

Outlier detection in multivariate functional data based on a geometric aggregation

Clément Lejeune*
 IRIT UMR 5505 CNRS,
 Airbus Commercial Aircraft
 Toulouse, France
 clement.lejeune@irit.fr

Josiane Mothe
 IRIT UMR 5505 CNRS, INSPE,
 Université de Toulouse
 orcid:0000-0001-9273-2193, France
 josiane.mothe@irit.fr

Olivier Teste
 IRIT UMR 5505 CNRS
 Toulouse, France
 olivier.teste@irit.fr

ABSTRACT

The increasing ubiquity of multivariate functional data (MFD) requires methods that can properly detect outliers within such data, where a sample corresponds to $p > 1$ parameters observed with respect to (w.r.t) a continuous variable (e.g. time). We improve the outlier detection in MFD by adopting a geometric view on the data space while combining the new data representation with state-of-the-art outlier detection algorithms. The geometric representation of MFD as paths in the p -dimensional Euclidean space enables to implicitly take into account the correlation w.r.t the continuous variable between the parameters. We experimentally show that our method is robust to various rates of outliers in the training set when fitting the outlier detection model and can detect outliers which are not detected by standard algorithms.

1 INTRODUCTION

1.1 Functional data context and taxonomy

In many fields (e.g. engineering, biology or medicine), detecting atypical behaviors of complex systems enables to better anticipate and understand both undesired and rare situations (e.g. engine failure, heart disease). Most of the time, detecting atypical behaviors requires the analysis of p system parameters ($p \geq 1$) measured by high sampling-rate sensors. The raw sensor measurements result in noisy data dependent on a continuous variable (e.g. time, wavelength) being discretized by the sampling process of the p sensors. Such data are referred as *univariate* or *multivariate functional data* depending on whether one ($p = 1$) or several parameters ($p > 1$) are analyzed, respectively.

Thus the observation of the parameters along the continuous variable is seen as the realization of an underlying (unknown) function that values in \mathbb{R}^p . We emphasize that in the functional data framework, a data set sample is represented as a function rather than a high-dimensional vector of different dimension containing the raw measurements. Dimension refers to the number of measurements which can be different from a sample to another). We refer to [13] for a comprehensive introduction to functional data analysis.

Here we adopt the following notations : the dependent continuous variable is denoted by $t \in \mathcal{T} \subset \mathbb{R}$ where \mathcal{T} is a closed interval of \mathbb{R} , the data samples are sub-scripted by $i \in \{1 \dots n\}$, univariate functional data (UFD) samples are denoted by lower case letter $x_i(t) \in \mathbb{R}$ and multivariate functional data (MFD) samples are denoted by capital letter $X_i(t) = (x_{i1}(t), \dots, x_{ik}(t), \dots, x_{ip}(t)) \in \mathbb{R}^p$. Thus a MFD is made up of p UFD which are potentially correlated.

*The author can also be contacted at: clement.lejeune@airbus.com

© 2020 Copyright held by the owner/author(s). Published in Proceedings of the 23rd International Conference on Extending Database Technology (EDBT), March 30-April 2, 2020, ISBN 978-3-89318-083-7 on OpenProceedings.org. Distribution of this paper is permitted under the terms of the Creative Commons license CC-by-nc-nd 4.0.

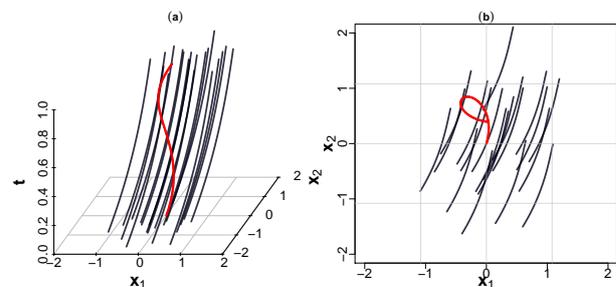


Figure 1: Example of 21 MFD ($p = 2$) with one shape persistent outlier in red. (a) (t, x_{i1}, x_{i2}) representation. (b) (x_{i1}, x_{i2}) representation i.e. projection along t -axis.

