimation (Figures 2, S1, and S7), out-of-sample reconstruction error (Figures 6, S5 and S11) and PC estimation error (Figures 7, S6, and S12). For these simulation models, the  $1/\sqrt{p}$  threshold generates a more sparse solution on average than the threshold selected via cross-validation (i.e., the sensitivity of EESPCA is uniformly larger than the sensitivity of EESPCA.cv as shown in Figures 3, S2 and S8 with the reverse holding for specificity as shown in Figures 4, S3 and S9). However, as demonstrated by the results for the limited sparsity model (Figures S13–S18), the performance of the  $1/\sqrt{p}$  threshold does not hold when the population sparsity is limited. In this limited sparsity scenario, the  $1/\sqrt{p}$  default is too large and the threshold should instead be selected via cross-validation to minimize out-of-sample reconstruction error.

### 3.6 Real data analysis

As detailed in Section 2.4, standard PCA, EESPCA, EESPCA.cv, SPC, SPC.1se, TPower and rifle were applied to a scRNA-seq data set that captured the expression values of the 1,000 genes with the largest estimated biological variance for 2,638 individual human immune cells. As illustrated in Figure 8, all methods produce a similar projection of these cells onto the first two standard or sparse PCs (the EESPCA.cv projection is very similar to the EESPCA projection so has been omitted to allow for a 3-by-2 panel plot). Similar to the simulation results, the EESPCA method was close to 2 orders-of-magnitude faster than SPC or SPC.1se (20.87 seconds vs. 25.92 minutes on a standard laptop). The TPower and EESPCA.cv methods were markedly slower, taking 57.12 and 54.07 minutes respectively. Because the rifle method is more than an order of magnitude slower than SPC, cross-validation was not performed for this data set and the optimal  $k$  values computed for TPower were used instead; even without cross-validation, execution of rifle still took 7.08 minutes. Qualitatively, all of the methods generate a very similar projection with the first PC separating cytotoxic from non-cytotoxic cells and the second PC capturing phenotypic differences among the non-cytotoxic populations.

Table 1 lists various statistics for first two normal or sparse PCs computed by each of these methods on the PBMC scRNA-seq data. Standard PCA assigned a non-zero loading to all genes on both PCs and GO enrichment analysis based on these loadings only produced significant findings for genes associated with PC 1 (44 significant terms for genes with a positive sign). The SPC method generated very little sparsity on this data set with GO enrichment results similar to those generated by standard PCA. As expected, the SPC.1se method produced sparser PCs than the SPC method, which increased the number of significant GO terms for PC 1 and allowed 47 GO terms to reach significance for PC 2 (33 for the positive direction and 14 for the negative direction). TPower produced few zero loadings for the first PC so had enrichment results similar to standard PCA. For PC 2, TPower generated some sparsity with 3 significant GO terms for the positive direction. Because rifle was executed with the same cardinality constraints as TPower, it also produced similar results for PC 1 and only moderate sparsity for PC 2 with similar numbers of significant GO terms. The EESPCA method generated the most sparse solution on



this scRNA-seq data, which enabled GO enrichment analysis to produce significant results for three of the four test cases. All of the sparse PCA methods had average out-of-sample reconstruction errors less than one percent larger than PCA with the SPC error slightly below the error produced by PCA. As expected, methods that produced sparser estimates had a larger out-of-sample reconstruction error. Table 2 below lists the five most significant GO terms for each test case for PCA, EESPCA, SPC and TPower (equivalent results for EESPCA.cv, SPC.1se and rifle can be found in the Supplementary Material). Importantly, these results mirror the expected interpretation of the first two PCs. These results also highlight a well known challenge with hierarchical gene set collections like the Gene Ontology, namely that the ranked result list will contain sequences of biologically similar and highly overlapping gene sets.

## 4 Conclusions

The EESPCA method is a novel sparse PCA technique based on an approximation of the recently rediscovered formula for calculating eigenvectors from eigenvalues for Hermitian matrices (Denton *et al.*, 2021). In this paper, we have explored the performance of two versions of this method: EESPCA and EESPCA.cv. The EESPCA version uses a fixed threshold of  $1/\sqrt{p}$  to induce sparsity in the estimated PC loadings and the EESPCA.cv version selects the threshold via cross-validation to minimize the out-of-sample reconstruction error. Compared to the state-of-the-art SPC (and SPC.1se variant) (Witten *et al.*, 2009), TPower (Yuan and Zhang, 2013) and rifle (Tan *et al.*, 2018) methods, the EESPCA version can more accurately estimate population PC sparsity across a range of covariance structures, sample sizes and variable dimensionality at a dramatically lower computational cost. Importantly, EESPCA provides improved estimation of PC sparsity and lower computational cost with an out-of-sample reconstruction error and PC estimation error close to the error achieved by the TPower method and significantly below the error for the SPC, SPC.1se and rifle techniques. In case of limited PC sparsity, i.e. when the proportion of PC loadings with a 0 population value is less than 0.5, the classification performance, out-of-sample reconstruction error and PC estimation error for the EESPCA method deteriorate significantly. In this scenario, the EESPCA.cv version provides the best sparsity estimation with PC estimation error and reconstruction error close to the error generated by TPower. To select the most appropriate sparse PCA method for a given analysis problem, users are encouraged to review the classification performance, PC estimation error, out-of-sample reconstruction error and execution time results for the four different simulation models explored in this paper. In general, for analysis problems where significant sparsity is assumed for the population PCs, EESPCA should be preferred over EESPCA.cv, SPC, SPC.1se, TPower and rifle. EESPCA is also generally preferable for computationally demanding problems, such as the analysis of very large data matrices or the use of statistical methods like resampling that require repeated application of sparse PCA (if  $p \gg n$ , the computational cost of SPC may be lower than EESPCA). If significant sparsity cannot be assumed for the population PCs and computational speed is not critical, then the EESPCA.cv version should be used over EESPCA. If minimization of PC estimation error or out-of-sample reconstruction error is the key goal and computational speed is not a major concern, then the TPower method provides the best overall performance.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

