Nordhausen, K.; Virta, J.

An overview of properties and extensions of FOBI

Published in:
KNOWLEDGE-BASED SYSTEMS

DOI:
10.1016/j.knosys.2019.02.026

Published: 01/06/2019

Document Version
Peer reviewed version

An overview of properties and extensions of FOBI

K. Nordhausen[a,*], J. Virta[b]

[a] *Institute of Statistics & Mathematical Methods in Economics, Vienna University of Technology, Wiedner Hauptstr. 7, 1040 Vienna, Austria*
[b] *Department of Mathematics and Systems Analysis, Aalto University, P.O. Box 11100, 00076 Aalto, Finland*

Abstract

Recently, Spurek, Tabor, Struksi and Smieja (2018, Knowledge-Based Systems, doi 10.1016/k.knosys.2018.07.027) suggested an independent component analysis method that is obtained via a generalized eigenvalue decomposition and stated that their estimator enjoys several useful properties. The estimator is however already known since 1989 as FOBI (Fourth order blind identification), and it indeed has many nice properties, even outside the independent component model. As there seems to be interest in the readers of Knowledge-Based Systems in the estimator, we shortly review it and various of its properties not mentioned by Spurek et al. (2018), along with some recent extensions of FOBI.

Keywords: Independent component analysis, Invariant coordinate selection, independence property

2010 MSC: 62H12

1. Independent Component Analysis

Independent component analysis (ICA) is a well-established multivariate method introduced in the mid-1980's (for an account of its early history, see [1]). The first driving force of ICA was the computer science and signal processing community but in the recent years also statisticians have gotten more and more

*Corresponding author
Email addresses: klaus.nordhausen@tuwien.ac.at (K. Nordhausen), joni.virta@aalto.fi (J. Virta)

Preprint submitted to Knowledge-Based Systems *February 19, 2019*

interested in ICA, i.e., deriving statistical properties of the existing ICA methods and suggesting new estimators. For some overviews, see for example [2, 3, 4].

The most basic model for ICA is

$$\mathbf{x} = \boldsymbol{\Omega}\mathbf{z} + \boldsymbol{\mu}, \tag{1}$$

where $\mathbf{x}$ is an observable p-variate vector which is a linear mixture of the latent p-variate vector $\mathbf{z}$ that has independent components. The full rank $p \times p$ matrix $\boldsymbol{\Omega}$ represents the mixing process and the p-vector $\boldsymbol{\mu}$ is the location of the data and usually considered a nuisance parameter. The goal of ICA is to find, based on $\mathbf{x}$ alone, an unmixing matrix $\boldsymbol{\Gamma}$ such that $\boldsymbol{\Gamma}^\top(\mathbf{x} - \boldsymbol{\mu})$ has independent components.

The literature knows meanwhile an abundance of ICA methods which almost always consist of the following steps:

1. Whitening: $\mathbf{x}_{st} = \mathrm{COV}(\mathbf{x})^{-1/2}[\mathbf{x} - \mathrm{E}(\mathbf{x})]$
2. Rotation: Find a $p \times p$ orthogonal matrix $\mathbf{U}$ such that $\mathbf{U}^\top \mathbf{x}_{st}$ has independent components.
3. Unmixing matrix: $\boldsymbol{\Gamma} = \mathrm{COV}(\mathbf{x})^{-1/2}\mathbf{U}$.

The difference between the methods is then in how the rotation $\mathbf{U}$ is obtained.

Recently, [5] suggested an ICA method which they denoted weighted ICA (WeICA) and which also follows the above principle. WeICA obtains $\mathbf{U}$ as the eigenvectors of the matrix $\mathbf{M}(\mathbf{x}_{st}) = \mathrm{E}\left(\mathbf{x}_{st}^\top \mathbf{x}_{st} \mathbf{x}_{st} \mathbf{x}_{st}^\top\right)$, the eigenvalues $\lambda_1, \ldots, \lambda_p$ of which are functions of the fourth moments of the corresponding independent components. The authors of [5] stated that the main advantages of WeICA are that (i) it is fast to compute (ii) it is affine equivariant, (iii) it has a closed-form solution and (iv) it can be used to conduct dimension reduction by retaining only a subset of components which are maximally non-Gaussian, i.e., the components whose eigenvalues most deviate from the value of a Gaussian component, $p+2$. In Section 2.3 we review the use of symmetrized scatter matrices in ICA and it turns out that the above properties are not particular to WeICA and by sacrificing property (i), a whole class of ICA-methods with properties (ii)-(iv) in addition to various others is obtained.