Detecting atypical behaviors is referred as *outlier* or *anomaly detection*. An outlier is defined as a sample which is rare and very different from the rest of the data set based on some measure [1]. A taxonomy of functional outliers into two classes has been proposed by Hubert *et al.* [8]. First, an *isolated outlier* is defined as a sample which exhibits an extreme behavior for very few points t . For instance a narrow vertical peak in the curve depicted by a parameter x_{ik} w.r.t t is named a *magnitude isolated outlyingness* and a high horizontal translation in the curve is referred as *shift isolated outlyingness*. Second, a *persistent outlier* is a sample which never exhibits extreme behavior but deviates from inliers for many points t , an example of shape persistent outliers is given in Fig.1. Persistent outliers can be divided into other sub-classes, see [8] for detailed examples. Note that an outlier can be of mixed type, i.e. a sample entailing several outlier classes. For an instance, one parameter has a shape persistent outlyingness and another one has an isolated outlyingness.

In this paper we focus on a geometric representation for outlier detection in MFD and highlight the situation of outliers of mixed type. The MFD case is more challenging than the UFD one since the potential correlation between the p parameters (i.e. how x_{ik} and $x_{ik'}$ are correlated w.r.t t) has to be taken into account additionally to the individual variations of the single parameters w.r.t t [8]. Indeed contrary to outliers in UFD, where the outlyingness of a sample only consists of an atypical variation w.r.t t of a single parameter, in MFD the outlyingness of sample might be hidden in an atypical variation of the *relationship* between some parameters [8, 11] as well as an atypical variation of one of the p parameters. Note that the representation of MFD we propose can also be used for other tasks than outlier detection (e.g. classification) as well as other geometric representations of 2D and 3D shapes which can be applied for the $p = 2$ and $p = 3$ (respectively) MFD cases [16].

1.2 Related work

The outlier detection in MFD is recent and has been addressed by statistical depth functions [18] originally proposed to provide

an outward-center ranking score, also named a *depth score* (e.g. in the interval $[0, 1]$), of multivariate data which are basically sample points in \mathbb{R}^d . In this general multivariate data context, where each sample is regarded as a point in a d -dimensional point cloud, the first ranked samples are the most central ones within such point cloud and are seen as most representative, whereas the last ranked samples are the least central ones and thus they are likely to be outliers. Such ranking is ensured through the monotonicity property of the depth function (see [18] for theoretical understanding of a depth function). Hence, the depth score can be viewed as an outlyingness score which reflects the degree of outlyingness at the sample level.

Some statistical depth functions have been extended first to handle UFD [3] and then MFD [2, 8]. The UFD extension consists in computing a depth function on $x_i(t)$, $\forall i$ at each t and then to compute the integral over $t \in \mathcal{T}$ of the resulting depth scores [3, 6] which in turn provides an average sample depth score for all i . Note that this extension is an aggregation of the depth function applied in a univariate manner since, for a given t , $\{x_i(t)\}_{i \leq n}$ is a point cloud in \mathbb{R} . Since $\{X_i(t)\}_{i \leq n}$ forms a point cloud in \mathbb{R}^p , the MFD extension relies on the application of a depth function in a multivariate manner and integrates the depth scores as in the UFD case [2]. Such an extension suffers from important issues :

- (1) First, it is not sensitive enough to persistent outliers because their point-wise depth scores (i.e. for each t) do not differ from those of inliers. One can augment the MFD samples by adding some derivatives functions of the parameters as supplementary (unobserved) parameters but it increases both computations and the complexity of the data analysis.
- (2) Second, even if the point-wise depth scores of an isolated outlier are different from those of an inlier, its sample depth score will be mixed with inliers because the integration of the point-wise depth scores acts as an average.
- (3) Furthermore, since the capacity of the depth function to capture different types of outlier is fundamental, outliers caused by abnormal correlation between the parameters (i.e. outliers of mixed type) are hard to detect. Such an abnormal correlation can result in outliers of mixed type.