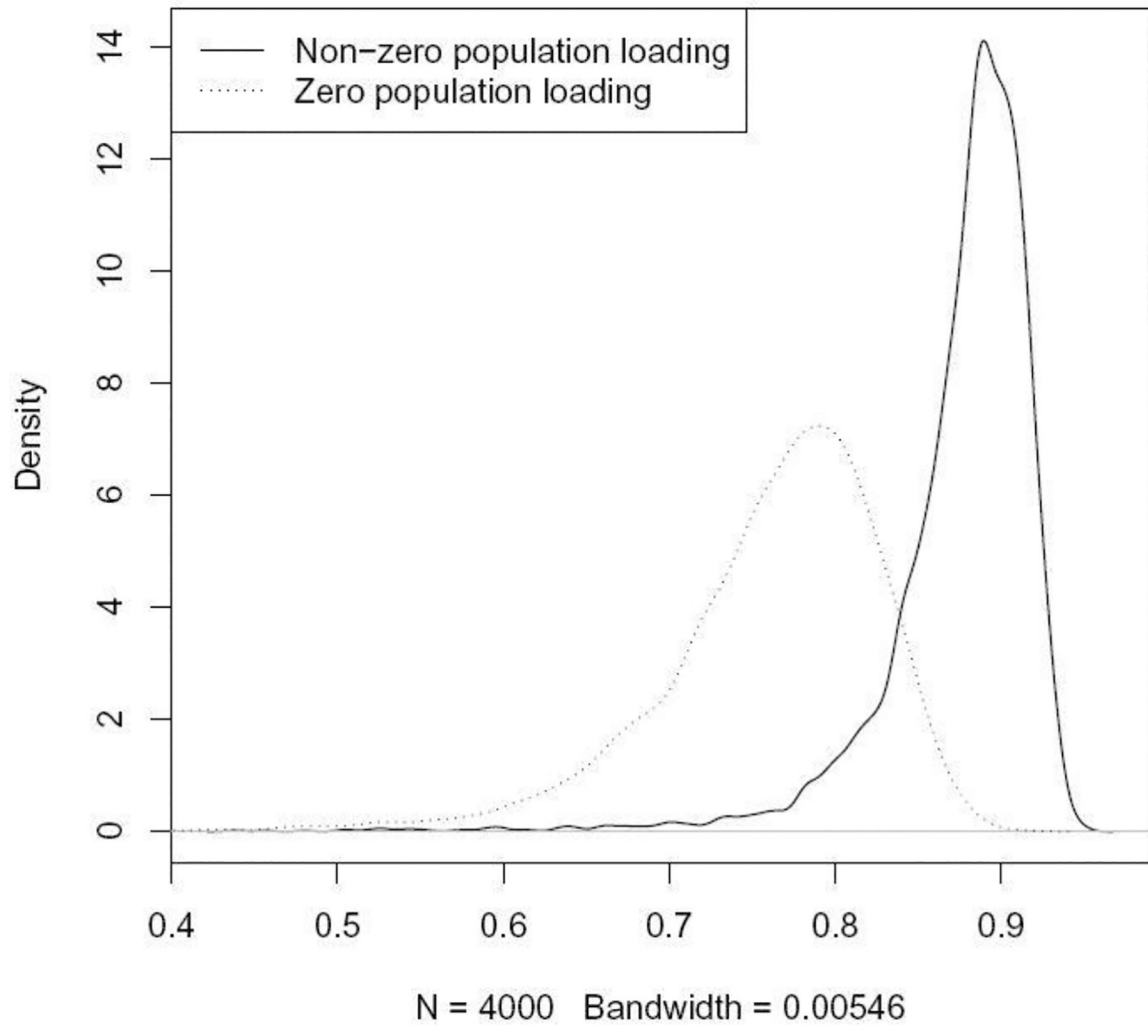
## Acknowledgments

The authors gratefully acknowledge funding support from National Institutes of Health grants K01LM012426, R21CA253408, P20GM130454 and P30CA023108.

