Physiological Measurement

IPEM
Institute of Physics and Engineering in Medicine

PAPER • **OPEN ACCESS**

A unified framework for multi-lead ECG characterization using Laplacian Eigenmaps

To cite this article: Amalia Villa *et al* 2023 *Physiol. Meas.* **44** 075006

View the article online for updates and enhancements.

You may also like

- Weak fault diagnosis of machinery using Laplacian eigenmaps and parallel sparse filtering
Shanshan Ji, Jinrui Wang, Baokun Han et al.
- Pairwise constraints cross entropy fuzzy clustering algorithm based on manifold learning and feature selection
Wei Cai, Shengbing Xu, LiangJun Zhang et al.
- The influence of environmental factors on the intensity of typhoon Mirinae in the South China Sea and its forecast analysis
Ying Huang, Xiaoyan Huang, Huasheng Zhao et al.

This content was downloaded from IP address 81.164.111.158 on 23/08/2023 at 14:19

IOP Publishing Physiol. Meas. 44 (2023) 075006 https://doi.org/10.1088/1361-6579/acdfb4

Physiological Measurement

IPEM Institute of Physics and Engineering in Medicine

CrossMark

OPEN ACCESS

RECEIVED
24 November 2022

REVISED
21 May 2023

ACCEPTED FOR PUBLICATION
19 June 2023

PUBLISHED
24 July 2023

Original content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence.

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

PAPER

A unified framework for multi-lead ECG characterization using Laplacian Eigenmaps

Amalia Villa[1], Sebastian Ingelaere[2], Ben Jacobs[3], Bert Vandenberk[2,4], Sabine Van Huffel[1], Rik Willems[2] and Carolina Varon[1,5]

[1] Department of Electrical Engineering (ESAT), STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, KU Leuven, Leuven, Belgium
[2] Department of Cardiovascular Diseases, Experimental Cardiology, KU Leuven, Leuven, Belgium
[3] Quality by Design NV (QbD), Wilrijk, Belgium
[4] Department of Cardiac Sciences, Libin Cardiovascular Institute, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada
[5] Microgravity Research Center, Universite Libre de Bruxelles, Brussels, Belgium

E-mail: villago.a@yahoo.com

Keywords: Laplacian Eigenmaps, ECG signal processing, dimensionality reduction, manifold learning

Abstract

Background. The analysis of multi-lead electrocardiographic (ECG) signals requires integrating the information derived from each lead to reach clinically relevant conclusions. This analysis could benefit from data-driven methods compacting the information in those leads into lower-dimensional representations (i.e. 2 or 3 dimensions instead of 12). *Objective.* We propose Laplacian Eigenmaps (LE) to create a unified framework where ECGs from different subjects can be compared and their abnormalities are enhanced. *Approach.* We conceive a normal reference ECG space based on LE, calculated using signals of healthy subjects in sinus rhythm. Signals from new subjects can be mapped onto this reference space creating a loop per heartbeat that captures ECG abnormalities. A set of parameters, based on distance metrics and on the shape of loops, are proposed to quantify the differences between subjects. *Main results.* This methodology was applied to find structural and arrhythmogenic changes in the ECG. The LE framework consistently captured the characteristics of healthy ECGs, confirming that normal signals behaved similarly in the LE space. Significant differences between normal signals, and those from patients with ischemic heart disease or dilated cardiomyopathy were detected. In contrast, LE biomarkers did not identify differences between patients with cardiomyopathy and a history of ventricular arrhythmia and their matched controls. *Significance.* This LE unified framework offers a new representation of multi-lead signals, reducing dimensionality while enhancing imperceptible abnormalities and enabling the comparison of signals of different subjects.

1. Introduction

The electrocardiogram (ECG) is a graphical representation of the electrical activity of the heart and an important diagnostic tool in cardiology. The standard ECG consists of 12 leads containing information of different regions in the heart. The analysis of these multi-dimensional (i.e. multi-lead) signals can be challenging, since the integration of the information from the different leads is not straightforward. It is done by visually exploring the ECG taking into account the spatial information associated to the different leads (Garcia 2013). Additionally, this multi-lead analysis can also be performed by deriving the vectorcardiogram (VCG) from 12-lead ECG. The VCG provides a graphical representation of the direction and magnitude of the heart dipole over time and can be calculated by applying a mathematical transformation to the 12-lead ECG signal (Edenbrandt and Pahlm 1988, Kors *et al* 1990, Vozda and Cerny 2015). The entries of the transformation matrix for these methods are fixed, hence ignoring specific characteristics of the signals under study and being susceptible to information loss (Kossmann *et al* 1967, Nelwan *et al* 2000).

© 2023 The Author(s). Published on behalf of Institute of Physics and Engineering in Medicine by IOP Publishing Ltd

Laplacian Eigenmaps (LE) are a data-driven dimensionality reduction technique that allows identifying low-dimensional structures in high-dimensional data (Belkin and Niyogi 2003). Based on manifold learning, the algorithm creates a new representation of the data, which can be reduced to a low-dimensional space referred to as LE space. While other approaches in literature have been proposed for dimensionality reduction, such as principal component analysis or singular value decomposition (SVD) (Acar and Koymen 1999, Castells *et al* 2007), one of the most relevant aspects of LEs is the enhancement of imperceptible changes in the signals thanks to the use of the Laplacian. Its use allows capturing nonlinear characteristics of the data, while preserving locality-information. This approach was first used for the analysis of dynamical changes in multi-lead electrogram signals in time by Erem *et al* (2012). Each heartbeat is represented by a trajectory in the LE space, and authors quantified the changes between consecutive trajectories, which were then related to the effects of ischemia on the electrical properties of the heart. Erem *et al* proposed an out-of-sample extension for the LE algorithm, which allowed to first build a LE space in a training phase, and later map other signals onto it (Erem *et al* 2016). This formulation has been applied to the identification of changes in electrograms caused by ischemia, comparing it to traditional signal-derived features (Good *et al* 2018, 2020).

While those studies have explored dynamic changes within the data of a single subject, the comparison of multiple subjects is not straight-forward. The LE algorithm derives a space from the signals, and the out-of-sample extensions allow mapping back future instances of these signals to this original map (Erem *et al* 2016, Good *et al* 2018, 2020). Nevertheless, this LE space is signal-dependent, and hence the morphologies of the trajectories of different subjects in their corresponding spaces cannot be compared.

The goal of this work is to propose a unified framework based on LE for the characterization of multi-lead ECG signals of different subjects in a common lower-dimensional space. The main novelty is the definition of a reference LE space onto which new subjects can be mapped, therefore generalizing the out-of-sample extension properties of the algorithm. This novel methodology merges the power of LEs to reduce dimensionality and identify abnormalities in multi-lead signals, with the utility of a generalized approach that allows comparing signals of different patients. This space is created using ECG signals recorded during sinus rhythm from presumed healthy subjects with comparable demographics. Our hypothesis is that ECGs of patients with cardiac disease will translate in different, more complex patterns than those of healthy subjects. These differences could be detected by analyzing the trajectories related to each type of subject, or even those of a particular subject at different moments in time. We propose a set of distance metrics to compare different trajectories, and other features to characterize their morphology.

This method was evaluated on the characterization of structural and arrhythmogenic signs in the ECG. First, the framework was applied to discriminate ECGs of healthy subjects from ECGs of patients with ischemic or non-ischemic cardiomyopathy. Additionally, it was applied for the discrimination between cardiomyopathy patients with and without ventricular arrhythmia. The goal was to capture the differences between these groups in the reference space. To highlight the relevance of data-driven approaches, these experiments were also evaluated using state-of-the-art VCG features, which were derived from the 12-lead ECG signals.

The LE framework is publicly available in https://github.com/avillago/LEnormalSpace, where the codes to generate the reference space are available, together with examples of how to map new data onto it and how to derive the distance metrics.