To address the first issue (1), and especially to detect shape persistent outliers, several depth functions have been proposed. Khunt and Rehage (*FUNTA*) [9] proposed a depth function based on the intersection angles between a curve sample depicted by x_{ik} and $\{x_{jk}\}_{j \leq n, j \neq i}$ and then average these angles over both their number and the parameters. Such method is not able to detect outliers caused by an abnormal correlation between the parameters and also isolated outliers because their depth function is only focused on shape persistent outliers.

To address the second issue (2) the integral can be replaced by the infimum as the aggregation of the point-wise depth scores, which avoids the masking of outliers having few different point-wise depth scores.

To address the third issue (3), Dai and Genton [4] proposed the *directional outlyingness* (*Dir.out*), a point-wise depth function based on the direction of $X_i(t)$ in \mathbb{R}^p toward the projection-depth [17] of $\{X_i(t)\}_{i \leq n}$. To compute the sample depth score, the point-wise depth scores are aggregated through an integral over \mathcal{T} which is further decomposed into an average component and a variance-like component. Such sample depth score decomposition enables to detect multiple outliers and also to identify their class by analyzing how the two depth components are distributed

according the other sample depth scores (e.g. samples with high variance-like component value are likely persistent shape outliers and samples with high average component value are likely isolated outliers). However, to detect persistent shape outliers, the direction of $X_i(t)$ is not a sufficient feature and further geometrical representation has to be considered.

1.3 Contribution

In this paper, we propose a different framework than the statistical depth to remedy these issues by treating MFD as trajectories in \mathbb{R}^p from which we extract geometrical features such as the curvature. Such geometrical features are computed by interpretable (from a geometric standpoint) aggregation functions, named *mapping functions* in the sequel, which combine some derived functions (e.g. derivatives, integral) from the MFD. We then apply a state-of-the-art algorithm on the geometrical MFD representation to achieve the outlier detection. Considering MFD in a geometric manner enables to implicitly capture the correlation between the p parameters w.r.t t and thus to detect different classes of outliers as well as mixed types. Moreover, such combination results in a more robust outlier detection method e.g. when there are more than 5% of outliers in the training set. Thus, we both take benefit from outlier detection algorithm for multivariate data as well as the geometry of the curve (i.e. the geometry of X_i in \mathbb{R}^p and the geometry of each parameter x_{ik} w.r.t t).

2 FUNCTIONAL DATA REPRESENTATION

The first step in functional data analysis is to approximate the unknown function, $X_i : \mathcal{T} \rightarrow \mathbb{R}^p$, underlying the noisy measurement samples $X_i(t_1), \dots, X_i(t_{m_i})$ where m_i is the number of measurements for each parameter of the sample i , by an approximation function \tilde{X}_i defined as X_i . Note that no assumption is made on the distribution of the measurement points $\{t_1 \dots t_{m_i}\} = t_{i\bullet}$, thus the functional data representation can deal with sparse measurements as well as uniform ones.

The functional approximation step aims at removing the noise and thus enables to achieve accurate evaluations of some derived functions that we need for the mapping function computation. This section introduces how $\tilde{X}_i = (\tilde{x}_{i1}, \dots, \tilde{x}_{ik}, \dots, \tilde{x}_{ip})$ is specified as well as it is inferred from the data.

