Dataset of Simulated Intracardiac Transmembrane Voltage Recordings and ECG Signals

Documentation

Mikel Landajuela[1] Rushil Anirudh[1] Robert Blake[1]

[1]Lawrence Livermore National Laboratory
June 28, 2022

Contents

1 Introduction

The dataset consists of pairs of computationally simulated intracardiac transmembrane voltage recordings and ECG signals. In total, 16140 organ-level simulations were conducted to create this dataset. Simulations were performed using *Lassen* supercomputer at Lawrence Livermore National Lab (LLNL) , concurrently utilizing 4 GPUs and 40 CPU cores. Each simulation produced pairs of 500×ms-by-10 ECG signals and 500×ms-by-75 transmembrane voltage signals. For convenience, we collect those signals in matrices $X \in \mathbb{R}^{500\times 10}$ and $V \in \mathbb{R}^{500\times 75}$, respectively. Each of these matrices is then stored as a `numpy` array. Figure 1 shows the data generation process. Figure 5 and Figure 6 show the structure of the dataset subfolders with number of files, memory size and examples of the conventions used for naming the numpy arrays. Given the matrix X, the standard 12 lead signal are obtained using the following transformation:

$$\begin{cases} X_{:,1} \rightarrow RA \\ X_{:,2} \rightarrow LA \\ X_{:,3} \rightarrow LL \\ X_{:,4} \rightarrow RL \\ X_{:,5} \rightarrow V1 \\ X_{:,6} \rightarrow V2 \\ X_{:,7} \rightarrow V3 \\ X_{:,8} \rightarrow V4 \\ X_{:,9} \rightarrow V5 \\ X_{:,10} \rightarrow V6 \end{cases} \text{and} \begin{cases} \text{Lead } I : LA - RA \\ \text{Lead } II : LL - RA \\ \text{Lead } III : LL - LA \\ \text{Lead } aVR : \frac{3}{2}(RA - Vw) \\ \text{Lead } aVL : \frac{3}{2}(LA - Vw) \\ \text{Lead } aVF : \frac{3}{2}(LL - Vw) \\ \text{Lead } V1 : V1 - Vw \\ \text{Lead } V2 : V2 - Vw \\ \text{Lead } V3 : V3 - Vw \\ \text{Lead } V4 : V4 - Vw \\ \text{Lead } V5 : V5 - Vw \\ \text{Lead } V6 : V6 - Vw \end{cases}$$

with the (Wilson Lead) $Vw : \frac{1}{3}(RA + LA + LL)$. Code to perform this transformation is provided in `notebooks/data_inspect.ipynb`. Further details on the simulation settings, such as the mathematical models, the anatomical geometries, and the parameter variations, are presented in the following section.

2 Simulation details

Cardiac simulations were carried out using Cardioid [17], a multiscale cardiac simulation code developed at Lawrence Livermore National Laboratory (LLNL). Cardioid uses a finite volume method with explicit timestep-