2. Methods

2.1. Data

This work used ECG signals from three groups. The first population includes healthy subjects, the second gathers patients with ischemic and non-ischemic cardiomyopathy; and the last one includes cardiomyopathy-patients either with or without ventricular arrhythmia. The different populations were matched for age and female to male ratio to the smallest group, i.e. the cardiomyopathy-patients with ventricular arrhythmia, more specifically, with a history of electrical storm (ES).

ES is defined as 3 or more appropriate interventions within 24 h for life threatening ventricular arrhythmia by an implantable cardioverter-defibrillator (ICD) (Eifling *et al* 2011). The ES patients were recruited from the database of patients with an ICD in the University Hospitals Leuven (UZL). This is a single-center registry containing information of all patients with ischemic heart disease (IHD) or dilated cardiomyopathy (DCM) implanted with an ICD between 01/01/1996 and 31/12/2018. The Ethical Committee of UZ Leuven approved retrospective analysis on this database. This database contains 12-lead ECG signals of 10 s, sampled at 250 or 500 Hz. Two different subsets of patients were extracted from this database according to their pathology and history of appropriate ICD interventions as surrogate for lethal ventricular arrhythmias. The first subset were patients with ES. Exclusion criteria were age at implantation $\leqslant$ 18 years, less than 6 months between ICD implantation and ES, ES induced by sepsis, and permanent ventricular pacing. In total, 32 ES patients were

Table 1. Baseline demographics of the ICD database from UZ Leuven and the matched normal ECGs from PTB-XL.

Mean (SD)	*H*	IHD	DCM	ES	Non-ES
Age(years)	63.5 (7.9)	65.7 (7.8)	62.2 (5.7)	64.0 (9.5)	63.0 (9.9)
LVEF(%)	NA	35.4 (13.3)[a]	27.9 (6.8)	38.4 (13.6)	37.8 (11.2)
HR(bpm)	70 (12)	68 (16)	71 (16)	65 (10)	65 (8)
QRS(ms)	78 (14.8)	87.7 (12.7)[b]	89.4 (12.1)[b]	120.0 (32.5)[b]	111.4 (32.9)[b]
Number (%)					
Male	29 (91)	28 (88)	22 (69)	29 (91)	29 (91)
IHD	NA	32 (100)	0 (0)	28 (88)	28 (88)
Sec.Prev.	NA	11 (34)	5 (16)	25 (78)	25 (78)

LVEF left ventricular ejection fraction, **HR** heart rate, **QRS** QRS duration, **Sec. Prev.** Secondary prevention.

[a] Sig. Diff. between IHD and DCM.

[b] Significant differences to the *H* group.

identified and matched for age, sex, underlying cardiomyopathy (including both IHD and DCM), and type of prevention with 32 ICD patients without any appropriate ICD intervention (non-ES). Baseline demographics are shown in table 1, under the columns ES and non-ES. The ECG signals considered for the ES patients correspond to the last visit prior to the ES event, in which the signal is in sinus rhythm or atrial fibrillation (AF). The average time to ES of the selected signals is 4 months, with a standard deviation of 2.5 months.

In order to evaluate the framework on structural ECG changes, a subset of signals from ICD patients without any appropriate ICD intervention at the moment of database closure is selected. To exclude QRS-width as confounding factor, patients with a narrow QRS (QRS duration $\leqslant$ 120 ms) were selected from the UZL registry. A total of 64 ECG signals on sinus rhythm or AF were selected, 32 from IHD patients and 32 from DCM patients. IHD patients have a documented history of myocardial ischemia, while those with DCM presented cardiac dysfunction in absence of significant coronary lesions. Baseline demographics of these subjects are presented in table 1.

The definition of the normal space required ECG signals from healthy subjects. These signals were extracted from the PTB-XL database (Wagner *et al* 2020), which contains 21837 12-lead ECG recordings of 10 sampled at 500 Hz. These recordings include annotations according to the SCP-ECG standards about diagnostic and rhythm; as well as a small description of abnormalities in the signals (ISO Central Secretary 2009). For our aim, we considered a subset of 865 signals from different subjects matching age and male to female ratio to the ES population, including only signals belonging to the normal class, sinus rhythms and without abnormalities. Out of these signals, 32 were selected to build the training space. The details of this group are shown in table 1, under the column H Albeit there is no detailed information about the comorbidities of the patients without abnormalities on their ECG, they were considered to be healthy subjects (Wagner *et al* 2020). In the remaining of this paper, the ECG signals of this group are referred to as normal ECGs. Despite the selection of patients with a QRS-width $\leqslant$ 120 ms for comparison, as shown in table 1, the QRS duration of the different groups was significantly different to that of the *H* group. Hence, this is considered in our experiments.

2.2. Preprocessing

The preprocessing of the signals comprises all the steps applied to the ECGs prior to the use of the LE algorithm, as shown in figure 1.

First, the heartbeats were detected and delineated. Since we focus on changes in ventricular depolarization, only the QRS complex was considered in the analysis. To segment the QRS complexes, the signals were first bandpass-filtered between 0.35 and 30 Hz, to remove baseline wander and reduce the impact of high-frequency noise. Then the absolute value of each lead was calculated, and the addition of these 12 absolute leads was used to delineate the QRS complexes. Due to the differences in electrode placement for each of the leads, there are small delays in the timing of the ECG waves. Taking this into account, performing QRS delineation in the absolute added signal guarantees that the complete QRS complex is included in each of the leads. First, the R-peaks were identified using the algorithm proposed in the R-DECO software (Moeyersons *et al* 2019), and the onset and offset of the QRS complexes were detected using the ECGkit (Martínez *et al* 2004).

Even though high frequencies were removed for the QRS delineation process, the information of those frequency bands is relevant in our analysis. The presence of high-frequency notches during the QRS complex

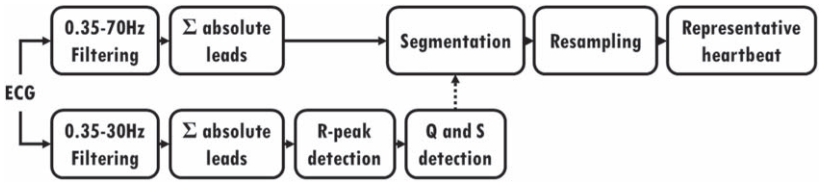


Figure 1. Preprocessing steps of the ECG signals prior to applying the LE algorithm.

can be associated to QRS fragmentation, which is an indicator of high risk of arrhythmia (Vandenberk *et al* 2017). Therefore, the raw 12-lead ECG signals were bandpass-filtered between 0.35 and 70 Hz, and the location of the Q and the S waves detected before were imported to this signal, as indicated by the dotted arrow in figure 1. Once the QRS complexes were segmented, they were resampled to a common length of N, here fixed to 200 samples. By doing so, segments could be comparable regardless of their original sampling frequency, which differs for each of the databases considered. Hence, this step of resampling allows the application of this framework to signals at any sampling frequency. A representative heartbeat per signal was selected, as the one with the highest average of the normalized lead per lead correlation with all the other QRS complexes of that recording.

2.3. Definition of the reference LE space

Using the signals of healthy subjects, a low-dimensional LE space was defined, where the underlying manifold structure of the data is captured. This means that the most relevant characteristics of the data are emphasized in this alternative space, while reducing the number of dimensions. The training and explicit mapping stages defined in Erem *et al* (2016) were used here to calculate the space, and further map new data onto it, with the novelty of enabling the comparison of signals from different subjects.