2.1 Functions as a basis expansion

First, we specify the functional form of the approximation function as a finite linear combination of basis functions, where each basis function depends on $t \in \mathcal{T}$. Suppose we want to approximate x_{ik} . Intuitively, it aims to represent \tilde{x}_{ik} with a small number of "specific functions", each one being able to capture some local features of X_i in hopes to recover it with a small approximation error. Hence, the following form is given for \tilde{x}_{ik} [13],

$$\forall t \in \mathcal{T}, \tilde{x}_{ik}(t) = \sum_{l=1}^{L_{ik}} \alpha_{ikl} \phi_l(t) = \alpha_{ik}^\top \phi(t) \quad (1)$$

where $\phi(t) = \{\phi_l(t)\}_{1 \leq l \leq L_{ik}}$ is a vector of orthonormal basis functions at t for some $L_{ik} \in \mathbb{N}^*$ (referred as the basis size) with fewer basis functions than sampled observation points ($L_{ik} \ll m_i$), and $\alpha_{ik}^\top = \{\alpha_{ikl}\}_{1 \leq l \leq L_{ik}}$ is the coefficient vector which element α_{ikl} is the importance of the l -th basis function. The choice of the basis of functions is data dependent.

Here we consider that x_{ik} are smooth and so we choose the B-spline basis of functions which are basically piece-wise polynomial functions. If the data were periodic data, one could choose the Fourier basis. We refer to [13] for a discussion on the choice of basis functions. Note that from the functional approximation Eq.1, one can easily compute some derivatives or integral based functional data since by linearity,

$$D^q \tilde{x}_{ik} = \sum_{l=1}^{L_{ik}} \alpha_{ikl} D^q \phi_l(t) \quad (2)$$

where $D^q = \frac{d^q}{dt^q}$ is the q -th derivative operator, provided that the basis functions ϕ_l are differentiable at the q -th order.

2.2 Inference

Assuming the data were sampled with a white noise ϵ_{ij} , i.e. $x_{ik}(t_{ij}) = \tilde{x}_{ik}(t_{ij}) + \epsilon_{ij}$ where ϵ_{ij} is independent from $\tilde{x}_{ik}(t_{ij})$, we can compute the coefficient vector α_{ik} by minimizing the following penalized least-squares criteria:

$$J_{\lambda_k}(\alpha_{ik}) = \|x_{ik}(t_{i\bullet}) - \Phi_{ik} \alpha_{ik}\|^2 + \lambda_k \alpha_{ik}^T \mathbf{R}_{ik} \alpha_{ik} \quad (3)$$

where $\|\cdot\|$ stands for the l_2 -norm, $\Phi_{ik} = \{\phi_l(t_{ij})\}_{1 \leq j \leq m_i, 1 \leq l \leq L_{ik}}$ is the $m_i \times L_{ik}$ matrix containing all the L_{ik} basis functions evaluated at the measurement points $t_{i\bullet}$ and $\mathbf{R}_{ik} = \{\int_{\mathcal{T}} D^q \phi_j(t) D^q \phi_m(t) dt\}_{1 \leq j, m \leq L_{ik}}$ is a $L_{ik} \times L_{ik}$ positive semi-definite matrix containing the inner products of the q -th derivative of the basis functions which enforces the approximation function to have a small q -th derivative i.e. to vary smoothly; $\lambda_k > 0$ is a hyper-parameter controlling the weight of the penalty and can be set to 0 for no penalization. In practice it is common to choose $q = 1$ or $q = 2$ (i.e. to penalize the velocity or the acceleration) and to compute λ_k by cross-validation.

Equating the gradient of J_{λ_k} to 0 w.r.t α_{ik} leads to the minimizer in Eq.4 [13] which is a special case of the ridge regression solution:

$$\alpha_{ik,\lambda}^* = (\Phi_{ik}^T \Phi_{ik} + \lambda_k \mathbf{R}_{ik})^{-1} \Phi_{ik}^T x_{ik}(t_{i\bullet}) \quad (4)$$

The estimated coefficient vector $\alpha_{ik,\lambda}^*$ can then be plugged in Eq.1 to evaluate \tilde{x}_{ik} over an arbitrary discretization of \mathcal{T} .

3 MAPPING FUNCTION

We propose to regard MFD as paths in \mathbb{R}^p to highlight some underlying shape outlyingness features corresponding to a change in the relationship between the parameters. We feature this change with a mapping function that we define as a geometric aggregation of the p parameters. We refer to [15] for an introduction to shape analysis from functional data.