2. Fourth order blind identification

The problem of finding the eigenvectors $\mathbf{U}$ of $\mathbf{M}(\mathbf{x}_{st})$ can be written as $\mathbf{U}^\top \mathrm{COV}(\mathbf{x})^{-1/2}\mathrm{COV}_4(\mathbf{x})\mathrm{COV}(\mathbf{x})^{-1/2}\mathbf{U} = \mathbf{\Lambda}$, where $\mathbf{\Lambda}$ is a diagonal matrix of eigenvalues in decreasing order and

$$\mathrm{COV}_4(\mathbf{x}) = \mathrm{E}\left\{[\mathbf{x} - \mathrm{E}(\mathbf{x})]^\top \mathrm{COV}(\mathbf{x})^{-1}[\mathbf{x} - \mathrm{E}(\mathbf{x})][\mathbf{x} - \mathrm{E}(\mathbf{x})][\mathbf{x} - \mathrm{E}(\mathbf{x})]^\top\right\}$$

is the so-called matrix of fourth moments (note that in some references $\mathrm{COV}_4(\mathbf{x})$ is divided by $p + 2$ to make it consistent for COV at the multivariate normal model). COV_4 can be seen as a member of the class of functionals of the form $\mathrm{E}\left\{g\left[\tilde{\mathbf{x}}^\top \mathrm{COV}(\mathbf{x})^{-1}\tilde{\mathbf{x}}\right]\tilde{\mathbf{x}}\tilde{\mathbf{x}}^\top\right\}$, where $\tilde{\mathbf{x}} = \mathbf{x} - \mathrm{E}(\mathbf{x})$ and $g : \mathbb{R}_+ \to \mathbb{R}_+$ is an arbitrary positive weight function (and these functionals are themselves members of the even wider class of one-step M-estimators of scatter). As $\mathbf{U}$ is constrained to have orthonormal columns, the problem can still be expressed in the guise of a generalized eigendecomposition as,

$$\mathbf{\Gamma}^\top \mathrm{COV}(\mathbf{x})\mathbf{\Gamma} = \mathbf{I}_p \quad \text{and} \quad \mathbf{\Gamma}^\top \mathrm{COV}_4(\mathbf{x})\mathbf{\Gamma} = \mathbf{\Lambda}, \tag{2}$$

where $\mathbf{\Gamma} = \mathrm{COV}(\mathbf{x})^{-1/2}\mathbf{U}$. The solution of the above *simultaneous diagonalization* problem is actually one of the oldest, and still one of the most popular, ICA estimators known as FOBI (fourth order blind identification), introduced already in [6].

In recent years, the properties of FOBI have been intensively studied and numerous generalizations of it can be found in the literature. These studies also suggest that the FOBI-methodology can be highly useful for complex non-standard data structures and outside of the independent component model. As it seems that FOBI is not that well-known in this community, we will in the following review these recent results grouped by theme. Note that in the software R [7], FOBI is available, for example, via the R packages ICS [8] and JADE [9] and the dimension tests discussed below are implemented in the package ICtest [10], the extension for time series data in the package tsBSS [11] and the extensions to tensorial observations and time series in the package tensorBSS [12].

2.1. Assumptions of FOBI

First, as the eigenvalues $\lambda_1 \geq \ldots \geq \lambda_p$ in $\mathbf{\Lambda}$ are given as $\kappa_i + p + 2$, where κ_i is the kurtosis of the ith independent component, it is clear that the FOBI-estimator is well-defined only if all independent components have finite fourth moments and distinct kurtosis values. These requirements are actually the two biggest points of criticism for FOBI as they exclude most heavy-tailed distributions and forbid two or more components having the same non-Gaussian distribution. Multiple extensions of ICA have been suggested that avoid one or both of these requirements. Some popular ones include, for example, JADE [13, 14] and fastICA [15, 16]. With the trade-off of increased computational complexity, the moment assumption can also be avoided by using the symmetrized scatter approach discussed in Section 2.3.

2.2. Statistical inference in FOBI

The large-sample properties of various ICA-estimators have also been discussed in the literature. [17, 14] established the limiting normality of the FOBI-estimator and give exact expressions for its limiting variances (simultaneously increasing the number of needed finite moments to eight). [14] showed that the asymptotic variances of FOBI are always equal or larger than the corresponding asymptotic variances of symmetric fastICA (using kurtosis as a non-linearity). Thus the performance of FOBI is asymptotically never better than that of symmetric fastICA and [14] argue also that FOBI is generally worse than JADE.

