MOX–Report No. 37/2011

Depth Measures For Multivariate Functional Data

Ieva, F.; Paganoni, A.M.

MOX, Dipartimento di Matematica "F. Brioschi"
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

mox@mate.polimi.it http://mox.polimi.it

Depth Measures For Multivariate Functional Data

Francesca Ieva[♯] and Anna Maria Paganoni[♯]

October 17, 2011

[♯] MOX– Modellistica e Calcolo Scientifico
Dipartimento di Matematica "F. Brioschi"
Politecnico di Milano
via Bonardi 9, 20133 Milano, Italy

francesca.ieva@mail.polimi.it, anna.paganoni@polimi.it

Keywords: Depth measures; multivariate functional data; rank tests; ECG signal.

Abstract

The statistical analysis of functional data is a growing interest research area. In particular more and more frequently in the biomedical context the output of many clinical examinations are complex mathematical objects like images or curves. In this work we propose, analyze, and apply a new concept of depth for multivariate functional observations, i.e. statistical units where each component is a curve, in order to study them from a statistical perspective. Robust statistics, such as the median function or trimmed mean, can be generalized to a multivariate functional framework using this new depth measure definition so that outliers detection and nonparametric tests can be carried out also within this more complex context. Mathematical properties of these new concepts are established and proved. Finally, an application to Electrocardiographic (ECG) signals is proposed, aimed at detecting outliers for identifying stable training set to be used in unsupervised classification procedures adopted to perform semi automatic diagnosis and at testing differences between pathological and physiological groups of patients.

1 Introduction and notation

Nowadays, an ever-increasing number of clinical and epidemiological studies leads to the necessity of dealing with functional data, since more and more often the analysis of data coming from biomedical fields meets problems where ideal units of observation are curves and a function is observed for each individual. This calls for the identification of suitable models and inferential techniques

for managing with complexity of such data. For example, a challenging task in functional data analysis is to provide an ordering within a sample of curves that allows the definition of order statistics, such as ranks and L-statistics (see Fraiman and Meloche, 1999).
A natural tool to analyze these functional data features is the idea of statistical depth, which provides a measure of centrality or outlyingness of an observation with respect to a given dataset or a population distribution. Several definition of depth measures for multivariate data have been proposed and analyzed in literature (see Mahlanobis, 1936; Tukey, 1975; Liu, 1990; Zuo and Serfling, 2000; Zuo, 2003 among others). A generalization to functional data is given in López-Pintado and Romo (2009), starting from depth measures for multivariate data. They also provide the extension of robust statistics to a functional framework, generalizing properties of depth measures which are proved to hold in multivariate case (see Liu, 1990; Zuo and Serfling, 2000; Serfling, 2004 for further details on multivariate setting). Finally a specific focus on trimmed means for functional data can be found in Fraiman and Muniz (2001), where a generalization of some issues treated in Fraiman and Meloche (1999) about multivariate L-estimation is proposed.
Once a depth measure is associated with each data within a sample, it is possible to rank them as well as to visualize graphically the result of ranking through Functional Boxplots, as proposed in the work of Sun and Genton (2011) and Ieva (2011).

There are lots of different aims which lead to rank curves according to suitable depth indexes. Indeed, several applications focus on classification of functions arising from different population and make inference about the latent differences among them analyzing the morphological effects they induce on the curves shape. This is usually carried out without parametric assumptions on the model which the sample of curves is associated with, like in Cuevas et al. (2007) and López-Pintado and Romo (2003). On the other hand, sometimes the interest is in making inference on specific summary statistics, as proposed in Li and Liu (2004) for the multivariate setting.
In this work we deal with multivariate functional observations, i.e. statistical units where each component is a curve. Then we firstly need to generalize the concept of depth for functional data to the multivariate functional case, then to define suitable generalizations of nonparametric statistics for ranking and classifying multivariate curves as well as making inference on them. So we propose, analyze, and apply a new concept of a multivariate index of depth, derived from averaging univariate centrality measures for functional data in a suitable multivariate index. We also widen the employment of the functional boxplots, adopting this graphical tool also in the more complex case of samples of multivariate functions. Then the Wilcoxon rank test based on the order induced by the multivariate functional depth is proposed to test differences between groups of multivariate curves.
In fact, two are the main goals of the analysis: the first one is to point out a