The first step is to concatenate the representative heartbeats of each of the 32 normal ECGs. This data can be referred to as a set of points $\mathbf{X} = \{x_i\}_{i=1}^{L} \in \mathbb{R}^{L\times d}$, with L the number of concatenated samples, i.e. 200 samples per heartbeat, 32 times; and d the number of dimensions (i.e. 12 channels) in the data. Both d and L can take different values according to the application, enabling the use of this methodology for signals of different number of leads as well as the selection of different number of samples to create the LE space. For the case of 12-lead ECG signals, the limiting factor in terms of computation is the number of samples L, since it determines the size of the matrix that needs to be stored in the memory and manipulated to derive the SVD. The structure of the data is shown in figure 2. To emphasize the nonlinear relations between the points X, a similarity matrix W is calculated using the pairwise kernel similarity between the points. Each entry of this matrix $\mathbf{W} \in \mathbb{R}^{L\times L}$ is calculated as:

$$\mathbf{W}_{i,j} = exp\left(-\frac{(\|x_i - x_j\|_2^2)}{\sigma^2}\right), \tag{1}$$

where i and j take values between 1,…,L; $W_{i,i} = 1$, $W_{i,j} = W_{j,i}$, and σ^2 is the radial basis function (RBF) kernel parameter. In W, those pairs of points that are closer in the high-dimensional space receive a higher value $W_{i,j}$ than those that are very distant. The value of σ is set to the maximum $\|x_i - x_j\|_2^2$ value as suggested in Erem *et al* (2016). Then the degree matrix D is computed, as a diagonal matrix in which each element is derived as $D_{i,i} = diag(\sum_j W_{i,j})$. This diagonal matrix captures the relevance of each point x_i in the manifold, since each entry $D_{i,i}$ is defined by the connectivity of the point x_i with the rest of the points in X.

These matrices allow the formulation of the SVD problem, defined as:

$$\mathbf{D}^{-1}\mathbf{W} = \mathbf{U}\mathbf{S}\mathbf{V}^T, \tag{2}$$

enabling the calculation of matrix V, where each column corresponds to one dimension of the LE space. These columns are sorted according to their associated singular values contained in S, from high to low values according to the amount of information they capture of the high-dimensional space. Since the first column is constant (all ones), the next ones are considered when constructing the LE space. In line with (Belkin and Niyogi 2003), the next three dimensions are used to define the space. These three dimensions capture >90% of the information contained in the data, according to the singular values, and they allow intuitive visualization of the data (Belkin and Niyogi 2003). The LE space is then defined by 3 vectors:

$$\mathrm{LE} = [V_2,\ V_3,\ V_4]. \tag{3}$$

Additionally, the LE space is characterized by a template trajectory or loop. The representative heartbeats of each of the normal ECGs contained in this LE space can be mapped onto it, creating a trajectory, as those observed in gray in figure 3. All the loops fall within a manifold that could be represented by a template loop.

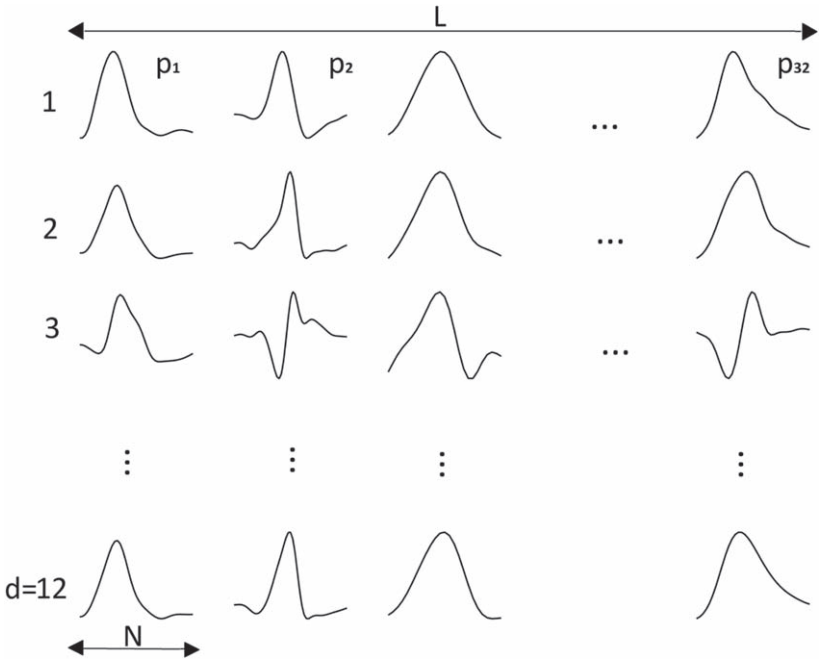


Figure 2. Structure of the data used to build the matrix required for the LE algorithm.

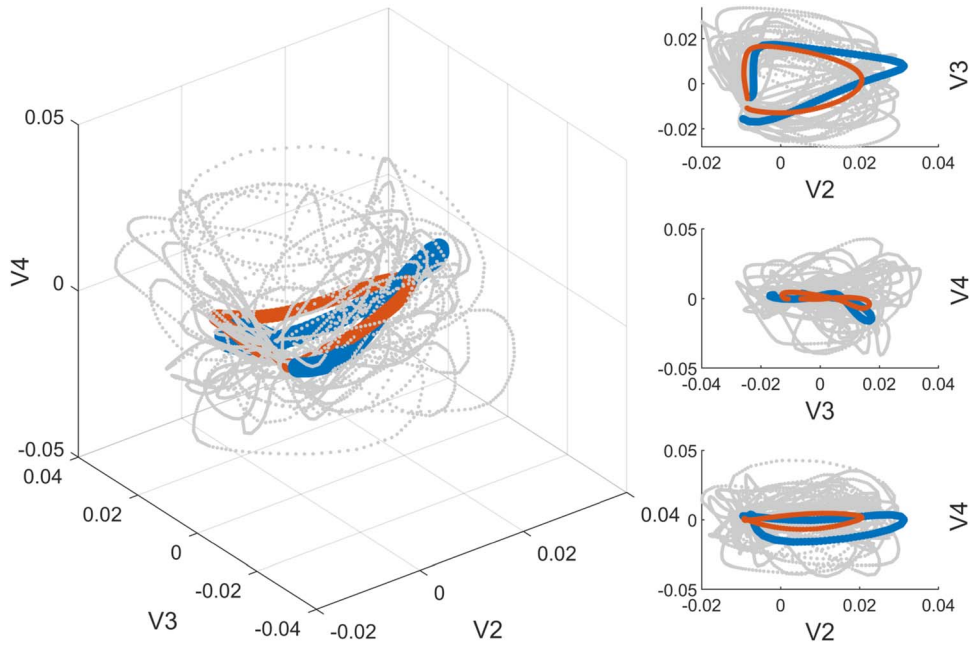


Figure 3. Trajectories of the ECGs used to build the reference space (gray). The template is shown in blue, and it corresponds to the trajectory that is closest to the average of all the training loops (in orange).

This template is selected as the normal loop out of the 32 closer to the mean of all according to the euclidean distance. This mean is calculated point by point, obtaining an average loop of the same length as those of the different subjects. Figure 3 presents in gray the loops of the normal ECGs used to build the LE space, the template loop in blue, and the average used to specify the template, in orange.

2.4. Explicit mapping

As suggested in Erem *et al* (2016), the decomposition formulated in equation (2) can be used to project multi-lead signals of new patients onto this manifold. In our case, the signal to be projected is the representative heartbeat of a new patient, extracted as explained in section 2.2.

Table 2. Features derived from the LE trajectories.

Features	Notation
Distance metrics	
Point-to-point	d_{P2P}
Dynamic time warping	d_{DTW}
Hausdorff	d_{Haus}
Loop characterization	
Range V_2	r_{V_2}
Range V_3	r_{V_3}
Range V_4	r_{V_4}
SD V_2	σ_{V_2}
SD V_3	σ_{V_3}
SD V_4	σ_{V_4}
Mean curvature	$\overline{\mathrm{curv}}$
SD curvature	σ_{curv}

The explicit mapping function is defined as

$$f(y) = S^{\dagger}U^{T}D^{-1}g(y), \tag{4}$$

where

$$g(y) = \left[exp\left(-\frac{\|y - x_1\|_2^2}{\sigma^2} \right) \ldots exp\left(-\frac{\|y - x_N\|_2^2}{\sigma^2} \right) \right]^{T}, \tag{5}$$

and $\mathbf{S}^{\dagger}$ is the Moore–Penrose pseudo-inverse of matrix $\mathbf{S}$. The 2nd, 3rd and 4th columns of the transposed output of $f(y)$ contain the low-dimensional representation of the test data $Y = \{y_i\}_{i=1}^{L} \in \mathbb{R}^{L\times d}$ (Erem *et al* 2016).