Note that even though independent components with identical kurtoses cannot be separated from each other by FOBI, it is nevertheless common to include multiple Gaussian components in $\mathbf{z}$ to represent noise. Indeed, one of the objectives of FOBI pointed out by [5] is its use in identifying and retaining only the non-Gaussian components in order to achieve dimension reduction.

An inferential treatise of this problem was given in [18, 19] via the use of the noisy independent component model,

$$\mathbf{x} = \mathbf{\Omega}_1 \mathbf{z}_1 + \mathbf{\Omega}_2 \mathbf{z}_2 + \boldsymbol{\mu},$$

where $\mathbf{z}_1$ is a d-vector with standardized non-Gaussian independent components ("the signal") and the vector $\mathbf{z}_2$ is independent of $\mathbf{z}_1$ and obeys the $p-d$-variate standard normal distribution ("the noise"). The mixing matrices $\mathbf{\Omega}_1$ and $\mathbf{\Omega}_2$ are of dimensions $p \times d$ and $p \times (p-d)$, respectively. The goal in the model is to estimate the unknown dimensionality d and the signal $\mathbf{z}_1$. The idea of [5] to retain only the components whose FOBI eigenvalues are most distinct from $p+2$ was already formalized in [18, 19]. To test the null hypothesis

$$H_{0,k}: \text{ exactly } p-k \text{ components are Gaussian,}$$

they suggested using as a test statistic the variance of the most Gaussian eigenvalues around the value $p+2$,

$$T_k = \frac{1}{p-k} \sum_{i=1}^{p-k} \left(\tilde{\lambda}_i - (p+2) \right)^2,$$

where $\tilde{\lambda}_i$ are the eigenvalues λ_i ordered in increasing order according to $(\lambda_i - (p+2))^2$. Under the noisy independent component model and assuming that no signal has zero kurtosis, [18] showed that if the sample size increases, $n \to \infty$,

$$n(p-k)T_k \rightsquigarrow 2\sigma_1 \chi^2_{(p-k-1)(p-k+2)/2} + (2\sigma_1 + 4(p-k))\, \chi^2_1,$$

where $\chi^2_{(p-k-1)(p-k+2)/2}$ and χ^2_1 are independent chi-squared random variables and $\sigma_1 = \sum_{i=1}^{p} E(z_i^4) - p + 8$. The constant σ_1 can now be consistently estimated from the data, allowing one to estimate the signal dimension by successively applying the hypothesis tests. [18, 19] proposed also bootstrap-based alternatives to the asymptotic test and [20] deployed a different bootstrapping strategy to estimate the signal dimension using FOBI. A special case of the above, a test for multivariate normality ($k = 0$), based on FOBI was already suggested in [21]. Moreover, the test can also be generalized to hold in a non-Gaussian component analysis (NGCA) model as shown in [19].

2.3. Generalizations of FOBI in an ICA context

The most obvious starting point for generalizations of FOBI is through the simultaneous diagonalization (2). [22, 23] showed that replacing COV and COV_4

with any two *scatter matrices* with the *independence property* produces a new affine equivariant ICA estimator. A scatter matrix $\mathbf{S}$ is any positive semi-definite matrix-valued functional which is affine equivariant in the sense that $\mathbf{S}(\mathbf{A}\mathbf{x}+\mathbf{b}) = \mathbf{A}\mathbf{S}(\mathbf{x})\mathbf{A}^\top$ for all full rank $p \times p$ matrices $\mathbf{A}$ and all p-vectors $\mathbf{b}$. A scatter matrix is said to possess the independence property if it is diagonal for all random vectors with independent components. Although examples of scatter matrices that have the independence property, besides COV and COV_4, are difficult to find, [22] showed that *symmetrizing* a scatter matrix gives it the independence property. E.g., robust (but computationally costly) ICA-estimators are then easily obtained through simultaneous diagonalization of two symmetrized robust scatter matrices, see [23, 24]. Moreover, also in this case Gaussian components will have equal eigenvalues, just that their value might not equal $p+2$ and therefore dimension reduction as in (iv) is again feasible.

Thus, depending on whether one counts (2) as a closed-form solution, the estimators produced by symmetrized scatter matrices have the properties (ii)-(iv) discussed in the introduction and can be seen to trade the property (i) for improved performance in other fields, such as robustness or efficiency.