In this section, we present the *curvature* as an example of mapping function. The curvature is a measure of how much bended a curve is, more formally how the curve locally deviates from the tangent line, see Fig.2. It is defined as:

$$\kappa(t) = \frac{\|D^1(\frac{D^1 X(t)}{\|D^1 X(t)\|})\|}{\|D^1 X(t)\|} \quad (5)$$

where $\|\cdot\|$ denotes the Euclidean norm in \mathbb{R}^p . One can interpret κ in Eq.(5) as follows: $\frac{D^1 X(t)}{\|D^1 X(t)\|}$ gives the direction vector (i.e. the normalized tangent vector), therefore $D^1 \frac{D^1 X(t)}{\|D^1 X(t)\|}$ gives the change direction vector and the denominator $\|D^1 X(t)\|$ aims to relate the change of direction w.r.t the tangent vector, i.e. how the direction vector varies w.r.t a tangent line. Consequently, the

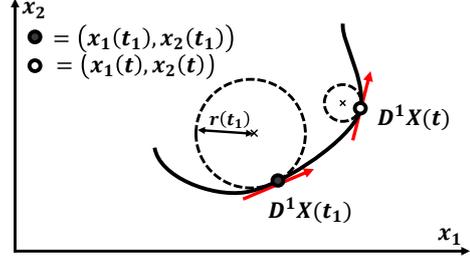


Figure 2: Curvature mapping κ . The curvature measures how large the radius of the tangent circle is. Here, in the neighbourhood of the curve at t_1 (dark-grey dot), the tangent vector $D^1 X(t_1)$ keeps the same direction, hence the tangent circle has a large radius ($r(t_1) = \frac{1}{\kappa(t_1)}$) resulting in a small curvature. In the neighbourhood of the curve at t (white dot), the tangent vector $D^1 X(t)$ quickly changes direction, hence the tangent circle has a lower radius i.e. a higher curvature than at t_1 .

curvature mapping can highlight functional outliers which curve exhibits a different bended shape than the other samples.

Thus if a curve abnormally changes direction (i.e. it deviates from a tangent line) w.r.t most of the data set, then the curvature mapping can highlight outliers. As a result, if the curve X_i depicts a line (i.e. the parameters are linearly correlated w.r.t t), then the curvature is constant w.r.t t since the directions do not vary in \mathbb{R}^p . Clearly, this is a geometric characterization of MFD.

From the reconstructed samples $\{\tilde{X}_i\}_{i \leq n}$, transformed to UFD by the mapping function, we detect the outliers with state-of-the-art algorithms initially proposed to deal with multivariate data (not functional). Here we use Isolation-Forest (*iFor*) [10] and One-class SVM (*OCSVM*) [14] which are both unsupervised.

4 NUMERICAL EXPERIMENTS

We conducted an experimental study on real data. We compare our approach with state-of-the-art depth-based methods, *FUNTA* and *Dir.out* [4, 9] (Sec.1.2) which take the MFD as input.

4.1 Experimental procedure

We experiment our method on a well-known real data set of electrocardiogram (ECG) time series [7] also used in outlier detection in [4]. Such data set correspond to time series of electrical activity and can reveal abnormal heartbeat. The time series are UFD (with number of measurements $m_i = 85, \forall i$) and in order to show the applicability of our approach in the MFD case, we augment the original UFD data to MFD ($p = 2$, bivariate) by adding the square of the initial time series. We did not add some derivatives-based functions as supplementary parameters since it is already considered by our mapping function (see Eq.5).

We evaluate our approach through multiple random splittings. We randomly split the data into a training and a test set. We generate the training set by setting the ratio of outliers (referred as the contamination level c) to 5, 10, 15, 20 and 25%. For each value of c , we repeat the random splitting 50 times, we fit *iFor* and *OCSVM* on the training set and compute the average and standard deviation Area Under (AUC) the Receive Operating Curve (ROC) on the corresponding test set. We present the results in Fig.3 and discuss them in Sec.4.3.