The new heartbeats mapped onto this space also form a trajectory that can be compared to that one of the normal template heartbeat to identify differences. Additionally, the differentiable characteristics of this formulation allow mapping points back from the LE space to the original multi-lead signal.

2.5. Feature extraction

A set of features is proposed to characterize the LE loops of different groups of subjects. Two types of features are proposed for this aim, namely distance metrics and features that intrinsically characterize the trajectories. They are summarized in table 2.

2.5.1. Distance metrics

These features characterize differences as the distance between two trajectories. Following our previous work (Jacobs *et al* 2020), three of the metrics proposed are further explored in this study. The first is the point-to-point distance which, given that all loops are resampled to the same length, calculates the euclidean distance between each pair of points. The final distance is the average of all the distances d_i, with $i = 1,\ldots,L$. The second distance is based on dynamic time warping (DTW) (Müller 2007). Similarly to the first metric, the final distance is calculated as the average of all the distances along the loop. The last metric is the Hausdorff distance, which is widely used to calculate the similarity between two sets of points (Zhang *et al* 2017).

2.5.2. Loop characterization

The trajectories in the LE space can also be independently characterized, i.e. without comparing them to any other. Given the formulation of the LE approach as a generalized eigenvector problem, the values of the trajectories in each dimension are related to their corresponding singular value and hence, to the amount of information that they capture. Therefore, the hypothesis here is that the simpler loops, corresponding to normal ECGs, should contain most of their information in V_2. In contrast, complex loops differ from a normal sinus rhythm, and they shall distribute their information along the other dimensions. To capture this, the metrics proposed are the range and the standard deviation of the loops in each of the 3 LE dimensions. The range r_{V_i} represents the difference between the maximum and minimum value of the points in each of the LE dimensions, and the standard deviation σ_{Vi} represents the dispersion of those values from the average. Additionally, the mean and the standard deviation of the curvature measured along the loop are proposed to characterize the morphology of the loops. The curvature for each point x_i of the loop is the inverse of the radius circumscribing points x_{i-1}, x_i and x_{i+1} (Are Mjaavatten 2021). Therefore, the average and the standard deviation of these values along the vector characterize the roundness of the loops in the space.

2.6. Experiments

The potential of the proposed reference LE space to characterize different groups of patients was evaluated in three experiments. The reference space for all these experiments is the same, since all the groups of patients were matched on a population level.

2.6.1. Capturing characteristics of normal ECGs

The first experiment consisted on verifying if the LE space was able to capture characteristics of normal ECGs. The hypothesis was that all normal sinus rhythm ECGs from healthy subjects of comparable age should present similar properties in this reference space.

To prove this, 10 separate groups of 32 normal ECGs, referred to as H_1,H_2,… H_{10}, were mapped onto the reference space. There was no repetition of subjects between the different groups, neither with those included in the creation of the reference space.

2.6.2. Identifying cardiovascular diseases: differences between IHD and DCM

The second experiment aimed to verify if the LE reference space could distinguish between ECGs from healthy subjects and those from patients of comparable age and gender with underlying cardiovascular diseases and a narrow QRS. To this end, the group of patients with no history of shocks from the UZL was mapped onto the reference space, including 32 IHD patients and 32 DCM patients.

The IHD patients considered have a history of myocardial infarction. Impaired coronary flow, mostly due to coronary atherosclerosis, can lead to hypoxia and subsequent cell death of cardiomyocytes in the region of the affected coronary artery (Institute of Medicine (US) 2010). This results in localized scar formation, which affects the conduction characteristics of the heart. Pathological Q-waves are a well-known ECG biomarker of previous infarction (de Luna and Fiol-Sala 2008). The second group of patients contains DCM patients. While in IHD the problem arises at the level of the coronaries, DCM primarily affects the heart muscle. There is a variety of causes that can underlie DCM, but they all lead to left or biventricular dilatation with systolic dysfunction (Jefferies and Towbin 2010). Imaging is necessary for diagnosis, but some ECG biomarkers have been proposed to aid in the diagnostic work-up, for example low QRS voltages due to loss of cardiomyocytes, presence of left bundle branch block and T-wave inversion (Finocchiaro *et al* 2020).

In this experiment, the goal was to evaluate if these two groups present different properties in the LE reference space, as well as if these are different than those of healthy subjects. The hypothesis is that IHD ECGs should be associated to the most complex loops in the LE space, in line with their morphology in the 12-lead format. On the contrary, DCM trajectories are hypothesized to differ from healthy subjects more subtly.

To assess the nonlinear characteristics of the embedded signal space, the ability of the RBF kernel to serve as a universal approximator was exploited. Different values for the kernel parameter σ were evaluated both for the definition of the LE map and for the mapping of signals onto it. As explained in section 2.3, σ is tuned as the maximum of $\mathbf{W}$. Additionally, two other values where evaluated, namely $\sigma_L = 10^*\sigma$, as an approximation to more linear kernel, and $\sigma_{NL} = 0.1^*\sigma$ to impose a more nonlinear case.

Our approach proposes a framework for data-driven dimensionality reduction of the ECG. However, as stated before, the derivation of the VCG from the 12-lead ECG can also be understood as a lower dimensional representation of the ECG. To identify the added value of our approach, two additional features based on VCG were included in this experiment.

The Kors transform was used to derive the VCG from the ECG signals (Kors *et al* 1990), since several studies suggest this approach as the one reconstructing Frank's VCG most accurately (Vozda and Cerny 2015, Jaros *et al* 2019). From this derived VCG, the spatial QRS-T angle and the standard deviation of the speed (*speedSD*) of the QRS loop were calculated for each of the three patient populations.

2.6.3. LE biomarkers for arrhythmia prediction: ES

The last experiment explored the use of the LE reference space to identify patients at risk of ES. For this aim, the 64 patients included in the ES study were mapped onto the reference space. Half of these patients had a history of ES and the other half did not. Both groups included IHD and DCM patients in a similar proportion. Signals from ES patients were hypothesized to present possible signs associated with arrhythmia, such as fragmentation (Das *et al* 2010, Villa *et al* 2020).

Additionally, the impact of the kernel parameter σ was evaluated; and the VCG features QRS-T angle and *speedSD* were calculated.

Table 3. Results for H_1, IHD, DCM, ES and non-ES. The average and standard deviation values of the features are presented, together with the statistical comparisons per experiment (i.e. structural or arrhyhmogenic differences compared to H_1). Differences are significant for $p < 0.05$.

Features	H_1	IHD	DCM	ES	Non-ES
d_{P2P}	3.4 ± 0.8	3.9 ± 0.8[a,b]	3.4 ± 0.8	4.3 ± 0.6[a]	3.9 ± 0.6
d_{Haus}	2.7 ± 0.5	3.0 ± 0.7	2.8 ± 0.6	3.2 ± 0.6[a]	2.9 ± 0.5
d_{DTW}	2.9 ± 0.5	3.5 ± 0.8[a,b]	3.0 ± 0.7	3.8 ± 0.7[a]	3.6 ± 0.6[a]
r_{V_2}	3.8 ± 0.5	2.8 ± 0.8[a]	3.0 ± 0.7[a]	2.8 ± 0.8[a]	2.7 ± 0.8[a]
r_{V_3}	4.0 ± 1.2	4.0 ± 0.9	4.5 ± 0.8	3.8 ± 1.2	3.7 ± 0.9
r_{V_4}	4.7 ± 0.2	5.0 ± 1.1	4.8 ± 1.2	4.7 ± 1.1	4.6 ± 1.4
σ_{V_2}	1.1 ± 0.6	0.7 ± 0.3[a]	0.8 ± 0.2[a]	0.7 ± 0.3[a]	0.7 ± 0.3[a]
σ_{V_3}	1.1 ± 0.4	1.1 ± 0.3	1.2 ± 0.3	1.0 ± 0.3	1.0 ± 0.2
σ_{V_4}	1.1 ± 0.1	1.2 ± 0.3	1.1 ± 0.3	1.2 ± 0.3	1.1 ± 0.3
$\overline{\mathrm{curv}}$	27.0 ± 7.3	29.8 ± 10[b]	34.6 ± 7.7[a]	30.0 ± 14.6	29.9 ± 12.0
σ_{curv}	29.7 ± 9.3	28.9 ± 9.3	31.2 ± 9.6	33.5 ± 27.8	36.2 ± 30.8

[a] Significant differences to the H group.