suitable method for performing outliers detection in a multivariate functional setting, within a sample of curves arising from the same population; the second one is to carry out non parametric test for comparing samples of multivariate curves and making inference on the corresponding populations.
A natural application of this theoretical framework comes from the biomedical context, and in particular from applications that deal with cardiovascular diseases diagnoses carried out using Electrocardiographic devices. In fact, ECG signals can be considered multivariate functional data with dependent components. In this context, some issues of interest are, for example, classification of groups of curves with similar morphological patterns, multivariate functional outliers detection within a homogeneous group and classical inference on mean and quantiles of subpopulations. From a clinical point of view, the first issue concerns how to carry out a semi automatic diagnosis based only on the morphological deviations from physiological patterns induced by the presence of the disease of interest; the second one leads to profile "typical" curve expression for each pathology; finally the third one allows for the investigation of the presence of statistically significant differences in the subpopulations of pathological units with respect to physiological ones.
The article is organized as follows. In Section 2 the definition of multivariate functional depth is presented and mathematical properties of this depth measure are proved in the more general framework of multivariate functional data. In Section 3 an application to ECG signals of patients affected and not affected by Bundle Brunch Block Infarction is presented. Finally, Section 4 contains conclusions, discussion of results and further developments. The proofs are included in the Appendix.

2 Band depth and inference for Multivariate Functional Data

In this Section, the new concept of multivariate functional depth measure is presented and natural properties are established and proved. Moreover, the modified version of the band depth is given, since it is used to implement the modified version of the functional boxplot. Finally, Wilcoxon non parametric rank test framework is adopted to make inference on samples of multivariate functional data, once a suitable ranking has been obtained.

2.1 Band depth for Multivariate Functional Data

As mentioned in the previous Section, a natural tool to analyze and rank functional data is the idea of statistical depth, which measures the centrality of a given curve within a group of trajectories providing center-outward orderings of the set of curves itself. In general, several different definitions of depth can be given (see Zuo and Serfling, 2000). In our case, we refer to the band depth measure for functional data proposed by López and Pintado (2009).

Let X a stochastic process with law P taking values on the space $\mathcal{C}(I)$ of real continuous functions on the compact interval I. The graph of a function $f \in \mathcal{C}(I)$ is the subset of the plane $G(f) = \{(t, f(t)) : t \in I\}$. The random band depth, of order $J \geq 2$, for a function $f \in \mathcal{C}(I)$ is then

$$BD^J_{P_X}(f) = \sum_{j=2}^{J} P_X\{G(f) \subset B(X_1, X_2, ..., X_j)\},$$

where $B(X_1, X_2, ..., X_j)$, for $j = 2, ..., J$ is the random band in $\mathbb{R}^2$ delimited by $X_1, ..., X_j$, independent copies of the stochastic process X, defined as

$$B(X_1, ..., X_j) = \{(t, y(t)) : t \in I, \min_{r=1,...,j} X_r(t) \leq y(t) \leq \max_{r=1,...,j} X_r(t)\}$$

In this paper we propose a new definition of a band depth measure for multivariate functional data, i.e. data generated by a stochastic process $\mathbf{X}$ taking values in the space $\mathcal{C}(I; \mathbb{R}^s)$ of continuous functions $\mathbf{f} = (f_1, ..., f_s) : I \to \mathbb{R}^s$.

Definition 1
Let $\mathbf{f}$ be a function on I taking values in $\mathbb{R}^s$. The *multivariate band depth measure* is defined as

$$BD^J_{P_{\mathbf{X}}}(\mathbf{f}) = \sum_{k=1}^{s} p_k BD^J_{P_{X_k}}(f_k), \qquad p_k > 0 \; for \; k = 1, ..., s, \qquad \sum_{k=1}^{s} p_k = 1. \quad (1)$$

Let $\mathbf{X}$ a multivariate random process such that $P(\min_{k=1,...,s} \|X_k\|_\infty > M) \to 0$ as $M \to \infty$, then it is easy to prove, using the properties of the functional depth measure summarized in López and Pintado (2009), the following results on the basic properties of the multivariate band depth measure defined in (1).

Proposition 1

(a) Let $T(\mathbf{f}) = \mathbf{A}(t)\mathbf{f}(t) + \mathbf{b}(t)$, where $\forall t \in I$ $\mathbf{A}(t)$ is a $s \times s$ diagonal matrix such that $\mathbf{A_{kk}}(t)$ are continuous functions in I, with $\mathbf{A_{kk}}(t) \neq 0$, for each $t \in I$, and $\mathbf{b}(t) \in \mathcal{C}(I; \mathbb{R}^s)$. Then $BD^J_{P_{T(\mathbf{X})}}(T(\mathbf{f})) = BD^J_{P_{\mathbf{X}}}(\mathbf{f})$.

