Generation of Realistic Synthetic Electrograms for Atrial Fibrillation Analysis

Rafael Costa de Almeida[1], Douglas Almonfrey[1], Fabien Squara[2], Vicente Zarzoso[3]

[1] Instituto Federal do Espírito Santo, Vitória, Espírito Santo, Brazil
[2] Hôpital Pasteur, Nice, France
[3] Université Côte d'Azur, CNRS, I3S Laboratory, Sophia Antipolis, France

Abstract

This work presents a pipeline for generating realistic synthetic electrograms (EGMs) simulating sustained atrial fibrillation (AF) using openCARP. Patient-specific cardiac meshes integrating fibrosis and fiber orientation, extracted from Late Gadolinium Enhancement Magnetic Resonance Imaging (LGE-MRI), enable high-fidelity datasets for downstream tasks, such as diagnostic and machine learning applications, complementing real-world clinical data. A multi-resolution protocol guarantees simulation stability, using the Courtemanche cell model with ionic modifications to replicate AF electrophysiology. Sustained AF is induced through multiple Archimedean spiral wavefronts, generating persistent reentrant activity. A virtual catheter randomly positioned on the atrial surface collects synthetic bipolar EGMs. Morphological features consistent with clinical AF are observed. In a blinded evaluation with five expert electrophysiologists, two experts could not reliably distinguish real from synthetic EGMs, while three others detected patterns suggestive of synthetic origin. These results indicate that, although synthetic data morphology aligns with clinical data, specific patterns warrant further refinement to enhance realism. This framework provides a reproducible methodology that addresses data scarcity and supports the development of personalized AF treatment and EGM analysis tools.

1. Introduction

Atrial fibrillation (AF) is the most prevalent sustained cardiac arrhythmia and a major contributor to stroke, heart failure, and a wide spectrum of secondary complications [1, 2]. Despite extensive clinical and experimental research, the condition remains difficult to manage effectively, as the underlying electrophysiological mechanisms are complex and heterogeneous across patients [3].

Electroanatomic mapping is still regarded as the clinical gold standard for identifying arrhythmogenic substrates and guiding ablation, but its utility is limited by its inherently invasive nature, equipment requirements, and data scarcity [4]. Computational modeling and digital twin approaches, informed by LGE-MRI for fibrosis and fiber orientation, provide new opportunities for personalized medicine and virtual trials [5, 6]. Synthetic data enables controlled exploration of AF mechanisms, including varying fibrosis patterns and stimulation protocols, and supports systematic evaluation of therapeutic targets. Realistic synthetic EGMs further reveal the relationship between fibrotic patterns and reentrant circuits [7].

Building on prior work in AF simulation and EGM generation [3, 8], this study introduces two key methodological enhancements to improve efficiency and realism. First, the stimulation and initialization protocol employs multiple small Archimedean spirals to induce AF and incorporates a multi-resolution strategy: low-resolution pseudo-bidomain simulations rapidly assess whether AF can be sustained over time, while high-resolution bidomain simulations are subsequently used for accurate synthetic EGM acquisition. Second, the virtual catheter acquisition extends previous approaches by modeling electrode dimensions and placement in full 3D space while incorporating dynamic tissue-contact variability, thereby reproducing amplitude fluctuations and signal loss observed in clinical recordings. Together, these developments yield a reproducible framework for generating personalized synthetic EGMs that complement real patient data for research and therapeutic development.

2. Methods

Synthetic EGMs are generated by simulating atrial electrophysiological behavior using the openCARP platform and patient-derived anatomical data from LGE-MRI. The pipeline induces sustained AF and records virtual EGMs using a simulated catheter for subsequent expert evaluation.

Electrical propagation is modeled using reaction-diffusion formulations for cardiac tissue, following established references [9, 10]. We adopt a computationally efficient pseudo-bidomain strategy, which combines an augmented monodomain formulation to reproduce bidomain-

like wavefront morphologies [11].

2.1. Virtual Catheter and EGM Acquisition

To simulate EGM recordings, a method that accounts for electrode geometry and dynamic tissue contact in a 3D tissue is employed, inspired by previous work [12]. The potential Φ_e at an electrode is computed using the transmembrane current I_m. Using a discrete approximation with N tissue points, the unipolar EGM is given by:

$$\Phi_e(t) = \frac{1}{4\pi\sigma_e} \sum_{i=1}^{N} R(\|p_i - e(t)\|) \cdot I_m(i,t) \quad (1)$$

Here, σ_e is the extracellular conductivity, p_i is the position of the i-th tissue point, $e(t)$ is the electrode position at time t, and $R(\cdot)$ is the tissue to electrode transfer function.