[b] Sig. Diff between IHD and DCM.

2.7. Statistical tests and performance metrics.

The LE features were calculated for each group of patients. Significant differences between these groups were evaluated using the Kruskal–Wallis test ($p < 0.05$) with Bonferroni correction for multiple comparisons. Correlation tests between features were evaluated using Pearson correlation.

Additionally, the classification performance of each feature for experiments 2 and 3 was evaluated using an SVM classifier with the RBF kernel, with five-fold cross-validation for training and evaluation. Due to the low amount of data, this approach was considered suitable as it gives an indication of how discriminative the features are. The area under the curve (AUC) of the receiver operating characteristics curve (ROC) was used to evaluate the performance of this classifier for each of the features considered. For experiments 2 and 3, 12 different AUC metrics were extracted: one for each of the features proposed in a uni-variate trend and a multi-variate one combining the 11 features. Given the multi-class nature of each of these experiments (H versus IHD versus DCM) and (H versus ES versus non-ES), the ROC values are provided as one class again the rest.

3. Results

3.1. Capturing characteristics of normal ECGs

The first experiment comparing 10 different groups of normal ECGs in the reference space did not find any significant difference between them. This confirmed that the reference space was able to reproducibly characterize normal ECGs. Therefore, in the following experiments the groups of patients are compared to the group H_1 of normal signals. The average and standard deviation of the LE features for the H_1 group are shown in table 3.

3.2. Identifying cardiovascular diseases: differences between IHD and DCM

The features obtained for the IHD and DCM group are shown in table 3. The symbol *indicates significant differences (p-value < 0.05) between IHD or DCM patients and normal ECGs; and † between the IHD and DCM groups. The violin plots related to the features showing significant differences are presented in figure 4, where these differences are indicated with $*$. Table 4 summarizes the AUC of the ROC values for single-feature and multi-feature classification.

The distance metrics identified significant differences between IHD and the other two groups, obtaining d_{DTW} the higher AUC values. The mean curvature $\overline{\mathrm{curv}}$ of the trajectories allowed to distinguish between all the three groups. However, the AUC values are only around 0.7 when comparing DCM against the other two classes. Additionally, features related to the first dimension of the LE space (r_{V_2} and σ_{V_2}) were significantly different between normal ECGs and the two groups of patients. The AUC metric for these two features was also specially high in these cases. However, this last distinction was also reflected in the significant differences between QRS duration of normal ECGs and patients shown in table 1.

In view of this potential relation between the features, the correlation between QRS duration and each feature was calculated, obtaining the highest values for σ_{V_2} (-0.27), and for the average curvature (-0.29). The upper plots of figure 5 show the relation between these features and QRS duration. The shaded areas give a visual intuition of the separation that could be done between normal trajectories (blue points) and the patient groups (purple and yellow symbols).

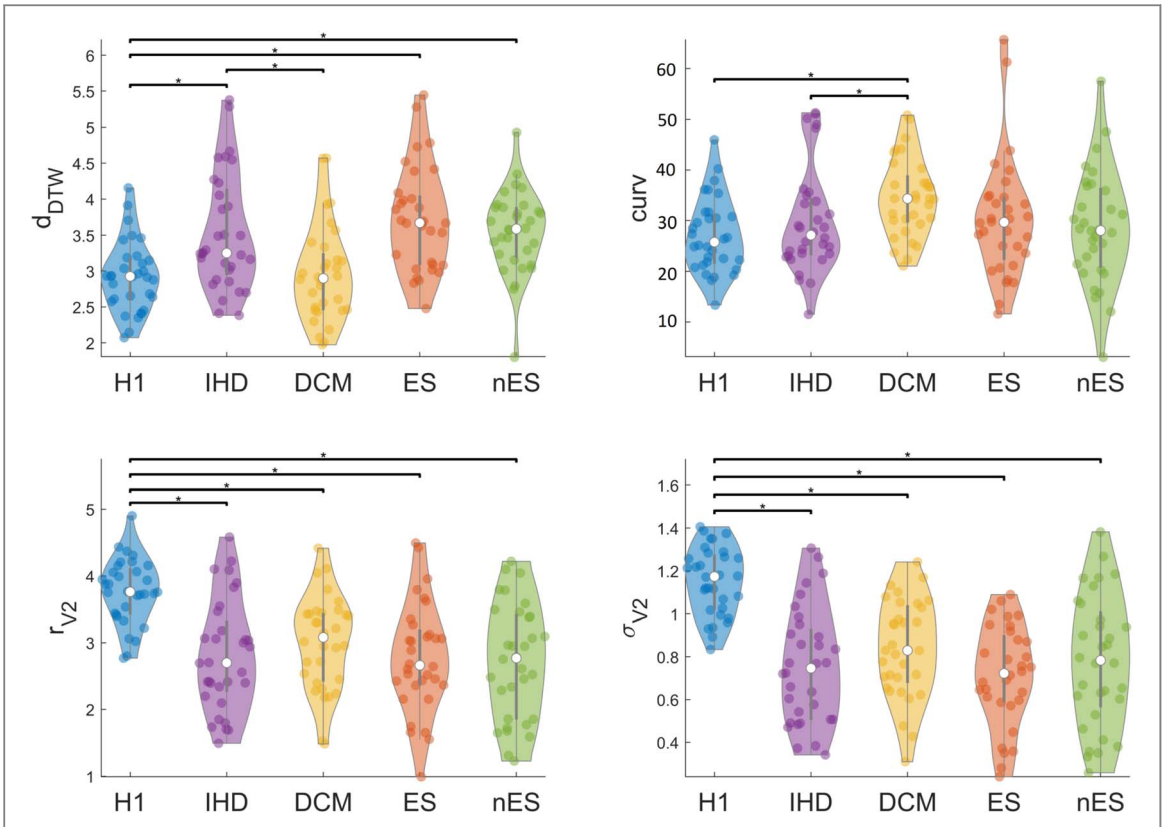


Figure 4. Violin plots for the features presenting significant differences for experiments 2 and 3. H_1 presents the values for normal ECGs, compared to those of IHD and DCM patients, and to ES and non-ES. Significant differences are indicated as * for p-value < 0.05.

Table 4. AUC values for classification between each of groups marked in the column and the other two, for experiments 2 (left block) and 3 (right block). The last row shows the results for the multi-feature case.

	AUC ROC for column class					
Feature	*H*	IHD	DCM	*H*	ES	Non-ES
d_{P2P}	0.48	0.60	0.52	**0.72**	0.66	0.65
d_{Haus}	0.44	0.52	0.41	0.66	0.59	0.60
d_{DTW}	0.57	**0.70**	0.52	**0.81**	0.61	0.60
r_{V2}	**0.80**	0.65	0.59	**0.84**	0.56	0.58
r_{V3}	0.58	0.50	0.57	0.50	0.40	0.50
r_{V4}	0.54	0.51	0.48	0.37	0.42	0.45
σ_{V2}	**0.85**	**0.69**	0.64	**0.83**	0.67	0.48
σ_{V3}	0.46	0.48	0.47	0.41	0.43	0.50
σ_{V4}	0.61	0.60	0.55	0.47	0.56	0.58
$\overline{curv}$	0.58	0.53	**0.69**	0.59	0.54	0.50
σ_{curv}	0.40	0.55	0.59	0.53	0.50	0.49
All	**0.72**	0.62	0.56	**0.89**	0.62	0.54

The evaluation of different σ values suggests a more linear behavior of the LE space, since the results for $10^*\sigma$ were similar to those obtained with σ. In contrast, the results for a smaller kernel parameter ($0.1^*\sigma$) resulted in overfitting. Due to the highly nonlinear behavior, the algorithm was unable to find any similarity between the features and groups.