FOBI has also been generalized in an alternative direction by adapting it to work with non-conventional types of data. Although ICA is a special case of blind source separation (BSS) and designed for independent and identically distributed (i.i.d.) data, many BSS applications use data that are actually time series. A generalization of FOBI (gFOBI) to multivariate time series was introduced in [25], where the fourth moment matrix $\mathbf{M}(\mathbf{x})$ was modified to incorporate serial information as $\mathbf{M}_\tau(\mathbf{x}_t) = E[\mathbf{x}_t^\top \mathbf{x}_t \mathbf{x}_{t+\tau} \mathbf{x}_{t+\tau}^\top]$. The rotation $\mathbf{U} = (\mathbf{u}_1, \ldots, \mathbf{u}_p)$ is then found as the orthogonal matrix which maximizes the joint diagonality condition,

$$\sum_{\tau \in \mathcal{T}} \sum_{i=1}^{p} (\mathbf{u}_i^\top \mathbf{M}_\tau(\mathbf{x}_t^{st}) \mathbf{u}_i)^2,$$

for a chosen set of lags $\mathcal{T} = \{\tau_1, \ldots, \tau_K\}$. The lag choice $\mathcal{T} = \{0\}$ reduces the method to the regular FOBI.

Besides multivariate time series, FOBI-based extensions have recently been suggested also for tensor-valued data, both in context of ICA and BSS. FOBI for tensor-valued i.i.d. data (TFOBI) was introduced in [26]. The tensorial FOBI applies to tensor data of all dimensions but due to the complexity of tensor notation, we consider in the following only the matrix version of the algorithm (the general tensor case can actually be reduced to the matrix case through the so-called tensor matricization of "flattening" [26]). We assume that the observed $p \times q$ random matrices are independent realization of the model

$$\mathbf{X} = \mathbf{\Omega}_1 \mathbf{Z} \mathbf{\Omega}_2^\top + \mu, \tag{3}$$

where $\mathbf{\Omega}_1 \in \mathbb{R}^{p \times p}$, $\mathbf{\Omega}_2 \in \mathbb{R}^{q \times q}$ are full rank mixing matrices, $\mu \in \mathbb{R}^{p \times q}$ is the location of the data and the latent matrix $\mathbf{Z} \in \mathbb{R}^{p \times q}$ has independent components. The TFOBI-solution for estimating $\mathbf{Z}$ given only $\mathbf{X}$ is based on treating both sides of the model separately and, for example, the solution of the left-hand side relies upon the following extensions of the covariance matrix and the matrix of fourth moments,

$$\mathrm{TCOV}(\mathbf{X}) = \mathrm{E}\left(\mathbf{X}\mathbf{X}^\top\right) \quad \mathbf{TM}(\mathbf{X}) = \mathrm{E}\left(\mathbf{X}\mathbf{X}^\top \mathbf{X}\mathbf{X}^\top\right)$$

where $\mathbf{X}$ is assumed to be centered. Interestingly, while seemingly naïve extensions, TCOV and $\mathbf{TM}$, and consequently the TFOBI-estimator, share many of the properties of the FOBI-estimator.

The time series method gFOBI was in turn generalized for tensor-valued time series (TgFOBI) in [27] and it is in the special case of matrices based on a model similar to (3), but where both $\mathbf{X}$ and $\mathbf{Z}$ depend on a time index t (i.e., they are matrix-valued time series). Analogously to the relationship between FOBI and gFOBI, the TgFOBI-algorithm is based on the joint diagonalization of the following lagged versions of $\mathbf{TM}$,

$$\mathbf{TM}_\tau(\mathbf{X}_t) = \mathrm{E}\left(\mathbf{X}_{t+\tau}\mathbf{X}_t^\top \mathbf{X}_t \mathbf{X}_{t+\tau}^\top\right)$$

for a given set of lags $\mathcal{T} = \{\tau_1, \ldots, \tau_K\}$. The lag choice $\mathcal{T} = \{0\}$ reverts the method then back to TFOBI.

Finally, FOBI for functional data was suggested in [28]. In functional FOBI we assume that the observed multivariate random function is a linear transformation of a latent multivariate random function with independent component functions, with the objective of estimating the latent function given only the observed one. We refrain from giving additional details here due to the large amounts of required extra notation, and the interested reader is referred to [28].