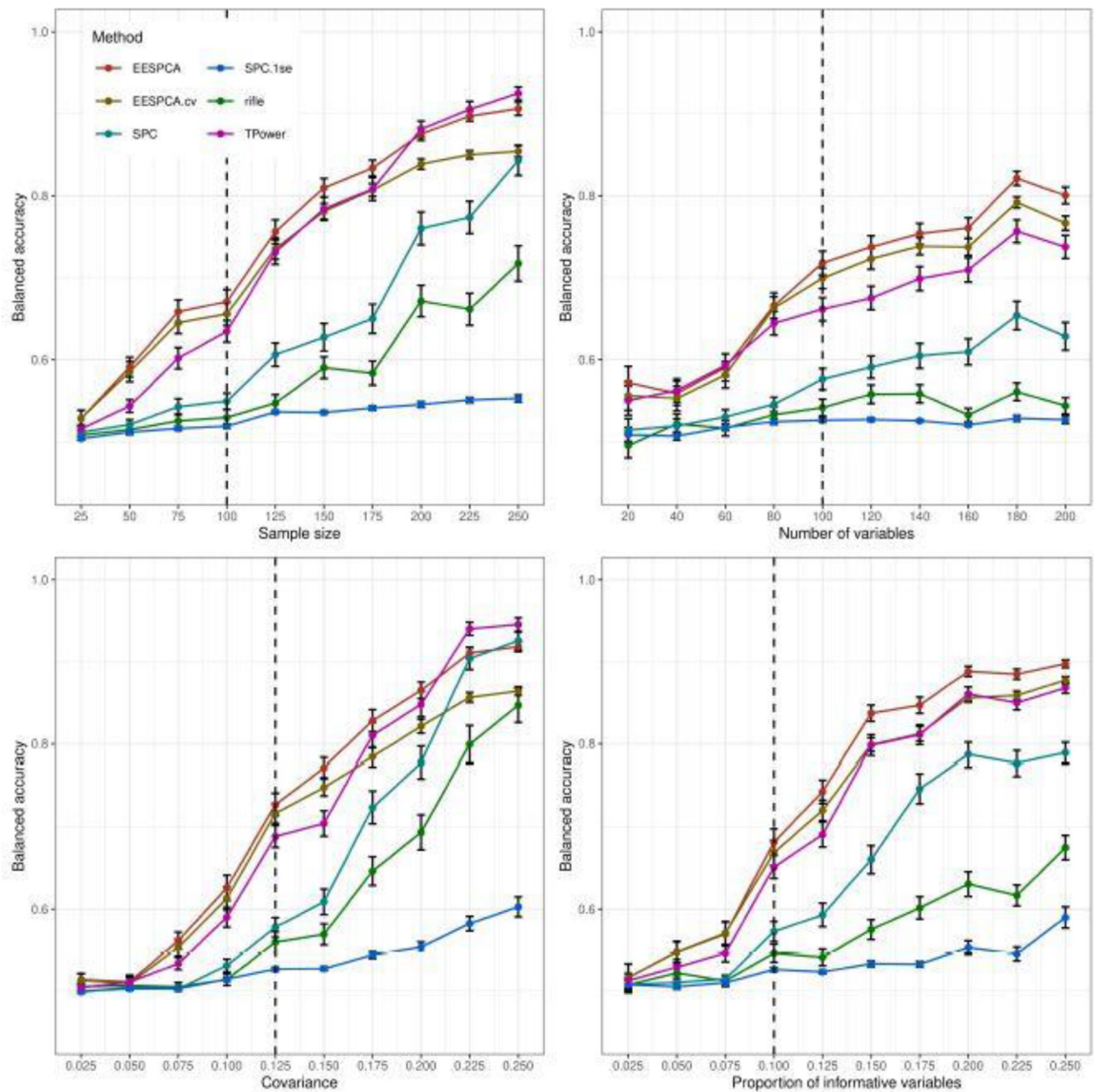
## References

- Benjamini Y and Hochberg Y (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, pages 289–300.
- Carlson M (2020). GO.db: A set of annotation maps describing the entire Gene Ontology. R package version 3.11.4.
- d'Aspremont A, El Ghaoui L, Jordan MI, and Lanckriet GRG (2007). A direct formulation for sparse PCA using semidefinite programming. *SIAM Review*, 49(3), 434–448.
- Denton PB, Parke SJ, Tao T, and Zhang X (2021). Eigenvectors from eigenvalues: A survey of a basic identity in linear algebra. *Bulletin of the American Mathematical Society*, page 1.
- Fan J, Salathia N, Liu R, Kaeser GE, Yung YC, Herman JL, Kaper F, Fan J-B, Zhang K, Chun J, and Kharchenko PV (2016). Characterizing transcriptional heterogeneity through pathway and gene set overdispersion analysis. *Nat Methods*, 13(3), 241–4. [PubMed: 26780092]
- Gao C, Ma Z, and Zhou HH (2017). Sparse cca: Adaptive estimation and computational barriers. *Ann. Statist*, 45(5), 2074–2101.
- Gene Ontology Consortium (2010). The gene ontology in 2010: extensions and refinements. *Nucleic Acids Res*, 38(Database issue), D331–5. [PubMed: 19920128]
- Hafemeister C and Satija R (2019). Normalization and variance stabilization of single-cell rna-seq data using regularized negative binomial regression. *Genome Biol*, 20(1), 296. [PubMed: 31870423]
- Hastie T, Tibshirani R, and Friedman JH (2009). *The elements of statistical learning: data mining, inference, and prediction*. Springer series in statistics. Springer, New York, NY, 2nd ed edition.
- Hotelling H (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24, 498–520.
- Jacobi C (1834). De binis quibuslibet functionibus homogeneis secundi ordinis per substitutiones lineares in alias binas transformandis, quae solis quadratis variabilium constant; una cum variis theorematis de transformatione et determinatione integralium multiplicium. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, pages 1–69.
- Johnstone IM and Lu AY (2009). On consistency and sparsity for principal components analysis in high dimensions. *Journal of the American Statistical Association*, 104(486), 682–693. [PubMed: 20617121]
- Jolliffe I (2002). *Principal Component Analysis*. Springer Series in Statistics. Springer, New York, NY.
- Jolliffe I, Trendafilov N, and Uddin M (2003). A modified principal component technique based on the LASSO. *Journal of Computational and Graphical Statistics*, 12(3), 531–547.
- Jolliffe IT and Cadima J (2016). Principal component analysis: a review and recent developments. *Philos Trans A Math Phys Eng Sci*, 374(2065), 20150202. [PubMed: 26953178]
- Jung S, Ahn J, and Jeon Y (2019). Penalized orthogonal iteration for sparse estimation of generalized eigenvalue problem. *Journal of Computational and Graphical Statistics*, 28(3), 710–721.
- Kluger Y, Basri R, Chang JT, and Gerstein M (2003). Spectral biclustering of microarray data: Co-clustering genes and conditions. *Genome Research*, 13(4), 703–716. [PubMed: 12671006]
- Lu J, Kerns RT, Peddada SD, and Bushel PR (2011). Principal component analysis-based filtering improves detection for affymetrix gene expression arrays. *Nucleic Acids Research*, 39(13), e86. [PubMed: 21525126]