In contrast to the LE features, neither the QRS-T angle nor the *speedSD* were able to identify significant differences between IHD and DCM patients. However, *speedSD* was significantly lower for these two groups when compared to the healthy group, i.e. the speed of the QRS loop was faster for healthy subjects. However, this feature presented a correlation of −0.42 with the QRS duration, which is higher than those obtained for the LE features.

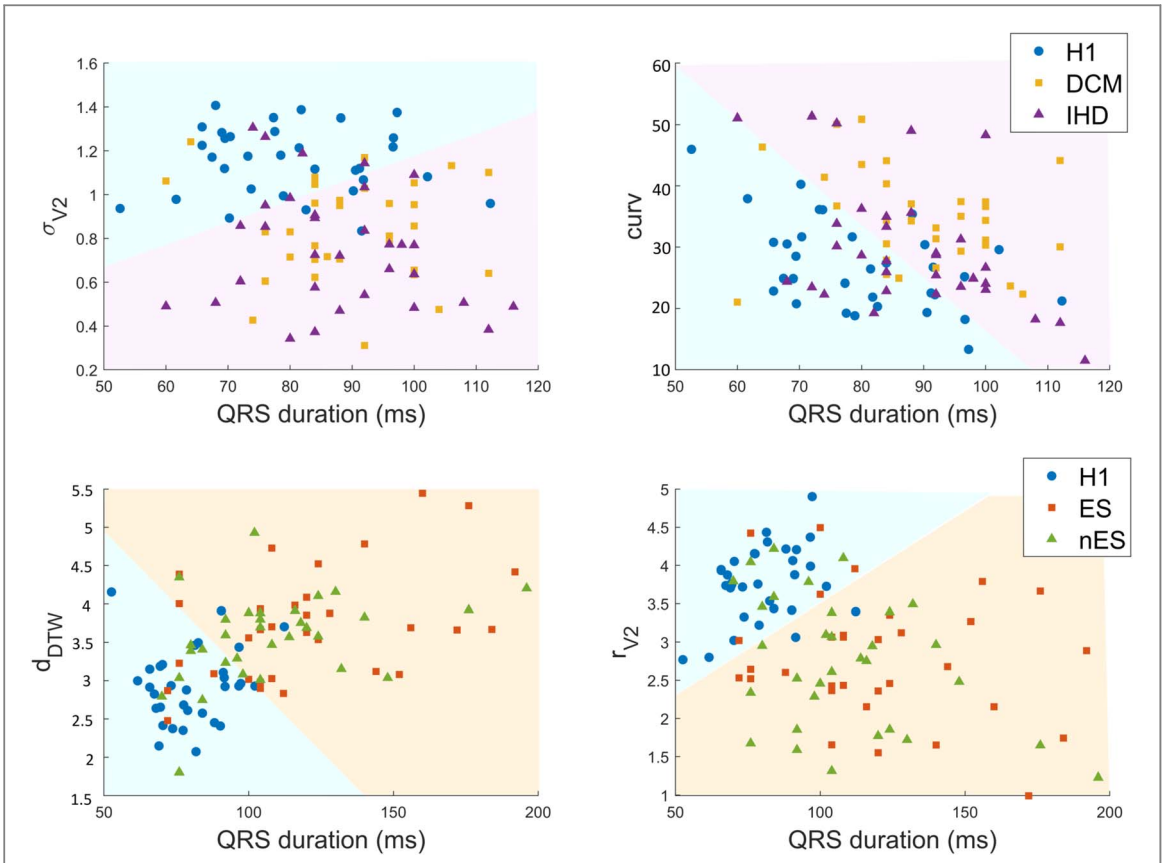


Figure 5. Relation between QRS duration and the features for the cases of higher correlations. In the top row, σ_{V_2} and $\overline{\text{curv}}$ for Experiment 2 exploring DCM and IHD patients, and on the lower row d_{DTW} and r_{V_2} for experiment 3, exploring ES and non-ES patients.

3.3. LE biomarkers for arrhythmia prediction: ES

The average values of the features for both ES patients and their control group are shown in the last columns of table 3, and the AUC results are shown in the right block of table 4. The LE features could not identify significant differences between the ES and non-ES group. This is also shown in the AUC values of table 4, where all AUC values are below 0.7. Nevertheless, the space was able to distinguish between normal ECGs and these patients. The most relevant features are the distance features, specially d_{DTW}, and those related to V_2, both the range and the standard deviation. However, in this case the correlation of these features with QRS duration is stronger than in the previous experiment, since these patients have much wider QRS complexes than healthy subjects. The correlations with QRS duration for d_{P2P}, d_{DTW}, r_{V_2} and σ_{V_2} ranged between 0.44 and 0.57. Figure 5 shows the relation of d_{DTW} and r_{V_2} with the QRS duration. Despite this weak correlation, the combination of QRS duration and these features contributes to differentiate between normal ECGs (blue circles) and both ES and non-ES signals (orange and green signs), especially for QRS duration below 100 ms.

The effect of increasing the value of the kernel parameter σ did not have an impact on the final results, suggesting a linear behavior of the LE space.

None of the VCG features were significantly different between ES and non-ES patients. Similarly to the previous experiment, *speedSD* was significantly lower for the two groups of patients compared to those obtained for the healthy group. Due to the longer QRS durations presented by these two patient groups, the correlation between this characteristic and *speedSD* was higher in this experiment, reaching a value of $\rho = -0.57$. This indicates how subtle the QRS speed is to the length of the QRS complex.

4. Discussion

This study explored the potential of building a reference LE space based on normal multi-lead ECG signals, and characterizing different groups of subjects mapped onto it.

The first experiment confirmed that the LE reference space capture the characteristics of normal ECGs. No significant differences were found between the 10 groups of signals from healthy subjects, confirming that the method was able to detect that they all belonged to the same population.

Secondly, the methodology was applied to identify differences between ECG signals belonging to patients with IHD, DCM, and normal ECGs. The distance metrics were significantly larger for IHD ECGs than for normal ECGs or DCM patients, achievening also the highest AUC for d_{DTW}. This indicates that the trajectories of IHD patients are more different from those derived from normal ECGs and DCM patients. While IHD and DCM are distinguishable using gadolinium enhanced cardiac MRI, where DCM often presents with patchy scars, this distinction is not yet possible using the QRS complexes of the ECG (McCrohon *et al* 2003). In IHD the electrical propagation over the heart is disturbed by scar tissue that affects the conduction through the ventricles, which leads to changes in the ECG (de Luna and Fiol-Sala 2008, Das *et al* 2010, Institute of Medicine (US) 2010). Nevertheless, most of the well-known characteristics of IHD ECGs involve irregularities at the end of the QRS complex, which are difficult to detect, often presenting high false-positive rates (Nadour *et al* 2014). However,in our experiments the framework is able to identify differences during ventricular depolarization (QRS complex). This suggests a big advantage of LE methods especially in noisy signals, since the low amplitude of the T-wave makes it more sensitive to artefacts (Clifford *et al* 2006, Bortolan *et al* 2015).

Additionally, the analysis of the mean curvature of the trajectories of the three groups showed significant intergroup differences, with higher AUC values for DCM signals. This metric can be associated with the complexity of the ECG signal: higher values of curvature describe more convex curves, while a regular QRS complex is expected to be characterized by a smooth rounded trajectory. Therefore, the most complex loops should be those of cardiomyopathy patients. The discovery of significant differences in this aspect between healthy and DCM ECGs is remarkable, since the characteristics of DCM ECGs are not evident by visual inspection (Schultheiss *et al* 2019). Nevertheless, the presence of diseased cardiomyocytes and patchy scar in the heart of patients with DCM can be the cause of micro-fragmentation in the QRS complex (Sha *et al* 2011), which could be captured by the LE algorithm.