Dynamic Electrode-Tissue Contact and 3D Position: To capture dynamic contact variations observed in clinical recordings, we model the electrode height above the tissue as a time-varying function, $h(t)$. The height is generated as:

$$h(t) = h_{\text{base}} + \sum_{k=1}^{K} A_k \sin(2\pi f_k t + \phi_k) + \eta(t) \quad (2)$$

where h_{base} is the baseline height, A_k, f_k, and ϕ_k are the amplitude, frequency, and phase of each sinusoidal component, and $\eta(t)$ is a noise term. The electrode's position at time t, $e(t)$, is calculated by displacing a position e_0 along the local tissue normal vector, $\hat{\mathbf{n}}$:

$$e(t) = \mathbf{e_0} + \mathbf{z}_{\text{max}} \cdot \mathbf{h(t)} \cdot \hat{\mathbf{n}} \quad (3)$$

where z_{max} is the maximum displacement height.

Electrode Transfer Function: The transfer function $R(\cdot)$ for an electrode with diameter d_0 accounts for both its physical characteristics and the time-varying height:

$$R(r) = 2\arcsin\left(\frac{d_0}{D(t)}\right) \quad (4)$$

where $r = \|p_i - e(t)\|$ is the Euclidean distance from a tissue point to the electrode, and the denominator $D(t)$ is defined as:

$$D(t) = \sqrt{(r - \frac{d_0}{2})^2 + z(t)^2} + \sqrt{(r + \frac{d_0}{2})^2 + z(t)^2} \quad (5)$$

and $z(t) = z_{\text{max}} \cdot h(t)$ is the time-varying height of the electrode.

This approach extends previous models by incorporating the full 3D geometry of both tissue and catheter while explicitly modeling dynamic contact.

2.2. Real Patient Data Modeling

Heart meshes derived from LGE-MRI provide the anatomical substrate, with fibrosis obtained via Image Intensity Ratio (IIR) and fiber orientation mappings. Electrophysiological behavior follows the Courtemanche atrial cell model with AF-remodeling parameters [3, 13]. IIR thresholds (e.g., IIR > 0.64) assign AF-remodeled properties, producing spatial heterogeneity in conduction [8]. We vary fibrosis intensity and ionic parameters to modulate conduction velocity and wavefront fragmentation, yielding diverse synthetic tissue patterns.

We leverage two sources of clinical data. For in-silico simulation inputs, we use patient-specific 3D meshes from 100 AF patients (paroxysmal 43, persistent 41, long-standing persistent 16) with fibers and fibrosis from LGE-MRI [8]. For validation, we use catheter EGMs from 53 patients with persistent AF acquired with the *PentaRay*® multielectrode mapping catheter at Nice Pasteur University. This anonymized dataset is an expansion of the dataset from 16 patients with persistent AF [14], containing over 10,000 samples of EGMs recorded before ablation procedures and is used during the blind-test evaluation.

2.3. Simulation Protocol

Persistent AF is simulated by defining electrical activity generated by the atrial tissue itself using a cell voltage map.

The Universal Atrial Coordinates system [8] enables the definition of electrical wavefronts by projecting them into 2D, where initial stimuli are defined and mapped back to 3D. This facilitates the creation of activation patterns of rotors and reentrant activity [13].

A version, inspired by the literature [8], employing eleven small radius Archimedean spirals with opposing rotations defined in 2D is used to induce sustained AF, generating rotors leading to persistent reentrant activity. By controlling their radius and positions, variants of persistent AF propagation can be created, increasing data variability.

A multi-resolution protocol is used. Once the parameters are chosen, an activation map with Archimedean spirals is created. Then, a low-resolution pseudo-bidomain simulation runs for 5 cycle lengths (CL) to assess AF stability. The activation map during the last cycle is used to determine if there is still activation in the mesh. If no activation is present, parameters are discarded as they are not suitable to maintain a heart rhythm. If sustained activation is confirmed, a high-resolution bidomain simulation is performed for 2 CL, ensuring the activation patterns no longer resemble the spirals, followed by 10 CL for synthetic EGM acquisition.

After the simulation, a synthetic PentaRay catheter is randomly placed on the atrial surface to record bipolar EGMs using the acquisition method described in section

2.1, where variations in catheter-tissue contact are modeled, allowing collected samples to contain patterns in clinical recordings.

2.4. Evaluation by Expert Electrophysiologists

To assess the realism of the synthetic EGMs generated by the framework, a blind-test evaluation is conducted with five expert electrophysiologists. A total of 20 samples were presented per expert (10 real and 10 synthetic), the same 20 samples were shown to each expert, each sample consisting of a PentaRay multi-lead set with 10 bipolar EGMs. Samples are randomly selected from the pool of real and synthetic samples. For real samples, we filtered the dataset for samples classified as AF. Each expert classified each sample as either "Real" or "Synthetic" alongside a justification for the classification. The objective was to determine if experts could distinguish between real and synthetic EGMs and gather feedback on the reasons for their classifications. The panel comprised experienced electrophysiologists, all of them with more than five years in practice, two with more than eight years, two with more than fifteen years, and one with more than five years.