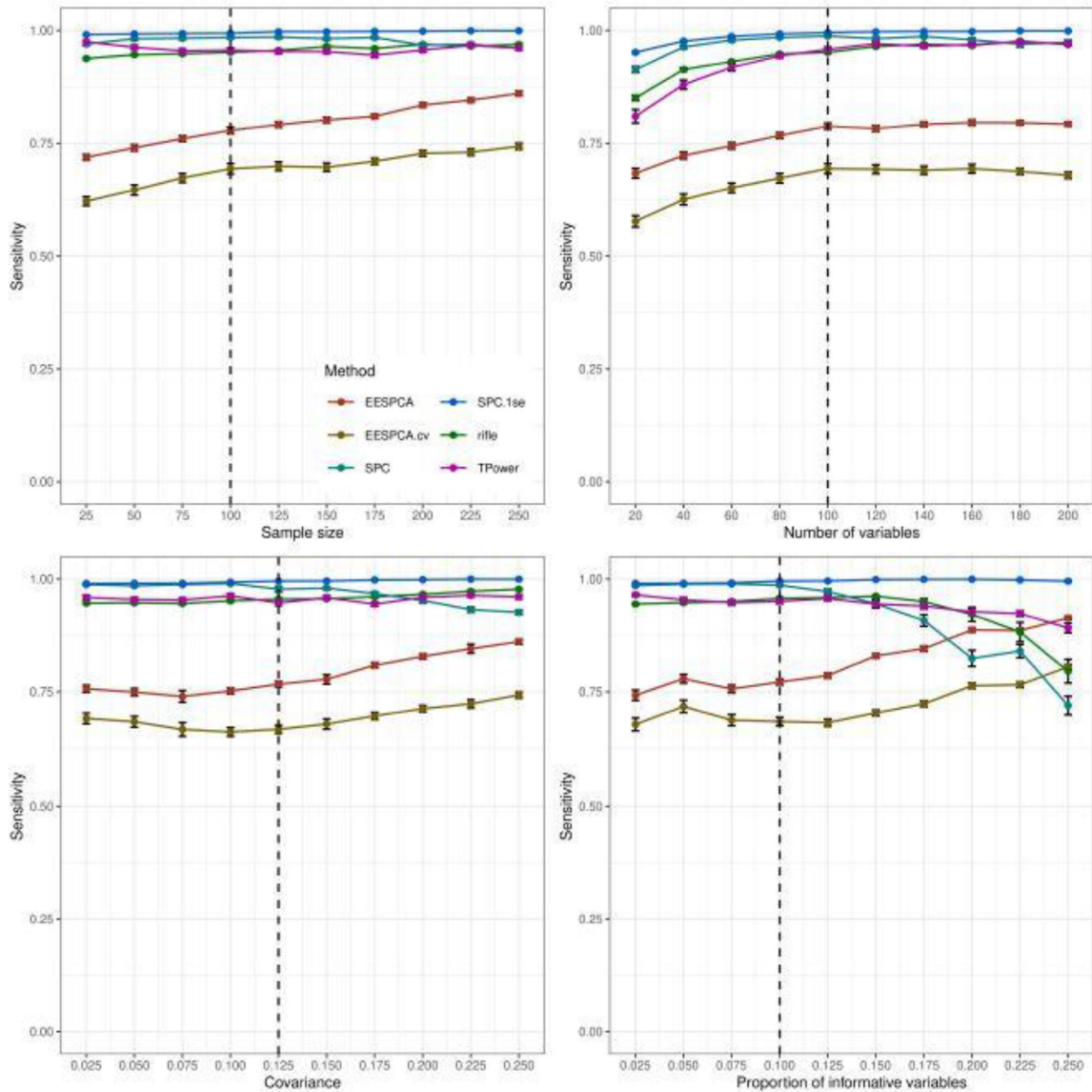
- Ma S and Dai Y (2011). Principal component analysis based methods in bioinformatics studies. *Briefings in Bioinformatics*, 12(6, SI), 714–722. [PubMed: 21242203]
- Ma Z (2013). Sparse principal component analysis and iterative thresholding. *Ann. Statist*, 41(2), 772–801.
- McInnes L, Healy J, and Melville J (2018). Umap: Uniform manifold approximation and projection for dimension reduction.
- Moghaddam B, Weiss Y, and Avidan S (2006). Spectral bounds for sparse pca: Exact and greedy algorithms. *Advances in neural information processing systems*, 18, 915.
- Patterson N, Price AL, and Reich D (2006). Population structure and eigenanalysis. *PLOS Genetics*, 2(12), e190. [PubMed: 17194218]
- Pearson K (1901). On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2(6), 559–572.
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, and Smyth GK (2015). limma powers differential expression analyses for rna-sequencing and microarray studies. *Nucleic Acids Res*, 43(7), e47. [PubMed: 25605792]
- Seurat (2020). Seurat guided clustering tutorial. [https://satijalab.org/seurat/v3.1/pbmc3k\\_tutorial.html](https://satijalab.org/seurat/v3.1/pbmc3k_tutorial.html). Accessed: 2020-02-10.
- Shen H and Huang JZ (2008). Sparse principal component analysis via regularized low rank matrix approximation. *Journal of Multivariate Analysis*, 99(6), 1015–1034.
- Sriperumbudur BK, Torres DA, and Lanckriet GRG (2011). A majorization-minimization approach to the sparse generalized eigenvalue problem. *Machine Learning*, 85(1–2), 3–39.
- Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM 3rd, Hao Y, Stoeckius M, Smibert P, and Satija R (2019). Comprehensive integration of single-cell data. *Cell*, 177(7), 1888–1902.e21. [PubMed: 31178118]
- Tan KM, Wang Z, Liu H, and Zhang T (2018). Sparse generalized eigenvalue problem: optimal statistical rates via truncated rayleigh flow. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 80(5), 1057–1086.
- Tanay A and Regev A (2017). Scaling single-cell genomics from phenomenology to mechanism. *Nature*, 541(7637), 331–338. [PubMed: 28102262]
- Tibshirani R (2011). Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 73(Part 3), 273–282.
- Tomfohr J, Lu J, and Kepler TB (2005). Pathway level analysis of gene expression using singular value decomposition. *BMC Bioinformatics*, 6, 225. [PubMed: 16156896]
- Vines S (2000). Simple principal components. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 49(Part 4), 441–451.
- Wagner A, Regev A, and Yosef N (2016). Revealing the vectors of cellular identity with single-cell genomics. *Nat Biotechnol*, 34(11), 1145–1160. [PubMed: 27824854]
- Waltman L and van Eck NJ (2013). A smart local moving algorithm for large-scale modularity-based community detection. *The European Physical Journal B*, 86(11).
- Witten DM, Tibshirani R, and Hastie T (2009). A penalized matrix decomposition, with applications to sparse principal components and canonical correlation analysis. *Biostatistics*, 10(3), 515–534. [PubMed: 19377034]
- Young MD, Wakefield MJ, Smyth GK, and Oshlack A (2010). Gene ontology analysis for rna-seq: accounting for selection bias. *Genome Biol*, 11(2), R14. [PubMed: 20132535]
- Yuan X-T and Zhang T (2013). Truncated power method for sparse eigenvalue problems. *J. Mach. Learn. Res*, 14(1), 899–925.
- Zou H, Hastie T, and Tibshirani R (2006). Sparse principal component analysis. *Journal of Computational and Graphical Statistics*, 15(2), 265–286.