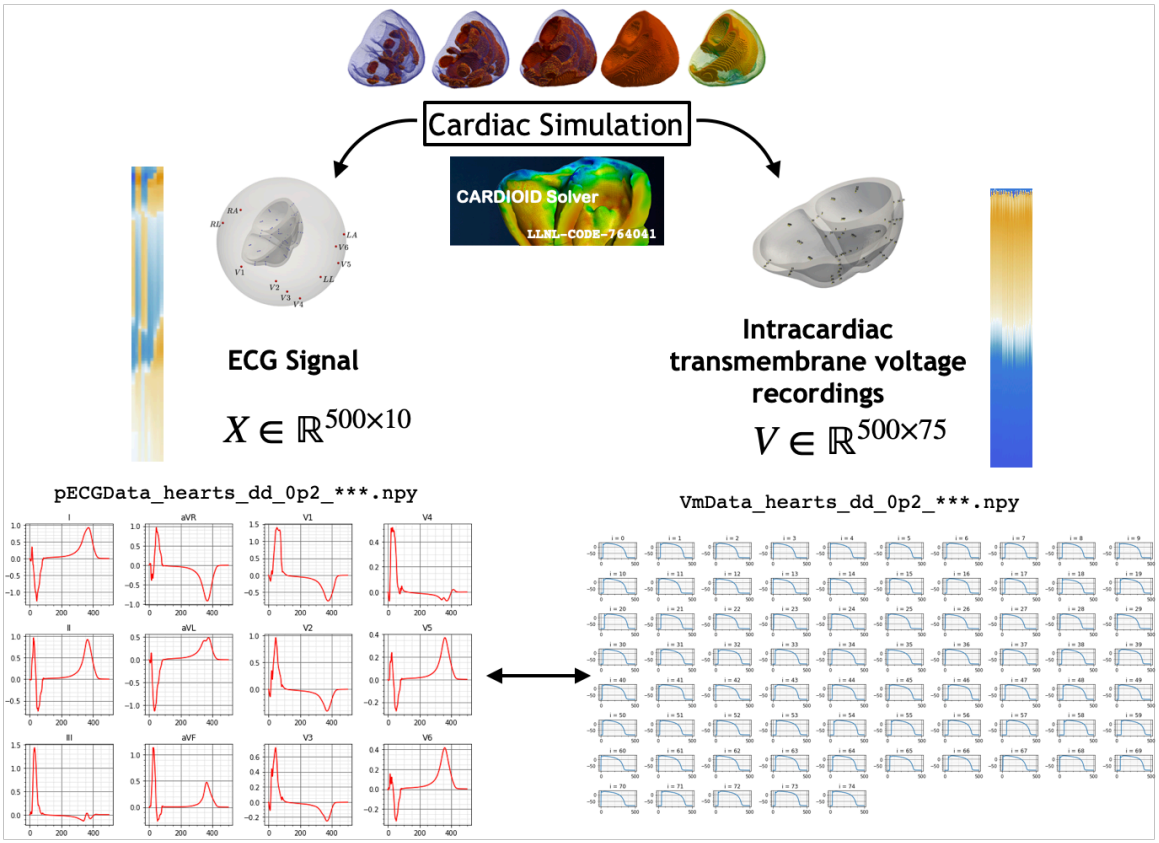


Figure 1: Illustration of the data generation process. The cardiac simulation produces 10 ECG signals and 75 transmembrane voltage signals of 500×ms duration. We collect each of those in a matrix and save them as numpy arrays.

ping to solve the monodomain model [11, 10], a system of reaction-diffusion equations that descibes the spatiotemporal evolution of the transmembrane voltage within the myocardium. These equations are coupled with cell models that describe the dynamics of ionic species flux through the cell membrane. In this work, the cell models proposed by Ten-Tusscher [19] were used for endocardial, midmyocardial, and epicardial cells.

The monodomain equations were solved in real bi-ventricular cardiac domains. The patient specific geometries were obtained from the publicly available database made available from Inria [7], in which those geometries were created using original MR images from the Stacom 2011 challenge [20]. Meshes were preprocessed to make them compatible with the Cardioid solver and resolved to a 200 μm resolution (see Figure 2). Myocardial fiber orientations were assigned based on rule-based laplacian-driven algorithms for interpolating fiber geometries in the absence of DTMRI data [2].

The high resolution simulations of the transmembrane voltages within the myocardium were used to compute the synthetic ECG signals. The simulation of realistic ECG signals using full organ models of the heart is a very active field of research [4, 12]. To reconstruct the ECG signal, a full heart-torso coupled problem can be solved for each time-step [4, 6] or, alternatively, the pseudo-ECG approach [16, 3, 8] can be followed; the latter was used in this work. The locations of the pseudo-ECG electrodes were chosen based on locations derived from an existing torso mesh and then normalized to a 100 mm radius around the center of each mesh (see Figure 3b).

The morphology of the ECG signal is sensitive to the initial endocardial stimulus [6]. In this work, activation patterns were extracted from the literature [9, 5]. In order to retrieve physiological T wave morphology in the ECG signals, apex-to-base action potential duration (APD) heterogeneity [12], and transmural APD heterogeneity [19] were included within the ionic models. In addition, heart conductivities were changed near endocardial and epicardial surfaces to account for the bath loading effects over the surfaces of the heart [3].

For the recording of the transmembrane voltages within the myocardium, 30 points were selected by hand for each mesh — 17 endocardial points were selected in the left ventricle (LV), corresponding to standard AHA17 segment locations [15], and 13 points were selected in the right ventricle (RV), according to an equivalent segmentation of that ventricle [21]. See Figure 3a for a schematic representation of the points over a Bull's-eye display of the heart. From these 30 points, 20 exterior wall points were programmatically identified based on minimum distances from the hand-selected endocardial points, and 25 mid-myocardial points were then found through interpolation. For each simulation performed, simulated transmembrane voltages were recorded for each of the 75 epicardial/midmyocardial/endocardial points.

Figure 2: Bi-ventricular cardiac geometries.