For each sample and each variable $x_{ik} \forall i, k$ we use a B-spline basis of functions (piece-wise polynomial functions, [13]) to achieve the functional approximations and we select the basis sizes L_{ik}

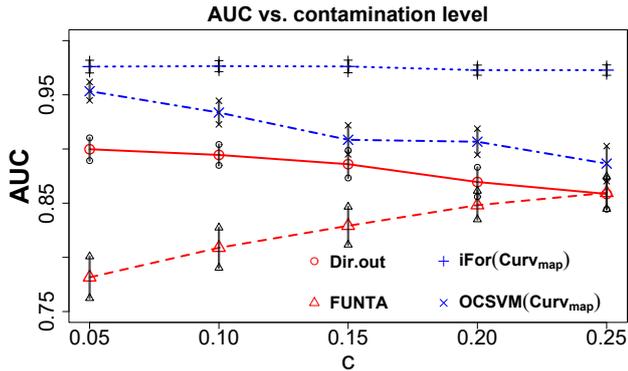


Figure 3: AUC (vertical axis) and standard deviation (vertical segments' length equals one standard deviation) from ECG data set - average results over 50 repetitions considering five contamination levels c (horizontal axis).

through a leave-one-out cross validation procedure. We evaluate each \tilde{X}_i on the same regular grid of \mathcal{T} with length $m_i = 85$. We compute the mapping function by combining the first and second derivatives, according to Eq.2 and Eq.5, and we apply *iFor* and *OCSVM* on the resulting UFD.

4.2 Outlier detection step

We use *iFor* and *OCSVM* as outlier detection algorithms on the UFD that our mapping function κ returns (Eq.5). *iFor* and *OCSVM* are unsupervised and, like the depth-based methods, output a normalized outlyingness score for each sample. In practice, in outlier detection one has not necessarily access to labeled samples, *i.e.* information depicting whether a sample is an outlier or not, but if he has, the labels can be combined with their corresponding outlyingness scores to learn an outlyingness threshold that can best discriminate outliers from inliers. Such a threshold can be learned from the ROC as well as an imbalanced classification algorithm [5, 12] in a one dimensional manner from the scores. Here, we do not learn any threshold and only consider the label information for empirical demonstration purpose, *i.e.* by computing the AUC on the test set.

4.3 Discussion of the results

From the results in Fig.3, we see that we outperform the two depth-based methods for all the contamination levels in average and perform equally in terms of standard deviation. Since *FUNTA* is only able to detect persistent shape outliers and *Dir.out* is expected to detect isolated as well as persistent outliers¹, we can deduce that the abnormal class (*i.e.* outliers) in the ECG data set not only contains persistent shape outliers but also isolated ones or outliers of mixed type which are well discriminated by the curvature mapping function. Thus the curvature mapping enables to detect mixed type outliers.

Moreover, we note that as c increases both *iFor*(*Curv_{map}*) and *OCSVM*(*Curv_{map}*) still outperform the baselines. Hence, we show that our combination of outlier detection algorithm with MFD mapped to a geometrical representation is more robust to the presence of outliers in the training set than the baselines. We note that *OCSVM* degrades as c increases. It is due to the ν hyper-parameter (we tune it on the training set with a 5-fold cross validation) corresponding to an estimate of contamination level in

¹Justification can be found in the experiments in [4] which were conducted on several synthetic data sets where each one contains a unique type of outlier.

the training set. We observed that such hyper-parameter is hard to tune as c increases and thus could decrease the performance w.r.t c .

5 CONCLUSION AND FUTURE WORK

We propose an approach to detect outliers in MFD. It consists in computing a geometrical representation of MFD followed by an outlier detection algorithm. We compare our approach with recent depth-based methods which handle MFD as input.