(b) $BD^J_{P_{\mathbf{X}(g(t))}}(\mathbf{f}(g(t))) = BD^J_{P_{\mathbf{X}(t)}}(\mathbf{f}(t)) =$ when g is a one-to-one transformation of the interval I.

(c) $\sup_{\min_{k=1,...,s} \|f_k\|_\infty > M} BD^J_{P_{\mathbf{X}}}(\mathbf{f}) \to 0$ as $M \to \infty$.

(d) If $\forall k = 1, ..., s$ the probability distribution P_{X_k} on $\mathcal{C}(I)$ has absolutely continuous marginal distributions, then $BD^J_{P_{\mathbf{X}}}$ is a continuous functional on $\mathcal{C}(I; \mathbb{R}^s)$.

If $\mathbf{X}_1, \ldots, \mathbf{X}_n$ are independent copies of the stochastic process $\mathbf{X}$, the sample version of (1) can be introduced in order to conduct descriptive and inferential statistical analyses on a set of multivariate functional data $\mathbf{f}_1, ..., \mathbf{f}_n$ generated by the process $\mathbf{X}$. For any $\mathbf{f}$ in the sample $\mathbf{f}_1, ..., \mathbf{f}_n$ we can compute the depth as

$$BD_n^J(\mathbf{f}) = \sum_{k=1}^{s} p_k BD_{n,k}^J(f_k),$$

where for the function $f_k \in \mathcal{C}(I)$

$$BD_{n,k}^J(f_k) = \sum_{j=2}^{J} \binom{n}{j}^{-1} \sum_{1 \le i_1 < i_2 < \cdots < i_j \le n} \mathbb{I}\{G(f_k) \subset B(f_{i_1;k}, ... f_{i_j;k})\}$$

and $\mathbb{I}\{G(f_k) \subset B(f_{i_1;k}, ..., f_{i_j;k})\}$ indicates if the band determined by $(f_{i_1;k}, ..., f_{i_j;k})$ contains the whole graph of f. The k component of the vector $\mathbf{f_i}$ is denoted by $f_{i;k}$.

Proposition 2
The sample version of multivariate functional depth is consistent, in fact

$$|BD_n^J(\mathbf{f}) - BD_{P_\mathbf{X}}^J(\mathbf{f})| \to 0, \ a.s. \ \ if \ \ n \to \infty \tag{2}$$

As proposed in López-Pintado and Romo(2009) also in this multivariate functional setting we can move to the analogous of the modified band depth:

$$MBD_n^J(\mathbf{f}) = \sum_{k=1}^{s} p_k MBD_{n,k}^J(f_k), \tag{3}$$

where for the function $f_k \in \mathcal{C}(I)$ the modified band depth measures the proportion of time that the curve f_k is in the band, i.e.

$$MBD_{n,k}^J(f_k) = \sum_{j=2}^{J} \binom{n}{j}^{-1} \sum_{1 \le i_1 < i_2 < \cdots < i_j \le n} \tilde{\lambda}\{E(f_k; f_{i_1;k}, ..., f_{i_j;k})\},$$

where $E(f_k) =: E(f_k; f_{i_1;k}, ..., f_{i_j;k}) = \{t \in I, \min_{r=i_1,...,i_j} f_{r;k}(t) \le f_k(t) \le \max_{r=i_1,...,i_j} f_{r;k}(t)\}$ and $\tilde{\lambda}(f_k) = \lambda(E(f_k))/\lambda(I)$ and λ is the Lesbegue measure on I. As stated in López and Pintado (2009) the values of the modified band depth measure are stable with respect to the choice of J, and in order to be computationally faster we set $J = 2$ and we denote $MBD_n^J(\mathbf{f})$ as $MBD(\mathbf{f})$ in the following. The use of the modified band depth measure avoids also having too many depth ties.
Given the multivariate band depth measure defined in (3), a sample of multivariate functional data $\mathbf{f}_1, ..., \mathbf{f}_n$ can be ranked. In the following we denote $\mathbf{f}_{[i]}$ the sample curve associated with the ith largest depth value, so $\mathbf{f}_{[1]} = \text{argmax}_{\mathbf{f} \in \{\mathbf{f}_1, ..., \mathbf{f}_n\}} MBD(\mathbf{f})$ is the *median* (deepest and more central) curve, and $\mathbf{f}_{[n]} = \text{argmin}_{\mathbf{f} \in \{\mathbf{f}_1, ..., \mathbf{f}_n\}} MBD(\mathbf{f})$ the most outlying one.