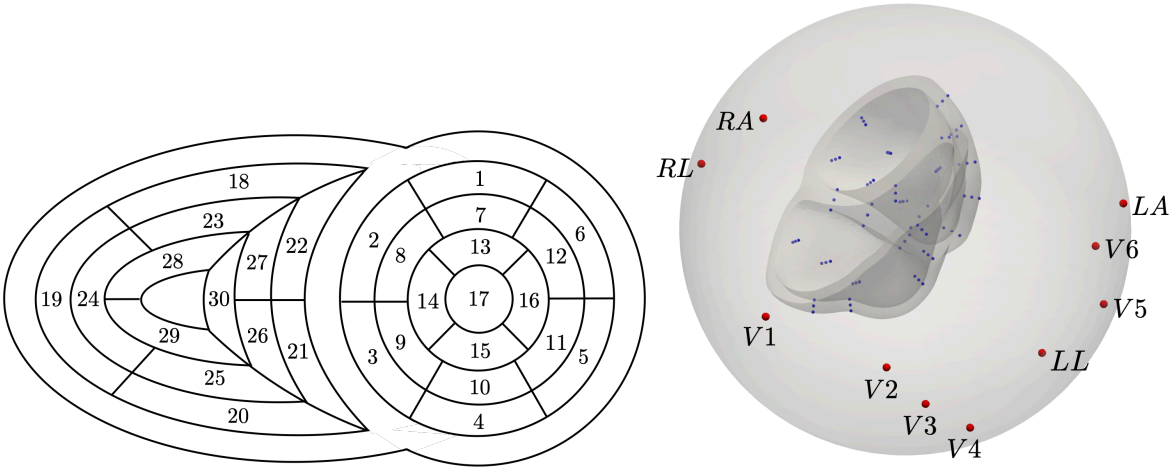
Human ventricular activation and ECG data exhibit a high level of morphological variability depending on physiological and pathophysiological factors [1]. In order to reproduce this variability and enrich the dataset of activation-ECG pairs, the following combinations of parameters were explored:

- A suite of 15 clinical bi-ventricular geometries was considered to model inter-subject variability in anatomical morphology and ventricular thickness. See Figure 2.
- A library of 29 clinically-inspired activation patterns was designed to account for variations across patient Purkinje systems. For each activation pattern, activation times over the previously identified 30 endocardial points were chosen according to the distributions presented in the literature [9, 5].
- A set of 3 different combinations of tissue conductivities (longitudinal, traversal and normal) was considered. Starting from the original values reported in [14], the library was designed to obtain 3:1, 4:1 and 6:1 ratios between longitudinal and transverse conduction velocities. These parameters were tested and exhibited conduction velocities that where within the physiological ranges reported in literature [10].
- In addition to the original value of the maximal conductance G_{Kr} of the rapid delayed rectifier current proposed in [19] for the cell model, two variations were considered: 0% (blocked) and 50% original value. This parameter is known to affects APD [13, 18].
- Basic Cycle Lengths (BCL) of 600 ms and 1000 ms were considered to produce different initial states for the cell models. The choice of initial BCL states affect APD [18].
- Randomized samples over the space of inner activation points (75 points inside the myocardium) were considered to capture early activation and pathological scenarios.

All simulations were performed for 500 ms of simulation time with 200 μm resolution meshes and a time-step of 5 μs. The ECG and transmembrane voltages were recorded at a resolution of 1 ms. All variations in the cell models parameters (apex-to-base and transmural APD heterogeneity, G_{Kr} and BCL) were pre-simulated with 100 beats in a single cell simulation in order to reach dynamic steady-state.

3 Data inspection

The data consists of simple pairs on `numpy` arrays. It can be loaded and displayed using the code below to produce Figure 4. In the repository, we provide Jupyter and Mathematica Notebooks, `notebooks/data_inspect.ipynb` and `notebooks/data_inspect.nb` respectively, to inspect the data.



(a) (b)

Figure 3: Recording points. (3a) Location of 30 manually selected endocardial points. (3b) Location of the pECG electrodes (red) and transmembrane voltage recording points (blue).

```
import numpy
import matplotlib.pyplot as plt
pECG = np.load(pECGData_hearts_dd_0p2_volunteer.v10_pattern.0.npy)
Vm = np.load(VmData_hearts_dd_0p2_volunteer.v10_pattern.0.npy)
plt.figure(figsize=(20, 5))
plt.subplot(1,2,1)
plt.plot(pECG)
plt.subplot(1,2,2)
plt.plot(Vm)
plt.show()
```

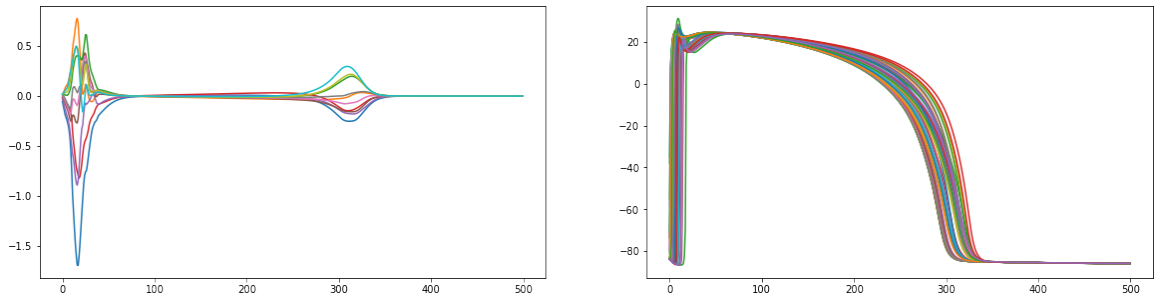


Figure 4: Example of processed numpy arrays.