Through one example of mapping function, we show that the geometrical representation of MFD is well suited to detect outliers of mixed type. However, it is hard to interpret what such mixed type outliers are made up: given a detected outlier, ideally one would like to access to the amount of the different outlyingness classes *e.g.* the amount of shape persistence and shift isolated outlyingness. As future work, a mean to achieve such an interpretability is first to detect some specific outliers with depth functions, second to train outlier detection algorithms (combined with a mapping function) on training sets containing each one a unique class of outlier previously detected and then to average all the models trained to form an ensemble one. As a result, one could know which model(s) in the ensemble most contribute to the outlyingness and deduce the outlyingness composition.

ACKNOWLEDGMENTS

This work is supported by ANRT within the CIFRE framework (grant N°2017-1391).

REFERENCES

- [1] Charu C Aggarwal and Philip S Yu. 2001. Outlier detection for high dimensional data. In *ACM Sigmod Record*, Vol. 30. ACM, 37–46.
- [2] Gerda Claeskens, M Hubert, Leen Slaets, and K Vakili. 2014. Multivariate Functional Halfspace Depth. *J. Amer. Statist. Assoc.* 109, 505 (2014), 411–423.
- [3] Antonio Cuevas and Manuel Febrero. 2007. Robust estimation and classification for functional data via projection-based depth notions. *Computational Statistics* 22, 3 (2007), 481–496.
- [4] Wenlin Dai and Marc G. Genton. 2019. Directional outlyingness for multivariate functional data. *Computational Statistics and Data Analysis* 131 (2019).
- [5] Shounak Datta and Swagatam Das. 2015. Near-Bayesian support vector machines for imbalanced data classification with equal or unequal misclassification costs. *Neural Networks* 70 (2015), 39–52.
- [6] Ricardo Fraiman and Graciela Muniz. 2001. Trimmed means for functional data. *Test* 10, 2 (2001), 419–440.
- [7] Ary L Goldberger, Luis AN Amaral, Leon Glass, Jeffrey M Hausdorff, Plamen Ch Ivanov, Roger G Mark, Joseph E Mietus, George B Moody, Chung-Kang Peng, and H Eugene Stanley. 2000. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* 101, 23 (2000), e215–e220.
- [8] Mia Hubert, Peter J. Rousseeuw, and Pieter Segaeert. 2015. Multivariate functional outlier detection. *Statistical Methods and Applications* 24, 2 (2015).
- [9] Sonja Kuhnt and André Rehage. 2016. An angle-based multivariate functional pseudo-depth for shape outlier detection. *Journal of Multivariate Analysis*. 146 (2016), 325–340.
- [10] Fei Tony Liu, Kai Ming Ting, and Zhi-Hua Zhou. 2008. Isolation Forest. In *ICDM*.
- [11] Sara López-pintado, Yin Sun, Juan K Lin, and Marc G. Genton. 2014. Simplicial band depth for multivariate functional data. *Advances in Data Analysis and Classification* 8, 3 (2014), 321–338.
- [12] Art B Owen. 2007. Infinitely imbalanced logistic regression. *Journal of Machine Learning Research*. 8 (2007), 761–773.
- [13] James O. Ramsay and B.W. Silverman. 2006. *Functional Data Analysis*. Springer.
- [14] Bernhard Schölkopf, John C Platt, John Shawe-Taylor, Alex J Smola, and Robert C Williamson. 2001. Estimating the support of a high-dimensional distribution. *Neural computation* 13, 7 (2001), 1443–1471.
- [15] Anuj Srivastava and Eric P Klassen. 2016. *Functional and Shape Data Analysis*. Springer.
- [16] Weiye Xie, Oksana Chkrebti, and Sebastian Kurtek. 2019. Visualization and Outlier Detection for Multivariate Elastic Curve Data. *IEEE Transactions on Visualization and Computer Graphics* (2019).
- [17] Yijun Zuo et al. 2003. Projection-based depth functions and associated medians. *The Annals of Statistics* 31, 5 (2003), 1460–1490.
- [18] Yijun Zuo and Robert Serfling. 2000. General Notions of Statistical Depth Function. *The Annals of Statistics*. 28, 2 (2000), 461–482.