2.2 Multivariate functional boxplot and outliers detection

The idea of generalizing the concept of functional boxplot to multivariate functional data is based on the new definition of multivariate functional depth measure given in (3) which takes into account simultaneously the behaviour of all the s components of $\mathbf{f}$ weighting in a suitable way the components in order to take into account correlations among them. Concerning the aim of performing outliers detection for robustifying training set adopted in unsupervised classification algorithms, the following steps should then be implemented on multivariate curves sample $\mathbf{f}_1, ..., \mathbf{f}_n$:

1. For each statistical unit j, compute the value of measure depth $MBD(\mathbf{f_j})$;

2. Rank the multivariate functions $\mathbf{f}_j(t)$ according to the value of multivariate depth measure and define outliers those curves that, for at least one t, are outside the fences obtained inflating the envelope of the $\alpha\%$ central region by h times the range of the $\alpha\%$ central region. In particular the $\alpha\%$ central region for the component f_k determined by a sample of curves is defined as

$$\mathcal{C}_\alpha = \left\{(t, y(t)) : \min_{r=1,\dots,\lceil \alpha n \rceil} f_{[r];k}(t) \leq y(t) \leq \max_{r=1,\dots,\lceil \alpha n \rceil} f_{[r];k}(t)\right\}$$

   where $\lceil \alpha n \rceil$ is the smallest integer greater than or equal to αn. In the following we set $\alpha\% = 50\%$ and $h = 1.5$.

3. Visualize the functional boxplot of each component, building the envelope of the 50% deepest functions and then the functional boxplot according to the ranking arising from the multivariate index previously pointed out.

Notice that this algorithm defines outliers according to a multivariate index of depth, which takes into account simultaneously the depth of all components of the multivariate function.

2.3 Robust statistics and rank test

Given the order in the sample of curves induced by the multivariate functional depth measure, the definition of *trimmed mean* following, for example, in Fraiman and Muniz (2001) can be extended to multivariate functional data straightforwardly. We can also widen to this framework a non parametric rank test to compare two samples of multivariate functions. In particular consider a sample $\mathbf{f}_1, ..., \mathbf{f}_n$ generated according to a distribution $P_{\mathbf{X}}$ and another sample $\mathbf{g}_1, ..., \mathbf{g}_m$ generated according to a distribution $P_{\mathbf{Y}}$. We want to test differences between the two populations; combine the two samples, that is, let $W = \mathbf{w}_1, ..., \mathbf{w}_{n+m} \equiv \mathbf{f}_1, ..., \mathbf{f}_n, \mathbf{g}_1, ..., \mathbf{g}_m$. We can assign to each element of the combined set a rank according to values of the multivariate functional depth, and in particular the higher the depth the lower the rank. The proposed test statistics R is the sum of the ranks of the second sample $\mathbf{g}_1, ..., \mathbf{g}_m$ with respect to the combined set W ($R = \sum_{j=1}^{m} \text{Rank}_W(\mathbf{g}_j) \equiv \sum_{j=1}^{m} r(\mathbf{g}_j)$). If there is no differences between the distributions generating the data (H_0), $(r(\mathbf{g}_1), ..., r(\mathbf{g}_m))$ can be viewed as a random sample size m drawn without replacement from the set $(1, ..., n + m)$, and we reject H_0 for values of R too small or too high. For large values of n and m it is possible to use a normal approximation (see Li and Liu, 2004).

Such test represents a quantitative method for carrying out inference in a supervised multivariate functional clustering framework. On the other hand, for the unsupervised clustering case, it can be also seen as a way to test if the process generating the outliers pointed out by the functional boxplot can be considered as different from the process generating the curves of the $\alpha\%$ most central region.

3 An application to ecg signals

In Ieva et al. (2011), a statistical framework for analysis and classification of ECG curves starting from their sole morphology is proposed. The main goal of the paper is to identify, from a statistical perspective, specific ECG patterns which could benefit from an early invasive approach. In fact, the identification of statistical tools capable of classifying curves using their shape only could support an early detection of heart failures, not based on usual clinical criteria. In order to do this, a real time procedure consisting of preliminary steps like reconstructing signals, wavelets denoising and removing biological variability in the signals through data registration is tuned and tested. Then, a multivariate functional k-means clustering of reconstructed and registered data is performed. Since when testing new procedures for classification the performances of classification method are to be validated through cross validation, it is mandatory a suitable training of the algorithm on data. This would lead to robustify classification algorithm and would improve reliability in prediction. The procedure proposed in the previous Section is an effective way to reach this goal. In fact, it leads to select for the training set the proportion of multivariate curves whose depth is greater. Considering the ECG of the j-th patient as a 8-variate function $\mathbf{f}_j = (f_{j;1}, ..., f_{j;8})$, the $f_{j;k}$, $(k = 1, ..., 8)$ correspond to the eight leads I, II, V1, V2, V3, V4, V5 and V6. Then the procedure discussed in Section 2 is applied in order to carry out functional boxplots and to perform outliers detection for two different groups: physiological and pathological patients, i.e. people affected by a particular kind of heart disease, called Bundle Branch Block (BBB). This is a pathology which is easy to detect through the observation of shape modifications it induces on ECG pattern and divides in Right Bundle Branch Block (RBBB) and Left Bundle Branch Block (LBBB) according to the heart side it affects. In the following, we will consider a sample of 100 physiological signals and 50 pathological ones, where the latter come from patients affected by LBBB.