4 Aacronyms and abbreviations

Here, we describe the aacronyms and abbreviations used in table headings or other data descriptors. The name conventions for the `numpy` arrays are

{`pECGData_`,`VmData_`} + `subfolder_name_` + {`volunteer.#`,`pattern.#`,`bcl.#`,`gkr.#`,`innerindex.#`}

with

File Flag	Description	Range for #
`volunteer.#`	Cardiac geometry	{v1,v2,v4,v5,v6,v7,v8,v9,v10,v11,v12,v13,v14,v15,v16}
`pattern.#`	Activation pattern	$0, 1, 2, \ldots, 28$
`bcl.#`	Basic cycle lengths	{600,1000}
`gkr.#`	Variation on G_{Kr}	{000,050}
`innerindex.#`	Inner activation point	$0, 1, 2, \ldots, 74$

The structure of the dataset subfolders with global parameters, number of files and memory size is presented in Figure 5. Examples of the name conventions for the `numpy` arrays inside each subfolder is shown in Figure 6.

5 Acknowledgments

We thank Livermore Computing (LC) at LLNL for the computational resources that enabled this work. We thank the Laboratory Directed Research and Development (LDRD) program at LLNL under the National Nuclear Security Administration (NNSA) for their support of this work, project number 18-LW-078. This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. LLNL-JRNL-799842.

References

[1] Mary G Adams-Hamoda, Mary A Caldwell, Nancy A Stotts, and Barbara J Drew. Factors to consider when analyzing 12-lead electrocardiograms for evidence of acute myocardial ischemia. *American Journal of Critical Care*, 12(1):9–16, 2003.

[2] Jason D Bayer, Robert C Blake, Gernot Plank, and Natalia A Trayanova. A novel rule-based algorithm for assigning myocardial fiber orientation to computational heart models. *Annals of biomedical engineering*, 40(10):2243–2254, 2012.

[3] Martin J Bishop and Gernot Plank. Bidomain ecg simulations using an augmented monodomain model for the cardiac source. *IEEE transactions on biomedical engineering*, 58(8):2297–2307, 2011.

[4] Muriel Boulakia, Serge Cazeau, Miguel A Fernández, Jean-Frédéric Gerbeau, and Nejib Zemzemi. Mathematical modeling of electrocardiograms: a numerical study. *Annals of biomedical engineering*, 38(3):1071–1097, 2010.

[5] Louie Cardone-Noott, Alfonso Bueno-Orovio, Ana Minchole, Nejib Zemzemi, and Blanca Rodriguez. Human ventricular activation sequence and the simulation of the electrocardiographic qrs complex and its variability in healthy and intraventricular block conditions. *EP Europace*, 18(suppl_4):iv4–iv15, 2016.

[6] Jonathan P Cranford, Thomas J O'Hara, Christopher T Villongco, Omar M Hafez, Robert C Blake, Joseph Loscalzo, Jean-Luc Fattebert, David F Richards, Xiaohua Zhang, James N Glosli, et al. Efficient computational modeling of human ventricular activation and its electrocardiographic representation: A sensitivity study. *Cardiovascular engineering and technology*, 9(3):447–467, 2018.

[7] Nicolas Duchateau, Maxime Sermesant, Hervé Delingette, and Nicholas Ayache. Model-based generation of large databases of cardiac images: synthesis of pathological cine mr sequences from real healthy cases. *IEEE transactions on medical imaging*, 37(3):755–766, 2017.

[8] Marie Dupraz, Simonetta Filippi, Alessio Gizzi, Alfio Quarteroni, and Ricardo Ruiz-Baier. Finite element and finite volume-element simulation of pseudo-ecgs and cardiac alternans. *Mathematical Methods in the Applied Sciences*, 38(6):1046–1058, 2015.

[9] Dirk Durrer, R Th Van Dam, GE Freud, MJ Janse, FL Meijler, and RC Arzbaecher. Total excitation of the isolated human heart. *Circulation*, 41(6):899–912, 1970.

[10] Piero Colli Franzone, Luca Franco Pavarino, and Simone Scacchi. *Mathematical cardiac electrophysiology*, volume 13. Springer, 2014.

[11] J Keener and J Sneyd. *Mathematical Physiology: II: Systems Physiology. 2009*. Springer, 2009.