2.4. Usage of FOBI outside of ICA

If one replaces COV and COV_4 in the FOBI-equation (2) with arbitrary scatter matrices $\mathbf{S}_1$ and $\mathbf{S}_2$, not necessary having the independence property, and steps outside of the IC-model (1), general results can still be obtained. Actually, the use of such procedures already has a quite long tradition in multivariate statistics. Early works are for example [29, 30, 31]. A unified framework was introduced as invariant coordinate selection (ICS) in [32] who considered (2) in a model-free context for any possible scatter combination.

Regardless of the chosen scatters, ICS creates an affine invariant coordinate system and can be used for model selection, dimension reduction, multivariate outlier detection or as a pre-processing step for clustering. For example, [32] proved that ICS finds Fisher's linear subspace without knowing the cluster labels in an elliptical mixture model. Also in ICS, the FOBI combination $\mathbf{S}_1 = \mathrm{COV}$ and $\mathbf{S}_2 = \mathrm{COV}_4$ is very popular, see for example [33, 31, 8, 34, 35, 36, 37] for further details of FOBI in this context.

3. Conclusion

Recently [5] rediscovered FOBI, a well-known ICA estimator with many nice properties established mainly in the statistics literature in the previous years. Despite being quite inefficient when compared to other ICA methods, the simplicity, ease of computation and numerous useful properties (both outside and inside the IC model) of FOBI help explain its persisting popularity. These aspects also make FOBI the perfect method for generalizing ICA to complex non-standard data structures, as the many extensions and generalizations proposed in recent years show. In this paper we gave a brief overview of FOBI and the recent findings related to it, which we think are also of interest outside the field of statistics.

Acknowledgements

The work of KN was partly supported by CRoNoS COST Action IC1408.

References

[1] C. Jutten, A. Taleb, Source separation: From dusk till dawn, in: Proceedings of the International Symposium on Independent Component Analysis and Blind Signal Separation, 2000, pp. 15–26.

[2] A. Hyvärinen, J. Karhunen, E. Oja, Independent component analysis, John Wiley & Sons, Inc.: New York, 2001.

[3] P. Comon, C. Jutten, Handbook of Blind Source Separation: Independent Component Analysis and Applications, Academic Press, 2010.

[4] K. Nordhausen, H. Oja, Independent component analysis: A statistical perspective, Wiley Interdisciplinary Reviews: Computational Statistics (2018) e1440.

[5] P. Spurek, J. Tabor, L. Struski, M. Smieja, Fast independent component analysis algorithm with a simple closed-form solution, Knowledge-Based Systems 161 (2018) 26–34. `doi:https://doi.org/10.1016/j.knosys.2018.07.027`.

[6] J.-F. Cardoso, Source separation using higher order moments, in: International Conference on Acoustics, Speech, and Signal Processing, 1989, 1989, pp. 2109–2112.

[7] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria (2017).

[8] K. Nordhausen, H. Oja, D. E. Tyler, Tools for exploring multivariate data: The package ICS, Journal of Statistical Software 28 (6) (2008) 1–31.

[9] J. Miettinen, K. Nordhausen, S. Taskinen, Blind source separation based on joint diagonalization in R: The packages JADE and BSSasymp, Journal of Statistical Software 76 (2) (2017) 1–31.

[10] K. Nordhausen, H. Oja, D. E. Tyler, J. Virta, ICtest: Estimating and Testing the Number of Interesting Components in Linear Dimension Reduction, R package version 0.3 (2017).

[11] M. Matilainen, C. Croux, J. Miettinen, K. Nordhausen, H. Oja, S. Taskinen, J. Virta, tsBSS: Blind Source Separation and Supervised Dimension Reduction for Time Series, R package version 0.5.2 (2018).

[12] J. Virta, B. Li, K. Nordhausen, H. Oja, tensorBSS: Blind Source Separation Methods for Tensor-Valued Observations, R package version 0.3.4 (2018).

[13] J.-F. Cardoso, A. Souloumiac, Blind beamforming for non-Gaussian signals, in: IEE Proceedings F-Radar and Signal Processing, Vol. 140, 1993, pp. 362–370.

[14] J. Miettinen, S. Taskinen, K. Nordhausen, H. Oja, Fourth moments and independent component analysis, Statistical Science 30 (3) (2015) 372–390.

[15] A. Hyvärinen, E. Oja, A fast fixed-point algorithm for independent component analysis, Neural Computation 9 (7) (1997) 1483–1492.