3. Results

The framework successfully generated a dataset of annotated EGM samples from 37 patient-specific meshes.

Figure 1 provides a visual overview of key simulation components. Panel (a) shows the IIR-based fibrosis distribution on a representative mesh, highlighting how patient-specific anatomical features, particularly fibrotic regions, are incorporated. The two black shapes indicate the positions of the virtual PentaRay catheter within the mesh. Panel (b) depicts the real PentaRay catheter, illustrating its physical structure, which is accurately reproduced in the virtual setup. The rotor dynamics during AF simulations are summarized in panel (c), showing the temporal evolution of electrical activity from the initial tissue condition to the induction and persistence of chaotic rotors and reentrant patterns. These dynamic simulations are initiated by multiple Archimedean spirals, demonstrating the capability of the framework to capture complex atrial fibrillation mechanisms.

The simulation triggers sustained AF patterns, capturing signals exhibiting fractionation and irregularity characteristic of sustained AF. As seen in Figure 2, the resulting EGMs exhibit irregularities and fractionation consistent with clinical observations.

Validation by Expert Electrophysiologists: Expert performance varied across individuals (Table 1), precision and recall reported are computed with respect to detecting

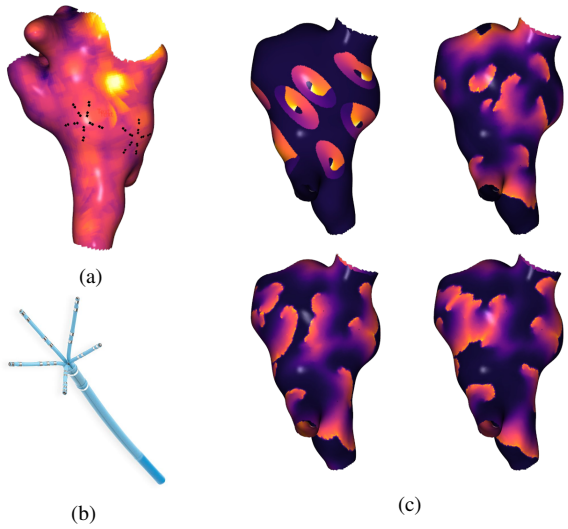


Figure 1: Combined illustration of fibrosis mapping (top-left), real Pentaray catheter (bottom-left), and 3D simulated rotor dynamics (right).

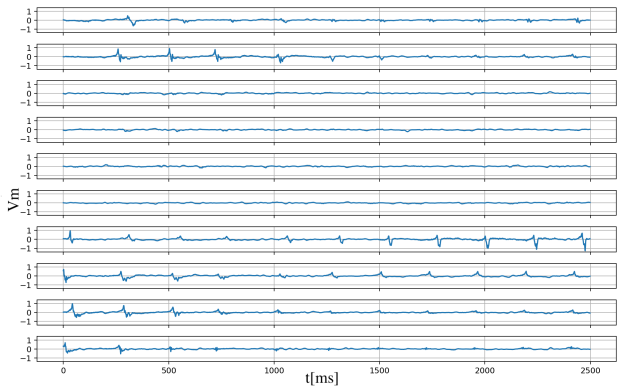


Figure 2: Example of synthetic EGMs under AF.

real samples (i.e., the positive class is "real"). Two experts operated near chance level, whereas three achieved high performance, indicating the presence of residual cues specific to the synthetic signals. Experts did not consistently flag the same samples, suggesting heterogeneous, sample-specific cues rather than a single obvious artifact. Collected feedback suggests that samples with low amplitude and change in polarity are not seen in clinical data, as well as some synthetic samples containing activation patterns not present in clinical data.

4. Conclusion

This work presented a reproducible framework for generating realistic synthetic atrial EGMs, addressing data scarcity in cardiovascular research. By integrating patient-

Table 1: Classification performance for detecting real samples (real = 1, synthetic = 0); $n = 20$ samples per expert

Expert	Accuracy (%)	Precision (%)	Recall (%)
Expert 1	55	53.8	70
Expert 2	55	54.5	60
Expert 3	90	90.0	90
Expert 4	95	90.9	100
Expert 5	80	87.5	70
Average	75	75.3	78

specific anatomical data, including fibrosis and fiber orientation, with electrophysiological modeling in openCARP combined with a strategy for catheter placement to emulate clinical settings, the approach produces morphologically plausible datasets consistent with clinical observations.