These differences were also captured by features r_{V_2} and σ_{V_2}, associated to the first dimension of the LE space. Out of the three dimensions of the LE space, V_2 is the one that contains most of the information when the trajectories are simple, i.e. more similar to a normal regular sinus rhythm. More complex heartbeats, which may include fragmentation or conduction delays, will contain more information in the second and third dimension. This intuition is shown in the results for r_{V_2} and σ_{V_2}, which are higher for normal ECGs than for DCM and IHD patients. While these features were slightly correlated to the duration of the QRS complex, as shown in figure 5, both aspects are complementary for the distinction between normal and abnormal ECGs.

Finally, the value of this methodology to identify features related to the future occurrence of arrhythmia was tested. None of the described features could detect differences between patients presenting with ES and those who did not. Currently no clear ECG biomarkers for risk prediction of this serious arrhythmic event are known. While previous work linked the manifestation of ES to the presence of QRS fragmentation (Villa *et al* 2020), this is not visible in this limited database. Further studies should explore comparable times between recordings, to reduce the impact of the time to event, which was variable and long in our series. Additionally, our dataset mixed patients with IHD and DCM, which have different LE trajectories. Therefore, future work should study the ECGs of both populations separately to discard confounding effects, similarly to how it was done with other non-ECG biomarkers (Arya *et al* 2006, Takigawa *et al* 2010).

Similarly to the IHD-DCM study, the ECGs of the patients included in the ES study were found to be significantly different from those of healthy subjects. These findings were moderately correlated to the duration of the QRS, which was larger for the group of patients than for the normal ECGs as reported in the literature (Chen *et al* 2021). Figure 6 shows the representative QRS complexes and the trajectories of the template loop (blue), its closest non-ES (green) and the farthest narrow ES (orange). The ES QRS complex in this case is fragmented in leads II and aVF, while the template and the non-ES signals are not. This could be the reason why the ES trajectory in the LE space contains more information in the third dimension, while the closest cases lay mostly in the horizontal plane (V_2 and V_3). This case enhances the relevance of the LE space to emphasize these differences, regardless of the QRS duration of the signals.

Regarding the comparison of the LE features to those derived from the VCG, we saw how differences between IHD and DCM signals were imperceptible for VCG features, while those coming from the LE framework could identify them. While in general the QRS-T angle was larger for all the patient groups than for healthy subjects, this feature took a wide range of values for all groups. This may be due to the calculation of the angle using the maximal QRS and T vectors, which is very sensitive to outliers and to the automatic methods of peak-detection. On the other hand, while the standard deviation of the speed along the QRS loop could identify differences between healthy and patient groups, it was shown that this feature was moderately correlated to the QRS duration. These results confirm the relevance of data-driven approaches and the proposed LE features to characterize ECG signals.

Limitations and future work

Despite the applicability of the LE framework for the analysis of multi-lead signals of different subjects in a common low-dimensional space, some research questions remain open.

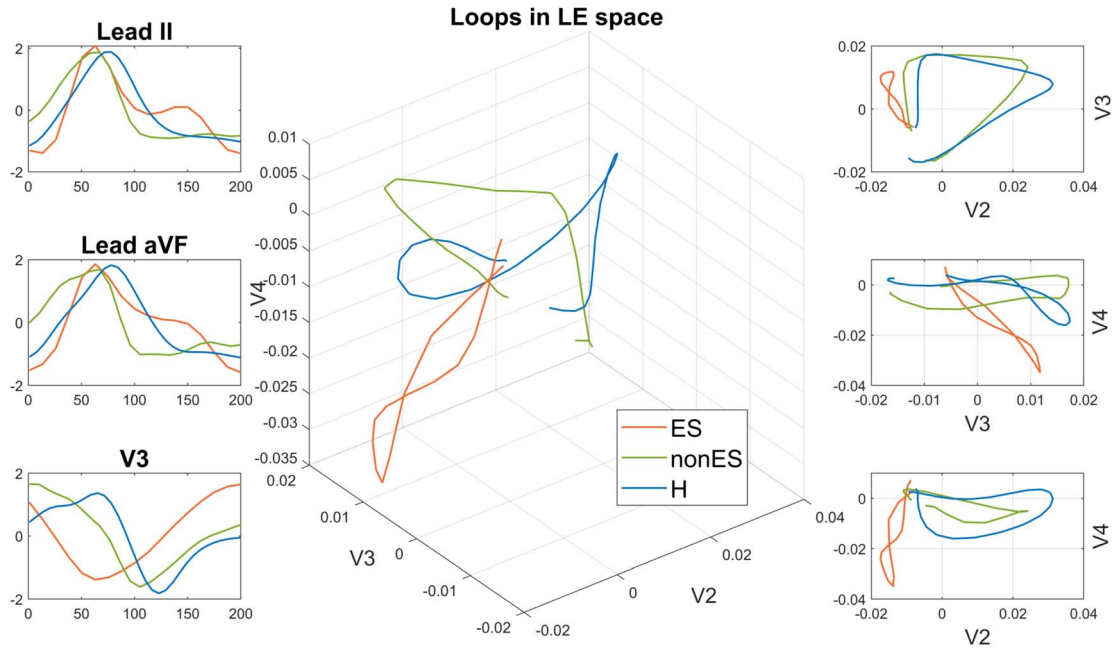


Figure 6. Example of 3 QRS complexes in the LE space and their corresponding ECGs. The ES signal (orange) corresponds to the farthest trajectory to the template, among the narrow QRS complexes. The green one is the closest narrow trajectory to the template (blue).

Firstly, the current work exclusively explores changes in ventricular depolarization, by only including the QRS complex in the analysis. The T wave should also be considered for further studies, as well as the analysis of different segments of the ECG. This would allow the characterization of other rhythms and pathologies. Secondly, the proposed reference space was built using signals of subjects aged between 50 and 75, to match these to the age of the ICD population. The behavior of these maps may differ for signals related to different ages, sex and other clinical characteristics, which deserve further investigation. Additionally, while the power of the LE methodology lies on the use of the Laplacian of the affinity matrix, a detailed analysis of its relevance in different scenarios should be done. This could highlight its added value for ECG signals, given that the LE space appears to be closely linear as reported by the results related to the kernel parameter. Finally, the proposed approach could be applied to other signals, such as electrograms or electroencephalograms by defining a common reference space built with normal signals in this context.

5. Conclusion

This work proposes a novel approach for the analysis of high-dimensional signals in an alternative space derived using the Laplacian Eigenmaps. This data-driven methodology allows reducing the dimensionality of the space spanned by multi-lead signals while emphasizing abnormalities and differences between subjects.

We evaluated this methodology in the context of 12-lead ECG signals, in the analysis of QRS complexes of signals recorded during sinus rhythm and AF. The methodology was able to capture characteristics of normal ECGs, and to distinguish these from those of patients with IHD and patients with DCM implanted with a defibrillator in the prevention of sudden death by cardiac arrhythmia. In particular, the identification of differences between normal and DCM ECGs highlights the possible relevance of the proposed approach to identify differences not directly visible in the ECG. This novel methodology opens the path for future analysis of multi-lead ECGs and other high-dimensional signals in a common low-dimensional space, providing a new field of exploration in which subtle ECG abnormalities can be emphasized.

Acknowledgments

Amalia Villa and Sabine Van Huffel received funding from the Flemish Government (AI Research Program) and are affiliated to Leuven.AI—KU Leuven institute for AI, B-3000, Leuven, Belgium. Rik Willems is supported as postdoctoral clinical researcher by the Fund for Scientific Research Flanders (FWO Vlaanderen). Carolina Varon received funding from the European Space Agency, Belspo (PRODEX-NEPTUNE).

Data availability statement

The data cannot be made publicly available upon publication because they are owned by a third party and the terms of use prevent public distribution. The data that support the findings of this study are available upon reasonable request from the authors.

Ethical statement

The research was conducted in accordance with the principles embodied in the Declaration of Helsinki and in accordance with KU Leuven statutory requirements. All participants gave written informed consent to participate in the study.