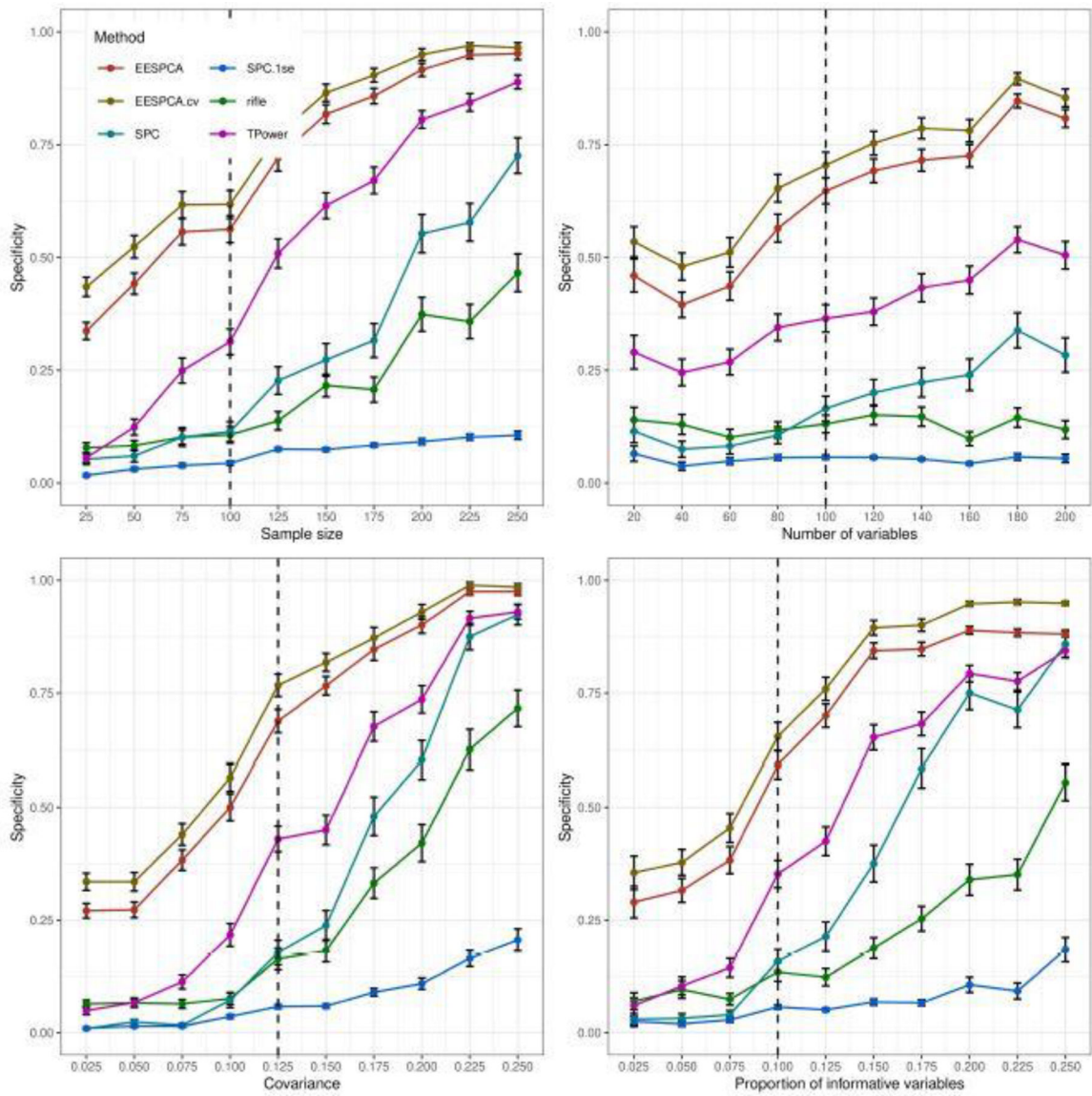
**Fig. 1.** Kernel density estimates of the distribution of the ratio of approximate-to-real PC loadings. The ratio distribution for variables that have a non-zero loading on the population PC is shown as a solid line. The ratio distribution for variables that have a zero population loading is shown as a dashed line.



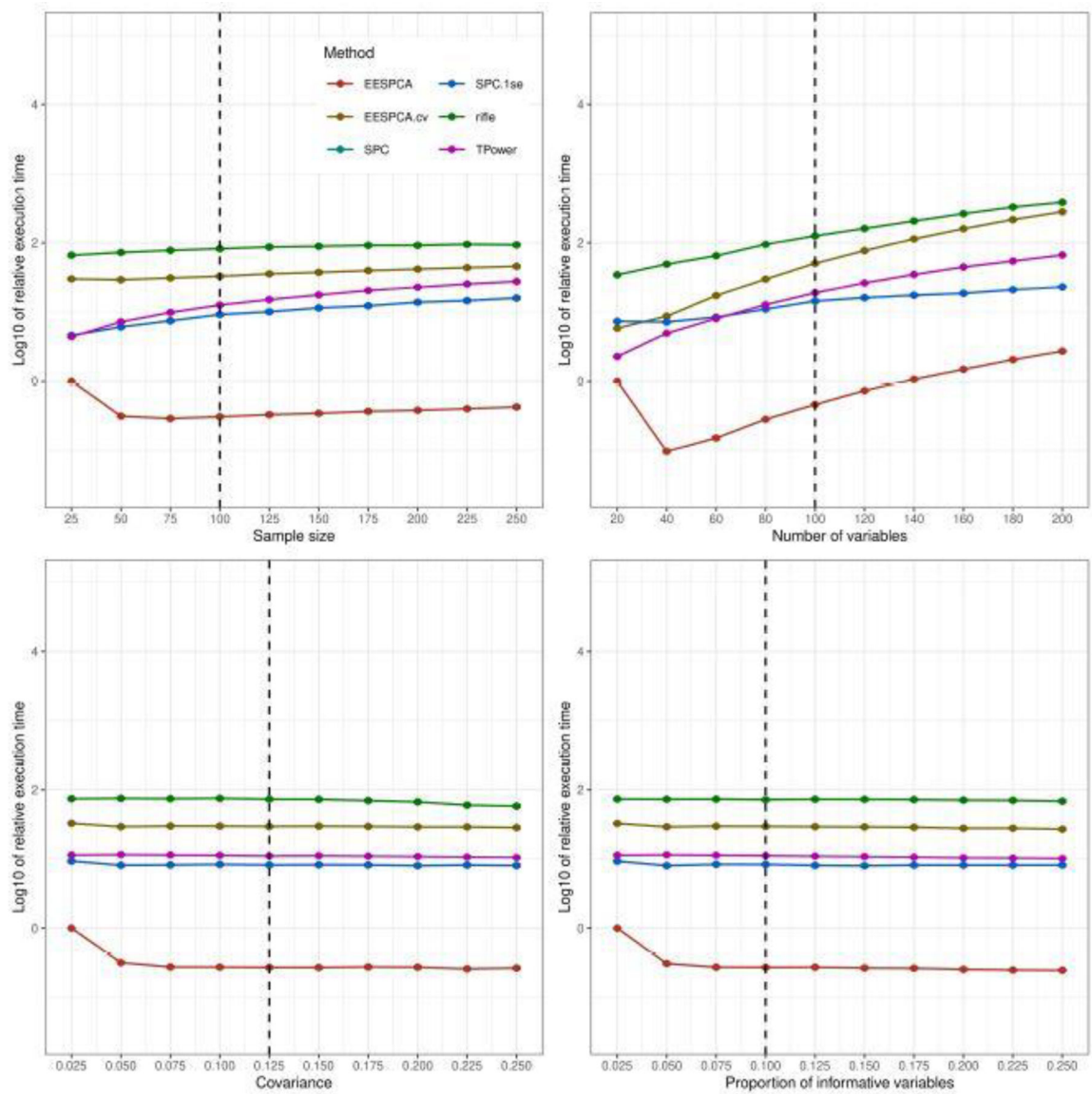
**Fig. 2.** Classification performance of EESPCA, EESPCA.cv, SPC, SPC.1se, TPower and rifle on data simulated using the basic model detailed in Section 2.3. Each panel illustrates the relationship between the balanced accuracy of the methods (i.e., the ability to correctly assign 0 loadings to variables with a 0 population loading) and one of the simulation parameters. The vertical dotted lines mark the default parameter value used in the other panels. Error bars represent  $\pm 1$  SE.