[12] Shankarjee Krishnamoorthi, Luigi E Perotti, Nils P Borgstrom, Olujimi A Ajijola, Anna Frid, Aditya V Ponnaluri, James N Weiss, Zhilin Qu, William S Klug, Daniel B Ennis, et al. Simulation methods and validation criteria for modeling cardiac ventricular electrophysiology. *PloS one*, 9(12):e114494, 2014.

[13] Gui-Rong Li, Jianlin Feng, Lixia Yue, Michel Carrier, and Stanley Nattel. Evidence for two components of delayed rectifier k+ current in human ventricular myocytes. *Circulation research*, 78(4):689–696, 1996.

[14] Steven A Niederer, Eric Kerfoot, Alan P Benson, Miguel O Bernabeu, Olivier Bernus, Chris Bradley, Elizabeth M Cherry, Richard Clayton, Flavio H Fenton, Alan Garny, et al. Verification of cardiac tissue electrophysiology simulators using an n-version benchmark. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 369(1954):4331–4351, 2011.

[15] American Heart Association Writing Group on Myocardial Segmentation, Registration for Cardiac Imaging:, Manuel D Cerqueira, Neil J Weissman, Vasken Dilsizian, Alice K Jacobs, Sanjiv Kaul, Warren K Laskey, Dudley J Pennell, John A Rumberger, Thomas Ryan, et al. Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart: a statement for healthcare professionals from the cardiac imaging committee of the council on clinical cardiology of the american heart association. *Circulation*, 105(4):539–542, 2002.

[16] Robert Plonsey and Roger C Barr. *Bioelectricity: a quantitative approach*. Springer Science & Business Media, 2007.

[17] David F Richards, James N Glosli, Erik W Draeger, Arthur A Mirin, Bor Chan, Jean-luc Fattebert, William D Krauss, Tomas Oppelstrup, Chris J Butler, John A Gunnels, et al. Towards real-time simulation of cardiac electrophysiology in a human heart at high resolution. *Computer methods in biomechanics and biomedical engineering*, 16(7):802–805, 2013.

[18] KHWJ Ten Tusscher, Denis Noble, Peter-John Noble, and Alexander V Panfilov. A model for human ventricular tissue. *American Journal of Physiology-Heart and Circulatory Physiology*, 286(4):H1573–H1589, 2004.

[19] Kirsten HWJ Ten Tusscher and Alexander V Panfilov. Alternans and spiral breakup in a human ventricular tissue model. *American Journal of Physiology-Heart and Circulatory Physiology*, 291(3):H1088–H1100, 2006.

[20] Catalina Tobon-Gomez, Mathieu De Craene, Kristin Mcleod, Lennart Tautz, Wenzhe Shi, Anja Hennemuth, Adityo Prakosa, Hengui Wang, Gerry Carr-White, Stam Kapetanakis, et al. Benchmarking framework for myocardial tracking and deformation algorithms: An open access database. *Medical image analysis*, 17(6):632–648, 2013.

[21] Liang Zhong, Like Gobeawan, Yi Su, Ju-Le Tan, Dhanjoo Ghista, Terrance Chua, Ru-San Tan, and Ghassan Kassab. Right ventricular regional wall curvedness and area strain in patients with repaired tetralogy of fallot. *American Journal of Physiology-Heart and Circulatory Physiology*, 302(6):H1306–H1316, 2011.