In Figures 1 and 2 the row data are shown, whereas Figures 3 and 4 show the corresponding functional boxplots, one for each lead of the ECG, (see Ieva et al. (2011) for details on statistical analysis and procedures). Functional Boxplots are produced according to the ranking induced by the multivariate functional index where the weights p_k, $(k = 1, \ldots, 8)$ are all equal to 1/8, weigthing in the same way all the leads.
Since there is a common ranking of all components of $\mathbf{f}_j$s, induced by the multivariate index of depth, the central band is defined with the same curves in each component of the functional boxplot, since the multivariate funtional index of depth defined in (3) takes jointly into account the order of each component (lead) of the multivariate function (ECG). This is the main and most important difference between functional boxplots reported in Figures 3 and 4 and those we would have obtained simply asking for functional boxplots of each lead. To be noticed is that the procedure is very easy to generalize to the adoption of any definition of multivariate index of depth for which properties like those established in Proposition 1 and 2 can be proved.
As described in Section 2.3, given the order in the sample of curves induced by the multivariate functional depth measure, it is possible to widen to this framework a non parametric rank test in order to compare two samples of multivariate functions. Actually, we will adopt the rank test to check for differences in the underlying process generating the LBBB curves with respect to the physiological ones. Then the combined dataset consists of 150 8-variate functional ECG signals. The $p-$value of the test carried out on these curves using the multivariate functional index computed on them all is equal to $3.38 * 10^{-16}$. The statistical evidence is still very strong ($p-$value $= 2.96 * 10^{-16}$) if we compute the depth measure (3) setting $(p_1, ..., p_8)$ equal to $(1/10, 1/10, 2/10, 1/10, 1/10, 1/10, 1/10, 2/10)$, stressing the weight of leads V1 and V6, since they are the most important for carrying out the LBBB diagnosis, as confirmed by cardiologists.

That is, a strong evidence for the LBBB to be considered as arising from a different latent process exists. This is also detectable looking at the functional boxplots arising from the new database, shown in Figure 5: almost all the outliers are those related to LBBB signals.

4 Conclusions

In this work, we generalize the notion of depth for functional data presented in López-Pintado and Romo (2009) to the multivariate functional case and define also a new multivariate functional index of depth which is able to take into account jointly the depth of the multivariate functional data on each component. This provides a center-outward ordering criterion for a sample of multivariate functions. Extensions and proofs of the properties of the new index are also provided, as well as for its modified version. A generalization of the non parametric test to this framework has been adopted to carry out inference in a supervised clustering context. Finally, the application of the new index to a real case of ECG signals has been proposed and discussed, highlighting how the methodology works effectively both in detecting outliers and in distinguishing between samples arising from different underlying processes.

Appendix

Proof of Proposition 1: (a) using Definition 1 and the property (1) of Theorem 3 in López and Pintado (2009) we have

$$BD^J_{P_{T(\mathbf{x})}}(T(\mathbf{f})) = \sum_{k=1}^{s} p_k BD^J_{P_{A_{kk}X_k+b_k}}(A_{kk}f_k+b_k) = \sum_{k=1}^{s} p_k BD_{X_k}(f_k) = BD^J_{P_{\mathbf{X}}}(\mathbf{f}).$$