**Fig. 3.** Sensitivity of EESPCA, EESPCA.cv, SPC, SPC.1se, TPower and rifle on data simulated using the basic model detailed in Section 2.3. Each panel illustrates the relationship between sensitivity (i.e., the proportion of variables with a zero population loading that are given a zero loading in the estimated sparse PC) and one of the simulation parameters. The vertical dotted lines mark the default parameter value used in the other panels. Error bars represent  $\pm 1$  SE.

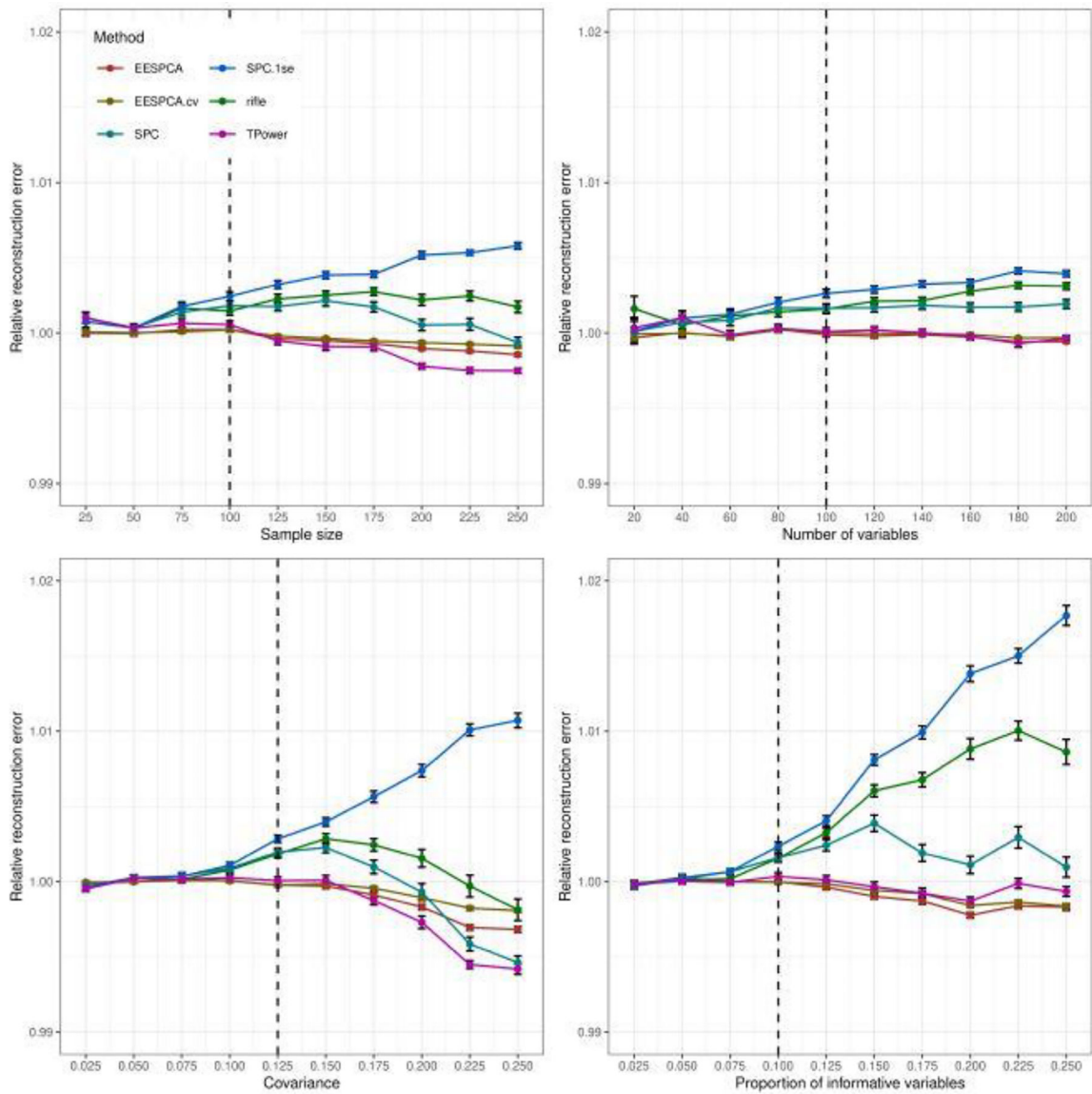


**Fig. 4.** Specificity of EESPCA, EESPCA.cv, SPC, SPC.lse, TPower and rifle on data simulated using the basic model detailed in Section 2.3. Each panel illustrates the relationship between specificity (i.e., the proportion of variables with a non-zero population loading that are given a non-zero loading in the estimated sparse PC) and one of the simulation parameters. The vertical dotted lines mark the default parameter value used in the other panels. Error bars represent  $\pm 1$  SE.