Name	Global parameters	Number_files	Size
data_hearts_dd_0p2	endocardial activation {gL ,gT,gN} = {0.133418, 0.035212, 0.035212} gkr is original value (0.153)	870	148.011 MB
data_hearts_dd_0p2_geo_act_1_bcl	endocardial activation {gL ,gT,gN} = {0.133418, 0.035212, 0.035212} bcl ∈ {1000, 600} gkr is original value (0.153)	1740	296.023 MB
data_hearts_dd_0p2_geo_act_2_bcl	endocardial activation {gL ,gT,gN} = {0.133418, 0.017606, 0.017606} bcl ∈ {1000, 600} gkr is original value (0.153)	1740	296.023 MB
data_hearts_dd_0p2_geo_act_3_bcl	endocardial activation {gL ,gT,gN} = {0.133418, 0.008803, 0.008803} bcl ∈ {1000, 600} gkr is original value (0.153)	1740	296.023 MB
data_hearts_dd_0p2_geo_act_1_bcl_gkr	endocardial activation {gL ,gT,gN} = {0.133418, 0.035212, 0.035212} bcl ∈ {1000, 600} gkrs ∈ [{0,0.5} ∗ original value (0.153)]	1398	237.839 MB
data_hearts_dd_0p2_geo_act_2_bcl_gkr	endocardial activation {gL ,gT,gN} = {0.133418, 0.017606, 0.017606} bcl ∈ {1000, 600} gkrs ∈ [{0,0.5} ∗ original value (0.153)]	1400	238.179 MB
data_hearts_dd_0p2_geo_act_3_bcl_gkr	endocardial activation {gL ,gT,gN} = {0.133418, 0.008803, 0.008803} bcl ∈ {1000, 600} gkrs ∈ [{0,0.5} ∗ original value (0.153)]	1398	237.839 MB
data_hearts_dd_0p2_geo_act_1_bcl_gkr_I	endocardial activation {gL ,gT,gN} = {0.133418, 0.035212, 0.035212} bcl ∈ {1000, 600} gkrs ∈ [{0,0.5} ∗ original value (0.153)] (The rest of simulations)	2080	353.866 MB
data_hearts_dd_0p2_geo_act_2_bcl_gkr_I	endocardial activation {gL ,gT,gN} = {0.133418, 0.017606, 0.017606} bcl ∈ {1000, 600} gkrs ∈ [{0,0.5} ∗ original value (0.153)] (The rest of simulations)	2080	353.866 MB
data_hearts_dd_0p2_geo_act_3_bcl_gkr_I	endocardial activation {gL ,gT,gN} = {0.133418, 0.008803, 0.008803} bcl ∈ {1000, 600} gkrs ∈ [{0,0.5} ∗ original value (0.153)] (The rest of simulations)	2080	353.866 MB
data_hearts_dd_0p2_geo_inn	inner activation points {gL ,gT,gN} = {0.133418, 0.035212, 0.035212} gkr is original value (0.153)	2250	382.788 MB
data_hearts_dd_0p2_geo_inn_act_1_bcl	inner activation points {gL ,gT,gN} = {0.133418, 0.035212, 0.035212} bcl ∈ {1000, 600} gkr is original value (0.153)	900	153.115 MB
data_hearts_dd_0p2_geo_inn_act_2_bcl	inner activation points {gL ,gT,gN} = {0.133418, 0.017606, 0.017606} bcl ∈ {1000, 600} gkr is original value (0.153)	900	153.115 MB
data_hearts_dd_0p2_geo_inn_act_3_bcl	inner activation points {gL ,gT,gN} = {0.133418, 0.008803, 0.008803} bcl ∈ {1000, 600} gkr is original value (0.153)	900	153.115 MB
data_hearts_dd_0p2_geo_inn_act_1_bcl_I	inner activation points {gL ,gT,gN} = {0.133418, 0.035212, 0.035212} bcl ∈ {1000, 600} gkr is original value (0.153) (The rest of simulations)	3562	605.996 MB
data_hearts_dd_0p2_geo_inn_act_2_bcl_I	inner activation points {gL ,gT,gN} = {0.133418, 0.017606, 0.017606} bcl ∈ {1000, 600} gkr is original value (0.153) (The rest of simulations)	3598	612.121 MB
data_hearts_dd_0p2_geo_inn_act_3_bcl_I	inner activation points {gL ,gT,gN} = {0.133418, 0.008803, 0.008803} bcl ∈ {1000, 600} gkr is original value (0.153) (The rest of simulations)	3598	612.121 MB

Figure 5: Name of dataset subfolders with golbal parameters, number of files and memory size.