The diagonality requirement on matrix A means that the multivariate functional depth measure $BD^J_{P_{\mathbf{X}}}(\mathbf{f})$ is invariant as regards affine transformations of each component taken one by one, without combining different elements of the multivariate function.
(b) follows directly from property (2) of Theorem 3 in López and Pintado (2009).
(c)

$$\sup_{\min_{k=1,\dots,s}\|f_k\|_\infty > M} BD^J_{P_{\mathbf{X}}}(\mathbf{f}) = \sup_{\min_{k=1,\dots,s}\|f_k\|_\infty > M} \sum_{k=1}^{s} p_k BD_{X_k}(f_k)$$

and each term in the sum over components goes to zero when M goes to infinity.
(d) also this point follows directly from property (4) of Theorem 3 in López and Pintado (2009).
Proof of Proposition 2:

$$|BD^J_n(\mathbf{f}) - BD^J_{P_{\mathbf{X}}}(\mathbf{f})| = |\sum_{k=1}^{s} p_k BD^J_{n,k}(f_k) - \sum_{k=1}^{s} p_k BD_{X_k}(f_k)|$$

$$\leq \sum_{k=1}^{s} p_k |BD^J_{n,k}(f_k) - BD_{X_k}(f_k)| \tag{4}$$

and each term of the sum in the last term of (4) goes to zero as stated in Theorem 4 of López and Pintado (2009).

Aknowledgements

This work is within PROMETEO (PROgetto sullarea Milanese Elettrocardiogrammi Teletrasferiti dallExtra Ospedaliero). The authors wish to thank Mortara Rangoni Europe s.r.l. for having provided data.

References

[1] Cuevas, A., Febrero, M., Fraiman, R. (2007), Robust Estimation and Classification for Functional Data via Projection-based Depth Notions, *Computational Statistics*, **22**, 481–496.

[2] Fraiman, R. and Meloche, J. (1999), Multivariate L-estimation, *Test*, **8**, 255–317.

[3] Fraiman, R. and Muniz G. (2001), Trimmed means for functional data. *Test*, **10**, 419–440.

[4] Ieva, F. (2011), Outlier detection for training sets in an unsupervised functional classification framework: an application to ECG signals, *Proceedings of the 17th EYSM, Lisbon, Portugal.*

[5] Ieva, F., Paganoni, A.M., Pigoli, D., Vitelli, V. (2011), Multivariate functional clustering for the analysis of ECG curves morphology, *Tech. Rep. MOX, Math. Dept.*, Politecnico di Milano.

[6] Li, J., and Liu, R. (2004), New Nonparametric Tests of Multivariate Locations and Scales using Data Depth, *Statistical Science*, **19**, 686-696.

[7] Liu, R. (1990), On a Notion of Data Depth Based on Random simplices, *The Annals of Statistics*, **18**, 405414.

[8] Lopez-Pintado, S., Romo, J. (2003), Depth-based classification for functional data, *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*

[9] Lopez-Pintado, S., Romo, J. (2009), On the Concept of Depth for Functional Data, *Journal of the American Statistical Association*, **104**, 486, 718–734.

[10] Mahalanobis, P. C. (1936), On the generalized distance in statistics, *Proceedings of National Academy of Science of India*, **12**, 49-55.

[11] Serfling, R. (2004), Depth Functions in Nonparametric Multivariate Inference, *DIMACS Series in Discrete Mathematics and Theoretical Computer Science.*

[12] Sun, Y., Genton, M.-G. (2011), Functional Boxplots, *Journal of Computation and Graphical Statistics*, to appear.

[13] Tukey, J. (1975), Mathematics and Picturing Data, *Proceedings of the 1975 International Congress of Mathematic* **2**, 523531.

[14] Zuo, Y.-Z., Serfling, R. (2000), General Notions of Statistical Depth Function, *The Annals of Statistics*, **28**, 2, 461–482.

[15] Zuo, Y. (2003), Projection Based Depth Functions and Associated Medians, *The Annals of Statistics*, **31**, 14601490.

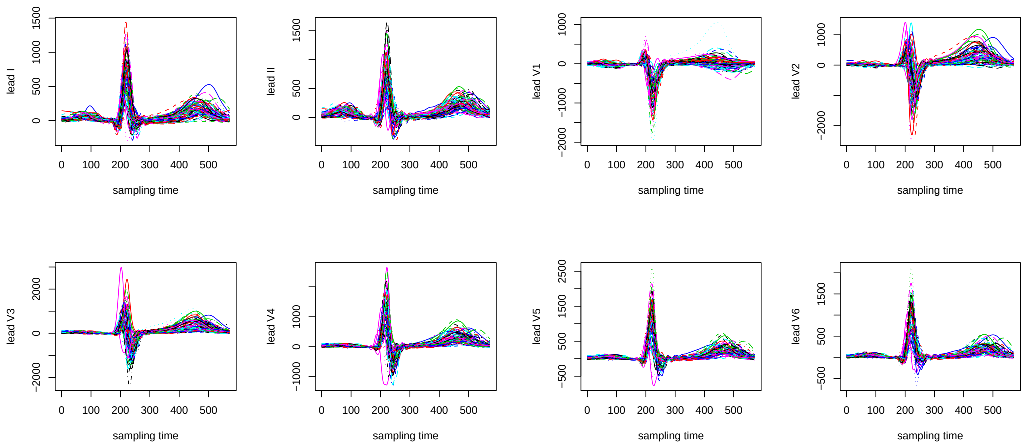

Figure 1: Row signals of the 100 physiological patients.