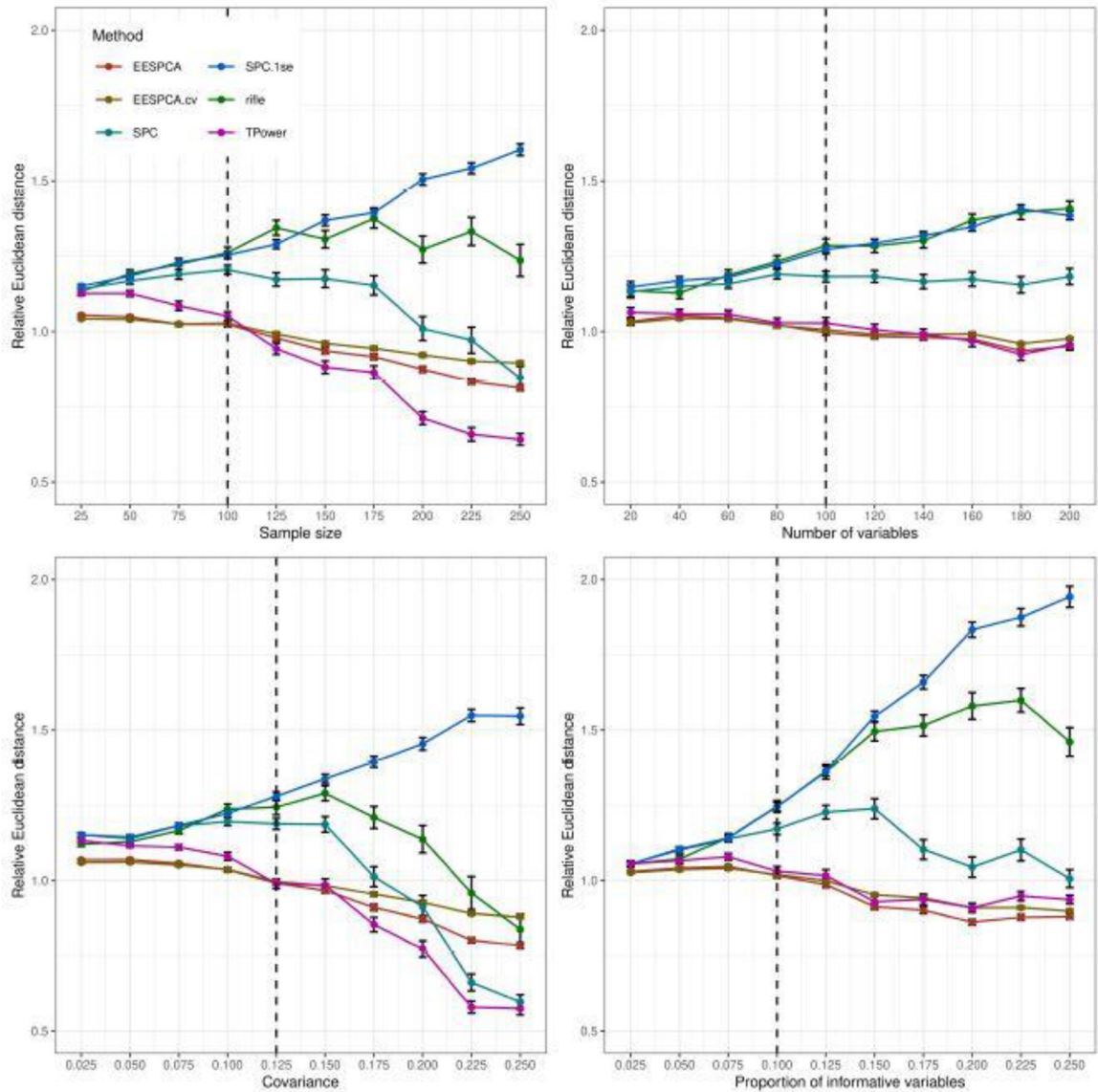


**Fig. 5.** Relative computational speed of EESPCA, EESPCA.cv, SPC, SPC.1se, TPower and rifle on data simulated using the basic model detailed in Section 2.3. Each panel illustrates the relationship between the relative computational speed of each method (i.e.,  $\log_{10}$  ratio of computational cost of the method and the cost of the EESPCA method on data simulated using the smallest parameter value shown on the x-axis) and one of the simulation parameters. The SPC and SPC.1se lines overlap given the identical computational times. The vertical dotted lines mark the default parameter value used in the other panels. Error bars represent  $\pm 1$  SE.

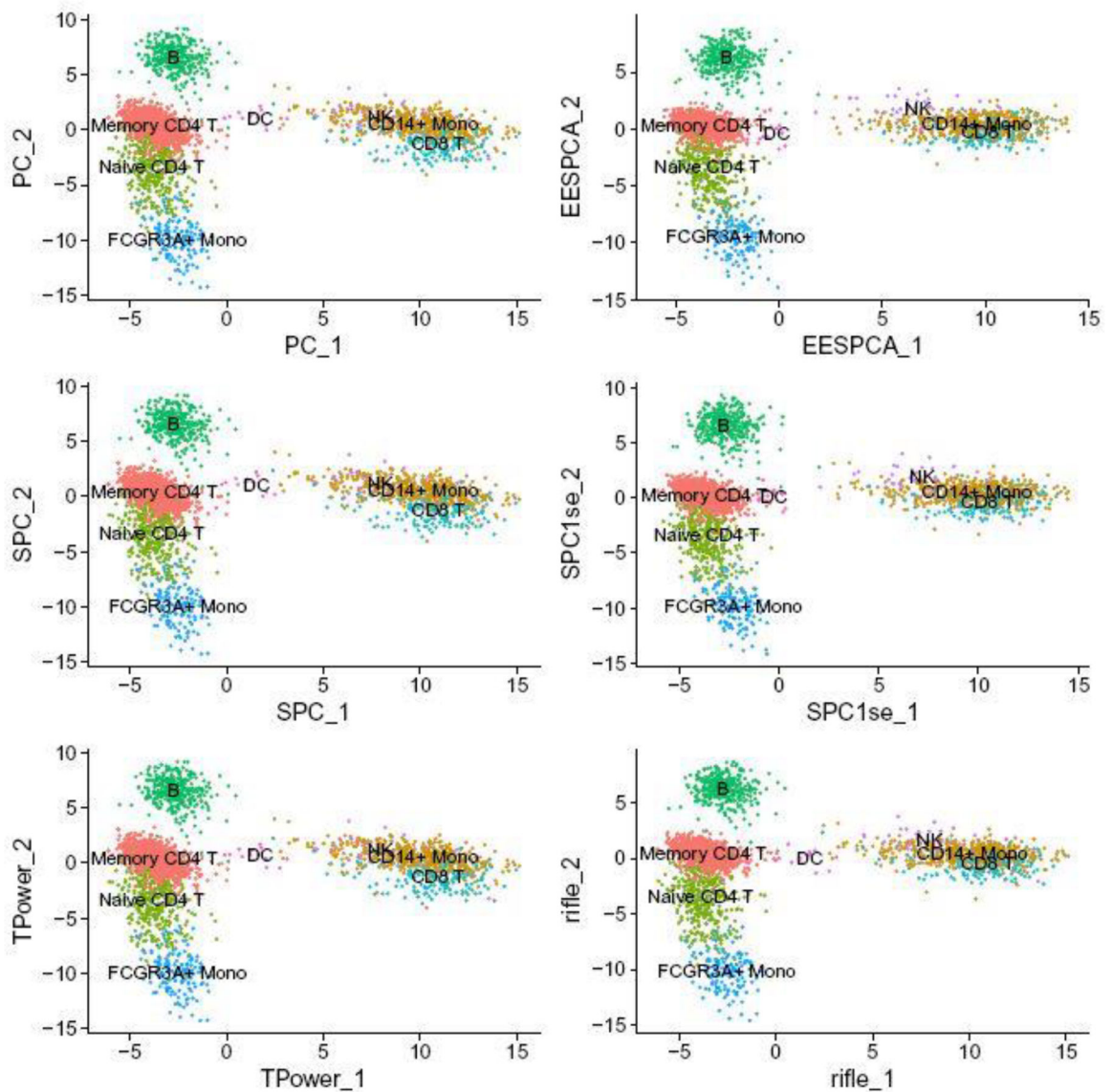




**Fig. 6.** Rank 1 out-of-sample reconstruction error of EESPCA, EESPCA.cv, SPC, SPC.1se, TPower and rifle on data simulated using the basic model detailed in Section 2.3. Each panel illustrates the relationship between the ratio of the rank 1 out-of-sample reconstruction error (i.e., squared Frobenius norm of the residual matrix formed by subtracting the rank 1 reconstruction matrix from the original matrix) for each method to the error achieved by standard PCA. The vertical dotted lines mark the default parameter value used in the other panels. Error bars represent  $\pm 1$  SE.