[16] J. Miettinen, K. Nordhausen, H. Oja, S. Taskinen, J. Virta, The squared symmetric FastICA estimator, Signal Processing 131 (2017) 402–411.

[17] P. Ilmonen, J. Nevalainen, H. Oja, Characteristics of multivariate distributions and the invariant coordinate system, Statistics & Probability Letters 80 (2010) 1844–1853.

[18] K. Nordhausen, H. Oja, D. E. Tyler, Asymptotic and bootstrap tests for subspace dimension, arXiv preprint arXiv:1611.04908.

[19] K. Nordhausen, H. Oja, D. E. Tyler, J. Virta, Asymptotic and bootstrap tests for the dimension of the non-Gaussian subspace, IEEE Signal Processing Letters 24 (6) (2017) 887–891.

[20] W. Luo, B. Li, Combining eigenvalues and variation of eigenvectors for order determination, Biometrika 103 (4) (2016) 875–887.

[21] A. Kankainen, S. Taskinen, O. H., Tests of multinormality based on location vectors and scatter matrices, Statistical Methods and Applications 16 (2007) 357–379.

[22] H. Oja, S. Sirkiä, J. Eriksson, Scatter matrices and independent component analysis, Australian Journal of Statistics 35 (2006) 175–189.

[23] K. Nordhausen, H. Oja, E. Ollila, Robust independent component analysis based on two scatter matrices, Austrian Journal of Statistics 37 (2008) 91–100.

[24] K. Nordhausen, D. E. Tyler, A cautionary note on robust covariance plug-in methods, Biometrika 102 (2015) 573–588.

[25] M. Matilainen, K. Nordhausen, H. Oja, New independent component analysis tools for time series, Statistics & Probability Letters 105 (2015) 80–87.

[26] J. Virta, B. Li, K. Nordhausen, H. Oja, Independent component analysis for tensor-valued data, Journal of Multivariate Analysis 162 (2017) 172–192.

[27] J. Virta, K. Nordhausen, Blind source separation of tensor-valued time series, Signal Processing 141 (2017) 204–216.

[28] J. Virta, B. Li, K. Nordhausen, H. Oja, Independent component analysis for multivariate functional data, arXiv preprint arXiv:1712.07641.

[29] H. Caussinus, A. Ruiz, Interesting projections of multidimensional data by means of generalized principal component analysis, in: K. Momirovic, V. Mildner (Eds.), Proceedings of COMPSTAT 90, Physica Verlag, Heidelberg, Germany, 1990, pp. 121–126.

[30] H. Caussinus, A. Ruiz-Gazen, Metrics for finding typical structures by means of principal component analysis, in: Y. Escoufier, C. Hayashi (Eds.), Data Science and Its Applications, Academic Press, Tokyo, 1995, pp. 177–192.

[31] D. Pena, F. J. Prieto, J. Viladomat, Eigenvectors of a kurtosis matrix as interesting directions to reveal cluster structure, Journal of Multivariate Analysis 101 (9) (2010) 1995–2007.

[32] D. E. Tyler, F. Critchley, L. Dümbgen, H. Oja, Invariant co-ordinate selection, Journal of the Royal Statistical Society B 71 (2009) 549–592.

[33] K. Nordhausen, H. Oja, E. Ollila, Multivariate models and the first four moments, in: D. R. Hunter, D. S. R. Richards, J. L. Rosenberger (Eds.), Nonparametric Statistics and Mixture Models, World Scientific, Singapore, 2011, pp. 267–287.

[34] F. Alashwali, J. Kent, The use of a common location measure in the invariant coordinate selection and projection pursuit, Journal of Multivariate Analysis 152 (2016) 145–161.

[35] D. Fischer, M. Honkatukia, M. Tuiskula-Haavisto, K. Nordhausen, D. Cavero, R. Preisinger, J. Vilkki, Subgroup detection in genotype data using invariant coordinate selection, BMC Bioinformatics 18 (2017) 173.

[36] A. Archimbaud, K. Nordhausen, A. Ruiz-Gazen, ICS for multivariate outlier detection with application to quality control, Computational Statistics & Data Analysis 128 (2018) 184–199.

[37] A. Archimbaud, K. Nordhausen, A. Ruiz-Gazen, ICSOutlier: Unsupervised Outlier Detection for Low- Dimensional Contamination Structure, The R Journal 10 (2018) 234–250.