Name	(X,V) pairs
data_hearts_dd_0p2	pECGData_hearts_dd_0p2_volunteer.v10_pattern.0.npy ⟶ VmData_hearts_dd_0p2_volunteer.v10_pattern.0.npy ⋮ pECGData_hearts_dd_0p2_volunteer.v9_pattern.9.npy ⟶ VmData_hearts_dd_0p2_volunteer.v9_pattern.9.npy
data_hearts_dd_0p2_geo_act_1_bcl	pECGData_hearts_dd_0p2_geo_act_1_bcl_bcl.1000.pattern.0.volunteer.v10.npy ⟶ VmData_hearts_dd_0p2_geo_act_1_bcl_bcl.1000.pattern.0.volunteer.v10.npy ⋮ pECGData_hearts_dd_0p2_geo_act_1_bcl_bcl.600.pattern.9.volunteer.v9.npy ⟶ VmData_hearts_dd_0p2_geo_act_1_bcl_bcl.600.pattern.9.volunteer.v9.npy
data_hearts_dd_0p2_geo_act_2_bcl	pECGData_hearts_dd_0p2_geo_act_2_bcl_bcl.1000.pattern.0.volunteer.v10.npy ⟶ VmData_hearts_dd_0p2_geo_act_2_bcl_bcl.1000.pattern.0.volunteer.v10.npy ⋮ pECGData_hearts_dd_0p2_geo_act_2_bcl_bcl.600.pattern.9.volunteer.v9.npy ⟶ VmData_hearts_dd_0p2_geo_act_2_bcl_bcl.600.pattern.9.volunteer.v9.npy
data_hearts_dd_0p2_geo_act_3_bcl	pECGData_hearts_dd_0p2_geo_act_3_bcl_bcl.1000.pattern.0.volunteer.v10.npy ⟶ VmData_hearts_dd_0p2_geo_act_3_bcl_bcl.1000.pattern.0.volunteer.v10.npy ⋮ pECGData_hearts_dd_0p2_geo_act_3_bcl_bcl.600.pattern.9.volunteer.v9.npy ⟶ VmData_hearts_dd_0p2_geo_act_3_bcl_bcl.600.pattern.9.volunteer.v9.npy
data_hearts_dd_0p2_geo_act_1_bcl_gkr	pECGData_hearts_dd_0p2_geo_act_1_bcl_gkr_bcl.1000.gkr.000.pattern.0.volunteer.v12.npy ⟶ VmData_hearts_dd_0p2_geo_act_1_bcl_gkr_bcl.1000.gkr.000.pattern.0.volunteer.v12.npy ⋮ pECGData_hearts_dd_0p2_geo_act_1_bcl_gkr_bcl.600.gkr.050.pattern.9.volunteer.v14.npy ⟶ VmData_hearts_dd_0p2_geo_act_1_bcl_gkr_bcl.600.gkr.050.pattern.9.volunteer.v14.npy
data_hearts_dd_0p2_geo_act_2_bcl_gkr	pECGData_hearts_dd_0p2_geo_act_2_bcl_gkr_bcl.1000.gkr.000.pattern.0.volunteer.v12.npy ⟶ VmData_hearts_dd_0p2_geo_act_2_bcl_gkr_bcl.1000.gkr.000.pattern.0.volunteer.v12.npy ⋮ pECGData_hearts_dd_0p2_geo_act_2_bcl_gkr_bcl.600.gkr.050.pattern.9.volunteer.v14.npy ⟶ VmData_hearts_dd_0p2_geo_act_2_bcl_gkr_bcl.600.gkr.050.pattern.9.volunteer.v14.npy
data_hearts_dd_0p2_geo_act_3_bcl_gkr	pECGData_hearts_dd_0p2_geo_act_3_bcl_gkr_bcl.1000.gkr.000.pattern.0.volunteer.v12.npy ⟶ VmData_hearts_dd_0p2_geo_act_3_bcl_gkr_bcl.1000.gkr.000.pattern.0.volunteer.v12.npy ⋮ pECGData_hearts_dd_0p2_geo_act_3_bcl_gkr_bcl.600.gkr.050.pattern.9.volunteer.v14.npy ⟶ VmData_hearts_dd_0p2_geo_act_3_bcl_gkr_bcl.600.gkr.050.pattern.9.volunteer.v14.npy
data_hearts_dd_0p2_geo_act_1_bcl_gkr_I	pECGData_hearts_dd_0p2_geo_act_1_bcl_gkr_I_bcl.1000.gkr.000.pattern.0.volunteer.v10.npy ⟶ VmData_hearts_dd_0p2_geo_act_1_bcl_gkr_I_bcl.1000.gkr.000.pattern.0.volunteer.v10.npy ⋮ pECGData_hearts_dd_0p2_geo_act_1_bcl_gkr_I_bcl.600.gkr.050.pattern.9.volunteer.v9.npy ⟶ VmData_hearts_dd_0p2_geo_act_1_bcl_gkr_I_bcl.600.gkr.050.pattern.9.volunteer.v9.npy
data_hearts_dd_0p2_geo_act_2_bcl_gkr_I	pECGData_hearts_dd_0p2_geo_act_2_bcl_gkr_I_bcl.1000.gkr.000.pattern.0.volunteer.v10.npy ⟶ VmData_hearts_dd_0p2_geo_act_2_bcl_gkr_I_bcl.1000.gkr.000.pattern.0.volunteer.v10.npy ⋮ pECGData_hearts_dd_0p2_geo_act_2_bcl_gkr_I_bcl.600.gkr.050.pattern.9.volunteer.v9.npy ⟶ VmData_hearts_dd_0p2_geo_act_2_bcl_gkr_I_bcl.600.gkr.050.pattern.9.volunteer.v9.npy