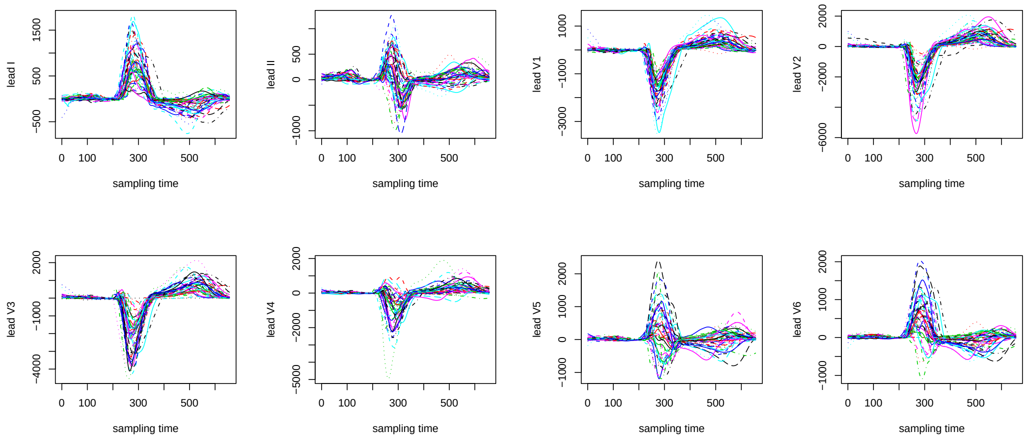

Figure 2: Row signals of the 50 pathological patients.

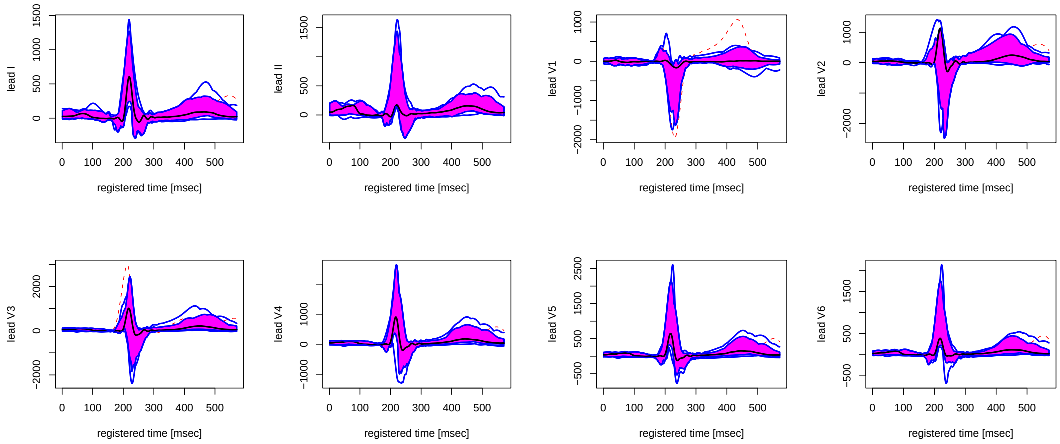

Figure 3: Functional boxplots of each component (*lead*) of the 100 physiological ECGs. The central bands (purple coloured area) and outliers (red dotted lines) of each lead are defined as described in Section 2.2, according to the ranking induced by $MBD_n^J(\mathbf{f})$ defined in (3).