**Fig. 7.** Euclidean distance between the estimated and population loadings of the first PC for the basic model detailed in Section 2.3. Each panel illustrates the relationship between the ratio of the distance for each method to the distance achieved by standard PCA. The vertical dotted lines mark the default parameter value used in the other panels. Error bars represent  $\pm 1$  SE.



**Fig. 8.**

Projections of the PBMC 2k single cell RNA-sequencing data onto the first two PC computed using standard PCA or the first two sparse PCs computed using the EESPCA, SPC, SPC.1se, TPower and rifle methods (EESPCA.cv results are very similar to the EESPCA results so have been omitted to enable a 3-by-2 panel plot). Each point represents a single cell is colored and labeled according to immune cell type as detailed in the Seurat Guided Clustering Tutorial Seurat (2020).

**Table 1**

Results from standard PCA, EESPC, EESPCA.cv, SPC, SPC.lse, TPower and rifle analysis of the PBMC scRNA-seq data following the procedure detailed in Section 2.4. The table lists the number of the 1,000 genes in the input data matrix that were assigned positive or negative loadings. Since standard PCA produces non-zero loadings for all genes, the magnitude of the difference between the PCA counts and the counts for the other methods reflects the relative sparsity of the solutions. The “sig. GO” columns capture the number of Gene Ontology Biological Process terms that were significantly enriched among the genes with either positive or negative loadings at an FDR of 0.1. The relative reconstruction error captures average out-of-sample reconstruction error measured over 10 random splits of the data relative to the error for standard PCA (standard deviation is included in parentheses). For each split, PCs were estimated using each method on half of the data and the squared Frobenius norm of the residual matrix for the first two PCs was computed for the other half of the data.

|           | PC 1     |         |          |         | PC 2     |         |          |         | Relative        |
|-----------|----------|---------|----------|---------|----------|---------|----------|---------|-----------------|
|           | Positive |         | Negative |         | Positive |         | Negative |         | reconstruction  |
|           | genes    | sig. GO | genes    | sig. GO | genes    | sig. GO | genes    | sig. GO | error           |
| PCA       | 558      | 44      | 442      | 0       | 316      | 0       | 684      | 0       | 1 (0)           |
| EESPCA    | 134      | 104     | 12       | 0       | 35       | 38      | 83       | 3       | 1.004 (0.0005)  |
| EESPCA.cv | 166      | 89      | 22       | 0       | 41       | 33      | 129      | 14      | 1.003 (0.0004)  |
| SPC       | 535      | 37      | 418      | 0       | 316      | 0       | 684      | 0       | 0.996 (0.003)   |
| SPC.lse   | 193      | 84      | 36       | 0       | 91       | 21      | 394      | 1       | 1.003 (0.0002)  |
| TPower    | 485      | 38      | 357      | 0       | 226      | 3       | 616      | 0       | 1.0002 (0.0003) |
| rifle     | 485      | 38      | 357      | 0       | 268      | 2       | 574      | 2       | 1.002 (0.002)   |

**Table 2**

Gene Ontology (GO) enrichment results for positive and negatives loadings on first two PCs generated by standard PCA, EESPC, SPC, and TPower analysis of the PBMC scRNA-seq data. The top five GO Biological Process terms with an FDR of 0.1 are listed. Similar results for EESPCA.cv, SPC.lse and rifle are included Table S1 in the Supplementary Results.

|                    | Gene Ontology term  | # genes<br>in term | # genes<br>in group | FDR      |
|--------------------|---|--------------------|---------------------|----------|
| PCA (PC 1, pos)    | secretion   | 139                | 113                 | 1.88e-07 |
|                    | secretion by cell   | 132                | 107                 | 1.06e-06 |
|                    | vesicle-mediated transport  | 172                | 133                 | 2.08e-06 |
|                    | exocytosis  | 101                | 84                  | 1.30e-05 |
|                    | immune effector process   | 129                | 102                 | 4.95e-05 |
| EESPCA (PC 1, pos) | immune response   | 204                | 70                  | 8.82e-15 |
|                    | immune system process   | 256                | 77                  | 4.54e-13 |
|                    | defense response  | 144                | 53                  | 5.22e-11 |
|                    | immune effector process   | 129                | 49                  | 2.04e-10 |
|                    | neutrophil activation   | 69                 | 34                  | 6.33e-10 |
| EESPCA (PC 2, pos) | adaptive immune response  | 51                 | 17                  | 1.69e-10 |
|                    | antigen receptor-mediated signaling pathway                                 | 35                 | 12                  | 2.11e-06 |
|                    | antigen processing and presentation of exogenous peptide antigen via MHC... | 16                 | 9                   | 2.34e-06 |
|                    | antigen processing and presentation of peptide antigen via MHC class II     | 17                 | 9                   | 4.85e-06 |
|                    | antigen processing and presentation of peptide or polysaccharide antigen... | 17                 | 9                   | 4.85e-06 |
| EESPCA (PC 2, neg) | regulation of transport   | 103                | 23                  | 0.0057   |
|                    | regulation of localization  | 154                | 28                  | 0.024    |
|                    | positive regulation of transport  | 59                 | 16                  | 0.027    |
| SPC (PC 1, pos)    | secretion   | 139                | 111                 | 6.94e-08 |
|                    | vesicle-mediated transport  | 172                | 131                 | 3.53e-07 |
|                    | secretion by cell   | 132                | 105                 | 4.76e-07 |
|                    | exocytosis  | 101                | 84                  | 7.21e-07 |
|                    | regulated exocytosis  | 97                 | 80                  | 5.16e-06 |
| TPower (PC 1, pos) | vesicle-mediated transport  | 172                | 126                 | 8.57e-09 |
|                    | secretion   | 139                | 105                 | 4.50e-08 |
|                    | secretion by cell   | 132                | 100                 | 1.23e-07 |
|                    | exocytosis  | 101                | 80                  | 3.79e-07 |
|                    | immune effector process   | 129                | 96                  | 2.14e-06 |
| TPower (PC 2, pos) | adaptive immune response  | 51                 | 27                  | 0.0083   |
|                    | humoral immune response   | 22                 | 14                  | 0.020    |
|                    | B cell receptor signaling pathway   | 8                  | 8                   | 0.044    |