ORCID iDs

Bert Vandenberk https://orcid.org/0000-0001-8296-920X

References

Acar B and Koymen H 1999 SVD-based on-line exercise ECG signal orthogonalization *IEEE Trans. Biomed. Eng.* **46** 311–21
Mjaavatten A 2021 Curvature of a 1D curve in a 2D or 3D space (https://mathworks.com/matlabcentral/fileexchange/69452-curvature-of-a-1d-curve-in-a-2d-or-3d-space), MATLAB Central File Exchange. Retrieved December 9, 2021
Arya A *et al* 2006 Prevalence and predictors of electrical storm in patients with implantable cardioverter-defibrillator *Am. J. Cardiol.* **97** 389–92
Belkin M and Niyogi P 2003 Laplacian eigenmaps for dimensionality reduction and data representation *Neural Comput.* **15** 1373–96
Bortolan G, Christov I, Simova I and Dotsinsky I 2015 Noise processing in exercise ECG stress test for the analysis and the clinical characterization of QRS and T wave alternans *Biomed. Signal Process. Control* **18** 378–85
Castells F *et al* 2007 Principal component analysis in ECG signal processing *EURASIP J. Adv. Signal Process.* **2007** 1–21
Chen X *et al* 2021 Incremental changes in QRS duration as predictor for cardiovascular disease: a 21-year follow-up of a randomly selected general population *Sci. Rep.* **11** 1–6
Clifford G D, Azuaje F and McSharry P 2006 *Advanced Methods and Tools for ECG Data Analysis* (Boston: Artech House) p 12
Das M K *et al* 2010 Fragmented QRS on twelve-lead electrocardiogram predicts arrhythmic events in patients with ischemic and nonischemic cardiomyopathy *Heart Rhythm* **7** 74–80
de Luna A B and Fiol-Sala M 2008 *Electrocardiography in Ischemic Heart Disease: Clinical and Imaging Correlations and Prognostic Implications* (John Wiley and Sons)
Edenbrandt L and Pahlm O 1988 Vectorcardiogram synthesized from a 12-lead ECG: superiority of the inverse Dower matrix *J. Electrocardiology* **21** 361–7
Eifling M, Razavi M and Massumi A 2011 The evaluation and management of electrical storm *Texas Heart Inst. J.* **38** 111
Erem B *et al* 2016 Extensions to a manifold learning framework for time-series analysis on dynamic manifolds in bioelectric signals *Phys. Rev.* E **93**
Erem B, Stovícek P and Brooks D H 2012 Manifold learning for analysis of low-order nonlinear dynamics in high-dimensional electrocardiographic signals *2012 9th IEEE Int. Symp. Biomedical Imaging (ISBI)* (IEEE) pp 844–7
Finocchiaro G. *et al* 2020 The electrocardiogram in the diagnosis and management of patients with dilated cardiomyopathy *Eur. J. Heart Failure* **22** 1097–107
Garcia T B 2013 *12-Lead ECG: The Art of Interpretation* (Jones and Bartlett Publishers)
Good W W *et al* 2020 Characterizing the transient electrocardiographic signature of ischemic stress using Laplacian Eigenmaps for dimensionality reduction *Comput. Biol. Med.* **127**
Good W W, Erem B, Zenger B, Coll-Font J, Brooks D H and MacLeod R S 2018 Temporal performance of laplacian eigenmaps and 3d conduction velocity in detecting ischemic stress *J. Electrocardiology* **51** S116–20
Institute of Medicine (US) Committee on Social Security Cardiovascular Disability Criteria. Cardiovascular Disability: Updating the Social Security Listings. Washington (DC): National Academies Press (US); 2010. 7, Ischemic Heart Disease. Available from: https://ncbi.nlm.nih.gov/books/NBK209964/
ISO Central Secretary. Health informatics—Standard communication protocol—Part 91064: computer-assisted electrocardiography. Standard ISO 11073-91064:2009, International Organization for Standardization, Geneva, CH (2009)
Jacobs B, Villa A, Moeyersons J, Van Huffel S, Willems R and Varon C 2020 Can Laplacian Eigenmaps be used for differentiation between healthy subjects and patients with corrected Tetralogy of Fallot? *2020 Computing in Cardiology* pp 1–4
Jaros R, Martinek R and Danys L 2019 Comparison of different electrocardiography with vectorcardiography transformations *Sensors* **19** 3072
Jefferies J L and Towbin J A 2010 Dilated cardiomyopathy *Lancet* **375** 752–62
Kors J A *et al* 1990 Reconstruction of the Frank vectorcardiogram from standard electrocardiographic leads: diagnostic comparison of different methods *Eur. Heart J.* **11** 1083–92
Kossmann C E *et al* 1967 Recommendations for standardization of leads and of specifications for instruments in electrocardiography and vectorcardiography *Circulation* **35** 583–602
Martínez J P, Almeida R, Olmos S, Rocha A P and Laguna P 2004 A wavelet-based ECG delineator: evaluation on standard databases *IEEE Trans. Biomed. Eng.* **51** 570–81
McCrohon J A *et al* 2003 Differentiation of heart failure related to dilated cardiomyopathy and coronary artery disease using gadolinium-enhanced cardiovascular magnetic resonance *Circulation* **108** 54–9
Moeyersons J, Amoni M, Van Huffel S, Willems R and Varon C 2019 R-DECO: an open-source Matlab based graphical user interface for the detection and correction of R-peaks *Peerj Comput. Sci.* **5** e226

Müller M 2007 Dynamic time warping *Information Retrieval for Music and Motion* (Heidelberg: Springer Berlin) pp 69–84
Nadour W *et al* 2014 Does the presence of Q waves on the EKG accurately predict prior myocardial infarction when compared to cardiac magnetic resonance using late gadolinium enhancement? A cross-population study of noninfarct vs infarct patients *Heart Rhythm* **11** 2018–26
Nelwan S P, Kors J A and Meij S H 2000 Minimal lead sets for reconstruction of 12-lead electrocardiograms *J. Electrocardiology* **33** 163–6
Schultheiss H P *et al* 2019 Dilated cardiomyopathy *Nat. Rev. Dis. Primers* **5** 1–19
Sha J *et al* 2011 Fragmented QRS is associated with all-cause mortality and ventricular arrhythmias in patient with idiopathic dilated cardiomyopathy *Ann. Noninvasive Electrocardiology* **16** 270–5
Takigawa M *et al* 2010 Predictors of electrical storm in patients with idiopathic dilated cardiomyopathy-how to stratify the risk of electrical storm *Circ. J.* **74** 1822–9
Tereshchenko L G *et al* 2015 Analysis of speed, curvature, planarity and frequency characteristics of heart vector movement to evaluate the electrophysiological substrate associated with ventricular tachycardia *Comput. Biol. Med.* **65** 150–60
Vandenberk B *et al* 2017 Inferior and anterior QRS fragmentation have different prognostic value in patients who received an implantable defibrillator in primary prevention of sudden cardiac death *Int. J. Cardiol.* **243** 223–8
Villa A, Ingelaere S, Van Huffel S, Willems R and Varon C 2020 Fragmented QRS dynamics towards electrical storm in ICD patients *2020 Computing in Cardiology* (IEEE) pp 1–4
Voulgari C, Pagoni S, Tesfaye S and Tentolouris N 2013 The spatial QRS-T angle: implications in clinical practice *Curr. Cardiol. Rev.* **9** 197–210
Vozda M and Cerny M 2015 Methods for derivation of orthogonal leads from 12-lead electrocardiogram: a review *Biomed. Signal Process. Control* **19** 23–34
Wagner P *et al* 2020 PTB-XL, a large publicly available electrocardiography dataset *Sci. Data* **7** 1–15
Zhang D, He F, Han S, Zou L, Wu Y and Chen Y 2017 An efficient approach to directly compute the exact Hausdorff distance for 3D point sets *Integr. Comput.-Aided Eng.* **24** 261–77