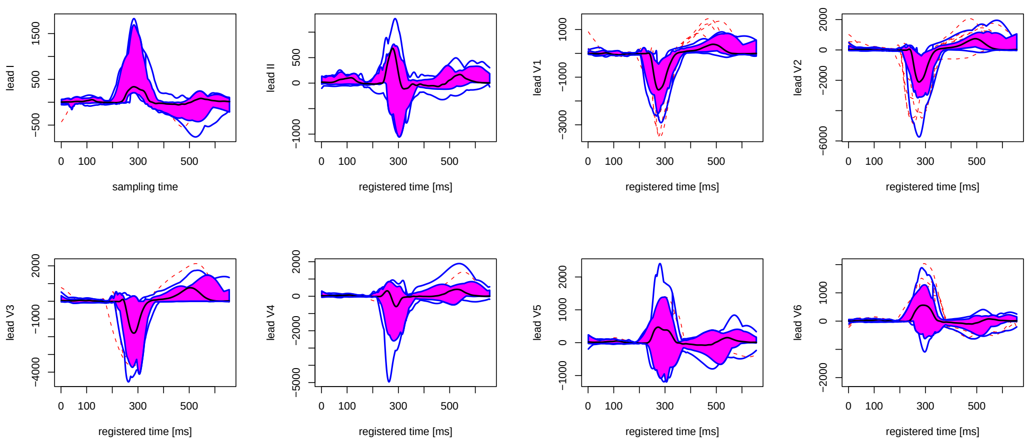

Figure 4: Functional boxplots of each component (*lead*) of the 50 pathological (Left Bundle Brunch Block) ECGs. The central bands (purple coloured area) and outliers (red dotted lines) of each lead are defined as described in Section 2.2, according to the ranking induced by $MBD_n^J(\mathbf{f})$ defined in (3).

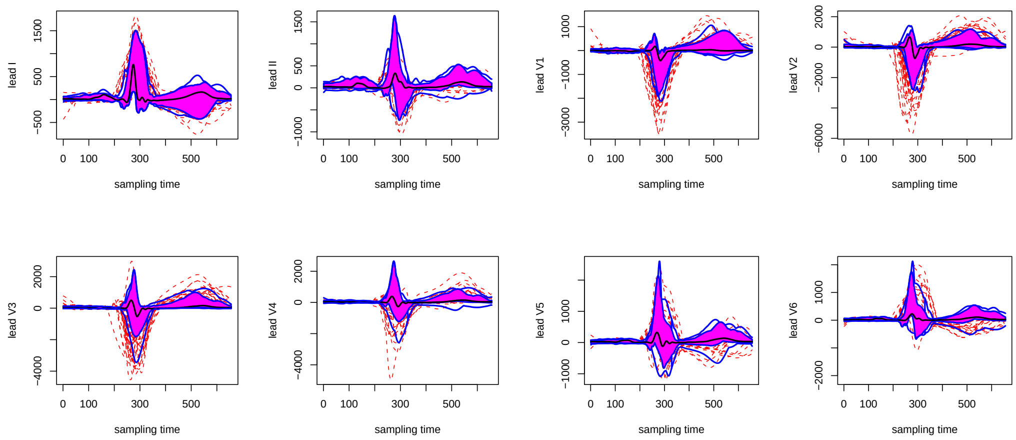

Figure 5: Functional boxplots of each component (*lead*) of the 150 physiological (100) and pathological (50 Left Bundle Brunch Block) ECGs. The central bands (purple coloured area) and outliers (red dotted lines) of each lead are defined as described in Section 2.2, according to the ranking induced by $MBD_n^J(\mathbf{f})$ defined in (3).

MOX Technical Reports, last issues

Dipartimento di Matematica "F. Brioschi",
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

37/2011 IEVA, F.; PAGANONI, A.M.
Depth Measures For Multivariate Functional Data

36/2011 MOTAMED, M.; NOBILE, F.; TEMPONE, R.
A stochastic collocation method for the second order wave equation with a discontinuous random speed

35/2011 IAPICHINO, L.; QUARTERONI, A.; ROZZA, G.
A Reduced Basis Hybrid Method for the coupling of parametrized domains represented by fluidic networks

34/2011 BENACCHIO, T.; BONAVENTURA, L.
A spectral collocation method for the one dimensional shallow water equations on semi-infinite domains

33/2011 ANTONIETTI, P.F.; BEIRAO DA VEIGA, L.; LOVADINA, C.; VERANI, M.
Hierarchical a posteriori error estimators for the mimetic discretization of elliptic problems

32/2011 ALETTI, G; GHIGLIETTI, A; PAGANONI, A.
A modified randomly reinforced urn design

31/2011 ASTORINO, M.; BECERRA SAGREDO, J.; QUARTERONI, A.
A modular lattice Boltzmann solver for GPU computing processors

30/2011 NOBILE, F.; POZZOLI, M.; VERGARA, C.
Time accurate partitioned algorithms for the solution of fluid-structure interaction problems in haemodynamics

29/2011 MORIN, P.; NOCHETTO, R.H.; PAULETTI, S.; VERANI, M.
AFEM for Shape Optimization

28/2011 PISCHIUTTA, M.; FORMAGGIA, L.; NOBILE, F.
Mathematical modelling for the evolution of aeolian dunes formed by a mixture of sands: entrainment-deposition formulation