data_hearts_dd_0p2_geo_act_3_bcl_gkr_I	pECGData_hearts_dd_0p2_geo_act_3_bcl_gkr_I_bcl.1000.gkr.000.pattern.0.volunteer.v10.npy ⟶ VmData_hearts_dd_0p2_geo_act_3_bcl_gkr_I_bcl.1000.gkr.000.pattern.0.volunteer.v10.npy ⋮ pECGData_hearts_dd_0p2_geo_act_3_bcl_gkr_I_bcl.600.gkr.050.pattern.9.volunteer.v9.npy ⟶ VmData_hearts_dd_0p2_geo_act_3_bcl_gkr_I_bcl.600.gkr.050.pattern.9.volunteer.v9.npy
data_hearts_dd_0p2_geo_inn	pECGData_hearts_dd_0p2_geo_inn_innerindex.0.volunteer.v10.npy ⟶ VmData_hearts_dd_0p2_geo_inn_innerindex.0.volunteer.v10.npy ⋮ pECGData_hearts_dd_0p2_geo_inn_innerindex.9.volunteer.v9.npy ⟶ VmData_hearts_dd_0p2_geo_inn_innerindex.9.volunteer.v9.npy
data_hearts_dd_0p2_geo_inn_act_1_bcl	pECGData_hearts_dd_0p2_geo_inn_act_1_bcl_bcl.1000.innerindex.0.volunteer.v13.npy ⟶ VmData_hearts_dd_0p2_geo_inn_act_1_bcl_bcl.1000.innerindex.0.volunteer.v13.npy ⋮ pECGData_hearts_dd_0p2_geo_inn_act_1_bcl_bcl.600.innerindex.9.volunteer.v4.npy ⟶ VmData_hearts_dd_0p2_geo_inn_act_1_bcl_bcl.600.innerindex.9.volunteer.v4.npy
data_hearts_dd_0p2_geo_inn_act_2_bcl	pECGData_hearts_dd_0p2_geo_inn_act_2_bcl_bcl.1000.innerindex.0.volunteer.v13.npy ⟶ VmData_hearts_dd_0p2_geo_inn_act_2_bcl_bcl.1000.innerindex.0.volunteer.v13.npy ⋮ pECGData_hearts_dd_0p2_geo_inn_act_2_bcl_bcl.600.innerindex.9.volunteer.v4.npy ⟶ VmData_hearts_dd_0p2_geo_inn_act_2_bcl_bcl.600.innerindex.9.volunteer.v4.npy
data_hearts_dd_0p2_geo_inn_act_3_bcl	pECGData_hearts_dd_0p2_geo_inn_act_3_bcl_bcl.1000.innerindex.0.volunteer.v13.npy ⟶ VmData_hearts_dd_0p2_geo_inn_act_3_bcl_bcl.1000.innerindex.0.volunteer.v13.npy ⋮ pECGData_hearts_dd_0p2_geo_inn_act_3_bcl_bcl.600.innerindex.9.volunteer.v4.npy ⟶ VmData_hearts_dd_0p2_geo_inn_act_3_bcl_bcl.600.innerindex.9.volunteer.v4.npy
data_hearts_dd_0p2_geo_inn_act_1_bcl_I	pECGData_hearts_dd_0p2_geo_inn_act_1_bcl_I_bcl.1000.innerindex.0.volunteer.v10.npy ⟶ VmData_hearts_dd_0p2_geo_inn_act_1_bcl_I_bcl.1000.innerindex.0.volunteer.v10.npy ⋮ pECGData_hearts_dd_0p2_geo_inn_act_1_bcl_I_bcl.600.innerindex.9.volunteer.v9.npy ⟶ VmData_hearts_dd_0p2_geo_inn_act_1_bcl_I_bcl.600.innerindex.9.volunteer.v9.npy
data_hearts_dd_0p2_geo_inn_act_2_bcl_I	pECGData_hearts_dd_0p2_geo_inn_act_2_bcl_I_bcl.1000.innerindex.0.volunteer.v10.npy ⟶ VmData_hearts_dd_0p2_geo_inn_act_2_bcl_I_bcl.1000.innerindex.0.volunteer.v10.npy ⋮ pECGData_hearts_dd_0p2_geo_inn_act_2_bcl_I_bcl.600.innerindex.9.volunteer.v9.npy ⟶ VmData_hearts_dd_0p2_geo_inn_act_2_bcl_I_bcl.600.innerindex.9.volunteer.v9.npy
data_hearts_dd_0p2_geo_inn_act_3_bcl_I	pECGData_hearts_dd_0p2_geo_inn_act_3_bcl_I_bcl.1000.innerindex.0.volunteer.v10.npy ⟶ VmData_hearts_dd_0p2_geo_inn_act_3_bcl_I_bcl.1000.innerindex.0.volunteer.v10.npy ⋮ pECGData_hearts_dd_0p2_geo_inn_act_3_bcl_I_bcl.600.innerindex.9.volunteer.v9.npy ⟶ VmData_hearts_dd_0p2_geo_inn_act_3_bcl_I_bcl.600.innerindex.9.volunteer.v9.npy

Figure 6: Name of dataset subfolders with examples of numpy arrays (for name convention clarification).