The evaluation yielded mixed results as detailed in Table 1. The outcome suggests that synthetic signals are not overtly artificial, but the high performance of some experts, indicates the presence of cues that distinguish samples from real data. Experts did not consistently flag the same samples, suggesting that these cues are heterogeneous and sample-specific. Feedback highlights that while the overall morphology is plausible, inconsistencies in activation dynamics remain. Future work will target the likely sources of residual cues, such as catheter-tissue contact variability, and improvements in the simulation pipeline, including the activation protocol.

Acknowledgments

This work has been supported in part by the French government, through the 3IA Cote d'Azur Investments in the project managed by the National Research Agency (ANR) with the reference number ANR-23-IACL-0001. V. Zarzoso holds the "IAblation" 3IA Chair. Support by the CAPES-COFECUB program through project "INTERACTION" (Ma-985-23) is also acknowledged.

The authors thank Christiano Lemos, Thaís Coutinho, Lucas Corcino, Pedro Victor Silva, and Fabricio Sarmento Vassallo for their technical support in this work.

References

[1] Joseph P, Leong D, McKee M, Anand SS, Schwalm JD, Teo K, Mente A, Yusuf S. Reducing the global burden of cardiovascular disease, part 1: The epidemiology and risk factors. Circulation Research September 2017;121(6):677–694.

[2] Brundel BJJM, Ai X, Hills MT, Kuipers MF, Lip GYH, de Groot NMS. Atrial fibrillation. Nature Reviews Disease Primers April 2022;8(1):21.

[3] Azzolin L, Schuler S, Dössel O, Loewe A. A reproducible protocol to assess arrhythmia vulnerability in silico: Pacing at the end of the effective refractory period. Frontiers Physiology April 2021;12:656411.

[4] Jaffery OA, Melki L, Slabaugh G, Good WW, Roney CH. A review of personalised cardiac computational modelling using electroanatomical mapping data. Arrhythmia and Electrophysiology Review May 2024;13:e08.

[5] Corral-Acero J, Margara F, Marciniak M, Rodero C, Loncaric F, Feng Y, Gilbert A, Fernandes JF, Bukhari HA, Wajdan A, Martinez MV. The 'digital twin' to enable the vision of precision cardiology. European Heart Journal December 2020;41(48):4556–4564.

[6] Wang Z, Yi R, Wen X, Zhu C, Xu K. Cardiovascular medical image and analysis based on 3d vision: A comprehensive survey. Meta Radiology 2024;2(4):100102. ISSN 2950-1628.

[7] Zahid S, Cochet H, Boyle PM, Schwarz EL, Whyte KN, Vigmond EJ, Dubois R, Hocini M, Haïssaguerre M, Jaïs P, Trayanova NA. Patient-derived models link re-entrant driver localization in atrial fibrillation to fibrosis spatial pattern. Cardiovascular Research June 2016;110(3):443–454.

[8] Roney CH, Sim I, Yu J, Beach M, Mehta A, Alonso Solis-Lemus J, Kotadia I, Whitaker J, Corrado C, Razeghi O, Vigmond E, Narayan SM. Predicting atrial fibrillation recurrence by combining population data and virtual cohorts of patient-specific left atrial models. Circulation Arrhythmia and Electrophysiology February 2022;ISSN 1941-3084.

[9] Plank G, Loewe A, Neic A, Augustin C, Huang YL, Gsell MA, Karabelas E, Nothstein M, Prassl AJ, Sánchez J, Seemann G, Vigmond EJ. The opencarp simulation environment for cardiac electrophysiology. Computer Methods and Programs in Biomedicine 2021;ISSN 0169-2607.

[10] Grandi E, Dobrev D, Heijman J. Computational modeling: What does it tell us about atrial fibrillation therapy? International Journal of Cardiology July 2019;287:155–161.

[11] Bishop MJ, Plank G. Bidomain ecg simulations using an augmented monodomain model for the cardiac source. IEEE Transactions on Biomedical Engineering August 2011;2297–2307. ISSN 1558-2531.

[12] Abdikivanani B. Atrial fibrillation fingerprinting. Ph.D. thesis, Delft University of Technology, 2021.

[13] Roney CH, Bayer JD, Zahid S, Meo M, Boyle PMJ, Trayanova NA, Haïssaguerre M, Dubois R, Cochet Hubert aAjEpOvmdykAle. Modelling methodology of atrial fibrosis affects rotor dynamics and electrograms ;.

[14] Ghrissi A, Almonfrey D, de Almeida RC, Squara F, Montagnat J, Zarzoso V. Data augmentation for automatic identification of spatiotemporal dispersion electrograms in persistent atrial fibrillation ablation using machine learning. In 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society. 2020; 406–409.

Address for correspondence:

Rafael Costa de Almeida.
Avenida Vitória, 1729, Jucutuquara, Vitória, Brazil.
contact@rafaelcostadealmeida.com